Privacy-Aware Trust-Based Recruitment in Social Participatory Sensing

Haleh Amintoosi(✉) and Salil S. Kanhere

The University of New South Wales, Sydney, Australia
{haleha,salilk}@cse.unsw.edu.au

Abstract. The main idea behind social participatory sensing is to leverage social networks as the underlying infrastructure for recruiting social friends to participate in a sensing campaign. Such recruitment requires the transmission of messages (i.e., tasks and contributions) between the requester and participants via routes consisting of social links. When selecting the routes, the recruitment scheme should consider two fundamental factors. The first factor is the level of trustworthiness of a route, which evaluates its reliability to ensure that the integrity of the message is preserved. The second factor is the privacy level of the route, which measures information leakage in the form of disclosure of private information contained in the message by intermediate nodes. The best route will be the route with maximum credibility, i.e., highest trust score and lowest likelihood of privacy breach. In this paper, we propose a privacy-preserving trust-based recruitment framework which is aimed at finding the best route from the requester to the selected participants. We propose to quantify the privacy score of a route by utilising the concept of entropy to measure the level of privacy breach in each intermediate node along the route. The trust score of the route is obtained by multiplying the mutual trust rates of all links along the route. Simulation results demonstrate the efficacy of our framework in terms of recruiting suitable participants through the most secure and trustable routes.

Keywords: Privacy · Trust · Social networks · Participatory sensing

1 Introduction

The widespread prevalence of sensor-rich smartphones has propelled the emergence of a novel sensing paradigm, known as *participatory sensing* [1]. In participatory sensing, a sensing task is defined by the *requester*, and ordinary citizens, called *participants*, volunteer to contribute by using their mobile phones. A plethora of applications have been proposed, ranging from personal health [2,3] and sharing prices of costumer goods [4] to air pollution monitoring [5].

The success of a typical participatory sensing application depends on overcoming several challenges: (i) evaluating the *trustworthiness* of contributions in an effort to weed out low fidelity/quality data (ii) recruiting sufficient number of

I. Stojmenovic et al. (Eds.): MOBIQUITOUS 2013, LNICST 131, pp. 262–275, 2014.
DOI: 10.1007/978-3-319-11569-6_21

well-suited participants and (iii) providing a secure and *privacy-aware* environment to contribute. To address the issue of participant sufficiency, one proposed idea is to leverage online social networks as the underlying infrastructure for participatory sensing applications [6]. The resulting paradigm, known as *social participatory sensing*, enables the usage of friendship relations for the identification and recruitment of participants.

In order to address the challenge of assessing contribution trustworthiness, we proposed a trust framework for social participatory sensing system [7], which attempts to recruit social friends as participants. The trust server separately evaluates the quality of contribution and the trustworthiness of participant, which are then combined via a fuzzy inference engine to arrive at a trustworthiness score for the contribution. To allow for better selection of well-suited participants, we extended this framework in [8] such that a reputation score is calculated for each participant by using the PageRank algorithm. In our most recent work [9], we proposed a trust-based recruitment framework to address the challenge of recruiting sufficient well-suited participants by leveraging multi-hop friendship relations.

The central focus of above discussed research is on addressing the first two aforementioned challenges, which are related to the fundamental question of trust without much regard to the other equally important issue of privacy. In other words, we assumed that all participants are trustworthy and do not attempt to infer others' private information. This assumption, however, is not realistic. There are always people who are curious about others' private information and may try to access sensitive information such as a person's whereabouts, by eavesdropping the private conversations. Without sufficient assurance about safeguarding their private information, it is likely that participants may be dissuaded from contributing in participatory sensing tasks. It is thus logical that we address these privacy issues, which is precisely the focus of this paper.

In order to preserve the privacy of participant's information which is embedded in exchanged messages, it is desirable to consider potential privacy breaches in selecting the route between the requester and participant. In other words, when multiple routes exist, a reasonable approach is to select the most secure and trustable route in a way that the likelihood of a privacy breach in intermediate nodes is minimal. The information may itself be a sensitive attribute such as participant's address or his telephone number, or it may be a combination of quasi-identifying attributes which would readily allow a malicious intermediary to infer the corresponding sensitive information. For example, according to a famous study [10] of the 1990 census data, 87 % of the US population can be uniquely identified by gender, ZIP code and full date of birth. Access to such private information thus naturally results in leakage of user privacy. So, the need for protecting such personal information is eminent.

A trivial solution to preserve privacy is encryption (e.g., HTTPS) which can be used to secure communication channels and protect against eavesdropping. However, this facility is not widely used by most online social networks. Only 3 of the top 5 online social networking services currently use HTTPS. Moreover,

they only make use of this security measure to protect login credentials. The rest of the communication happens unencrypted and is visible to everyone along the communication path [11]. The primary reason for not using HTTPS for all communication is to minimize the hardware and connectivity costs. Moreover, public key cryptography needs additional computations and components for key management, which makes it computationally expensive for multi-hop social networks with extremely large number of nodes. To sum up, encryption-based methods are most likely too complicated or expensive for general adoption.

In this paper, we address the challenge of privacy in social participatory sensing systems by proposing a privacy-preserving recruitment framework. In fact, we aim at removing the barriers against active participation by assuring the users about the privacy preservation of their sensitive information along the communication routes. In particular, as an integration with our previous work [9], for each potential route between the requester and participant, we quantify the *credibility* of the route by quantifying and combining the trust and privacy scores of the route. The route with maximum credibility is selected for message exchange. To quantify the privacy score, we use *entropy* to quantify the privacy leak of sensitive information at each intermediate node. To quantify the trust score, we assume that the friendship links are weighted by mutual trust ratings, which is a dynamic value and is continuously updated by parameters such as the trustworthiness of the provided contributions. The trust score of the route is obtained by multiplying the mutual trust rates of all links along the route. A credibility score is then computed by combining the trust and privacy scores, which is then used to select the best routes between the requester and participant. We investigate two different methods for combining these scores to arrive at a final credibility score: the first one is *geometric mean* which simply is the root square of the multiplication of these two parameters, and the second one is *fuzzy combination*, which leverages fuzzy inference engine to address different states of such combination.

The rest of the paper is organised as follows. Related work is discussed in Sect. 2. We present the details of our architecture in Sect. 3. The routing algorithm is described in Sect. 4. Simulation results are discussed in Sect. 5. Finally, Sect. 6 concludes the paper.

2 Related Work

To the best of our knowledge, the issue of privacy in social participatory sensing has not been addressed in prior work. Moreover, prior work on privacy in participatory sensing such as [12,13] address orthogonal issues which do not cover the social aspects and relations that are inherent in social participatory sensing. As such, we discuss related research focusing on social based routing in literature and then specifically look at the privacy-preserving routing algorithms.

Social-Based Routing. Recent studies have focused on mobile social networks and analyzed the social network properties of these networks to assist the design of efficient routing algorithms. In [14], the social similarity (to detect nodes

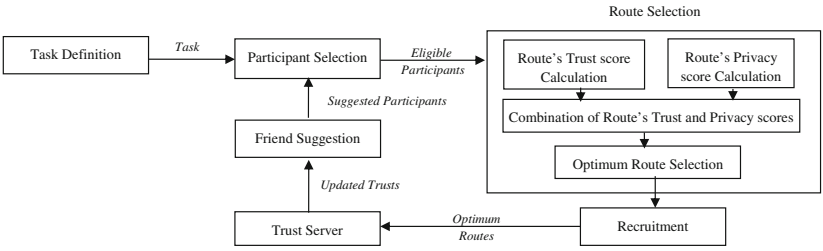


Fig. 1. Framework architecture

that are part of the same community) and ego-centric betweenness (to identify bridging nodes between different communities) has been used as two metrics to increase the performance of routing. When two nodes encounter each other, they calculate the joint utility function comprised of these two metrics for each of their destinations. The message is given to the node having higher utility for the message's destination. In [15], each node is assumed to have a global ranking which denotes the popularity of the node in the entire society, and a local ranking which denotes its popularity within its own community. Messages are forwarded to nodes having higher global ranking until a node in the destination's community is found. Then, messages are forwarded to nodes having higher local ranking within destination's community. Mashhadi et al. [16] combines social network and location information to build an information-dissemination system for mobile devices. In [17], instead of determining social network information from encounter patterns, they use pre-existing social ties.

Privacy-Preserving Routing. In [18], the authors propose a method that leverages a key management system, where users are divided into communities and public-key cryptography is used to secure communication within a community. A major drawback is that community members can observe the routing tables of all other members, resulting in eavesdropping attacks. In the context of opportunistic networks, Shikfa et al. [19] proposed an approach that uses group-based cryptography, using multiple levels of cryptography to prevent data from being accessed by different groups. They concentrate on protecting the application-layer data payload, rather than the routing information. In [20], the authors present methods to improve anonymity within an ad hoc network by compressing and obscuring a packet's routing list. Another popular mechanism is onion routing [21], where packets are routed through a group of collaborating nodes, thus making it difficult to determine the source of a communication. Onion routing, however, does not prevent eavesdropping attacks from intermediate nodes. Moreover, it requires a Public Key Infrastructure (PKI).

In this paper, we aim to propose a privacy-preserving routing algorithm that does not require encryption or PKI. As mentioned in Sect. 1, encryption is not widely supported in online social networks, and is computationally expensive for multi-hop social networks with a large number of members. So it is reasonable and also cost effective to preserve the privacy without relying on encryption.

3 Framework Architecture

Figure 1 illustrates the architecture of the proposed recruitment framework. The social network serves as the underlying publish-subscribe substrate for recruiting friends and friends of friends (FoFs) as participants. We abstract an online social network (e.g., Facebook) as a weighted directed graph, in which, social network members serve as graph nodes and friendship relations denote the edges of graph, with weights equal to the trust rating between users. A person wishing to start a participatory sensing campaign acts as a requester and defines the task, including the specification of task's main requirements such as needed expertise or location. Then, participant selection component crawls the social graph up to L levels (friends and FoFs) to determine eligible participants who can fulfil the task's requirements. The selection is performed by comparing the participants' profile information with the task's requirements.

Next, route selection component traverses the social graph to find the best route from requester to each of eligible participants. We claim that the best route will be the route with highest trustworthiness and highest privacy preservation. To do so, a credibility score, which is the combination of the trust score and privacy score of the route, is calculated (we investigate two different methods, geometric mean and fuzzy, for such combination). In case of multiple routes between the requester and an eligible participant, the route with the highest credibility is chosen. Those eligible participants for whom, the credibility of the route is greater than a predefined threshold are considered as *selected* participants. Once selected, the task is routed along the specified routes and delivered to the selected participants. The same path in reverse is used to transmit contributions back. The trust server then evaluates an objective trust rating for each contribution (which we call it the Trustworthiness of Contribution (ToC)). Based on ToC, mutual trusts between members along the route are updated (more details about the trust server functionality can be found in [7,8]).

Periodically (where a period typically spans a certain number of campaigns), the suggestion component builds a recommended participant group for each requester, containing a list well behaved participants, which is further used for recruitment or friendship establishment. It should be noted that all the computations including the evaluation of the trust and privacy scores is done inside the trust server.

4 Privacy-Preserving Trust-Based Route Selection

Once eligible participants are selected, route selection component finds the best routes to them. To do so, the trust and privacy scores of the route are considered. In the following, we elaborate how these scores are computed.

4.1 Trust Score of the Route

As mentioned in Sect. 3, we assume that the social links are labelled with weights equal to the trust ratings between two parties. If intermediate nodes exist,

the trust score of the route is a combination of trust ratings of each pair nodes along the route. We leverage multiplication for the combination since it has been shown in [22] to be an effective strategy for trust propagation. In other words, we assume the set R is the set of all possible routes between the requester and a specific participant. The route R_i ($R_i \in R$) has been defined with (N_{R_i}, E_{R_i}) in which, N_{R_i} is the set of nodes within this route and E_{R_i} is the set of edges of R_i. In that case, the trust score of R_i, denoted by $Trust(R_i)$ is calculated as:

$$Trust(R_i) = \prod_{k=1}^{l} w(e_k), e_k \in E_{R_i} \tag{1}$$

where l is the length of the R_i and $w(e_k)$ is the weight of the edge e_k. $Trust(R_i)$ is in the range of $[0, 1]$.

4.2 Privacy Score of the Route

When discussing about privacy in social networks, it is important to specify what defines failure to preserve privacy. A privacy breach occurs when a piece of sensitive information about an individual is disclosed to an adversary. Traditionally, two types of privacy breaches have been studied: *identity disclosure* and *attribute disclosure*. Identity disclosure occurs when an adversary is able to map a profile in the social network to a specific real-world entity. Attribute disclosure, on the other hand, occurs when an adversary is able to determine the value of a sensitive user attribute, one that the user intended to stay private. There are three sets of personal attributes [23]:

1- Identifying attributes: attributes such as social security number (SSN) which uniquely identify a person. To avoid identity disclosure, identifying attributes should be removed from profiles.
2- Quasi-identifying attributes: a combination of attributes which can identify a person uniquely. It has been observed that 87 % of individuals in the U.S. can be uniquely identified based on their date of birth, gender and zip code [10].
3- Sensitive attributes: those that users tend to keep hidden from the public, such as politic view, location, and sexual orientation.

The messages exchanged between the requester and participant (including task or contribution) may contain private information such as sensitive attributes and quasi-identifiers, which may leak in intermediate nodes. To prevent such privacy leakage, it is reasonable to select the routes which contain intermediate nodes that are least likely to cause privacy breaches.

In order to quantify the privacy leak, we leverage the concept of entropy. Entropy is a measure of the uncertainty in a random variable [24]. In fact, our model aims to maximize the entropy which means the maximization of the unpredictability of information for an adversary node. Higher entropy means better privacy for the information contained inside a message. Since identifying attributes such as SSN are not normally kept in profiles, we assume that privacy leakage may happen if two types of information are leaked: sensitive attributes

and quasi-identifiers. With this assumption, we aim at calculating the amount of uncertainty of a node about private information inside a message.

For the intermediate node $m \in N_{R_i}$, we have the following definitions:

Let $S = \{s_1, s_2, s_3, ..., s_k\}$ is the set of sensitive attributes and $Q = \{q_1, q_2, q_3, ..., q_t\}$ is the set of quasi-identifying attributes that may exist in a user's message. We aim at quantifying the privacy leakage of message in the intermediate node m who wishes to know Q in order to uniquely identify the user. So, we first measure the "initial uncertainty" of m about Q. The initial uncertainty, denoted by $H(Q)$ is generally defined using Shannon entropy [24]:

$$H(Q) = -\sum_{i=1}^{t} p(q_i) log(p(q_i)) \tag{2}$$

where $p(q_i)$ is the probability mass function.

Next, we assume the situation where m knows the value of some sensitive values of S from the user. In this case, the uncertainty of m about Q after knowing some information from S, denoted by the "remaining uncertainty" is measured using conditional Shannon entropy:

$$H(Q \mid S) = \sum_{i=1}^{t} p(q_i) H(Q \mid q_i \in S) \tag{3}$$

It is natural to measure the *privacy leakage* ($\mathcal{L}$) of a message by comparing the m's uncertainty about Q, before and after knowing the value from S. So, we use the concept of *mutual information* as: leakage = initial uncertainty − remaining uncertainty [25]. In other words, the privacy leakage ($\mathcal{L}$) of a message is:

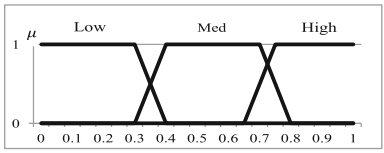
$$\mathcal{L} = H(Q) - H(Q \mid S) \tag{4}$$

In general, if n messages have passed through node m, then the amount of privacy leakage in node m, denoted by $\mathcal{L}_m$, is calculated as: $\mathcal{L}_m = \frac{1}{n} \sum_{j=1}^{n} \mathcal{L}_{m,j}$ in which, $\mathcal{L}_{m,j}$ is the privacy leakage of message j in intermediate node m. As this equation shows, $\mathcal{L}_m$ keeps a history of privacy leakage upon each message originator. It is obvious that the privacy score of node m, ($Privacy(m)$), is inversely related to the privacy leakage ($\mathcal{L}_m$) in this node. However, in order to have the privacy score value in the range of $[0, 1]$, we divide the value of $\mathcal{L}_m$ by $log(n)$. The reason is that $H(Q)$ has a value less than $log(n)$. So, the maximum value for $\mathcal{L}_{m,n}$ and $\mathcal{L}_m$ will also be $log(n)$, which results in $\mathcal{L}_m / log(n)$ in the range of $[0, 1]$. To summarize:

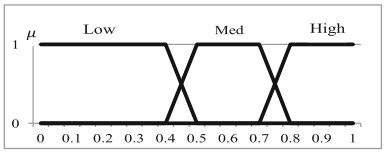
$$Privacy(m) = 1 - \frac{\mathcal{L}_m}{log(n)} \tag{5}$$

in which, n are number of messages that have passed through intermediate node m. So, for each route R_i consisting of a set of nodes, the privacy of the route is:

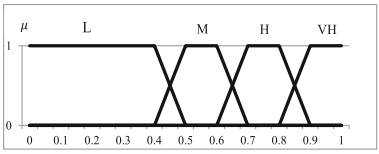
$$Privacy(R_i) = min\big(Privacy(m)\big) \text{ where } m \in N_{R_i} \tag{6}$$



(a) Membership function for trust



(b) Membership function for privacy



(c) Membership function for credibility

Rule no.	trust score	privacy score	credibility
1	Low	Low	L
2	Low	Med	L
3	Low	High	L
4	Med	Low	M
5	Med	Med	H
6	Med	High	H
7	High	Low	M
8	High	Med	VH
9	High	High	VH

(d) Fuzzy rule base

Fig. 2. Rule base and Membership functions of input and output linguistic variables

4.3 Credibility Score of the Route

In order to select the best route (R_{best}) among the set of routes, we combine the privacy score of each route with its trust score via a combination function F to reach to a single value for the credibility score of the route. In other words,

$$Credibility(R_{best}) = max\Big(F(Trust(R_i), Privacy(R_i))\Big) \text{ where } R_i \in R \quad (7)$$

The selection of a proper combination function F is important. F should be efficient enough to handle possible conflicts between the trust score and the privacy score in a reasonable manner. The decision on how to combine these factors affects the performance of the route selection module. In this paper, we have considered two combination functions:

Geometric Mean. The geometric mean is defined as the n_{th} root (n:count of numbers) of the product of the numbers. It is often used for comparing different items and finding a single "figure of merit" for these items, when each item has multiple properties. A geometric mean, unlike an arithmetic mean, tends to dampen the effect of very high or low values, which might bias the mean if a straight average (arithmetic mean) were calculated. This property makes geometric mean suitable for our situation since there may be situations where the trust score is high but privacy score is low (or vice versa). The combination of trust and privacy scores of the route via the geometric mean is:

$$Credibility(R_i) = \sqrt{Trust(R_i) * Privacy(R_i)} \quad (8)$$

Fuzzy Combination. Another possible option is to employ fuzzy logic to calculate a comprehensive credibility score for the route. Consider a situation where

the trust score is high but the privacy score is low. The use of fuzzy logic allows us to achieve a meaningful balance between these two scores. The inputs to the fuzzy inference system are the crisp values of the trust and privacy scores of a route. In the following, we describe the fuzzy inference system components.

Fuzzifier: The fuzzifier converts the crisp values of input parameters into a linguistic variable according to membership functions. The fuzzy sets for the input and output variables are defined as: T(trust score) = T(privacy score) = {Low, Med, High}, T(route's credibility) = {L, M, H, VH}. The membership function which represents a fuzzy set A is usually denoted by μ_A. The membership degree $\mu_A(x)$ quantifies the grade of membership of the element x to the fuzzy set A. Figure 2(a)–(c) represents the membership functions of parameters.

Inference Engine: The role of inference engine is to convert fuzzy inputs to the fuzzy output (route's credibility) by leveraging If-Then type fuzzy rules. The combination of the above mentioned fuzzy sets create 3*3 = 9 different states, addressed by 9 rules as shown in Fig. 2(d). The rule based design is based on the experience on how the system should work by leveraging *max-min* composition method. The result is the credibility score which is a linguistic fuzzy value.

Defuzzifier: A defuzzifier converts the credibility fuzzy value to a crisp value in the range of $[0, 1]$. We employed the Centre of Gravity (COG) method.

In Sect. 5, the effect of each combination function on the performance of route selection module will be investigated.

5 Experimental Evaluation

This section presents simulation-based evaluation of the proposed system. The simulation setup is outlined in Sect. 5.1 and the results are in Sect. 5.2.

5.1 Simulation Setup

To undertake the preliminary evaluations outlined herein, we chose to conduct simulations, since real experiments in social participatory sensing are difficult to organise. We developed a custom Java simulator for this purpose.

The data set that we use for our experiment is the real web of trust of Advogato.org [26]. Advogato.org is a web-based community of open source software developers in which, site members rate each other in terms of their trustworthiness. Trust values are one of the three choices master, journeyer and apprentice, with master being the highest level in that order. The result of these ratings is a rich web of trust, which comprises of 14,019 users and 47,347 trust ratings. In order to conform it to our framework, we map the textual ratings to the range of $[0, 1]$ as master = 0.8, journeyer = 0.6, and apprentice = 0.4.

Whenever a task is launched, one of the Advogato users is selected to be the requester. The participant selection component traverses the Advogato graph beginning from the requester until level L ($L = 3$) to find suitable participants (i.e., those whose profile information such as their expertise match the task's

requirements). Next, the route selection component finds the best routes (details in Sect. 4.3). Tasks and contributions are then exchanged and trust server calculates the Trustworthiness of Contribution (ToC) for each receiving contribution. Trust ratings along the routes are then updated based on the ToC achieved. If above a threshold, all mutual trusts along the route are increased; otherwise, if less than a threshold, mutual trust is decreased (details have been presented in [8]). We run the simulation for 20 periods, each consisting of launching 30 tasks.

The amount of private information contained in the exchanged messages (tasks and contributions) may vary. Some messages may contain more sensitive information than others. To simulate these differences, we assume that the total number of attributes contained in a message are 6 and a message may belong to one of three privacy classes: (i) *Class* 0: messages in this class contain 4 sensitive attributes; (ii) *Class* 1: messages in this class contain 2 sensitive attributes; (iii) *Class* 2: messages in this class contain 1 sensitive attribute. Greater number of sensitive attributes implies more private information. Whenever a message is created, a set of sensitive attributes, defined by numbers in the range of $[1, 6]$, is randomly assigned to it. The credibility of the route is then computed via Eq. 7. The route with highest credibility score is chosen for message exchange.

In order to observe the performance of the system in the presence of noise, we artificially create situations in which, the privacy score of a specific node reduces for a period of time. This may happen in reality when a participant starts to reveal private information about another user. Our goal is to observe whether the system is able to rapidly detect such behavioural change and demonstrate a reasonable reaction accordingly. The duration of behavioural change has been set to be between the 5th and 10th periods. We investigate the average credibility score of all routes passing through this malicious node. We expect to see that the proposed method is able to rapidly reduce the credibility score of these specific routes and thus, eliminate them from being selected for message exchange.

As mentioned in Sect. 2, since there is no related work in the area of social participatory sensing, we compare the performance of our method against the one described in our previous work [9]. To be more specific, we compare the following: (1) trust based recruitment, in which, the route selection is based only on the trust score of the route (the method in [9]). (2) privacy-trust recruitment, our proposed method in which, the best route is selected based on both privacy score and trust score.

As mentioned in Sect. 3, a ToC rating is calculated for each contribution in the trust server. We consider the overall trust as the evaluation metric. The overall trust of a campaign is defined as $OverallTrust = \frac{\sum_{i=1}^{n} ToC}{n}$ in which, n is the number of non-revoked contributions. Greater overall trust demonstrates better ability to achieve highly trustable contributions and revoke untrusted ones. Overall trust has a value in the range of $[0, 1]$. The overall trust values obtained for all tasks will be averaged to make a single value as the *average overall trust* for the entire simulation.

As explained in Sect. 4, the participant selection component determines a set of eligible participants. A subset of eligible participants, known as selected participants, are those for whom, the credibility score of the route is greater than a predefined credibility threshold (set to 0.6). Relevant to this selection process, we define the notion of *participation score* as the ratio of selected participants to eligible participants. A higher value of participation score implies better ability to recruit suitable participants via optimum routes. Participation score has a value in the range of $[0, 1]$.

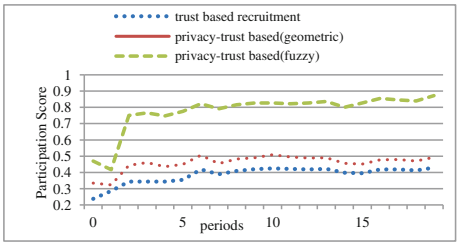


Fig. 3. Evolution of participation score

5.2 Simulation Results

Figure 3 demonstrates the evolution of participation score for trust based and privacy-trust based recruitment methods with both geometric and fuzzy combination (details about combination functions can be found in Sect. 4.3). As this figure shows, our proposed recruitment framework achieves a higher score in comparison with the trust based method, thus implying better performance in terms of recruiting more participants. Note that participant selection process in both methods is limited to L levels (L = 3). Since our proposed method takes participant's privacy into account, it results in the selection of a diverse set of routes which in turn, allows selecting well-suited participants from a broader group. In other words, our proposed scheme achieves greater diversity than the scheme that purely relies on trust.

Figure 4 illustrates the average privacy score of the routes selected for message exchange between a specific requester and a set of selected participants, calculated separately for different privacy classes. As this chart demonstrates, our proposed method achieves higher privacy score for all types of messages originated from a specific requester. For instance, the average privacy score obtained in our proposed method for class 0 messages is 12 % higher than the one obtained in the trust-based method. This is because our scheme considers the privacy score of the route as an effective issue in evaluating the route's credibility. Note that this improved performance is consistent for all privacy classes (and more explicit for class 0), since the route selection method is identical for all types of messages. This implies that our proposed method is able to achieve higher privacy scores for all types of exchanged messages containing different levels of private information. The importance of this outperformance is better understood when obtained average overall trust is also considered in conjunction (details in Sect. 5.1). The average overall trust is 0.89 for trust-based recruitment method. Our scheme achieves a score of 0.89 with geometric mean and 0.84 with the fuzzy combination. This comparison shows that although our proposed method does not consider the trust score as the only determinant factor for the route

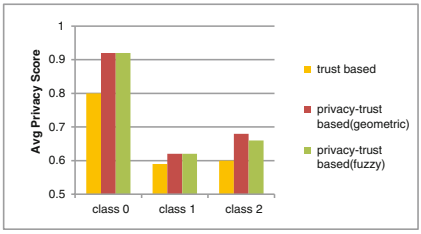


Fig. 4. Average privacy score of the selected routes for different privacy classes

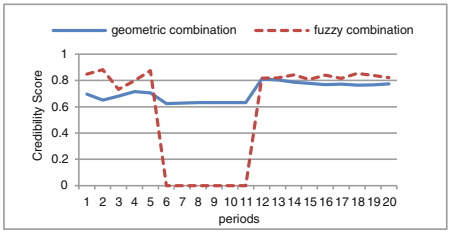


Fig. 5. Evolution of average route's credibility score passing from a malicious node

selection, the achieved overall trust is still similar to trust based recruitment method, while better preserving the privacy of sensitive information.

As mentioned in Sect. 5.1, we also consider a scenario where the privacy score of an intermediate node decreases for a certain time interval (between 5th and 10th periods). The aim is to investigate the sensitivity of our recruitment method in promptly detecting and reacting to such fluctuation. Figure 5 shows the evolution of average credibility score for all routes passing through this malicious node. Observe that the credibility score for both combination methods decreases in the transition period. However, fuzzy method demonstrates better performance in early detection and severe punishment by a sudden decrease (to zero) in the credibility score. This is due to the adjustment of fuzzy rules such as rule no.1, 2 in Fig. 2(d). In other words, according to Eq. 6, low privacy score for a malicious node results in the low privacy score for all routes passing through it. We set the fuzzy rules in a way that when the privacy score of the route is low, the resulting credibility score will be L (Low). This will result in the exclusion of this route from the set of candidate routes. This is not always true for geometric mean. There may be cases in geometric mean combination (as observed in Fig. 5) where the privacy score is low, but its credibility score is above the threshold, resulting in inclusion of this route for message exchange.

6 Conclusions

In this paper, we proposed a privacy-preserving trust-based recruitment framework for social participatory sensing system. Our system leverages friendship relations to recruit participants via the optimum routes with highest level of trustworthiness and privacy preservation. Simulations demonstrated that our scheme preserves better privacy for participants while achieving acceptable overall trust as compared to the trust-based method, and provides the system with better recruitment of participants.

References

1. Burke, J., et al.: Participatory sensing. In: WSW Workshop, ACM SenSys'06 (2006)
2. Reddy, S., et al.: Image browsing, processing, and clustering for participatory sensing: lessons from a DietSense prototype. In: ACM EmNets'07, pp. 13–17 (2007)
3. Stuntebeck, E.P., et al.: HealthSense: classification of health-related sensor data through user-assisted machine learning. In: HotMobile'08, pp. 1–5 (2008)
4. Dong, Y.F., Kanhere, S.S., Chou, C.T., Bulusu, N.: Automatic collection of fuel prices from a network of mobile cameras. In: Nikoletseas, S.E., Chlebus, B.S., Johnson, D.B., Krishnamachari, B. (eds.) DCOSS 2008. LNCS, vol. 5067, pp. 140–156. Springer, Heidelberg (2008)
5. Rana, R.K., Chou, C.T., Kanhere, S.S., Bulusu, N., Hu, W.: Ear-phone: an end-to-end participatory urban noise mapping. In: ACM/IEEE IPSN'10 (2010)
6. Krontiris, I., Freiling, F.: Urban sensing through social networks: the tension between participation and privacy. In: ITWDC'10 (2010)
7. Amintoosi, H., Kanhere, S.S.: A trust framework for social participatory sensing systems. In: Zheng, K., Li, M., Jiang, H. (eds.) MobiQuitous 2012. LNICST, vol. 120, pp. 237–249. Springer, Heidelberg (2013)
8. Amintoosi, H., Kanhere, S.S.: A reputation framework for social participatory sensing systems. J. Mobile Netw. Appl. **19**(1), 88–100 (2014)
9. Amintoosi, H., Kanhere, S.S.: A trust based recruitment framework for social participatory sensing. In: IEEE DCOSS'13 (2013)
10. Sweeney, L.: Achieving k-anonymity privacy protection using generalization and suppression. Int. J. Uncertain. Fuzz. Knowl.-Based Syst. **10**(05), 571–588 (2002)
11. Huber, M., Mulazzani, M., Weippl, E.: Who on earth is "Mr. Cypher": automated friend injection attacks on social networking sites. In: Rannenberg, K., Varadharajan, V., Weber, C. (eds.) SEC 2010. IFIP AICT, vol. 330, pp. 80–89. Springer, Heidelberg (2010)
12. Huang, K.L., Kanhere, S.S., Hu, W.: A privacy-preserving reputation system for participatory sensing. In: IEEE LCN'12 (2012)
13. Christin, D., Reinhardt, A., Kanhere, S.S., Hollick, M.: A survey on privacy in mobile participatory sensing applications. J. Syst. Softw. **84**, 1928–1946 (2011)
14. Daly, E., Haahr, M.: Social network analysis for routing in disconnected delay-tolerant MANETs. In: ACM MobiHoc' 07 (2007)
15. Hui, P., Crowcroft, J., Yoneki, E.: BUBBLE Rap: social based forwarding in delay tolerant networks. In: ACM MobiHoc'08 (2008)
16. Mashhadi, A.J., et al.: Habit: leveraging human mobility and social network for efficient content dissemination in delay tolerant networks. In: WoWMoM'09 (2009)
17. Mtibaa, A., et al.: PeopleRank, social opportunistic forwarding. In: INFOCOM'10 (2010)
18. Boldrini, C., et al.: Exploiting users' social relations to forward data in opportunistic networks: the HiBOp solution. Perv. Mobile Comput. **4**(5), 633–657 (2008)
19. Shikfa, A., Önen, M., Molva, R.: Privacy and confidentiality in context-based and epidemic forwarding. Comput. Commun. **33**(13), 1493–1504 (2010)
20. Aad, I., et al.: Packet coding for strong anonymity in ad hoc networks. In: IEEE SecureComm'06 (2006)
21. Goldschlag, D., et al.: Onion routing. Commun. ACM **42**, 39–41 (1999)
22. Hasan, O., et al.: Evaluation of the iterative multiplication strategy for trust propagation in pervasive environments. In: ACM ICPS'09 (2009)

23. Zheleva, E., Getoor, L.: Privacy in social networks: a survey. In: Aggarwal, C.C. (ed.) Social Network Data Analytics, pp. 277–306. Springer, New York (2011)
24. Shannon, C.E.: A mathematical theory of communication. Bell Syst. Tech. J. **27**(3), 379–423 (1948)
25. Smith, G.: Quantifying information flow using min-entropy. In: IEEE QEST'11 (2011)
26. Levien, R., Aiken, A.: Attack-resistant trust metrics for public key certification. In: 7th USENIX Security Symposium (1998)