Research Article

Privacy-Preserving Self-Helped Medical Diagnosis Scheme Based on Secure Two-Party Computation in Wireless Sensor Networks

Yi Sun,¹ Qiaoyan Wen,¹ Yudong Zhang,² and Wenmin Li¹

¹ State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

² Brain Image Processing, Columbia University, New York, NY 10032, USA

Correspondence should be addressed to Yi Sun; sybupt@bupt.edu.cn

Received 10 April 2014; Accepted 23 June 2014; Published 9 July 2014

Academic Editor: David A. Elizondo

Copyright © 2014 Yi Sun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the continuing growth of wireless sensor networks in pervasive medical care, people pay more and more attention to privacy in medical monitoring, diagnosis, treatment, and patient care. On one hand, we expect the public health institutions to provide us with better service. On the other hand, we would not like to leak our personal health information to them. In order to balance this contradiction, in this paper we design a privacy-preserving self-helped medical diagnosis scheme based on secure two-party computation in wireless sensor networks so that patients can privately diagnose themselves by inputting a health card into a self-helped medical diagnosis ATM to obtain a diagnostic report just like drawing money from a bank ATM without revealing patients' health information and doctors' diagnostic skill. It makes secure self-helped disease diagnosis feasible and greatly benefits patients as well as relieving the heavy pressure of public health institutions.

1. Introduction

With the rapid development of science, more and more advanced technologies such as the internet of things and cloud computing are utilized in the area of modern medicine and this trend further pushes healthcare into the digital era [1–3]. Currently, numerous healthcare devices such as heart rate monitor, blood pressure monitor, and electrocardiogram are already popular in people's normal life. It makes it convenient for people to be aware of their health situation by viewing the reports of these devices. Especially, by the growing use of sensor technology in telecare, the new field known as wireless body area networks (WBAN) [1, 4] has designed various sensor devices that can be used to supervise critical body parameters and activities anytime and anywhere. People can easily and conveniently get the health data by these advanced sensor devices [5] such as temperature measurement, respiration monitor, heart rate monitor, pulse oximeter SpO₂, blood pressure monitor, pH monitor, glucose sensor, cardiac arrhythmia monitor/recorder, brain liquid pressure sensor, and endoscope capsule. What is more, these

devices are becoming more functional and portable. More and more mobile medical monitors have already been used to serve us [2].

Therefore, people no longer worry about how to obtain the health data but are concerned about how to securely deal with these sensitive data to have disease diagnosis with a medical institution. Traditionally, the issue of privacy of medical data has been dealt with primarily as a policy problem [6, 7]. Many related laws have been issued to protect the privacy of patients. However, it is still far away from satisfactory and people still fear the leakage of their private data. Hence, the most efficient solution to this problem is to protect patients' privacy in technology rather than in policy alone. In this aspect, most of previous literatures have introduced homomorphic encryption (HE) [8–10] to protect patients' privacy in some privacy-preserving medical applications [11]. However, HE will inevitably introduce tremendous cost and is not applicable to practical large-scale applications. Therefore, in this paper, we focus on building a secure and practical privacy-preserving medical diagnosis system that can serve us in our daily life. Starting from the aspiration of

the patient, the most secure and plausible diagnostic method is to apply the processed data rather than the original data to interact with the hospital which owns a disease database to diagnose the health status privately. Moreover, it requires that after diagnosis, the hospital gets nothing about the patient's health data and the patient has no idea of the hospital's disease database.

Inspired by daily used bank automated teller machine (ATM), we introduce the privacy-preserving self-helped medical diagnosis ATM (MD-ATM) so that after obtaining a healthcare card that stores some information about the health data which is collected by various sensor medical devices, patients can privately diagnose himself by inserting the health card into the MD-ATM to obtain diagnostic report just like drawing money from a bank ATM without revealing patient's health information and the disease database or doctors' diagnostic skill. When needing local computing, storing, or inputting some information, the patient uses his own portable device, called portable medical diagnostic device (PMDD).

In this paper, we will show how to realize this modern diagnosis system without HE. The main idea and technology we used in this scheme are secure two-party computation (STC) and oblivious transfer (OT). Firstly, we assume that patients themselves collect related data by various wireless sensor medical devices and further process and store them in their own health cards using PMDD. When diagnosing, the patient firstly transforms the original data locally and then inserts the card into the MD-ATM of the hospital to check up his health. Operating following the instructions of the MD-ATM, the patient will finally obtain a diagnostic report through OT and the patient then completes the self-helped diagnosis. In brief, our main contributions can be summarized as follows.

Our Contributions.

- (i) We build a new "patient-centered" medical diagnosis model in wireless sensor networks where patients themselves collect health data by various sensor medical devices while the hospital provides a disease database to help patients to complete disease diagnosis by themselves. Compared with traditional "doctor-centered" medical diagnosis model where patients have to depend on the doctor, our system is more appropriate especially when people pay more and more attention to privacy in wireless sensor networks.
- (ii) We firstly propose the privacy-preserving self-helped MD-ATM to construct a secure medical diagnosis scheme following the idea of STC. It makes secure self-helped medical diagnosis feasible and convenient just like drawing money from a bank ATM. It will greatly benefit patients as well as relieving the heavy pressure of public health institutions.
- (iii) We construct the self-helped medical diagnosis system based on OT without expensive HE. It provides us with another perspective to consider the problem of secure medical diagnosis for patients.

The rest of this paper is organized as follows. In Section 2, we briefly give an overview of secure two-party computation and oblivious transfer, and then we present our medical diagnosis system model in Section 3. In Section 4, we propose our privacy-preserving self-helped medical diagnosis scheme in detail and give a strict proof based on real-ideal simulation paradigm in Section 5. Finally, we summarize our work of this paper in the last section.

2. Preliminaries

2.1. Secure Two-Party Computation. Secure multiparty computation (SMC) is dedicated to deal with the problem of secure computation among distrustful participants. It was first introduced by Yao in 1982 [12] and then was extended by Goldreich et al. [13] and many other researchers [14–19]. Generally speaking, SMC is a method to implement cooperative computation with participants' private data, ensuring the correctness of the computation as well as not disclosing additional information except the necessary results. It has become a research focus in the international cryptographic community due to its wide applications in various areas and a mass of research results have been published one after another. Secure two-party computation (STC) [20] is a special case in SMC where there are only two participants. The well-known millionaires' problem [12] put forward by Yao is the representative problem of STC. In our discussing, we will consider the two-party case.

Generally speaking, STC is dedicated to computing a certain function between two mutually distrusted participants on their private inputs without revealing their private information. Informally, assuming that there are 2 participants, P_1, P_2 , each of them has a private number, x_1, x_2 , respectively. They want to cooperate to compute the function $y = f(x_1, x_2)$. A STC protocol is dubbed secure if no participant can learn more from the description of the public function and the result of the global calculation than what he can learn from his own information.

Formally, we usually analyze the security of a STC protocol using the real-ideal paradigm in the semihonest model where both of the two parties act semihonestly, following the protocol but making effort to gain more information about other parties' inputs, intermediate results, or overall outputs by the transcripts of the protocol [15]. We can overview the real-ideal paradigm as follows.

Firstly, in the ideal world, we assume that the computation of the functionality F on users private inputs is conducted by an additional trusted party, who receives x_i from user P_i , $i = 1, 2$, and returns the result $f(x_1, x_2)$ to P_i , $i = 1, 2$. However, there is no trusted party in the real world and so the two parties have to run a protocol Π to get the desired result. During executing protocol Π , both parties act semihonestly. Herein, the view of the i th party during an execution of Π on x_1, x_2 is denoted as $\text{VIEW}_i^\Pi(x_1, x_2)$, which contains P_i 's input, random tape, and the messages received from the other party. For a deterministic private function f , we say that Π privately computes f if there exist probabilistic polynomial-time algorithms S_1, S_2 , such that the simulated distribution

	OT_1^m
Inputs:	S inputs a set of m messages: $(r_1, r_2, \dots, r_m)$; R inputs an index: z ;
Outputs:	S obtains λ , which means <i>Null</i> . R obtains r_z .

ALGORITHM 1

$\{S_i(x_i; f(x_1, x_2))\}$ is indistinguishable to $\{VIEW_i^\Pi(x_1, x_2)\}$, $i = 1, 2$. That is,

$$\begin{aligned} \{S_1(x_1; f(x_1, x_2))\} &\equiv \{VIEW_1^\Pi(x_1, x_2)\}, \\ \{S_2(x_2; f(x_1, x_2))\} &\equiv \{VIEW_2^\Pi(x_1, x_2)\}. \end{aligned} \quad (1)$$

2.2. Oblivious Transfer. In cryptography, OT is a type of protocol in which a sender transfers one of potentially many pieces of information to a receiver but remains oblivious as to which piece has been transferred. It was firstly introduced by Rabin [21] in 1981. Therein, the sender sends a message to the receiver with probability 1/2, while the sender remains oblivious as to whether or not the receiver received the message. Rabin's oblivious transfer scheme is based on the RSA cryptosystem. In 1985, Even et al. [22] proposed a more useful OT called 1-out-of-2 OT (OT_1^2) to build protocols for secure multiparty computation.

Afterwards, it has been generalized to 1-out-of- m OT (OT_1^m) [23] where the receiver gets exactly one message without the sender getting to know which message was queried and the receiver getting to know anything about the other messages that were not retrieved. OT_1^m has become a fundamental tool in cryptography and is usually used as a black-box when constructing protocols.

Formally, we can describe an OT_1^m protocol as follows. There are 2 participants called the sender S and the receiver R . Specifically, S has m messages, and R has an index z . R wishes to receive the z th message of the sender's m messages without leaking z to S , while knowing nothing about the rest $m - 1$ messages. A simplified OT_1^m protocol can be presented as in Algorithm 1.

3. System Model

In this section, we present the system model including the goals we aim to achieve in detail.

In this paper, we consider the privacy-preserving medical diagnosis system with two participants: the patient and the hospital. We assume that each patient can collect his own health data such as heart beat and blood pressure, in the form of a vector, called query vector, easily by various advanced medical devices. Herein, we call the heart beat, blood pressure, and so forth, as parameter items and the health data corresponding to heart beat, blood pressure, and so forth, as parameter values. For example, $q_P = (q_{P1}, \dots, q_{Pn})$ is the query vector of the patient P , where all $\{q_{Pw}\}_{w=1, \dots, n}$ are

the necessary parameters the hospital needs for diagnosis, and q_{Pw} is the parameter value of the parameter item heart beat. Each patient has a health card to store related data and a portable device PMDD to read the data stored in the card and to do some related computations after inserting the card. The hospital has a disease database $DB = \{d_i\}_{i=1, \dots, m}$, which in fact is the standard to determine which disease the patient has got. Each record of the disease database is presented as a triple $d_i = (i, t_i, r_i)$, $i = 1, \dots, m$, where m is the capacity of the disease database; i is the index of a disease; t_i , called the trait vector of the disease d_i , is a vector that covers all necessary parameters the hospital needs for diagnosis; and r_i is the disease diagnostic report including the disease name, doctors' advices, and prescriptions corresponding to the i th disease d_i . Concerning these parameters, we have some illustrations as follows.

- (i) q_P : it includes all necessary parameter items the hospital needs for diagnosis such as heart beat and blood pressure. The query vectors of the same patient P are different if P goes to different hospitals since their medical levels are different. The query vectors of the same patient P may be also different if P goes to the same hospital at different time since the medical level of the hospital has been always keeping improving. The dimension n and every parameter item $\{q_{Pw}\}_{w=1, \dots, n}$ of the query vector are determined by the trait vector t_i of the hospital. After registering to the hospital, patient can know what parameter items are needed in this diagnosis by reading the health card.
- (ii) $DB = \{d_i\}_{i=1, \dots, m}$: it concludes all diseases a hospital can diagnose. Different hospitals have different disease databases and the same hospital has different disease databases at different time since its medical level has been keeping improving. The dimension m is determined by the hospital's medical level.
- (iii) t_i : it includes all necessary parameter items the hospital needs for diagnosis such as heart beat and blood pressure. Different hospitals have different trait vectors and the trait vectors of the same hospital at different time may be different since the medical level of the hospital has been always keeping improving. The dimension n and every parameter item $\{t_{ij}\}_{j=1, \dots, n}$ of the trait vector are determined by the hospital's medical level. In order to improve the precision of our diagnosis system, the hospital can consider as many

FIGURE 1: Self-help medical diagnosis model of our scheme.

factors as possible such as adding more personal feelings, symptoms, and previous medical features from the patient as parameter items. Although we only can diagnose some simple diseases currently, it is believed that it will be feasible for more complicated diseases in the future by extending the dimension n of the parameter items.

- (iv) r_i : it includes the disease name, doctors' advices, and prescriptions corresponding to the i th disease d_i . Each report may conclude many doctors' advices and prescriptions. Herein, we assume that every report obtained from the MD-ATM following the self-help medical diagnosis is authorized by the hospital and all advices and prescriptions of a report are signed by corresponding doctors. After receiving the diagnostic report, patient can choose one doctor's advice and prescription to treat himself.

In this paper, the system makes medical diagnosis according to the Euclidean distances of two vectors. Specifically, taking the query vector of the patient and a trait vector of the database as an example, given a patient's query vector $q_p = (q_{p1}, \dots, q_{pn})$ and a disease trait vector $t_i = (t_{i1}, \dots, t_{in})$, $i \in \{1, \dots, m\}$, their Euclidean distance [3] denoted by $\text{dist}_{q_p \leftrightarrow t_i}$ is

$$\text{dist}_{q_p \leftrightarrow t_i} = \sqrt{\sum_{w=1}^n (q_{pw} - t_{iw})^2}. \quad (2)$$

Herein, we compare the squares of the Euclidean distances,

$$\text{dist}_{q_p \leftrightarrow t_i}^2 - \text{dist}_{q_p \leftrightarrow t_j}^2 = \sum_{w=1}^n (q_{pw} - t_{iw})^2 - \sum_{w=1}^n (q_{pw} - t_{jw})^2. \quad (3)$$

It is obvious that we can figure out which one has smaller distance with patient's query vector just by checking the sign of (3) without exact result of $\text{dist}_{q_p \leftrightarrow t_i}$ or $\text{dist}_{q_p \leftrightarrow t_j}$. Assuming that the report r_z corresponding to the trait vector

t_z , $z \in \{1, \dots, m\}$, is the diagnosed disease report, we have the following result, for all $j = 1, \dots, m$, $j \neq z$:

$$\text{dist}_{q_p \leftrightarrow t_z}^2 - \text{dist}_{q_p \leftrightarrow t_j}^2 < 0. \quad (4)$$

In our scheme, we will compare the squares of the Euclidean distances of the query vector and the trait vectors to find the diagnostic report that satisfies (4).

In real application, the hospital provides a MD-ATM, which is connected with the disease database and can read the data of the card, to direct patients to complete self-help disease diagnosis. Specifically, we assume that each patient registers to the hospital for the first time and gets a health card. The hospital provides a self-help MD-ATM in public just like a bank ATM. Whenever P wants to have a diagnosis, inserting his health card into the MD-ATM and following the instructions, P can complete the self-help diagnosis by himself. The basic model can be illustrated in Figure 1.

Apart from the above, to enable a privacy-preserving medical diagnosis system, our scheme should simultaneously fulfill the following two security goals.

- (i) Confidentiality of disease database should be protected during the self-help diagnosis process.
- (ii) Confidentiality of patient's private health data should be protected during the self-help diagnosis process.

4. Our Scheme

In this section, we propose our privacy-preserving self-help medical diagnosis scheme (PP-SH-MDS) in detail to show how a patient can diagnose by himself using his PMDD and the self-help MD-ATM. The core of our construction can be summarized in Figure 2.

Specifically, the patient P executes as follows to make a self-help diagnosis using his PMDD and the MD-ATM.

In the setup phase, P registers to a hospital as traditional medical diagnosis and gets a health card.

In the diagnosis phase, there are three subphases.

FIGURE 2: PP-SH-MDS.

(1) *Local Preprocessing.* Whenever P wants to have a diagnosis, he firstly conducts the following two transformations on PMDD locally.

(i) Vector-to-Vector.

(a) P firstly extends his original health data from an n -vector $q_P = (q_{P1}, \dots, q_{Pn})$ to an $(n+2)$ -vector $Q_P = (q_{P1}, \dots, q_{Pn}, q_{P,n+1}, q_{P,n+2})$, where $q_{P,n+1} = -(1/2) \sum_{w=1}^n q_{Pw}^2$ and $q_{P,n+2} = 1$.

(ii) Vector-to-Matrix.

(a) P randomly chooses a password $s_P = (s_{P1}, \dots, s_{Pk})$ and then generates a $k \times (n+2)$ matrix

$$B_P = \begin{bmatrix} b_{11} & \cdots & b_{1,n+2} \\ \vdots & & \vdots \\ b_{k1} & \cdots & b_{k,n+2} \end{bmatrix}, \quad (5)$$

where $\sum_{u=1}^k s_{Pu} \cdot b_{uw} = 1, w = 1, \dots, n+2$.

(b) By blinding Q_P using the matrix B_P , P further extends Q_P to a matrix

$$M_P = \begin{bmatrix} b_{11} \cdot q_{P1} & \cdots & b_{1,n+2} \cdot q_{P,n+2} \\ \vdots & & \vdots \\ b_{k1} \cdot q_{P1} & \cdots & b_{k,n+2} \cdot q_{P,n+2} \end{bmatrix}. \quad (6)$$

After completing the above steps, P stores the matrix M_P in the health card.

(2) *Diagnosis.*

(i) After local preprocessing, P inserts his health card into the MD-ATM and then the MD-ATM reads the card to get the matrix M_P and randomly chooses two trait vectors $t_i = (t_{i1}, \dots, t_{in})$ and $t_j = (t_{j1}, \dots, t_{jn})$ and, respectively, extends them to $(n+2)$ -vectors $T_i = (t_{i1}, \dots, t_{in}, t_{i,n+1}, t_{i,n+2})$ and $T_j = (t_{j1}, \dots, t_{jn}, t_{j,n+1}, t_{j,n+2})$, where $t_{i,n+1} = 1, t_{i,n+2} = -(1/2) \sum_{w=1}^n t_{iw}^2$ and $t_{j,n+1} = 1, t_{j,n+2} = -(1/2) \sum_{w=1}^n t_{jw}^2$. Then the MD-ATM computes $\Delta'_{ij} = M_P \cdot (T_i - T_j)^T, i, j = 1, \dots, m, i \neq j$, and writes (Δ'_{ij}, i, j) in the card and indicates the patient to get back his card.

(ii) Following the instructions of the MD-ATM, the patient gets the card back and inserts it into PMDD. After inputting his password, PMDD begins to compute $\Delta_{ij} = s_P \cdot \Delta'_{ij}, i, j = 1, \dots, m, i \neq j$, and finds the index z so that for all $j = 1, \dots, m, j \neq z, \Delta_{ij} > 0$. z is the input of the following OT₁^m protocol.

(3) *1-out-of- m OT Protocol.*

(i) P inserts his card into the MD-ATM and invokes an OT₁^m protocol, where P 's input is the index z and the MD-ATM's input is the diagnostic report set $(r_1, r_2, \dots, r_m)$ of the database.

After executing the OT₁^m protocol, P gets the diagnostic report r_z corresponding to the disease d_z according to the index z , while the MD-ATM gets *Null*, denoted by λ .

5. Analysis

In this section, we analyze our scheme in detail. We firstly have a look at the correctness and then give a strict security proof following the real-ideal simulation paradigm of STC in the scenarios of semihonest adversaries.

5.1. Correctness. In this aspect, we follow the steps of our scheme and make sure that the patient indeed finds out the most possible disease from the disease database of the hospital using his health data by comparing Euclidean distances.

Following the scheme, we can see that the patient transforms the health data in two steps. Firstly, he extends his original health data from an n -vector $q_P = (q_{P1}, \dots, q_{Pn})$ to an $(n + 2)$ -vector $Q_P = (q_{P1}, \dots, q_{Pn}, q_{P,n+1}, q_{P,n+2})$, where $q_{P,n+1} = -(1/2) \sum_{w=1}^n q_{Pw}^2$ and $q_{P,n+2} = 1$ and then blinds and extends Q_P using the matrix B_P to a matrix

$$M_P = \begin{bmatrix} b_{11} \cdot q_{P1} & \cdots & b_{1,n+2} \cdot q_{P,n+2} \\ \vdots & & \vdots \\ b_{k1} \cdot q_{P1} & \cdots & b_{k,n+2} \cdot q_{P,n+2} \end{bmatrix}. \quad (7)$$

On the other hand, the MD-ATM randomly selects two trait vectors $t_i = (t_{i1}, \dots, t_{in})$ and $t_j = (t_{j1}, \dots, t_{jn})$ and, respectively, extends them to $(n + 2)$ -vectors $T_i = (t_{i1}, \dots, t_{in}, t_{i,n+1}, t_{i,n+2})$ and $T_j = (t_{j1}, \dots, t_{jn}, t_{j,n+1}, t_{j,n+2})$, where $t_{i,n+1} = 1$, $t_{i,n+2} = -(1/2) \sum_{w=1}^n t_{iw}^2$ and $t_{j,n+1} = 1$, $t_{j,n+2} = -(1/2) \sum_{w=1}^n t_{jw}^2$. After receiving M_P , for $i, j = 1, \dots, m$, $i \neq j$, the MD-ATM computes

$$\begin{aligned} \Delta'_{ij} &= M_P \cdot (T_i - T_j)^T \\ &= \begin{bmatrix} b_{11} \cdot q_{P1} & \cdots & b_{1,n+2} \cdot q_{P,n+2} \\ \vdots & & \vdots \\ b_{k1} \cdot q_{P1} & \cdots & b_{k,n+2} \cdot q_{P,n+2} \end{bmatrix} \cdot \begin{bmatrix} t_{i1} - t_{j1} \\ \vdots \\ t_{i,n+2} - t_{j,n+2} \end{bmatrix}. \end{aligned} \quad (8)$$

After receiving the returned message Δ'_{ij} , for $i, j = 1, \dots, m$, $i \neq j$, P uses PMDD to compute

$$\begin{aligned} \Delta_{ij} &= s_P \cdot \Delta'_{ij} = s_P \cdot M_P \cdot (T_i - T_j)^T \\ &= (s_{P1}, \dots, s_{Pk}) \cdot \begin{bmatrix} b_{11} \cdot q_{P1} & \cdots & b_{1,n+2} \cdot q_{P,n+2} \\ \vdots & & \vdots \\ b_{k1} \cdot q_{P1} & \cdots & b_{k,n+2} \cdot q_{P,n+2} \end{bmatrix} \\ &\quad \cdot \begin{bmatrix} t_{i1} - t_{j1} \\ \vdots \\ t_{i,n+2} - t_{j,n+2} \end{bmatrix} \end{aligned}$$

$$\begin{aligned} &= \sum_{u=1}^k s_{Pu} \cdot b_{u1} \cdot q_{P1} \cdot (t_{i1} - t_{j1}) + \cdots \\ &\quad + \sum_{u=1}^k s_{Pu} \cdot b_{u,n+2} \cdot q_{P,n+2} \cdot (t_{i,n+2} - t_{j,n+2}) \\ &= \sum_{w=1}^{n+2} q_{Pw} \cdot (t_{iw} - t_{jw}). \end{aligned} \quad (9)$$

Thus, we have

$$\begin{aligned} \Delta_{ij} &= \sum_{w=1}^{n+2} q_{Pw} \cdot t_{iw} - \sum_{w=1}^{n+2} q_{Pw} \cdot t_{jw} \\ &= \left(\sum_{w=1}^n q_{Pw} \cdot t_{iw} + q_{P,n+1} \cdot t_{i,n+1} + q_{P,n+2} \cdot t_{i,n+2} \right) \\ &\quad - \left(\sum_{w=1}^n q_{Pw} \cdot t_{jw} + q_{P,n+1} \cdot t_{j,n+1} + q_{P,n+2} \cdot t_{j,n+2} \right) \\ &= \left(\sum_{w=1}^n q_{Pw} \cdot t_{iw} + q_{P,n+1} \cdot t_{i,n+1} \right) \\ &\quad - \left(\sum_{w=1}^n q_{Pw} \cdot t_{jw} + q_{P,n+1} \cdot t_{j,n+1} \right) \\ &= \left(\sum_{w=1}^n q_{Pw} \cdot t_{iw} - \frac{1}{2} \sum_{w=1}^n q_{Pw}^2 - \frac{1}{2} \sum_{w=1}^n t_{iw}^2 \right) \\ &\quad - \left(\sum_{w=1}^n q_{Pw} \cdot t_{jw} - \frac{1}{2} \sum_{w=1}^n q_{Pw}^2 - \frac{1}{2} \sum_{w=1}^n t_{jw}^2 \right) \\ &= \frac{1}{2} \left[\sum_{w=1}^n (q_{Pw} - t_{jw})^2 - \sum_{w=1}^n (q_{Pw} - t_{iw})^2 \right] \\ &= \frac{1}{2} \left(\text{dist}_{q_P \leftrightarrow t_j}^2 - \text{dist}_{q_P \leftrightarrow t_i}^2 \right). \end{aligned} \quad (10)$$

Obviously, if for all $j = 1, \dots, m$, $j \neq z$, $\Delta_{zj} = (1/2) (\text{dist}_{q_P \leftrightarrow t_j}^2 - \text{dist}_{q_P \leftrightarrow t_z}^2) > 0$, then t_z , $z \in \{1, \dots, m\}$, is the trait vector of the diagnosed disease. The report r_z corresponding to the trait vector t_z is the diagnosed report. Taking the index z as the input of the following OT_1^m protocol, the patient can finally get the disease report r_z from the set $(r_1, r_2, \dots, r_m)$ of the database.

Therefore, our scheme is correct.

5.2. Security. In this subsection, we strictly prove the security of our scheme. From the whole process, we can specify that the two parties in our system are the patient P and the hospital. They cooperate to compute the function $f(q_P, (d_1, d_2, \dots, d_m)) = r_z$, where r_z is the disease diagnostic report corresponding to the disease d_z and the distance $\text{dist}_{q_P \leftrightarrow t_z}$ satisfies the condition $\text{dist}_{q_P \leftrightarrow t_z}^2 = \min\{\text{dist}_{q_P \leftrightarrow t_j}^2\}_{j=1, \dots, m}$. As mentioned in Section 3, we should

achieve two security goals, that is, keeping both parties' inputs private. We apply the real-ideal simulation paradigm to prove that our scheme has achieved the two goals in the scenarios of semihonest adversaries assuming the OT_1^m protocol we used is secure.

Theory 1. *Our privacy-preserving self-helped medical diagnosis scheme is secure against semihonest adversaries if the OT_1^m protocol is secure.*

Proof. Notice that the view of P_i , $\{\text{VIEW}_i(q_p, \{d_1, \dots, d_m\})\}_{i=1,2}$, in the real execution consists of three parts, the private input, random tape, and the messages received from the other party including the output. Therefore, we can get the views of P_1 and P_2 , respectively, in the real execution as follows:

$$\begin{aligned} & \{\text{VIEW}_1(q_p, \{d_i\}_{i=1,\dots,m})\} \\ &= \{q_p, s_p, B_p, \{\Delta'_{ij}\}_{i,j=1,\dots,m, i \neq j}, \end{aligned} \quad (11)$$

$$\begin{aligned} & \{\text{VIEW}_1^{\text{OT}_1^m}(z, \{r_i\}_{i=1,\dots,m})\}, r_z\}; \\ & \{\text{VIEW}_2(q_p, \{d_i\}_{i=1,\dots,m})\} \\ &= \{\{d_i\}_{i=1,\dots,m}, M_p, \{\text{VIEW}_2^{\text{OT}_1^m}(z, \{r_i\}_{i=1,\dots,m})\}, \lambda\}, \end{aligned} \quad (12)$$

where $\{\text{VIEW}_1^{\text{OT}_1^m}(z, \{r_i\}_{i=1,\dots,m})\}$ and $\{\text{VIEW}_2^{\text{OT}_1^m}(z, \{r_i\}_{i=1,\dots,m})\}$ are the views produced in the execution of OT_1^m protocol.

From the definition of security, we need to construct the probabilistic polynomial-time algorithm S_1/S_2 so that given the input and output of the patient P_1 /the hospital P_2 , $(q_p; r_z)/(\{d_i\}_{i=1,\dots,m}; \lambda)$, it can output a simulated view $\{S_1(q_p; r_z)/\{S_2(\{d_i\}_{i=1,\dots,m}; \lambda)\}$, which is indistinguishable to the view $\{\text{VIEW}_1(q_p, \{d_i\}_{i=1,\dots,m})\}/\{\text{VIEW}_2(q_p, \{d_i\}_{i=1,\dots,m})\}$ in the real execution of the scheme; that is,

$$\begin{aligned} & \{S_1(q_p; r_z)\} \cong \{\text{VIEW}_1(q_p, \{d_i\}_{i=1,\dots,m})\}, \\ & \{S_2(\{d_i\}_{i=1,\dots,m}; \lambda)\} \cong \{\text{VIEW}_2(q_p, \{d_i\}_{i=1,\dots,m})\}. \end{aligned} \quad (13)$$

In the following discussion, we follow the real-ideal simulation paradigm to construct such probabilistic polynomial-time algorithms S_1, S_2 . We separately prove the case when P_2 is semihonest and when P_1 is semihonest.

Case 1 (P_2 is semihonest). In this case, we only need to construct a simulator S_2 so that, given P_2 's input $\{d_i\}_{i=1,\dots,m}$ and output λ , S_2 can simulate P_2 's view in the real execution presented above as (12).

Firstly, since we assume that the OT_1^m protocol used in our scheme is secure and can be taken as a black-box, there exists an algorithm $S_2^{\text{OT}_1^m}$: given the input $\{r_i\}_{i=1,\dots,m}$ and the output λ , it can simulate P_2 's view of the OT_1^m execution and output $\{S_2^{\text{OT}_1^m}(\{r_i\}_{i=1,\dots,m}; \lambda)\}$ so that

$$\{S_2^{\text{OT}_1^m}(\{r_i\}_{i=1,\dots,m}; \lambda)\} \cong \{\text{VIEW}_2^{\text{OT}_1^m}(z, \{r_i\}_{i=1,\dots,m})\}. \quad (14)$$

Next, notice that S_2 is given $(\{d_i\}_{i=1,\dots,m}; \lambda)$; it can easily simulate the remaining parts of (12) by randomly choosing a $k \times (n+2)$ matrix M which is indistinguishable to the blinded matrix M_p . Then, S_2 outputs the simulated view,

$$\begin{aligned} & \{S_2(\{d_i\}_{i=1,\dots,m}; \lambda)\} \\ &= \{\{d_i\}_{i=1,\dots,m}, M, \{S_2^{\text{OT}_1^m}(\{r_i\}_{i=1,\dots,m}; \lambda)\}, \lambda\}. \end{aligned} \quad (15)$$

Obviously, we can conclude that

$$\{S_2(\{d_i\}_{i=1,\dots,m}; \lambda)\} \cong \{\text{VIEW}_2(q_p, \{d_i\}_{i=1,\dots,m})\}. \quad (16)$$

Case 2 (P_1 is semihonest). Similar to Case 1, we only need to construct a simulator S_1 so that given P_1 's input q_p and output r_z , S_1 can simulate P_1 's view in the real execution presented above as (11).

As discussed above, since the OT_1^m protocol is secure, there exists an algorithm $S_1^{\text{OT}_1^m}$: given the input z and the output r_z , it can simulate P_1 's view of the OT_1^m execution and output $\{S_1^{\text{OT}_1^m}(z; r_z)\}$ so that

$$\{S_1^{\text{OT}_1^m}(z; r_z)\} \cong \{\text{VIEW}_1^{\text{OT}_1^m}(z, \{r_i\}_{i=1,\dots,m})\}. \quad (17)$$

Next, given $(q_p; r_z)$, S_1 then simulates the remaining parts of P_1 's view in the real execution as follows.

Firstly, as in the real execution, S_1 extends the original health data from an n -vector $q_p = (q_{p1}, \dots, q_{pn})$ to an $(n+2)$ -vector $Q_p = (q_{p1}, \dots, q_{pn}, q_{p,n+1}, q_{p,n+2})$, where $q_{p,n+1} = -(1/2) \sum_{w=1}^n q_{pw}^2$ and $q_{p,n+2} = 1$. Then, S_1 randomly chooses a password $s'_p = (s'_{p1}, \dots, s'_{pk})$ and then generates a $k \times (n+2)$ matrix

$$B'_p = \begin{bmatrix} b'_{11} & \dots & b'_{1,n+2} \\ \vdots & & \vdots \\ b'_{k1} & \dots & b'_{k,n+2} \end{bmatrix}, \quad (18)$$

where $\sum_{u=1}^k s'_{pu} \cdot b'_{uw} = 1$, $w = 1, \dots, n+2$. By blinding Q_p using the matrix B'_p , S_1 further extends Q_p to a matrix

$$M'_p = \begin{bmatrix} b'_{11} \cdot q_{p1} & \dots & b'_{1,n+2} \cdot q_{p,n+2} \\ \vdots & & \vdots \\ b'_{k1} \cdot q_{p1} & \dots & b'_{k,n+2} \cdot q_{p,n+2} \end{bmatrix}. \quad (19)$$

Then S_1 randomly selects m vectors, $T'_j = (t'_{j1}, \dots, t'_{jn}, t'_{j,n+1}, t'_{j,n+2})$, $t'_{j,n+1} = 1$, $t'_{j,n+2} = -(1/2) \sum_{w=1}^n t'^2_{jw}$, $j = 1, \dots, m$, and for all $j = 1, \dots, m$, $j \neq z$, $\sum_{w=1}^n (q_{pw} - t'_{zw})^2 < \sum_{w=1}^n (q_{pw} - t'_{jw})^2$; otherwise, S_1 reselects T'_j , $j = 1, \dots, m$. Afterwards, S_1 computes $\Lambda'_{ij} = M'_p \cdot (T'_i - T'_j)^T$, $i, j = 1, \dots, m$, $i \neq j$. Thus, for $i, j = 1, \dots, m$, $i \neq j$, $\Lambda_{ij} = s'_p \cdot \Lambda'_{ij} = \sum_{w=1}^{n+2} q_{pw} \cdot (t'_{iw} - t'_{jw})$.

Therefore, we have

$$\begin{aligned}
\Lambda_{ij} &= \sum_{w=1}^{n+2} q_{Pw} \cdot t'_{iw} - \sum_{w=1}^{n+2} q_{Pw} \cdot t'_{jw} \\
&= \left(\sum_{w=1}^n q_{Pw} \cdot t'_{iw} + q_{P,n+1} \cdot t'_{i,n+1} + q_{P,n+2} \cdot t'_{i,n+2} \right) \\
&\quad - \left(\sum_{w=1}^n q_{Pw} \cdot t'_{jw} + q_{P,n+1} \cdot t'_{j,n+1} + q_{P,n+2} \cdot t'_{j,n+2} \right) \\
&= \left(\sum_{w=1}^n q_{Pw} \cdot t'_{iw} + q_{P,n+1} \cdot t'_{i,n+2} \right) \\
&\quad - \left(\sum_{w=1}^n q_{Pw} \cdot t'_{jw} + q_{P,n+1} \cdot t'_{j,n+2} \right) \quad (20) \\
&= \left(\sum_{w=1}^n q_{Pw} \cdot t'_{iw} - \frac{1}{2} \sum_{w=1}^n q_{Pw}^2 - \frac{1}{2} \sum_{w=1}^n t_{iw}^{\prime 2} \right) \\
&\quad - \left(\sum_{w=1}^n q_{Pw} \cdot t'_{jw} - \frac{1}{2} \sum_{w=1}^n q_{Pw}^2 - \frac{1}{2} \sum_{w=1}^n t_{jw}^{\prime 2} \right) \\
&= \frac{1}{2} \left[\sum_{w=1}^n (q_{Pw} - t'_{jw})^2 - \sum_{w=1}^n (q_{Pw} - t'_{iw})^2 \right] \\
&= \frac{1}{2} \left(\text{dist}_{q_P \leftrightarrow t'_j}^2 - \text{dist}_{q_P \leftrightarrow t'_i}^2 \right).
\end{aligned}$$

Since for all $j = 1, \dots, m$, $j \neq z$, we have $\sum_{w=1}^n (q_{Pw} - t'_{zw})^2 < \sum_{w=1}^n (q_{Pw} - t'_{jw})^2$. Obviously, for all $j = 1, \dots, m$, $j \neq z$, $\Lambda_{zj} = (1/2)(\text{dist}_{q_P \leftrightarrow t'_j}^2 - \text{dist}_{q_P \leftrightarrow t'_z}^2) > 0$ and t_z , $z \in \{1, \dots, m\}$, is the trait vector of the diagnosed disease. The report r_z corresponding to the trait vector t_z is the diagnosed report, which matches the relationship in the real execution.

Now, S_1 can output the simulated view,

$$\begin{aligned}
\{S_1(q_P; r_z)\} \\
= \left\{ q_P, s'_P, B'_P, \{\Lambda'_{ij}\}_{i,j=1,\dots,m,i \neq j}, \{S_1^{\text{OT}_1^m}(z; r_z)\}, r_z \right\}. \quad (21)
\end{aligned}$$

Since s'_P and B'_P are randomly chosen and $\sum_{u=1}^k s'_{Pu} \cdot b'_{uw} = 1$, $w = 1, \dots, n+2$, as s_P and B_P in the real view $\{\text{VIEW}_1(q_P, \{d_i\}_{i=1,\dots,m})\}$. Due to the randomness and relationship, it is easy to find that (s'_P, B'_P) is indistinguishable to (s_P, B_P) . From the construction process of $\{\Lambda'_{ij}\}_{i,j=1,\dots,m,i \neq j}$, it is obvious to conclude that $\{\Lambda'_{ij}\}_{i,j=1,\dots,m,i \neq j}$ is indistinguishable to the set $\{\Lambda'_{ij}\}_{i,j=1,\dots,m,i \neq j}$. Combined with (17), we have

$$\{S_1(q_P; r_z)\} \cong \{\text{VIEW}_1(q_P, \{d_i\}_{i=1,\dots,m})\}. \quad (22)$$

□

6. Conclusions

In this paper, we consider the problem of how to securely make diagnosis without leaking patient's health data, diagnosed result, and hospital's disease database in wireless sensor networks. By applying the idea of secure two-party computation and the technology of oblivious transfer, we propose a privacy-preserving self-helped medical diagnosis scheme so that patients can privately diagnose themselves by inserting a health card into a self-helped MD-ATM to obtain the diagnostic report just like drawing money from a bank ATM. We also have a detailed analysis about the correctness and further strictly prove the security following the real-idea simulation paradigm. We expect to provide people another perspective on future medical care.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is supported by NSFC (Grant nos. 61300181, 61272057, 61202434, 61170270, 61100203, and 61121061) and the Fundamental Research Funds for the Central Universities (Grant nos. 2012RC0612 and 2011YB01).

References

- [1] J. Wan, S. Ullah, C. Lai et al., "Cloud-enabled wireless body area networks for pervasive healthcare," *IEEE Network*, vol. 27, no. 5, pp. 56–61, 2013.
- [2] C. Doukas, T. Pliakas, and I. Maglogiannis, "Mobile healthcare information management utilizing Cloud Computing and Android OS," in *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '10)*, pp. 1037–1040, September 2010.
- [3] J. Yuan and S. Yu, "Efficient privacy-preserving biometric identification in cloud computing," in *Proceedings of the 32nd IEEE International Conference on Computer Communications (INFOCOM '13)*, pp. 2652–2660, 2013.
- [4] J. Liu, Q. Wang, J. Wan, J. Xiong, and B. Zeng, "Towards key issues of disaster aid based on wireless body area networks," *KSII Transactions on Internet and Information Systems*, vol. 7, no. 5, pp. 1014–1035, 2013.
- [5] M. Al Ameen, J. Liu, and K. Kwak, "Security and privacy issues in wireless sensor networks for healthcare applications," *Journal of Medical Systems*, vol. 36, no. 1, pp. 93–101, 2012.
- [6] P. M. Schwartz and J. R. Reidenberg, *Data Privacy Law: A Study of United States Data Protection*, LEXIS Law, Dayton, Ohio, USA, 1996.
- [7] H. Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life*, Stanford University, Stanford, Calif, USA, 2010.
- [8] C. Gentry, *A fully homomorphic encryption scheme [doctoral dissertation]*, Stanford University.
- [9] Z. Brakerski and V. Vaikuntanathan, "Efficient fully homomorphic encryption from (standard) LWE," in *Proceedings of the IEEE 52nd Annual Symposium on Foundations of Computer*

- Science (FOCS '11)*, pp. 97–106, Palm Springs, Calif, USA, October 2011.
- [10] A. López-Alt, E. Tromer, and V. Vaikuntanathan, “On-the-fly multiparty computation on the cloud via multikey fully homomorphic encryption,” in *Proceedings of the 44th Annual ACM Symposium on Theory of Computing (STOC '12)*, pp. 1219–1234, ACM, May 2012.
 - [11] J. Bringer, H. Chabanne, and A. Patey, “Privacy-preserving biometric identification using secure multiparty computation: an overview and recent trends,” *IEEE Signal Processing Magazine*, vol. 30, no. 2, pp. 42–52, 2013.
 - [12] A. C. Yao, “Protocols for secure computations,” in *Proceedings of the 23rd Annual IEEE Symposium on Foundations of Computer Science*, pp. 160–164, Chicago, Ill, USA, 1982.
 - [13] O. S. Goldreich, S. Micali, and A. Wigderson, “How to play any mental game,” in *Proceedings of the 19th Annual ACM Symposium on Theory of Computing (STOC '87)*, pp. 218–229, ACM, New York, NY, USA, 1987.
 - [14] O. S. Goldreich, “Secure multiparty computation,” Manuscript, Preliminary version, 1998.
 - [15] O. Goldreich, *Foundations of Cryptography: Volume 2, Basic Applications*, Cambridge University Press, Cambridge, UK, 2004.
 - [16] M. M. Prabhakaran and A. Sahai, Eds., *Secure Multiparty Computation*, IOS Press, 2013.
 - [17] D. Chaum, C. Crepeau, and I. Damgård, “Multi-party unconditionally secure protocols (extended abstract),” in *Proceedings of the STOC*, pp. 11–19, ACM, 1988.
 - [18] I. Damgård, V. Pastro, N. P. Smart, and S. Zakarias, “Multiparty computation from somewhat homomorphic encryption,” in *Advances in Cryptology—Crypto 2012*, vol. 7417 of *Lecture Notes in Computer Science*, pp. 643–662, Springer, 2012.
 - [19] Y. Lindell and B. Pinkas, “An efficient protocol for secure two-party computation in the presence of malicious adversaries,” in *Advances in Cryptology—EUROCRYPT 2007*, vol. 4515 of *Lecture Notes in Computer Science*, pp. 52–78, Springer, Berlin, Germany, 2007.
 - [20] B. Pinkas, T. Schneider, N. P. Smart, and S. C. Williams, “Secure two-party computation is practical,” in *Advances in Cryptology—ASIACRYPT 2009*, vol. 5912 of *Lecture Notes in Computer Science*, pp. 250–267, Springer, Berlin, Germany, 2009.
 - [21] M. O. Rabin, “How to exchange secrets by oblivious transfer,” Tech. Rep. TR-81, Aiken Computation Laboratory, Harvard University, 1981.
 - [22] S. Even, O. Goldreich, and A. Lempel, “A randomized protocol for signing contracts,” *Communications of the ACM*, vol. 28, no. 6, pp. 637–647, 1985.
 - [23] M. Naor and B. Pinkas, “Oblivious transfer with adaptive queries,” in *Advances in Cryptology—CRYPTO'99*, vol. 1666 of *Lecture Notes in Computer Science*, pp. 573–590, Springer, 1999.