This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2352253, IEEE Transactions on Mobile Computing

Privacy and Quality Preserving Multimedia Data Aggregation for Participatory Sensing Systems

Fudong Qiu, *Student Member, IEEE,* Fan Wu, *Member, IEEE,* and Guihai Chen, *Senior Member, IEEE*

Abstract—With the popularity of mobile wireless devices equipped with various kinds of sensing abilities, a new service paradigm named participatory sensing has emerged to provide users with brand new life experience. However, the wide application of participatory sensing has its own challenges, among which privacy and multimedia data quality preservations are two critical problems. Unfortunately, none of the existing work has fully solved the problem of privacy and quality preserving participatory sensing with multimedia data. In this paper, we propose *SLICER*, which is the first k-anonymous privacy preserving scheme for participatory sensing with multimedia data. SLICER integrates a data coding technique and message transfer strategies, to achieve strong protection of participants' privacy, while maintaining high data quality. Specifically, we study two kinds of data transfer strategies, namely transfer on meet up (TMU) and minimal cost transfer (MCT). For MCT, we propose two different but complimentary algorithms, including an approximation algorithm and a heuristic algorithm, subject to different strengths of the requirement. Furthermore, we have implemented SLICER and evaluated its performance using publicly released taxi traces. Our evaluation results show that SLICER achieves high data quality, with low computation and communication overhead.

Index Terms—Participatory Sensing, Privacy Preservation, K-Anonymity, Erasure Coding.

1 Introduction

The wide application of mobile communication equipments and the fast advance of sensing technologies have led to the wide availability of privately-held, low-cost, advanced-processing, and big-storage mobile wireless devices, that are equipped with a number of embedded sensors (*e.g.*, microphone, camera, accelerometer, gyroscope, and GPS). On one hand, modern wireless communication technologies (*e.g.*, 2G/3G/4G, Wi-Fi, and Bluetooth) make the communication between mobile devices and infrastructure, as well as between mobile devices themselves, convenient and fast. On the other hand, the mobile devices, especially smart phones, are no longer a tool only for communication, but "computers" with multifunction.

Participatory sensing [1] emerged as a new service paradigm using human-carried mobile devices, such as smart phones, for distributed data collection, exchange, analysis, and sharing. With an estimated number of 6.8 billion mobile-cellular subscriptions worldwide [2], participatory sensing may provide an unprecedented spatial coverage, with very low or even no deployment cost. Compared with traditional decentralized data collection methods (*e.g.*, wireless sensor networks), participatory sensing demonstrates several outstanding advantages, including larger coverage, lower cost, mobile capability, more sufficient energy supply, and more flexible interactive capability. Attracted by the practical and commercial value of participatory sensing, many participatory sensing applications have appeared. For instance, GreenGPS [3] provides the most fuel-efficient routes to drivers; PEIR [4] presents a personal environmental impact report for every individual; PEPSI [5] [6] introduces a privacy enhanced infrastructure for participatory sensing system; ARTSense [7] proposes an anonymous reputation and trust mechanism for participatory sensing; and Ikarus [8] uses sensor data collected during cross-country flights via participatory sensing applications to study thermal effects in the atmosphere, and PoolView [9] gives a privacy preserving architecture for stream data collection. In addition, participatory sensing has been widely used in many practical situations [1], for instance, environment measurement, health care, traffic monitoring, community service, crowdsourcing, and so on.

However, the application of participatory sensing has a number of challenges. One of the major challenges is on privacy preservation [10]–[17]. Sensing record sent to the service provider, is usually attached with spatio-temporal tags indicating the location and time information of the data collected. However, a corrupt service provider may infer private information of the participants, such as identity, home and office addresses, traveling paths, as well as participants' habits and lifestyles, from the sensing records. In turn, many users are reluctant to contribute any sensing record if proper privacy preservation scheme is not applied. Without sufficient number of participants, participatory sensing applications cannot guarantee their quality of services at the expected level. Therefore, designing privacy preserving schemes for participatory sensing is highly important. Another major challenge is on the variety of sensing data. Most of existing applications

- *F. Qiu, F. Wu, and G. Chen are with the Department of Computer Science and Engineering, Shanghai Key Laboratory of Scalable Computing and Systems, Shanghai Jiao Tong University, Shanghai 200240, P.R. China. Emails: fdqiu@sjtu.edu.cn, {fwu, gchen}@cs.sjtu.edu.cn.*
- *F. Wu is the corresponding author.*

- *This work was supported in part by the State Key Development Program for Basic Research of China (973 project 2014CB340303 and 2012CB316201), in part by China NSF grant 61422208, 61472252, 61272443 and 61133006, in part by Shanghai Science and Technology fund 12PJ1404900 and 12ZR1414900, and in part by Program for Changjiang Scholars and Innovative Research Team in University (IRT1158, PCSIRT) China. The opinions, findings, conclusions, and recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies or the government.*

1536-1233 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2352253, IEEE Transactions on Mobile Computing

of participatory sensing only collect small pieces of sensing data (*e.g.*, temperature, velocity, and geographic location). However, more and more newly emerged applications rely on collecting information of surrounding environment in the format of multimedia (*e.g.*, digital image and video) [18], which result in much higher volume of sensing data. Simply applying existing privacy preserving schemes to participatory sensing with multimedia data is not satisfactory, since existing schemes either induce unacceptable amount of communication cost, or degrade the utility/quality of the data badly, in case of multimedia sensing.

In this paper, we present *SLICER*, which is a coding-based k-anonymous privacy preserving scheme, working on application layer, for participatory sensing with multimedia data. Intuitively, k-anonymity means that the service provider cannot identify the contributor of each sensing record from a group of at least k participants. SLICER integrates a data coding technique and message exchanging strategies, to achieve strong protection of participants' privacy, while maintaining high data quality and inducing low communication and computation overhead.

The contributions of this work are listed as follows:

- We propose SLICER for participatory sensing with multimedia data, to achieve both k-anonymous privacy preservation and high data quality, with low communication and computation overhead.
- We design an erasure coding based sensing record coding scheme to encode each sensing record into a number of data slices, each of which can be delivered to the service provider through the other participants or the record's generator herself. When a proper data slice exchanging strategy is applied, the contributor of each particular sensing record is hidden in a group of at least k participants.
- We propose two kinds of strategies for slice transfer. The first and straightforward strategy is named *Transfer on Meet Up* (TMU), which is to transfer a slice upon meeting another participant. The latter delivers the slice to the service provider. The second kind contains two complementary sub-optimal strategies to transfer the slices to a set of participants that might be met within a required period of time, minimizing the total cost while guaranteeing that the sensing record can be delivered to the service provider with guaranteed high probability, which is named *Minimal Cost Transfer* (MCT). The cost difference can be resulted from the wireless communication fee, available bandwidth, battery power, and so on.
- We have implemented SLICER and evaluated its performance using publicly released real traces of taxis [19]. Evaluation results show that SLICER achieves high data quality, with low computation and communication overhead.

The rest of this paper is organized as follows. In section 2, we briefly introduce some technical preliminaries, including the system model, privacy model, and design objectives. In section 3, we describe our coding-based privacy preserving scheme (SLICER), illustrate the basic rationale and detailed design processes, propose the well designed algorithms of slices transfer, and give the necessary analysis and proof of privacy preserving. In section 4, we present the evaluation results. In section 5, we talk about the related work and make some comparison with ours. Finally, we conclude our article and point out our potential directions of future work in section 6.

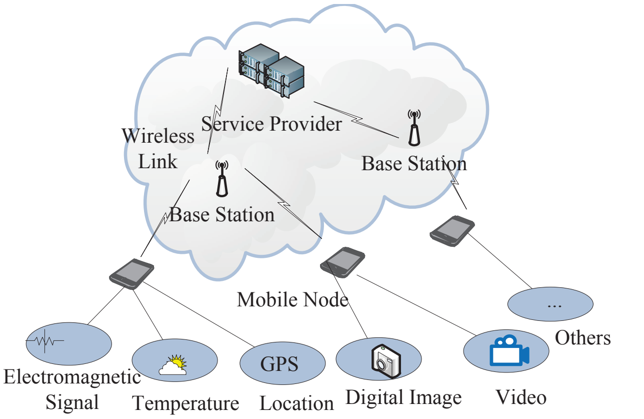


Fig. 1. The Architecture of Cloud-Based Participatory Sensing.

2 Technical Preliminaries

In this section, we present the system model, privacy models, as well as objectives of our design.

2.1 System Model

We consider a cloud-based participatory sensing and service framework as shown in Fig. 1, in which there is a service provider and a number of mobile nodes/participants equipped with different kinds of sensors.

The service provider aggregates, classifies, analyzes, and stores sensing records reported from the participants, and provides query services based on the records. A mobile node/participant is a user carrying a portable and wireless-enabled device (*e.g.*, smart phone, tablet, and laptop). In this paper, we use mobile node and participant interchangeably. Participants can use their sensing devices to collect various kinds of environmental information, such as geographical location, temperature, electromagnetic signal, digital image, video, and so on. In contrast to most of the existing work, which focus on short sensor readings, we consider a participatory sensing system that adapts to multimedia information, such as digital image, audio, and video. We assume that the participants can directly report sensing records through pre-existing communication infrastructure, including GSM, 3G/4G, and Wi-Fi, or indirectly report the records with the help of the other participants.

In this paper, we consider one service provider and a set $N = \{a_1, a_2, \ldots, a_n\}$ of participants. Each participant $a_i \in N$ would like to contribute her sensing records $R_i = \{< t_1, l_1, d_1 >, < t_2, l_2, d_2 >, \ldots\}$ to the service provider, only when her privacy is properly protected. The triple $< t, l, d >$ denotes a sensing record including timestamp, location info, and data info. To facilitate reading, the summary of the notations appeared in this paper is presented in Table 1.

1536-1233 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2352253, IEEE Transactions on Mobile Computing

2.2 Privacy Model

Although participatory sensing provides a new service paradigm, its functionality relies on the contribution of participants. Existing work [1], [11]–[13], [16], [17], [20]–[22] show that contributed information may be misused to reveal the participants' privacy [23]. Most users are not willing to join participatory sensing applications, unless their sensitive information is well protected from both service provider and neighboring participants [12], [24], [25].

In this paper, we consider the problem of privacy preserving in a semi-honest model, in which the adversary correctly follows the protocol specification, but attempts to learn additional information by analyzing the transcript of messages received during the execution [20], [26]–[31]. We classify the attacks in the semi-honest model into two categories: external attack and internal attack. The external attack aims to obtain private information of participants by overhearing the message passing through the wireless communication network. Such attack can be prevented by end-to-end cryptographic schemes. Different from the external attack, designing a scheme to prevent the internal attack is much more challenging. The internal attack may come from two different kinds of entities, including the service provider and the participants.

- Service provider's attack: The service provider has full access to the sensing records reported by the participants. It might infer considerable amount of sensitive information about the participants (*e.g.*, home address, frequently visited places, traveling path, and even the lifestyle), if a proper privacy-preserving scheme is not provided. For instance, the sensor readings collected by a user who drives from home to work might reveal the participant's traveling path as well as her home address. In this work, we focus on protecting users' location/path privacy against the service provider, while assuming that the service provider does not have other background or correlated information about participants. It is also important to consider the privacy protection of the content of multimedia data. However, it is out of the scope of this work. For interested readers, please refer to the previous literatures [32] [33] [34] for privacy processing techniques.
- Participants' attack: Participants may receive some sensing records, when they serve as relays for other participants (*e.g.*, in [35]). Semi-honest participants might position themselves to some critical locations in order to collect sensitive information by pretending to be relays. In this work, we assume that the participants do not collude with the service provider, and there is no collusion among different participants.

2.3 Design Objectives

The design of a privacy preserving scheme should prevent both the external and the internal attacks. Specifically, first, the design needs to prevent external eavesdroppers from obtaining any meaningful information. Second, the design needs to prevent service provider from recognizing the identity of the participant who contributes a particular sensing record, and to prevent the participants from knowing the content of the relayed sensing record. Especially, we require the privacy protection scheme be k-anonymous [36] against the service provider. Here, k-anonymity is reached when the service provider can only identify a particular participant that contributes a sensing record with probability no more than $1/k$.

Symbol	Description
$N = \{a_1, a_2, \ldots, a_n\}$	The participants set
$< t, l, d >$	An original sensing record
$R_i = \{< t_1, l_1, d_1 >, \ldots\}$	The sensing records set
m	Number of encoded slices from one record
k	Minimal number needed to construct record
$EC(\cdot)$	Erasure coding algorithms.
$H(\cdot, \cdot)$	Cryptographic hash function
r_{ij}	Encoded slice
r'_{ij}	Encrypted slice
$ENCRYPT(\cdot, \cdot)$	Asymmetric encryption function
$p(a_j)$	Meeting probability
$c(a_j)$	Cost of a_j for delivering a slice
P	Threshold possibility
x_i	Boolean parameter
$DECRYPT(\cdot, \cdot)$	Asymmetric decryption function
$EC^{-1}(\cdot)$	Decoding function

TABLE 1
Notations

Definition 1 (K-Anonymous Participatory Sensing): A privacy preserving participatory sensing scheme satisfies k-anonymity against the service provider, if for any sensing record reported to the service provider, the service provider cannot distinguish the generator of the record from a group of at least k participants.

Besides the objective on privacy preservation, the design should also satisfy the following requirements:

- The design should maintain high quality of the sensor readings.
- The design should be tolerant of packet/message loss.
- The design can only induce low computation and communication overhead.

3 Coding-based Privacy Preserving Scheme

In this section, we present the design of our coding-based k-anonymous privacy preserving scheme — *SLICER*. We first outline the general idea of SLICER, and then explain the details of each component. Finally, we analyze the privacy preservation properties of SLICER.

3.1 Design Rationale

The main idea of SLICER is to hide the generator of each sensing record among a group of at least k participants, through which all parts of the sensing record are reported to the service provider. Thus, the service provider cannot identify the generator of the original sensing record from at least k participants. We will illustrate the designing challenges and our idea in this section.

(1) Sensing Record Coding

1536-1233 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

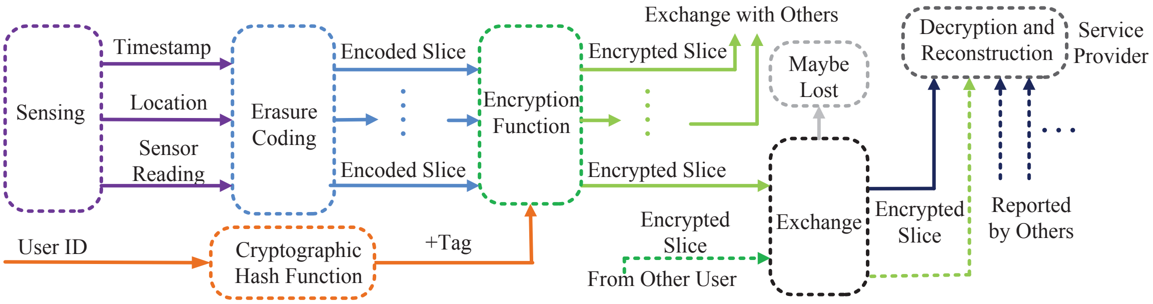


Fig. 2. Work Flow of SLICER

If we simply transfer the (encrypted) sensing record to k participants, then the communication overhead is k times the size of the sensing record, which is unacceptable especially when the sensing record contains multimedia data. Therefore, we incorporate erasure coding to encode each sensing record into a number of small slices. Then each of the slices can be transferred to a participant, and the latter reports the slice to the service provider. Once the service provider receives enough number of slices, not necessarily all the slices, it can decode the original sensing record. The usage of erasure coding has two advantages. One is to greatly reduce the communication overhead needed to transfer the sensing record (slices in this paper) to other participants. The other is to increase the reliability of the system, when the slices may be lost due to various reasons.

(2) Transfer Strategy

Since the slices need to be transferred to a set of participants, carefully selecting the participants to transfer to may affect the performance of the scheme. The straightforward strategy is to transfer a slice whenever another participant is met. However, when the participants in the system have different capabilities, the straightforward way may not be the best strategy. In this paper, we consider the case, in which the participants have different cost to deliver the same slice. The cost difference can be resulted from the wireless communication fee, available bandwidth, battery power, and so on. Through analysis, we also propose two sub-optimal slice transfer strategies to minimize the total cost for delivering the slices in section 3.3.2.

Fig. 2 shows the general work flow of our SLICER. Specifically, a sensing record contains the sensor reading and spatio-temporal information. Then, SLICER encodes the sensing record using an erasure coding technique(*e.g.*, Tornado [37]), encrypts the encoded slices, and attaches an unique tag, to generate encrypted slices. Next, SLICER selectively transfers the encrypted slices to the target participants, following one of its transfer strategies. The slices are delivered to the service provider through different participants. Finally, the service provider decrypts the slice and reconstructs the original sensing record, when enough number of slices are received.

In the following subsections, we present the design details of SLICER's major components, including Coding, Transferring, and Reconstructing.

3.2 Coding

Algorithm 1 Sensing Record Coding Algorithm

```
Input: A sensing record $< t, l, d >$ from participant $a_i \in N$, and coding rate $k/m$.
Output: Encrypted slices $\{r'_{ij} | 1 \leq j \leq m\}$.
$\{r_{ij} | 1 \leq j \leq m\} \leftarrow EC(< t, l, d >)$;
$nonce \leftarrow random()$;
$tag = H(i, nonce)$;
for all $j = 1$ to $m$ do
    $r'_{ij} = ENCRYPT(r_{ij} || tag, KEY_{pub})$;
end for
return $\{r'_{ij} | 1 \leq j \leq m\}$;
```

Algorithm 1 shows the pseudo-code of our sensing record coding algorithm. Given a sensing record $< t, l, d >$ from participant $a_i \in N$, we encode it into a number of slices, each of which will be delivered to the service provider through different participants. We encode the record $< t, l, d >$ using erasure coding (*e.g.*, Reed-Solomon [38] and Tornado [37]). Basically, erasure coding breaks a sensing record into fragments, expands and encodes with redundant data pieces into m slices. The original record can be reconstructed from any k out of m encoded slices, where $m > k$. The ratio k/m is the coding rate. Here, the combined size of any k slices is approximately equal to the size of the original record, according to Tornado Codes [37]. Intuitively, if the service provider decodes the record from k slices reported by k different participants, the real generator of the record is hidden in a group of k participants, which provides a privacy guarantee of k-anonymity. Furthermore, SLICER inherits the property of loss tolerance from erasure coding to achieve high record reconstruction ratio with relatively lower communication overhead. We denote the encoded slices by $\{r_{ij} | 1 \leq j \leq m\}$:

$$\{r_{ij} | 1 \leq j \leq m\} = EC(< t, l, d >),$$

where $EC(\cdot)$ is one of the erasure coding algorithms.

Since the service provider may receive a large number of encoded slices originating from various participants' sensing records, we have to tag the slices to clearly indicate which slices belong to the same record. Since directly tagging a slice with its generator's ID and a sequence number will reveal the identity privacy of the generator to the service provider,

1536-1233 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2352253, IEEE Transactions on Mobile Computing

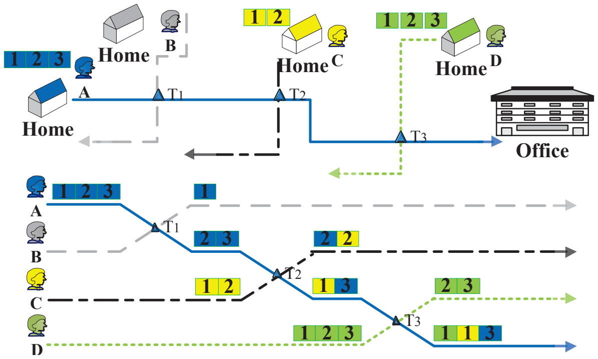


Fig. 3. An Example of Transfer on Meet Up

we adopt a cryptographic hash function (*e.g.*, SHA-1 [39]) to create the tag:

$$tag = H(i, nonce),$$

where $H(\cdot,\cdot)$ is a cryptographic hash function and $nonce$ is an arbitrary number. Since the pseudo-random number generator usually takes discrete time as the seed in practice, if multiple participants happen to initialize their pseudo-random number generators at the same time, then the same sequences of numbers will be generated as the nonce, resulting in encoded slices originating from different sensing records having the same tag. This will cause failure in the process of sensing record reconstruction. Therefore, we append the participant's ID to the randomly generated nonce, in order to eliminate the harm of nonce collision. Noting that the use of IDs in the form of plaintext reveals the participants privacy, we hash the combination of the generators ID and the nonce.

To prevent the content of encoded slices being revealed to external attacker and neighboring participants, we encrypt the encoded slices and the tag using the public key KEY_{pub} of the service provider and get the encrypted slices:

$$r'_{ij} = ENCRYPT(r_{ij}||tag, KEY_{pub}), 1 \leq j \leq m,$$

where $ENCRYPT(\cdot,\cdot)$ is an asymmetric encryption function, and $||$ is string concatenation operation.

3.3 Transferring

To prevent the service provider from recognizing participants' identities with the collected sensing records, not all slices of a sensing record can be directly sent to the service provider by the generator. To guarantee k-anonymity, at least $k-1$ slices need to be delivered by participants other than the generator. We note that although all the slices can be transferred to and delivered by participants other than the generator, SLICER requires the generator to report (at least) one slice to the service provider by herself, in order to guarantee the integrity of the sensing record.

In this paper, we consider two kinds of slice transferring strategies: *transfer on meet up (TMU)* and *minimal cost transfer (MCT)*.

3.3.1 Transfer on Meet Up (TMU)

This is the straightforward way to spread the encrypted slices. One slice of each sensing record is transferred, when the generator meets another participant. Later, all the participants, including the generator, report the slices(and received slices) to the service provider.

Fig. 3 shows a toy example of applying the strategy of TMU. Assume that there is a participant A who is going to office from her home. She meets other participants B, C, and D in sequence on her way to the office. The upper part of Fig. 3 shows the path that A travels, and the lower part shows the slices each of the users hold with advance of time. Assume that A, B, C, and D initially have 3, 0, 2 and 3 slices of their own, respectively, and meetings occurs at T_1, T_2, and T_3, at where a participant transfers one slice to the one met. For example, at T_1, A transfers one slice to B. After that, A has 2 slices left, and B holds 1 slice from A. Finally, after three meetings, A has 1 own slice and 2 slices from C and D, B has 1 slice from A, C has 1 own slices and 1 from A, and D has 2 own slices.

3.3.2 Minimal Cost Transfer (MCT)

In this section, we consider the case that different participants consume different costs to deliver a slice. The cost difference can be resulted from the wireless communication fee, available bandwidth, battery power, and so on. Intuitively, high cost will reduce people's enthusiasm to participate in the sensing activities. Here, we present our algorithms for the problem of Minimal Cost Transfer (MCT).

Each sensing record has an expiration time, before which the record has to be delivered to the service provider. We assume that each participant $a_i \in N$ knows a set $N(a_i) \subset N$ of participants that might be met before the expiration of the sensing record. For each participant $a_j \in N(a_i)$, let $p(a_j)$ and $c(a_j)$ be the meeting probability before the expiration time and the cost of the participant a_j for delivering a slice. As we mentioned before, the cost can be resulted from the wireless communication fee, available bandwidth, battery power, and so on. We assume that there is a mobility prediction module ([40]–[42]) to provide the prediction of $N(a_i)$, $(p(a_j))_{a_j \in N(a_i)}$, and $(c(a_j))_{a_j \in N(a_i)}$, based on historical event logs.

The objective of MCT is to pick a subset of participants $F \subseteq N(a_i)$ as forwarders of the slices to minimize the cost for delivering the slices, satisfying one of the following requirements.

- Requirement 1: It is expected to meet at least $m-1$ participants from the forwarder set F, namely MCT-EXP problem;
- Requirement 2: The (expected) probability of meeting at least $m-1$ participants from F is at least P $(0 \leq P \leq 1)$, namely MCT-PRO problem.

Next, we will present our approaches to solve the above two problems, MCT-EXP problem and MCT-PRO problem.

Solution to MCT-EXP Problem

We first consider the MCT-EXP problem (*i.e.*, MCT problem with requirement 1), which can be formulated as a binary program with an objective of minimizing the expected

1536-1233 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2352253, IEEE Transactions on Mobile Computing

delivery cost of the slices, as follows:

Objective:

$$Minimize \qquad \sum_{a_j \in N(a_i)} (c(a_j)p(a_j)x_j)$$

Subject to:

$$\sum_{a_j \in N(a_i)} (p(a_j)x_j) \geq m-1, \tag{1}$$

$$x_j \in \{0,1\}, \qquad \forall a_j \in N(a_i) \tag{2}$$

Here, constraint (1) guarantees that the participant a_i is expected to meet at least other $m-1$ participants in the selected forwarder set $F = \{a_j \in N(a_i) | x_j = 1\}$. Constraint (2) indicates the possible values of x_j. If a_j is selected to be a candidate for delivering a slice, then $x_j = 1$; otherwise, $x_j = 0$.

We note that the above formulation of MCT-EXP Problem can be reduced to the 0-1 Knapsack Problem [43] with an objective of maximizing the expected cost of the complimentary of the forwarder set. The re-formulated equation can be written as follows:

Objective:

$$Maximize \qquad \sum_{a_j \in N(a_i)} (c(a_j)p(a_j)(1-x_j))$$

Subject to:

$$\sum_{a_j \in N(a_i)} (p(a_j)(1-x_j)) \leq \sum_{a_j \in N(a_i)} p(a_j) - (m-1), \tag{3}$$

$$x_j \in \{0,1\}, \qquad \forall a_j \in N(a_i) \tag{4}$$

In the reduced 0-1 Knapsack Problem, $p(a_j)$ and $c(a_j)p(a_j)$ are the weight and value of the jth item, respectively, while the capacity of the knapsack is $\sum_{a_j \in N(a_i)} p(a_j) - (m-1)$. Here, constraint (3) guarantees that the sum of the weights must be less than the knapsack's capacity. Constraint (4) is exactly the same as constraint (2). Consequently, we can have a Fully Polynomial Time Approximation Scheme (FPTAS) [43], which runs in polynomial time and is correct within $1-\epsilon$ percent of the optimal solution, to solve the MCT-EXP problem. Due to limitations of space, we refer the reader to [43] for the detailed solution.

Solution to MCT-PRO Problem

Although we can have an FPTAS solution to the MCT-EXP problem, it is still not satisfactory, because the probability of meeting $m-1$ participants cannot be guaranteed at a high level. Therefore, we further consider the MCT-PRO problem, which strictly require that the probability of meeting at least $m-1$ participants from the forwarder set F is at least at a preset level P. Again, we formulate the MCT-PRO problem as a binary program, which aims to minimize the average delivery cost of the $m-1$ slices, as follows:

Objective: Minimize

$$\frac{\displaystyle\sum_{\vec{y}:\sum_{a_g \in N(a_i)}(x_g y_g)=m-1} \left(\sum_{a_j \in N(a_i)} (c(a_j)x_j y_j) \prod_{a_j \in N(a_i)} p(a_j)^{y_j} \right)}{\displaystyle\sum_{\vec{y}:\sum_{a_g \in N(a_i)}(x_g y_g)=m-1} \prod_{a_j \in N(a_i)} p(a_j)^{y_j}}$$

Subject to:

$$\sum_{t=m-1}^{\sum_{a_g \in N(a_i)} x_g} \sum_{\vec{y}:\sum_{a_g \in N(a_i)}(x_g y_g)=t} \prod_{a_j \in N(a_i)} \Big(p(a_j)^{y_j} \cdot (1-p(a_j))^{1-y_j}\Big) \geq P, \tag{5}$$

$$x_j \in \{0,1\}, \qquad \forall a_j \in N(a_i) \tag{6}$$

Here, the numerator of objective formula calculates the total "weighted" cost of all possible combinations of $m-1$ participants from a selected set of forwarders $F = \{a_j \in N(a_i) | x_j = 1\}$, while the denominator denotes the total "weight" of these combinations. The "weight" of a combination of $m-1$ participants here is the possibility of meeting exactly all of them by a_i. Consequently, the objective formula is to minimize the weighted-average cost for delivering the slices. Constraint (5) guarantees that a_i can meet at least $m-1$ participants in the selected forwarder set $F = \{a_j \in N(a_i) | x_j = 1\}$ with probability at least P. Constraint (6) is exactly the same as constraint (2). In the binary program, $\vec{y}$ is a binary vector with $|N(a_i)|$ bits. However, since the above binary program cannot be efficiently solved in polynomial time, we propose a polynomial time greedy algorithm, which can achieve good performance in most of the cases.

We first sort the participants in set $N(a_i)$ by $p(a_j)/c(a_j), a_j \in N(a_i)$ in non-increasing order β:

$$\beta : a'_1, a'_2, \ldots, a'_{|N(a_i)|},$$

such that

$$\frac{p(a_j)}{c(a_j)} \geq \frac{p(a_g)}{c(a_g)}, \forall 1 \leq j < g \leq |N(a_i)|.$$

Then, we find the smallest number α of participants in the front of the ordered list β, such that the probability of meeting at least $m-1$ of them is at least P (*i.e.*, constraint (5) is satisfied). We call the last selected participant in this process as *critical participant* and α as *critical number*. The pseudo-code for finding the critical participant is shown by Algorithm 2.

In Algorithm 2, we first check whether there are enough participants (Lines 1-3). If not, then there is no feasible solution; otherwise, we use a dynamic programming-based method to find the critical participant a'_α (Lines 4-14). In this process, we first initialize a one-dimensional array ρ for storing intermediate results (Line 4). Each element $\rho[j]$ $(0 \leq j \leq |N(a_i)|)$ means the probability of meeting j participant(s), given the first α participant(s) in the list β. Then we test the participants in list β one by one (Lines 5-9) and update the array elements up to $\rho[\alpha]$ (Lines 10-13), until the critical participant a'_α is identified. If no critical participant is

1536-1233 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2352253, IEEE Transactions on Mobile Computing

Algorithm 2 Finding Critical Participant

```
Input: Set of participants $N(a_i)$, profile of meeting probabilities $(p(a_j))_{a_j \in N(a_i)}$, profile of delivery costs $(c(a_j))_{a_j \in N(a_i)}$, ordered list $\beta$, and the minimal probability $P$.
Output: Critical participant $a'_\alpha$.
 1: if $|N(a_i)| < m - 1$ then
 2:     return "No feasible solution.";
 3: end if
 4: $\rho \leftarrow 0^{|N(a_i)|+1}$; $\rho[0] \leftarrow 1 - p(a'_1)$; $\rho[1] \leftarrow p(a'_1)$; $\alpha \leftarrow 1$;
 5: while $\sum_{j=m-1}^{\alpha} \rho[j] < P$ do
 6:     if $\alpha = |N(a_i)|$ then
 7:         return "No feasible solution.";
 8:     end if
 9:     $\alpha \leftarrow \alpha + 1$;
10:     for $g = \alpha$ to 1 do
11:         $\rho[g] \leftarrow \rho[g-1]p(a'_\alpha) + \rho[g](1 - p(a'_\alpha))$;
12:     end for
13:     $\rho[0] \leftarrow \rho[0](1 - p(a'_\alpha))$;
14: end while
15: return $a'_\alpha$;
```

Algorithm 3 Forwarder Set Selection

```
Input: Set of participants $N(a_i)$, profile of meeting probabilities $(p(a_j))_{a_j \in N(a_i)}$, profile of delivery costs $(c(a_j))_{a_j \in N(a_i)}$, ordered list $\beta$, and critical participant $a'_\alpha$.
Output: Set of forwarders $F$.
 1: $\rho \leftarrow 0^{|N(a_i)|+1,|N(a_i)|+1}$; $cost \leftarrow$ MAX_REAL;
 2: for $\gamma = 1$ to $|N(a_i)|$ do
 3:     for $j = 1$ to $|N(a_i)|$ do
 4:         for $g = \gamma$ downto 2 do
 5:             if $j = \gamma$ then
 6:                 $\rho[j][g] \leftarrow \rho[j][g-1]p(a'_1) + \rho[j][g]$;
 7:             else
 8:                 $\rho[j][g] \leftarrow \rho[j][g-1]p(a'_g) + \rho[j][g]$;
 9:             end if
10:         end for
11:         $\rho[j][1] \leftarrow p(a'_j)$; $\rho[j][0] \leftarrow 1$;
12:     end for
13:     if $\gamma \geq \alpha$ then
14:         $cost' \leftarrow \sum_{j=1}^{\gamma} c(a'_j)\rho[j][m-1]$;
15:         if $cost' < cost$ then
16:             $F \leftarrow$ first $\gamma$ participants in $\beta$; $cost \leftarrow cost'$;
17:         end if
18:     end if
19: end for
20: return $F$;
```

found, then return with no feasible solution (Lines 6-8). The runtime of Algorithm 2 is $O(n^2)$, where $n = |N(a_i)|$.

Noting that having more than α participants in the front of the ordered list β, constraint (5) is always satisfied. Consequently, after locating the critical participant a'_α, if any, each set with $\gamma \in \{\alpha, \alpha+1, \ldots, |N(a_i)|\}$ participants in the front of the ordered list β is a feasible solution of the MCT-PRO problem. So, our next job is to find the $\gamma \in \{\alpha, \alpha+1, \ldots, |N(a_i)|\}$ that minimize the objective function of the MCT-PRO problem formulation. Algorithm 3 shows our pseudo-code for selecting forwarder set F, given the critical participant a'_α found by Algorithm 2.

Algorithm 3 maintains a two-dimensional matrix ρ to store intermediate results. Each element $\rho[j][g]$ $(0 \leq j, g \leq |N(a_i)|)$ represents the probability of meeting g participants, under the condition that participant a'_j is met, given the first γ participants in the list β (during the process, the position of a'_1 and a'_j is switched for calculating the probabilities of row $\rho[j]$). After initialization (Line 1), we iterate each of the possible values of γ from 1 to $|N(a_i)|$ (Lines 2-19). For the iterations of γ from 1 to $\alpha - 1$, we only update the dynamic matrix ρ (Lines 3-12) without checking the average delivery cost, because the necessary number of participants has not been reached. From the iteration with $\gamma = \alpha$ on, we check the average delivery cost with $m - 1$ participants (Line 14), after updating the dynamic matrix ρ (Lines 3-12). If a lower average delivery cost is found (*i.e.*, $cost' < cost$), we update the current smallest average delivery cost and its corresponding forwarder set (Lines 14-17). Finally, Algorithm 3 returns the forwarder set F. The running time of Algorithm 3 is $O(n^3)$, where $n = |N(a_i)|$.

Algorithm 3 can return a feasible result if there are sufficient number of meeting opportunities with other participants. However, we note that it is possible that a sensing record generator cannot meet enough participants to transfer each of the encoded slices from a record to a different participant. In this case, we use the prediction model based on the history to estimate the number of encounters beforehand. For participants who do not have sufficient slice transfer opportunities, we allow them to transfer more than one slice during each meeting. Suppose h slices are transferred each time, then the record generator is hidden in $\lceil k/h \rceil$ participants.

3.4 Reconstructing

After receiving at least k slices encoded from the same sensing record, the service provider can reconstruct the original sensing record. Besides maintaining a database storing the sensing records, the service provider also keeps a table T caching slices that have not been decoded.

Algorithm 4 shows the pseudo-code of our sensing record reconstructing algorithm. Upon receiving a reported slice s, the service provider decrypts the slice using her private key KEY_{priv} to get the encoded slice s' and a tag that uniquely identifies the record it is encoded from:

$$(s', tag) = DECRYPT(s, KEY_{priv}),$$

where $DECRYPT(\cdot, \cdot)$ is an asymmetric decryption function.

The service provider adds the encoded slice s' into the caching table T with index tag, and then check whether there are k encoded slices with the same tag. Then, the service provider checks the integrity of the k slices. If these slices pass the integrity check, service provider extracts the k encoded slices with the same tag, and then decodes the original sensing record:

$$< t, l, d >= EC^{-1}(\{\bar{s} | < \bar{s}, \bar{t} > \in T \wedge \bar{t} = tag\}),$$

1536-1233 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2352253, IEEE Transactions on Mobile Computing

Algorithm 4 Sensing Record Reconstructing Algorithm

```
Input: Caching table $T$.
Output: Each original sensing record $< t, l, d >$.
while $TRUE$ do
    Receive slice $s$;
    $(s', tag) \leftarrow DECRYPT(s, KEY_{priv})$;
    Add $(s', tag)$ into $T$;
    if $|\{\bar{s}| < \bar{s}, \bar{t} > \in T \wedge \bar{t} = tag\}| \geq k$ then
        if IntegrityCheck($\{\bar{s}| < \bar{s}, \bar{t} > \in T \wedge \bar{t} = tag\}$)=$true$ then
            $< t, l, d > \leftarrow EC^{-1}(\{\bar{s}| < \bar{s}, \bar{t} > \in T \wedge \bar{t} = tag\})$;
            Remove $\{\bar{s}| < \bar{s}, \bar{t} > \in T \wedge \bar{t} = tag\}$ from $T$;
            Store sensing record $< t, l, d >$;
        else
            Remove $\{\bar{s}| < \bar{s}, \bar{t} > \in T \wedge \bar{t} = tag\}$ from $T$;
        end if
    end if
end while
```

where $EC^{-1}(\cdot)$ is the decoding function corresponding to $EC(\cdot)$. Otherwise, the collected slices marked with tag are removed from the caching table.

3.5 Analysis

In this section, we show that SLICER can provide strong privacy protection against the external and internal attacks.

3.5.1 Protection Against External Attacks

The external attacker eavesdrops messages passed in the participatory sensing system, in order to collect sensitive information about particular participants. In SLICER, we employ an end-to-end cryptographic encryption scheme, such that the external attacker cannot decrypt the slices transferred among participants, as well as that reported to the service provider. Although the external attacker may extract some information from the eavesdropped packets to uniquely identify the participant, she cannot get the concrete content of the sensing record. Because the eavesdropped content is under the protection of the end-to-end encryption, such that the eavesdropper cannot decrypt it unless she colludes with service provider. Therefore, SLICER provides privacy protection against the external attacks.

3.5.2 Protection Against Internal Attacks

The internal attack may come from both the participants and the service provider. We distinguish two cases:

Protection against participants' attack

Each participant may receive some slices, when she is selected as a slice deliver for participants met. Similar with the external attacker, the participant cannot decrypt the slice for delivering.

Protection against service provider's attack

Since the service provider has full access to the sensing records contributed by the participants, she can easily infer private information about the participants, if proper privacy-preserving scheme is not provided. However, SLICER can achieve the k-anonymity and protect participants' privacy information against the service provider. Therefore, we can draw the following theorem.

Theorem 1: SLICER achieves k-anonymity, when there are k participants who deliver slices to the service provider.

Proof: In SLICER, we isolate the participants' identity and the sensing records, by encoding each sensing record into m slices and letting at least k different slices be delivered to the service provider through different participants. To achieve this, we designed three different algorithms (TMC, MCT-EXP, and MCT-PRO) in section 3.3 according to different situations to select at least m participants (including the generator itself) as forwarders to transfer m slices to the service provider. Then, the original sensing record can be decoded by the service provider if and only if receiving at least k different slices. Therefore, the identity of the record generator is hidden among a group of at least k participants. □

We note that SLICER's privacy guarantee degrades to $\lceil k/h \rceil$-anonymity, when a sensing record generator cannot meet enough participants to transfer slices and thus has to transfer h slices during each meeting. Further, if the sensing record generator is completely isolated and cannot meet any other participant (*i.e.*, $h = k$), SLICER cannot preserve the privacy on linkage between identity and location. In this case, an alternative privacy preserving scheme (*e.g.*, [11], [21], [44]) can be applied.

4 Evaluation

We have implemented the SLICER and evaluated its performance on taxi traces collected from practice. In this section, we specify evaluation setups and metrics, and present evaluation results.

4.1 Setup and Metrics

Our evaluation is based on the realistic GPS mobility traces of 500 taxi cabs over 30 days in San Francisco, USA, which were collected by Cabspotting Project [19] and can be accessed from the CRAWDAD [45] website. In this real world deployment, each cab is outfitted with a GPS tracking device that is used by dispatchers to efficiently reach customers. Each cab sends a location-update triplet (timestamp, identifier, geo-coordinates) to a central server in a period varied from 30 to 60 seconds, which forms the mobility traces we used in this paper. We extend this scenario to a participatory sensing situation by assuming that the cabs are participants equipped with mobile devices.

We consider a mobile infrastructure with the whole 500 participants. We set that every participant generates one record per day, and the period of validity of the record is 24 hours. The loss possibility of the slices varies from 0.2 to 0.4.

We evaluate the performance of SLICER using the following four metrics.

- *Reconstruction Ratio*: The percentage of sensing records successfully reconstructed by the service provider. This reflects the loss tolerance of SLICER.
- *Communication Overhead*: The total amount of data transmitted to guarantee required reconstruction ratio.

1536-1233 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2352253, IEEE Transactions on Mobile Computing

- *Computation Overhead*: The time consumed to process a sensing record.
- *Total Transfer Cost*: The sum of the cost for delivering a sensing record (*i.e.*, $m-1$ slices) to the service provider.

4.2 Evaluation Results on Reconstruction Ratio

We compare the performance of SLICER implemented with the three transfer strategies proposed in Section 3 (*i.e.*, T-MU, MCT-EXP, and MCT-PRO), with an existing privacy preserving schemes for participatory sensing, namely Simple Exchanging [35], in which the sensing records are transferred among participants as a whole without coding. We should note that we did not compare with [11], [12], [21], because the setup of these work are significantly different with ours.

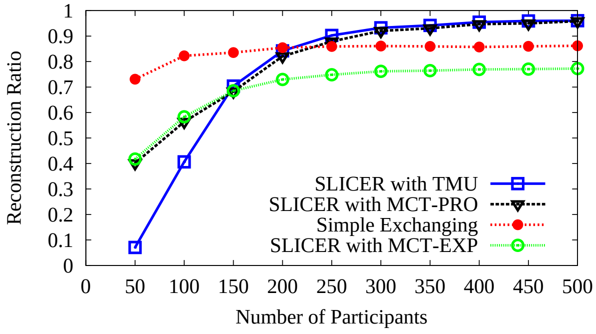
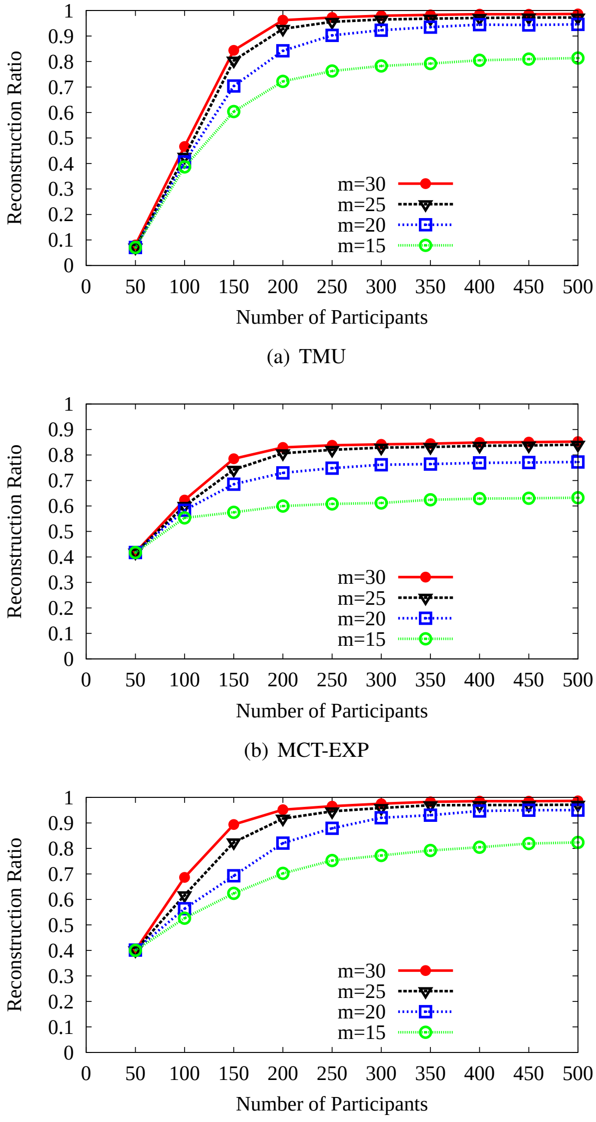


Fig. 4. Impact of Participant Number on Reconstruction Ratio

Fig. 4 shows the reconstruction ratios achieved by the four schemes with growing number of participants, which are selected from the public taxi trace dataset. We set the coding rate to $10/20$ and the probability of slice loss to 0.2 in this simulation. To be fair, we let the four evaluated schemes have the same communication overhead, and then compare their achieved reconstruction ratios. Specifically, given that the coding rate of our three SLICER strategies is 10/20, the total size of encoded slices is doubled from the original sensing record. So, we let the Simple Exchanging scheme transfer twice for each sensing record. We can see from Fig. 4 that SLICER with TMU and SLICER with MCT-PRO perform better than Simple Exchanging, when there are sufficient number of participants (*i.e.*, $>$ 200 participants). This is because SLICER inherits high loss tolerant capability from erasure coding technique. Specifically, the reconstruction ratio of SLICER with TMU, SLICER with MCT-PRO reaches 0.97 when there are 400 participants or more. In contrast, Simple Exchanging has relatively stable reconstruction ratio (about 0.86). However, we can see that SLICER with MCT-EXP performs not well, due to the fact that the MCT-EXP strategy may not guarantee the probability of meeting $m-1$ participants at a high level. In addition, when the number of participants is less than 200, Simple Exchanging performs the best. This is because Simple Exchanging only needs one other participant to deliver the sensing record, while SLICER needs $m-1$ participants. However, Simple Exchanging cannot improve its reconstruction ratio with the help of increasing number of participants, and loses its advantage when the number of participants grows beyond 200. Furthermore, Simple Exchanging cannot provide the strong guarantee of k-anonymity. So the results of this simulation confirms that SLICER with TMU or MCT-PRO is preferred when there are sufficient number of participants in the participant sensing system.



(a) TMU

(b) MCT-EXP

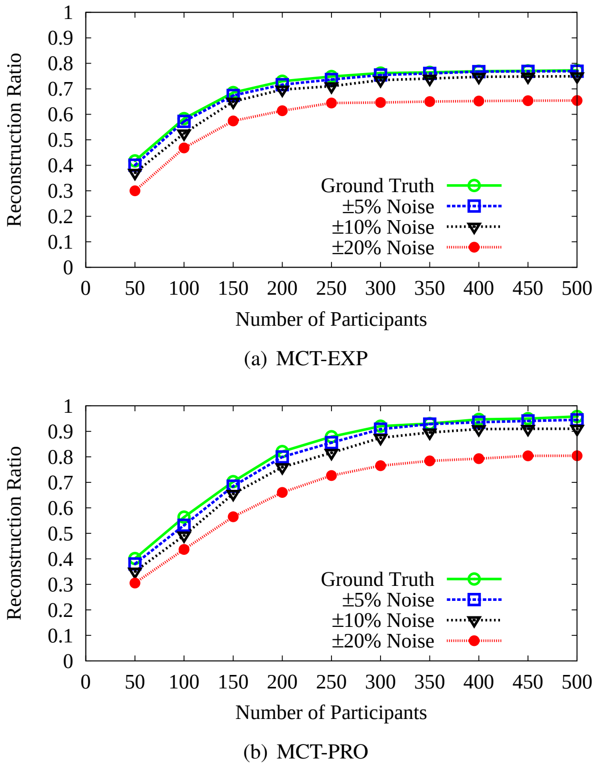
(c) MCT-PRO

Fig. 5. Impact of Coding Rate k/m on Reconstruction Ratio (We fix $k=10$, and vary m in this evaluation.)

Then, we evaluate the impact of coding rate (k/m) on reconstruction ratio of our transfer strategies, including TMU, MCT-EXP, and MCT-PRO. The evaluation results are shown in Fig. 5. Here, we fix $k = 10$, and vary the value of m from 15 to 30 with a step of 5 in this evaluation. The slice losing probability is again set to 0.2. From Fig. 5, we can see that the reconstruction ratios achieved by the three transfer strategies increase with the decrement of coding rate (*i.e.*, increment of m in the evaluation) and increment of the number of participants. Having coding rates of 10/25 and 10/30, each

1536-1233 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

of the three transfer strategies produces close reconstruction ratios, which are clearly higher than those in cases of 10/15 and 10/20. This indicates that coding the sensing record into at least 25 slices can achieve relatively good reconstruction ratio on the dataset used in our evaluation. We note that the coding ratio still need to be carefully set for different application scenarios in order to obtain high reconstruction ratios with appropriate costs.



(a) MCT-EXP

(b) MCT-PRO

Fig. 6. Impact of Inaccurate Mobility Prediction Module on the Reconstruction Ratios of Our Designs

Furthermore, we evaluate the impact of inaccurate mobility prediction module on the performance of our designs. In this set of evaluations, we directly add noises to the meeting probabilities generated by the mobility prediction module to make them deviate from the ground-truth prediction. Fig. 6 shows the evaluation results. By adding ±5% (±10% and ±20%) noise, we mean the meeting probabilities are randomly increased or decreased by up to 5% (10% and 20%) from their ground truth values, respectively. In this evaluation, the coding rate is set to 10/20, and the probability of slice loss is 0.2. Fig. 6(a) shows the results for MCT-EXP. We can observe that the reconstruction ratios achieved by MCT-EXP with ±5% and ±10% noise are very close to the case with ground-truth prediction. Specifically, when ±10% noise is added, reconstruction ratio is only decreased by 4.92% from the result on ground truth, given 500 participants. Only when the noise is as large as ±20%, the reconstruction ratio is decreased by 15.28% for 500 participants. Besides, the results shown in Fig. 6(b) for MCT-PRO is quite similar to those for MCT-EXP. Reconstruction ratios of MCT-PRO with ±5% and ±10% noise have good approximations to that of MCT-PRO with ground-truth prediction, while MCT-PRO with ±20% noise suffers from 16.1% decrement on construction ratio for 500 participants. These results show that our approaches can tolerate small amount of prediction inaccuracy

4.3 Evaluation Results on Overhead

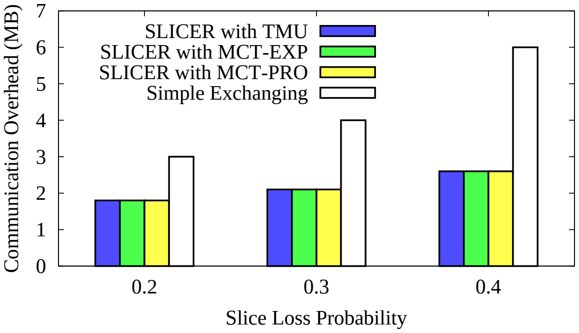


Fig. 7. Communication Overhead to Achieve Reconstruction Ratio of 0.99

We evaluate the communication overhead of four schemes (TMU, MCT-EXP, MCT-PRO, and Simple Exchanging) to achieve a targeted reconstruction ratio of 0.99, under different slice losing probabilities. We set the sensing record size to $1MB$. Three loss probabilities are evaluated. To achieve the reconstruction ratio of 0.99, the coding rate of SLICER needs to reach 10/18, 10/21, and 10/26, when the loss probability is 0.2, 0.3, and 0.4, respectively. Similarly, we also set proper transmission redundancies for the Simple Exchanging for different loss probabilities. As shown in Fig. 7, we can see that the communication overhead of SLICER is always lower than Simple Exchanging under different losing probabilities, showing that SLICER has better loss tolerance. Although the communication overheads of the four schemes increase with the loss probability, the growth speed of SLICER is much slower. In addition, the performance of SLICER implemented with different transfer strategies has subtle differences due to the reason that participants selected by SLICER with MCT-EXP and MCT-PRO may not be met in some probability. This result confirms that SLICER can achieve low communication overhead.

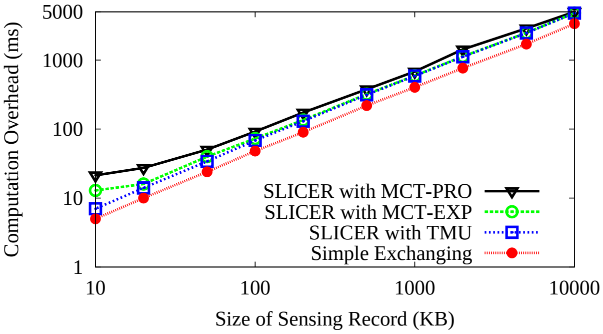


Fig. 8. Computation Overhead

We also evaluate the computation overhead of SLICER with different transfer strategies (as shown in Fig. 8, which is a

1536-1233 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

log-log scale plot), comparing with the traditional encryption only scheme, which is the Simple Exchanging [35] scheme (RSA is adopted in this simulation). What we consider in SLICER are only the computations needed at mobile device side, including erasure coding (Tornado [37]), hashing (SHA-1 [39]), encryption (RSA [46]), and running the three different transfer strategies. Our schemes are evaluated in Windows 7 OS environment, with C++ programmed simulator running on a computer with a CPU speed of 2.40GHz. In Fig. 8, we can see that SLICER induces some extra computation overhead compared with the Simple Exchanging scheme, when dealing with the same size of data. This is caused mainly by the usage of erasure coding. Although, SLICER with TMU, SLICER with MCT-EXP, and SLICER with MCT-PRO consume 42.5%, 48.3% and 42.7% more time than the Simple Exchanging method when dealing with a $10MB$ sensing record, respectively, the per kilobyte computation overheads are still very small and can be afforded by mobile devices. Specifically, for a 10MB sensing record, SLICER with TMU, SLICER with MCT-EXP, and SLICER with MCT-PRO consume 0.479ms/KB, 0.480ms/KB, and 0.499ms/KB, respectively.

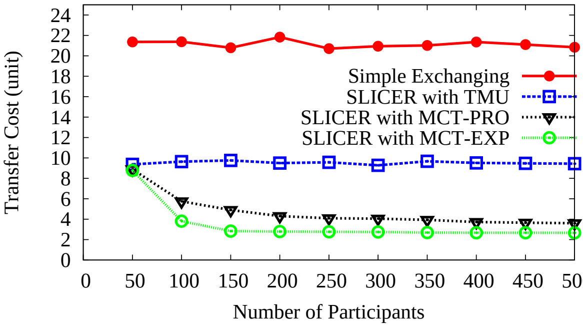


Fig. 9. Transfer Cost of Different Strategies

Then, we compare the total slice transfer cost when using different strategies (Simple Exchanging, TMU, MCT-EXP, and MCT-PRO) to achieve an expected construction ratio of 0.9994. The transfer cost on each participant is generated randomly from (0, 1], the meeting probability (used in Minimal Cost Transfer) comes from the statistics of the 500 participants' trace data, the probability of slice loss is set to 0.2, and the coding rate is set to 10/20. As shown in Fig. 9, Simple Exchanging performs the worst, and SLICER with TMU also suffers from relatively high transfer cost, which is close to the excepted value (*i.e.*, $0.5 \times 19 = 9.5$). Minimal Cost Transfer performs better than the previous two when the participants are sufficient, due to the well designed algorithms, especially the MCT-EXP. We believe the MCT-PRO is more reasonable due to its threshold probability of P (0.9). In addition, the transfer cost of SLICER converges with the growth of number of participants, because more participants will provide more meeting opportunities, higher meeting probability, and more low-cost relays to select. For example, the total cost for transferring one record is lower than 3.6 by SLICER with MCT-PRO, when there are 500 participants. This means a per participant cost of 0.19. The results of this simulation confirm that our algorithm for minimal cost transfer can reduce transfer cost.

4.4 Summary of Evaluation Results

We summarize the above evaluation results as follows:

- SLICER with TMU provides the best reconstruction ratio when the participants are sufficient in the participatory sensing system. It also has relatively low communication and computation overhead compared with other SLICER strategies. However, TMU has the highest transfer cost.
- SLICER with MCT-EXP is more sensitive to inaccurate mobility predictions than MCT-PRO (and actually TMU is not sensitive at all), but has a lower computation overhead.
- SLICER with MCT-PRO achieves the lowest transfer cost and (near) the best reconstruction ratio, but has a little bit higher computation overhead.

Generally speaking, SLICER with MCT-PRO provides good reconstruction ratio with appropriate overhead most of the time, while SLICER with TMU can be a good alternative when the mobility prediction module is not available or inaccurate.

5 Related Work

In this section, we first review some related work on privacy preserving techniques for participatory sensing, and then review the work on data aggregation. Finally, we analyse some key differences with the closely related previous work.

5.1 Privacy Preserving Techniques

In the current state-of-the-art, a number of privacy preserving techniques for participatory sensing systems, especially the location-based services (LBSs), have been proposed by previous researchers, mainly to address the privacy of data source identity, user location, user trajectory, and sensing data content itself. These techniques can be classified into the following four categories.

5.1.1 Randomization Based Techniques

Randomization (noise) based technique [13], [47]–[49], where noise (*e.g.*, Gaussian noise) may be added into the original data, can hide the real value of sensitive information (*e.g.*, the trend of the data over time). This method was widely studied and used in data mining field. However, the loss of data quality is a significant shortcoming.

5.1.2 Generalization

The k-anonymity [36] model, which aims to hide each user's sensitive information among $k-1$ others', is a universal metric for privacy preservation, and has been applied to participatory sensing in several previous work [11], [50]. However, this kind of method usually needs an honest third-party as the anonymizer, which is not allowed in ubiquitous semi-honest models. Therefore, when a more severe situation of semi-honest third-party is considered, these approaches cannot meet requirements.

1536-1233 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2352253, IEEE Transactions on Mobile Computing

5.1.3 Cloaking Techniques

Cloaking techniques usually use generalization or perturbation to replace the actual location with larger area or to cloak real location using some functions (*e.g.*, [11], [44], [50], [51]). However, while spatial cloaking techniques can well protect single location information, they fail to protect the trace privacy, with which user's identity is also inferable [22], [52]. Recently, several work were proposed aiming to solve the trajectory privacy problems [21], [35], [53]. However, same questions exist that the protection of privacy reduces the quality of reported data.

5.1.4 Cryptography Based Solutions

End-to-end encryption, which can guarantee the high security of reported data, is widely used for the privacy preservation [12], [54]–[57]. However, encryption can only protect participants' privacy from external attacks (*e.g.*, the eavesdropper). When the encrypted data arrives at the service provider side, service provider can decrypt ciphertext and obtain the corresponding plaintext. Therefore, encryption technique fails to prevent the service provider from inferring users sensitive features. Since internal attacks are also undesirable, designing privacy preserving schemes for participatory sensing against both external and internal attacks is highly important.

5.2 Data Aggregation Protocols

Data aggregation is a widely used technique in wireless sensor networks. Data aggregation algorithms are designed to gather and aggregate data in an energy efficient manner so that the network lifetime is enhanced [58]–[60]. Cam *et al.* [61] presented a multi-stage real-time alert aggregation technique over mobile networks that greatly reduces the amount data transmission and attempts to maximize the bandwidth utilization. Kumar *et al.* [62] proposed a learning automata-based opportunistic data aggregation and forwarding scheme for alert generation in vehicle ad hoc networks (VANETs), which overcomes the challenges of high velocity and constant topological changes in VANETs and can adaptively select the next hop for data forwarding and aggregation from the other nodes. However, security issues are not considered in these data aggregation protocols.

As mentioned above, security is an important issue in the process of data aggregation. Secure data aggregation protocols (*e.g.*, [59], [63], [64]) try to achieve security requirements (*e.g.*, data integrity, data confidentiality, authentication, and etc.) along with data aggregation. Aviv *et al.* [65] proposed a privacy-aware geographic message exchange protocol for Human Movement Networks (HumaNets). However, they only consider static networks. Therefore, these methods are not suitable for participatory sensing, where the network changes dynamically.

5.3 Differences with Existing Work

In the literature, [35] and [12] are the two most closely related work to ours. Christin *et al.* [35] proposed to hide participants' travel paths via collaborative message exchanging in physical proximity. However, although by carefully setting the exchange strategies and the reporting strategies, various levels of privacy preservation against the application administrator can be achieved (e.g., k-anonymity), the approach is still vulnerable to privacy breach from malicious participants, since the triplets encapsulating the whole sensor readings are directly transferred to the encountered participants. Shi *et al.* [12] elegantly implemented a data aggregation method for supporting various aggregation functions on numerical data. However, their method cannot be applied to multimedia sensing data. In contrast, SLICER proposed in this paper is a coding-based k-anonymous privacy preserving scheme for high quality multimedia data aggregation in participatory sensing systems.

6 Conclusion and Future Work

In this paper, we have presented a coding-based privacy preserving scheme, namely SLICER, which is a k-anonymous privacy preserving scheme for participatory sensing with multimedia data. SLICER integrates the technique of erasure coding and well designed slice transfer strategies, to achieve strong protection of participants' private information as well as high data quality and loss tolerance, with low computation and communication overhead. We have studied two kinds of data transfer strategies, including TMU and MCT. While TMU is a simple and straightforward strategy, MCT contains two complimentary algorithms, including an approximation algorithm and a heuristic algorithm, designed for satisfying different levels of delivery guarantee. We also implement SLICER and evaluate its performance using publicly released taxi traces. Our evaluation results confirm that SLICER achieves high data quality, strong robustness, with low computation and communication overhead.

For future work, one possible direction is to study the problem of privacy preservation in the query process [66] [67], and design new privacy preserving query schemes based on SLICER. We also think about the lost-packet authentication in server side to increase the construction ratio and further reduce the communication overhead. Another possible direction is to design efficient slice transfer algorithm, considering the limitation of mobile devices' battery power, storage space, availability, computation ability, and communication bandwidth.

Acknowledgements

A preliminary version of this paper appeared at the IEEE MASS 2013 [68]. We would like to thank the anonymous reviewers of for their insightful comments and suggestions. We would also thank Pan Shao and Chen Jiang for their effort in improving the quality of this work.

References

[1] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B. Srivastava, "Participatory sensing," in *First Workshop on World-Sensor-Web (WSW), co-located with the 4th ACM Conference on Embedded Networked Sensor Systems (SenSys)*, Boulder, Colorado, USA, Oct. 2006.

[2] "The world in 2013: ICT Facts and Figures," International Telecommunication Union. [Online]. Available: http://www.itu.int

1536-1233 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2352253, IEEE Transactions on Mobile Computing

[3] R. K. Ganti, N. Pham, H. Ahmadi, S. Nangia, and T. F. Abdelzaher, "GreenGPS: a participatory sensing fuel-efficient maps application," in *Proceedings of The 8th International Conference on Mobile Systems, Applications, and Services (MobiSys)*, San Francisco, California, USA, Jun. 2010.

[4] M. Mun, S. Reddy, K. Shilton, N. Yau, J. Burke, D. Estrin, M. Hansen, E. Howard, R. West, and P. Boda, "PEIR, the personal environmental impact report, as a platform for participatory sensing systems research," in *Proceedings of The 7th International Conference on Mobile Systems, Applications, and Services (MobiSys)*, Krakw, Poland, Jun. 2009.

[5] E. D. Cristofaro and C. Soriente, "Pepsi: Privacy enhancing participatory sensing infrastructure," in *Proceedings of 4th ACM Conference on Wireless Network Security (WiSec)*, Hamburg, Germany, Jun. 2011.

[6] ——, "Extended capabilities for a privacy-enhanced participatory sensing infrastructure (pepsi)," *IEEE Transactions on Information Forensics and Security (TIFS)*, vol. 8, no. 12, 2013.

[7] X. O. Wang, W. Cheng, P. Mohapatra, and T. Abdelzaher, "Artsense: anonymous reputation and trust in participatory sensing," in *Proceedings of The 32nd IEEE International Conference on Computer Communications (INFOCOM)*, Turin, Italy, Apr. 2013.

[8] M. von Kaenel, P. Sommer, and R. Wattenhofer, "Ikarus: large-scale participatory sensing at high altitudes," in *Proceedings of The 12th Workshop on Mobile Computing Systems and Applications (HotMobile)*, Phoenix, Arizona, USA, Mar. 2011.

[9] R. K. Ganti, N. Pham, Y.-E. Tsai, and T. F. Abdelzaher, "PoolView: stream privacy for grassroots participatory sensing," in *Proceedings of The 6th International Conference on Embedded Networked Sensor Systems (SenSys)*, Raleigh, NC, USA, Nov. 2008.

[10] N. Xia, H. H. Song, Y. Liao, M. Iliofotou, A. Nucci, Z.-L. Zhang, and A. Kuzmanovic, "Mosaic: Quantifying privacy leakage in mobile networks," in *Proceedings of The ACM Special Interest Group on Data Communication (SIGCOMM)*, Hong Kong, China, Aug. 2013.

[11] K. Vu, R. Zheng, and J. Gao, "Efficient algorithms for k-anonymous location privacy in participatory sensing," in *Proceedings of The 31st Annual IEEE International Conference on Computer Communications (INFOCOM)*, Orlando, Florida USA, Mar. 2012.

[12] J. Shi, R. Zhang, Y. Liu, and Y. Zhang, "PriSense: privacy-preserving data aggregation in people-centric urban sensing systems," in *Proceedings of The 29th IEEE International Conference on Computer Communications (INFOCOM)*, San Diego, CA, USA, Mar. 2010.

[13] B. Hoh, M. Gruteser, H. Xiong, and A. Alrabady, "Preserving privacy in GPS traces via uncertainty-aware path cloaking," in *Proceedings of The 14th ACM Conference on Computer and Communications Security (CCS)*, Alexandria, VA, USA, Oct. 2007.

[14] R. Chen, I. E. Akkus, and P. Francis, "SplitX: High-performance private analytics," in *Proceedings of The ACM Special Interest Group on Data Communication (SIGCOMM)*, Hong Kong, China, Aug. 2013.

[15] E. D. Cristofaro and C. Soriente, "Participatory privacy: Enabling privacy in participatory sensing," *IEEE Network*, vol. 27, no. 1, 2013.

[16] C. Cornelius, A. Kapadia, D. Kotz, D. Peebles, M. Shin, and N. Triandopoulos, "Anonysense: Privacy-aware people-centric sensing," in *Proceedings of The 6th International Conference on Mobile Systems, Applications, and Services (MobiSys)*, Breckenridge, CO, USA, Jun. 2008.

[17] Q. Li and G. Cao, "Efficient and privacy-preserving data aggregation in mobile sensing," in *Proceedings of the 20th IEEE International Conference on Network Protocols (ICNP)*, Austin, TX, USA, Oct. 2012.

[18] J. Soldatos, M. Draief, C. Macdonald, and I. Ounis, "Multimedia search over integrated social and sensor networks," in *Proceedings of The International World Wide Web Conference (WWW)*, Lyon, France, Apr. 2012.

[19] "Cabspotting Project." [Online]. Available: http://cabspotting.org/

[20] Z. Yang, S. Zhong, and R. N. Wright, "Anonymity-preserving data collection," in *Proceedings of The 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, Chicago, Illinois, USA, Aug. 2005.

[21] J. Meyerowitz and R. Roy Choudhury, "Hiding stars with fireworks: location privacy through camouflage," in *Proceedings of The 15th Annual International Conference on Mobile Computing and Networking (MobiCom)*, Beijing, China, Sep. 2009.

[22] C. Y. T. Ma, D. K. Y. Yau, N. K. Yip, and N. S. V. Rao, "Privacy vulnerability of published anonymous mobility traces," in *Proceedings of The 16th Annual International Conference on Mobile Computing and Networking (MobiCom)*, Chicago, Illinois, USA, Sep. 2010.

[23] K. Shilton, J. Burke, D. Estrin, M. Hansen, and M. B. Srivastava, "Participatory privacy in urban sensing," in *Workshop on Mobile Device and Urban Sensing*, Apr. 2008.

[24] N. Xia, H. H. Song, Y. Liao, M. Iliofotou, A. Nucci, Z.-L. Zhang, and A. Kuzmanovic, "Mosaic: Quantifying privacy leakage in mobile networks," in *Proceedings of The ACM Special Interest Group on Data Communication (SIGCOMM)*, Hong Kong, China, Aug. 2013.

[25] S. Han, V. Liu, Q. Pu, S. Peter, T. Anderson, A. Krishnamurthy, and D. Wetherall, "Expressive privacy control with pseudonyms," in *Proceedings of The ACM Special Interest Group on Data Communication (SIGCOMM)*, Hong Kong, China, Aug. 2013.

[26] Y. Lindell and B. Pinkas, "Privacy preserving data mining," *The Journal of Cryptology*, vol. 15, no. 3, pp. 177–206, 2002.

[27] J. Vaidya and C. Clifton, "Privacy preserving association rule mining in vertically partitioned data," in *Proceedings of The 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, Edmonton, Alberta, Canada, Jul. 2002.

[28] ——, "Privacy preserving naïve bayes classifier for vertically partitioned data," in *Proceedings of The 4th SIAM International Conference on Data Mining (SDM)*, Orlando, Florida, USA, Apr. 2004.

[29] R. Wright and Z. Yang, "Privacy-preserving bayesian network structure computation on distributed heterogeneous data," in *Proceedings of The 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, Seattle, Washington, USA, Aug. 2004.

[30] Z. Yang, S. Zhong, and R. N. Wright, "Privacy-preserving classification of customer data without loss of accuracy," in *Proceedings of The 5th SIAM International Conference on Data Mining (SDM)*, Newport Beach, CA, USA, Apr. 2005.

[31] A. W.-C. Fu, R. C.-W. Wong, and K. Wang, "Privacy-preserving frequent pattern mining across private databases," in *Proceedings of The 5th IEEE International Conference on Data Mining (ICDM)*, Houston, Texas, USA, Nov. 2005.

[32] W. Lu, A. L. Varna, and M. Wu, "Security analysis for privacy preserving search of multimedia," in *Proceedings of The 17th IEEE International Conference on Image Processing (ICIP)*, Hong Kong, China, Sep. 2010.

[33] A. Adams and M. Sasse, "Taming the wolf in sheeps clothing: Privacy in multimedia communications," in *Proceedings of The 7th ACM International Conference on Multimedia (ACM Multimedia)*, Orlando, FL, USA, Oct. 1999.

[34] Z. Erkin, M. Franz, J. Guajardo, S. Katzenbeisser, R. L. Lagendijk, and T. Toft, "Privacy-preserving face recognition," in *Proceedings of The 9th International Symposium on Privacy Enhancing Technologies (PETS)*, Seattle, WA, USA, Aug. 2009.

[35] D. Christin, J. Guillemet, A. Reinhardt, M. Hollick, and S. S. Kanhere, "Privacy-preserving collaborative path hiding for participatory sensing applications," in *Proceedings of The 8th IEEE International Conference on Mobile Ad-hoc and Sensor Systems (MASS)*, Valencia, Spain, Oct. 2011.

[36] L. Sweeney, "k-Anonymity: a model for protecting privacy," *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, vol. 10, no. 5, pp. 557–570, 2002.

[37] J. W. Byers, M. Luby, M. Mitzenmacher, and A. Rege, "A digital fountain approach to reliable distribution of bulk data," in *Proceedings of The ACM Special Interest Group on Data Communication (SIGCOMM)*, Vancouver, B.C., Canada, Aug. 1998.

[38] I. S. Reed and G. Solomon, "Polynomial codes over certain finite fields," *Journal of the Society for Industrial and Applied Mathematics*, vol. 8, no. 2, pp. 300–304, 1960.

[39] D. Eastlake, 3rd and P. Jones, "US Secure Hash Algorithm 1 (SHA1)," RFC 3174, September 2001. [Online]. Available: http://www.rfc-editor.org/in-notes/rfc3174.txt

[40] G. Lin, G. Noubir, and R. Rajaraman, "Mobility models for ad hoc network simulation," in *Proceedings of The 23rd Annual IEEE International Conference on Computer Communications (INFOCOM)*, Hong Kong, China, Mar. 2004.

[41] K. Lee, S. Hong, S. J. Kim, I. Rhee, and S. Chong, "SLAW: a new mobility model for human walks," in *Proceedings of The 28th Annual IEEE International Conference on Computer Communications (INFOCOM)*, Rio de Janeiro, Brazil, Apr. 2009.

[42] X. Wang, W. Huang, S. Wang, J. Zhang, and C. Hu, "Delay and capacity tradeoff analysis for MotionCast," *IEEE/ACM Transactions on Networking*, vol. 19, no. 5, pp. 1354–1367, Oct. 2012.

[43] V. V. Vazirani, *Approximation Algorithms*. Springer Verlag Press, 2001.

[44] B. Hoh, M. Gruteser, R. Herring, J. Ban, D. Work, J.-C. Herrera, A. M. Bayen, M. Annavaram, and Q. Jacobson, "Virtual trip lines for distributed privacy-preserving traffic monitoring," in *Proceedings of The 6th International Conference on Mobile Systems, Applications, and Services (MobiSys)*, Breckenridge, CO, USA, Jun. 2008.

[45] "CRAWDAD: a community resource for archiving wireless data at dartmouth." [Online]. Available: http://crawdad.cs.dartmouth.edu/

1536-1233 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2352253, IEEE Transactions on Mobile Computing

[46] R. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Communications of the ACM*, vol. 21, no. 2, pp. 120–126, 1978.

[47] R. Agrawal and R. Srikant, "Privacy-preserving data mining," in *Proceedings of The ACM SIGMOD International Conference on Management of Data (SIGMOD)*, Dallas, Texas, USA, May 2000.

[48] R. Brand, *Microdata Protection through Noise Addition*. Springer Berlin Heidelberg, 2002.

[49] F. Zhang, L. He, W. He, and X. Liu, "Data perturbation with state-dependent noise for participatory sensing," in *Proceedings of The 31st Annual IEEE International Conference on Computer Communications (INFOCOM)*, Orlando, Florida USA, Mar. 2012.

[50] K. L. Huang, S. S. Kanhere, and W. Hu, "Preserving privacy in participatory sensing systems," *Computer Communications*, vol. 33, no. 11, pp. 1266–1280, Jul. 2010.

[51] Y. Wang, D. Xu, X. He, C. Zhang, F. Li, and B. Xu, "L2P2: location-aware location privacy protection for location-based services," in *Proceedings of The 31st Annual IEEE International Conference on Computer Communications (INFOCOM)*, Orlando, Florida USA, Mar. 2012.

[52] H. Zang and J. Bolot, "Anonymization of location data does not work: a large-scale measurement study," in *Proceedings of The 17th Annual International Conference on Mobile Computing and Networking (MobiCom)*, Las Vegas, Nevada, USA, Sep. 2011.

[53] T. Xu and Y. Cai, "Feeling-based location privacy protection for location-based services," in *Proceedings of The 16th ACM Conference on Computer and Communications Security (CCS)*, Chicago, Illinois, USA, Nov. 2009.

[54] V. Rastogi and S. Nath, "Differentially private aggregation of distributed time-series with transformation and encryption," in *Proceedings of The ACM SIGMOD International Conference on Management of Data (SIGMOD)*, Indiana, USA, Jun. 2010.

[55] J. Manweiler, R. Scudellari, and L. P. Cox, "SMILE: encounter-based trust for mobile social services," in *Proceedings of The 16th ACM Conference on Computer and Communications Security (CCS)*, Chicago, IL, USA, Nov. 2009.

[56] M. Bechler, H.-J. Hof, D. Kraft, F. Pählke, and L. Wolf, "A cluster-based security architecture for ad hoc networks," in *Proceedings of The 23rd Conference of the IEEE Communications Society (INFOCOM)*, Hong Kong, China, Mar. 2004.

[57] K. Argyraki, S. Diggavi, M. Duarte, C. Fragouli, M. Gatzianas, and P. Kostopoulos, "Creating secrets out of erasures," in *Proceedings of The 19th International Conference on Mobile Computing and Networking (MobiCom)*, Miami, Florida, USA, Sep. 2013.

[58] J. Heidemann, F. Silva, C. Intanagonwiwat, R. Govindan, D. Estrin, and D. Ganesan, "Building efficient wireless sensor networks with low-level naming," in *Proceedings of 18th ACM Symposium on Operating Systems Principles (SOSP)*, Banff, Alberta, Canada, Oct. 2001.

[59] S. Ozdemir and Y. Xiao, "Secure data aggregation in wireless sensor networks: A comprehensive overview," *Computer Networks*, vol. 53, pp. 2022–2037, 2009.

[60] L. Mottola and G. P. Picco, "MUSTER: Adaptive energy-aware multi-sink routing in wireless sensor networks," *IEEE Transactions on Mobile Computing (TMC)*, vol. 10, no. 12, pp. 1694–1709, 2011.

[61] H. Cam, P. A. Mouallem, and R. E. Pino, "Alert data aggregation and transmission prioritization over mobile networks," *Network Science and Cybersecurity, Advances in Information Security*, vol. 55, pp. 205–220, 2014.

[62] N. Kumar, N. Chilamkurti, and J. J. Rodrigues, "Learning automata-based opportunistic data aggregation and forwarding scheme for alert generation in vehicular ad hoc networks," *Computer Communications*, vol. 39, pp. 22–32, 2014.

[63] C.-M. Chen, Y.-H. Lin, Y.-C. Lin, and H.-M. Sun, "RCDA: Recoverable concealed data aggregation for data integrity in wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems (TPDS)*, vol. 23, no. 4, pp. 727–734, 2012.

[64] D. Westhoff, J. Girao, and M. Acharya, "Concealed data aggregation for reverse multicast traffic in sensor networks: Encryption, key distribution, and routing adaptation," *IEEE Transactions on Mobile Computing (TMC)*, vol. 5, no. 10, pp. 1417–1431, 2006.

[65] A. J. Aviv, M. Blaze, M. Sherr, and J. M. Smith, "Privacy-aware message exchanges for HumaNets," *Computer Communications*, 2014.

[66] E. D. Cristofaro and R. D. Pietro, "Preserving query privacy in urban sensing systems," in *Proceedings of 13th International Conference on Distributed Computing and Networking (ICDCN)*, Hong Kong, China, Jan. 2012.

[67] T. Dimitriou, I. Krontiris, and A. Sabouri, "PEPPeR: A queriers privacy enhancing protocol for participatory sensing," *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol. 107, pp. 93–106, 2012.

[68] F. Qiu, F. Wu, and G. Chen, "SLICER: A slicing-based k-anonymous privacy preserving scheme in participatory sensing," in *Proceedings of The 10th IEEE International Conference on Mobile Ad-hoc and Sensor Systems (MASS)*, Hangzhou, China, Oct. 2013.

Fudong Qiu is a Ph.D candidate in the Department of Computer Science and Engineering, Shanghai Jiao Tong University, P. R. China. He received his B.S. degree in Computer Science and Technology from Xi'an Jiao Tong University in 2012. His research interests include security and privacy preservation in wireless networks. He is a student member of ACM, CCF, and IEEE.

Fan Wu is an associate professor in the Department of Computer Science and Engineering at Shanghai Jiao Tong University, P. R. China. He received his B.S. in Computer Science from Nanjing University in 2004, and Ph.D. in Computer Science and Engineering from the State University of New York at Buffalo in 2009. He has visited the University of Illinois at Urbana-Champaign (UIUC) as a Post Doc Research Associate. His research interests include wireless networking and mobile computing, algorithmic network economics, and privacy preservation. He received Excellent Young Scholar award of Shanghai Jiao Tong University in 2011, and Pujiang Scholar award in 2012. He is a member of ACM, CCF, and IEEE. For more information, please visit http://www.cs.sjtu.edu.cn/~fwu/.

Guihai Chen earned his B.S. degree from Nanjing University in 1984, M.E. degree from Southeast University in 1987, and Ph.D. degree from the University of Hong Kong in 1997. He is a distinguished professor of Shanghai Jiao Tong University, China. He had been invited as a visiting professor by many universities including Kyushu Institute of Technology, Japan in 1998, University of Queensland, Australia in 2000, and Wayne State University, USA during September 2001 to August 2003. He has a wide range of research interests with focus on sensor networks, peer-to-peer computing, high-performance computer architecture and combinatorics. He has published more than 200 peer-reviewed papers, and more than 120 of them are in well-archived international journals such as IEEE Transactions on Parallel and Distributed Systems, Journal of Parallel and Distributed Computing, Wireless Networks, The Computer Journal, International Journal of Foundations of Computer Science, and Performance Evaluation, and also in well-known conference proceedings such as HPCA, MOBIHOC, INFOCOM, ICNP, ICPP, IPDPS and ICDCS.

1536-1233 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.