



PHDA: A priority based health data aggregation with privacy preservation for cloud assisted WBANs



Kuan Zhang^{a,*}, Xiaohui Liang^a, Mrinmoy Baura^a, Rongxing Lu^b, Xuemin (Sherman) Shen^a

^a Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

^b School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore 639798, Singapore

ARTICLE INFO

Article history:

Received 26 August 2013

Received in revised form 26 May 2014

Accepted 4 June 2014

Available online 20 June 2014

Keywords:

Wireless body area network

Cloud

Priority

Aggregation

Privacy preservation

ABSTRACT

Wireless Body Area Networks (WBANs), as a promising health-care system, can timely monitor human physiological parameters. Due to the limitation of communications, power, storage and computation in WBANs, a cloud assisted WBAN flourishes and provides more reliable, real-time, and intelligent health-care services for patients and mobile users. However, it is still critical to efficiently aggregate the different types of WBAN data to the cloud server. In addition, security and privacy concerns are also of paramount importance during the communications between WBAN and cloud. In this paper, we propose a priority based health data aggregation (PHDA) scheme with privacy preservation for cloud assisted WBANs to improve the aggregation efficiency among different types of health data. Specifically, we first explore social spots to help forward health data and enable users to select the optimal relay according to their social ties. According to different data priorities, the adjustable forwarding strategies can be selected to forward the user's health data to the cloud servers with the reasonable communication overheads. The security analysis demonstrates that the PHDA can achieve identity and data privacy preservation, and resist the forgery attacks. Finally, the performance evaluation shows that the PHDA achieves the desirable delivery ratio with reasonable communication costs and lower delay for the data in different priorities.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Wireless Body Area Networks (WBANs) which can real-timely monitor patients or users' health status play an essential role in health-care systems as the phenomenon of aging population and the demands of remote health monitoring in our daily life [24]. WBANs provide a variety of services in diverse fields including medical or personal health monitoring, consumer electronics, entertainment, sports or fitness, and military applications. Different physiology parameters, such as temperature, blood pressure, and electrocardiography (ECG) can be collected by WBANs [18]. With the increasing demands from customers and patients, the sensing data is required to be timely processed and the feedback from the doctors is also desirable. Since it requires more network resources, i.e., storage, computation and communication power, it is difficult to achieve these goals only relying on the traditional WBANs [9]. Therefore, the cloud computing is introduced to assist WBANs to store and process the sensing data in a real time fashion.

* Corresponding author. Tel.: +1 5198884567.

E-mail addresses: k52zhang@bcr.uwaterloo.ca (K. Zhang), x27liang@bcr.uwaterloo.ca (X. Liang), mbarua@bcr.uwaterloo.ca (M. Baura), rxlu@ntu.edu.sg (R. Lu), xshen@bcr.uwaterloo.ca (X. (Sherman) Shen).

Taking the advantage of the cloud server to store the large volume of sensing data and process them for doctor's diagnosis [2,8], cloud assisted WBANs become more robust and provide the desirable services for patients and users. For example, in a gym or conference environment, many people have some social activities [27], and wear WBANs to sense their health data and to periodically report them to the cloud servers. The hospital or doctors to access the data stored in the cloud servers in a real time pattern. Then, the doctors (trusted authorities) are able to timely detect the abnormal phenomenon and feedback the corresponding diagnosis. Once a user has an emergency, WBANs can help him call the hospital and continuously upload the real-time health data. However, when a large number of users located at the same place upload their data at the same time, the connection between WBANs and cloud servers might be intermittent. The available bandwidths from WBANs to cloud servers for each individual user are also limited so that the network performance is considerably degraded. Therefore, the communications between WBANs and cloud servers is the bottleneck with the perspective of efficiency and reliability.

Some existing research works [13,15] utilize cooperation among users to improve the reliability. Recent emergency call schemes [12] for health-care applications usually adopt the epidemic dissemination to deliver the general emergency information to the cloud server or hospitals. Even though it can guarantee the emergency call's delivery ratio and minimal delay, the communication costs are still very high. In the above example, some detailed physiology parameters of the patients with the emergency should be continuously uploaded to the cloud server for the further diagnosis and monitoring. If this portion of data is still epidemically disseminated in the network, it consumes an extreme large number of network resources. Therefore, the health data should be classified into different categories with different requirements (i.e., delay) and communication strategies. As communications are deeply involved in cloud assisted WBAN, security and privacy are of paramount importance [23]. All the data transmitted in health-care applications should be authenticated and secure against malicious modification. For example, an attacker might forge a fake emergency call and make it distributed in the network to degrade the network performance. In addition, privacy is also a primary concern from customers point of view, as health data is highly relevant to users themselves, for example, the ECG can reflect people's some specific behaviors, such as sleeping, having meals etc. As a result, the reveal of such health data might violate user's privacy. Therefore, how to efficiently aggregate different types of data and preserve user's privacy is still challenging in cloud assisted WBANs.

In this paper, we propose a priority based health data aggregation scheme (PHDA) with privacy preservation for cloud assisted WBANs to reduce the aggregation overheads and preserve user privacy. The health data is divided into different types, and each type of data is assigned a specific priority. When a user wants to upload his data, he can select different forwarding strategies according to his data's priority. The intuition is that the data with higher priority can be forwarded in a smaller delay. Furthermore, the data with the same priority can be efficiently aggregated which significantly reduces the communication overheads. Specifically, the major contributions of this paper are threefold.

- Firstly, we propose a priority based data aggregation scheme (PHDA) for cloud assisted WBANs. The health data is divided into different types assigned corresponding priorities. Different forwarding strategies are selected according to the data priority. Furthermore, the PHDA enables social spots to help mobile users forward the data to the cloud servers. An eligible relay can be selected based on his social tie to the social spots, which reflects the relay's forwarding capability.
- Secondly, we investigate a lightweight privacy-preserving aggregation scheme with aggregate authentication. The cloud servers can only learn the statistical information without knowing the exact data of individual user. The proposed aggregate authentication scheme can validate the data priority for users which resists the forgery attack, while reducing the authentication overhead.
- Finally, we provide the security and privacy analysis to show that the PHDA can achieve identity and data privacy preservation and resist the forgery attack. In additional, the performance evaluation shows that the PHDA satisfies the delay and delivery ratio requirements for the data with different priorities and consumes lower communication overheads compared with other schemes.

The remainder of this paper is organized as follows: The related works are investigated in Section 2. Network model and design goals are presented in Section 3. In Section 4, we propose the detailed PHDA, followed by the security analysis and performance evaluation in Sections 5 and 6, respectively. Finally, Section 7 concludes the paper.

2. Related work

Recently, there are several research works [1,28] on data forwarding different applications. An effective approach is to pre-deploy some fixed nodes in the network to help mobile users forward their data. Aviv et al. [1] investigate the human mobility patterns and propose a forwarding protocol, named Return-to-Home, which enables the fixed social spots to help mobile users store-and-forward the packets and improves the forwarding efficiency. Lu et al. [15] propose a social spot aided packet forwarding protocol (SPRING) in vehicular ad hoc networks. The SPRING follows the Return-to-Home principle and preserves user privacy at the same time. Zhang et al. [28] investigate a novel social spot deployment to preserve the location privacy for both the users and social spots. Despite the social spots, our PHDA also enables mobile users to help forward the data to social spots so that the data forwarding efficiency is further improved. In addition, some research efforts are paid to investigate data forwarding in health-care systems. Borrego et al. [7] investigate a new paradigm, called store-carry-process-and-forward, based on mobile code to improve the integration of wireless sensor networks and grid computing

infrastructures. Liang et al. [12] propose a privacy-preserving emergency call scheme, named PEC, for mobile health-care social networks. The PEC exploits the epidemic dissemination for emergency call and provides fine-grained access control on the emergency data. In terms of aggregation, Yager [26] propose the priority based data aggregation scheme with multiple criteria aggregation. In [26], the trade-off between the data priority and satisfaction to criteria is investigated. Yager also propose two schemes to formulate the priority based aggregation with multiple criteria. Misra et al. [16] consider the bandwidth shifting and redistribution problems for mobile cloud with the QoS guarantee. They introduce the gateway to aggregate the demands from the mobile users, and formulate it as an utility maximization problem. Misra et al. [17] propose a lightweight energy-efficient routing scheme for wireless sensor network to increase the network life time. The neighbor nodes with higher energy aggregates the data from other nodes and forwards the aggregated data packet to the destination.

Privacy-preserving aggregation schemes are also widely investigated in recent years. Shi et al. [21] introduce a privacy-preserving aggregation of time series data which slices the data to mix them together and confines the aggregator's decryption capability, where it enables the aggregator to only decrypt the sum of the data without learning any exact data value. Lu et al. [14] utilize the increasing sequence to mix the user's multi-dimensional data together which reduces the communication and computation overheads for the aggregation. Shi et al. [22] also present a privacy-preserving aggregation scheme which supports a wide range of statistical additive and non-additive aggregation functions. Further, it can resist the collusion attack during the aggregation. To improve the robustness of the privacy-preserving aggregation, Chan et al. [10] upgrade the existing aggregation with fault tolerance. The TA assigns the aggregator N capabilities corresponding to the N low level users. The fault tolerant aggregation scheme explores a binary tree and establishes groups for the low level users to improve the robustness. However, the overheads during the user-aggregator communications are not negligible when multi-hop forwarding is involved in cloud assisted WBANs. Therefore, the priority based privacy-preserving aggregation scheme is very important in terms of efficiency.

3. Problem definition

In this section, we first describe the network model and identify our primary design goals to establish a reliable and secure connection between WBANs and cloud servers. Then, we present the security model and illustrate the security requirements.

3.1. Network model

We consider a cloud assisted WBAN consisting of a trusted authority, L social spots, a small portion of semi-trusted cloud servers, and N mobile patients/users as shown in Fig. 1. The details of these entities are presented as follows.

- **Trusted Authority (TA)** is a trustable, powerful, and storage-rich entity, and bootstraps the whole system in the initialization phase. In the real world, the TA could be a certificated hospital having the responsibility to manage the users' health data. When bootstrapping, the TA generates secret keys for each legitimate user, and users' certificates for further authentication. After the aggregation, the TA can decrypt the data from each individual user for diagnosis. Upon receiving attack reports from residential users, the TA revokes the malicious users and adjusts the users' encryption keys.

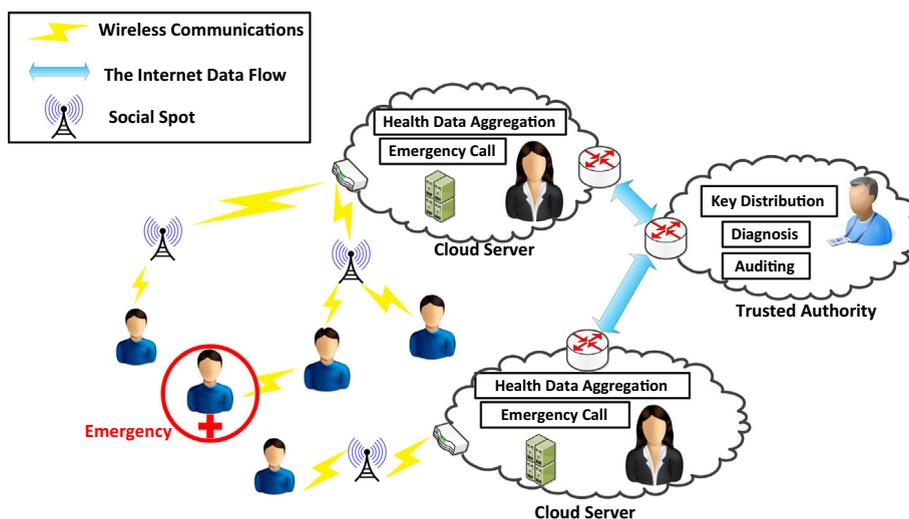


Fig. 1. Network model for cloud assisted WBAN.

- **Social Spot (SP)** is a pre-deployed local gateway and equipped with storage-rich and powerful communication devices. According to the user's behaviors, totally L social spots are located at the intersections or spots where a large portion of mobile users visit frequently. The SP directly collects the health sensing data from each individual user via WBAN communications. Finally, the SPs upload the aggregated data to the cloud servers via the Internet.
- **Cloud Server (CS)** stores the large volume of health sensing data from mobile users, and processes some data, such as ECG, to produce the useful information for doctor's diagnosis. Since some third parties, e.g., insurance company, can access the CS for query and some other operations, the cloud server is a semi-trusted entity in cloud assisted WBAN. To achieve data confidentiality and user's privacy, the data stored in the CSs are of ciphertexts.
- **Mobile users** are denoted by $\mathbb{U} = \{u_1, u_2, \dots, u_N\}$. Each mobile user is equipped with body area sensors which monitors the personal health sensing data in a real time fashion and periodically uploads these health data to the CS via the user's smartphone or PDA [29]. At the beginning, the individual user u_i should firstly register to the TA for the profiles (unique identity), certificates and key materials. Then, u_i should keep them secure and generate session keys in each time slot. When u_i obtains sensing data or faces an emergency, u_i only needs to forward the corresponding data to any one of SPs.

3.2. Security model

Malicious users might exist in the network and launch attacks to violate legitimate user's identity and data privacy, and degrade the network performance. Some inside users might forge the data priority, i.e., making a fake emergency call, or increasing their data priority, to degrade the network performance. Furthermore, the cloud server is semi-trusted, and some third parties might launch attacks on the cloud servers to violate user's data privacy.

3.3. Design goals

Our design goal is to develop a priority based privacy-preserving health data aggregation scheme for cloud assisted WBANs to improve the aggregation efficiency.

3.3.1. Efficiency goals

We intend to reduce the communication overhead of the aggregation, and guarantee the delivery ratio and delay according for the data in different priorities. The health sensing data should be classified into different types with specific priorities. For different data priorities, the data forwarding strategies should be different and maximize the network resource usage with satisfaction of the minimum requirements.

3.3.2. Security goals

Our primary security goal is to protect the individual user's data from disclosure and resist the forgery attack.

- **Data Privacy:** The proposed scheme should not only refine the cloud server's decryption capability [3] but also protect the user's data from eavesdropping during the communications. Therefore, the individual user's data privacy should be protected.
- **Identity Privacy:** The legitimate users might not want to disclose their unique identity information, especially when they are close to the social spots. Therefore, the proposed scheme should be able to prevent the malicious users or attackers from identifying them.
- **Resistance to Forgery Attack:** A malicious user could forge a false emergency call, or increase his data priority so that his data can be preferentially uploaded to the cloud server. The proposed scheme should be able to detect the forged data and block them in the network.

4. Proposed PHDA protocol

In this section, we first provide an overview of the PHDA. Then, we present the details of our proposed PHDA scheme, which mainly consists of initialization, health data generation, and priority based data aggregation.

4.1. Overview of PHDA

To efficiently aggregate user's health sensing data, we investigate the priority based data forwarding in cloud assisted WBANs. Towards a variety of sensing data, different forwarding strategies should be selected to not only forward data within the given delay but also consumes the reasonable network resources.

First of all, we classify the health data into three categories: emergency call, vital health data, and regular health data. As depicted in Table 1, the emergency call is the highest priority data and should be successfully delivered to the cloud server as fast as possible. In addition, the time line is divided into many small time periods. At the beginning of each period, every user obtains his/her health data from the wearable WBANs. The vital health data are the requested data by doctors for continuous monitoring on the user with the emergency. The regular data are not for the emergency user so that the delay requirement is not that tough. Usually, they should be delivered to the cloud server before the next time period. We further divide the vital

Table 1
Data priority in cloud assisted WBANs.

Priority	Data category	Data size
P ₅	Emergency call	Small
P ₄	Vital physiology parameter	Small
P ₃	Vital image data	Large
P ₂	Regular physiology parameter	Small
P ₁	Regular image data	Large

and regular data into small data and big data. The small data, such as physiology parameters with the size of 10–100 bytes, should be delivered to the cloud server within a given delay. On the other hand, the big data, such as ECG or images, are of large size, and should be uploaded in time without consuming too much network resources (see Table 2).

A mobile user u_i sets his data priority with the data priority detection module as shown in Fig. 2. According to the data priority, u_i has different forwarding strategies to forward his data. With our proposed relay selection algorithm, the optimal relay can be selected for different data priorities. When the mobile users visits any one of the pre-deployed social spots, the data can be forwarded to SPs and finally uploaded to the cloud servers (see Fig. 3).

For the cloud server, the CS first authenticates and classifies the aggregated data. Then, the data can be accessed by different entities, including hospital, doctors, insurance companies. The doctors can send request to some specific mobile users via the CSs for vital data monitoring. With the request from the doctors, the mobile users can make their vital data verified when forwarding them to other relays.

Specifically, the PHDA can be proceeded in the following phases: initialization phase, health data generation, priority based data aggregation, and data decryption.

4.2. Initialization

The TA, who could be the authorized hospital or health-care center, initializes the network and audits the aggregated data. We utilize bilinear pairing [6] and Paillier cryptograph [19] to achieve the privacy-preserving aggregation.

Let \mathbb{G}, \mathbb{G}_1 be two additive cyclic groups of the same prime order q , and P be the generator of \mathbb{G} . There is a bilinear map $e: \mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}_1$. A bilinear pairing exists if it is computationally efficient for $e(aP_1, bP_2) = e(P_1, P_2)^{ab} \in \mathbb{G}_1$ for any $P_1, P_2 \in \mathbb{G}$ and all $a, b \in \mathbb{Z}_q^*$, and $e(P, P) \neq 1_{\mathbb{G}_1}$. A bilinear key generation algorithm $Gen(\kappa)$ is used to produce the key materials, where κ is the system security parameter of bilinear pairing.

During the system initialization, the TA first generates $(q, P, \mathbb{G}, \mathbb{G}_1, e)$ by running the key generation algorithm $Gen(\kappa)$. Then, the TA selects the Paillier cryptographic security parameter κ' and two large primes p', q' where $|p'| = |q'| = \kappa'$. The public keys of Paillier cryptograph are: (1) $n = p' \cdot q'$ and (2) $g \in \mathbb{Z}_{n^2}^*$ as the generator. The secret keys are: (1) $\lambda = \text{lcm}(p' - 1, q' - 1)$ where lcm the least common multiple of $p' - 1$ and $q' - 1$ and (2) $\mu = \frac{1}{L(g^2 \bmod n^2)} \bmod n$ where L is a defined function and $L(x) = \frac{x-1}{n}$.

Suppose the maximum number of health data with each priority from N users is smaller than a constant ϕ . The data value for each priority is less than a constant θ . Then, the TA builds up a superincreasing sequence [14] $\vec{b} = (b_1 = 1, b_2, \dots, b_N)$, where b_i is a large prime, the length $|b_i| \geq \kappa$. $\sum_{j=1}^{i-1} b_j \cdot \phi \cdot \theta < b_i$ for $i = 2, \dots, N$, and $\sum_{j=1}^N b_j \cdot \phi \cdot \theta < n$. Similarly, the TA builds up another superincreasing sequence $\vec{a} = (a_1 = 1, a_2, \dots, a_5)$, where a_2, \dots, a_5 are large primes and the length $|a_i| \geq \kappa$. Let $\sum_{i=1}^N b_i = \gamma$. We have $\sum_{j=1}^{i-1} a_j \cdot \gamma \cdot \theta < a_i$ for $i = 2, \dots, 5$, and $\sum_{j=1}^5 a_j \cdot \gamma \cdot \theta < n$. Finally, the TA obtains (g_1, g_2, \dots, g_5) where $g_i = g^{a_i}$ for $i = 1, 2, \dots, 5$.

Afterwards, the TA selects a random number $x \in \mathbb{Z}_q^*$ to compute $Y = xP$ as the public key. $H: \{0, 1\}^* \rightarrow \mathbb{G}$ and $H_1: \{0, 1\}^* \rightarrow \mathbb{Z}_q^*$ are cryptographic hash functions. The secret keys are $\{\lambda, \mu, \vec{a}, \vec{b}, \alpha, x\}$. The public keys are $\{q, P, \mathbb{G}, \mathbb{G}_1, e, n, g_1, g_2, g_3, g_4, g_5, Y, H, H_1\}$.

Table 2
Frequently used notations.

Notation	Definition
B_i	Buffer size of user u_i
d_{ij}	Data of Priority j from user u_i
EM_i	Emergency on user u_i
$ P_{j,i} $	Data size of Priority j from user u_i
ST_{u_i}	Social spot tie of u_i
TH_ν	Threshold of forwarding P_4 and P_3 data
TH_r	Threshold of forwarding P_3 and P_2 data
TH_E	Threshold of forwarding emergency call

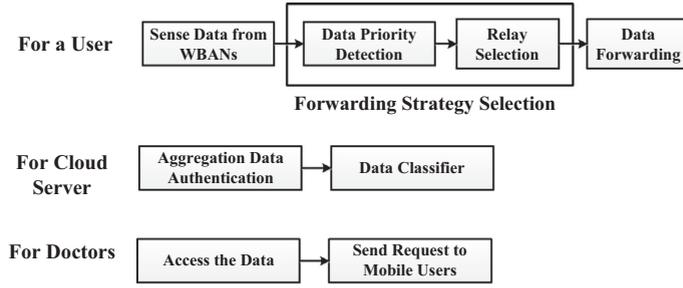


Fig. 2. Overview of PHDA.

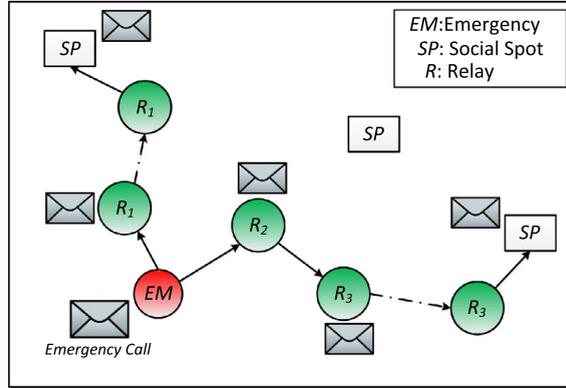


Fig. 3. Forwarding emergency call.

When a user u_i registers to the TA, u_i obtains his secret key b_i . With the multiple pseudonym techniques [11], u_i is also assigned with a set of asymmetric key pairs and uses the alternatively changing public keys as the user's pseudonyms PID_i for the communications. The unique identity u_i can be protected as only literally-meaningless pseudonyms are exposed to the public. u_i selects a random number $x_i \in \mathbb{Z}_q^*$ as his private key.

4.3. Health data generation

WBANs worn on or in the user's body sense the physiology parameters and some large-size sensing data (i.e., ECG). The user u_i should forward the data with a specific priority to the cloud server within a given deadline which is the maximum delay for the specific data priority. Here, the data $(d_1, d_2, d_3, d_4, d_5)$ are generated with different priorities. u_i first chooses a random number $r_i \in \mathbb{Z}_q^*$ and computes

$$C_{ij} = g_j^{b_i d_{ij}} \cdot r_i^n \text{ mod } n^2, \text{ where } j \in \{1, 2, 3, 4, 5\}. \tag{1}$$

Here, $j \in \{1, 2, 3, 4, 5\}$ is the priority number. Note that the ciphertexts for different data priorities cannot be combined together because the forwarding strategies are different. But the ciphertext from the same data priority can be combined together. u_i then signs on the data with his private key x_i to generate the signature. For the regular data, u_i records the system time and makes the signature R_{ij} as

$$R_{ij} = x_i H(C_{ij} || PID_i || Time) \text{ mod } n^2 \tag{2}$$

Regarding the vital data, the TA chooses a random number $s \in \mathbb{Z}_q^*$, and computes $S = sP$. Then, the TA sends $REQ_i || S$ to user u_i where $REQ_i = H_1(DataType || Time)$. With REQ_i , u_i can authenticate his data priority as the vital level (P3 or P4). u_i makes the signature V_{ij} on the vital data d_{ij} with his private key x_i as

$$V_{ij} = x_i S + x_i H_1(C_{ij}) Y. \tag{3}$$

4.4. Priority based data aggregation

After the data generation, a user u_i wants to forward his data as soon as possible. From the view of the network, we have to balance the traffic and optimize the network recourses. Therefore, we propose a priority based data aggregation scheme to not only guarantee the forwarding delay but also reduce the communication overheads. We provide the different forwarding strategies for the data with different priorities.

Algorithm 1. Relay Selection

```

1: Two users  $u_s$  and  $u_r$  are encountered, and  $u_s$  has an emergency.
2: if  $ST_{u_r} > ST_{u_s}$  OR  $ST_{u_s} - ST_{u_r} > TH_E$  then
3:    $u_s$  forwards the emergency call  $\mathbb{EM}_s$  to  $u_r$  AND  $u_r$  verifies the emergency call  $\mathbb{EM}_s$ .
4:   if  $\mathbb{EM}_s$  is valid then
5:      $u_r$  stores  $\mathbb{EM}_s$  and forwards it to AP or another relay if possible AND  $B_r = B_r - |\mathbb{EM}_s|$ .
6:     if  $u_s$  has  $P_4$  data AND  $B_r > 0$  then
7:       if  $ST_{u_r} > ST_{u_i}$  OR  $ST_{u_i} - ST_{u_r} < TH_v$  then
8:          $u_s$  forwards its  $P_4$  data to  $P_{2,s}$   $u_r$  AND  $B_r = B_r - |P_{4,s}|$ .
9:       end if
10:      end if
11:      if  $u_s$  has  $P_3$  data AND  $B_r > |P_{3,s}|$  then
12:        if  $ST_{u_r} > ST_{u_i}$  OR  $ST_{u_i} - ST_{u_r} < TH_v$  then
13:           $u_s$  forwards its  $P_3$  data  $P_{3,s}$  to  $u_r$  AND  $B_r = B_r - |P_{3,s}|$ .
14:        end if
15:      end if
16:      if  $u_s$  has  $P_2$  data AND  $B_r > 0$  then
17:        if  $ST_{u_r} > ST_{u_i}$  OR  $ST_{u_i} - ST_{u_r} < TH_r$  then
18:           $u_s$  forwards its  $P_2$  data to  $u_r$  AND  $B_r = B_r - |P_{2,s}|$ .
19:        end if
20:      end if
21:      if  $u_s$  has  $P_1$  data AND  $B_r > |P_{5,s}|$  then
22:        if  $ST_{u_r} > ST_{u_i}$  OR  $ST_{u_i} - ST_{u_r} < TH_r$  then
23:           $u_s$  forwards its  $P_1$  data to  $u_r$  AND  $B_r = B_r - |P_{5,s}|$ .
24:        end if
25:      end if
26:    else
27:       $u_r$  reports  $u_s$  as a malicious user to the TA.
28:    end if
29:  end if
30: End Procedure

```

4.4.1. Emergency call

When a user u_i has an emergency event denoted as $\mathbb{EM}_i = (u_i || Des_i || Time || Location)$, where Des_i is the general description of the emergency, u_i sets his data priority as P5 in Table 1. When u_i meets another user u_r , u_i first checks whether the social spot tie ST_{u_r} is larger than u_i 's or the difference between ST_{u_r} and ST_{u_i} is less than the threshold of emergency call TH_E . If one of the conditions holds, u_r is selected as an emergency relay. Then, u_i makes short group signature [5,4] $G \cdot \text{sign}(u_i)$ and forwards $(\mathbb{EM}_i || G \cdot \text{sign}(u_i))$ to u_r .

Receiving $(\mathbb{EM}_i || G \cdot \text{sign}(u_i))$, u_r first checks the signature $G \cdot \text{sign}(u_i)$ with $G \cdot \text{verify}(G \cdot \text{sign}(u_i))$. Here, we use $G \cdot \text{sign}$ and $G \cdot \text{verify}$ to denote the group signature and verification algorithms. If invalid, u_r reports u_i to the SP or TA. If valid, u_r carries and forwards \mathbb{EM}_i to any social spot SP . Once u_r meets another user u_{r_1} before u_r forwards \mathbb{EM}_i to SP , u_r follows the steps that u_i does and determines whether \mathbb{EM}_i should be forwarded to u_{r_1} or not.

When u_r visits any social spot SP , u_r forwards \mathbb{EM}_i to SP . Since all SP s are connected via the Internet, the data can be successfully uploaded if one SP receives the data.

4.4.2. Vital data forwarding

When u_i is encountered with u_r and has a piece of vital data C_{ij} where $j = 3$ or 4 , u_i checks whether u_r meets the following criteria (1) the social spot tie ST_{u_r} is larger than u_i 's or the difference between ST_{u_r} and ST_{u_i} is less than the threshold (TH_v) of P4 or P3 and (2) the available buffer size of u_r is larger than the transmitting data size. If both conditions hold, u_r is selected as a relay. Then, u_i forwards data to u_r .

After receiving the data, u_r first checks the signature. If invalid, u_r reports u_i to the SP or TA. If valid, u_r computes $C_{rj} = C_{rj} C_{ij} \bmod n^2$, and forwards C_{rj} to any social spot SP if possible. Once u_r meets another user u_{r_1} before u_r forwards the data to SP , u_r follows the steps that u_i does and determines whether the data should be forwarded to u_{r_1} or not.

4.4.3. Regular data forwarding

For P1 and P2 data, u_i also needs to check whether a relay u_r is eligible or not by following: (1) the available buffer size of u_r is larger than u_i 's data and (2) the social spot tie ST_{u_r} of u_r is larger than ST_{u_i} , or $ST_{u_i} - ST_{u_r} \leq TH_r$. If and only if both conditions hold, u_i can forward the data to u_r . The detailed relay selection steps are depicted in Algorithm 1.

4.5. Data aggregation for the cloud server

(1) **Aggregate Authentication:** When the CS receives $M \leq N$ vital data in time period t , the CS first verifies the authenticity of the data. First, the CS computes the sum of all the signatures $\sum_{i=1}^M V_i$, and checks

$$e\left(\sum_{i=1}^M V_i, P\right) \stackrel{?}{=} e\left(S, \sum_{i=1}^M PK_i\right) \cdot \prod_{i=1}^M e(Y, PK_i)^{H_1(C_{ij})}. \quad (4)$$

If Eq. (4) holds, all M vital data packets are authenticated. The correctness can be proved as

$$\begin{aligned} e\left(\sum_{i=1}^M V_{ij}, P\right) &= \prod_{i=1}^M e(x_i S + x_i H_1(C_{ij}) x P, P) = \prod_{i=1}^M e(S, x_i P) \cdot \prod_{i=1}^M e(x H_1(C_{ij}) P, x_i P) = e\left(S, \sum_{i=1}^M PK_i\right) \cdot \prod_{i=1}^M e(x H_1(C_{ij}) P, PK_i) \\ &= e\left(S, \sum_{i=1}^M PK_i\right) \cdot \prod_{i=1}^M e(Y, PK_i)^{H_1(C_{ij})} \end{aligned} \quad (5)$$

Here, S is distributed by the TA in the time period t , PK_i is u_i 's public key, and Y is the TA's public key. The pairing operations $e\left(S, \sum_{i=1}^M PK_i\right)$ and $e(Y, PK_i)$ can be pre-computed. During each aggregate authentication, only one pairing operation is required so that the authentication efficiency is considerably improved.

For the regular data, the TA can do the batch verification to efficiently verify the signatures as

$$e\left(\sum_{i=1}^N R_{ij}, P\right) = e\left(\sum_{i=1}^N x_i H(C_{ij} \| PID_i \| Time), P\right) = \prod_{i=1}^N e(H(C_{ij} \| PID_i \| Time), PK_i) \quad (6)$$

When the CS receives all the data from N mobile users at the end of every time slot (one time period is divided into several time slots), the CS aggregates all the ciphertexts C_{ij} together, and sends $C = \prod_{j=1}^5 \prod_{i=1}^N C_{ij} \bmod n^2$ to the TA. In addition, the CS generates the signature of the aggregated data C as $Sign_{CS} = x_{CS} H(C \| CS \| Time) \bmod n^2$, where x_{CS} is the private key of the CS.

4.6. Data decryption by the TA

After receiving the aggregated data from the CS, the TA first verifies the signature $Sign_{CS}$. The TA checks whether $e(Sign_{CS}, P) \stackrel{?}{=} e(H(C \| CS \| Time), PK_{CS})$. If it holds, the received data are valid.

The TA has the data as

$$\begin{aligned} C &= \prod_{j=1}^5 g_j^{\sum_{i=1}^N} \left(\prod_{i=1}^N r_i\right)^n \bmod n^2 = g_1^{\sum_{i=1}^N} \cdot g_2^{\sum_{i=1}^N} b_i d_{i2} \cdots g_5^{\sum_{i=1}^N} b_i d_{i5} \cdot \left(\prod_{i=1}^N r_i\right)^{5n} \bmod n^2 \\ &= g^{a_1 \sum_{i=1}^N b_i d_{i1}} \cdot g^{a_2 \sum_{i=1}^N b_i d_{i2}} \cdots g^{a_5 \sum_{i=1}^N b_i d_{i5}} \cdot \left(\prod_{i=1}^N r_i\right)^{5n} \bmod n^2 \\ &= g^{a_1 \sum_{i=1}^N b_i d_{i1} + a_2 \sum_{i=1}^N b_i d_{i2} + \cdots + a_5 \sum_{i=1}^N b_i d_{i5}} \cdot \left(\prod_{i=1}^N r_i\right)^{5n} \bmod n^2 \end{aligned} \quad (7)$$

Let $M = a_1 \sum_{i=1}^N b_i d_{i1} + a_2 \sum_{i=1}^N b_i d_{i2} + \cdots + a_5 \sum_{i=1}^N b_i d_{i5}$, and $r = \left(\prod_{i=1}^N r_i\right)^5$, we have $C = g^M r^n \bmod n^2$. The TA can use his secret key (λ, μ) to decrypt the aggregated data according to Paillier cryptograph.

Algorithm 2. Recover the Aggregated Data

-
- 1: Input: $\vec{a} = (a_1 = 1, a_2, \dots, a_N)$ and M
 - 2: Output: D_1, D_2, \dots, D_l
 - 3: Set $X_l = M$
 - 4: **for** $j = l$ to 2 **do**
 - 5: $X_{j-1} = X_j \bmod a_j$
 - 6: $D_j = \frac{X_j - X_{j-1}}{a_j} = \sum_{i=1}^N b_i d_{ij}$
 - 7: **end for**
 - 8: $D_1 = X_1 = \sum_{i=1}^N b_i d_{i1}$
 - 9: Return (D_1, D_2, \dots, D_l)
 - 10: **End Procedure**
-

Having the aggregated data, the TA runs [Algorithm 2](#) to obtain the data for each priority. The correctness of [Algorithm 2](#) can be achieved as

$$X_l = a_1 \sum_{i=1}^N b_i d_{i1} + a_2 \sum_{i=1}^N b_i d_{i2} + \cdots + a_{l-1} \sum_{i=1}^N b_i d_{il} < a_1 \sum_{i=1}^N \theta + a_2 \sum_{i=1}^N \theta + \cdots + a_{l-1} \sum_{i=1}^N \theta = \sum_{j=1}^{l-1} a_j N \theta < a_l \quad (8)$$

Therefore, we have $X_{l-1} = X_l \bmod a_l = a_1 \sum_{i=1}^N b_i d_{i1} + a_2 \sum_{i=1}^N b_i d_{i2} + \cdots + a_{l-1} \sum_{i=1}^N b_i d_{i(l-1)}$ and

$$\frac{X_l - X_{l-1}}{a_l} = \frac{a_l \sum_{i=1}^N d_{il}}{a_l} = \sum_{i=1}^N d_{il} = D_l, \quad \text{where } l = 1, 2, \dots, 5. \quad (9)$$

X_l is the sum for the data with the l -th priority. Similarly, d_{ij} is obtained by the TA.

5. Security analysis

In this section, we discuss the security properties of our proposed PHDA scheme. We focus on the aforementioned security requirements in [Section 3](#).

5.1. Data privacy

The user's data privacy can be achieved based on the assumption that Decisional Diffie–Hellman (DDH) problem or Decisional Bilinear Diffie–Hellman (DBDH) [[25](#)] problem is hard. The passive eavesdropping can be resisted since all the transmitted data are encrypted by Paillier cryptograph. Furthermore, the whole superincreasing sequences \vec{a} and \vec{b} are the secret keys which are securely kept by the TA. Each user u_i can only obtain b_i for the encryption. As a result, the cloud server and other entities including the attackers who do not know all the secret keys cannot recover the exact data for different priorities. Therefore, the user's data privacy can be achieved. Due to the properties of the superincreasing sequence, the TA can recover the data even though they are aggregated together.

5.2. Identity privacy

The user's identity privacy can be preserved with the multiple pseudonym techniques. In each time period, a user u_i changes his pseudonym PID_i to protect his identity privacy. Only the meaningless pseudonyms are exposed to the other users. No entity except the TA can trace the pseudonyms of u_i and link them together. By frequently changing his pseudonyms, u_i can protect his identity privacy due to the unlinkability of the current pseudonym and the previous ones. On the other hand, if u_i launches some attacks, the TA is able to trace u_i 's pseudonyms PID_i and link them together to identify the attacker.

5.3. Resistance to forgery attack

The malicious insider users cannot launch forgery attack to tamper with the data priority since every desired vital sensing data is transmitted with a request $REQ_i || S$ from the TA. All the collected vital data should be verified by the TA and authenticate that the data type is exactly the one the doctors require. If the malicious user \mathcal{U} forges a signature

$$V_{\mathcal{U}} = x_{\mathcal{U}} S' + x_{\mathcal{U}} H_1(C_{\mathcal{U}}) Y. \quad (10)$$

The other users can verify it and have

$$e(V_{\mathcal{U}}, P) \neq e(S, PK_{\mathcal{U}}) \cdot e(Y, PK_{\mathcal{U}})^{H_1(C_{\mathcal{U}})}. \quad (11)$$

Then, \mathcal{U} is drawn to the revocation list by the TA.

In addition, the emergency call cannot be forged by the outside attackers since the group signature is adopted. Only the registered user can obtain the key materials from the TA to produce the valid emergency call signature. If an attacker A forges an emergency call \mathbb{EM}_A , other legitimate users can verify A 's signature with $G.verify(G.sign(A))$ and detect the attack.

In summary, from the above analysis, the PHDA can resist the forgery attack from the inside malicious users and the outside attackers.

6. Performance evaluation

6.1. Computational complexity

We compare the computational complexity of the PHDA with the non-aggregate scheme. In the PHDA, an individual user u_i encrypts the health data with 6 exponentiation operations in \mathbb{Z}_{n^2} . For the signature generation, u_i performs 1 and 2 multiplication operations \mathbb{G} for regular health data and vital data, respectively. The cloud server CS needs to verify the

signatures of the received health data. The vital data verification requires $M + 2$ pairing operations, which are the primary computational costs, and M exponentiation operations in \mathbb{G}_1 , and M multiplication operations in \mathbb{G}_1 . Meanwhile, CS performs $N + 1$ pairing operations and N multiplication operations in \mathbb{G}_1 . When sending the aggregated data to the TA, the CS generates the signature with 1 multiplication operation in \mathbb{G} . The multiplication operations in \mathbb{Z}_{n^2} can be considered negligible compared with the exponentiation, pairing operations. Therefore, the overhead of data aggregation can be negligible. TA verifies the signature with 2 pairing operations, and decrypts the data from CS with 1 exponentiation operations in \mathbb{Z}_{n^2} .

We compare the proposed PHDA scheme with a non-aggregate scheme where the data are directly sent to the TA in the separate type. The individual user u_i needs to separately encrypt 5 types of health data with 10 exponentiation operations in \mathbb{Z}_{n^2} , and generate signatures with 5 multiplication operations in \mathbb{G}_1 . At the TA end, it requires $10N$ pairing operations for verification and $5N$ exponentiation operations in \mathbb{Z}_{n^2} to decrypt the data.

The computational complexity of PHDA and non-aggregate scheme is depicted in Table 3. We denote C_e as the exponentiation operation in \mathbb{Z}_{n^2} , C_m as the multiplication operation in \mathbb{G} , $C_{m,1}$ as the multiplication operation in \mathbb{G}_1 , C_p as the pairing operation. As depicted in Table 3, the computation overhead of the TA is significantly reduced with the assistance of the cloud server.

6.2. Simulation setup

For the simulation, we utilize a real world human trace – Infocom06 [20] trace, where 78 mobile users attend a conference within four days. Every two mobile users' encounter in the proximity can be detected via their attached Bluetooth devices. There are several fixed nodes in the trace, and we use them as the social spots according to their contacts with mobile users. Finally, we select 10 fixed nodes as social spots in our simulation. The contacts of all users and fixed nodes are recorded in the log file. For the simulation, we collect 128,979 useful contacts, and divide them into two portions: the first one third of the data set as a training set producing user's social ties and the residual data as the experiment set used for the simulation. We implement the PHDA and some other schemes under the Matlab simulator to evaluate the performance. Basically, we utilize delivery ratio, average delay and number of copies as metrics for the comparison.

6.3. Simulation results

To evaluate the emergency call's forwarding efficiency of the PHDA, we implement the PHDA, Epidemic and SPRING schemes for comparison. The Epidemic forwarding, which enables every encountered user to forward the data, is also adopted in some other emergency call schemes [12]. The SPRING [15] only relies on mobile users to forward their own data to the social spots. Totally 78 emergency calls are generated randomly. The comparison results shown in Fig. 4 with the comparison among PHDA, Epidemic and SPRING schemes in terms of delivery ratio, average delay and the number of copies. From Fig. 4(a), the delivery ratio of the PHDA is less than that of the Epidemic at the beginning of the emergency event. However, with the PHDA, 85% emergency calls can be successfully forwarded to the servers within 2 min, while the percentage for the Epidemic is around 90%. The PHDA and Epidemic can achieve the same delivery ratio after 6 min and finally reach 100% delivery. Regarding the SPRING, it consumes less communication overhead but cannot achieve the desirable delivery

Table 3
The comparison of computational complexity between PHDA and non-aggregate scheme.

	PHDA	Non-aggregate scheme
Individual user	$6C_e + 3C_m$	$10C_e + 5C_m$
Cloud server	$(M + N + 3)C_p + (M + N)C_{m,1} + C_m$	N/A
TA	$2C_p + C_e$	$10NC_p + 5NC_e$

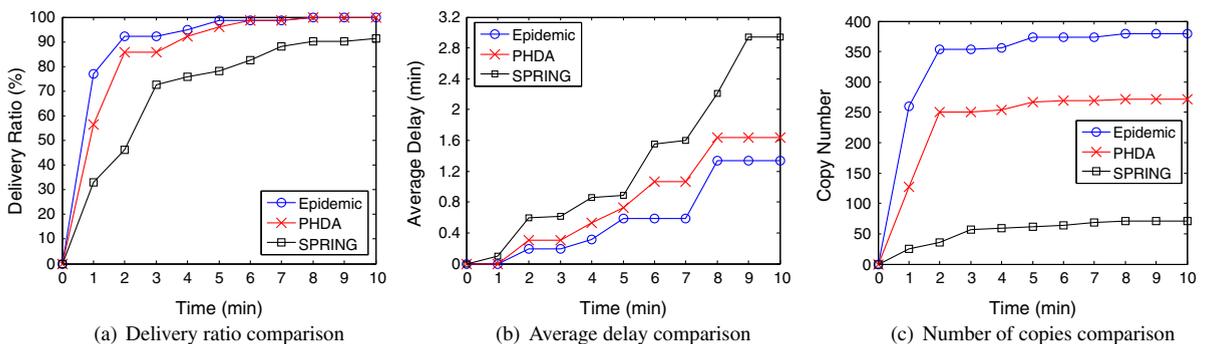


Fig. 4. Emergency call performance between PHDA and epidemic.

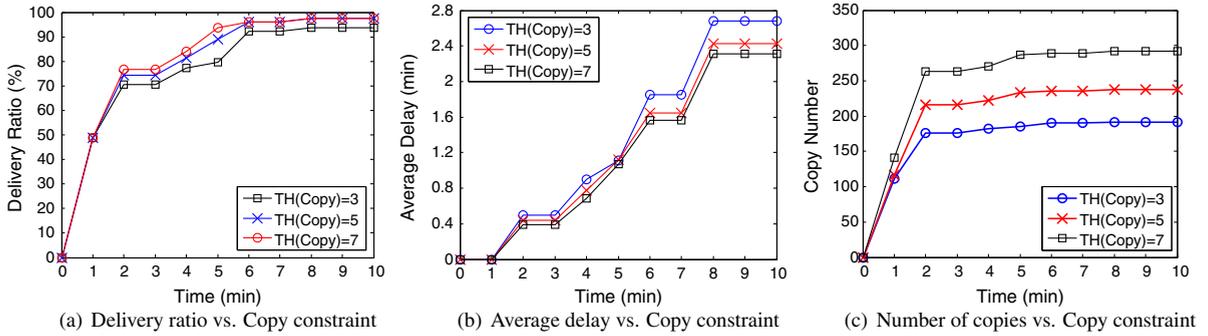


Fig. 5. Impact of copy constraints on performance of PHDA.

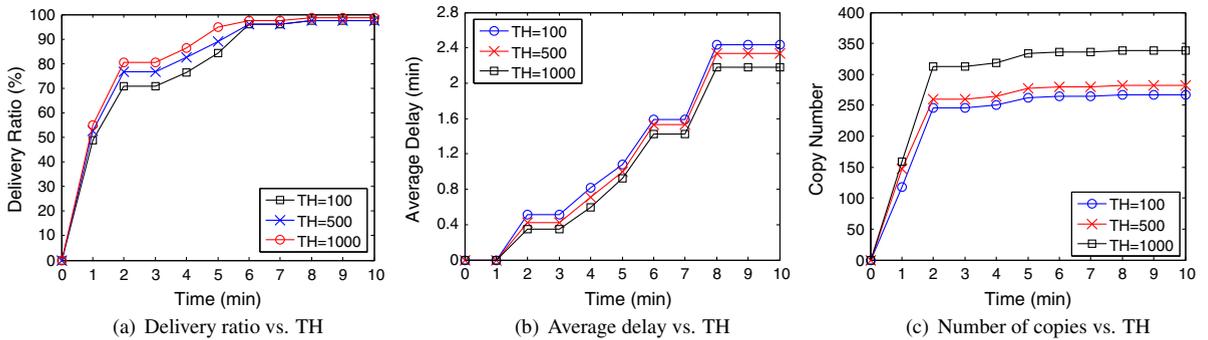


Fig. 6. Impact of TH on performance of PHDA.

ratio which is not suitable for health-care applications. From Fig. 4(c), we can see that the communication overhead of the PHDA is significantly reduced compared with the Epidemic. The reason is that the PHDA utilizes the fixed social spots to help mobile users store-and-forward the data so that the fixed social spots provide more opportunities for mobile users to forward their data. Furthermore, the deployment of the social spots is selected at the location where a lot of mobile users visit frequently. In addition, the PHDA enables the mobile users to select the active mobile users which further improve the connections between the mobile users and social spots. Therefore, the delivery ratio of the PHDA is close to the Epidemic with much lower communication overhead.

In Fig. 5, we show the impact of the copy constraints on the PHDA with a constant social tie constraint TH . Here, the copy constraint is the maximum number of copies that a user can hold. With this constraint, any mobile user cannot take too many copies which significantly save each individual user's storage and energy consumption. Therefore, the network resources are fairly utilized. With a lower copy constraint, for example, at most 3 packets can be held by a user, the delivery ratio is less than that with a higher copy constraint from Fig. 5(a). But after copy constraint reaches 7, the delivery ratio varies a little because the number of eligible relay is bounded by the social tie constraint. On the other hand, with a lower available buffer size (the maximum number of copies), the communication overhead is considerably reduced.

The impact of the social tie threshold TH on the performance of the PHDA is shown in Fig. 6. We set the copy constraint as 5. From Fig. 6(a) and (b), with a larger TH , the PHDA achieves better performance in terms of delivery ratio and average delay. But the improvement is not that high. The number of copies increases when TH is larger from Fig. 6(c). This is because the larger TH causes the increased number of eligible relays which correspondingly increase the number of copies.

7. Conclusions

In this paper, we have proposed a priority based privacy-preserving health data aggregation scheme (PHDA) for cloud assisted WBANs to improve the aggregation efficiency and preserve identity and data privacy. The PHDA utilizes the fixed social spots and the social tie between users and social spots to select the optimal relay and provides reliable data aggregation. With different data priorities, the forwarding strategies are adjustable and the corresponding delay requirements can be satisfied with the minimum communication overheads. The security analysis demonstrates that the PHDA can preserve identity and data privacy, while it also resists the forgery attack from inside malicious users and outside attackers. The performance evaluation shows that the PHDA satisfies the delay and delivery ratio requirements for the data with different priorities, and reduces the communication overheads at the same time. In our future work, we intend to investigate the lightweight homomorphic aggregation scheme to further reduce the communication and computation overheads.

Acknowledgement

This research has been supported by a research grant from the Natural Science and Engineering Research Council (NSERC), and Care In Motion, Canada.

References

- [1] A. Aviv, M. Sherr, M. Blaze, J. Smith, Evading cellular data monitoring with human movement networks, in: *USENIX Workshop on Hot Topics in Security (HotSec)*, 2010, pp. 1–6.
- [2] A. Azadeh, I.M. Fam, M. Khoshnoud, M. Nikafrouz, Design and implementation of a fuzzy expert system for performance assessment of an integrated health, safety, environment (HSE) and ergonomics system: the case of a gas refinery, *Elsevier Inf. Sci.* 178 (22) (2008) 4280–4300.
- [3] M. Barua, X. Liang, R. Lu, X. Shen, ESPAC: enabling security and patient-centric access control for ehealth in cloud computing, *Int. J. Secur. Networks* 6 (2/3) (2011) 67–76.
- [4] D. Boneh, X. Boyen, *Short Signatures Without Random Oracles and the SDH Assumption in Bilinear Groups*, Springer-Verlag, 2008.
- [5] D. Boneh, X. Boyen, H. Shacham, *Short Group Signatures*, 2004. <<http://hovav.net/dist/groupsigs.ps>>.
- [6] D. Boneh, M. Franklin, Identity based encryption from the weil pairing, *IACR Cryptol. ePrint Arch.* 2001 (2001) 90.
- [7] C. Borrego, S. Robles, A store-carry-process-and-forward paradigm for intelligent sensor grids, *Elsevier Inf. Sci.* 222 (2013) 113–125.
- [8] N. Botts, B. Thoms, A. Noamani, T. Horan, Cloud computing architectures for the underserved: public health cyberinfrastructures through a network of healthATMs, in: *Proc. of HICSS*, 2010, pp. 1–10.
- [9] J. Caldeira, J. Rodrigues, P. Lorenz, Toward ubiquitous mobility solutions for body sensor networks on healthcare, *IEEE Commun. Mag.* 50 (5) (2012) 108–115.
- [10] T. Chan, E. Shi, D. Song, Privacy-preserving stream aggregation with fault tolerance, *IACR Cryptol. ePrint Arch.* 2011 (2011) 655.
- [11] J. Freudigery, M. Manshaei, J. Hubaux, D. Parkes, On non-cooperative location privacy: a game-theoretic analysis, in: *Proc. of CCS*, 2009, pp. 324–337.
- [12] X. Liang, R. Lu, L. Chen, X. Lin, X. Shen, PEC: a privacy-preserving emergency call scheme for mobile healthcare social networks, *J. Commun. Networks* 13 (2) (2011) 102–112.
- [13] C. Liu, J. Wen, Q. Yu, B. Yang, W. Wang, HealthKiosk: a family-based connected healthcare system for long-term monitoring, in: *Proc. of IEEE Infocom*, 2011, pp. 241–246.
- [14] R. Lu, X. Liang, X. Li, X. Lin, X. Shen, EPPA: an efficient and privacy-preserving aggregation scheme for secure smart grid communications, *IEEE Trans. Parallel Distrib. Syst.* 23 (9) (2012) 1621–1631.
- [15] R. Lu, X. Lin, X. Shen, SPRING: a social-based privacy-preserving packet forwarding protocol for vehicular delay tolerant networks, in: *Proc. of IEEE INFOCOM*, 2010, pp. 632–640.
- [16] S. Misra, S. Das, M. Khatua, M. Obaidat, Qos-guaranteed bandwidth shifting and redistribution in mobile cloud environment, *IEEE Trans. Cloud Comput.* (in press).
- [17] S. Misra, P. Dias, A simple, least-time, and energy-efficient routing protocol with one-level data aggregation for wireless sensor networks, *J. Syst. Softw.* 83 (5) (2010) 852–860.
- [18] U. Mitra, B. Emken, S. Lee, M. Li, V. Rozgic, G. Thatté, H. Vathsangam, D. Zois, M. Annavaram, S. Narayanan, M. Levorato, D. Spruijt-Metz, G. Sukhatme, KNOWME: a case study in wireless body area sensor network design, *IEEE Commun. Mag.* 50 (5) (2012) 116–125.
- [19] P. Paillier, Public-key cryptosystems based on composite degree residuosity classes, in: *Proc. of EUROCRYPT*, 1999, pp. 223–238.
- [20] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, A. Chaintreau, CRAWDAD trace cambridge/haggle/imote/infocom (v. 2006-01-31).
- [21] E. Shi, T. Chan, E. Rieffel, R. Chow, D. Song, Privacy-preserving aggregation of time-series data, in: *Proc. NDSS*, 2011.
- [22] J. Shi, R. Zhang, Y. Liu, Y. Zhang, PriSense: privacy-preserving data aggregation in people-centric urban sensing systems, in: *Proc. IEEE Infocom*, 2010, pp. 758–766.
- [23] M. Valero, S. Jung, A. Uluagac, Y. Li, R. Beyah, Di-Sec: a distributed security framework for heterogeneous wireless sensor networks, in: *Proc. of IEEE Infocom*, 2012, pp. 585–593.
- [24] H. Viswanathan, B. Chen, D. Pompili, Research challenges in computation, communication, and context awareness for ubiquitous healthcare, *IEEE Commun. Mag.* 50 (5) (2012) 92–99.
- [25] L. Wang, L. Wang, Y. Pan, Z. Zhang, Y. Yang, Discrete logarithm based additively homomorphic encryption and secure data aggregation, *Elsevier Inf. Sci.* 181 (16) (2011) 3308–3322.
- [26] R. Yager, On prioritized multiple-criteria aggregation, *IEEE Trans. Syst. Man Cybern. Part B: Cybern.* 42 (5) (2012) 1297–1305.
- [27] K. Zhang, X. Liang, R. Lu, X. Shen, Exploiting multimedia services in mobile social network from security and privacy perspectives, *IEEE Commun. Mag.* 52 (3) (2014) 58–65.
- [28] K. Zhang, X. Liang, R. Lu, X. Shen, H. Zhao, VSLP: voronoi-socialspot-aided packet forwarding protocol with receiver location privacy in MSNs, in: *Proc. of GLOBECOM*, 2012, pp. 348–353.
- [29] Medical Body Area Networks First Report and Order, 2009. <<http://www.fcc.gov/document/medical-body-area-networks-first-report-and-order>>.