Contents lists available at ScienceDirect

Computer Communications

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

Privacy-preserving scheme in social participatory sensing based on Secure Multi-party Cooperation

Ye Tian[a], Xiong Li[*,b], Arun Kumar Sangaiah[c], Edith Ngai[d], Zheng Song[e], Lanshan Zhang[f], Wendong Wang[*,a]

[a] The State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, 100876, China
[b] School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan 411201, China
[c] School of Computer science and Engineering, VIT University, Vellore-632014, Tamilnadu, India
[d] The Department of Information Technology, Uppsala University, Uppsala, Sweden
[e] Virginia Tech University, Virginia, US
[f] Beijing Key Laboratory of Network System and Network Culture, Beijing University of Posts and Telecommunications, Beijing, 100876, China

ARTICLE INFO

Keywords:
Social participatory sensing
Privacy-preserving
Secure Multi-party Cooperation

2010 MSC:
00–01
99–00

ABSTRACT

Social participant sensing has been widely used to collect location related sensory data for various applications. In order to improve the Quality of Information (QoI) of the collected data with constrained budget, the application server needs to coordinate participants with different data collection capabilities and various incentive requirements. However, existing participant coordination methods either require participants to reveal their trajectories to the server which causes privacy leakage, or tradeoff the location accuracy of participants for privacy, thereby leading to lower QoI. In this paper, we propose a privacy-preserving scheme, which allows application server to provide quasi-optimal QoI for social sensing tasks without knowing participants' trajectories and identity. More specifically, we first suggest a Secure Multi-party Cooperation (SMC) based approach to evaluate participant's contribution in terms of QoI without disclosing each individual's trajectory. Second, a fuzzy decision based approach which aims to finely balance data utility gain, incentive budget and inferable privacy protection ability is adopted to coordinate participant in an incremental way. Third, sensory data and incentive are encrypted and then transferred along with participant-chain in perturbed way to protect user privacy throughout the data uploading and incentive distribution procedure. Simulation results show that our proposed method can efficiently select appropriate participants to achieve better QoI than other methods, and can protect each participant's privacy effectively.

1. Introduction

The ubiquity of various sensors within smart devices has inspired a new wave of research towards social participatory sensing, which was first discussed in [1]. As a people-centric and spatial-based sensing task, social participatory sensing fully utilizes the idea of crowdsourcing [2]. The major difference between social participatory sensing and traditional sensing lies in the fact that, each participant being regarded as a sensor, called social sensor, sensing the surrounding environment to upload data [3]. A group of mobile users subscribe to an application server and a number of task publishers who publish to the application server both task requirements and corresponding incentive budgets.

Quality-of-information (QoI) is a widely used index to describe publisher's requirement, or evaluate the actual performance of social participatory sensing tasks. Broadly speaking, QoI relates to the ability to judge whether information is fit for use for a particular purpose [2,4,5]. Actually, QoI is characterized by a number of attributes, including sensing locations, sensing time period, required amount of data at each location. Above all, coverage rate and redundancy of sensory data in sensing areas are two key attributes in sensing tasks. Low coverage or high redundancy will lead to deficiency of valid data, and finally affects the performance of social participatory sensing task. An evenly-distributed sensory data with good balance in data coverage and redundancy is of great importance to QoI, thus borrowing the concept which was proposed in [2], we adopt *data utility* to evaluate QoI (detail definition please refer to Section 4.2). To achieve better QoI for sensing tasks under budget constraints, application server needs to coordinate appropriate participants for data collection. Existing approaches

[*] Corresponding author.
E-mail addresses: yetian@bupt.edu.cn (Y. Tian), lixiongzhq@163.com, lixiong@hnust.edu.cn (X. Li), edith.ngai@it.uu.se (E. Ngai), zls326@sina.com (L. Zhang), wdwang@bupt.edu.cn (W. Wang).

http://dx.doi.org/10.1016/j.comcom.2017.10.007
Received 15 January 2017; Received in revised form 12 September 2017; Accepted 7 October 2017

assume that the application server knows the exact locations of all potential participants as a prior condition, it selects a portion of participants to collect more uniformly-distributed sensing data within the incentive budget constraint, avoiding redundant data. Besides, the data tagged with locations are required to be connected with its collector, so that the application server can evaluate the contributions of the selected participants and reward them accordingly.

However, potential privacy disclosure extends far beyond the temptation of incentives, which may prevent part of people from joining social participatory sensing tasks [1]. The key steps in traditional social participatory sensing scenarios [2] can be summarized in five steps: (1) Application server first publishes sensing task with detail requirements, including task area, task time period, required amount of data in each region and incentive budget; (2) Mobile users report their trajectory and requested incentive for each piece of data; (3) Application server selects optimal users as participants according to their data collection capability (how many regions can be covered by his/her trajectory) and requested incentive; (4) Participants report their data with location tag and user ID; (5) Application server evaluates the gathered data and distributes incentives to each participant.

Conspicuously, privacy disclosure and security risks occur in the above steps. First, users' trajectory information may leak in the following aspect: (1) at participant selection stage, application server needs mobile users' trajectory information to compare their coverage on task area, for selecting optimal candidate with best data collecting capability. (2) application server or other third-party server keeps mobile users' ID and IP address for data uploading or incentive distribution. (3) reported sensing data is tagged with location and user ID for incentive distribution. Second, besides the trajectory information, users' sensitive identity information (such as gender, age, income, political tendency and etc.) may leak. An adversary (for example, the application server) can obtain background knowledge from the pieces of reported sensory data (especially for those semantically-rich data, like photo, video and etc), and identify with high confidence the sensitive value of an individual through association rules based background knowledge attack. Third, at incentive distribution stage, it may involve an important security issue, i.e., incentives may be misappropriated by malicious users. Based on the above analysis, the privacy issues are classified into two categories: namely ***visible privacy leakage*** and ***inferable privacy leakage***. More concretely, visible privacy refers to the visible sensitive information which relates to individual's location or trajectory privacy. Correspondingly, inferable privacy refers to those sensitive information about participant's identity which are deduced by adversary through background knowledge attack [6,7].

There has been two kinds of approaches to resolve the conflict between the server's requirement of knowing participants' locations and the participants' requirements of keeping their location private. The first approach assumes that there is a trustful third party (TTP) server, which is responsible for connecting locations and identifications [8,9]. However, this approach relies too much on the TTP as argued by recent approaches [10,11]. Since the TTP knows too much sensitive information of users, it may become the single target of attacks easily. Therefore, most recent solutions are based on the second approach. The main idea is to tradeoff the location accuracy of uploaded data for location privacy. K-anonymity is a representative approach which guarantees that a user is indistinguishable from at least $k - 1$ other users, and widely used in privacy of social network [12,13]. To achieve k-anonymity, a location-based query is submitted to server via a centralized location anonymizer, which enlarges the queried location into a bigger region, geographically covering at least $k - 1$ other users [14,15]. However, not knowing the accurate location of uploaded data may affect the coordination phase and the incentive distribution phase, and cause redundant data collection or misjudgement of participants' uploaded data.

For the concern of inferable privacy leakage, differential privacy is an emerging technology to provide means for maximizing the accuracy of queries from statistical databases while minimizing the chances of identifying records. Differential privacy is most used in situations when a trusted party holds a dataset of sensitive information (e.g., transaction records, medical records, voter registration information, and etc.) with the goal of providing global, statistical information about the data publicly available, while preserving the privacy of users whose information the data set contains. Unlike the situation mentioned above, the objective of application server in participatory sensing is not for providing public accessible data while preserving participants's sensitive information. On the contrary, the application server itself is not a completely trusted party in participants's eyes. So differential privacy, which is designed for providing secure data release mechanism does not fit well of the scenario that participatory sensing focus on.

Motivated by the application scenario proposed in[2], we first propose a privacy-preserving participant selection approach based on Secure Multi-party Cooperation. The basic concept behind such scheme is to replace centralized computation involving participants' sensitive information with distributed cooperation among participants. On the application server side, it iteratively selects participants according to processed non-sensitive data instead of raw location related data, and finally constructs a participants-chain. On participants side, participants jointly compute their own contribution for a participatory sensing task while keeping each one's location and identity information private. In addition, a distributed mechanism for data aggregation and incentive distribution is also designed based on the constructed participants-chain. The proposed scheme can achieve quasi-optimal QoI for sensing task, guarantee the robustness of data collection, and above all, preserve participants' both location privacy and identity privacy. The major contribution of our work is four-fold:

- As far as we know, the proposed privacy-preserving scheme is the first to address both visible privacy and inferable privacy problems of participatory sensing task.
- We propose a secure multi-party approach to calculate participants' data sensing ability cooperatively among candidates, which provides essential decision-making basis for participants selection while keeping visible privacy private.
- We propose a multi-criteria ranking based participant selection algorithm to achieve quasi-optimal quality of sensing task. Participants are selected iteratively by explicitly considering their data sensing ability, incentive requirements and impact on inferable privacy preservation ability.
- We design a distributed mechanism to support data aggregation and incentive distribution that works with the proposed privacy-preserving participant coordination method.

The rest of this paper is organized as follow: Section 2 reviews the related literatures. Section 3 establishes the architecture. Section 4 elaborates the proposed participant selection mechanism based on Secure Multi-party Cooperation and fuzzy multi-criteria ranking. Section 5 analyzes the supporting mechanism for data aggregation and incentive distribution. Section 6 conducts privacy analysis and evaluates the performance of the proposed scheme by simulations using real mobility traces. Finally, Section 7 concludes the paper.

2. RELATED WORK

Privacy-preserving is an important issue in many systems, such as in cloud computing environment, Fu et al. [16] and Xia et al. [17] proposed two efficient privacy-preserving search schemes over encrypted outsourced data, respectively. Wang et al. [35] presented an agent-based model of manipulating prices in finacial markets through spoofing, which provided way to quantify the effect of spoofing on trading behavior and efficiency. Besides the tradeoff between the quality requirement of sensing task and the budget constraint of incentives in most existing works, as Krumm [18] discussed in their work,