Morality-driven Data Forwarding with Privacy Preservation in Mobile Social Networks

Xiaohui Liang[†], *Student Member, IEEE*, Xu Li[*], Tom H. Luan[†], Rongxing Lu[†], *Member, IEEE*, Xiaodong Lin[‡], *Member, IEEE*, and Xuemin (Sherman) Shen[†], *Fellow, IEEE*
[†]Department of Electrical and Computer Engineering, University of Waterloo, Canada
[*]INRIA Lille - Nord Europe, France
[‡]Faculty of Business and Information Technology, University of Ontario Institute of Technology, Canada
Email: {x27liang, hluan, rxlu, xshen}@bbcr.uwaterloo.ca; xu.li@inria.fr; xiaodong.lin@uoit.ca

***Abstract*—Effective data forwarding is critical for most mobile social network applications, such as content distribution and information searching. However, it could be severely interrupted or even disabled when privacy preservation of users is applied, because that users become unrecognizable to each other and the social ties and interactions are no longer traceable to facilitate cooperative data forwarding. Therefore, how to enable efficient user cooperation in mobile social networks (MSNs) without intruding user privacy is a challenging issue. In this paper, we address this issue by introducing the *social morality* – a fundamental social feature of human society – to MSNs, and accordingly design a three-step protocol suite to achieve both privacy preservation and cooperative data forwarding. Firstly, the developed protocol adopts a novel privacy-preserving route-based authentication scheme which notifies a user's *anonymized* mobility information to the public. Secondly, it measures the proximity of the user's mobility information to a specific packet's destination and evaluates the user's forwarding capacity for the packet. Thirdly, using a game theoretical approach, it determines the optimal data forwarding strategy according to users' morality level and payoff. Using analysis and examples, we show that the developed protocol suite can effectively protect user personal information such as identity and visited locations. Lastly, we conduct extensive trace-based simulations and show that the proposed protocol suite is effective to explore the user cooperation efficiently and attain near-optimal performance in data forwarding.**

***Index Terms*—Mobile social networks, data forwarding, privacy preservation, social theory, game theory**

I. Introduction

MOBILE social networks (MSNs) are emerging social networking platforms over which participants are able to communicate with each other using Bluetooth or P2P WiFi enabled wireless handheld devices [1]–[5]. With MSNs, proximity services can be supported to strengthen the social interactions of geographically-close users. For example, by probing others in proximity through bluetooth communication, PeopleNet [6] enables the human-like information search among mobile phones. [7] facilitates the carpool and ride sharing in proximity based on message relay among mobile social users. A micro-blog system [8] is built on MSNs to enable users to locally record multimedia blogs on-the-fly, enriched with inputs from other physical sensors. In general, MSNs rely on the opportunistic contacts among users for cooperative data forwarding, and they are alternatively known as Pocket Switched Networks (PSNs). Unlike conventional wireless relay networks assuming end-device to be insensate, MSNs consider mobile devices to be pertained with human users and have specific social features. As such, MSN applications place great emphasis on user social behavior such as selfishness and social proximity, and explore the social features of devices for efficient data forwarding protocol design.

In MSNs, allowing the exchange of personal information to some extent is inevitable to enable social-related cooperative communications. This, however, should be strictly controlled at the prerequisite of effective user privacy preservation. Of all the user privacy requirements, the *identity privacy* and *location privacy* [9] are of paramount importance. Specifically, Identity privacy dictates that the identities of users as source, relay or destination in cooperative data forwarding are not disclosed. Location privacy implies that a user's future mobility route cannot be predicted or inferred from its current and past route information. In other words, the exposed location information of a user should not be linkable to its future location. The state-of-the-art privacy-preserving techniques in MSNs include a *multiple-pseudonym* technique [10], [11] and a *hotspot* technique [12]. The former assigns each user with a set of asymmetric key pairs, and uses the alternatively changing public keys as the user's pseudonyms for data communication. The user identity can be protected as only literally-meaningless pseudonyms are exposed to the public. By frequently changing its pseudonym for authentication over time, the user achieves location privacy due to the unlinkability of old and new pseudonyms. As a result, the *multiple-pseudonym* technique can guarantee both identity privacy and location privacy by making the social interactions of users anonymous. The *hotspot* technique guarantees the receiver location privacy while achieving efficient data forwarding. Specifically, it defines some hotspots that are common and with high population, and then makes use of these hotspot information to assist the data forwarding without revealing sensitive locations of receivers.

In this paper, we address a fundamental tradeoff between the privacy preservation and the data forwarding efficiency in MSNs. Specifically, with the *multiple pseudonym* technique applied for privacy preservation, an unpleasant ac-

companying side-effect is that users are unable to identify their social friends because of the anonymity of users. This directly impedes the cooperative data forwarding as social ties among users are interrupted. Since users are anonymous, the malicious behaviors (e.g., selfish and free-riding) can no longer be tracked and punished on time using traditional mechanisms. This may discourage user cooperations and deteriorate the data forwarding efficiency. Therefore, privacy preservation protects and hides the identities of users to the public, which, however, hinders the social-based cooperative data forwarding. Our goal is *to resolve the two conflicting goals in one framework by proposing a privacy preserving social-based cooperative data forwarding protocol*. We exploit the social morality for cooperative data forwarding design. Specifically, the morality of human beings is a common social phenomenon in real-world which provides the rules for people to act upon and grounds the moral imperatives. It is the fundament of a cooperative and mutually beneficial social life in the real-world society. Our main contributions are three-fold:

First, we identify the conflicting nature between privacy preservation and cooperative data forwarding in MSNs. On addressing this issue, we are the first to leverage social morality to incentivize the user cooperation and accordingly promote the communication efficiency.

Second, we propose a three-step protocol suite to attain the privacy-preserving data forwarding. Firstly, we introduce a privacy-preserving route-based authentication scheme. It enables users to expose the mobility information to each other for cooperation, yet with location privacy preserving. Based on the mobility of users, we evaluate the forwarding capability of individual users on a given packet. Lastly, a game-theoretic approach taking account of both the morality and forwarding capability is designed to adaptively determine the optimal data forwarding strategy for individual users[1]. The final optimal solution of the protocol suite stands for a balance of moral peace and benefit maximization for all users.

Third, we evaluate the performance of the proposed protocols through extensive trace-based simulations. Our simulations validate the efficiency of the proposed data forwarding protocol with location privacy preservation.

The remainder of this paper is organized as follows: we introduce some related works in Section II. We describe the system model and provide an overview of the three-step protocol suite in Section III. We present a privacy-preserving route-based authentication scheme and proximity measurement as the first two steps in Section IV, and use game-theoretic analysis to derive the optimal forwarding strategies for users as the third step in Section V. We conduct trace-based simulations to evaluate the performance of the proposed protocol in Section VI. Finally, we draw our conclusion in Section VII, respectively.

[1]Assume users are rational and selfish to maximize their own utility.

II. Related Work

A. Data Forwarding Protocol

Data forwarding protocols have been extensively investigated in delay-tolerant networks. Due to the sparse and dynamic nature of delay-tolerant networks, user-to-user data forwarding often relies on the mobility and random contacts of users. For example, Lindgren et al. [13] evaluated the forwarding capability of a user by the historic contact information. Under the similar framework, [14]–[18] used social metrics calculated from the contact information to evaluate the forwarding capability of a user. Hui et al. [16] demonstrated that community and centrality social metrics can be effectively used in data forwarding protocol. Li et al. [17] introduced the social-selfish effect into user behavior, i.e., a user gives preference to packets received from other users with stronger social relations. Yuan et al. [18] proposed a data forwarding protocol enabling two users to share their historical mobility information. Based on the opponent's past mobility information, a user is able to predict the future location that the opponent will visit.

Though significantly improving the data forwarding effectiveness, most contact-based data forwarding protocols require a contact calculation phase in which each user must have a unique identity and reveal it to others. In this phase, user behaviors are very easy to be linked together and user's identity privacy and location privacy are completely compromised. In the contact-based data forwarding protocol, a sender must exchange the contact and unique identity with a relay user. In [13], [14], [17], to improve the forwarding effectiveness, a sender can evaluate the forwarding capability of a relay user based on both the relay user's contact probability and forwarding willingness. However, the required contact probability and unique identity information are privacy-sensitive to the relay user and not available in a privacy-preserving environment. The conventional contact-based data forwarding protocols do not provide effective privacy preservation and can hardly be accepted by the privacy-aware mobile users. In this paper, we aim to solve the privacy preservation and security issues of cooperative data forwarding in MSNs.

Recently a rich body of literature [19]–[24] addressed the cooperation stimulation issue from a game-theoretic perspective. Yu and Liu [19] proposed a game-based approach to stimulate cooperation in mobile ad hoc networks, where two participating users set a threshold on the number of forwarded packets in each forwarding round and they alternatively forward each other's packets. The setting of the threshold can stimulate cooperation and also limit the possible damage caused by the opponent's defection. If the opponent defects, a user immediately terminates the interaction and his maximum damage is bounded by the threshold setting in the previous round. Li and Shen [24] proposed an integrated system over an individual reputation system and a price-based system which demonstrates a superiority in terms of the effectiveness of cooperation incentives and selfish node detection. However, their works do not address user privacy and are not applicable in the

privacy-sensitive MSNs.

B. Privacy-preserving and Social Perspective

The studies in MSNs mainly focus on exploring the human factors and behaviors for communications in a distributed and autonomous environment. Privacy preservation as a fundamental user requirement is however neglected in previous research. Recent proposals [25] indicated that one or few snapshots of a user's location over time might assist an adversary to identify the user's trace, and an effective attack was presented to identify victims with high probability. As a defense technique, the *multiple-pseudonym* technique providing both identity and location privacy is widely applied in literatures [9], [10], [26]. Freudiger et al. [10] developed a user-centric location privacy model to measure the evolution of location privacy over time, and they derive the equilibrium strategies on changing pseudonyms for each user from the game-theoretic perspective. With the *multiple-pseudonym* technique applied, conventional cooperation stimulation mechanisms without privacy preservation [19], [27]–[29] are no longer applicable in the considered environment.

This work is inspired by the extensive literature in social theory related to human behaviors and their subjective morality. From social perspective, Keterlaar and Au [30] introduced an "affect-as-information" model to investigate how the past human behaviors influence on the future behaviors of human according to the internalized human rationalities. Based on this result, we develop the guilty model for cooperation stimulation in MSNs. In addition, some social-related works exploited a social graph of human to improve the protocol efficiency. For example, Li et al. [17] observed that if two users have a social relation in a social graph, they have more contacts than those who do not have a relation. Based on this observation, they demonstrated that introducing a social graph into the routing protocol results in a better network performance. However, the construction of such a social graph requires users to be fully cooperative without privacy concerns which is not feasible in MSNs. Other social-related works [30]–[32] considered the social relations including not only inter-user relations in a social graph but also the relations between users and established social organizations. Following this idea, we introduced a sociality strength for each user to model their cooperation behavior influenced by the social factor.

III. System Model

We consider a homogenous MSN composed of a set $\mathcal{V}$ of mobile users with the network size $|\mathcal{V}| = N$. Users follow the same behavior model: they are selfish, tending to maximize their individual utilities during data forwarding, and do not perform irrational attacks. A specific utility function will be given in Section IV. Users have equal communication range, denoted by R_t. The communication between any two users i and j, $i, j \in \mathcal{V}$, is bidirectional, i.e., user i can communicate to user j if and only if user j can also communicate to user i. A trusted authority (TA) is available at the initialization phase for generating pseudonyms and secret keys for MSN users, but it will not be involved in the data forwarding. Users continuously change their pseudonyms to preserve their identity and location privacy. The pseudonym change breaks any relation previously established between two users and as a result they can no longer recognize each other.

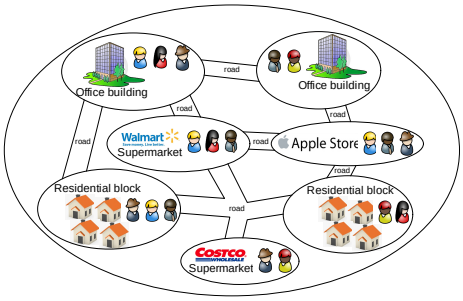


Fig. 1. A mobile social network

A. User-to-Spot Data Forwarding

We assume that there exists a set $\mathcal{A} = \{a_1, \cdots, a_l\}$ of social hotspots in the network. They are located in regions such as supermarkets, restaurants, office buildings and residential blocks with high population density as shown in Fig. 1. Different users have different sets of favored hotspots that they frequently visit. The hotspots that a user visited in the past indicate the personal preference of the user and thus may relate to the user's future locations [18]. In addition, the hotspots can be categorized into sensitive hotspots, e.g., office buildings, residential blocks, and non-sensitive hotspots, e.g., supermarkets, restaurants. Sensitive hotspots are tightly related to users' private lives. The access to sensitive hotspots needs to be protected according to users' privacy needs. In this work, we apply the hotspot technique [12] to preserve receiver location privacy.

We propose a user-to-spot data forwarding protocol to achieve privacy preservation and user cooperative data forwarding. Specifically, each hotspot is equipped with a non-compromised and communicable storage device which buffers the packets for receivers to fetch. A data sender/forwarder leaves packets at selected hotspots, and receivers can fetch the packets upon their later access to the same hotspots. Compared with the contact-based

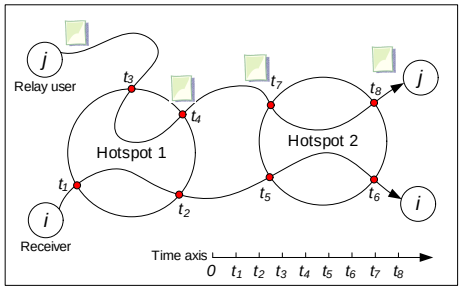


Fig. 2. An illustration of effective data forwarding by the proposed protocol

data forwarding protocols where users swap data upon their contacts, the user-to-spot data forwarding protocol would have more successful deliveries in special cases as shown in Fig. 2. In this figure, relay user j has no contact with receiver i but they enter the common hotspots during different time periods. By making use of this property, the user-to-spot data forwarding protocol enables j to deliver the packet to i. This user-to-spot data forwarding protocol is practical due to the following facts:

- Social users often have specific preferences on common social hotspots, such as supermarkets, office buildings, etc. They are likely to choose part of these hotspots and visit them frequently.
- In the MSN, data sender often has certain social relationship with receiver. The sender is likely to be partially aware of the social behaviors and frequently-visited hotspots of the receiver.
- Hotspot buffers are low-cost and can be pervasively available data storage resources [33]. They are not interconnected and will not be involved in cooperative data forwarding. They act as static receivers to temporarily store user data and allow authorized wireless access of the data when users come into their wireless communication range.

In this work, the identity of the receiver is implicitly contained (thus protected) in the packet, and the receiver can fetch the packet from the hotspot buffer after a simple authentication operation, e.g., using the scheme in [12].

B. Model of Social Morality

Social theory [34] indicates that in a fully autonomous system users behave independently based on the rational calculation of expediency. The decision on how to act in social interactions is viewed as either primarily economic and motivated by self-interest, or non-economic and motivated by collective interest and moral obligation. Different norms govern users' behavior in economic and non-economic spheres of activity, and appropriate behavior varies according to the context and the nature of the goods being exchanged. These norms are not just learned, but are incorporated by social users into their personalities. In reality, if users violate a deeply internalized norm, they would feel guilty to certain extent regardless of whether or not anyone else knew of their actions, and would likely punish themselves in some manner. This is known as social morality.

Social study [30] also indicates that individuals who experience feeling of guilt (compared to individuals who feel no guilt) after pursuing a non-cooperative strategy in the first round of play, display higher levels of cooperation in the subsequent round of play. Experimental results demonstrate that non-cooperative individuals who experience a certain level of guilt in a social bargaining game may use this feeling state as "information" about the future costs of pursuing a non-cooperative strategy. Their findings that the guilt manipulation interacted with social motives (e.g., guilt tended to have its intense effect on non-cooperative individuals) also suggest that these results do not occur merely because of pre-existing individual differences in the tendency to cooperate or defect. Instead, it would appear that guilt actually provokes non-cooperative individuals to depart from their "typical" strategy of non-cooperative behavior.

Observing the unique social features in the MSN, we exploit the morality factor of the MSN by mimicking the morality-centric human society. We emphasize that the morality factor should be counted into the calculation of users' utility. To this end, we instantiate two forms of social morality, i.e., guilt and high-mindedness, in the context of MSN-based data forwarding where cooperation is highly desirable: users feel *guilty* when they defect (i.e., refuse to forward a packet), and they feel *high-minded* when choosing to cooperate (i.e., help to forward a packet). Guilt creates a feeling of indebtedness, which directs them to cooperate, while high-mindedness alleviates the guilty feeling of users.

A self-regulated morality factor g, internalized for each user that quantitatively depicts the internal moral force, is based on two elements:

- *Morality state* x: The morality state reflects the behavior history of a user. It increases by one level for a single cooperation behavior and decreases by one level due to a single defection conduct.
- *Sociality strength* st: The sociality strength st is related to a user's personal experience, such as education and habitation. It is stabilized and less independent with short-term behavior changes. If the sociality strength of a user is significant, the user feels a correspondingly significant increment of guilt towards a single defection behavior and a correspondingly significant increment of high-mindedness towards a single cooperation behavior.

Each user i has a sociality strength denoted by st_i, and a varying morality state x_i. Following social theory [30], [32], we depict the morality state x_i by a Markov chain model with the state space and non-null transitions shown in Fig. 3. Let $P_i(j, j+1)$ and $P_i(j, j-1)$ denote the transition probabilities from the j-th state to the $(j+1)$-th and the $(j-1)$-th states, respectively. The state $j = 0$ is the initial neutral state (neither guilty nor high-minded). The states with a positive index are high-minded states, and those with a negative index are guilty states. Being in a high-minded state implies frequent cooperation behavior in the past; being in a guilty state indicates overwhelming defection conduct in the past. The morality factor g_i of user i is evaluated by a function $f(x_i, st_i)$ that increases as x_i decreases or st_i increases. Later, in Section VI when we present our performance evaluation, we will define a specific $f()$.

C. Overview of the Proposed Protocol

With the user-to-spot data forwarding protocol deployed, in the following sections, we concentrate on how to forward

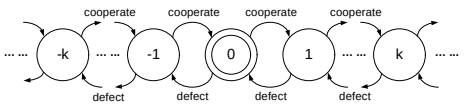


Fig. 3. Markov chain model for morality state

packets to the hotspots for effective and efficient data forwarding with privacy preservation. This delivery is enabled in three steps:

1) privacy-preserving route-based authentication,
2) proximity measurement,
3) morality-driven data forwarding.

In the first step, the privacy-preserving route-based authentication enables two encountered users to exchange partial route information. The route information can be constructed in a privacy-preserving structure determined by users themselves. The use of an authentication scheme is to resist user manipulation attacks, i.e., users have to honestly tell about their hotspots. In the second step, each user measures a proximity score between the destination and the route information provided by the relay user. The proximity score reflects the forwarding capability of a relay node with respect to a specific destination. The larger a proximity score is, the more effective a relay's forwarding is. In addition, the proximity score also affects the morality factor of the relay node. The rationale is that a user would feel more guilty if he/she demonstrates more capability to deliver a packet (with a large proximity score), and yet, drops the packet. In the third step, the morality factor is incorporated into the utility calculation of a data forwarding game in which users act selfishly and preserve their privacy. We elaborate these three steps in the subsequent sections. Note that, we do not consider irrational attacks in this paper. Users tend to be rational and selfish to maximize their own utility.

IV. Authentication and Proximity Measurement

In this section, we describe the first two steps of the proposed protocol, i.e., privacy-preserving route-based authentication and proximity measurement. The first step relies on a novel tree structure that provides limited user route information to the public to boost the data forwarding efficiency. The second step calculates the shortest distance between a destination and a hotspot that will be visited by a user. The distance implies the forwarding capability when this user attempts to forward packets to the destination.

A. Privacy-Preserving Route-Based Authentication

We first show how to construct a privacy-preserving *routing tree* which describes the route of user i between hotspots. At an initial stage, the TA associates user i to a subset of hotspots $\mathcal{A}_i = \{a_y | y = (2, 3, 6, 7)\} \subseteq \mathcal{A}$, which represents the hotspots frequently visited by user i. We consider that user i is located at hotspot a_1 and moving towards a_8, as shown in Fig. 4. Suppose that user i moves along the route $a_1 \rightarrow a_2 \rightarrow a_3 \rightarrow a_6 \rightarrow a_7 \rightarrow a_8$. Users

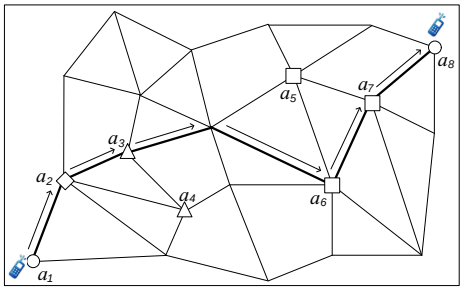


Fig. 4. Geographical view of user i's route

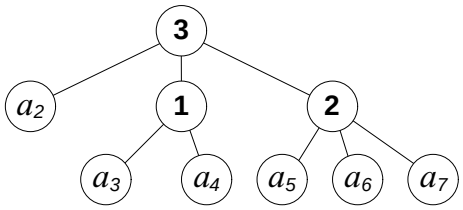


Fig. 5. Tree structure of user i's route

neighboring i have already known i's current location a_1. But i has no intention to reveal a_8 to them for privacy reason. In addition, it is unwilling to authenticate the entire hotspot set $\{a_y | y = (2, 3, 6, 7)\}$, which contains privacy-sensitive hotspots $\{a_3, a_6, a_7\}$. Then i creates a tree for its mobility route $\mathcal{T}_i$ as "a_2 AND (a_3 OR a_4) AND (2 of (a_5, a_6, a_7))" and only authenticates this tree to others. The authentication reveals the following fuzzy information instead of the precise route: user i will visit a_2, one of (a_3, a_4), and at least two hotspots from (a_5, a_6, a_7).

We present the routing tree structure $\mathcal{T}$ as shown in Fig. 5, where each non-leaf node represents a threshold gate and each leaf node represents a hotspot in $\mathcal{A}_u$. We use $\mathcal{A}_\mathcal{T} = \{a_{z_1}, a_{z_2}, \cdots, a_{z_\tau}\} \subseteq \mathcal{A}_u$ to denote the hotspot set corresponding to all leaf nodes in $\mathcal{T}$. Note that, if we assign 0 or 1 to the hotspots $(a_{z_1}, a_{z_2}, \cdots, a_{z_\tau})$ of leaf nodes in $\mathcal{T}$, $\mathcal{T}$ will be transformed into a Boolean function $F(a_{z_1}, a_{z_2}, \cdots, a_{z_\tau})$. For example, in Fig. 5, $F(a_1, a_2, \cdots, a_7) = a_2(a_3 + a_4)(a_5a_6 + a_5a_7 + a_6a_7)$. We say that a hotspot set $\mathcal{A}_i$ satisfies both $\mathcal{T}$ and function $F(a_{z_1}, a_{z_2}, \cdots, a_{z_\tau})$ if and only if $F(a_{z_1}, a_{z_2}, \cdots, a_{z_\tau}) = 1$, where for each a_y, $y \in \{z_1, z_2, \cdots, z_\tau\}$,

$$a_y = \begin{cases} 1, & \text{if } a_y \in \mathcal{A}_i, \\ 0, & \text{if } a_y \notin \mathcal{A}_i. \end{cases}$$

The routing tree preserves user privacy by making sensitive hotspots anonymous, and at the same time it provides certain information of the mobility route that can be used to evaluate the user's forwarding capability. We are now ready to present our privacy-preserving route-based authentication scheme which supports a single threshold gate (maximum threshold value d) for a routing tree. A multiple-threshold tree can be semantically converted to multiple single-threshold trees. The proposed scheme is built on the bilinear pairing technique [35], [36].

Initialization: Let $\mathbb{G}$ and $\mathbb{G}_T$ be two finite cyclic groups of the same large order n, where $n = pq$ is a

product of two large primes p and q. Suppose $\mathbb{G}$ and $\mathbb{G}_T$ are equipped with a pairing, i.e., a non-degenerated and efficiently computable bilinear map $e : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_T$ such that i) $\forall g, h \in \mathbb{G}$, $\forall a, b \in \mathbb{Z}_n$, $e(g^a, h^b) = e(g, h)^{ab}$; and ii) $\exists g \in \mathbb{G}$, $e(g, g)$ has order n in $\mathbb{G}_T$.

TA chooses a redundant hotspot set $\mathcal{A}_r = \{a_{l+1}, a_{l+d-1}\}$, two generators (g, u) of $\mathbb{G}$, a generator h of $\mathbb{G}_q$ ($\mathbb{G}_q$ is a subgroup of $\mathbb{G}$ with order q), a secure cryptographic hash function $H : \{0,1\}^* \to \mathbb{Z}_n^*$, and random number $\delta \in \mathbb{Z}_n^*$. For all $1 \le y \le l + d - 1$, TA chooses random numbers $t_y \in \mathbb{Z}_n^*$ and computes $T_y = g^{t_y}$. TA also computes $\Delta = e(g, u)^{\delta}$. With these settings, TA keeps the master key $(\delta, (t_y)_{1 \le y \le l+d-1})$ secretly, and publishes the public parameter pub $= (n, g, u, h, \mathbb{G}, \mathbb{G}_T, e, H, \Delta, T_y (1 \le y \le l+d-1), \mathcal{A} \cup \mathcal{A}_r)$.

USER REGISTRATION: TA chooses a unique random number $t \in \mathbb{Z}_n^*$ and a random polynomial $q(x) = \kappa_{d-1} x^{d-1} + \kappa_{d-2} x^{d-2} + \cdots + \kappa_1 x + \delta$, and generates $E_i = \langle k_d, (d_y)_{a_y \in \mathcal{A}_i \cup \mathcal{A}_r} \rangle$, where $k_d = t$ and $d_y = u^{\frac{q(y)}{t+t_y}}$. It informs the registering user i about the secret key E_i.

Let users i and j denote the signer and verifier respectively. Denote user i's routing tree (with a single threshold) by $\mathcal{T}_i$. Let k be the threshold value of the root of $\mathcal{T}_i$ and Θ_i a hotspot set corresponding to $\mathcal{T}_i$'s leaf nodes. $\Phi_i \subseteq \mathcal{A}_i \cap \Theta_i$ is a hotspot set of size k.

SIGNING BY USER i: User i first chooses a subset $\mathcal{A}_{r'} \subseteq \mathcal{A}_r$ ($|\mathcal{A}_{r'}| = d - k$). Let $\mathcal{A}_{r'}$ be $\{a_{l+1}, \cdots, a_{l+d-k}\}$. Then, for each hotspot $a_y \in \Psi = \Phi_i \cup \mathcal{A}_{r'}$, user i computes the Lagrange coefficient $\omega_y = \sum_{w | a_w \in \Psi, w \ne y} \frac{0-w}{y-w}$. It randomly selects $r_t, r_p, r_y \in \mathbb{Z}_n^*$ for $a_y \in \Theta_i \cup \mathcal{A}_{r'}$ and computes S_y for $a_y \in \Theta_i \cup \mathcal{A}_{r'}$ as

$$S_y = \begin{cases} d_y^{\omega_y} \cdot h^{r_y}, \text{ if } a_y \in \Psi \\ h^{r_y}, \text{ if } a_y \in \Theta_i \setminus \Phi_i \end{cases} \quad (1)$$

It outputs the signature

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

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

$$\pi_1 = S_p^{r_t} (g^{H(pid_i)} g^{k_d})^{r_p},$$

$$\text{and } \pi_2 = \prod_{a_y \in \Psi} (d_y^{\omega_y})^{r_t} \prod_{a_y \in \Theta_i \cup \mathcal{A}_{r'}} (S_t T_y)^{r_y}.$$

VERIFICATION BY USER j: User j receives σ_i and checks

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

If the above equations hold, user j confirms that user i has pseudonym pid_i and a hotspot set satisfying $\mathcal{T}_i$. The correctness of the verification is from the following mathematical manipulation:

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

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

Privacy discussion: For user privacy preservation, the route-based authentication scheme mixes the hotspot $a_y \in \Psi$ that user i has with the hotspot $a_y \notin \Psi$ that user i does not have from the equation (1) by multiplying a subgroup element h. This achieves full-anonymity, i.e., any other user cannot trace the hotspots which are used to generate the signature, because the element h cannot be distinguished from either $\mathbb{G}_p$ or $\mathbb{G}_q$ without p or q known as a priori. The theoretical proof can be found in [35], [36]. Consider that an adversarial user may use the authenticated route information to identify the signer's trace. Without precaution, such misbehavior may violate location privacy. An effective defense mechanism against this privacy violation is to let each user change the routing tree structures of their route information as frequently as the change of their pseudonyms, and also include redundant hotspots into their routing tree. As a result, different users may generate the same routing tree, and the signature cannot be used to link the past/future locations and behaviors of any specific user.

B. Proximity Measurement

In this section, we develop a novel proximity measurement for implementing the user-to-spot data forwarding protocol. Consider a packet originated from user j and destined to $\mathcal{D}_j$, which is a hotspot that its intended receiver frequently visits. When j meets a user i, it computes a forwarding score $e_{j,i}$. This score implies i's forwarding capability of bringing the packet to $\mathcal{D}_j$. It is subject to multiple factors such as the time-to-live period of the packet, the probability that i drops the packet due to limited storage buffer, how close that i can be to $\mathcal{D}_j$, when the closest distance will occur, and so on. However, the more factors used, the more personal information revealed, and the less privacy preserved.

To avoid any additional privacy leakage, we define that $e_{j,i} = \psi(r_{j,i})$, where $r_{j,i}$ is the smallest distance between $\mathcal{D}_j$ and the hotspots that i will visit and $\psi()$ is a monotonically decreasing function of $r_{j,i}$. The smaller $r_{j,i}$, the more closely i can deliver the packet to $\mathcal{D}_j$, the larger $e_{j,i}$ by this definition. A particular case is shown in Fig. 6. Even if user h appears to move away from $\mathcal{D}_j$, its forwarding, when used, will still be effective since it is going to encounter user i who will visit $\mathcal{D}_j$ afterwards. Given that no global knowledge is available and any user can be an effective forwarder, $\psi()$ always returns a positive value.

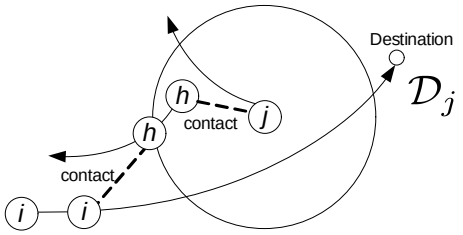


Fig. 6. A user moving towards opposite direction of the destination can still provide effective forwarding

Algorithm 1 Smallest radius calculation by user j

Input: $\mathcal{T}_i$ and $\mathcal{D}_j$.
Transform $\mathcal{T}_i$ to $F_i(a_{z_1}, a_{z_2}, \cdots, a_{z_\tau})$.
Calculate $\tilde{F}_i(a_{z_1}, a_{z_2}, \cdots, a_{z_\tau}) = \overline{F_i(\overline{a_{z_1}}, \overline{a_{z_2}}, \cdots, \overline{a_{z_\tau}})}$.
Calculate $D_s = \{d_{z_1}, d_{z_2}, \cdots, d_{z_i}\}$, where d_y is the distance between $\mathcal{D}_j$ and a_y for $y \in \{z_1, z_2, \cdots, z_\tau\}$.
Sort D_s in an ascending order $\{d_{z_1^*}, d_{z_2^*}, \cdots, d_{z_\tau^*}\}$ corresponding to spots $\{a_{z_1^*}, a_{z_2^*}, \cdots, a_{z_i^*}\}$.
Initialize $\tilde{\mathcal{A}} = \{a_{z_1^*}\}$, $\mu = 1$.
while ($\tilde{\mathcal{A}}$ does not satisfy $\tilde{F}_i(a_{z_1}, a_{z_2}, \cdots, a_{z_\tau})$) **do**
    $\mu = \mu + 1$,
    $\tilde{\mathcal{A}} = \tilde{\mathcal{A}} \cup \{a_{z_\mu^*}\}$.
end while
Let $r^*_{j,i} = d_{z_\mu^*}$ and $\mathcal{A}_{\mathcal{D}_j, r^*_{j,i}} = \tilde{\mathcal{A}}$.
Output $r^*_{j,i}$.

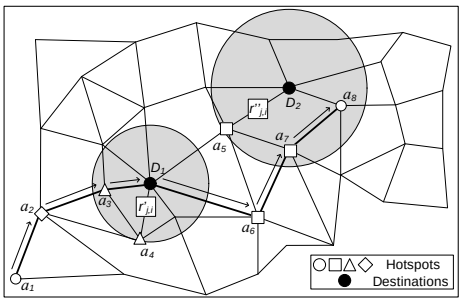


Fig. 7. An example of the smallest radius calculation

Since user i only exposes partial information $\mathcal{T}_i$ of its mobility route to user j during route-based authentication, user j cannot compute $r_{j,i}$ accurately. We devise an approximation algorithm for user j to obtain an approximate value $r^*_{j,i}$ with the inputs $\mathcal{T}_i$ and $\mathcal{D}_j$. In this algorithm, we first transform $\mathcal{T}_i$ to a Boolean function $F_i(a_{z_1}, a_{z_2}, \cdots, a_{z_\tau})$. We denote a self dual function of F_i as $\tilde{F}_i(a_{z_1}, a_{z_2}, \cdots, a_{z_\tau}) = \overline{F_i(\overline{a_{z_1}}, \overline{a_{z_2}}, \cdots, \overline{a_{z_\tau}})}$. Let $\mathcal{A}_{\mathcal{D}_j, r}$ denote a set of hotspots located in a circular area centered at the destination $\mathcal{D}_j$ with radius r. For a user i neighboring user j, we can find the smallest radius $r^*_{j,i}$ such that $\mathcal{A}_{\mathcal{D}_j, r^*_{j,i}}$ satisfies function $\tilde{F}_i(a_{z_1}, a_{z_2}, \cdots, a_{z_i})$. The algorithm finally outputs an approximate value $r^*_{j,i}$. The algorithmic detail is given in Algorithm 1. User j will then use this value $r^*_{j,i}$ to calculate the forwarding score of user i.

We use an example to illustrate how proximity score is computed, in accordance with the scenario given in Fig. 7. User i encounters user j. User i generates a routing tree $\mathcal{T}_i$ and the corresponding Boolean function $F_i(a_2, a_3, \cdots, a_7)$ ="a_2 AND (a_3 OR a_4) AND (2 of (a_5, a_6, a_7))". User j has two packets with destinations $\mathcal{D}_1$ and $\mathcal{D}_2$, respectively. We have $\tilde{F}_i(a_2, a_3, \cdots, a_7)$ ="a_2 OR (a_3 AND a_4) OR (2 of (a_5, a_6, a_7))". According to the Algorithm 1, with $\mathcal{T}_i$ and $\mathcal{D}_1$ as inputs, $\tilde{\mathcal{A}}$ is initialized to $\{a_3\}$ since a_3 is the hotspot closest to $\mathcal{D}_1$. Then, $\{a_4\}$ will be added into $\tilde{\mathcal{A}}$ since $\{a_3\}$ does not satisfy $\tilde{F}_i(a_2, a_3, \cdots, a_7)$ and a_4 is the second closest to $\mathcal{D}_1$. $\tilde{\mathcal{A}} = \{a_3, a_4\}$ now satisfies $\tilde{F}_i(a_2, a_3, \cdots, a_7)$. The algorithm finally outputs the distance $r'_{j,i}$ between a_4 and $\mathcal{D}_1$. Similarly, with $\mathcal{T}_i$ and $\mathcal{D}_2$ as inputs, the algorithm outputs the distance $r''_{j,i}$ between a_5 and $\mathcal{D}_2$, where $\tilde{\mathcal{A}} = \{a_5, a_7\}$ satisfying $\mathcal{T}_i$.

V. Morality-Driven Data Forwarding

After finishing the first two steps, users can perform morality-driven data forwarding. Note that the mobile social users are autonomous and intelligent individuals. It is reasonable to assume that they are rational and their behaviors are driven by personal profit and morality. On one hand, they tend to act defection in order to reduce their forwarding costs. On the other hand, they offer cooperation from time to time so as to counteract the guilty feelings brought by the past selfish deeds. During MSN-based data forwarding, social users implement the best strategy to balance cost and payoff. In this section, we apply game theory to model individual user behavior and obtain the optimal data forwarding strategy.

Consider a scenario where users move along independently and randomly determined mobility routes. Upon the contact with another user, a user would either cooperate or defect for data forwarding. We assume that, for two users that both have packets to send, cooperation is reciprocal. Due to the random mobility and privacy preservation, users' future contacts are unpredictable. A user thus derives the optimal data forwarding strategy based on its self-related information, including its own mobility route, destination of its own packet, and morality factor, as well as the opponent information, including the morality factor, mobility route and packet destination of the encountered user.

From a user's perspective, among a series of cooperations with different encountered opponents, due to the privacy preservation, the opponent information of current contact is always independent from that of previous contacts, and thus the decision on cooperation or defection depends only on the self-related information and the opponent information of the current contact. We thus model the interplay upon each contact, namely cooperation game, as a nonzero sum two-player game.

A. Basic/Extended Cooperation Games

We first define a basic cooperation game, called B-game (B stands for Basic), as a 3-tuple $(\mathcal{N}, \mathcal{S}, \mathcal{P})$, where $\mathcal{N}$ is a pair of users, $\mathcal{S}$ is a set of strategies and $\mathcal{P}$ is a set of payoff functions. According to Section III, users continuously change their pseudonyms to preserve their privacy. Pseudonym change breaks any relation previously established between two users and as a result they no longer

TABLE I
PAYOFF MATRIX

(a) Payoff matrix of two-player B-game

$i \setminus j$	Cooperate (C)	Defect (D)
Cooperate (C)	$(b-c, b-c)$	$(-c, b)$
Defect (D)	$(b, -c)$	$(0, 0)$

(b) Payoff matrix of two-player E-game

$i \setminus j$	C	D
C	$(b-c, b-c)$	$(-c, b-g_j)$
D	$(b-g_i, -c)$	$(-g_i, -g_j)$

(c) Payoff matrix of two-player S-game

$i \setminus j$	C	D
C	$(e_{i,j}b-c, e_{j,i}b-c)$	$(-c, e_{j,i}b-e_{i,j}g_j)$
D	$(e_{i,j}b-e_{j,i}g_i, -c)$	$(-g_i, -g_j)$

recognize each other. Therefore, B-game is a non-repeated game which can be described as follows:

- **Players:** Two users i and j belong to the universal user set $\mathcal{V}$. User j can also be denoted as $-i$. The two users are within the transmission range of each other, and they decide to cooperate or defect, aiming at maximizing their individual payoff.
- **Strategy:** Upon the forwarding request of the opponent user, each user has two strategies: Cooperate (C) and Defect (D). Denote user i's strategy by s_i. Then $s_i = C$ means that user i forwards user j's packet, and $s_i = D$ that user i drops user j's packet.
- **Payoffs:** The cost c of forwarding on one packet is a value, the same for both users. If user i's data is forwarded by user j, the profit acquired by user i is b, which is also a constant. We set $b \geq c > 0$ since the profit acquired from each forwarding should be at least equal to the incurred cost. The user payoffs under different strategies are shown in Table I(a).

From the payoff matrix in Table I(a), it is observed that the B-game is a typical prisoner-dilemma game, where the only Nash Equilibrium (NE) is (D, D) for non-repeated version. In other words, no matter what the opponent's strategy is, the best strategy for a user is to defect. This is because that $b > b-c, 0 > -c$.

Next, we introduce an E-game (E stands for Extended), where the payoff matrix is shown in Table I(b). This game considers users i and j's behaviors affected by morality factors g_i and g_j. The morality factors are introduced as the costs of defection behaviors into the payoff functions. The best strategy of the E-game for user i is: cooperate if $g_i > c$; defect if $g_i \leq c$. Based on the Markov chain model given in Section III-B, there exists a morality state $x^* < 0$ such that $f(st_i, x^*+1) < c < f(st_i, x^*)$. After a finite series of defections, user i will reach state x^*, and then alternatively chooses to cooperate.

B. Social Cooperation Game

In the following, we extend the E-game to a complex S-game (S stands for Social), which is also denoted by a 3-tuple $(\mathcal{N}, \mathcal{S}, \mathcal{P})$. $\mathcal{S}$-game further incorporates the forwarding scores $e_{i,j}$ and $e_{j,i}$ computed in the previous two steps (see Section IV) into the payoff function.

- **Players:** Two users i and j with different sociality strength st_i, st_j and current morality factors g_i, g_j.
- **Strategy:** The strategy is the same as that of the B-game. User i's strategy is denoted by s_i.
- **Payoffs:** The payoff of user i is evaluated by

$$p_i^s = \begin{cases} e_{i,j}b - c, & \text{if } s_i = C,\ s_j = C, \\ -c, & \text{if } s_i = C,\ s_j = D, \\ e_{i,j}b - e_{j,i}g_i, & \text{if } s_i = D,\ s_j = C, \\ -g_i, & \text{if } s_i = D,\ s_j = D. \end{cases} \quad (2)$$

In payoff formula (2), the forwarding scores $e_{i,j}$ and $e_{j,i}$ are used to measure user i's profit and morality factor. If user j forwards user i's data, the profit that user i acquires is $e_{i,j}b$ instead of b. If user i drops user j's data, depending on user j's strategy, user i acquires different morality factors, $e_{j,i}g_i$ or g_i. Note that, when users i and j both drop each other's packets, the morality factor on user i's payoff is independent of the forwarding score $e_{j,i}$. This is because users i and j treat each other equally and do not further consider their forwarding capability.

1) S-game with complete information: We first analyze the S-game in the case that two players have complete information including the sociality strength and morality state of each other. Each player can calculate the morality factor by $\psi()$ as defined in Section III-B and determine the payoff before deciding whether to cooperate or defect, according to Table I(c). We use *Theorem* 1 to identify the NE strategies of the S-game.

Theorem 1: When the two players have complete information of each other in the S-game, there are multiple pure-strategy NE in different cases and one mixed-strategy NE (x_i, x_j), where $x_i = \frac{c-g_{-\theta}}{e_{\theta,-\theta}g_{-\theta}-g_{-\theta}}$ is the probability that user n_θ chooses to cooperate, as shown in Fig. 8(b).

Proof: For $\theta = i$ or j, we have the following three cases to consider.

• $g_\theta < \frac{c}{e_{-\theta,\theta}}$: We have $e_{\theta,-\theta}b - e_{-\theta,\theta}g_\theta > e_{\theta,-\theta}b - c$ and $-g_\theta > -\frac{c}{e_{-\theta,\theta}} \geq -c$ due to $e_{-\theta,\theta} \geq 1$. As a result, when $g_{-\theta} < c$, $(s_\theta = D, s_{-\theta} = D)$ is a NE; and when $g_{-\theta} > c$, $(s_\theta = D, s_{-\theta} = C)$ is a NE.

• $g_\theta > c$: We have $-g_\theta < -c$ and $e_{\theta,-\theta}b - e_{-\theta,\theta}g_\theta < e_{\theta,-\theta}b - c$ due to $e_{-\theta,\theta} \geq 1$. As a result, when $g_{-\theta} > \frac{c}{e_{\theta,-\theta}}$, $(s_\theta = C, s_{-\theta} = C)$ is a NE; when $g_{-\theta} < \frac{c}{e_{\theta,-\theta}}$, $(s_\theta = C, s_{-\theta} = D)$ is a NE.

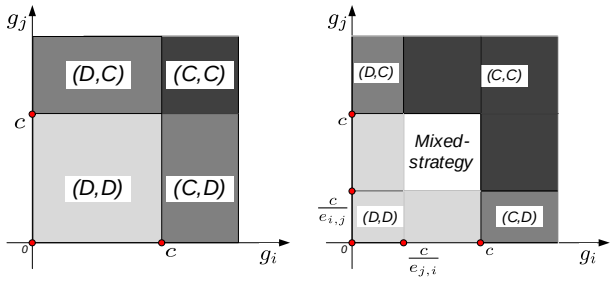
• $\frac{c}{e_{-\theta,\theta}} < g_\theta < c$: Let x_θ denote the forwarding probability of user n_θ. For $s_{-\theta} = C$ or $s_{-\theta} = D$, we separately calculate the payoff for $n_{-\theta}$ as follows:

$$p_{-\theta}^s|C = x_\theta \times (e_{-\theta,\theta}b - c) + (1 - x_\theta) \times (-c)$$

$$p_{-\theta}^s|D = x_\theta \times (e_{-\theta,\theta}b - e_{\theta,-\theta}g_{-\theta}) + (1 - x_\theta) \times (-g_{-\theta})$$

If x_θ is the best strategy of n_θ, we have $p_{-\theta}^s|C = p_{-\theta}^s|D$ which gives $x_\theta = \frac{c-g_{-\theta}}{e_{\theta,-\theta}g_{-\theta}-g_{-\theta}}$. ■

2) S-game with incomplete information: We consider the case that the two players have incomplete information of each other. Specifically, user i obtains sociality strength st_i, morality state g_i, forwarding scores $e_{i,j}$ and $e_{j,i}$, but it does not obtain the sociality strength st_j and morality factor g_j of user j. As a supplementary information, we



(a) The best strategy of E-game (b) The best strategy of S-game

Fig. 8. The best strategy for different games

assume that user i obtains the probability distribution ϱ of the morality factor of all users. Based on this, user i can estimate the morality factor g_j of user j. Then, user i follows the following steps according to the best strategy shown in Fig. 8(b):

- If $0 \leq g_i < \frac{c}{e_{j,i}}$, then user i chooses to defect regardless of user j's strategy.
- If $c \leq g_i$, then user i chooses to cooperate regardless of user j's strategy.
- If $\frac{c}{e_{j,i}} \leq g_i < c$, then there exists a pure-strategy NE (D, D) for $g_j < \frac{c}{e_{i,j}}$, a pure-strategy NE (C, C) for $g_j > c$, and a mixed-strategy NE for $\frac{c}{e_{i,j}} < g_j < c$. For the pure strategy NE, we calculate the defection probability $\Pr_1$ and cooperation probability $\Pr_2$:

$$\Pr_1 = \Pr(0 \leq g_j < \frac{c}{e_{i,j}}) = \int_0^{\frac{c}{e_{i,j}}} \varrho(\alpha)d\alpha,$$
$$\Pr_2 = \Pr(c \leq g_j) = \int_c^{+\infty} \varrho(\alpha)d\alpha.$$

In addition, user i makes a mixed-strategy NE with probability $\Pr_3$, which is given by

$$\Pr_3 = \Pr(\frac{c}{e_{j,i}} \leq g_i < c) = \int_{\frac{c}{e_{j,i}}}^{c} \varrho(\alpha)d\alpha.$$

For the mixed-strategy NE with probability $\Pr_3$, *Theorem* 1 indicates the best strategy of user i is to forward the data with probability $\frac{c-g_j}{e_{i,j}g_j-g_j}$ if g_j is known by user i. In this case, the probability that user i chooses to cooperate is

$$\Pr_4 = \int_{\frac{c}{e_{j,i}}}^{c} (\frac{c-\alpha}{e_{i,j}\alpha-\alpha})\varrho(\alpha)d\alpha. \quad (3)$$

Overall, user i decides to cooperate with probability $\Pr_F = \Pr_2 + \Pr_4$ and to defect with probability $\Pr_D = 1 - \Pr_F$.

C. S-Game Based Data Forwarding

Notice that the S-game with incomplete information emulates MSN environments in reality, where the opponent's morality factor cannot be directly obtained. We use the optimal strategy of this game in our protocol for users to make the optimal data forwarding strategies. As we defined in Section III-B, user morality factor would vary with both sociality strength and morality state. However, revealing such information violates user privacy since other adversarial users can utilize the information to track user behavior. In this case, we do not require an accurate calculation of morality factor in the S-game. Instead, we examine the proposed strategy by using a probability distribution function ϱ of morality factor. This function ϱ can be either observed by a trusted authority or reported by individual users. Further analysis is presented in Section VI-B3.

A user who has packets to forward starts the data forwarding protocol with a randomly selected neighbor. Consider two neighboring users i and j that are running the protocol, i.e., they are both able to provide cooperative data forwarding to each other and any forwarding/defection decision in the two-user game will impact their social morality. Let $S_i = \{p_{i_1}, p_{i_2}, \cdots, p_{i_\alpha}\}$ and $S_j = \{p_{j_1}, p_{j_2}, \cdots, p_{j_\beta}\}$ be the packet sets held by i and j, respectively. We summarize the protocol as follows. User i first randomly selects a packet p_x (destined to $\mathcal{D}_i$) from its local repository. It then calculates the digest of the packet $d_i = H(p_x)$, where H is the cryptographic hash function. Lastly, it sends d_i to j. In the meantime, user j executes a similar procedure locally and sends i the digest d_j of a packet of p_y (destined to $\mathcal{D}_j$). According to d_i (d_j), if j (resp., i) finds that it already has p_x (resp., p_y), it will inform i (resp., j) to re-select p_x (resp., p_y). Through exhaustive packet re-selection, if they cannot find any exclusively owned packet, the protocol will terminate. Otherwise, they proceed to exchange $(p_x, \mathcal{D}_i)$ and $(p_y, \mathcal{D}_j)$, together with their own routing trees $\mathcal{T}_i$ and $\mathcal{T}_j$. Then, they validate each other's routing trees (see Section IV-A). After that, they evaluate each other's forwarding scores (see Section IV-B), and finally make the forwarding strategy for each other (see Section V-B2).

VI. Performance Evaluation

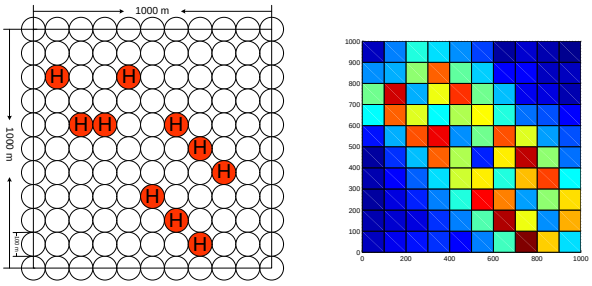
In this section, we conduct trace-based custom simulations of the MSN to evaluate the proposed data forwarding protocol.

A. Simulation Settings

1) User mobility, hotspots, and packet generation: We generate user mobility model according to the real-world trace of pedestrian runners provided in [37]. In the real trace set, $N = 100$ mobile users are randomly deployed in a 1000×1000 m^2 square region with the velocity randomly distributed with a mean value of 1 m/s. The communication range R_t of users is set to 50 m. The log contains the user locations in successive $T = 900$ time slots.

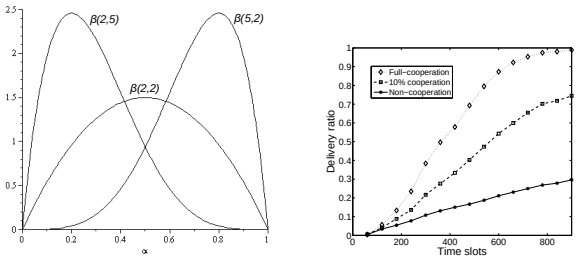
We divide the network field into 10×10 grids, where each grid is a square with side length 100 m. We create a circle of radius R_t around each grid point, and there are totally 121 circles. The areas enclosed by these circles are called spots and denoted by $(a_1, a_2, \cdots, a_{121})$ as shown in Fig. 9; no any two spots overlap.

We aggregate user route information to determine the most popular spots as follows. Let $d_{m,n}$ denote the number



(a) Hotspots (denoted by H) (b) Population density

Fig. 9. Hotspots and population density



(a) Sociality strength (b) Cooperation effect

Fig. 10. Preliminary results

of users in hotspot a_m at time slot n, where integers $m \in [1, 121]$ and $n \in [1, 900]$. We sort the spots in an descending order according to $d_m = \sum_{n=1}^{T} d_{m,n}$, and choose the top-ten spots as hotspots (i.e., $l = 10$). Figure 9(a) shows the selected hotspots, and Fig. 9(b) shows the population density of spots according to d_m. At the middle of each hotspot, we place a wireless storage device which has a communication range equal to R_t. Once a user enters a hotspot, it can access the storage device of the hotspot via wireless communication.

For each simulation run, there are totally 1000 packets generated for transmissions, 100 packets per each user, with uniformly selected hotspots as the packet destinations. In each time slot, a user i randomly selects a neighboring user j to play a two-player cooperation game. In the cooperation game, we consider the communication cost of data forwarding to be much greater than the computational cost of the associated authentication. As such, the authentication scheme imposes negligible influence on user behavior. Upon each contact, users uniformly select one available packet from their buffers to transmit. In order to focus on the impact of cooperation on the data forwarding effectiveness, we consider packets do not expire during the simulations and hotspot buffers and user device buffers are not limited in size.

2) Sociality Strength and Morality Function: The sociality strength st_i of user i $(1 \leq i \leq 100)$ is selected from the range of $[0, 1]$. The greater st_i is, the more intense social morality impact on user i's cooperation. In this section, we adopt different models of sociality strength represented by three beta distributions $\beta(2, 5), \beta(2, 2), \beta(5, 2)$ shown in Fig.10(a), respectively, to evaluate the performance of the proposed protocol in the cases of low, medium and high users' sociality strength, respectively.

The morality function f is used to calculate the morality factor of each user i using the user's sociality strength st_i and current morality state x. From Section III-B, we define three morality functions: *linear* function f_1, *natural logarithm* function f_e and *common logarithm* function f_{10}. They outputs 0 if $x \geq 0$, and otherwise,

$$f_1(st_i, x) = k \cdot st_i \cdot (-x)$$
$$f_e(st_i, x) = k \cdot \ln(1 + st_i \cdot (-x))$$
$$f_{10}(st_i, x) = k \cdot \log_{10}(1 + st_i \cdot (-x))$$

where k is a tunable coefficient in the range of $(0, +\infty)$. For simplicity, we fix $k = 1$ in our simulation.

The three morality functions represent three different levels of morality force affecting user cooperation behavior, respectively. They always output a non-negative value. The *common logarithm* function f_{10} generates a smaller morality factor, compared with the other two functions. If it is adopted, we can expect to see more defection behaviors.

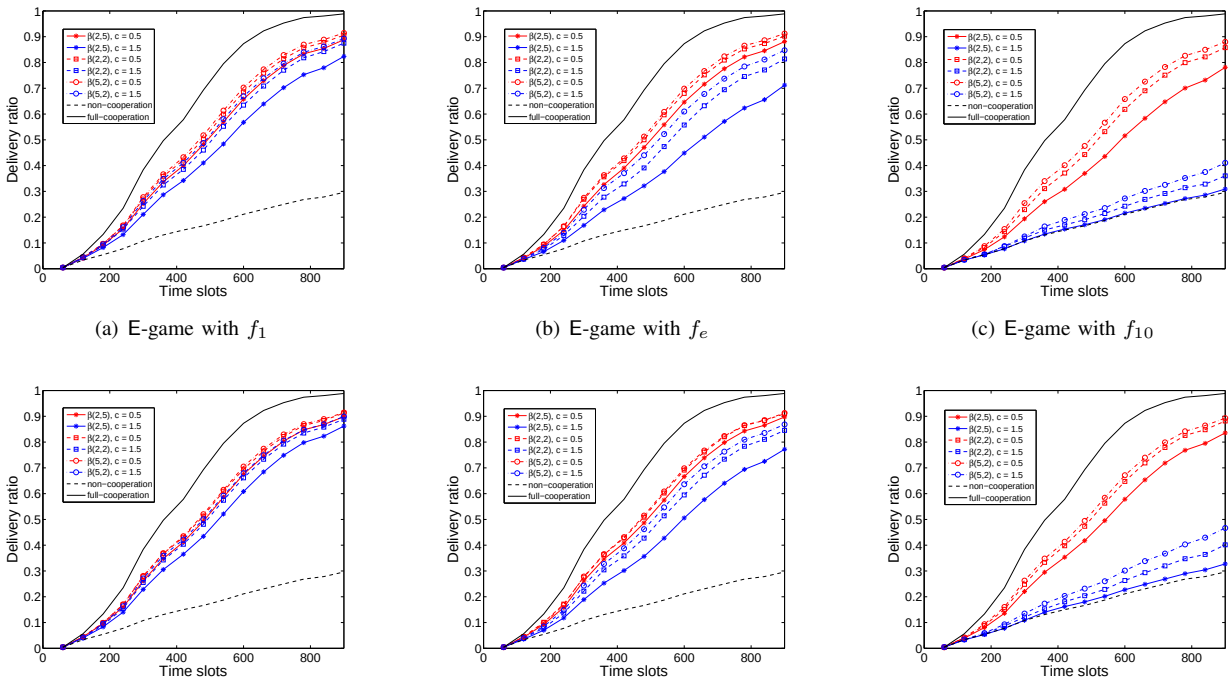
3) Routing tree and forwarding capability: Recall that a user's routing tree preserves user privacy by making the sensitive hotspots anonymous, and in the meantime provides partial information of user mobility route in order to facilitate cooperative data forwarding. With 10 hotspots in simulations, each user i may have at most 10 hotspots and at least 0 hotspot in $\mathcal{A}_i$. We generate a simplified routing tree structure $\mathcal{T}$ in the following way: if $|\mathcal{A}_i| = 0$, the tree cannot be created; if $0 < |\mathcal{A}_i| < 5$, we set the threshold as $|\mathcal{A}_i|$, and the leaf nodes as all the hotspots of $\mathcal{A}_i$ and other $5 - |\mathcal{A}_i|$ ones from $\mathcal{A}_u \setminus \mathcal{A}_i$; if $|\mathcal{A}_i| \geq 5$, we set the threshold as 4, and the leaf nodes as four randomly selected hotspots from $\mathcal{A}_i$ and another different hotspot. In short, for every user, the tree structure can be written as "t of 5", where $1 \leq t \leq 4$.

In Section IV-B, a function ψ is used to compute the forwarding capability of a given user i for a packet with a specific destination. We set the lower bound of ψ as 1. In the network grid, $r^*_{i,j}$ can be $1000 \times \sqrt{2} = 1415$ meters at most and 0 at least. Intuitively, if $r^*_{i,j} = 1415$, the forwarding capability $e_{i,j}$ reaches the minimum value; and if $r^*_{i,j} = 0$, $e_{i,j}$ reaches the maximum value. We define $\psi(r^*_{i,j}) = e^{k' - k' r^*_{i,j}/1415}$ and set $k' = 3$ as an example to illustrate the effect of forwarding capability.

B. Simulation Results

The performance metrics used in the simulation are: i) the delivery ratio, which is the fraction of packets that are correctly delivered to the hotspots as their destinations; and ii) the average morality state, which reflects the intention of users to cooperate over time. The delivery ratio examines the overall cooperation of users in the MSN, while the average morality state denotes the long-term cooperation strategies for a single user. For each simulation, we conduct 50 simulation runs and report the average results.

1) B-game: We first examine the B-game, where users always choose defection as the best strategy as discussed



(a) E-game with f_1 (b) E-game with f_e (c) E-game with f_{10}

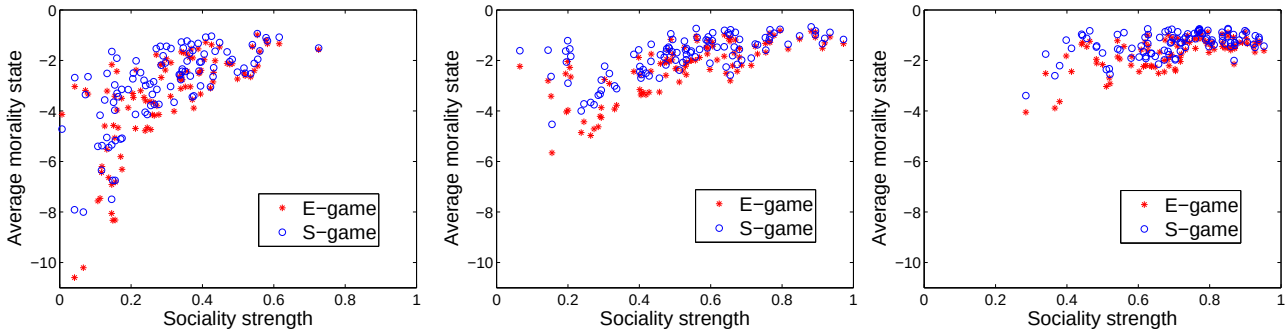
(d) S-game with f_1 (e) S-game with f_e (f) S-game with f_{10}

Fig. 11. Delivery ratio in E-game and S-game with complete information

in Section V-A. Fig. 10(b) shows three delivery ratios in the following three cases: a) users do not cooperate (i.e., B-game); b) users stochastically cooperate to forward packet with the probability of 10%; and c) users fully cooperate. It can be seen that at time slot 900, the full-cooperation strategy achieves 99% delivery ratio while the non-cooperation strategy achieves only 30%. Furthermore, Fig. 10(b) indicates that the probabilistic cooperation strategy provides a significant improvement to the delivery ratio up to 74%. However, without effective incentive and appropriate exploration of their social feature, users will not take cooperation due to the selfishness. Successful delivery happens only when the data senders arrive at their selected hotspots. This inevitably results in a low delivery ratio in the B-game.

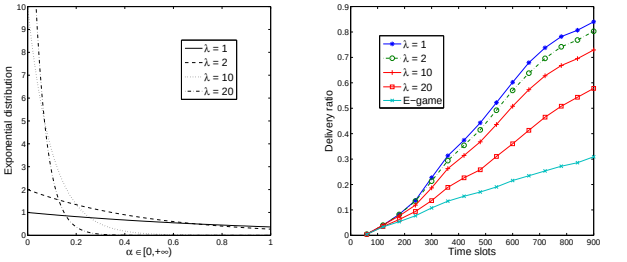
2) E-game and S-game: The E-game extends the B-game by embedding the morality factor into the payoff function as shown in Table I(b), while the S-game further considers the forwarding capability into the payoff of the E-game. Fig. 11 shows the delivery ratio of both the E-game and the S-game with complete information, with red lines representing the performance of forwarding cost $c = 0.5$, blue lines representing that of $c = 1.5$, and black lines depicting those of full-cooperation and non-cooperation as the best and worst case. It is clearly observed that the strategies with $c = 0.5$ can achieve higher delivery ratio than the strategies with $c = 1.5$. The rationale is that a large forwarding cost $c = 1.5$ hinders the cooperation performed by users who have limited resources and thus limits guilty incentive. In particular, when f_{10} is adopted, the cooperation condition in case of $c = 1.5$ approaches to the worst case. This is because that the guilty function f_{10} returns the smallest morality factor resulting in the least incentive to cooperate, compared to the function f_e and f_1. Fig. 11 shows that the strategies with the sociality strength $\beta(5, 2)$ perform much better than those with $\beta(2, 2)$ and $\beta(2, 5)$ in terms of delivery ratio. This is because that, compared to cases $\beta(2, 2)$ and $\beta(2, 5)$, users will be initialized with larger sociality strength in case $\beta(5, 2)$ as shown in Fig. 10(a), and as discussed in Section III-B, more users feel intense guilt towards their defection and choose to cooperate, which leads to a better performance.

Fig. 11 shows the performance comparisons between the E-game and the S-game under the same parameters. It can be seen that the delivery ratio can be further improved by enabling privacy-preserving route-based authentication. But since the route information is limited due to the privacy preservation, the improvements are not significant, e.g., when choosing $\beta(2, 5)$ and $c = 1.5$, the delivery ratio increases from 0.309 as shown in Fig. 11(c) to 0.327 as shown in Fig. 11(f). To further investigate the impact of the route information on the data forwarding cooperation, we randomly select 100 users in the network and examine their average morality states. Fig. 12 shows the average morality state of each selected user in terms of the user sociality strength in three settings of social strength $\beta(2, 5), \beta(2, 2)$, and $\beta(5, 2)$, respectively. The blue circle represents a user which adopts the best strategy from the S-game, and the red star represents a user which adopts the best strategy from the E-game. It can be seen that with the same sociality strength, the users represented by the red star have smaller



(a) Sociality strength $\beta(2,5)$ (b) Sociality strength $\beta(2,2)$ (c) Sociality strength $\beta(5,2)$

Fig. 12. Average morality states of all users in E-game and S-game with complete information, $c = 0.8$, and common logarithm



(a) Morality factor (b) Delivery ratio, f_{10}, $c = 1.5$, $st_i \sim \beta(2,5)$

Fig. 13. $\mathcal{S}$-game with incomplete information

morality states than users represented by the blue circle. This is to say, the incentive to defect in the cooperation game can be further reduced by enabling privacy-preserving route-based authentication.

3) S-game with Incomplete Information: For the S-game with incomplete information, the morality factor cannot be obtained directly in our morality model due to the lack of sociality strength and morality state information about the opponent user. As such, the morality factor will be estimated by a probability distribution function ϱ. In our simulation, we use exponential distribution with parameter $\lambda = \{1, 2, 10, 20\}$ to generate the morality factors for all users. The probability distribution function ϱ is shown in Fig. 13(a).

Fig. 13(a) shows that most users in case of $\lambda = 1$ may have relatively large morality factor. As we make $st_i \sim \beta(2,5)$, most users would have the weak sociality strength. Thus, the large morality factors of users indicate that they have already adopted a large amount of defections. Accordingly, they would have intense guilty feeling so that their following behaviors are probably cooperative. Besides that, it can be seen that when $\lambda = 20$ most users with the weak sociality strength have smaller morality factors, and without enough guilt as cooperation incentives their future behaviors would likely be defections. The performance results from Fig. 13(b) validate the above analysis, where the delivery ratio largely decreases if λ changes from 1 to 20. By investigating the proposed strategy, it can be seen that when $\lambda = 20$, from user i's perspective, the opponent user j has a morality factor $g_j < \frac{c}{e_{i,j}}$ with a large probability. In this case, user i chooses to cooperate if $g_i \geq c$ and defect if $g_i < c$. The best strategy of the S-game with incomplete information is thus almost equal to that of the E-game; both games indicate users to cooperate or defect mostly based on user self morality factors. However, the S-game with incomplete information outperforms the E-game since it has an additional mixed-strategy space shown in Fig. 8(b) to encourage user cooperation.

VII. Conclusion

In mobile social networks (MSNs), the two fundamental design goals – privacy preservation and cooperative data forwarding – would severely conflict with each other if carelessly designed. This is because that concealing and protecting user information may prohibit tracking the social behavior of users, which impedes the cooperative data forwarding and effective incentive mechanism. In this paper, we have attained the two conflicting design goals in one framework by exploiting social morality. Specifically, we first have proposed a novel user-to-spot data forwarding protocol where each packet is destined to a hotspot associated with the receiver and then retrieved by the receiver upon its access to the hotspot. With this protocol, not only can receiver location privacy be preserved, but the packet delivery ratio is also enhanced. In addition, a privacy-preserving route-based authentication scheme has been integrated to allow users to reveal anonymized route information to the public. Based on the information, a user is able to evaluate the data forwarding capability of each relay for a given packet with a specific destination. Game-theoretic models then have been adopted to derive the best data forwarding strategy for users, with respect to user morality factor, interest and forwarding capability. Through extensive trace-based simulations, we have demonstrated the data forwarding effectiveness of the proposed protocol in terms of packet delivery ratio. Particularly, the embedded privacy-preserving route-based authentication scheme makes important contribution to the protocol performance. For the future work, we will extend this work by studying

a more general and complicated situation in which mobile social users may have diverse behavior models.

REFERENCES

[1] A. Miklas, K. Gollu, K. Chan, S. Saroiu, P. Gummadi, and E. Lara, "Exploiting social interactions in mobile systems," in *Ubicomp*, 2007, pp. 409–428.

[2] S. Ioannidis, A. Chaintreau, and L. Massoulié, "Optimal and scalable distribution of content updates over a mobile social network," in *Proc. IEEE INFOCOM*, 2009, pp. 1422–1430.

[3] R. Lu, X. Lin, and X. Shen, "Spring: A social-based privacy-preserving packet forwarding protocol for vehicular delay tolerant networks," in *Proc. IEEE INFOCOM*, 2010, pp. 632–640.

[4] W. He, Y. Huang, K. Nahrstedt, and B. Wu, "Message propagation in ad-hoc-based proximity mobile social networks," in *PERCOM workshops*, 2010, pp. 141–146.

[5] D. Niyato, P. Wang, W. Saad, and A. Hjørungnes, "Controlled coalitional games for cooperative mobile social networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 4, pp. 1812–1824, 2011.

[6] M. Motani, V. Srinivasan, and P. Nuggehalli, "Peoplenet: engineering a wireless virtual social network," in *MobiCom*, 2005, pp. 243–257.

[7] M. Brereton, P. Roe, M. Foth, J. M. Bunker, and L. Buys, "Designing participation in agile ridesharing with mobile social software," in *OZCHI*, 2009, pp. 257–260.

[8] S. Gaonkar, J. Li, R. R. Choudhury, L. P. Cox, and A. Schmidt, "Micro-blog: sharing and querying content through mobile phones and social participation." in *ACM MobiSys*, 2008, pp. 174–186.

[9] M. Li, K. Sampigethaya, L. Huang, and R. Poovendran, "Swing & swap: user-centric approaches towards maximizing location privacy," in *WPES*, 2006, pp. 19–28.

[10] J. Freudiger, M. Manshaei, J.-P. Hubaux, and D. Parkes, "On non-cooperative location privacy: a game-theoretic analysis," in *ACM CCS*, 2009, pp. 324–337.

[11] R. Lu, X. Lin, H. Luan, X. Liang, and X. Shen, "Pseudonym changing at social spots: An effective strategy for location privacy in vanets," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 1, pp. 86–96, 2012.

[12] X. Lin, R. Lu, X. Liang, and X. Shen, "Stap: A social-tier-assisted packet forwarding protocol for achieving receiver-location privacy preservation in vanets," in *Proc. IEEE INFOCOM*, 2011, pp. 2147–2155.

[13] A. Lindgren, A. Doria, and O. Schelén, "Probabilistic routing in intermittently connected networks," in *SAPIR*, 2004, pp. 239–254.

[14] E. M. Daly and M. Haahr, "Social network analysis for routing in disconnected delay-tolerant manets," in *MobiHoc*, 2007, pp. 32–40.

[15] W. Gao, Q. Li, B. Zhao, and G. Cao, "Multicasting in delay tolerant networks: a social network perspective," in *MobiHoc*, 2009, pp. 299–308.

[16] P. Hui, J. Crowcroft, and E. Yoneki, "Bubble rap: Social-based forwarding in delay-tolerant networks," *IEEE Transactions on Mobile Computing*, vol. 10, no. 11, pp. 1576–1589, 2011.

[17] Q. Li, S. Zhu, and G. Cao, "Routing in socially selfish delay tolerant networks," in *Proc. IEEE INFOCOM*, 2010, pp. 857–865.

[18] Q. Yuan, I. Cardei, and J. Wu, "Predict and relay: an efficient routing in disruption-tolerant networks," in *MobiHoc*, 2009, pp. 95–104.

[19] W. Yu and K. J. R. Liu, "Game theoretic analysis of cooperation stimulation and security in autonomous mobile ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 6, no. 5, pp. 507–521, 2007.

[20] H. Zhu, X. Lin, R. Lu, Y. Fan, and X. Shen, "Smart: A secure multilayer credit-based incentive scheme for delay-tolerant networks," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 8, pp. 4628–4639, 2009.

[21] M. H. Manshaei, Q. Zhu, T. Alpcan, T. Basar, and J.-P. Hubaux, "Game theory meets network security and privacy," Ecole Polytechnique Fédérale de Lausanne (EPFL), Tech. Rep., September 2010, epfl-report-151965.

[22] M. Raya, R. Shokri, and J.-P. Hubaux, "On the tradeoff between trust and privacy in wireless ad hoc networks," in *WISEC*, 2010, pp. 75–80.

[23] M. Mahmoud and X. Shen, "Pis: A practical incentive system for multihop wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 8, pp. 4012–4025, 2010.

[24] Z. Li and H. Shen, "Game-theoretic analysis of cooperation incentive strategies in mobile ad hoc networks," *IEEE Transactions on Mobile Computing*, 2011, preprint.

[25] C. Y. T. Ma, D. K. Y. Yau, N. K. Yip, and N. S. V. Rao, "Privacy vulnerability of published anonymous mobility traces," in *MobiCom*, 2010, pp. 185–196.

[26] D. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," *Commun. ACM*, vol. 24, no. 2, pp. 84–88, 1981.

[27] K. Hoffman, D. Zage, and C. Nita-Rotaru, "A survey of attack and defense techniques for reputation systems," *ACM Comput. Surv.*, vol. 42, no. 1, 2009.

[28] S. Zhong, E. L. Li, Y. G. Liu, and Y. R. Yang, "On designing incentive-compatible routing and forwarding protocols in wireless ad-hoc networks: an integrated approach using game theoretical and cryptographic techniques," in *MobiCom*, 2005, pp. 117–131.

[29] R. Lu, X. Lin, H. Zhu, X. Shen, and B. Preiss, "Pi: A practical incentive protocol for delay tolerant networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 4, pp. 1483–1493, 2010.

[30] T. Ketelaar and W. T. Au, "The effects of feelings of guilt on the behaviour of uncooperative individuals in repeated social bargaining games: An effect-as-information interpretation of the role of emotion in social interaction," *Cognition and Emotion*, vol. 17, no. 3, pp. 429–453, 2003.

[31] A. Colman, "Cooperation, psychological game theory, and limitations of rationality in social interaction," *Behavioral and Brain Sciences*, vol. 26, no. 02, pp. 139–153, 2003.

[32] M. Wubben, *Social Functions of Emotions in Social Dilemmas*. Rotterdam, 2010.

[33] H. Luan, L. Cai, J. Chen, X. Shen, and F. Bai, "Vtube: Towards the media rich city life with autonomous vehicular content distribution," in *SECON*, 2011, pp. 359 – 367.

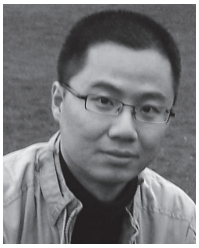
[34] F. Fukuyama, *Trust: Social Virtues and the Creation of Prosperity*. NY: Free Press, 1995.

[35] X. Boyen and B. Waters, "Full-domain subgroup hiding and constant-size group signatures," in *Public Key Cryptography (PKC)*, 2007, pp. 1–15.

[36] X. Liang, Z. Cao, J. Shao, and H. Lin, "Short group signature without random oracles," in *International Conference on Information and Communications Security (ICICS)*, 2007, pp. 69–82.

[37] X. Li, N. Mitton, and D. Simplot-Ryl, "Mobility prediction based neighborhood discovery for mobile ad hoc networks," in *IFIP International Conference on Networking (NETWORKING)*, 2011, pp. 138–151.

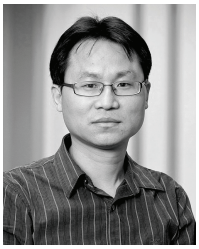
Xiaohui Liang currently working toward a Ph.D. degree with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. He is currently a research assistant with the Broadband Communications Research (BBCR) Group, University of Waterloo. His research interests include security and privacy for e-healthcare system and mobile social networks.



Xu Li is a research scientist at Inria, France. Prior to joining Inria, he worked as postdoc fellow at several locations: the University of Waterloo, Canada; Inria/CNRS, France; and the University of Ottawa, Canada. He received a PhD (2008) degree from Carleton University, Canada, an MSc (2005) degree from the University of Ottawa, Canada, and a BSc (1998) degree from Jilin University, China, all in computer science. During 2004.1-8, he held a visiting researcher position at National Research Council Canada (NRC). His research interests are in the areas of machine-to-machine communications and mobile social networks, with over 50 different works published in refereed journals, conference proceedings and books. He is on the editorial boards of the Wiley Transactions on Emerging Telecommunications Technologies, Ad Hoc & Sensor Wireless Networks, and Parallel and Distributed computing and Networks. He is/was a guest editor of the IEEE Transactions on Parallel and Distributed Systems, Mobile Networks and Applications, Peer-to-Peer Networking and Applications, Journal of Communications, Computer Communications, and Ah Hoc & Sensor Wireless Networks. He was a recipient of NSERC PDF awards and a number of other awards.

Tom H. Luan received the B.E. degree in Xi'an Jiaotong University, China in 2004 and the M.Phil. degree in Electronic and Computer Engineering from the Hong Kong University of Science and Technology, Kowloon, Hong Kong in 2007. He is now pursuing the Ph.D. degree at the University of Waterloo, ON, Canada. His current research interests focus on wired and wireless multimedia streaming and content distribution in vehicular networks.

Rongxing Lu received the Ph.D. degree in computer science from Shanghai Jiao Tong University, Shanghai, China in 2006 and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2008. He is currently a Postdoctoral Fellow with the Broadband Communications Research (BBCR) Group, University of Waterloo. His research interests include wireless network security, applied cryptography, and trusted computing.



Xiaodong Lin received the Ph.D. degree in information engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 1998 and the Ph.D. degree (with Outstanding Achievement in Graduate Studies Award) in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2008. He is currently an assistant professor of information security with the Faculty of Business and Information Technology, University of Ontario Institute of Technology, Oshawa, ON, Canada. His research interests include wireless network security, applied cryptography, computer forensics, software security, and wireless networking and mobile computing. Dr. Lin was the recipient of a Natural Sciences and Engineering Research Council of Canada (NSERC) Canada Graduate Scholarships (CGS) Doctoral and the Best Paper Awards of the 18th International Conference on Computer Communications and Networks (ICCCN 2009), the 5th International Conference on Body Area Networks (BodyNets 2010), and IEEE International Conference on Communications (ICC 2007).

Xuemin (Sherman) Shen received the BSc degree from Dalian Maritime University, China, in 1982 and the MSc and PhD degrees from Rutgers University, New Jersey, in 1987 and 1990, all in electrical engineering. He is a professor and a university research chair in the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research focuses on resource management in interconnected wireless/wired networks, UWB wireless communications networks, wireless network security, wireless body area networks, and vehicular ad hoc and sensor networks. He is a co-author of three books and has published more than 400 papers and book chapters in wireless communications and networks, control, and filtering. He is the Editor-in-Chief of IEEE Network, and will serve as a Technical Program Committee Co-Chair for 2014 IEEE Infocom. He is the Chair for IEEE ComSoc Technical Committee on Wireless Communications, and P2P Communications and Networking, and a voting member of GITC. He was a Founding Area Editor for IEEE Transactions on Wireless Communications, and a Guest Editor for IEEE JSAC, IEEE Wireless Communications, and IEEE Communications Magazine. He also served as the Technical Program Committee Chair for Globecom'07, the Tutorial Chair for ICC'08, and the Symposia Chair for ICC'10. Dr. Shen received the Excellent Graduate Supervision Award in 2006, and the Outstanding Performance Award in 2004, 2007 and 2010 from the University of Waterloo, the Premier's Research Excellence Award (PREA) in 2003 from the Province of Ontario, Canada. He is a registered Professional Engineer of Ontario, Canada, an IEEE Fellow, a Fellow of the Engineering Institute of Canada, a Fellow of Canadian Academy of Engineering and was a ComSoc Distinguished Lecturer.