Computer Communications xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

HealthShare: Achieving secure and privacy-preserving health information sharing through health social networks

Xiaohui Liang [a,*], Mrinmoy Barua [a], Rongxing Lu [a], Xiaodong Lin [b], Xuemin (Sherman) Shen [a]

[a] Department of Electrical and Computer Engineering, University of Waterloo, Canada
[b] Faculty of Business and Information Technology, University of Ontario Institute of Technology, Ontario, Canada

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Health social networks
Attribute-oriented authentication
Attribute-oriented transmission

ABSTRACT

In this paper, we propose two attribute-oriented authentication and transmission schemes for secure and privacy-preserving health information sharing in health social networks (HSNs). HSN users are tagged with formalized attributes. The attribute-oriented authentication scheme enables each HSN user to generate an attribute proof for itself, where its sensitive attributes are anonymized. By verifying provided attribute proof, other users are able to know what attributes an HSN user has. The attribute-oriented transmission scheme enables an HSN user to encrypt his/her health information into a ciphertext bonded with a customized access policy. The access policy is defined by a target set of attributes. Only users who satisfy the access policy are able to decrypt the ciphertext. Through security analysis, we show that the proposed schemes can effectively resist various attacks including forgery attack, attribute-trace attack, eavesdropping attack, and collusion attack. Through extensive simulation studies, we demonstrate that both schemes can offer satisfactory performance in helping HSN users to share health information.

1. Introduction

As the ever-increasing demand for a secure and efficient healthcare communication, health social networks (HSNs), a promising and pervasive platform where patients and doctors are able to easily share health information, is widely adopted today, such as Patientslikeme [1], AT&T Healthcare Community Online [2]. HSNs easily and efficiently integrates the collective experience and expertise of patients and doctors by using the healthcare communities built upon social networks. By connecting to HSNs, patients can receive fast and accurate healthcare services from senior doctors, and meanwhile doctors are able to timely check the feedbacks from their patients. The significance of HSNs is to provide patients with social support from individuals who are geographically distant but emotionally and experientially close to the patients. Statistical reports [3,4] indicate that 13 percent of informants benefit from HSNs on their own or another's healthcare management, and 42 percent state that they or someone they know have been helped by following medical advice or using health information from HSNs.

Recent advances in body sensors and wireless communications have revealed the possibility of providing pervasive health monitoring to patients [5–9]. The body sensors deployed in, on or around human body are able to capture physical phenomena from a human body and contextual information from the environment in a situation where large-sized and standard medical examination equipments are not available. Patients are able to use handheld devices to completely control and generate the personal health reports. The provision of such information via HSNs may help their friends and relatives to know what social support the patients need. It is envisioned that with the body sensors largely deployed, HSNs will become the primary communication platform for patients and doctors due to its easy-to-use and pervasive access.

The fundamental principle of most HSN applications is to share health information. An HSN user often needs to share privacy-sensitive health information with the acquaintances that the user trusts and needs to communicate with. A patient may have a need to share either good or bad anonymized health information with the strangers who are able to provide suggestions and reliefs. However, before the health information sharing through HSNs can be adopted by patients or doctors in reality, security and privacy preservation issues must be considered and resolved. Generally, health information sharing involves two security and privacy preservation issues. First, when HSN users receive the information from the networks, they need to estimate the accuracy of an anonymous content and determine how much weight they can rely on the suggestions. Health-related information is very critical to patients' lives; a good advice from a well-behaved doctor may be used to

[*] Corresponding author. Tel.: +1 2268085610.
E-mail addresses: x27liang@bbcr.uwaterloo.ca (X. Liang), mbarua@ecemail.uwaterloo.ca (M. Barua), rxlu@bbcr.uwaterloo.ca (R. Lu), xiaodong.lin@uoit.ca (X. Lin), xshen@bbcr.uwaterloo.ca (X. Shen).

0140-3664/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2012.01.009

improve a patient's health condition significantly, while following an inappropriate instruction from a misunderstood person may put the lives of patients in danger. Therefore, for health information sharing, it is a must to design an authentication scheme that allow receivers to authenticate the sender's credibility. Second, when HSN users share their health information with others, they often make restrictions on who have the authority to read the information. If the health information can be obtained by unauthorized users, the privacy of HSN users would be completely violated and the health information sharing cannot be adopted [10].

In this paper, to achieve secure and privacy preserving health information sharing through HSNs, we propose an attribute-oriented authentication scheme and an attribute-oriented transmission scheme. The main contributions of this paper are three-folds. First, we enable an HSN user to authenticate another user's attributes by proposing an attribute-oriented authentication scheme. In specific, a user first creates an authentication policy to reveal partial of its attributes, and then generates an attribute proof corresponding to this authentication policy. By verifying the attribute proofs, two users are able to know each other's attribute information while their privacy are preserved; second, we propose an attribute-oriented transmission scheme to help HSN users achieving the confidentiality of the shared health information. With the scheme, a user can flexibly choose the eligible receivers by creating an access policy on the attributes. Any unauthorized users are unable to read the information; and third, we evaluate the proposed schemes by extensive simulations. Simulation results show that the proposed schemes can help users to effectively share their health information in HSNs.

The remainder of this paper is organized as follows. In Section 2, network model, adversary model and design goals are introduced. Some preliminaries are presented in Section 3. Then, an attribute-oriented authentication scheme and an attribute-oriented transmission scheme are proposed in Section 4, followed by the security analysis in Section 5. In Section 6, the performance evaluation results of the schemes through simulations are given. Finally, related work and conclusion are presented in Section 7 and Section 8, respectively.

2. Models and design goals

In this section, we define the network model, the adversary model, and the design goals of the proposed schemes.

2.1. Network model

We consider a typical HSN consisting of τ networked social users denoted by $\mathcal{V} = \{n_1, n_2, \ldots, n_\tau\}$, where the communication between any two users are bidirectional. If users n_i and n_j are able to directly communicate with each other, they are each other's social friend. Specifically, the communication can be multiple types, e.g., two geographically-close users may use the handheld devices to setup an end-to-end wireless communication, or two geographically distant users may use Internet to connect each other online.

In the HSN, users can play different roles, such as patients, doctors. Patients may have specific symptoms and diseases, while doctors may have specialties and professional levels as shown in Fig. 1. We denote all these detailed and common characteristics as attributes. At the initial step, an attribute trusted authority (ATA) is needed to formalize the universal attribute set and assign users with the attributes respect to the characteristics they have. Users register to ATA in order to obtain their secret keys corresponding to their attributes. The secret keys can be used to authenticate other users' attributes and access to confidential health information. We denote the universal attribute set by $\mathcal{A}_u = \{a_1, a_2, \ldots, a_l\}$ and the attribute set of user n_i by $\mathcal{A}_i (\subseteq \mathcal{A}_u)$.

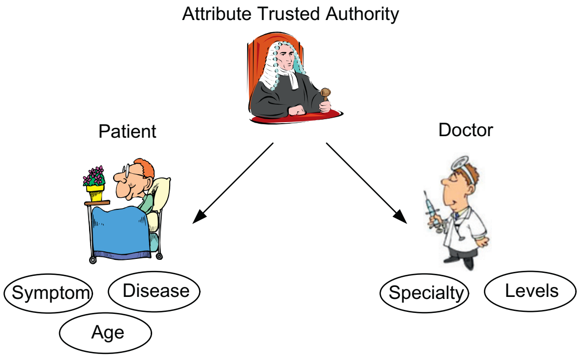


Fig. 1. Attribute trusted authority, patients and doctors.

HSN users have rich social relationships; they autonomously recognize and connect to one another. For example, they may get in touch with others via an online healthcare community, where patients and doctors both share experiences towards a specific disease or symptom; they may make new friendship with another who is introduced by their social friends. Observing the characteristics of such social relations, we consider an attribute-based social links pervasively existing among HSN users. Specifically, if users n_i and n_j are each other's social friends, they must share at least one attribute; if they have more common attributes, it is more likely that they have a social link inbetween. The probability that two users have a attribute-based social link is proportional to the number of their shared attributes.

For each HSN user n_i, we further consider a *social-active factor* $0 < \varrho_i \leqslant 1$ [11,12] to represent the intrinsical social level. It can be used to determine the probability that the user is going to forward the information for the social friends. If $\varrho_i = 1$, user n_i is maximally social-active and will forward the packets for its social friends with 100%; if $\varrho_i = 0$, user n_i is minimally social-active and will not forward any packet for its social friends.

2.2. Adversary model

The health information sharing application is subject to the following four security threats: eavesdropping attack, collusion attack, forgery attack, and attribute-tracing attack, which are further discussed as below.

Forgery attack: A user n_i sends an attribute proof to user n_j indicating that n_i does have some attributes. A forgery attack aims to cheat other users by making a false claim about the attributes it has. In addition, multiple users $n_x\{x \in X\}$ with attribute sets $\mathcal{A}_x (x \in X)$, where $\mathcal{A}_x \subset \mathcal{A}_y (x \in X)$ and $\mathcal{A}_y \subseteq \bigcup_{x \in X} \mathcal{A}_x$, cannot pretend a user with $\mathcal{A}_y$.

Attribute-trace attack: In the attribute proof, user n_j is able to anonymize some sensitive attributes to preserve user privacy. An attribute-trace attacker aims to check if user n_j has a specific attribute that has been anonymized in the attribute proof. If an attribute-trace attacker succeeds, the user privacy will be violated.

Eavesdropping attack: Two users n_i and n_j want to share some sensitive information with each other. Meanwhile, they do not allow other unauthorized users to read the information. An eavesdropping attacker without having the appropriate secret keys aims to obtain the transmitted data and decrypt it.

Collusion attack: A collusion attack is launched by multiple eavesdropping attackers. These eavesdropping attackers are unable to decrypt the confidential information individually. However, they are fully collaborative and share every secret they have. The aim of a collusion attack is the same as that of the eavesdropping attack, i.e., to obtain the confidential information.

2.3. Design goals

Our design goal is to develop novel schemes to achieve secure and privacy-preserving health information sharing for the HSN users. Specifically, we propose two schemes to attain the following design objectives.

- *Attribute-oriented authentication*: It enables a user to create an attribute proof of its attributes. The attribute proof cannot be forged, i.e., if user n_j successfully verifies the proof, user n_j confirms that user n_i does have the attributes as it claims. In order to achieve privacy preservation, user n_i should be able to anonymize the sensitive attributes in the attribute proof.
- *Attribute-oriented transmission*: It enables a user to securely transmit personal health information to other users based on a self-defined access policy. By having an attribute set that satisfies the access policy, a receiver is able to read the content; otherwise, the receiver cannot learn any meaningful information from the ciphertext.

3. Preliminaries

In this section, we briefly introduce the bilinear groups of a composite order and the corresponding bilinear maps [13,14]. After that, we introduce two satisfying relations which work as the basis of the proposed schemes.

3.1. Bilinear pairing

Let two finite cyclic groups $\mathbb{G}$ and $\mathbb{G}_T$ of the same composite order $n = pq$, where p and q are two different large primes. $\mathbb{G}_p$ and $\mathbb{G}_q$ are two subgroups of $\mathbb{G}$ with respective orders p and q. Suppose $\mathbb{G}$ and $\mathbb{G}_T$ are equipped with a pairing, i.e., an efficiently computable bilinear map $e : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_T$, with the following properties:

- Bilinearity: $\forall g,\ h \in \mathbb{G},\ \forall a, b \in \mathbb{Z}_n,\ e(g^a, h^b) = e(g, h)^{ab}$, where the product in the exponent is defined modulo n;
- Non-degeneracy: $\exists g \in \mathbb{G}$ such that $e(g,g)$ has order n in $\mathbb{G}_T$. $e(g,g)$ is a generator of $\mathbb{G}_T$, whereas g generates $\mathbb{G}$.

Definition 1. A bilinear parameter generator $\mathcal{G}en$ is a probabilistic algorithm that takes a security parameter p_k as input and outputs a 9-tuple $(n, p, q, g, u, h, \mathbb{G}, \mathbb{G}_T, e)$ where (p,q) are two large primes, $n = pq$, $\mathbb{G}$ and $\mathbb{G}_T$ are two finite cyclic groups with order n, (g,u) are two generators of $\mathbb{G}$ and h is a generator of $\mathbb{G}_q$ (a subgroup of $\mathbb{G}$ with order q), and $e : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_T$ is a bilinear map.

3.2. An attribute set satisfying a tree structure

An HSN user is able to construct a tree structure $\mathcal{T}$ with attributes from $\mathcal{A}_u$. The tree structure extensively used in the proposed schemes is formally defined as follows. Let $\mathcal{T}$ represent a tree structure with a root r, and each non-leaf node x represents a threshold gate with threshold value k_x. If x has c_x child nodes, the condition $0 < k_x \leqslant c_x$ must be satisfied. For each leaf node x, we use att(x) to denote the attribute associated with node x.

An attribute set γ satisfies a tree structure $\mathcal{T}$ (denoted by $\mathcal{T}(\gamma) = 1$): let $\mathcal{T}_x$ be a subtree rooted at a node x in $\mathcal{T}$. If an attribute set γ satisfies the tree structure $\mathcal{T}_x$, we have $\mathcal{T}_x(\gamma) = 1$. If x is a non-leaf node and there are at least k_x child nodes z of x so that $\mathcal{T}_z(\gamma) = 1$, we have $\mathcal{T}_x(\gamma) = 1$. If x is a leaf-node, then $\mathcal{T}_x(\gamma) = 1$ if and only if $\text{att}(x) \in \gamma$. An example is illustrated in Fig. 2.

Note that a tree structure semantically equals to a Boolean function, e.g.,

$$2 \text{ of } (A, B, C) \Longleftrightarrow (A \text{ AND } B) \text{ OR } (B \text{ AND } C) \text{ OR } (A \text{ AND } C).$$

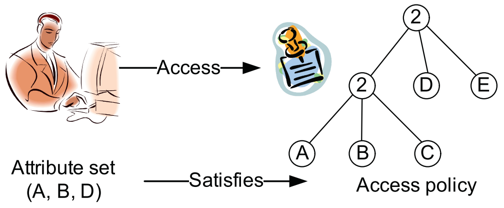


Fig. 2. An attribute set satisfying a tree structure.

3.3. An attribute set satisfying a linear secret sharing structure (LSSS)

Suppose that a linear secret sharing structure $\mathbb{A} = (M, \rho)$ can be satisfied by an attribute set γ as shown in Fig. 3, where M is a $l \times n$ matrix and ρ is an injective function from $\{1, \ldots, l\}$ to any attribute. Let $\mathcal{I} = \{i | \rho(i) \in \gamma\}$. Therefore, there exist constants $\{\omega_i \in \mathbb{Z}_q\}$ such that $\sum_{i \in \mathcal{I}} \omega_i M_i = (1, 0, \ldots, 0)$, where M_i is the ith row of matrix M. On the other hand, if γ does not satisfy $\mathbb{A}$, those constants $\{\omega_i\}$ do not exist. From the work [15], the constants $\{\omega_i\}$ can be found in polynomial time with the size of the matrix M. Moreover, let a vector $\bar{v} = (s, r_2, \ldots, r_n)$, where $s \in \mathbb{Z}_q$ is the secret to be shared, $r_2, \ldots, r_n \in \mathbb{Z}_q$ are random numbers. The inner product $M\bar{v}^T = (\lambda_1, \ldots, \lambda_l)^T$ can be regarded as the linear secret sharing. Given an attribute set γ and its corresponding rows $\mathcal{I} = \{i | \rho(i) \in \gamma\}$ in the matrix M, finding $\{\omega_i \in \mathbb{Z}_q\}$ so that $\sum_{i \in \mathcal{I}} \omega_i \cdot \lambda_i = s$ is called linear secret reconstruction.

4. Proposed schemes

In this section, we propose two attribute-oriented authentication and attribute-oriented transmission schemes for secure and privacy preserving health information sharing.

4.1. Design rationale

Attribute-oriented authentication: As shown in Fig. 4, user n_1 has an attribute set $\{a_2, a_4\}$ and the corresponding secret keys. User n_1 wants to find other users in the network who have similar attributes. Meanwhile, user n_1 does not want to overly expose his sensitive attributes, since too much disclosure of these sensitive attributes will help adversaries to track user n_1's behavior. Therefore, in order to preserve the privacy, user n_1 constructs a tree structure as an authentication policy. For example, in Fig. 4, the authentication policy is set to $a_4 \wedge (a_2 \vee a_8)$. By authenticating this policy to others, user n_1 proves that it has a_4 and one of $\{a_2, a_8\}$. Thus, user n_i can do effective authentication to reveal its attribute information while preserving the privacy on some sensitive attributes. If such attribute proof is delivered along the way $n_1 \to n_2 \to n_3 \to n_4$, n_3 and n_4 confirm that they individually share at least one attribute with n_1. They may add n_1 as one of their social friends.

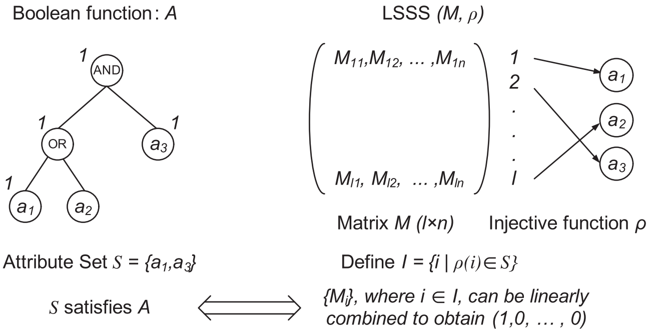


Fig. 3. An attribute set satisfying an LSSS.

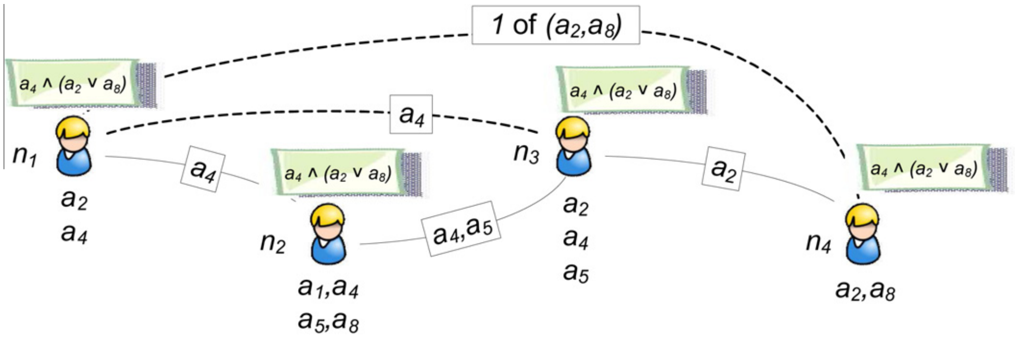


Fig. 4. Attribute-oriented authentication.

Attribute-oriented transmission: Attribute-oriented transmission has two modes, direct and indirect. For direct mode, users create the access policies by themselves; for indirect mode, a delegated user may help to create an appropriate access policy for the received ciphertext without having an access to the content.

(1) *Direct mode*: As shown in Fig. 5, user n_1 shares health information with other users, requiring that the user being able to read the health information must have an attribute set to satisfy an access policy. The access policy defined by n_1 is $a_2 \wedge a_8$, i.e., the users must have both a_2 and a_8 to read the content. The health information is delivered along the path $n_1 \rightarrow n_2 \rightarrow n_3 \rightarrow n_4$. During the data transmission, n_2 and n_3 receive the transmitted data packets but they are unable to read the health information since they individually do not have both a_2 and a_8. On the contrary, n_4 can successfully read the information by using the corresponding secret keys of a_2 and a_8. Thus, the direct mode of attribute-oriented transmission can well protect the confidentiality of users' health information according to their self-defined requirements.

(2) *Indirect mode*: As shown in Fig. 6, some users in the network can help user n_1 to create an appropriate access policy. Specifically, user n_1 first encrypts the health information with the public keys of a delegated user. After receiving the encrypted health information, the delegated user will re-encrypt the ciphertext and change the access policy to a standard one. With the re-encryption, the health information can be shared with other trusted users who satisfy the standard access policy. In addition, during the re-encryption, the delegated user is unable to read the content and thus the privacy of health information is still preserved. In Fig. 6, the delegated user transmits the re-encrypted health information to its social friends n_2, n_3, n_4 and n_5. Users n_2 and n_4 can read the information since their attribute sets satisfy the standard access policy. Note that the indirect mode is important in health information sharing applications, since a delegated user with more medical knowledge and richer social relations compared to the sender can provide help in appropriately encrypting and effectively distributing the health information.

4.2. Constructions

We introduce the schemes in three steps. First, in the initialization part, ATA generates the public parameters and secret keys for users; second, an attribute-oriented authentication scheme is presented which enables users to authenticate their attributes according to an authentication policy while preserving their privacy; and third, an attribute-oriented transmission scheme is proposed which enables users to protect the health information according to an access policy. For simplicity, the authentication policy and access policy considered are both single-threshold tree structures. In the following, we present the details of each step.

4.2.1. Attribute initialization

ATA runs $\mathcal{G}en(p_k)$ to generate the bilinear parameter $(n, p, q, g, u, h, \mathbb{G}, \mathbb{G}_T, e)$. It also initializes the universal attribute set $\mathcal{A}_u$ with size l. Suppose that the attribute-based authentication scheme supports tree structures with maximum threshold d for the authentication policy. ATA chooses a symmetric encryption algorithm $Sym = \langle Sym.enc,Sym.dec \rangle$, a secure cryptographic hash function $H : \{0,1\}^* \rightarrow \mathbb{Z}_n^*$, random numbers $\alpha, \delta, a, b, \tilde{t} \in \mathbb{Z}_n^*$ and different numbers $t_y \in \mathbb{Z}_n^*$ for all the $a_y \in \mathcal{A}_u$. ATA further chooses a redundant attribute set $\mathcal{A}_r$ indexed from $l+1$ to $l+d-1$. ATA additionally computes $\Lambda = e(g,g)^\alpha$, $\Delta = e(g,u)^\delta$, $A = g^a$, $B = g^b$, $\widetilde{T} = g^{\tilde{t}}$, and $T_y = g^{t_y}$ $(1 \leqslant y \leqslant l+d-1)$. ATA keeps the master key α, a, $t_y(1 \leqslant y \leqslant l+d-1)$ to itself, and publishes the system public parameters $\mathsf{pubs} = (n, g, u, h, \mathbb{G}, \mathbb{G}_T, e, H, Sym, \Lambda, \Delta, A, B, \widetilde{T}, T_y(1 \leqslant y \leqslant l+d-1))$.

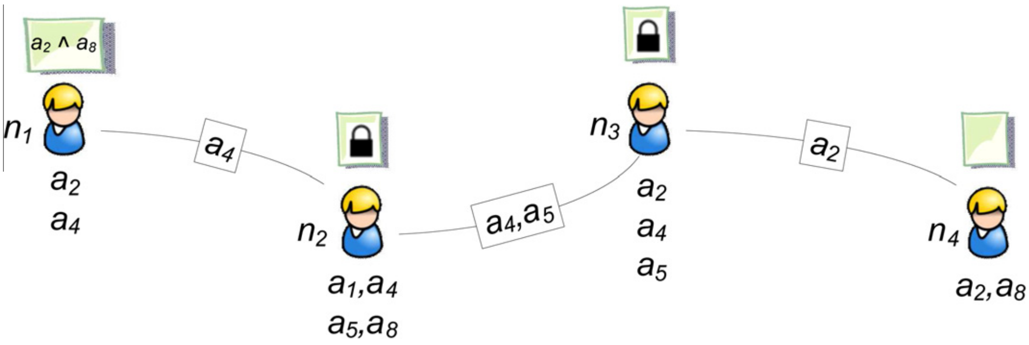


Fig. 5. Direct mode of attribute-oriented transmission.

Please cite this article in press as: X. Liang et al., HealthShare: Achieving secure and privacy-preserving health information sharing through health social networks, Comput. Commun. (2012), doi:10.1016/j.comcom.2012.01.009

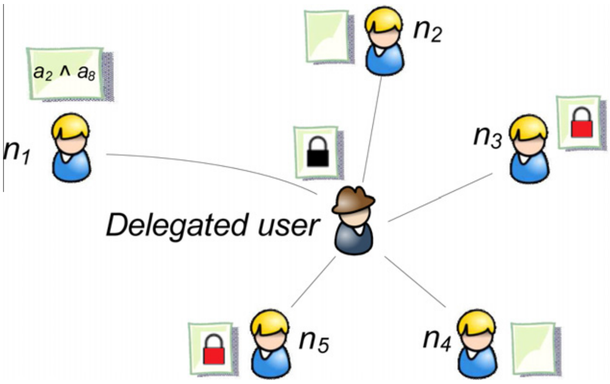


Fig. 6. Indirect mode of attribute-oriented transmission.

If user n_i having an attribute set $\mathcal{A}_i$ registers to ATA, ATA will generate a secret key corresponding to $\mathcal{A}_i$. Specifically, ATA chooses random numbers $t, t' \in \mathbb{Z}_n^*$ and a random polynomial $q(x) = \kappa_{d-1}x^{d-1} + \kappa_{d-2}x^{d-2} + \cdots + \kappa_1 x + \delta$ with degree $d-1$. ATA then calculates E_i as follows.

$$E_i = \langle K_e, K_d, L, (e_y)_{a_y \in \mathcal{A}_i}, (d_y)_{a_y \in \mathcal{A}_i \cup \mathcal{A}_r} \rangle,$$

where $K_e = g^{\alpha} g^{at}$, $K_d = t'$, $L = g^t$, $e_y = T_y^t$ and $d_y = u^{\frac{q(y)}{t'+t_y}}$. ATA secretly delivers the secret key E_i to user n_i. In addition, ATA assigns pseudonyms and corresponding pseudonym keys to users who will periodically change the pseudonyms in the communication to preserve their identity privacy [16].

4.2.2. Attribute-oriented authentication

Let users n_i and n_j denote a signer and a verifier respectively. Let AUP_i denote user n_i's authentication policy $\mathcal{T}_i$, where $\mathcal{T}_i$ is a single-threshold tree structure. Let the threshold value of AUP_i be k and Θ_i be an attributes set of $\mathcal{T}_i$. Since $\mathcal{A}_i$ satisfies AUP_i, it is able to find a k-size attribute set $\Phi_i \subseteq \mathcal{A}_i \cap \Theta_i$.

- *Signing algorithm performed by user n_i*: User n_i first chooses a subset $\mathcal{A}_{r'} \subseteq \mathcal{A}_r$ $(|\mathcal{A}_{r'}| = d - k)$. Let $\mathcal{A}_{r'}$ be $\{a_{l+1}, \ldots, a_{l+d-k}\}$. Then, for each attribute $a_y \in \Psi = \Phi_i \cup \mathcal{A}_{r'}$, user n_i computes the Lagrange coefficient $\omega_y = \sum_{w|a_w \in \Psi, w \neq y} \frac{0-w}{y-w}$. User n_i randomly selects $r_t, r_p, r_y \in \mathbb{Z}_n^*$ for $a_y \in \Theta_{i,k} \cup \mathcal{A}_{r'}$ and computes S_y for $a_y \in \Theta_i \cup \mathcal{A}_{r'}$ as follows

$$S_y = \begin{cases} d_y^{\omega_y} \cdot h^{r_y}, & \text{if } a_y \in \Psi, \\ h^{r_y}, & \text{if } a_y \in \Theta_i \setminus \Phi_i. \end{cases}$$

User n_i outputs the attribute proof

$$\sigma_i = \langle \mathcal{T}_i, S_t, S_p, (S_y)_{a_y \in \Theta_i \cup \mathcal{A}_{r'}}, \pi_1, \pi_2 \rangle,$$

where $S_t = g^{K_d} \cdot h^{r_t}$, $S_p = g^{\frac{1}{K_d + H(pid_i)}} \cdot h^{r_p}$ and

$$\pi_1 = S_p^{r_t} (g^{H(pid_i)} g^{K_d})^{r_p}, \pi_2 = \prod_{a_y \in \Psi} \left(d_y^{\omega_y}\right)^{r_t} \prod_{a_y \in \Theta_i \cup \mathcal{A}_{r'}} (S_t T_y)^{r_y}.$$

- *Verification algorithm performed by user n_j*: User n_j receives σ_i and checks

$$\begin{cases} e(S_t g^{H(pid_i)}, S_p) \stackrel{?}{=} e(g,g) \cdot e(h, \pi_1), \\ \prod_{a_y \in \Theta_i \cup \mathcal{A}_{r'}} e(S_y, S_t T_y) \stackrel{?}{=} \Delta \cdot e(h, \pi_2). \end{cases}$$

The correctness of the verification algorithm is proved as follows:

$$\begin{aligned} e(S_t g^{H(pid_i)}, S_p) &= e(g^t \cdot h^{r_t} \cdot g^{H(pid_i)}, g^{\frac{1}{t+H(pid_i)}} \cdot h^{r_p}) \\ &= e(g,g) \cdot e\left(h, \left(g^{\frac{1}{t+H(pid_i)}} \cdot h^{r_p}\right)^{r_t} \cdot (g^t \cdot g^{H(pid_i)})^{r_p}\right) \\ &= e(g,g) \cdot e\left(h, S_p^{r_t} (g^{H(pid_i)} g^t)^{r_p}\right) \\ &= e(g,g) \cdot e(h, \pi_1), \end{aligned}$$

$$\begin{aligned} \prod_{a_y \in \Theta_i \cup \mathcal{A}_{r'}} e(S_y, S_t T_y) &= \prod_{a_y \in \Psi} e\left(d_y^{\omega_y}, S_t T_y\right) \cdot \prod_{a_y \in \Theta_i \cup \mathcal{A}_{r'}} e(h^{r_y}, S_t T_y) \\ &= \prod_{a_y \in \Psi} e\left(u^{\frac{\omega_y q(y)}{K_d + t_y}}, g^{K_d} \cdot h^{r_t} g^{t_y}\right) \\ &\quad \cdot \prod_{a_y \in \Theta_i \cup \mathcal{A}_{r'}} e(h^{r_y}, S_t T_y) \\ &= e(g,u)^{\delta} \prod_{a_y \in \Psi} e\left(u^{\frac{r_t \omega_y q(y)}{K_d + t_y}}, h\right) \\ &\quad \cdot \prod_{a_y \in \Theta_i \cup \mathcal{A}_{r'}} e(h^{r_y}, S_t T_y) \\ &= \Delta \cdot e(h, \prod_{a_y \in \Psi} \left(d_y^{\omega_y}\right)^{r_t} \prod_{a_y \in \Theta_i \cup \mathcal{A}_{r'}} (S_t T_y)^{r_y}) \\ &= \Delta \cdot e(h, \pi_2) \end{aligned}$$

If the above equations hold and user n_i further proves that it has pid_i using the pseudonym keys, user n_j then confirms that the user with pseudonym pid_i has attributes to satisfy AUP_i.

4.2.3. Attribute-oriented transmission

Let users n_i and n_j denote a sender and a receiver, respectively, m denote user n_i's health information, and an LSSS (M, ρ) denote user n_i's access policy ACP_i, where M is a $\mu \times \theta$ matrix and ρ is a mapping from $\{1,2,\ldots,\mu\}$ to the attribute index $\{1,2,\ldots,l\}$. Let M_y denote the y-th row of M. Suppose that user n_j's attribute set $\mathcal{A}_j$ satisfies the LSSS (M, ρ) as defined in Section 3.3.

- *Encryption algorithm performed by user n_i*: User n_i chooses a random vector $\vec{v} = (v_1 = s, v_2, \ldots, v_\theta) \in \mathbb{Z}_n^\theta$, random numbers $v_1, \ldots, v_\theta \in \mathbb{Z}_n$ and a symmetric key $sk \in \mathbb{G}_T$. For $1 \leqslant y \leqslant \mu$, user n_i calculates $\lambda_y = \vec{v} \cdot M_y$. The ciphertext is $\mathcal{C} = \langle C, C', C_s, (C_y, D_y)_{1 \leqslant y \leqslant \mu} \rangle$, where $C = sk \cdot \Lambda^s$, $C' = g^s$, $C_y = A^{\lambda_y} T_{\rho(y)}^{-r_x}$, $D_y = g^{r_y}$ and $C_s = Sym.enc(sk, m)$.
- *Decryption algorithm performed by user n_j*: User n_j receives the ciphertext C from user n_i. User n_j has $E_j = \langle K_e, K_d, L, (e_y, d_y)_{a_y \in \mathcal{A}_j} \rangle$ and let $\mathcal{I} = \{y | a_{\rho(y)} \in \mathcal{A}_j\}$. It is computationally feasible to find $\{\omega_y \in \mathbb{Z}_n\}_{y \in \mathcal{I}}$ so that $\sum_{y \in \mathcal{I}} \omega_y \lambda_y = s$. User n_j then calculates $C \cdot \prod_{y \in \mathcal{I}} (e(C_y, L) \cdot e(D_y, e_{\rho(y)}))^{\omega_y} / e(C', K_e) = sk$ and $Sym.dec(sk, C_s) = m$.

  The correctness of the decryption algorithm is proved as follows:

$$\begin{aligned} & C \cdot \prod_{x \in I} (e(C_x, L) \cdot e(D_x, e_{\rho(x)}))^{\omega_x} / e(C', K_e) \\ & = sk \cdot \Lambda^s \cdot \prod_{x \in I} (e(C_x, L) \cdot e(D_x, e_{\rho(x)}))^{\omega_x} / e(g^s, g^{\alpha} g^{at}) \\ & = sk \cdot \prod_{x \in I} (e(A^{\lambda_x} T_{\rho(x)}^{-r_x}, g^t) \cdot e(g^{r_x}, T_{\rho(x)}^t))^{\omega_x} / e(g^s, g^{at}) \\ & = sk \cdot \prod_{x \in I} (e(A^{\lambda_x}, g^t))^{\omega_x} / e(g^s, g^{at}) = sk \cdot e(A^{\lambda_x}, g^t)^s / e(g^s, g^{at}) \\ & = sk. \end{aligned}$$

Please cite this article in press as: X. Liang et al., HealthShare: Achieving secure and privacy-preserving health information sharing through health social networks, Comput. Commun. (2012), doi:10.1016/j.comcom.2012.01.009

- *Delegated encryption algorithm performed by user n_i and a delegated user n_d:* In this algorithm, a delegated user n_d selected by user n_i will help user n_i to control the access of user n_i's health information. The delegated user n_d cannot access the health information during the whole process. Specifically, n_d first obtains a transforming key $\mathsf{TK} = \langle K, L, K_o, L_o\rangle$ from ATA, where $K = g^{\alpha} g^{at} g^{b\beta},\ L = g^t,\ K_o = \widetilde{T}^t$, and $L_o = \mathbb{E}(g^{\beta})$ ($\mathbb{E}$ is the encryption algorithm with a standard access policy). User n_i computes $\mathcal{C} = \langle C,\ C', C'', C_s, C_1, D_1\rangle$, where $C = sk \cdot \Lambda^s,\ C' = g^s,\ C'' = B^s,\ C_1 = A^s \widetilde{T}^{-r},\ D_1 = g^r$ and $C_s = Sym.enc(sk, m)$. Then, after receiving $\mathcal{C}$ from user n_i, n_d transforms $\mathcal{C}$ to $\mathcal{C}'$ with the standard access policy. n_d computes $C_t = e(K, C')/(e(C_1, L)e(D_1, K_o))$, and re-organizes $\mathcal{C}' = \langle C,\ C', C_t, C_s, L_o\rangle$.
- *Delegated decryption algorithm performed by user n_j*: The ciphertext $\mathcal{C}'$ can be decrypted by user n_j if $\mathcal{A}_j$ satisfies the standard access policy. User n_j decrypts L_o to obtain g^{β} and calculates $C \cdot e(C'', g^{\beta})/C_t = sk$ and $Sym.dec(sk, C_s) = m$.

The correctness of the delegated decryption algorithm is proved as follows:

$$\begin{aligned} C \cdot e(C'', g^{\beta})/C_t &= sk \cdot \Lambda^s \cdot e(B^s, g^{\beta}) e(C_1, L) e(D_1, K_o)/e(K, C') \\ &= sk \cdot \Lambda^s \cdot e(B^s, g^{\beta}) e(A^s \widetilde{T}^{-r}, g^t) e(g^r, \widetilde{T}^t)/e(g^{\alpha} g^{at} g^{b\beta}, g^s) \\ &= sk \cdot \Lambda^s \cdot e(B^s, g^{\beta})/e(g^{\alpha} g^{b\beta}, g^s) = sk \cdot \Lambda^s/e(g^{\alpha}, g^s) \\ &= sk. \end{aligned}$$

5. Security analysis

In this section, we analyze the security and privacy properties of the proposed schemes. In particular, following the adversary model discussed earlier, our analysis will focus on how the proposed scheme can resist the forgery attack, attribute-trace attack, eavesdropping attack, and collusion attack.

- *The attribute-oriented authentication is resistant to the forgery attack.* A forgery attacker $\mathcal{K}$ pretends to have some attributes that it actually does not have. Specifically, $\mathcal{K}$ uses the secret keys of an attribute set $\mathcal{A}_i$ to forge an attribute proof $\mathcal{S}$ with an authentication policy AUP, where $\mathcal{A}_i$ does not satisfy AUP. From the algorithm, it can be seen that if $\mathcal{K}$ succeeds, ATA can use q to calculate ϵ such that $\epsilon \equiv 1\ (\text{mod} p)$ and $\epsilon \equiv 0\ (\text{mod} q)$ and then have $\sigma_i^* = \left\langle \mathcal{T}_i, S_t^*, S_p^*, \left(S_y^*\right)_{a_y \in \Theta_i \cup \mathcal{A}_{r'}} \right\rangle$, where $\prod_{a_y \in \Theta_i \cup \mathcal{A}_{r'}} e\left(S_y^*, S_t^* T_y\right) = A$ and $e\left(S_t^* \cdot g^{H(pid_i)}, S_p^*\right) = e(g, g)$. The first equation confirms that (S_y, S_t) are generated by a registered user with enough attributes. The unforgeability is guaranteed by the short signature scheme [17]. The second equation confirms that $\mathcal{K}$ must have a secret key K_d to relate the signature to pid_i. The unforgeability is guaranteed by the group signature scheme [18] and a general signature scheme used on pseudonyms. Therefore, the forgery attack can be effectively resisted by the attribute-oriented authentication scheme.
- *The attribute-oriented authentication scheme is resistant to the attribute-trace attack.* An attribute-trace attacker $\mathcal{K}$ aims to trace the real attributes of a user from its attribute proof. We consider a target user n_i and its attribute proof $\mathcal{S}$. Recall the signing algorithm, we have $S_y = d_y^{\omega_y} \cdot h^{r_y}$ if $a_y \in \Phi_i$ or $S_y = h^{r_y}$ if $a_y \notin \Phi_i$. $\mathcal{K}$ is unable to check whether ATA chooses $h \in \mathbb{G}$ or $h \in \mathbb{G}_q$ in the initialization phase. If $h \in \mathbb{G},\ S_y$ is a random number from the attacker's perspective. Therefore, the capability of $\mathcal{K}$ is stronger than the capability of distinguishing the element $h \in \mathbb{G}_q$ from $\mathbb{G}$. However, it is shown that the later is a computational hard problem in recent works [14,18]. Thus, the attribute-trace attacker $\mathcal{K}$ cannot succeed in the attribute-oriented authentication scheme.
- *The attribute-oriented transmission scheme is resistant to the eavesdropping attack.* An eavesdropping attacker $\mathcal{K}$ aims to obtain the health information from the intercepted ciphertext $\mathcal{C}$. Suppose $\mathcal{K}$ has an attribute set $\mathcal{A}_i$ and $\mathcal{C}$ is with an access policy (M, ρ). $\mathcal{A}_i$ does not satisfy (M, ρ) or the standard access policy.
  For the ciphertext $\mathcal{C}$ output by the *Encryption Algorithm*, in order to decrypt m, the attacker must have the symmetric key sk which is encrypted by a $\mathbb{G}_T$ group element Λ^s. In the algorithm, the random number s has been divided into multiple shares λ_y, where each share λ_y can be only revealed by a tuple (C_y, D_y). In other words, $\mathcal{K}$ must have enough shares λ_y to recover Λ^s. In addition, for each share, since a unique random number r_y is embedded into both C_y and D_y, the attacker must have the secret key $e_{\rho(y)}$ to delete λ_y-related element from $C_y = A^{\lambda_y} T_{\rho(y)}^{-r_y}$ and then obtain the share. However, since $\mathcal{A}_i$ does not satisfy (M, ρ), the attacker cannot get enough secret keys to obtain the secret shares. Therefore, the ciphertext generated by *Encryption Algorithm* can resist the eavesdropping attack.
  For the ciphertext $\mathcal{C}'$ output by the *Delegated Encryption Algorithm*, in order to get message $m, \mathcal{K}$ must decrypt g^{β} from L_o and use g^{β} to remove the β-related element from C_t. The decryption of L_o requires the users to have an attribute set satisfying the standard access policy. This requirement protect the health information from being accessed by an unauthorized user. Therefore, the ciphertext generated by *Delegated Encryption Algorithm* also is resistant to the eavesdropping attack.
- *The attribute-oriented transmission scheme is resistant to the collusion attack.* The collusion attack is launched by multiple users. Suppose they individually cannot decrypt a target ciphertext, but they try to decrypt it by sharing the secret keys. In the following, we show that the scheme is resistant to two-user collusion attack. Its resilience against more-than-two-user collusion attack can be derived similarly. Recall that ATA assigns each user with a random and unique number t. Suppose two users n_i and n_j are corresponding to t_i and t_j respectively. They are able to compute $e(g, g)^{at_i \lambda_{y_i}}$ and $e(g, g)^{at_j \lambda_{y_j}}$, where $a_{\rho(y_i)} \in \mathcal{A}_i$ and $a_{\rho(y_j)} \in \mathcal{A}_j$. From the above results, it can be seen that even if $\mathcal{A}_i \cup \mathcal{A}_j$ satisfies the access policy, the independent relation between t_i and t_j can prevent the combination of $e(g, g)^{at_i \lambda_{y_i}}$ and $e(g, g)^{at_j \lambda_{y_j}}$. Therefore, the attribute-oriented transmission scheme can effectively resist the collusion attack.

6. Performance analysis and evaluation

In this section, we analyze the proposed schemes from two perspectives. First, we analyze the computational costs of the proposed schemes. We count the operation times of each algorithm and estimate the time costs. Second, we simulate the HSNs where users are connected via attribute-based social links. We conduct a set of custom simulations in the HSN to demonstrate the efficiency in sharing health information by adopting the proposed schemes.

6.1. Efficiency analysis

The attribute-oriented authentication and the attribute-oriented transmission schemes include six algorithms. Denote the computational cost of a bilinear pairing operation and a point multiplication operation on $\mathbb{G}$ by C_p and C_m, respectively. Note that since the pairing and point multiplication operations are more costly than other types of operations, we only count the times of these operations used in each algorithm. The summarized computational costs of signing and verification algorithms are given in Table. 1. It can be seen that the computational costs of the two algorithms increase as $|\Theta_i|$ increases and k decreases. This is because when $|\Theta_i|$ increases, the parameters of more attributes are

Table 1
Computational cost of attribute-oriented authentication scheme.

	Signing	Verification
Costs	$(4d - 3k + 7 + 3\lvert\Theta_i\rvert) \cdot C_m$	$(4 + \lvert\Theta_i\rvert + d - k) \cdot\ C_p + C_m$

needed to be included in the computation; when k decreases, the parameters from more redundant attributes are included in the computation. From Table 2, it can be seen that the computational cost of encryption algorithm increases as μ increases. The reason is that when μ increases, the access policy becomes more complicated and the parameters of more attributes are included. The computational cost of decryption algorithm mainly depends on the size of $|\mathcal{I}|$ which represents the number of attributes that the receiver uses in the decryption. The delegated encryption and decryption algorithms have constant computational costs, where C_{std} denotes the computational costs spent on decrypting a ciphertext encrypted with a standard access policy.

We further conduct the experiments with PBC [19] and MIRACL [20] libraries on a 3.0 GHz-processor 512 MB-memory computing machine to study the time costs. The experimental results indicate that a single point multiplication operation in $\mathbb{G}$ with 160 bits costs 6.4 ms and the corresponding pairing operation costs 20 ms. We choose $d = 5$ to restrict the maximum threshold that HSN users sets in their policies as 5, and $|\Theta_i| = 15$ to indicate that the number of attributes used in policies is 15. (We will use the same settings in the simulation). Thus, in the case of $k = 1$, the time costs of Signing and Verification algorithms reach the maximum values 441.6 ms and 466.4 ms, respectively. If constructing the access policy as the authentication policy, we have $\mu = 15$ and $|\mathcal{I}| = 5$. Thus, we obtain the estimated time costs of Encryption, Decryption, Delegated Encryption, and Delegated Decryption as 288 ms, 252 ms, 60 ms, and 272 ms, respectively. It can be seen that these algorithms are very efficient in comparison with the communication delay. In addition, since users can perform these algorithms to guarantee data security and user privacy in a distributed manner, their secure and privacy-preserving communications do not rely on the protection mechanism from a centralized third party and are more reliable and efficient.

6.2. Simulation settings

In the simulation, we consider a typical HSN where 1000 users $\{n_1, \ldots, n_{1000}\}$ are interconnected via attribute-based social links. Denote the universal attribute set by $\mathcal{A}_u$ ($|\mathcal{A}_u| = 100$) and the attribute set of user n_i by $\mathcal{A}_i$. We let $|\mathcal{A}_u| = 100$ and $1 \leqslant |\mathcal{A}_i| \leqslant 15$ for $1 \leqslant i \leqslant 1000$. We divide 1000 users into fifteen categories $Cat_z (1 \leqslant z \leqslant 15)$, where the users in category Cat_z all have z attributes.

Attribute-based social link: We generate the attribute-based social links among users to simulate their social communications. Basically, if two users share more attributes, they are connected with a larger probability. We treat every attribute equally and independently, and the probability of two users being connected is $1 - (1 - \eta)^k$ where $0 < \eta < 1$ is the probability related to one shared attribute and k is the number of shared attributes. Under the above setting, we generate the social links in the case of $\eta = 0.95$ and show the minimum and maximum number of social friends in terms of the number of users' attribute in Fig. 7(a). It can be seen that a user would have more social friends if he is more social-active (i.e., has more attributes).

Forwarding probability: HSN users can directly communicate with their social friends. When they receive the forwarding requests from their social friends, they would help to forward the data with certain probability. Such probability can be represented by a social-active factor which is a constant value internalized for users. In the simulation, we generate a probability ϱ for each user according to the Beta distributions. Three considered distributions are shown in Fig. 7(b), where most users are highly social-active in $\beta(2,5)$, most users are averagely social-active in $\beta(2,2)$ and most users are lowly social-active in $\beta(5,2)$.

Access policy and authentication policy: In the simulation, we demonstrate the effectiveness of the proposed schemes by using simple policies. We choose the policies in the form of "*th* out of k attributes". We set $k = 15$ and $1 \leqslant th \leqslant 5$ for all the cases. The attribute set of user n_i's authentication policy AUP includes $\mathcal{A}_i$ and $15 - |\mathcal{A}_i|$ other attributes in $\mathcal{A}_u \setminus \mathcal{A}_i$. In the following, we show the simulation results of attribute-oriented authentication and attribute-oriented transmission schemes, respectively.

6.3. Simulation results of attribute-oriented authentication

The performance metric used to evaluate attribute-oriented authentication is the number of satisfied users. A user is called a satisfied user of an authentication policy AUP if the attribute set of that user satisfies a AUP. We first show how many users in the network are satisfied users. As shown in Fig. 8(a), the number of satisfied user significantly decreases as the threshold value of AUP increases. If the threshold value is larger than five, the number is extremely small. This result implies that if a user chooses an authentication policy with a threshold value larger than 5, the user's behaviors may be easily traced and the privacy may be violated. In addition, we use probability theory to confirm the performance results. Let $F(t)$ be the probability that a user satisfies a AUP with threshold t. $F(t)$ can be represented as follows:

$$F(t) = \Pr[\text{user } n_i\text{'s attribute set satisfying AUP}]$$
$$= \frac{1}{15}\sum_{i=th}^{15} \frac{\sum_{j=th}^{i} \binom{15}{j}\binom{85}{i-j}}{\binom{100}{i}}.$$

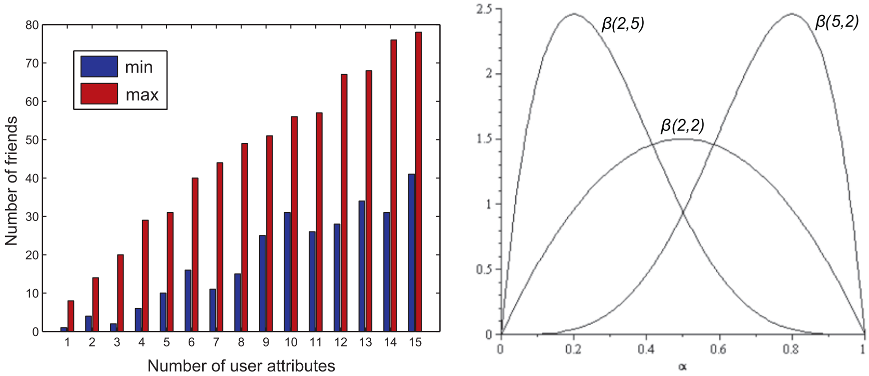
We have $F(1) = 0.667$, $F(2) = 0.342$, $F(3) = 0.138$, $F(4) = 0.042$, $F(5) = 0.010$ and $F(6) < 0.002$. Therefore, if the number of users is 1000, the threshold value of AUP can be 5 at the maximum for privacy preservation.

Since the social links are tightly related to the number of shared attributes, we mathematically derive the probability of sharing at least one common attribute between two users after they finish an attribute-oriented authentication. The probability can be used to measure the effectiveness of the attribute-oriented authentication algorithm. Let users n_i and n_j have attribute sets $\mathcal{A}_i$ and $\mathcal{A}_j$, respectively. Suppose that user n_j proves to user n_i that his attribute set satisfies an authentication policy AUP_j. Denote the attribute set of the AUP_j by ϕ_j ($|\phi_j| = q_j$) and the threshold of the AUP_j by th_j. Denote $|\mathcal{A}_i \cap \phi_j| = th_{i,j}$. We use q_j, th_j and $th_{i,j}$ to approxi-

Table 2
Computational cost of attribute-oriented transmission scheme.

	Encryption	Decryption	Delegated encryption	Delegated decryption
Costs	$3\mu \cdot C_m$	$(2\lvert\mathcal{I}\rvert + 1) \cdot C_p + \lvert\mathcal{I}\rvert \cdot C_m$	$3 \cdot C_p$	$C_p + C_{std}$

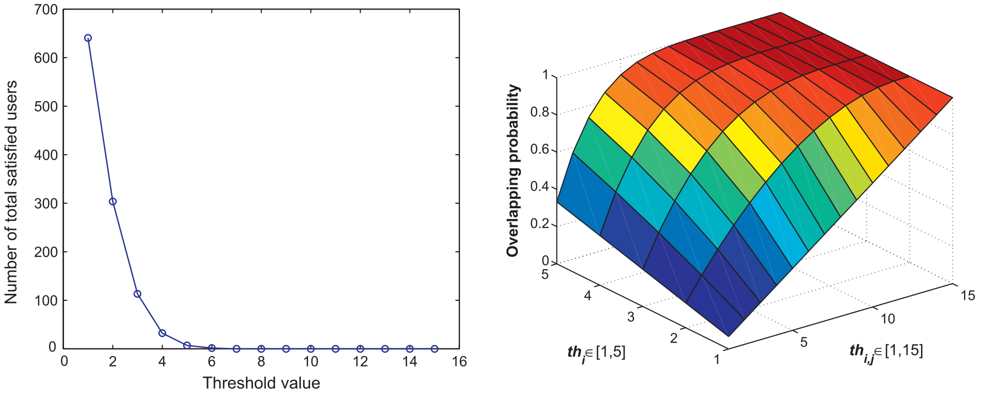
Please cite this article in press as: X. Liang et al., HealthShare: Achieving secure and privacy-preserving health information sharing through health social networks, Comput. Commun. (2012), doi:10.1016/j.comcom.2012.01.009



(a) Number of friends ($\eta = 0.95$)

(b) Social-active factor ϱ

Fig. 7. Parameter settings.



(a) Number of satisfied users vs. th

(b) Overlapping Probability between $\mathcal{A}_i$ and $\mathcal{A}_j$

Fig. 8. Performance of attribute-oriented authentication.

mately derive the probability $\Pr_o$ that users n_i and n_j share at least one attribute. ($|\mathcal{A}_u| \gg |\mathcal{A}_i|, |\mathcal{A}_j|$)

$$\Pr_o = \begin{cases} 1 - \dfrac{\binom{q_j - th_{i,j}}{th_j}}{\binom{q_j}{th_j}}, & \text{if } 1 \leqslant th_{i,j} \leqslant q_j - th_{i,j} \\ 1, & \text{if } q_j - th_{i,j} < th_{i,j} \end{cases}$$

In Fig. 8(b), the overlapping probability is shown in terms of th_i and $th_{i,j}$. For simplicity, we choose $q_j = 15$, th_j from 1 to 5, and $th_{i,j}$ from 1 to 15. It can be seen that the overlapping probability dramatically increases when either th_j or $th_{i,j}$ increases. Therefore, the proposed attribute-oriented authentication scheme enables two users to check if they have common attributes while preserving their privacy. The effectiveness of the scheme is significant especially when th_j and $th_{i,j}$ are large.

6.4. Simulation results of attribute-oriented transmission

The performance metrics used to evaluate attribute-oriented transmission scheme are the number of visited users, the number of visited eligible users, and the number of total eligible users. A user is called: a visited user if it receives the packet; an eligible user if it is able to access the content of the packet. In the simulation, we choose user n_i who generates a data packet m_i with an access policy ACP. Any user forwards a received data packet to its social friends with the probability ϱ. In Fig. 9, we show the three metrics in terms of the number of user n_i's attributes.

6.4.1. Impact of number of user attributes

From Fig. 9, it can be seen that if user n_i has more attributes, there are more visited users and more visited eligible users. The reason is that a user with more attributes will have more social friends in the network. The social friends will help forwarding m_i to more users. In addition, Fig. 9 shows that the number of total eligible users is a constant. In the initialization phase, the attributes are uniformly and randomly assigned to every user. Therefore, given a specific ACP with constant-sized attribute set and constant threshold value, the number of the users that satisfy ACP is a constant.

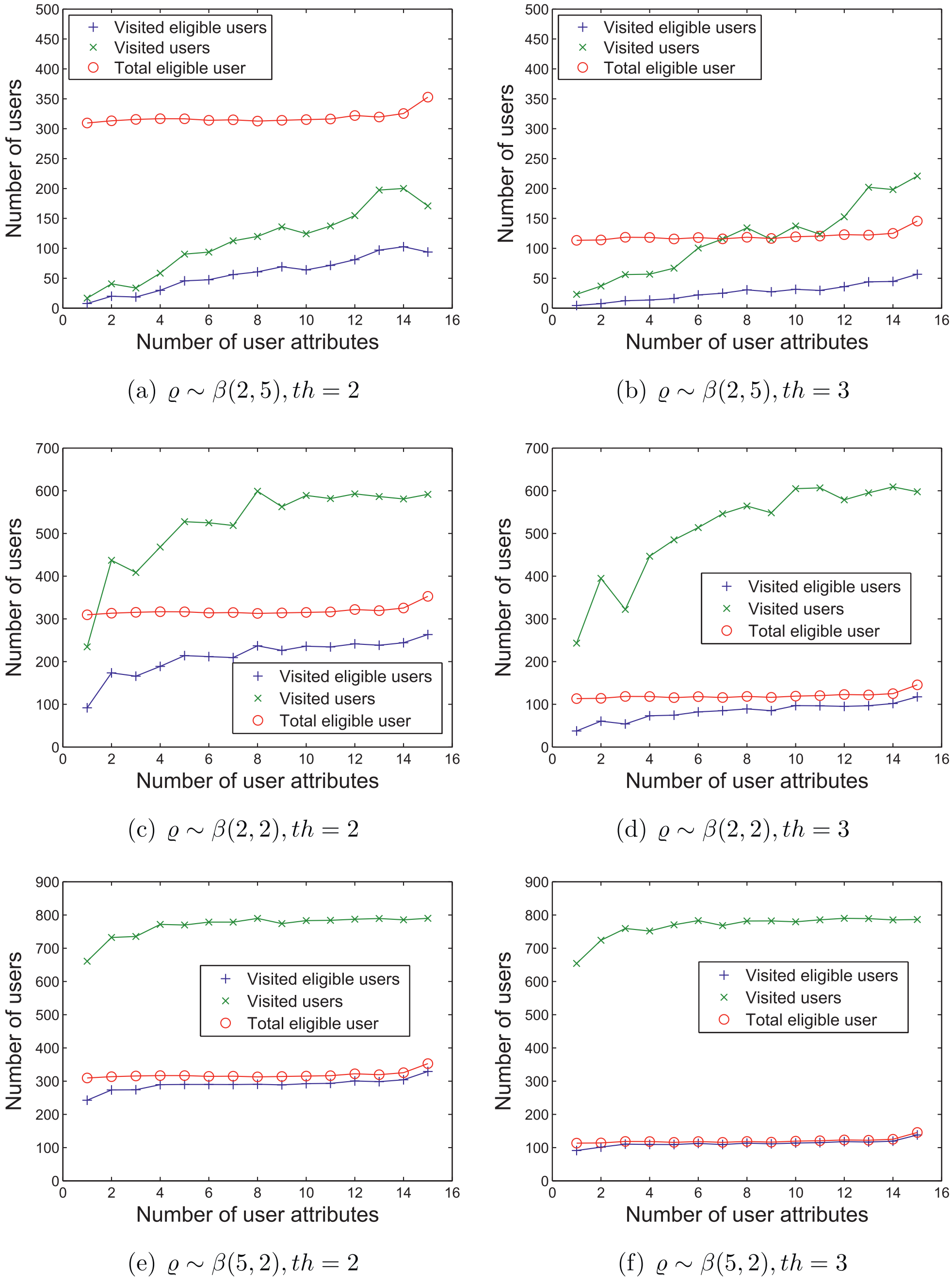
6.4.2. Impact of threshold value th

By comparing any two subfigures in a row of Fig. 9, we can see that the number of total eligible users decreases from 300 to 100 if the threshold increases from 2 to 3. This is because if the threshold increases, the access policy will become more strict and the number of total eligible users will decrease. Furthermore, the threshold change does not affect the number of visited users. However, an increase of threshold value leads to a decrease of the number of visited eligible users.

6.4.3. Impact of social-active factor ϱ

By comparing any three subfigures in a column of Fig. 9, when most users have large social-active factors in the case of $\beta(5,2)$, the number of visited users and visited eligible users increases significantly. This is because most users help forwarding data packets to their social friends and more users receive m_i. However, the number of total eligible users remains unchanged.

Please cite this article in press as: X. Liang et al., HealthShare: Achieving secure and privacy-preserving health information sharing through health social networks, Comput. Commun. (2012), doi:10.1016/j.comcom.2012.01.009



(a) $\varrho \sim \beta(2,5), th = 2$

(b) $\varrho \sim \beta(2,5), th = 3$

(c) $\varrho \sim \beta(2,2), th = 2$

(d) $\varrho \sim \beta(2,2), th = 3$

(e) $\varrho \sim \beta(5,2), th = 2$

(f) $\varrho \sim \beta(5,2), th = 3$

Fig. 9. Performance of attribute-oriented transmission.

In summary, from the above analysis, we have demonstrated that the attribute-based authentication scheme can enable users to exchange authenticated information while preserving their privacy, and the attribute-based transmission scheme can enable users to efficiently deliver their health information to their specified users.

7. Related work

Recent research works related to health information sharing applications include (i) health social networks and health information sharing application, (ii) access control of sensitive health information, and (iii) security and privacy preservation in social networks.

Most research works related to information sharing in health social networks focus on framework design or architectural details [21–24]. Rahman et al. proposed a framework, called SenseFace, that works as a bridge between Wireless Body Area Networks (WBANs) and health social networks [21]. Moncur et al. [22] showed that people's required information can be predicted from their self-reported emotional proximity and their gender. They built a model of social intelligent communication of health information across the social network. Massey et al. [23] further analyzed techniques to mitigate data loss in social networks by modifying the sampling rate and leveraging patterns in human mobility and reception rate in WBANs. Moncur et al. [25] indicated that a patient would require more information from their social friends who have a personal and trusted relation with the patient. Based on this observation, they proposed a model for socially intelligent communication of health information across the social networks. Our proposed schemes also consider the social proximity of HSN users and aim to improve the communication efficiency

Please cite this article in press as: X. Liang et al., HealthShare: Achieving secure and privacy-preserving health information sharing through health social networks, Comput. Commun. (2012), doi:10.1016/j.comcom.2012.01.009

by utilizing the social characteristics. In addition, our proposed scheme considers security and privacy preservation issues which so far have not been addressed.

Since the applications of Online Social Networks (OSNs) currently dominate the internet users, it is unsurprisingly that the security and privacy preservation of these applications has been paid a great attention [26–30]. Baden et al. [26] indicated an attribute-based encryption technique could be used to fine-grained control the access to personal information for OSNs. Besides, the authors [27–29] addressed the disclosure of identifiable information in OSNs. Zheleva et al. studied an adversary who predicts the private attributes of users based on the mixture of public and private user profiles [27]. Krishnamurthy and Wills [28] measured the leakage of personally identifiable information (PII) over the OSNs and identified potential solutions to eliminate it. Puttaswamy et al. [29] defined a social intersection attacker who aims at identifying the source of the shared content. They developed a new graph structure named StarClique, where the addition of redundant social links guarantees the *k*-anonymity of users. In this paper, we address the data access control problem by proposing an attribute-oriented transmission scheme and implement the privacy preservation into the proposed attribute-oriented authentication scheme.

For data access control, Massacci et al. [31] presented a goal-oriented role-based access control model in which users' predefined security policy determines whether they are endangered or not and to send an emergency request to monitoring and emergency response center for help. Mohan et al. [32] proposed a framework for electronic health record sharing that provides patients with the fine-grained control on the private information. The framework also uses the attribute-based techniques to improve the communication efficiency. Barua et al. [33] categorized patient's health information into different privacy levels before using attribute-based encryption to ensure patient-centric access control. Liang et al. [34] studied the data access control problem in a healthcare emergency case. In addition, Benaloh et al. [35] built an efficient healthcare system in which patients are able to not only share partial access rights with others, but also perform searches over their records. In this paper, the proposed attribute-oriented transmission scheme further improves the fine-grained access control by offering alternative options to users and enabling them to choose either the direct mode or the indirect mode for the access control.

For privacy preservation, Decker et al. [36] proposed a patient-privacy preserving protocol for the prescription handling process in the current Belgian healthcare practise, achieving both information unlinkability and identity untraceability. Kotz et al. [37] integrally surveyed the current framework for mobile healthcare system and identified their key differences and shortcomings. They proposed 11 privacy properties and indicated the future research directions, e.g., consent management and information anonymization. In this paper, we address the privacy preservation problem by proposing an attribute-oriented authentication scheme which enables users to share attribute information without violating user privacy.

8. Conclusion

In this paper, we have proposed an attribute-oriented authentication scheme and an attribute-oriented transmission scheme for Health Social Newtork (HSN) users to achieve secure and privacy preserving health information sharing. The proposed schemes can help HSN users to create more social relations with trusted users and share health information with them. By security and efficiency analysis, and simulation evaluation, the proposed schemes have been demonstrated to resist various attacks including forgery attack, attribute-trace attack, eavesdropping attack, and collusion attack. In addition, They are able to preserve user privacy while realizing highly-efficient health information sharing applications. The proposed schemes can be easily adopted to any existed HSN framework due to their simplicity and efficiency. For the future work, the proposed schemes will be extended to fit a more complicated social environment where HSN users may have distinctive social behaviors and social requirements.

References

[1] Patientslikeme. <http://www.patientslikeme.com/>.
[2] AT&T Healthcare Community Online (HCO). <http://www.corp.att.com/healthcare/hco/>.
[3] K.J. Leonard, One patient, one record: Report on one-day symposium to promote patient ehealth, Technique reports, 2010. <http://patientdestiny.typepad.com/OPOR%20Report%20-%20Ottawa.pdf>.
[4] M. Domingo, Managing healthcare through social networks, IEEE Computer Magazine 43 (7) (2010) 20–25.
[5] M. Chen, S. Gonzalez, V. Leung, Q. Zhang, M. Li, A 2G-RFID-based e-healthcare system, IEEE Wireless Communications Magazine 17 (1) (2010) 37–43.
[6] M. Chen, S. González-Valenzuela, A.V. Vasilakos, H. Cao, V.C.M. Leung, Body Area Networks: A Survey Mobile Networks and Applications (MONET) 16 (2) (2011) 171–193.
[7] X. Liang, X. Li, Q. Shen, R. Lu, X. Lin, X. Shen, W. Zhang, Exploiting Prediction to Enable Secure and Reliable Routing in Wireless Body Area Networks, in: IEEE INFOCOM 2012.
[8] X. Shen, N. Kato, X. Lin (guest editors), IEEE Wireless Communications, Special Issue on Wireless Technologies for E-healthcare 17 (1) (2010).
[9] X. Shen, J. Misic, N. Kato, P. Langendorfer, X. Lin (guest editors), IEEE/JCN, Special Issue on Emerging Technologies and Applications of Wireless Communication in Healthcare, 2011.
[10] X. Lin, R. Lu, X. Shen, Y. Nemoto, N. Kato, Sage: a strong privacy-preserving scheme against global eavesdropping for ehealth systems, IEEE Journal on Selected Areas in Communication 27 (4) (2009) 365–378.
[11] V. Borrel, F. Legendre, M.D. de Amorim, S. Fdida, Simps: using sociology for personal mobility, IEEE/ACM Transactions on Networking 17 (3) (2009) 831–842.
[12] Q. Li, S. Zhu, G. Cao, Routing in socially selfish delay tolerant networks, in: IEEE INFOCOM, 2010, pp. 857–865.
[13] D. Boneh, M.K. Franklin, Identity-based encryption from the weil pairing, in: CRYPTO, 2001, pp. 213–229.
[14] X. Boyen, B. Waters, Full-domain subgroup hiding and constant-size group signatures, in: Public Key Cryptography, 2007, pp. 1–15.
[15] B. Waters, Ciphertext-policy attribute-based encryption: an expressive, efficient, and provably secure realization, in: Public Key Cryptography, 2011, pp. 53–70.
[16] J. Freudiger, M.H. Manshaei, J.P. Hubaux, D.C. Parkes, On non-cooperative location privacy: a game-theoretic analysis, in: ACM Conference on Computer and Communications Security (CCS), 2009, pp. 324–337.
[17] D. Boneh, X. Boyen, Short signatures without random oracles and the sdh assumption in bilinear groups, Journal of Cryptology 21 (2) (2008) 149–177.
[18] X. Liang, Z. Cao, J. Shao, H. Lin, Short group signature without random oracles, in: International Conference on Information and Communications Security (ICICS), 2007, pp. 69–82.
[19] B. Lynn, PBC Library, <http://crypto.stanford.edu/pbc/>.
[20] Multiprecision Integer and Rational Arithmetic C/C++ Library. <http://www.shamus.ie/>.
[21] M. Rahman, M. Alhamid, A. El Saddik, W. Gueaieb, A framework to bridge social network and body sensor network: an e-health perspective, in: International Conference on Multimedia and Expo (ICME), 2009, pp. 1724–1727.
[22] W. Moncur, E. Reiter, J. Masthoff, A. Carmichael, Modeling the socially intelligent communication of health information to a patient's personal social network, IEEE Transactions on Information Technology in Biomedicine 14 (2) (2010) 319–325.
[23] T. Massey, G. Marfia, A. Stoelting, R. Tomasi, M. Spirito, M. Sarrafzadeh, G. Pau, Leveraging social system networks in ubiquitous high-data-rate health systems, IEEE Transactions on Information Technology in Biomedicine 15 (3) (2011) 491–498.
[24] R. Lu, X. Lin, X. Liang, X. Shen, Secure handshake with symptoms-matching: the essential to the success of mhealthcare social network, in: BodyNets 2010.
[25] W. Moncur, E. Reiter, J. Masthoff, A. Carmichael, Modeling the socially intelligent communication of health information to a patient's personal social network, IEEE Transactions on Information Technology in Biomedicine 14 (2) (2010) 319–325.
[26] R. Baden, A. Bender, N. Spring, B. Bhattacharjee, D. Starin, Persona: an online social network with user-defined privacy, in: SIGCOMM, 2009, pp. 135–146.
[27] E. Zheleva, L. Getoor, To join or not to join: the illusion of privacy in social networks with mixed public and private user profiles, in: International World Wide Web Conference (WWW), 2009, pp. 531–540.
[28] B. Krishnamurthy, C.E. Wills, On the leakage of personally identifiable information via online social networks, in: Workshop on Online Social Networks (WOSN), 2009, pp. 7–12.

[29] K.P.N. Puttaswamy, A. Sala, B.Y. Zhao, Starclique: Guaranteeing user privacy in social networks against intersection attacks, in: ACM Conference on emerging Networking EXperiments and Technologies (CoNEXT), 2009, pp. 157–168.

[30] X. Li, R. Lu, X. Liang, J. Chen, X. Lin, X. Shen, Smart community: an internet of things application, IEEE Communications Magazine – Feature Topic on The Internet of Things 49 (11) (2011) 68–75.

[31] F. Massacci, V.H. Nguyen, A. Saidane, No purpose, no data: goal-oriented access control for ambient assisted living, in: ACM workshop on Security and Privacy in Medical and Home-Care Systems (SPIMACS), 2009, pp. 53–58.

[32] A. Mohan, D. Bauer, D.M. Blough, M. Ahamad, B. Bamba, R. Krishnan, L. Liu, D. Mashima, B. Palanisamy, A patient-centric, attribute-based, source-verifiable framework for health record sharing, Technique reports, 2009. <www.cercs.gatech.edu/tech-reports/tr2009/git-cercs-09-11.pdf>.

[33] M. Barua, X. Liang, R. Lu, X. Shen, Peace: An efficient and secure patient-centric access control scheme for ehealth care system, in: INFOCOM Workshop on Security in Computers, Networking and Communications, 2011, pp. 970 –975.

[34] X. Liang, L. Chen, R. Lu, X. Lin, X. Shen, Pec: a privacy-preserving emergency call scheme for mobile healthcare social networks, IEEE/KICS Journal Communications and Networks 13 (2) (2011) 102–112.

[35] J. Benaloh, M. Chase, E. Horvitz, K. Lauter, Patient controlled encryption: ensuring privacy of electronic medical records, in: ACM workshop on Cloud Computing Security, 2009, pp. 103–114.

[36] B.D. Decker, M. Layouni, H. Vangheluwe, K. Verslype, A privacy-preserving ehealth protocol compliant with the belgian healthcare system, in: EuroPKI, 2008, pp. 118–133.

[37] D. Kotz, S. Avancha, A. Baxi, A privacy framework for mobile health and home-care systems, in: ACM workshop on Security and Privacy in Medical and Home-Care Systems (SPIMACS), 2009, pp. 1–12.

Please cite this article in press as: X. Liang et al., HealthShare: Achieving secure and privacy-preserving health information sharing through health social networks, Comput. Commun. (2012), doi:10.1016/j.comcom.2012.01.009