# PSaD: A Privacy-preserving Social-assisted Content Dissemination Scheme in DTNs

Linke Guo, *Student Member, IEEE,* Chi Zhang, *Member, IEEE,* Hao Yue, *Student Member, IEEE,* Yuguang Fang, *Fellow, IEEE*

***Abstract*—Content dissemination is very useful for many mobile applications, like instant messaging, file sharing, and advertisement broadcast, etc. In real life, for various kinds of time-insensitive contents, such as family photos and video clips, the process of content dissemination forms a Delay Tolerant Networks (DTNs). To improve the data forwarding performance in DTNs, several social-based approaches have been proposed, most of which leverage mobile users' social information, including contact history, moving trajectory, and personal profiles as metrics to design routing schemes. However, although the social-based approaches provide better performance, the revealing of mobile users' information apparently compromises their privacy. Moreover, users' contents may only be shared with a particular group of users rather everyone in the system. In this paper, we propose the PSaD: a Privacy-preserving Social-assisted content Dissemination scheme in DTNs. We apply users' verifiable attributes to establish their social relationships in terms of identical attributes in a privacy-preserving way. Besides, to provide the confidentiality of contents, our approach enables users to encrypt contents before the dissemination process, and only allows users who have particular attributes to decrypt them. By trace-driven simulations and experiments, we show the performance, privacy preservation, and efficiency of our proposed scheme.**

***Index Terms*—Content dissemination, Social networks, Delay tolerant networks, Privacy.**

## I. INTRODUCTION

Delay Tolerant Networks (DTNs) have drawn great attentions in recent years due to the increasing number of applications that are delay-tolerant, such as disaster recovery, military networks, and vehicular networks, all of which are in accordance with the salient nature: intermittent connectivity and unpredictable network topology. Lately, with the deployment of a tremendous number of mobile devices, e.g., laptops, smartphones, and tablets, people use these devices to opportunistically exchange and forward contents to expecting as many receivers and feedbacks as possible. The dissemination of contents potentially forms a new paradigm of DTNs, Pocket Switched Networks (PSNs) [2] that feature both opportunistic communication and mobile users' mobility. In such networks, especially for mobile users' time-insensitive data with large volume, they seek the opportunistic device-to-device dissemination instead of uploading and downloading directly from Internet in order to reduce communication costs.

For the content dissemination with device-to-device (D2D) exchange, mobile users' privacy concerns should be well addressed with respect to their contents and the way that they interact with each other. On the one hand, in some cases, mobile users may be willing to share contents with groups of receivers with specific requirements and remain undisclosed to others. For instance, imagine that Alice, who takes a high-definition photo containing at Time Square, wants to disseminate her photo to users who have the similar interests on *photography* and *travel* for social interactions. Rather than uploading to the online social networking sites, she prefers to use D2D exchange to dissemination the photo due to the expensive cellular data rate and rare free Wi-Fi hotspots, both of which are hardly to access in a heavy crowded area. Since the photo contains private information of Alice, such as identity, location, close or intimate relationships, for which she wants to keep undisclosed during the content dissemination. However, without any authentication and verification mechanism, the D2D content exchange in the crowd would incur physical and cyber attacks based on her photo, e.g., stalking her according to her current location, stealing her eye-catching belongings, or even posting her photo in the cyberspace for misuse. Unfortunately, due to the lack of pre-established end-to-end routing path and unpredictable receivers, some of existing public-key cryptographic schemes cannot preserve the confidentiality of contents while maintaining the dissemination functionality. On the other hand, considering the way mobile users interact with each others, several works apply social-based approaches to design efficient content dissemination strategies. Most of these schemes rely on users' personal social profiles and the corresponding information to determine forwarding metrics, such as the number of users they contacted before, or how frequently they have met each other, which obviously compromise mobile users' privacy. In a word, existing social-based content dissemination schemes fail to jointly address the confidentiality of contents and the privacy of users' social information.

To solve the above privacy breaches, a possible social-assisted solution would be leveraging users' social attributes to design both the dissemination strategy and privacy mechanism for contents. In sociology [3], the *homophily phenomenon* shows that people with more similar attributes contact more frequently than complete strangers. Intuitively, if users with more similarities exchange their contents with each other, the dissemination performance would become better due to their frequently contacts. We take a step further to consider users' potential social relationships as shared identical attributes, which not only enables users to determine their social-assisted dissemination strategies, but also renders a way to preserve the privacy of their social profiles by using existing cryptographic mechanisms, e.g., secure multiparty computation (SMC) [4]. Meanwhile, the confidentiality of contents can be well kept undisclosed by applying attribute-based encryption (ABE) schemes [5]. Thus, in this paper, we propose the PSaD, a social-assisted content dissemination scheme that provides the privacy of both users' contents and their social information in terms of users' attributes.

**Related Works:**

**Social-assisted Routing:** A thorough survey [6] on social-

L. Guo, H. Yue, and Y. Fang are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, 32611, USA. E-mail: {blackglk@, hyue@, fang@ece.}ufl.edu

C. Zhang is with the School of Information Science and Technology, University of Science and Technology of China, Hefei 230026, China. E-mail: chizhang@ustc.edu.cn

The preliminary version of this paper [1] was published on The 32nd IEEE International Conference on Computer Communications (INFOCOM), Turin, Italy, 2013.

based routing schemes in DTNs describes the current trends and development in this area, where Zhu *et al.* mainly compare the positive (community, centrality, similarity and friendship) and negative (selfishness) social effects in designing the routing schemes. In [7]–[10], a set of social based approaches have been proposed to improve the efficiency and effectiveness of content dissemination. Gao *et al.* study a set of multicast routing methods using both centrality and community for relay selection in [11]. One of their later works [12] designs a preference-aware data dissemination protocol using users' centrality values. However, the above approaches have two obvious drawbacks. First, the above schemes use users' social profiles, e.g., contact history, contact duration and friendships, to design their routing metrics, which is even impractical in daily life, in the sense that one may not render his/her contact history to a stranger. Second, all their schemes rely on the a posteriori knowledge (e.g., centrality values, betweenness utility, etc) to design the routing protocols. Unfortunately, as a mobile user, he/she will not know such crucial information before the experimental analysis is done, and hence they cannot forward data to the "best" relay. Wu *et al.* propose a social feature-based multi-path routing scheme in DTN [13], which is similar to our idea in using users' features (attributes) to route packets, but they do not consider the distribution of social features together with the trace file.

**Privacy-preserving Dissemination in DTNs:** In terms of privacy preservation, Lu *et al.* in [14] propose a privacy-preserving relay filtering scheme for DTN, which intends to enable contact users to filter junk packets for the destination node. Their scheme considers the privacy of the relayed packets, but it lacks of practical applications to maintain connection with spam filter users. They also propose a privacy-preserving scheme for vehicular DTNs in [15], which consider the existence of infrastructures that handle the parameter distribution and content dissemination. In [16], Hason *et al.* propose a dissemination scheme with consideration on preserving the contact history privacy, and it lacks of the discussion on the content privacy.

**Secure Profile Matching:** Besides, there are several works discussing the correlation between similarity of social profile and contact frequency [17]–[19], but none of these works jointly considers the social-assisted dissemination and privacy. Particularly, Costantino *et al.* in [20] addresses the tradeoff between privacy preservation and forwarding accuracy, which is quite similar to our work on designing the scheme for preserving the profile or attributes privacy. The major difference between our work and this work is we additionally provide the privacy of the disseminated content. One of their another work in [21] shares the similar idea with us on providing the profile privacy during the pairwise verification between mobile devices. Their scheme applies secure two-party computation for determine the interest similarity differences, while our scheme uses one-by-one zero-knowledge proof for proving the equality of each private attribute.

**Our Contributions:** Our major contributions are as follows:

- We propose a privacy-preserving mutual authentication scheme that uses the verified identical attributes to establish potential social relationships between arbitrary users.
- We design an attribute-based privacy-preserving mobile dissemination scheme that provides the confidentiality of users' content. Our scheme enables users who have the corresponding attributes required by the content owner to decrypt the content.
- To better capture the characteristics of the network, we make extensions to better fit practical scenarios in terms of setting different weights on users' attributes based on the analysis of real-world trace file.
- We apply real-world dataset to evaluate the routing performance of PSaD scheme, which outperform several existing approaches.
- Extensive simulations and experiments are conducted to verify the performance of our scheme on the aspects of security, efficiency, and feasibility.

The remainder of this paper is organized as follows. Section II introduces our intuitions and cryptographic primitives in designing the scheme. We describe the system model in Section III, along with the design objectives. The proposed scheme is presented in detail in Section IV, followed by the protocol evaluation in Section V. Finally, Section VI concludes the paper.

## II. Motivation and Preliminaries

In this section, we introduce our motivation in designing our protocols and cryptographic primitives. Moreover, by analyzing the real-world trace file, we give our preliminary results on verifying the *homophily phenomenon* and give our results on the weights of users' attributes.

### A. Motivation

To better understand the content exchange and dissemination process, we highlight our intuitive idea in the social-assisted approach. Given a pair of strangers, one cannot push another to help forward his/her message if they do not have any **pre-established relationship** or **without any incentive mechanism**. However, comparing with completely strangers, people may intend to help the one that shares some similarities in terms of attributes, e.g., language, nationality, affiliation, etc. Thus, a potential social relationship can be set up based on the similarity. It has been shown in [3] that individuals often befriend those who have similar interests, reform similar actions and frequently meet each other. Such an observation is called *homophily phenomenon*. According to [22], users who share similar interests in their profiles intend to form groups, and they can forward messages to the groups more efficiently in terms of delivery rate and delay. If users apply attribute similarity to form attribute-similar groups, the social-assisted content dissemination would be more efficient given the assumption that the attributes of contents are identical with the content owner (***homogeneity***). Here, based on our following experimental results, we assume that the attributes chosen for the dissemination have the property of homogeneity, which can be used for dissemination based on social interactions.

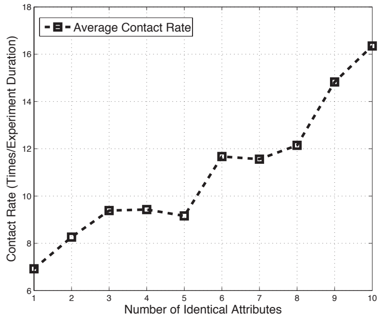


Fig. 1. Contact Rate vs. Identical Attributes

We conduct an experiment on analyzing the contact rate with the number of identical attributes based on the trace file collected during INFOCOM 2006 [23] for a period of 337,417 seconds. As we can see from Fig.1, the contact rate in terms of the number of contacts between two users increases with the increment of identical attributes. The above results validate the existence of *homophily phenomenon* and further show that the dissemination process would benefit from the content exchange between users having more identical attributes. Therefore, if the assumption ***homogeneity*** holds, leveraging identical attributes as a forwarding strategy would enhance the reliability of content dissemination process.

### B. Analysis on Weighted Attributes

In our scheme, we assume users are characterized by a unique set of attributes. We refer the term *weight* to represent the normalized ratio of contact frequency resulting from each attribute.

TABLE I
ATTRIBUTE WEIGHT VS. CONTACT DURATION

| $\omega_i$ | 0s | 60s | 120s | 300s | 600s |
|---|---|---|---|---|---|
| Language | 56.58% | 53.21% | 51.91% | 46.07% | 41.79% |
| Position | 22.69% | 21.52% | 20.8% | 18.19% | 16.28% |
| Country | 5.62% | 6.71% | 7.18% | 9.18% | 10.15% |
| City | 3.64% | 4.75% | 5.28% | 7.42% | 8.58% |
| Nationality | 3.55% | 4.03% | 4.26% | 5.08% | 5.73% |
| Affiliation | 3.13% | 3.89% | 4.20% | 5.64% | 6.60% |
| University | 2.45% | 3.06% | 3.29% | 4.37% | 5.07% |
| Stay(Hotel) | 2.31% | 2.81% | 3.05% | 4.05% | 4.77% |

According to what we observed from the above trace file, the weights on different attributes highly depend on the characteristics of both the designated network type and users. The previous work [13] concludes that the feature *Affliation* ranks the top among all features (weighs the most), and followed by *City*, *Nationality* and *Language*, etc. However, since their approach only relies on social feature sets without considering the real contact trace file, the results cannot reflect the real situation on how frequently users contact with each other using their specific attributes. We analyze 65, 536 contacts together with two involved users' attribute profiles and discover the positive correlation between identical attributes and their weights. Based on the contact trace file, we show the weights of 8 out of total 10 attributes, where we exclude the *Association Membership* and *Research Interest* for their inaccurate information. Different from what Wu and Wang observed in [13], we find out the attribute *Language* ranks the top among all attributes as shown in Table. I, which indicates that attendees mainly speak in English during the conference and seldom use other languages in formal occasions. Moreover, we observe the impact on different contact durations, due to the fact that some of the contact duration may not be sufficiently enough to exchange contents. Therefore, for a specific type of networks, the weights of users' different attributes vary from one to another in accordance with the increasing contact durations.

## III. System Model

### A. Network Model

We first give a brief introduction to the network model. As we can see from Fig. 2, the system mainly consists of a trust authority (TA) and multiple users with mobile devices. Here, TA can be any service provider which is able to run the whole system and verify the corresponding attributes, and it can be found in real life, e.g., LinkedIn on verifying users' occupations and professionals. For mobile users, they can mutually communicate with each other and TA via wireless interfaces, such as WiFi, Bluetooth, Zigbee, etc. We simply assume users stay in the transmission range of each other when they have contacts. TA can be considered as a fully-trusted infrastructure, and it is responsible for parameter distribution and attribute validation, both of which can be done before the scheme starts. Then, TA can go offline after the scheme starts except when there is a new valid user comes in. Our PSaD scheme is a two-step process, which incorporates attribute verification and content dissemination, and we will present it in Sec. IV-A in detail.

### B. Design Objectives

The main design objective of our proposed scheme is to apply users' verified identical attributes to establish relationships and disseminate content by opportunistic contacts between users. In our scheme, we assume the potential social relationship established by shared attributes can be used as incentives to help disseminate the attribute-related contents. To achieve better dissemination performance, users forward the received contents to their contact users as much as they can. For the same objective, we require users' contents should follow ***homogeneity*** assumption, such that users can disseminate contents using our social-assisted approach. Another design objective is to provide the confidentiality of user-generated content during the dissemination process. We require that users who do not have the required attributes cannot obtain contents, and hence contents should be encrypted before the dissemination. We also need to preserve the privacy of users' attributes when two users establish their social relationships.

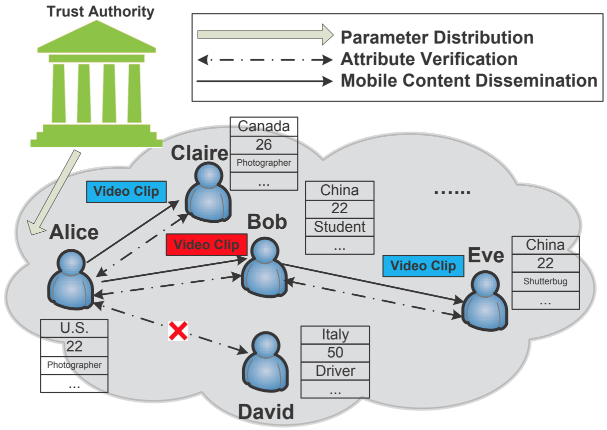


Fig. 2. System Model

## IV. Proposed Scheme

In this section, we present our PSaD scheme in detail. Based on the assumptions and intuitions listed in previous sections, we first give the scheme overview and present our scheme along with security mechanisms step by step.

### *A. Overview*

In our system, we assume each user has $|\mathcal{S}_j| = \varsigma$ attributes[1], where $\mathcal{S}_j$ represents the attribute set of a valid user $j$. TA issues certificates $\sigma$ and public/private key pairs $pk/sk$ after it verifies the authenticity of a user's claimed attributes, where $\sigma$ corresponds to users' attributes. Then, TA may go off-line.

We continue to use Alice's photo dissemination as an example to demonstrate the dissemination process. As shown in Fig. 2, Alice takes a photo $\mathcal{M}$ and only wants to share with users whose attribute set $\mathcal{S}_U$ has the following property: $\mathcal{R}_A = \mathcal{S}_U \bigcap \mathcal{S}_A$, where $\mathcal{R}_A$ is Alice's attribute requirements, e.g.,"photographers" and "teenagers". Hence, the requirement of Alice would be like "I only want to share my photo to users who have the number of identical attributes greater than $t$ as mine", in which $t$ is a number set by Alice. Then, Alice encrypts her content using threshold attribute-based encryption (*th*-ABE) [24] which allows users who have $|\mathcal{R}_A| \geqslant t$ to decrypt the ciphertext $C := E_{pk}(\mathcal{M})$. We design the following format of packet: $\{sid, E_{pk}(\mathcal{M}), TTL, Sig\}$, where *sid* is the session ID, *TTL* is the time-to-live of the packet, and *Sig* is the signature of the packet.

Our scheme is a two-step process as shown in Fig.2: attribute verification and content dissemination. To increase the delivery ratio, we ask contact users who cannot decrypt the content to store and forward encrypted content to next contact users, in order to increase the possibility of reaching users who satisfied Alice's requirements as given in Algorithm. 1 in the below.

**Algorithm 1** Pseudo Code of PSaD Scheme

```
Each pair of contact users in the dataset;
for All nodes $U_i$ contacts do
    Label each node as $U_{ij}$;
    Verify the authenticity and equality of $\sigma$;
    Determine the policy on attributes as $\mathcal{R}_i$;
    if $|\mathcal{R}_i \bigcap \mathcal{S}_{U_{ij}}| \geqslant t$ then
        Count hop = 1; Set Delay = 0;
    end
    else if $|\mathcal{R}_i \bigcap \mathcal{S}_{U_{ij}}| \geqslant \theta$ then
        Forward packets to $U_{ij}$;
        Label $U_{ij}$'s contact users as $U_{ijk}$;
        if $|\mathcal{R}_i \bigcap \mathcal{S}_{U_{ijk}}| \geqslant t$ then
            Count hop += hop ;
            Set Delay += InterContactTime;
        end
        else
            repeat
                Forward packets to next contact users;
                Count hop += hop ;
                Set Delay += InterContactTime;
            until No available user satisfy $(\theta, t)$;
        end
    else
        Drop the packet;
    end
end
```

**Attribute Verification:** First, when Alice opportunistically meets a stranger Bob who may potentially help her forward packets, she first uses the certificate $\sigma$ to mutually check the attribute similarity between Bob and her using non-interactive witness-indistinguishable proof. For example, if she requires the attribute *position:photographer* or *age:22* as the forwarding criterion, she will only give packets to users who have that verified $|\mathcal{S}_{U,\sigma}| = \theta$, where $\mathcal{S}_{U,\sigma}$ is the certificate set of user $U$'s verified attributes. Note this criterion is different from $t$, where $\theta$ is designed for checking the forwarding requirement, while $t$ is the encryption policy.

**Content Dissemination:** If Bob successfully proves to Alice that he has $\sigma$, Alice forwards packets to Bob. Otherwise, Alice aborts the protocol and waits for another contact user. Having Alice's packet, Bob first attempts to decrypt it using his private key $sk$ corresponding to his attribute set $\mathcal{S}_B$. If $|\mathcal{S}_B \bigcap \mathcal{S}_A| \geqslant t$, he can view the video clip $\mathcal{M}$ and he will not disseminate $\mathcal{M}$ to his next contact user, in the sense we consider the dissemination process completes. Otherwise, Bob stores and forwards the packets to the next contact user. As we can see from Fig. 2, Bob only has the certificate of *age* that meets Alice's criterion, and thus he is not able to decrypt $C$. Then, he forwards the packets to Eve, whom he opportunistically contacts, after they mutually verify their certificates on attribute *age*. Here, we assume Bob can only verify the attributes which are successfully verified between him and Alice with the next contact user. Otherwise, Bob may maliciously add unnecessary attributes on Alice's contents, and further disables receivers from decrypting the contents. For practical concern, if Bob does not hold the certificates of *photographer*, he is not able to even initiate the verification process. Finally, since Eve's attributes satisfy $\mathcal{R}_A \geqslant t$, she is able to decrypt the ciphertext.

### *B. Setup*

*1) Global parameter initialization:* Given the security parameter $\xi$, TA first generates a parameter tuple $(p, G_1, G_2, G_T, e) \leftarrow 1^{\xi}$, where $e$ is a bilinear map $e : G_1 \times G_2 \rightarrow G_T$ which has the properties of *bilinearity*, *computability*, and *non-degeneracy* [25], [26]. Then, TA randomly selects three generators $g_0, g_1 \in G_1$, $g_2 \in G_2$ as well as two random numbers $\beta, \gamma \in Z_p^*$, and sets $w = g_0^{\beta\gamma} \in G_1$ and $z = e(g_0^{\beta}, g_2) \in G_T$. TA generates a common reference string (*crs*) as follows, sets $u_1 := u_2^r \in G_1^2$ and $v_1 := v_2^s \in G_2^2$, where $u_2 = (g_1, g_1^a) \in G_1^2$ and $v_2 = (g_2, g_2^b) \in G_2^2$ and $a, b, r, s \in Z_p$ are randomly chosen by TA. Finally, TA distributes the $crs := (p, G_1, G_2, G_T, e, g_1, g_2, u_1, u_2, v_1, v_2)$ to every valid user for the purpose of attribute verification.

*2) Key distribution:* Based on the system parameters, TA sets master secret key as $\mathsf{msk} := (g_0, \beta, \gamma)$. Then, we use $m \in Z_p^*$ to denote user's attribute values, and use $\mathcal{S}$ to represent the universe set of attribute values in the system, where $|\mathcal{S}| = \alpha$. Moreover, TA chooses a set of dummy attributes $\mathcal{D} = \{\delta_1, ..., \delta_{\alpha-1}\}$ which consists of $\alpha - 1$ pairwise different elements of $Z_p^*$. Then, TA issues public/private key pair to users as $pk := \{\mathcal{S}, \alpha, w, z, \{g_2^{\beta\gamma^i}\}_{i=0,\ldots,2\alpha-1}, \mathcal{D}\}$. For each valid user $j$ with corresponding verified attributes $\mathcal{S}_j$, TA picks a random number $\kappa_j \in Z_p^*$, and renders user $j$ a private key as $sk_j := \{\{g_0^{\frac{\kappa_j}{\gamma+m_i}}\}_{m_i \in \mathcal{S}_j, i=1,\ldots,\varsigma}, \{g_2^{\kappa_j\gamma^i}\}_{i=0,\ldots,\alpha-2}, g_2^{\frac{\kappa_j-1}{\gamma}}\}$. We apply the constant-size ciphertext threshold attribute-based encryption scheme in [24] to deploy our secure content dissemination. The encryption scheme relies on the intractability of augmented multi-sequence of exponents decisional Diffie-Hellman ($\mathsf{mse-DDH}$) problem, which is generated from the hardness problem DDH [27].

[1] If $X$ is a set, $|X|$ means its cardinality; if $X$ is a number, $|X|$ denotes the length of bits representing the number.

*3) Attribute Validation and Certificate Issuance:* For each verified attribute, apart from distributing corresponding public/private key pairs, TA also issues users certificates $\sigma$ on different values of attributes after it validates the authenticity of attribute values. Here, we use Boneh-Boyen signature scheme [28] to sign each attribute value, which is secure under weak chosen message attacks based on the $q-\mathsf{SDH}$ assumption. TA uses $x_\imath \in Z_p^*$ to sign different values of one specific attribute, e.g., *photographer* and *artist* in attribute classification *position*. We can rewrite each user's attribute as $m_{\imath.\jmath} \in Z_p^*$, where $\imath$ is the general classification of attributes and $\jmath$ denotes the specific value of the attribute. Note that both $\imath$ and $\jmath$ are the indices in its own set. Given a verified attribute $m_{\imath.\jmath}$, TA outputs a certificate as $\sigma_{\imath.\jmath} = g_1^{1/(x_\imath + m_{\imath.\jmath})}$ and a verification key $k_\imath = g_2^{x_\imath}$ and gives them to valid users. In what follows, we use $m_{\imath.\jmath}^i$ to denote user $i$'s attribute value $\jmath$ on attribute classification $\imath$. For privacy concerns, users cannot reveal the certificates and keys for verification, otherwise malicious users may perform impersonation attacks. Moreover, for the same category of attribute values, the verification keys remain the same and they should be kept undisclosed as well.

### C. Attribute Verification

The attribute verification process takes place when Alice opportunistically contacts a user, say Bob, and decides whether forwards the ciphertext to him based on Alice's requirements of $\sigma$. In what follows, we design a mutual verification scheme by applying non-interactive wittiness-indistinguishable (NIWI) proof.

*1) Commitment and Proof Generation:* We apply part of the non-interactive proof system for bilinear map discussed in [29], [30] and make extensions to fit our proposed scenario and scheme. To check the validity of corresponding certificates, the verifier needs to check whether the following equation is satisfied:

$$e(\sigma_{\imath.\jmath}, k_\imath \cdot g_2^{m_{\imath.\jmath}}) = e(g_1, g_2).^{2} \qquad (1)$$

For privacy concern, we require that the verification should not reveal the certificates and verification keys. We use the following approach to generate commitments and proofs on the certificate and verification key, and further verify the equality of the above equation without exposing the secret information. Alice first picks random numbers $r_1, r_2, s_1, s_2 \in Z_p$, then she commits $\sigma_{\imath.\jmath}$ and $k_\imath$ in the following way: $c^A = (1, \sigma_{\imath.\jmath}) u_1^{r_1} u_2^{r_2}$ and $d^A = (1, k_\imath g_2^{m_{\imath.\jmath}}) v_1^{s_1} v_2^{s_2}$ without exposing the value of $\sigma_{\imath.\jmath}$ and $k_\imath$ of Alice[3]. Note that $u_1, u_2 \in G_1^2, v_1, v_2 \in G_2^2$ come from pre-loaded *crs* and $r_1, r_2, s_1, s_2$ are randomly chosen from $Z_p$. Based on the commitment $\{c^A, d^A\}$, Alice continues to generate NIWI proofs as follows,

$$\pi_1^A := v_1^{t_{11}} v_2^{t_{12}} d_i^{r_1}, \pi_2^A := v_1^{t_{21}} v_2^{t_{22}} d_i^{r_2};$$

$$\psi_1^A := u_1^{-t_{11}} u_2^{-t_{21}} \sigma_{\imath.\jmath}^{s_1}, \psi_2^A := u_1^{-t_{12}} u_2^{-t_{22}} \sigma_{\imath.\jmath}^{s_2}.$$

where $t_{11}, t_{12}, t_{21}, t_{22} \in Z_p$ are randomly selected by Alice. Finally, Alice sends Bob a set of commitments along with proofs as $\langle c^A, d^A, \pi_1^A, \pi_2^A, \psi_1^A, \psi_2^A \rangle$ for further verification.

*2) Authenticity Verification:* The verification is a two-step process. First, the verifier (contact user, say Bob) checks the authenticity of the certificate set of Alice. Similarly, Bob sends back his commitment and proof set required by Alice. Then, Alice also checks the authenticity of Bob's verified attributes. Without loss of generality, we give the verification process performed by Bob when he obtains Alice's commitments and proof set.

With the entry-wise multiplication of $G_1$ and $G_2$, we obtain $Z_p$-modules $G_1^2$ and $G_2^2$. Then, given the map $G_1 \times G_2 \to G_T$, the entry-wise multiplication also makes $G_T^4$ a $Z_p$-module. Thus, there is a bilinear map $\hat{e} : G_1^2 \times G_2^2 \to G_T^4$,

$$\hat{e}\left(\begin{pmatrix} g \\ h \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix}\right) \to \begin{pmatrix} e(g,x) & e(g,y) \\ e(h,x) & e(h,y) \end{pmatrix}.$$

To verify the certificates satisfying Eq. (1), for each attribute, Bob checks whether the following equation is hold,

$$\hat{e}(c^A, d^A) = \begin{pmatrix} 1 & 1 \\ 1 & \Gamma \end{pmatrix} \circ \hat{e}(u_1, \pi_1^A) \circ \hat{e}(u_2, \pi_2^A) \circ \hat{e}(\psi_1^A, v_1) \circ \hat{e}(\psi_2^A, v_2) \quad (2)$$

where $\Gamma = e(g_1, g_2)$ is known to all users in the system. It has been proven in [29] that Eq. (1) holds when Eq. (2) can be successfully deduced. We also refer the matrix operation $(\circ)$ in Eq. (2) as the Schur product, in which given two $n \times n$ matrixes $\mathbf{A}$ and $\mathbf{B}$, the Schur product is $(\mathbf{A} \circ \mathbf{B})[i,j] = \mathbf{A}[i,j]\mathbf{B}[i,j]$.

Due to page limit, we only deduce the verification process of the element in the second row and second column of the corresponding matrix in Eq. (2). Thus, the left hand side can be written as,

$$\begin{aligned} LHS_{22} &= e(g_1^{\frac{1}{x_\imath + m_{\imath.\jmath}} + arr_1 + ar_2}, g_2^{x_\imath + m_{\imath.\jmath} + bss_1 + bs_2}) \\ &= e(g_1, g_2)^{1 + \frac{bss_1 + bs_2}{x_\imath + m_{\imath.\jmath}} + (x_\imath + m_{\imath.\jmath} + bss_1 + bs_2)(arr_1 + ar_2)}. \end{aligned}$$

Meanwhile, the element in the same position on right hand side can be derived as,

$$\begin{aligned} RHS_{22} &= \Gamma \cdot e(g_1^{ar}, g_2^{bst_{11} + bt_{12} + (x_\imath + m_{\imath.\jmath} + bss_1 + bs_2) r_1}) \\ &\cdot\ e(g_1^{a}, g_2^{bst_{21} + bt_{22} + (x_\imath + m_{\imath.\jmath} + bss_1 + bs_2) r_2}) \\ &\cdot\ e(g_1^{-art_{11} - at_{21} + \frac{s_1}{x_\imath + m_{\imath.\jmath}}}, g_2^{bs}) \\ &\cdot\ e(g_1^{-art_{12} - at_{22} + \frac{s_2}{x_\imath + m_{\imath.\jmath}}}, g_2^{b}) \\ &= \Gamma \cdot e(g_1, g_2)^{(x_\imath + m_{\imath.\jmath} + bss_1 + bs_2)(arr_1 + ar_2) + \frac{bss_1 + bs_2}{x_\imath + m_{\imath.\jmath}}} \end{aligned}$$

which shows the equality of both sides. Till now, Bob is convinced that Alice has the verified attributes. Similarly, since our protocol is executed bilaterally, Alice can also be convinced that Bob has a set of verified attributes.

*3) Equality Verification:* The second step of verification process is to verify the equality between two users' attributes, which requires only users who share the same attributes can know the comparison results, otherwise both of them learn nothing about both certificates and verification keys.

Suppose Alice wants to make sure that Bob has her required attributes. She sends her randomly selected parameters used to generate $\langle c^A, d^A, \pi_1^A, \pi_2^A, \psi_1^A, \psi_2^A \rangle$ to Bob, which are $\langle r_1, r_2, s_1, s_2 \rangle$. Then, Bob is able to generate another set of NIWI proofs to show the equality of their attributes. During the previous process, Bob has obtained Alice's commitments on $\sigma_{\imath.\jmath}$ and $k_\imath g_2^{m_{\imath.\jmath}}$, then he can construct the equality NIWI proofs as follows,

$$\tilde{\pi}_1 := v_1^{t'_{11}} v_2^{t'_{12}} (d^A)^{r'_1}, \tilde{\pi}_2 := v_1^{t'_{21}} v_2^{t'_{22}} (d^A)^{r'_2};$$

$$\tilde{\psi}_1 := u_1^{-t'_{11}} u_2^{-t'_{21}} (\sigma_{\imath.\jmath}^B)^{s_1}, \tilde{\psi}_2 := u_1^{-t'_{12}} u_2^{-t'_{22}} (\sigma_{\imath.\jmath}^B)^{s_2}$$

where the parameters $r'_1, r'_2, t'_{11}, t'_{12}, t'_{21}, t'_{22} \in Z_p$ are randomly

[2] The operation $(\cdot)$ is multiplication operation on the corresponding groups, e.g., $G_1$, $G_2$, and $G_T$.

[3] The multiplication on $G_1^2$ and $G_2^2$ is defined as entrywise multiplication.

chosen by Bob for generating his commitments and proofs.

The equality verification process is similar to the authenticity verification. When Bob sends the equality NIWI proofs to Alice, she checks whether the following equation is satisfied,

$$\hat{e}(c^A, d^B) = \begin{pmatrix} 1 & 1 \\ 1 & \Gamma \end{pmatrix} \circ \hat{e}(u_1, \tilde{\pi}_1) \circ \hat{e}(u_2, \tilde{\pi}_{i2}) \circ \hat{e}(\tilde{\psi}_1, v_1) \circ \hat{e}(\tilde{\psi}_2, v_2). \quad (3)$$

If the equation holds, it implies that $e(g_1, g_2)^{\frac{x_\imath + m^B_{\imath.\jmath}}{x_\imath + m^A_{\imath.\jmath}}} = e(g_1, g_2)$, such that Alice's and Bob's certificates on this attribute are the same. Then, they are aware that they share the same attribute values after the equality verification process. On the other hand, if they do not have the identical attributes, the output of equality verification is a random number in $G_T$, which prevents users from knowing other users' certificates and verification keys. Users may compare more attributes by creating multiple commitment and proof sets based on their verified certificates and verification keys. Therefore, the potential social relationships can be established when users share identical attributes.

### *D. Privacy-Preserving Content Dissemination*

The dissemination process happens when Alice and Bob have established potential social relationships based on Alice's criterion, she wants Bob to help her disseminate contents. Once users whose private keys can decrypt the encrypted contents, we consider the dissemination process is accomplished for these users. The dissemination process can be divided into the following steps: *policy setup*, *content encryption*, and *user decryption scheme*. Apart from our basic privacy-preserving dissemination scheme, we also give our advanced scheme based on the analysis on attribute weights in Sec. II-B.

*1) Policy setup:* We continue to use $\imath$ to represent the classification, where $\mathcal{S}^\imath$ is the set of attribute values in attribute $\imath$, and $m_{\imath.\jmath}$ is the elements of $\mathcal{S}^\imath$. Before presenting our scheme in detail, we first list the following definitions,

*Def 1*: Uniqueness: $\forall$ user $i$, given $m_{\imath.\jmath} \in \mathcal{S}^\imath$, we require $\forall 1 \leqslant \jmath, \jmath' \leqslant |\mathcal{S}^\imath|$, $\nexists\ m^i_{\imath.\jmath}, m^i_{\imath.\jmath'} \in \mathcal{S}^\imath_i$, where $\jmath \neq \jmath'$;

*Def 2*: Completeness: $\forall$ user $i$, $\forall 1 \leqslant \imath \leqslant |\mathcal{S}_i|$, $\exists\ m^i_{\imath.\jmath} \in \mathcal{S}^\imath_i$.

The uniqueness property requires each user has a specific value of an attribute, and he/she cannot simultaneously have multiple values on one attribute. Moreover, we require that users have corresponding values on all of their attributes. To create the encryption policy, user $i$ chooses subsets from $\mathcal{S}^\imath$, and combines them together as her policy to encrypt her contents. Then, we define user $i$'s policy as

$$\mathcal{P}_i := \bigcup_{\imath=1}^{|\mathcal{S}_i|} \bigcup_{\jmath=1}^{|\tilde{\mathcal{S}}^\imath_i|} m_{\imath.\jmath},$$

where $\tilde{\mathcal{S}}^\imath_i$ denotes user $i$ attribute value requirement on classification $\imath$, e.g., he/she requires *photographer*, *artist* or *movie director* under attribute classification *position*. Then, user $i$ applies the policy $\mathcal{P}_i$ to encrypt $\mathcal{M}$ using the following encryption scheme.

*2) Content Encryption:* Before presenting the encryption scheme, we first list the requirements for encrypting the generated content. As we illustrated before, the content could be video clips, high-quality images or a large chunk of data, all of which consume not only a large amount of storage, but also the transmission bandwidth and power. Since users apply attributes as the encryption scheme, we need to consider the cost-effectiveness of our proposed scheme in practical scenarios. If the size of $|\mathcal{P}|$ is large, traditional encryption scheme [5] makes the ciphertext size of contents grow linearly, which impedes the possibility of content exchange in DTNs. Thus, we apply the constant-size ciphertext threshold attribute-based encryption scheme [24] in our proposed protocol to avoid large ciphertexts.

In our basic scheme, users treat every attribute equally without considering attribute weights. According to *Def 1*, users only have one value on one specific classification of attributes, which implies that the threshold of $t$ is the number of attributes from different classifications. Taking Alice as an example, she wants to disseminate her content $\mathcal{M} \in G_T$ throughout the whole network with a corresponding policy $\mathcal{P}_A$. We denote that the cardinality of $|\mathcal{P}_A| = \rho_A$, and the threshold $t_A$ satisfies $1 \leqslant t_A \leqslant \varsigma_A \leqslant \rho_A \leqslant \alpha$. The encryption scheme is as follows. Alice first chooses a random number $\mu \in Z^*_p$ and computes,

$$C_1 = w^{-\mu}, C_2 = g_2^{\mu\beta \prod_{m^A_{\imath.\jmath} \in \mathcal{P}_A} (\gamma + m^A_{\imath.\jmath}) \prod_{\delta \in \mathcal{D}_\eta} (\gamma+\delta)},$$

$$K = z^\mu, C_3 = K \cdot \mathcal{M},$$

where the ciphertext is $(C_1, C_2, C_3)$ and $K$ is the decryption key. Note that the ciphertext $C_2$ includes a set of $\eta = \alpha + t_A - \rho_A - 1$ dummy attributes, in order to obtain a polynomial of degree $\alpha + t_A - 1$ in the exponent of $g_2$. Alice uses the parameters $\{g_2^{\beta\gamma^i}\}_{i=0}^{2\alpha-1}$ in $pk$ to construct $C_2$. Since the threshold $t_A \leqslant \alpha$, the maximum degree of the exponent of $g_2^{\beta\gamma^i}$ is $2\alpha - 1$, then,

$$\begin{aligned} C_2 &= g_2^{\mu\beta \prod_i (\gamma + l_i) \prod_{\delta_j \in \mathcal{D}_{\alpha+t_A-\rho_A-1}} (\gamma+\delta_j)} \\ &= \left( g_2^{\beta(\sum_{i=0}^{\rho} \binom{\rho}{i} \gamma^{\rho-i} l_i^i)} \cdot g_2^{\beta(\sum_{j=0}^{\eta} \binom{\eta}{j} \gamma^{\eta-j} \delta_j^j)} \right)^\mu \\ &= \left( g_2^{\beta\gamma^{\rho+\eta}} \cdots g_2^{\beta \cdot l_\rho^\rho \delta_\eta^\eta} \right)^\mu, \end{aligned}$$

where $\max(\rho+\eta) = 2\alpha - 1$. Then, Alice forwards the ciphertext together with a signature (ref. Sec. V-A3) to Bob.

*3) User Decryption Scheme:* The decryption process is as follows. First, Bob runs the Aggregate algorithm using $\{g_0^{\frac{\kappa_B}{\gamma+m_i}}\}_{m_i \in \mathcal{S}_B, i=1,\ldots,\varsigma}$ and attribute values $m^B_{\imath.\jmath}$. For the ease of description, we refer $x_\imath := m^B_{\imath.\jmath}$ for all $\imath \in \mathcal{S}_B$ and $|\mathcal{S}_B| = \varsigma$. Define for any $(j, l)$ such that $1 \leqslant j < l \leqslant \varsigma$, $P_{j,l} = g_0^{\frac{1}{\gamma+x_l} \cdot \frac{\kappa}{\prod_j (\gamma+x_j)}}$. The Aggregate algorithm consists in computing sequentially $P_{j,l}$ for $j = 1, 2, ..., \varsigma - 1$ and $l = j+1, j+2, ..., \varsigma$ using the induction

$$P_{j,l} = \left( \frac{P_{j-1,j}}{P_{j-1,l}} \right)^{\frac{1}{x_l - x_j}}$$

and setting $P_{0,l} = g_0^{\frac{\kappa}{\gamma+x_l}}$ for $l = 1, 2, ..., \varsigma$. The algorithm finally outputs $P_\varsigma = P_{\varsigma-1,\varsigma} = g_0^{\frac{\kappa}{\prod_i (\gamma+x_i)}}$. Note that this computation process can be accomplished before the protocol run.

Second, if $|\mathcal{S}_A \bigcap \mathcal{S}_B| \geqslant t_A$, Bob can compute the following parameter $L := e(g_0^{\frac{\kappa_B}{\prod_{m^B_{\imath.\jmath}} (\gamma + m^B_{\imath.\jmath})}}, C_2)$. Then, he uses $\{g_2^{\kappa_B \gamma^i}\}_{i=0}^{\alpha-2}$ to compute the following parameters. Since the degree of Bob's $sk$ is less than $\alpha - 2$, we need to guarantee he is able to construct the polynomial with degree no more than $\alpha - 2$. Thus, we first define a polynomial,

$$Y(\gamma) = \frac{1}{\gamma} \left( \frac{\prod\limits_{m_{\imath.\jmath} \in \mathcal{P}_A} (\gamma + m_{\imath.\jmath}) \prod\limits_{\delta \in \mathcal{D}_\eta} (\gamma + \delta)}{\prod\limits_{m_{\imath.\jmath} \in \mathcal{S}_A \bigcap \mathcal{S}_B} (\gamma + m_{\imath.\jmath})} - X_{(\mathcal{S}_A \bigcap \mathcal{S}_B, \mathcal{P}_A)} \right)$$

where $X_{(\mathcal{S}_{AB},\mathcal{P}_A)}=\frac{\prod_{m_{\imath.\jmath}\in\mathcal{P}_A} m_{\imath.\jmath}\prod_{\delta\in\mathcal{D}_\eta}\delta}{\prod_{m_{\imath.\jmath}\in\mathcal{S}_A\bigcap\mathcal{S}_B} m_{\imath.\jmath}}$. Note that $\deg(Y(\gamma)) < \alpha-2$, which indicates the parameters in $sk$ are enough to generate the decryption key $K$ if $|\mathcal{S}_A\bigcap\mathcal{S}_B| \geqslant t_A$.

Then, Bob uses $\{g_2^{\kappa_B\gamma^i}\}_{i=0}^{\alpha-2}$ to calculate,

$$e(C_1, g_2^{\kappa_B Y(\gamma)})\cdot L \quad = \quad e(g_0,g_2)^{\mu\cdot\kappa_B\cdot\beta X_{(\mathcal{S}_A\cap\mathcal{S}_B,\mathcal{P}_A)}} \tag{4}$$

and,

$$e(C_1, g_2^{\frac{\kappa_B-1}{\gamma}}) \quad = \quad e(g_0,g_2)^{\beta\mu-\beta\mu\kappa_B} \tag{5}$$

As we can see, if $|\mathcal{S}_A\bigcap\mathcal{S}_B| < t_A$, which indicates the degree of $X_{(\mathcal{S}_{AB},\mathcal{P}_A)}$ is greater than $\alpha-2$, Bob fails to construct $g_2^{\kappa_B Y(\gamma)}$. Then, Bob can derive the decryption key $K$ by combining two equations Eq. (4) and Eq. (5),

$$\begin{aligned} K &= (e(C_1, g_2^{\kappa_B Y(\gamma)})\cdot L)^{1/X_{(\mathcal{S}_{AB},\mathcal{P}_A)}} e(C_1, g_2^{\frac{\kappa_B-1}{\gamma}}) \\ &= e(g_0,g_2)^{\beta\mu} = z^\mu. \end{aligned}$$

Finally, Bob uses the derived key $K$ to recover the content by computing $\mathcal{M} = C_3\cdot K^{-1}$.

For users who cannot decrypt the contents, they further forward encrypted contents to the next contact user after repeating the equality verification on certificates.

*4) Advanced scheme for weighted attributes:* Our basic approach defines the threshold $t$ as the number of identical attributes between Alice and Bob. However, based on our observations, some of users' attributes become more important in resulting the contacts. For example, in Sec. II-B, we discuss the weighted attributes in each contact, in which the attribute *Language* weighs the top according to the INFOCOM 2006 trace file. Here, we distinguish the difference between *Network-determined Weight (NDW)* $\{\omega_i\}_{i=1}^{\alpha}$ and *User-defined Weight (UDW)* $\{\nu_i\}_{i=1}^{\alpha}$, where *Network-determined Weight* denotes the characteristic weight of designated network and it varies depending on different types of networks. *User-defined Weight* is the weight policy that the content generator defined, e.g., he/she only wants users who have attributes *Professor* and *College of Engineering* to decrypt the content, so he/she assigns more weight on those attributes.

- *Network-determined Weight Encryption Scheme*

For *NDW*, we redefine the universal attribute set as $\mathcal{S}' := \{m_1||1, m_1||2, ...m_1||\Omega, ..., m_\alpha||1, ..., m_\alpha||\Omega\}$, where $\Omega$ is the maximum integral weight of the corresponding attribute and $||$ is the concatenation between attribute classification and its weighed value. To better fit the encryption and decryption scheme, we define a mapping $\omega : Q \rightarrow Z^+$, which maps the original weight $w_i$ to positive integral values. Depending on different types of networks, users are connected or contacted based on different attribute set, e.g., the trace file collected at INFOCOM 2006 shows that people with same affiliation or nationality contact more frequently. Users may design their policies $\mathcal{P}$ only using *NDW* to better fit the characteristic of the corresponding networks. Thus, we modify our basic scheme for weighted attributes as follows. Based on *NDW* $\{\omega_i\}$, TA distributes elements in secret keys as $\{g_0^{\frac{\kappa}{\gamma+m_i^j}}\}_{i=1}^{\varsigma}$, where $m_i^j = m_i||j$ for $j = 1, 2, ..., \Omega$. According to the *NDW* value $\omega_i$, TA assigns different values of $j$ to attribute classification, which enables attributes with more weights performing crucial roles in decrypting the ciphertext. For each attribute classification $i$ with maximum weight $\omega_i := \Omega_i$, TA renders a set of secret key elements $\{g_0^{\frac{\kappa}{\gamma+m_i^j}}\}_{j=1}^{\Omega_i}$ other than a single value $g_0^{\frac{\kappa}{\gamma+m_i}}$. Comparing to the basic scheme, for the receiver, he/she carries more $g_0^{\frac{\kappa}{\gamma+m_i^j}}$ elements to construct $L'$, which makes the degree of the exponential part of $L'$ less than $\alpha-2$ only with the condition $\sum_{i=1}^{|\mathcal{P}|}\Omega_i \geqslant t$. Thus, he/she is able to decrypt the ciphertext if he/she has several crucial attributes that satisfy sender's encryption policy $\mathcal{P}$ other than requiring no less than $t$ same attribute values.

- *User-determined Weight Encryption Scheme*

Since our basic scheme does not specify the weights of particular attributes, it would be possible that several receivers do not really satisfy the content generators' requirements on specific attributes, but they are able to decrypt the encrypted content because they share more than $t$ unspecified attributes. Although we achieve the *NDW* by expanding the private key set, it is infeasible for us to apply the same technique in the user-determined encryption scheme, because, in our scheme, the content generator and receiver do not each other beforehand. Therefore, they cannot exchange information based on their encryption policy. To achieve the fine-grained access policy, we intend to solve the above problem by introducing our advanced scheme, *User-determined Weight Encryption Scheme (UDW)*, which prioritizes required attributes in decrypting the content.

To better extend existing encryption scheme, we apply the CP-ABE scheme proposed in [31] and make several significant extensions in order to realize *UDW*. Here, we provide the encryption scheme with *UDW* as $\nu_i := \{1, 1, ..., -1, -1, ...*, *, ..\}$, where $1$ denotes required attributes, $-1$ is the attributes that the content generator excludes and $*$ represents that the content generator does not specify it and it can be any value. For example, Alice can weigh *Position: Professor* as $1$ for required attributes, while assigning $*$ on *Nationality: United States* for *do not care* attributes. Therefore, if Alice and Bob have more than $t_A$ same attributes, Bob can decrypt the packet when he has required attributes, which further filters out desired users.

By applying the scheme in [31], we modify the ciphertext in our basic scheme as $(C_1, C_2, C_3', C_4)$, where $C_3' = C_3\cdot K'$, and $C_4$ is ciphertext of $K'$ generated by *UDW* scheme. Note the setup and key generation process can be done before the protocol run, and TA is responsible for distributing the public parameter and private keys for users with corresponding attributes. If the potential receiver is able to use his/her private keys to obtain $C_3$, he/she satisfies *UDW* required by the content generator. Then, he/she can use threshold-based decryption scheme to see whether he/she is able to disclose the content $\mathcal{M}$. We continue to use Alice's content dissemination as an illustrative example, and we assume DBDH assumption [25] is hard in our advanced scheme. The advanced *UDW* scheme is run as follows,

**Setup:** Based on the public parameter defined in previous section, TA chooses the master secret key for *UDW* scheme as $\mathtt{mk} := \langle y, h_k\rangle \in Z_p$, where $k \in \{1, .., 3\varsigma\}$ denotes the index of each attributes on different weights. Then, it sets the public key as $\mathtt{pk} := \langle Y, H_k\rangle$, where $Y = e(g_0,g_2)^y$ and $H_k = g_2^{h_k}$. $H_k, H_{\varsigma+k}, H_{2\varsigma+k}$ correspond to three types of occurrences of $k$, positive, negative and do not care.

**KeyGen:** For every $m \in \mathcal{S}$, TA chooses a random $l_k \in Z_p$, and sets $l := \sum_{k=1}^{\alpha} l_k$. Let $\hat{D} := g_2^{y-\lambda_U l}$ and let $D_k = g_0^{\frac{\lambda_U l_k}{h_k}}$ if $k \in \mathcal{S}_U$ of a particular user $U$, where $\lambda_U \in Z_p^*$ is a private parameter selected by TA for each user; $D_k = g_0^{\frac{\lambda_U l_k}{h_{\varsigma+k}}}$ for $k \in$

$\bar{\mathcal{S}}_U$, where $\bar{\mathcal{S}}_U$ denotes attributes that user $U$ does not have. Set $D_k := g_0^{\frac{\lambda_U l_k}{h_{2\varsigma+k}}}$ for else attributes. Finally, TA renders private keys to user $U$ as $\texttt{sk} := \langle \hat{D}, D_k, E = g_0^{\lambda_U} \rangle$.

**Encryption:** Alice first chooses her required attributes in her attribute set $\mathcal{S}$, and labels other attributes as excluded attributes and do not care attributes. Then, Alice chooses $c \in Z_p$, encryption key $K' \in G_T$, to construct the ciphertext as,

$$C_4 := \Big( \bigcap_{m_k \in \mathcal{P}_A} m_k, C' := K' \cdot Y^c, \hat{C} := g_0^c, \{\tilde{C}_k | m_k \in \mathcal{S}\} \Big).$$

Here, $\tilde{C}_k = H_k^c$ for $m_k \in \mathcal{P}_A$ and $\tilde{C}_k = H_{\varsigma+k}^c$ for $m_k$ Alice does not share with. In particular, $\tilde{C}_{\varsigma+k} = H_{2\varsigma+k}^c$ for the attributes that Alice does not care.

**Decryption:** For $m_k \in \mathcal{P}_A$, the receiver, say Bob, computes $e(\tilde{C}_k, D_k) = e(g_0^{h_k \cdot c}, g_2^{\frac{\lambda_B l_k}{h_k}}) = e(g_0, g_2)^{\lambda_B l_k c}$; For $m_k \in \bar{\mathcal{P}}_A$, he computes $e(\tilde{C}_k, D_k) = e(g_0^{h_{\varsigma+k} \cdot c}, g_2^{\frac{\lambda_B l_k}{h_{\varsigma+k}}}) = e(g_0, g_2)^{\lambda_B l_k c}$; For other do not care attributes, Bob computes $e(\tilde{C}_k, D_k) = e(g_0^{h_{2\varsigma+k} \cdot c}, g_2^{\frac{\lambda_B l_k}{h_{2\varsigma+k}}}) = e(g_0, g_2)^{\lambda_B l_k c}$. Then, based on these results, Bob derives the following results,

$$\begin{aligned} Y^c &= e(\hat{C}, \hat{D}) \cdot \prod_{k=1}^{\alpha} e(g_0, g_2)^{\lambda_B \cdot l_k \cdot c} \\ &= e(g_0^c, g_2^{y - \lambda_B l}) \cdot e(g_0, g_2)^{\lambda_B l c} \\ &= e(g_0, g_2)^{y \cdot c}. \end{aligned}$$

The encryption key $K'$ can be consequently derived as $K' = C'/Y^c$. Bob can obtain the content if and only if his attribute set satisfy both the requirement on *UDW* and Alice's identical number of attributes. The deriving process is as follows,

$$C_3' \cdot (K')^{-1} \cdot (K)^{-1} = \mathcal{M} \cdot e(g_0, g_2)^{\beta\mu} \cdot (K)^{-1} = \mathcal{M}.$$

We continue to consider the situation where both *NDW* and *UDW* affect the forwarding process. Based on the previous approaches, we can simply apply the *NDW* and *UDW* encryption and decryption scheme to guarantee the privacy-preservation dissemination, which means TA distributes a set of secret key elements and public parameters to let sender re-encrypt the ciphertext of content instead of directly sending it, and further enable contact users use *NDW* private keys to obtain the plaintext of content.

## V. Protocol Evaluation

In this section, we evaluate our proposed scheme from the aspects of security, efficiency and feasibility. Our security analysis studies how the design objectives are achieved in scheme based on the adversarial model. The efficiency of the proposed system in terms of computation and storage cost is also discussed. We also conduct simulation based on the real-world trace file to verify the feasibility of our proposed scheme.

### A. Security Analysis

*1) Adversarial Model:* For active attacks, adversaries can launch impersonation attacks to compromise the attribute privacy. We allow adversaries to change their attributes to meet the requirements of content. However, since all the user attributes should be verified by TA, adversaries who launch this kind of attack may fail to perform from the very beginning. As an active attacker, he/she may modify and inject bogus data during the transmissions. On the other hand, we are also concerned with the collusion attack among a group of malicious users who share different private keys of attributes. They may intend to decrypt contents by collecting enough valid attributes. For passive attacks, we allow adversaries to eavesdrop the communication channels during the forwarding and dissemination process. On the other hand, we do not consider the possibility of sharing secrets with others. Similarly, we exclude attackers who successfully steal other valid users' mobile devices perform attacks based on the stolen secrets. The insider attack is also not considered in our proposed scheme.

*2) Attribute Privacy in the Verification Process:* Our verification scheme relies on the issued certificates on attribute values. During the process of attribute verification, users need to mutually verify the authenticity and equality of attribute certificates. Malicious users fail to learn the detail of both $k_\imath$ and $\sigma_{\imath.\jmath}$, because they cannot obtain the plaintext of these during the verification process. Instead, users use the parameters $u_1, u_2 \in G_1^2, v_1, v_2 \in G_2^2$ to commit above verification keys and certificates. Based on the DDH assumption, no user is able to reveal the plaintext values from the commitments. For the same reason, adversaries will not discover certificates by comparing multiple ciphertexts which contain user-chosen random numbers. For the verifier, he/she can only check the equality of Eq. (2). If the check process succeeds, he/she learns that the corresponding verification keys and certificates are authentic without disclosing $k$ and $\sigma$. Otherwise, the verifier knows nothing about the detail of whether the commitments or proofs are authentic or not.

For the equality verification, we consider two possible attacks that may compromise the attribute privacy. The first type of attack comes from sending random numbers used for equality verification. In our example, Alice sends back to Bob the random number set $\{r_1, r_2, s_1, s_2\}$ for Bob generating equality proofs $\{\tilde{\pi}_1, \tilde{\pi}_2, \tilde{\psi}_1, \tilde{\psi}_2\}$. It is obvious that Bob can use the received random number set to construct $c_i' = (1,1)u_1^{r_1}u_2^{r_2}$ and $d_i' = (1,1)v_1^{s_1}v_2^{s_2}$. Comparing to the received $c_i^A = (1, \sigma_{\imath.\jmath})u_1^{r_1}u_2^{r_2}$ and $d_i^A = (1, k_\imath g_2^{m_{\imath.\jmath}})v_1^{s_1}v_2^{s_2}$, it is difficult for Bob to derive $k_\imath$ and $\sigma_{\imath.\jmath}$ due to the assumption that DDH and DL problems are hard. Thus, the above approach guarantees the attribute privacy when two users send random numbers back and forth. The other possible attack is performed on the outcome of comparison results. The verification result returns $1$ if and only if $e(g_1, g_2)^{(x_\imath + m_{\imath.\jmath}^B)/(x_\imath + m_{\imath.\jmath}^A)} = e(g_1, g_2)$. Otherwise, the result is an arbitrary string in $G_T$ instead of explicitly listing the values of $m_{\imath.\jmath}^A$ and $m_{\imath.\jmath}^B$, which prevents adversaries from knowing the detail regarding the different attribute values. Here, it is possible for users to return incorrect random numbers and/or proofs, which renders the inconsistent comparison results on two sides. However, we exclude this kind of attack if users agree to compare their attributes in advance for the content dissemination.

We also highlight the collusion attack during the whole verification process, in which adversaries collude to compare commitments and proofs from the same user in order to obtain the certificate. If a group of malicious users want to find out the certificate and verification key of a particular user, they have to launch a number of queries to this user. According to the above analysis, adversaries fail to find credential and certificate from multiple commitments and proofs. Moreover, as an extension to our proposed scheme, we allow users to use frequently changed collision-resistance pseudonyms to generate and forward contents. A user can generate different commitments representing the same attribute value by using different pseudonyms, and thus attackers cannot tell the relationship between commitments and the users' pseudo-identities.

*3) Analysis on Content Dissemination Process:* Based on our assumptions on adversarial model, attackers may inject bogus data during the content dissemination process. We can apply the signature scheme in [32] to construct a signature on $(C_1, C_2, C_3)$. Define a set of collision-resistance hash functions, where $H_1 : G_1 \to \{0,1\}^*$, $H_2 : G_2 \to \{0,1\}^*$, $H_3 : G_T \to \{0,1\}^*$ and $H : \{0,1\}^* \to \{0,1\}^n$, where $n$ denotes the length of output hash value. When Alice transmits a packet including $(C_1, C_2, C_3)$, he/she generates a signature $Sig_{sk_A}(H(H_1(C_1)||H_2(C_2)||H_3(C_3)))$ to guarantee the non-repudiation of the ciphertext, where $sk_A$ is a pre-assigned private key of Alice, and users can use the corresponding public key to verify the signature. The man-in-the-middle attack also can be prevented because we have the attribute verification mechanism to prevent the attacker from even obtaining the ciphertext. Based on the mutual authentication results, we can easily defeat this type of attack. The above attacks cannot succeed because the adversary does not have the corresponding private keys.

By applying advanced scheme *UDW*, another possible attack may happen on deriving the content when adversaries have correct $K$ and $e(g_0, g_2)^y$. However, it is infeasible due to our assumption that decisional bilinear Diffie-Hellman Problem (DBDH) is assumed hard. Therefore, adversaries fail cannot determine $K'$ and $\mathcal{M}$ from $C_3'$.

*4) Collusion Attacks on Decrypting Encrypted Contents:* The collusion attack is a very powerful attack and will severely threaten the security and privacy requirements of our scheme if not thwarted during the content dissemination process. For our basic scheme and *NDW*, It mostly happens when one of the adversaries does not have sufficient number of attributes satisfying $\mathcal{P}_A$. He/she intends to collude with users who have enough attributes and combines them in order to decrypt the ciphertext. However, the decryption scheme perfectly prevents this attack as follows: for each user's private key, it includes a set of elements as $\{g_0^{\frac{\kappa_j}{\gamma+m_i}}\}_{m_i \in \mathcal{S}_j, i=1,\ldots,\varsigma}$ based on the user's verified attributes. It has the unique parameter $\kappa_j$ for user $j$, which is selected by TA. Moreover, according to the DLP assumption, user $j$ fails to deduce $\kappa_j$ from $g_0^{\frac{\kappa_j}{\gamma+m_i}}$ even if $m_i$ is given. Hence, given different elements from different users, the colluded users fail to construct $g_0^{\frac{\kappa_j}{\Pi(\gamma+m_{i.j}^j)}}$ with the same $\kappa_j$, and hence colluded users cannot decrypt the ciphertext.

In particular to the advanced scheme *UDW*, we modify the original scheme in [31] to prevent possible collusion attacks among adversaries. As similar as the collusion attack on the basic scheme, the attack objective of adversaries is trying to obtain the content that they cannot decrypt by using their own private keys $\texttt{sk}$. In our advanced scheme *UDW*, the private key elements $\hat{D} := g_2^{y-\lambda_U l}$ and $D_k := g_0^{\frac{\lambda_U l_k}{h_k}}$ have the unique parameter $\lambda_U$ for each particular user $U$. Even though adversaries colluding with each other, they fail to reconstruct $e(g_0, g_2)^{\lambda_U lc}$ due to their different values on $\lambda$. Furthermore, they cannot derive the encryption key $K'$ by dividing the reconstructed element $Y^c$. Therefore, the collusion attacks on encrypted contents are prevented in our scheme.

### B. Performance of PSaD

Regardless of efficiency of the privacy preservation technique used in PSaD, we first analyze the performance of the attribute-based content dissemination process. Based on the trace file collected during INFOCOM 2006, we use the following metrics in the simulation, which try to specifically evaluate our proposed scheme. In addition, we compare our scheme with several existing social-assisted approaches in order to show the advantages of our scheme.

- **Delivery ratio**, the ratio of the number of delivered destinations to the total number of destinations.
- **Delivery delay**, the average delay for all the delivered destinations to receive the content.
- **Average cost**, the average number of relays used for one delivered destination to receive a content.

*1) Analysis on Identical Attributes:* We first analyze the distribution of number of identical attributes between two arbitrary users in the dataset, which helps us determine the forwarding strategy of PSaD scheme. According the system setting in the dataset, the total number of possible pair-wise contacts is $A_{78}^2 = 78!/(78-2)! = 6006$ given the number of participants is 78. As similar as the simulation setting in Sec.II, we consider the maximum number of identical attributes is 10. In Fig. 3, there are more than 1800 pairwise users do not share any similar attributes. Besides, among all users with identical attributes, the number of users sharing three attributes has more than 800 pairs. Regarding the analysis of distribution of identical attributes among users, it is easier to disseminate content with threshold $t = 3$, but it may not directly share with the desired group of users who are expected to have the attributes threshold $t > 3$. Moreover, we need to consider the parameter of $\theta = |\mathcal{S}_{U,\sigma}|$, which is the total number of attributes required by the content generator for attribute verification.

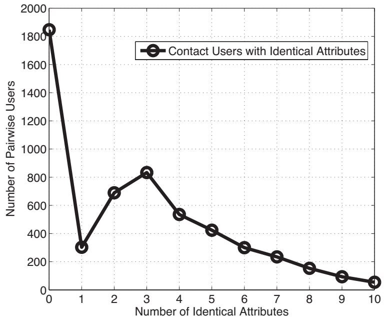


Fig. 3. Distribution of Users' Identical Attributes (INFOCOM 2006)

In what follows, we intend to discuss the tradeoff between the selection of $(\theta, t)$-pair and the performance of PSaD scheme, where $1 \leqslant \theta \leqslant t \leqslant 10$ for practical concern. In addition to the selection of the number of identical attributes, we also investigate the impact of contact duration between contact users. The contact duration $T$ between users reflects the length of their interaction time. Due to the ways of data collection, users' contacts may be mistakenly recorded, such as the Bluetooth connection establishment time and users' unintentional interactions. As one of the parameter in the experiment settings, we gradually increase the threshold of contact duration $T$ to exclude some of the contacts in the datasets. As the threshold of $T$ increases, the contacts can be used as real social interactions, and further can be applied for content exchange and dissemination.

*2) Performance Evaluation:* We give our simulation results along with the discussion in the following subsection.

- **Analysis on Delivery Ratio**

We first discuss the delivery ratio of the proposed scheme. The delivery ratio is calculated in the following way to keep consistent with other approaches [7], [9], [11],

$$\text{Delivery Ratio} = \frac{\sum_i \#\text{Receivers for } i \text{ hop}}{\#\text{Users in the system}} \quad (6)$$

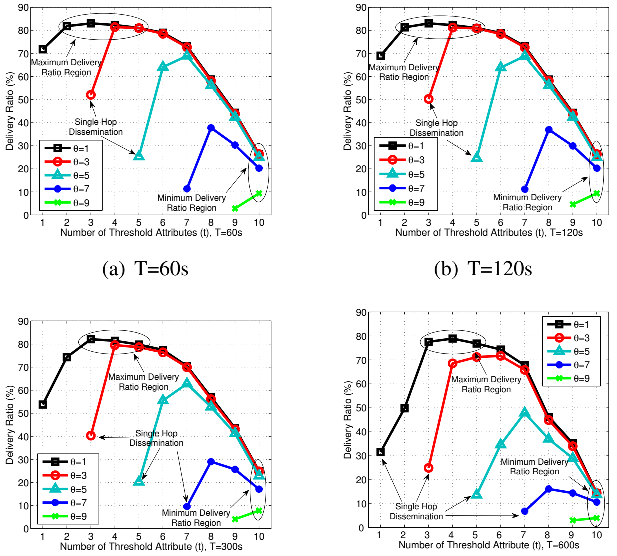
where the purpose of content generator is to maximally disseminate the content based on the selection of $(\theta, t)$-pair. Note that content owners will not be able to the number of receivers before they disseminate the content, and also it is infeasible for them to determine the potential receivers given different selections on $(\theta, t)$-pair. Although with more restriction on $(\theta, t)$-pair, it may reduce the number of potential receivers, we have to keep pace with content owners sharing purposes, and to help them maximally disseminate the content.

As shown in Fig. 4, we show the delivery ratio of PSaD scheme for different $(\theta, t)$-pair with consideration of the requirement on the contact duration $T$. For the maximum delivery ratio, we refer the region with more than $75\%$ as the maximum delivery ratio region. Meanwhile, the minimum delivery ratio region is referred as the selection of $(\theta, t)$-pair in resulting the delivery ratio smaller than $30\%$. Apparently, the number of the selection of $(\theta, t)$-pairs in maximum delivery region is reduced when the time constraint $T$ grows, which indicates that the increasing contact durations will filter out the unintentional user interactions. Moreover, if we set $T = 600s$ as shown in Fig. 4(d), most of the contacts can be seen as real communication, and they can be further used as content dissemination for practical information exchange. We also highlight $(\theta, t)$-pairs selections for $\theta = t$, which can be characterized as the single-hop dissemination. In Fig. 4(a), the highest delivery ratio of single-hop dissemination is around $70\%$, and it significantly reduced to $32\%$ when $T = 600s$. To increase the delivery ratio given a threshold value $T$, the content generator may increase the value of $t$, which indicates that more contact users may satisfy the requirement of the content generator, and may further relay the content. Rather than the situation where $\theta = t$ (indicates the dissemination neglects the possible values of $t$ which is greater than $\theta$), the selection of $t > \theta$ forms the multi-hop dissemination process and it outperforms single-hop dissemination.

- **Analysis on Delivery Delay**

As similar as the performance with respect to the delivery ratio, the average delay of PSaD scheme grows when the time constraint increases as shown in Fig. 5. For $T < 300s$, the average delay increases linearly when the content generator sets $\theta < 5$. Meanwhile, it keeps flat when the time constrains are greater than 300s. The maximum delay points increases from 48,984s second to 53,671 seconds in Fig. 5(a) and Fig. 5(d). An interesting observation is that the trends of average delay varies for $\theta \leqslant 4$ and $\theta \geqslant 5$. In Fig. 5(a) and Fig. 5(b), the average delay increases for $\theta = 1, 3$ when the threshold number of attributes $t$ increases. Comparatively, the average delay decreases after $t$ reaches a maximum delay point for $\theta \geqslant 5$. As what we discussed in Sec. II, we refer this region as homophily phenomenon region, where the increase of contact rate facilitates the communication between users with more identical attributes. Therefore, the average delay of content dissemination dramatically reduces based on their frequently contacts, while the selection of smaller $\theta$ keeps the average delay growing.

- **Analysis on Average Cost**



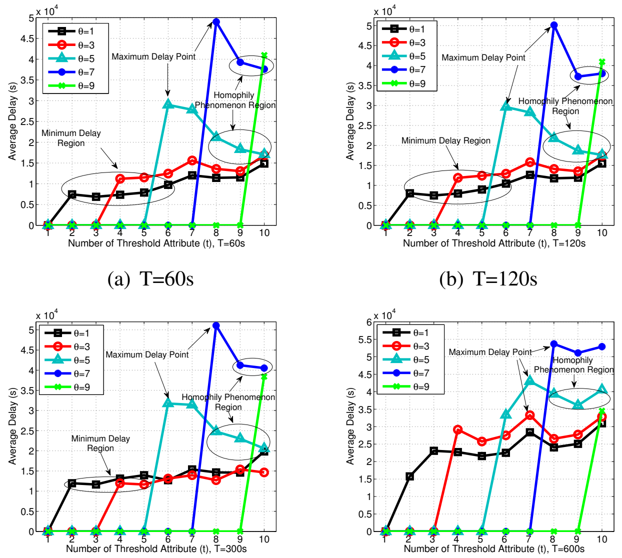
(a) T=60s

(b) T=120s

(c) T=300s

(d) T=600s

Fig. 4. Delivery Ratio of PSaD Scheme



(a) T=60s

(b) T=120s

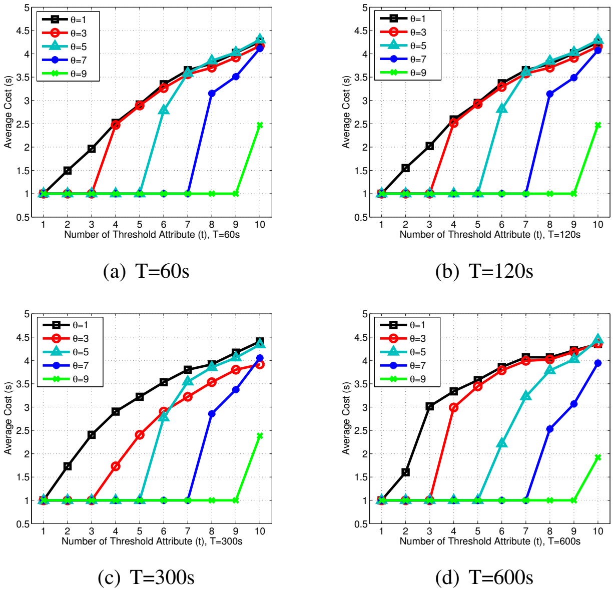
(c) T=300s

(d) T=600s

Fig. 5. Average Delay of PSaD Scheme

We also give analysis on the average cost in terms of average number of relays during the dissemination process. With the number of threshold attributes growth, the average cost increases linearly in all the selections of $(\theta, t)$-pairs as shown in Fig. 6. As the same as our analysis on delivery delay, the average cost decreases for $\theta \geqslant 5$ when the time constraint increases. The maximum cost for the selection $\theta \leqslant 7$ is between 3.9 and 4.3 relays. Moreover, to achieve the best delivery ratio and lower delivery delay of PSaD scheme, the average cost is between 2.1 and 3.2 relays for $T \leqslant 600s$.

*3) Discussion on Homophily Phenomenon:* Based on the previous analysis, we show the existence of homophily phenomenon, which verifies our intuition on designing the PSaD scheme. However, it is obvious that the selection of $(\theta, t)$-pair will impact the increase and decrease of delivery delay and



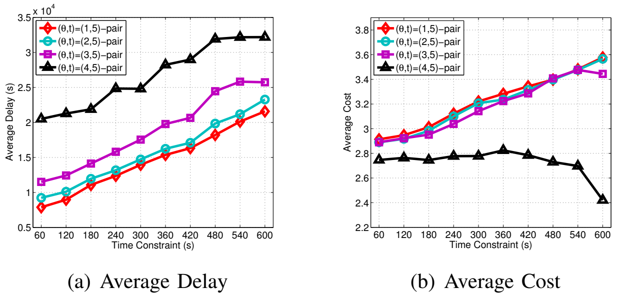
(a) T=60s

(b) T=120s

(c) T=300s

(d) T=600s

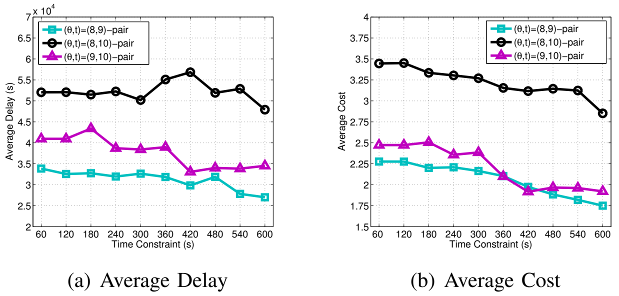
Fig. 6. Average Cost of PSaD Scheme



(a) Average Delay

(b) Average Cost

Fig. 7. Average Delay and Cost of PSaD Scheme for $\theta \leqslant 4, t \leqslant 5$



(a) Average Delay

(b) Average Cost

Fig. 8. Average Delay and Cost of PSaD Scheme for $\theta \geqslant 8, t \geqslant 9$

average cost. In what follows, we investigate the performance of PSaD scheme with respect to the time constraint and the selection of $(\theta, t)$-pair, which help content generators set their verification and encryption policy.

We pick two sets of representative simulation results, in which we choose $t \leqslant 5$ and $t \geqslant 9$ as the decryption policy as shown in Fig. 7 and Fig. 8. For the selection of $t \leqslant 5$, with the increment of $\theta$, the average delay increases in Fig. 7(a), which indicates that the content generator requires more certificates for attribute verification. Consequently, the average delay will grow when $T$ is increasing. The average cost for $t \leqslant 5$ is shown in Fig. 7(b). For $\theta \leqslant 3$, it grows linearly with the increment of $T$. However, the average cost decreases when we set $\theta = 4$, in which we consider the homophily phenomenon appears among users with identical attributes greater than 4. To further verify this phenomenon, we analyze the delivery delay and average cost for $t \geqslant 8$ which implies that the contact users have nearly all identical attributes. Compared with the results in Fig.7, we find the performance of PSaD scheme has an opposite result in Fig. 8. As shown in Fig. 8(a), the average delay decreases when $T$ grows. This results indicates that the content is disseminated within a group of people with a lot of identical attributes, and they may frequently contact with each other, which lower the average delay for approximately 30% compared with $\theta = 4, t = 5$ at $T = 600s$. Moreover, the average cost also decreases for $t \geqslant 8$ in Fig. 8(b), where it only has 1.76 relays when the content generator sets $\theta = 9, t = 10$ at $T = 600s$.

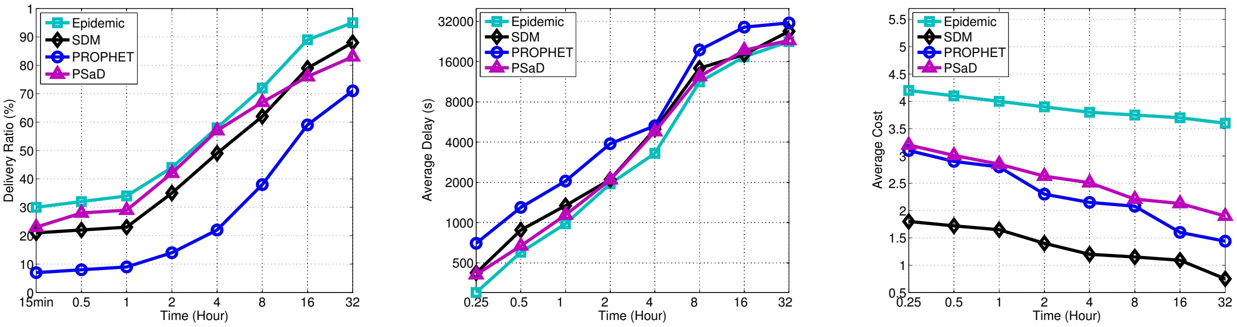
Therefore, we verify the existence of the well-known homophily phenomenon by real-world trace file. In addition, in terms of content dissemination, content generators may choose different selections of $(\theta, t)$-pair, to achieve desired performance. If they intend to have better delivery ratio and smaller delivery delay, they can set a lower $(\theta, t)$-pair value, but their receivers may not be as desired as they want; or, they can disseminate to a particular group of users with almost all identical attributes in order to achieve low cost on relays, but the price would be the low delivery ratio.

*4) Protocol Comparison:* In what follows, we compare our PSaD scheme with several existing social-assisted routing protocol in DTNs as shown in Fig. 9. We mainly consider the routing protocol Epidemic [7], SDM [11], and PROPHET [9]. In Fig. 9(a), the best delivery ratio of PSaD scheme is about 4% lower than the Epidemic routing scheme, but it outperforms SDM and PROPHET for 7% and 21%, respectively. The delivery delay of PSaD and SDM are more or less similar for the time length smaller than 16 hours as shown in Fig. 9(b). For the overall performance (3 days experiment time), the PSaD scheme is 792 seconds longer than Epidemic, but it is smaller than SDM and PROPHET for an average of 3,911 seconds and 8,235 seconds. In terms of average cost, the PSaD scheme is better than the Epidemic, but it is larger than PROPHET and SDM scheme, as shown in Fig. 9(c).

### C. Efficiency Analysis

*1) Simulation-based Analysis:* First, we use Pairing-based Cryptography (0.5.12) Library to implement our simulation on computational and storage cost. We continue to use the real-world trace file collected in INFOCOM 2006 to verify our proposed scheme, including number of users and their corresponding attributes. We take Tate pairing as our basic pairing operation. The elliptic curve we use for the our scheme is type D159. A curve of such type has the form of $y^2 = x^3 + ax + b$. The base field of the curve is 159 bits, and it has the same security level as 1024-bit RSA. For the experiments, we use a laptop with an Intel processor 2.8GHz and 4GB RAM under the platform Ubuntu 11.10. All the timing reported below are averaged over 100 randomized runs.

**Computational Cost:** We consider the computational cost during the attribute verification process and the content dissemination process. During the attribute verification process, Alice generates 2 commitments and 4 proofs for the secrets in Eq. (1). The commitment generating process takes 2 group elements from each of $G_1$ and $G_2$, while the proofs have 4 group elements from



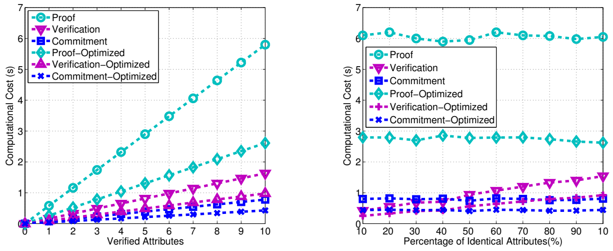
(a) Delivery Ratio

(b) Delivery Delay

(c) Average Cost

Fig. 9. Comparison with Other Social-assisted Protocols, $(\theta, t) = (2, 3)$

both $G_1$ and $G_2$. For the verification process, each user computes 20 pairing operations to verify the validity of the corresponding attributes. For the equality verification, the verification process takes another 20 pairing operations in total. As we can see from Fig.10(a), the computational costs of generating commitments, proofs, and verification grow nearly linearly with the increment of the number of compared attributes. To verify total 10 attributes takes less than 6 seconds for each side. On the other hand, we also need to consider the impact of identical attributes during the verification process. As shown in Fig.10(b), if the total number of compared attributes is 10, the computation time of commitment and proof generation process keep stable when the identical attribute ratio grows. On the contrary, the verification time increases linearly with the growth of identical attribute ratio. To reduce the computational cost, we revise parts of our scheme in the simulation, in oder to allow users pre-compute some costly operations and store them for future use. As shown in Fig.10, the computational costs reduce for 52.3% on average during the verification process.



(a) Computational Cost

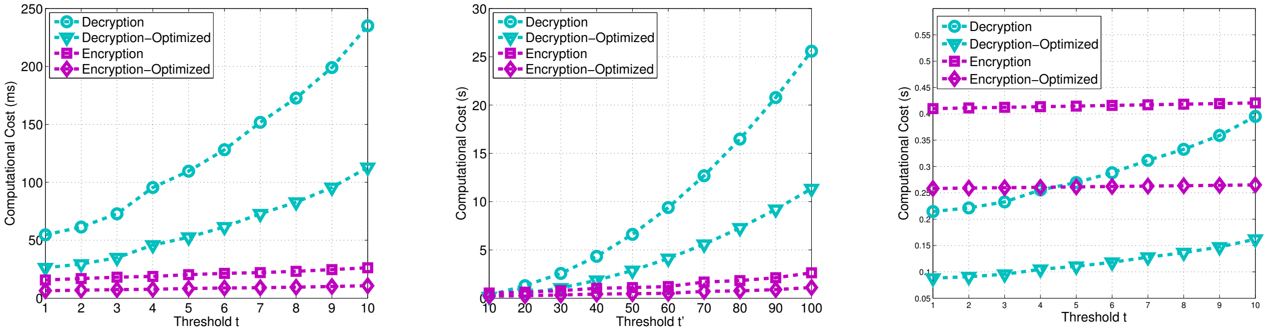
(b) Identical Attribute Ratio

Fig. 10. Computational Cost in Attribute Verification Process

In the content dissemination process, we only consider the computational cost in terms of time in our encryption and decryption scheme. As we can see from the previous section, the encryption scheme only takes $\alpha + t - 1$ exponentiations on the group elements over $G_2$, and 1 exponentiation operation over each of $G_2$ and $G_T$. For the decryption scheme, it requires 3 pairing operations and $O(t^2 + \alpha)$ exponentiations. In the following simulation settings, we fix the number $\alpha = 10$ (one attribute value per attribute) and mostly focus on the computation time in our encryption and decryption scheme. Based on the simulation results, the average computation time for the exponentiation operation over $G_1$ is 1.14ms, and 2.53ms over group $G_2$ and $G_T$. The average time in computing a pairing operation costs 11.9ms under our simulation settings. As we can see from Fig. 11(a), the encryption time increases with the growth of threshold $t$, and it takes about 27ms for generating the ciphertext when $t = 10$. The decryption time increases dramatically compared with encryption, which takes approximately 241ms for $t = 10$. By using our optimized approach, the decryption costs reduces to 112.4ms for $t = 10$. For our advanced scheme using *NDW*, the number of $\alpha' = 10 \cdot \Omega$ increases depending on the granularity of attribute weights. In our simulation, we define $\Omega = 10$. In Fig. 11(b), it is shown that the decryption time increases rapidly compared to the time consumed in the encryption scheme, while the optimized decryption scheme reduces for 53.7%. In *UDW*,to generate the ciphertext of $C_4$ costs 1 exponentiation operation over group $G_1$, $4\alpha$ exponentiation operation on $G_T$ and 1 multiplication operation in $G_T$. For the decryption scheme of *UDW*, it requires $\alpha + 1$ pairing operations and $\alpha$ multiplication in $G_1$ for decrypting $\tilde{C_4}$, where the average cost of bilinear pairing is 10.56ms. The decryption time costs less than the encryption time for $1 \leqslant t \leqslant 10$ as shown in Fig. 11(c). In particular, the computational cost of encryption using optimized scheme decreases more than 50% compared with original scheme.

**Storage Cost:** In our storage analysis, we choose $|p| = 160$, and the element in $G_1$ is 170 bits, while elements of $G_2$ and $G_T$ will be 510 bits as introduced in [33], [34]. We first consider the storage cost in the attribute verification process. The *crs* costs $|p| + 5|G_1| + 5|G_2|$ bits, where $|G_1|$ denotes the size of the element in $G_1$. To store the verification keys and certificates of users' attributes, they take $\varsigma|G_1| + \varsigma|G_2|$ bits in total for one user. For the storage cost during the dissemination process, both the basic scheme and *NDW* require $|G_T| + |G_1| + (2\alpha - 1)|G_2|$ bits on public key and $|G_1| + |G_2| + |G_T|$ bits as the constant-size ciphertext of contents. The private keys for the basic scheme have the length of $\varsigma|G_1| + (\alpha - 1)|G_2|$ bits, while *NDW* scheme requires $\varsigma\Omega|G_1| + (\alpha - 1)|G_2|$ bits for each user to store private keys. For the advanced scheme $UDW$, each user stores additional $(\varsigma + 1)|G_1| + |G_2|$ bits for the private key, while the public key costs $\varsigma|G_2| + |G_T|$ bits.

*2) Experiment-based Analysis:* Besides the simulation on laptops, we also conduct practical experiments on mobile devices. We use Java Pairing Based Cryptography (jPBC-1.2.0) Library to implement our scheme. We continue to use curve D159 as the elliptic curve. For the experiment, we use the smartphone Nexus S with a Samsung Exynos 3110 processor. The smartphone has 1 GHz ARM Cortex A8 core, and 512 MB RAM. The following experiment is built on the platform Android 2.3.2. Since the



(a) Basic Scheme (b) *NDW* (c) *UDW*

Fig. 11. Computational Cost in Content Dissemination

storage cost in the experiment is the same as it in the simulation, we omit this analysis. For the computational cost analysis, we show the experimental results on attribute verification process and the basic scheme of content dissemination process. As we can see from Fig.12(a), the total cost of verifying 10 attributes takes up to 60s on the smartphone platform, which is almost 10 times than the results in our simulation. Due to the insufficient computation capability of smartphones, it takes at most 4s to decrypt the content with the length no more than 510 bits. By applying our optimized scheme, generating proof, commitment, and verification process reduce for 47.4%, 32%, and 41.2%, respectively. We also show our simulation results on the content dissemination. As shown in Fig.12(b), the encryption time keeps flat when $t$ increases, while the decryption time grows dramatically. For our advanced scheme *NDW*, the encryption time with *NDW* reaches 79.2s in Fig.12(c), which is approximately 3 times more than the time in Fig.11(b), while the optimized approach help the decryption time reduce to 29.8s for $t = 10$. If we apply *UDW* on our basic scheme, the maximum decryption time is 5.13s compared with the encryption time, which is 5.84s. The optimized scheme also achieves better performances as shown in Fig.12(d).

## VI. Conclusion

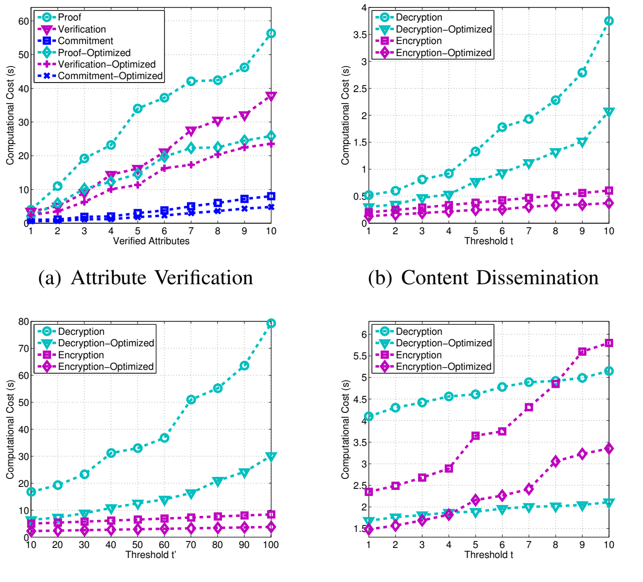
In this paper, we propose the PSaD, a privacy-preserving social-assisted content dissemination scheme in DTNs. The scheme mainly consists of attribute verification and privacy-preserving content dissemination process. In our scheme, mobile users first identify their social relationships in terms of identical attributes, then disseminate their encrypted contents by each contact, in order to let users who have the corresponding attributes decrypt the content. Based on the protocol evaluation, we show both the security and efficiency of the proposed scheme.

## Acknowledgement

This work was partially supported by the U.S. National Science Foundation under Grants CNS-0916391. The work of Chi was partially supported by the National Natural Science Foundation of China under Grant 61202140.

## References

[1] L. Guo, C. Zhang, H. Yue, and Y. Fang, “A privacy-preserving social-assisted mobile content dissemination scheme in dtns,” in *The 32nd IEEE International Conference on Computer Communications*, ser. INFOCOM 2013. Turin, Italy: IEEE, 2013, pp. 2349–2357.



(a) Attribute Verification (b) Content Dissemination

(c) Content Dissemination w/ NDW (d) Content Dissemination w/ UDW

Fig. 12. Experiment-based Results on Computational Cost

[2] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, and C. Diot, “Pocket switched networks and human mobility in conference environments,” in *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, ser. WDTN ’05. New York, NY, USA: ACM, 2005, pp. 244–251.

[3] M. McPherson, L. Smith-Lovin, and J. M. Cook, “Birds of a feather: Homophily in social networks,” *Annual Review of Sociology*, vol. 27, no. 1, pp. 415–444, 2001.

[4] A. C.-C. Yao, “How to generate and exchange secrets,” *Proceedings of the 27th Annual Symposium on Foundations of Computer Science –SFCS 1986*, pp. 162–167, 1986.

[5] J. Bethencourt, A. Sahai, and B. Waters, “Ciphertext-policy attribute-based encryption,” *Proceedings of the 2007 IEEE Symposium on Security and Privacy*, pp. 321–334, 2007.

[6] Y. Zhu, B. Xu, X. Shi, and Y. Wang, “A survey of social-based routing in delay tolerant networks: Positive and negative social effects,” *Communications Surveys Tutorials, IEEE*, vol. PP, no. 99, pp. 1 –15, 2012.

[7] A. Vahdat and D. Becker, “Epidemic routing for partially-connected ad hoc networks,” Duke Univeristy, Tech. Rep., 2000.

[8] P. Hui, J. Crowcroft, and E. Yoneki, “Bubble rap: social-based forwarding in delay tolerant networks,” in *Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing*, ser. MobiHoc ’08. New York, NY, USA: ACM, 2008, pp. 241–250.

[9] A. Lindgren, A. Doria, and O. Schelén, “Probabilistic routing in intermittently connected networks,” *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 7, no. 3, pp. 19–20, Jul. 2003.

[10] E. M. Daly and M. Haahr, “Social network analysis for routing in disconnected delay-tolerant manets,” in *Proceedings of the 8th ACM international*

*symposium on Mobile ad hoc networking and computing*, ser. MobiHoc '07. New York, NY, USA: ACM, 2007, pp. 32–40.

[11] W. Gao, Q. Li, B. Zhao, and G. Cao, "Multicasting in delay tolerant networks: a social network perspective," in *Proceedings of the tenth ACM international symposium on Mobile ad hoc networking and computing*, ser. MobiHoc '09. New York, NY, USA: ACM, 2009, pp. 299–308.

[12] W. Gao and G. Cao, "User-centric data dissemination in disruption tolerant networks," in *INFOCOM, 2011 Proceedings IEEE*, april 2011, pp. 3119 –3127.

[13] J. Wu and Y. Wang, "Social feature-based multi-path routing in delay tolerant networks," in *INFOCOM, 2012 Proceedings IEEE*, march 2012, pp. 1368 –1376.

[14] R. Lu, X. Lin, T. Luan, X. Liang, X. Li, L. Chen, and X. Shen, "Prefilter: An efficient privacy-preserving relay filtering scheme for delay tolerant networks," in *INFOCOM, 2012 Proceedings IEEE*, march 2012, pp. 1395 –1403.

[15] R. Lu, X. Lin, and X. Shen, "Spring: A social-based privacy-preserving packet forwarding protocol for vehicular delay tolerant networks," in *INFOCOM, 2010 Proceedings IEEE*, 2010, pp. 1–9.

[16] O. Hasan, J. Miao, S. Mokhtar, and L. Brunie, "A privacy preserving prediction-based routing protocol for mobile delay tolerant networks," in *Advanced Information Networking and Applications (AINA), 2013 IEEE 27th International Conference on*, 2013, pp. 546–553.

[17] Y. Zhang and J. Zhao, "Social network analysis on data diffusion in delay tolerant networks," in *Proceedings of the tenth ACM international symposium on Mobile ad hoc networking and computing*, ser. MobiHoc '09. New York, NY, USA: ACM, 2009, pp. 345–346.

[18] A. Mei, G. Morabito, P. Santi, and J. Stefa, "Social-aware stateless forwarding in pocket switched networks," in *INFOCOM, 2011 Proceedings IEEE*, april 2011, pp. 251 –255.

[19] E. Bulut and B. Szymanski, "Friendship based routing in delay tolerant mobile social networks," in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, dec. 2010, pp. 1 –5.

[20] G. Costantino, F. Martinelli, and P. Santi, "Investigating the privacy vs. forwarding accuracy tradeoff in opportunistic interest-casting," *Mobile Computing, IEEE Transactions on*, vol. PrePrint, no. 99, pp. 1–1, 2013.

[21] G. Costantino, F. Martinelli, P. Santi, and D. Amoruso, "An implementation of secure two-party computation for smartphones with application to privacy-preserving interest-cast," in *Proceedings of the 18th annual international conference on Mobile computing and networking*, ser. Mobicom '12. New York, NY, USA: ACM, 2012, pp. 447–450.

[22] W.-j. Hsu, D. Dutta, and A. Helmy, "Profile-cast: behavior-aware mobile networking," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 12, no. 1, pp. 52–54, Jan. 2008.

[23] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau, "CRAWDAD trace cambridge/haggle/imote/infocom2006 (v. 2009-05-29)," May 2009.

[24] J. Herranz, F. Laguillaumie, and C. Ràfols, "Constant size ciphertexts in threshold attribute-based encryption," in *Proceedings of the 13th international conference on Practice and Theory in Public Key Cryptography*, ser. PKC'10. Berlin, Heidelberg: Springer-Verlag, 2010, pp. 19–34.

[25] D. Boneh and M. Franklin, "Identity-based encryption from the weil pairing," *Advances in Cryptology —CRYPTO 2001*, pp. 213–229, 2001.

[26] E.-J. Goh, "Encryption Schemes from Bilinear Maps," Ph.D. dissertation, Department of Computer Science, Stanford University, Sep 2007.

[27] C. Delerablée and D. Pointcheval, "Dynamic threshold public-key encryption," in *Proceedings of the 28th Annual conference on Cryptology: Advances in Cryptology*, ser. CRYPTO 2008. Berlin, Heidelberg: Springer-Verlag, 2008, pp. 317–334.

[28] D. Boneh, X. Boyen, and H. Shacham, "Short group signatures," *CRYPTO'04, LNCS*, pp. 41–55, 2004.

[29] J. Groth and A. Sahai, "Efficient non-interactive proof systems for bilinear groups," in *Proceedings of the theory and applications of cryptographic techniques 27th annual international conference on Advances in cryptology*, ser. EUROCRYPT'08, 2008, pp. 415–432.

[30] L. Guo, C. Zhang, J. Sun, and Y. Fang, "PAAS: Privacy-preserving attribute-based authentication system for ehealth networks," in *The 32nd IEEE International Conference on Distributed Computing Systems*, ser. ICDCS 2012. Macau, China: IEEE, 2012.

[31] L. Cheung and C. Newport, "Provably secure ciphertext policy abe," in *Proceedings of the 14th ACM conference on Computer and communications security*, ser. CCS '07. New York, NY, USA: ACM, 2007, pp. 456–465.

[32] P. S. L. M. Barreto, B. Libert, N. McCullagh, and J.-J. Quisquater, "Efficient and provably-secure identity-based signatures and signcryption from bilinear maps," in *ASIACRYPT*, 2005, pp. 515–532.

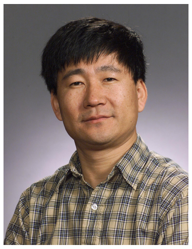
[33] B. Lynn. [Online]. Available: http://crypto.stanford.edu/pbc/

[34] M. Pirretti, P. Traynor, P. McDaniel, and B. Waters, "Secure attribute-based systems," *In ACM Conference on Computer and Communications Security. (CCS'06)*, pp. 99–112, 2006.

**Linke Guo (S'10)** received his B.E. degree in electronic information science and technology from Beijing University of Posts and Telecommunications in 2008 and M.S. degree in electrical and computer engineering from University of Florida in 2011. He is currently working towards his Ph.D. degree in University of Florida. His research interests include network security and privacy, social networks, and applied cryptography.

**Chi Zhang** received the B.E. and M.E. degrees in Electrical and Information Engineering from Huazhong University of Science and Technology, China, in 1999 and 2002, respectively, and the Ph.D. degree in Electrical and Computer Engineering from the University of Florida in 2011. He joined the University of Science and Technology of China in September 2011 as an Associate Professor of the School of Information Science and Technology. His research interests are in the areas of network protocol design and performance analysis, and network security particularly for wireless networks and social networks. He has published over 60 papers in journals such as IEEE/ACM Transactions on Networking, IEEE Journal on Selected Areas in Communications, and IEEE Transactions on Mobile Computing and in networking conferences such as IEEE INFOCOM, ICNP, and ICDCS. He has served as the Technical Program Committee (TPC) members for several conferences including IEEE INFOCOM, ICC, GLOBECOM, WCNC and PIMRC. He is the recipient of the 7th IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award.

**Hao Yue (S'11)** received his BSc degree in Telecommunication Engineering from Xidian University, China, in 2005. He has been working towards the Ph.D. degree in the Department of Electrical and Computer Engineering at University of Florida, Gainesville since August 2009. His research interests include wireless networks and mobile computing, cyber physical systems and security and privacy in distributed systems.



**Yuguang Fang (F'08)** received a Ph.D. degree in Systems Engineering from Case Western Reserve University in January 1994 and a Ph.D degree in Electrical Engineering from Boston University in May 1997. He was an assistant professor in the Department of Electrical and Computer Engineering at New Jersey Institute of Technology from July 1998 to May 2000. He then joined the Department of Electrical and Computer Engineering at University of Florida in May 2000 as an assistant professor, got an early promotion to an associate professor with tenure in August 2003 and a professor in August 2005. He has published over 350 papers in refereed professional journals and conferences. He received the National Science Foundation Faculty Early Career Award in 2001 and the Office of Naval Research Young Investigator Award in 2002. He won the Best Paper Award at IEEE ICNP'2006. He has served on many editorial boards of technical journals including IEEE Transactions on Communications, IEEE Transactions on Wireless Communications, IEEE Transactions on Mobile Computing, Wireless Networks, and IEEE Wireless Communications (including the Editor-in-Chief). He is currently the Editor-in-Chief of IEEE Transactions of Vehicular Technology. He is also serving as the Technical Program Co-Chair for IEEE INFOCOM'2014. He is a fellow of the IEEE.