Ad Hoc Networks 43 (2016) 56–70

Contents lists available at ScienceDirect

Ad Hoc Networks

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

Credible and energy-aware participant selection with limited task budget for mobile crowd sensing☆

Wendong Wang[a], Hui Gao[b], Chi Harold Liu[c,*], Kin K. Leung[d]

[a] State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China
[b] School of Software Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
[c] School of Software, Beijing Insitute of Technology, Beijing 100081, China and Department of Computer Information and Security, Sejong University, Seoul 143-747, South Korea
[d] Electrical and Electronic Engineering Department, Imperial College, London SW7 2BT, UK

ARTICLE INFO

Article history:
Received 1 September 2015
Revised 8 January 2016
Accepted 1 February 2016
Available online 11 February 2016

Keywords:
Crowd sensing
Incentive
Reputation
Quality of information
Difficulty of task

ABSTRACT

Crowd sensing campaigns encourage ordinary people to collect and share sensing data by using their carried smart devices. However, new challenges that must be faced have arisen. One of them is how to allocate tasks to the most appropriate participants when considering their different incentive requirements and credibility, in order to best satisfy the quality-of-information (QoI) requirements of multiple concurrent tasks, with different, and limited budget constraints. Another challenge is how to maximize participants' rewards to encourage them to contribute sensing data continuously. To this end, in this paper, we first propose a crowd sensing system, that aims to address the above two challenges, where the system considers the benefits of both platform and participants. Then, a participant reputation definition and update method is proposed, that takes participant's willingness and contributed data quality into consideration. Last, we introduce two metrics called "QoI satisfaction" and "Difficulty of Task (DoT)". The former quantifies how much collected sensing data can satisfy the multi-dimensional task's QoI requirements in terms of data quality, granularity and quantity, and the later aids participants to choose proper tasks to maximize their rewards. Finally, we compare our proposed scheme with existing methods via extensive simulations based on a real dataset. Extensive simulation results well justify the effectiveness and robustness of our approach.

1. Introduction

Smart devices, such as smartphones and iPad, are used not only as a means of communication mobile devices of choice, but also as powerful sensing units with a rich set of embedded sensors, such as accelerometer, digital compass, gyroscope, GPS, microphone, camera, etc. [1]. Moreover, smart wearable solutions and even the emerging smart vehicles can enable a new and fast-growing sensing paradigm: the ability to acquire local knowledge through sensor-enhanced mobile devices, e.g., location, personal and surrounding contexts, noise level, traffic conditions, and in the future more specialized information such as pollution, and even the possibility to share this knowledge within the social sphere, healthcare providers, and utility providers [2]. This series of campaign is called *Crowd Sensing* [3], where people/entities who need sensing data are called as "task publishers", and when they request to

☆ This work was supported by National Natural Science Foundation of China (grant nos.61370197 and 61300179), and supported by BUPT Excellent Ph.D. Students Foundation.

* Corresponding author. Tel.: +8613718763233.
E-mail addresses: wdwang@bupt.edu.cn (W. Wang), gaohui786@bupt.edu.cn (H. Gao), chiliu@bit.edu.cn, liuchi02@gmail.com, 36630212@qq.com (C.H. Liu), kin.leung@imperial.ac.uk (K.K. Leung).

http://dx.doi.org/10.1016/j.adhoc.2016.02.007
1570-8705/

collect some types of sensing data, such as noise level data or some photos that we have mentioned above, we refer them as “sensing tasks”, or simply “tasks”, with multiple requirements [4,5]. Ordinary people, who participate in collecting sensing data before any task, are called as “participants”, and we assume that they have advertised their availabilities to the central platform before any task is published (e.g., through periodic information exchange). Normally, there is also a central “platform” to recruit participants, process their reported sensing data and send results back to task publishers.

However, some crowd sensing systems, such as [6] and [7], are prototyped for a single sensing task, or simply: tasks. They do not explicitly consider the co-existence of *multiple* concurrent tasks. On the contrary, approaches such as Song et al. [8] and PRISM [9] can provide sensing data simultaneously for multiple concurrent tasks. On the other hand, as data are sensed by participants' mobile devices, and inevitably this sensing campaign will incur monetary costs, network bandwidth usage, and consume battery lifetime [10]. Therefore, it is necessary to introduce a reward mechanism, normally paid by task publishers, to compensate their costs [4]. To this end, support for multiple sensing tasks with rewards is critical for future crowd sensing systems, and is our research path in its own right.

As some researchers have divided existing incentive strategies into two categories [11], namely, (a) participant-centric approaches focusing on how to encourage participants to continue to contribute sensing data, recruit more participants and improve their motivation [12,13], and (b) platform-centric approaches focusing on how to improve the information gain of the platform and reduce the overall sensing costs [14,15]. Both types of approaches have certain benefits as well as drawbacks, and in this paper, we aim to design a crowd sensing system that considers both the platform and participants, where the platform can receive sensing data with satisfactory quality, and meanwhile participant earns good amount of reward to keep him/her contributing sensing data to the system.

From the platform side, participant selection scheme has always been a major challenge in crowd sensing systems, due to the diversity of participants' incentive requirements and their sensing data quality. Therefore, we aim to find a subset of trustable participants whose data can best satisfy the quality-of-information (QoI) requirements of multiple concurrent tasks with limited task budgets. Here we define a user's reputation to represent his/her past behaviors, by using a *reputation value* to select most credible ones. In this way, we minimize the damage and threat of their dishonest or manipulative behaviors, and protect systems from possible misuses and abuses [16]. On the other hand, broadly speaking, QoI relates to the ability to judge whether information is fit-for-use for a particular purpose [8]. For the purposes of this paper, we assume that QoI is characterized by a number of attributes including the sensing region, data granularity and quantity requirements, and the incentive budget that a task publisher is willing to afford. Furthermore, we employ the economics knowledge to define a novel QoI satisfactory metric, where participant's sensing cost and reputation value are employed to forecast the quality level of sensing data that he/she can contribute.

From the participant side, how to maintain an appreciable number of participants is critical. Imagine that if a task costs too much energy of a participant's smart device, but in the end he/she only earns a little reward that cannot compensate the cost, he/she may not be reluctant to perform the next task, or even quit the crowd sensing campaign. Therefore, in this paper, we introduce a novel metric called “Difficulty of Task (DoT)” to weight the difficulty level of each task, of which we explicitly consider the required sensor types of each task, sensing time slot, and device remaining energy, as the attributes to quantify the DoT level. We allow participants to choose which task to perform, when considering his/her requested reward and the DoT level of a task, by a ratio. It is worth noting that, here we refer “energy-awareness” to minimize the energy cost per unit reward; or in other words, maximizing reward per unit energy cost for selected participants (instead of decreasing the overall used energy), since every task requires fixed sensing duration, fixed amount of sensing data and fixed types of sensors to guarantee data quality even before the task is notified to the crowd. That is, it is impossible to shorten sensing duration or decrease the amount of sensing data, neither does change sensor types. Furthermore, to secure more participants in the campaign, we allow each participant only to perform one task at a time.

The basic work flow of our crowd sensing system is shown in Fig. 1. First, task publishers publish their sensing tasks with affordable budgets to the platform. It is notable that the platform may receive some different tasks, which require sensing the same region nearly at meantime, and the different tasks may require different types of sensors. Especially, each task is associated with certain QoI requirements, calculated by the platform after a task publisher sends the budget and sensing data requirements. Then, the platform selects several participants, based on their locations, credibility and rewards, in the sensing region for each task. After the selected participants choose the tasks to perform according to the DoT value and their rewards of each task, they will collect sensing data and upload to the platform. Finally, the platform processes and sends the sensing data to the task publishers. To ease the task publishers accurately indicate the QoI requirement (i.e., data distribution) from spatial domains, the entire sensing region is considered as to be divided into several equal sized blocks which is mentioned in [4,17]. This is because that if the pieces of sensing data are *all* received from one area block that is only a part of the region, then it is not accurate to represent the whole region. However, if the pieces of sensing data are received from different blocks, then it is more accurate to reflect the entire region. The size of block depends on the nature of tasks, normally sensing data contributed in the same block do not vary enormously (e.g., Mendez et al. argued that environmental variables did not change drastically in short periods of time and space). For example, if a task requests to measure temperature of a 3.9 km^2 region, then it will be divided into around 180 blocks [17]. The size of blocks used in this paper will be provided in detail in Section 7.1.