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/TVT.2015.2490679, IEEE Transactions on Vehicular Technology

Quality-aware Sensing Coverage in Budget Constrained Mobile Crowdsensing Networks

Maotian Zhang, *Student Member, IEEE,* Panlong Yang, *Member, IEEE,*
Chang Tian, *Member, IEEE,* Shaojie Tang, *Member, IEEE,*
Xiaofeng Gao, *Member, IEEE,* Baowei Wang, *Member, IEEE,* and Fu Xiao, *Member, IEEE*

***Abstract*—Mobile crowdsensing has emerged to show elegant capacity in data collection and give rise to numerous applications. In the sense of coverage quality, marginal works have considered the efficient (less cost) and effective (considerable coverage) design for mobile crowdsensing networks. We investigate the optimal quality-aware coverage in mobile crowdsensing networks. The difference between ours and conventional coverage problem is that we only select a subset of mobile users so that the coverage quality is maximized with constrained budget. To address this new problem which is proved NP-hard, we first prove the set function of coverage quality is nondecreasing submodular. By leveraging the favorable property in submodular optimization, we then propose an $(1-\frac{1}{e})$ approximation algorithm with $O(n^{k+2})$ time complexity, where k is an integer that is greater than or equal to 3. Finally, we conduct extensive simulations for the proposed scheme, and the results demonstrate that ours outperforms the random selection scheme and one of the state-of-arts in terms of total coverage quality by at most $2.4\times$ and $1.5\times$, and by averagely $1.4\times$ and $1.3\times$, respectively. Additionally, ours achieves near optimal solution comparing with the brute-force search results.**

***Index Terms*—Mobile crowdsensing networks, coverage, quality-aware sensing, approximation algorithm.**

I. Introduction

The proliferation of smartphones and other mobile devices embedded with a number of sensors (*e.g.*, accelerator and camera), gives rise to a new frontier called *mobile crowdsensing* [1], [2]. In a mobile crowdsensing system, individuals with sensing and computing devices collectively share data and extract information to measure phenomenon of common interest. A growing and important class of crowdsensing systems are designed to provide place-related information and focus on locations that participants routinely visit as targets for collecting and analyzing data. For example, several systems have been developed for reconstructing floor plan using crowdsourced images [3], autonomously naming downtown places [4] or counting the number of individuals of a crowd [5]. To fulfill these system execution quality and efficiency, the sensing data (such as images and audio) should be adequate, available and affordable [6].

Consider an example in downtown places naming [4]. As the task starts, the recruited users will upload the collected sensing data (such as images and audio) of places that they routinely visit (*e.g.*, cafes, shopping malls and offices), to the crowdsensing center. On one hand, the values of various places are different, taking the spatial properties of places into consideration. For example, places like a shopping mall provide more valuable information than an office for a recommendation application [7]. Meanwhile, mobile users cover different number of places along their trajectories. Some of them, such as graduated students, may stay in the campus of most of the working time, while others, such as couriers, may travel to many places during a day [8]. Therefore, how to quantify the quality of mobile users associated with different Places of Interest (POIs) becomes a challenge.

On the other hand, to fulfill the system task, the crowdsensing center should adopt the incentive mechanism to recruit adequate participants [9]. Even worse, there could be a strict budget constraint, which limits the number of selected mobile users. Then, the other challenge is how to make use of the constrained budget to select the qualified mobile users to achieve good coverage. Besides, the areas where the crowdsensing task to monitor, are distributed in an even larger area. Sensing coverage for the relatively remote area would need more participants. There should be a tradeoff between the sensing quality and the number of participants involved.

In response to the aforementioned requirements and challenges, traditional schemes on coverage [10]–[12] and user selection [13] in sensor networks can not work well, due to the specific properties of mobile crowdsensing networks, such as the dynamic conditions of the set of mobile users [1] and the variety of POIs. Nevertheless, very little effort has been

This research is partially supported by NSF China under Grants No. 61190114, 61272487, 61232016, U1405254, 61173136, 61232018, 61502520, 61202024, BK20150030, and CMMI 1436786, PAPD fund, and CICAEET, and the NSF of Shanghai under Grant No. 12ZR1445000, Shanghai Educational Development Foundation under Chenguang Grant No. 12CG09, Shanghai Pujiang Program under Grant No. 13PJ1403900.

M. Zhang and C. Tian are with the College of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China. e-mail: maotianzhang@gmail.com, tianchang163@163.com.

P. Yang (*Corresponding author*) is with the College of Computer Science and Technology, University of Science and Technology of China, Hefei 230027, China. e-mail: panlongyang@gmail.com.

S. Tang is with the Naveen Jindal School of Management, the University of Texas at Dallas, Richardson, TX 75080, USA. e-mail: shaojie.tang@utdallas.edu

X. Gao is with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China. e-mail: gaoxf@cs.sjtu.edu.cn.

B. Gao is with the School of Computer and Software, Nanjing University of Information Science and Technology, China.

F. Xiao is with the Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks and the School of Computer, Nanjing University of Posts and Telecommunications, China. e-mail: xiaof@njupt.edu.cn.

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/TVT.2015.2490679, IEEE Transactions on Vehicular Technology

devoted to this new field. One systematic study of the coverage properties of place-centric crowdsensing is conducted by Chon *et al.* [6]. In their study, they discovered some properties of mobile crowdsensing networks, *e.g.*, the number of places covered by participants follows a power-law distribution. In terms of the problem of coverage, a relatively related work was presented by Wang *et al.* [14], in which the coverage of photos obtained via crowdsourcing was investigated. However, the mobile users are opportunistic, dynamic and sociable, which is different from the photos.

In this paper, we design an effective and efficient scheme to achieve considerable sensing coverage in a budget constrained mobile crowdsensing network. We first propose *coverage quality*, to evaluate the quality of the mobile user. It measures how much weight of POIs to be covered by mobile users.

Then, we deal with challenges brought by the budget constraint, which is referred to as the maximum coverage quality with budget constraint problem, *i.e.*, given a set of mobile users and POIs, how to select a subset of mobile users such that the coverage quality is maximized under the constraint of a predefined budget? To address this problem which is proved NP-hard, we first prove the submodularity of the set function of the coverage quality. And then, leveraging the favorable properties in submodular optimization, we present an $(1-\frac{1}{e})$ approximation algorithm. Finally, we conduct the simulations to evaluate the performance of our algorithms compared to other selection schemes.

Particularly, the contributions of this work are summarized as follows:

- We present the maximum coverage quality with budget constraint problem in the mobile crowdsensing networks. To address this NP-hard problem, we first prove that the set function of coverage quality is nondecreasing submodular. Then, employing the favorable properties in submodular optimization, we design an effective, efficient and near optimal algorithm, where the approximation ratio is $(1-\frac{1}{e})$ and the time complexity is $O(n^{k+2})$ in which k is an integer that is greater than or equal to 3.
- We evaluate the proposed scheme with extensive simulations. The results of random-walk-based trace-driven simulations show that our algorithm achieves the better performance than the random selection scheme and one of the state-of-arts [13] in terms of total coverage quality by at most $2.4\times$ and $1.5\times$, and by averagely $1.4\times$ and $1.3\times$, respectively. Besides, when using real-world mobility traces, ours also outperforms the other methods.

The rest of the paper is organized as follows. Section II describes the scenario and system model. The definition and solution of the maximum coverage quality with budget constraint problem are presented in Section III. In Section IV, we use simulations to evaluate our algorithms. Section V introduces the related work. Finally, we conclude our work in Section VI.

II. System Model

In this section, we describe the scenario that we focus on and present the system model. We also introduce the definitions and denote frequently used notations.

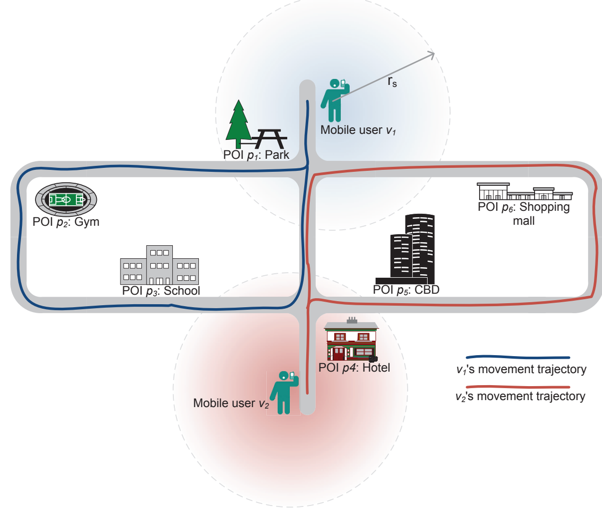


Fig. 1. Scenario and model.

A. Scenario and Model Description

We consider that, the crowdsensing center distributes the task that collects the information of the logical *Places of Interest* (POIs) (such as a restaurant and a user's home) to the mobile users who own sensor-enabled mobile devices. The set of POIs are denoted by $P=\{p_1,p_2,...,p_m\}$, where m is the number of POIs. Let $\vec{P}=\{\vec{p}_1,\vec{p}_2,...,\vec{p}_m\}$ be the locations of P, where each single POI includes a 2D location information $\vec{p}_i=\{x_i,y_i\}$. Particularly, the crowdsensing center allocates a predefined weight w_{p_i} for the POI p_i, taking the various values of different POIs into consideration, and correspondingly, the set of weights for P is $W=\{w_{p_1},w_{p_2},...,w_{p_m}\}$. Namely, we consider the spatial properties of POIs, to improve the sensing coverage. The weight of a POI indicates how much useful information can be provided by this POI. For example, places like a shopping mall or a restaurant could provide more valuable information than a residence or a workplace, for most of crowdsensing applications, such as Jigsaw [3] and a place naming system presented in [4].

Meanwhile, we assume that there are n candidate users, and the user set is given by $V=\{v_1,v_2,...,v_n\}$. The locations of user v_j could be denoted as $\vec{l}_{v_j}=\{x_{v_j},y_{v_j}\}$. Each mobile user has an arbitrary movement trajectory, and visits different places along the trajectory. Once the mobile user is close enough to the POI, we consider that a POI is properly covered if it could be identified (or detected) by analyzing multi-modal sensing data (such as images and audio) [6]. The sensing data could be directly uploaded by participants through cellular networks. When the monitoring areas are not covered by cellular network (e.g., 3G), the sensing data could be forwarded using opportunistic networks in a delay tolerant manner via short-range communications, such as Bluetooth and WiFi. Formally, let r_s^j be the user v_j's sensing range, and we regard that a POI p_i is covered by a mobile user v_j if and only if the POI is in the sensing range of this mobile user, *i.e.*,

$$||\vec{l}_{v_j}-\vec{p}_i||\le r_s^j$$

. Let $f_{v_j}(p_i)$ represent whether a mobile user v_j covers a POI

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/TVT.2015.2490679, IEEE Transactions on Vehicular Technology

p_i or not, *i.e.*,

$$f_{v_j}(p_i) = \begin{cases} 1 & p_i \text{ is covered by } v_j \\ 0 & \text{otherwise} \end{cases}$$

We note that $f_{v_j}(p_i)$ is independent of how many times a POI p_i is covered by v_j. The set of POIs covered by the user v_j is given by $C_{v_j} = \{p_i : f_{v_j}(p_i) = 1, \forall p_i \in P\}$. Accordingly, the set of POIs covered by all mobile users V is $C_V = \bigcup_{j=1}^{n} C_{v_j}$.

Fig. 1 illustrates the scenario and our model. $\{p_1, p_2, p_3\}$ and $\{p_4, p_5, p_6\}$ are covered by two mobile users v_1 and v_2, respectively. The crowdsensing center could select these two mobile users to collect the information (such as sensing data and location-based social media data) from those POIs [6].

B. Coverage Quality

Different from traditional sensor networks [15] [16], the users with mobile sensing devices are endowed with uncertain but better mobility in a mobile crowdsensing network, which imposes new challenges on assessing the quality of the mobile user in terms of coverage. For instance, for a place naming system [4], the platform wants to name a variety of urban places online, and so it prefers to recruit the mobile users who are capable of covering more distinct places. For a location-based recommendation system [17], places such as a shopping mall and a cafe provide more recommendable information, and so, the mobile users who often visit such places tend to be selected by the mobile crowdsensing platform.

In dealing with above challenges, we measure the quality of the mobile user in terms of coverage by *coverage quality*. The coverage quality of each mobile user is independent, only associated with the movement trajectory in which the mobile user could visit different POIs. The formal definition of coverage quality is defined as follows.

Definition 1: (Coverage Quality): Given a POI p_i and a mobile user v_j, the coverage quality of v_j on p_i, denoted by $U_{v_j}(p_i)$, is the weight of p_i covered by v_j, i.e., $U_{v_j}(p_i) = w_{p_i} f_{v_j}(p_i)$.

Then, the coverage quality of a set of mobile users $V^{'} = \{v_j : 1 \leq j \leq k\}$ respecting to POI p_i is the total weights of p_i covered by the mobile users of $V^{'}$, i.e., $U_{V'}(p_i) = \sum_{j=1}^{k} U_{v_j}(p_i)$.

Finally, the total coverage quality of the mobile users regarding all POIs $P = \{p_1, p_2, ..., p_m\}$ is the sum of the coverage quality respecting to each POI. By dividing the total number of POIs, it is normalized as $U_{V'}(P) = \frac{1}{m}\sum_{i=1}^{m} U_{V'}(p_i)$.

For example in Fig.1, there are two participants v_1 and v_2 with different movement trajectories respectively, and six places of interest $P = \{p_1, p_2, p_3, p_4, p_5, p_6\}$ with the weights $\{w_{p_1}, w_{p_2}, w_{p_3}, w_{p_4}, w_{p_5}, w_{p_6}\}$. The coverage quality of v_1 is $U_{v_1}(C_{v_1}) = \frac{1}{3}(w_{p_1}+w_{p_2}+w_{p_3})$, which is different with that of v_2, $U_{v_2}(C_{v_2}) = \frac{1}{3}(w_{p_4}+w_{p_5}+w_{p_6})$. Then, the total coverage quality of these two mobile users regarding all six POIs is, $U_{v_1,v_2}(C_{v_1} \cup C_{v_2}) = \frac{1}{6}(w_{p_1}+w_{p_2}+w_{p_3}+w_{p_4}+w_{p_5}+w_{p_6})$.

III. Maximum Coverage Quality with Budget Constraint

In this section, we describe the maximum coverage quality with budget constraint problem, and present corresponding algorithm to address it.

A. Problem Statement

In practical scenarios, the budget of the crowdsensing center to recruit participants is limited, determining the number of mobile users that could be selected. Intuitively, the budget could be directly represented by the reward. The crowdsensing center should pay some rewards to motivate participants for sensing tasks. We assume the set of mobile users $V = \{v_1, v_2, ..., v_n\}$ is associated with different costs $\{b(v_j)\}_{j=1}^{n}$. The total cost of selecting mobile users should not exceed a given budget B. We define the problem as follows.

Definition 2: (Maximum Coverage Quality with Budget Constraint Problem): Given a set of m POIs $P = \{p_1, p_2, ..., p_m\}$ and n mobile users $V = \{v_1, v_2, ..., v_n\}$, and also given a predefined budget $B(> 0)$, the maximum coverage quality with budget constraint problem asks for a subset S ($S \subseteq V$), such that the total cost of S is less than B, i.e., $b(S) = \sum_{v_j \in S} b(v_j) \leq B$, and the total coverage quality of the selected users' $U_S(P)$ is maximized.

Formally, the optimization problem is given by:

$$\max_{S \subseteq V} U_S(P), \text{ subject to } b(S) \leq B.$$

Solution overview: In tackling the problem, we first claim that the problem could be transformed to the budget maximum coverage problem which is proved NP-hard [18]. Then, we prove that the set function of the total coverage quality $U_S(P)$, given the set of selected mobile users S and the set of POIs P, is nondecreasing submodular. Finally, leveraging the favorable properties of submodular optimization [19], we present an $(1-\frac{1}{e})$ approximation algorithm for the problem.

B. Conversion to Budgeted Maximum Coverage

Refer to [18], the budgeted maximum coverage problem is defined as follows. A collection of sets $\mathcal{S} = \{S_1, S_2, ..., S_n\}$ with associated costs $\{b_i\}_{i=1}^{n}$ is defined over a domain of elements $X = \{x_1, x_2, ..., x_m\}$ with associated weights $\{w_i\}_{i=1}^{m}$. The goal is to find a collection of sets $\mathcal{S}^{'} \subseteq \mathcal{S}$, such that the total cost of elements in $\mathcal{S}^{'}$ does not exceed a given budget B and the total weight of elements covered by $\mathcal{S}^{'}$ is maximized.

To transform to above problem, without loss of generalization, given a POI set P, we first consider a single mobile user v_j that covers a subset of POIs $C_{v_j} = \{p_i : f_{v_j}(p_i) = 1, \forall p_i \in P\}$. Then, a collection of cover set for all mobile users is $\mathcal{C} = \{C_{v_1}, C_{v_2}, ..., C_{v_n}\}$ with associated costs $\{b(v_j)\}_{j=1}^{n}$. The coverage quality of each user U_{v_i} is associated with the predefined weight of POIs, in which a domain of elements (*i.e.*, POIs) $P = \{p_1, p_2, ..., p_m\}$ is associated with $W = \{w_{p_1}, w_{p_2}, ..., w_{p_m}\}$. Accordingly, the simple corresponding relation between these two set systems is $\mathcal{C} \longrightarrow \mathcal{S}$, $P \longrightarrow X$, and the targeted set $S \longrightarrow \mathcal{S}^{'}$, respectively. Therefore, the proposed problem could be reduced to the budgeted maximum coverage problem, which was proved to be NP-hard [18].

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/TVT.2015.2490679, IEEE Transactions on Vehicular Technology

C. Submodularity

To solve above problem, we prove that the set function of the coverage quality is nondecreasing submodular, such that we could leverage the favorable property in submodular optimization. Before presenting the proof of the submodularity, we first provide preliminary knowledge on the submodular set function.

Definition 3: Given a finite set E, a real-valued function $f(\cdot)$ on the set of subsets of E is called submodular if

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B), \forall A, B \subseteq E$$

We often make use of the incremental value of adding element u to a set A, *e.g.*, let $\rho_u(A) = f(A \cup u) - f(A)$. Besides, if the function satisfies the *diminishing returns* rule, it is called submodular. That is, the difference from adding an element u to a subset A is at least as large as the one from adding the same element to a superset B of A:

$$f(A \cup \{u\}) - f(A) \geq f(B \cup \{u\}) - f(B)$$

for all element u and all pairs $A \subseteq B$. Furthermore, $f(\cdot)$ is said to be *nondecreasing* if $f(A) \leq f(B), \forall A \subseteq B \subseteq E$. Lovász has shown that submodular functions could be understood as set functions with convexity [20].

Based on above preliminaries, we acquire the following lemma and further give its proof.

Lemma 1: Given a set of POIs P and a subset of mobile users S ($S \subseteq V$), the set function of the coverage quality $U_S(P)$ is nondecreasing submodular.

Proof: It is straightforward that $U_\emptyset(P) = 0$. Consider V's two arbitrary subsets S and S', $S \subseteq S' \subseteq V$, we have $U_S(P) \leq U_{S'}(P)$ (the equality holds if $C_{S'-S} = \emptyset$) and so $U_S(P)$ is nondecreasing.

Consider a mobile user $u \in V - S'$, $|C_{S\cup\{u\}}| - |C_S|$ is the number of covered POIs in C_u that are not in the C_S. we note that $\bigcup_{v_j \in S'} C_{v_j}$ is at least as large as $\bigcup_{v_j \in S} C_{v_j}$. The coverage quality $U_S(P)$ is associated with the sum of weights of POIs in C_S, we then have

$$U_{S\cup\{u\}}(P) - U_S(P) \geq U_{S'\cup\{u\}}(P) - U_{S'}(P)$$

. It is satisfied with the *diminishing returns* rule in which the difference from adding an element to a set S is at least as large as the one from adding the same element to a superset S' of S [21], and hence $U_S(P)$ is submodular. As a result, $U_S(P)$ is nondecreasing submodular with $U_\emptyset(P) = 0$. ■

D. Approximation Algorithm

Motivated by the submodular property of coverage quality [19], we propose an approximation algorithm to address the problem with guaranteed performance, as Algorithm 1 shows. The algorithm uses enumeration technique partially and then leverages greedy heuristic, so as to obtain better approximation ratios and output the candidate solution having the greatest coverage quality.

We note that let G be a subset of V, and k be some fixed integer in Algorithm 1. We consider all subsets of V of cardinality of k which have cost at most B, and each subset is completed to a candidate solution using greedy heuristic. Another set of candidate solution contains all subsets of V of cardinality less than k which have cost at most B. Concretely, Algorithm 1 first enumerates all subsets of up to k users for some constant $k > 0$, and then complements these subsets using the modified greedy algorithm (line 4-9).

Algorithm 1 Approximation Algorithm for Maximum Coverage Quality with Budget Constraint Problem

Input:
PoI set P and its weight set W, candidate user set V, and a predefined budget B

Output:
The target subset $S, S \subseteq V$

1: $H_1 \leftarrow \arg\max \{U_G(P)$, such that $G \subseteq V, |G| < k$, and $b(G) \leq B\}$; $H_2 \leftarrow \emptyset$
2: For all $G \subseteq V$, such that $|G| = k$, and $b(G) \leq B$ do
3: $\quad M \leftarrow V \setminus G$
4: $\quad$ Repeat
5: $\quad\quad$ select $v_j \in M$ that maximizes $\frac{U_{v_j}(C_{v_j} - C_{v_j} \cap C_G)}{b(v_j)}$
6: $\quad\quad$ if $b(G) + b(v_j) \leq B$ then
7: $\quad\quad\quad G \leftarrow G \cup \{v_j\}$
8: $\quad\quad M \leftarrow M \setminus \{v_j\}$
9: $\quad$ Until $M = \emptyset$
10: $\quad$ if $U_G(P) > U_{H_2}(P)$ then $H_2 \leftarrow G$
11: End For
12: If $U_{H_1}(P) > U_{H_2}(P)$, $S \leftarrow H_1$, otherwise, $S \leftarrow H_2$

Theorem 1: For $k \geq 3$, Algorithm 1 achieves an approximation ratio of $(1 - \frac{1}{e})$ for the maximum coverage quality with budget constraint problem, *i.e.*,

$$U_S \geq (1 - \frac{1}{e})U_{OPT} \quad \text{for} \quad k \geq 3$$

where U_{OPT} is the optimal value of the total coverage quality that can be achieved by any user set S.

Proof: The coverage quality derives from the weight of POIs covered by selected users, and it is maximized *iff* the corresponding subsets has the maximum total weight. According to the theorem presented in [18], the subsets selected by Algorithm 1 yield a total weight that is at least $(1 - \frac{1}{e})$ times the optimal value for $k \geq 3$. As a result, the total coverage quality of the selected user set has a lower bound by $(1 - \frac{1}{e})$ times the maximum total coverage quality. (The reader could refer [18] for detailed proof.) ■

Complexity analysis: At first, straightforwardly, the enumeration of all subsets of V (line 2) takes $O(|V|^k)$ time. Then, using the modified greedy selection (line 4-9) takes $O(|V|^2)$ time. Since the greedy selection is embedded into the enumeration part, associated with these two parts, the time complexity of the algorithm is given by:

$$O(|V|^k \times |V|^2) = O(|V|^{k+2}) = O(n^{k+2})$$

. Therefore, the time complexity of our algorithm is $O(n^{k+2})$, where k is an integer that is greater than or equal to 3. For Algorithm 1, If sets at least up to cardinality k = 3 are enumerated, it achieves an approximation guarantee of $(1-\frac{1}{e})$. However, the complexity of the proposed algorithm increases with increasing k.

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/TVT.2015.2490679, IEEE Transactions on Vehicular Technology

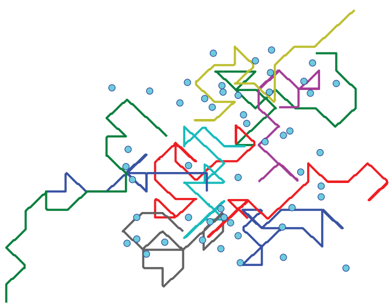


Fig. 2. Simulation scenario. The colored lines and the cyan circles represent the trajectories of mobile users and the POIs, respectively (ten mobile users and fifty POIs).

IV. Performance Evaluation

In this section, we evaluate the performance of proposed algorithm, and present the results of trace-driven simulations using both synthetic random-walk-based trace and real-world mobile traces. Moreover, we make analysis on these evaluation results.

A. Random-walk based Simulation

1) Simulation Settings: We create a simulator in Matlab to generate synthetic random-walk based mobility traces. We set different numbers of mobile users and POIs in a $3000m \times 3000m$ square field. POIs are assumed to be randomly distributed in the field. Each mobile user performs the random walk. The walking speed is set to 1.5m/s, and the time interval for direction changing is set to 100s. For simplicity, when the POI is in the user's sensing range, say, $r_s = 45m$ (*e.g.*, detecting a noise source) [10], this POI is recorded by the user. The traces record each user's ID and the covered POIs. Fig. 2 illustrates our simulation scenario, in which 10 trajectories (the colored lines) are generated by random walk of the users, and 50 POIs (labeled with cyan circles) are uniformly distributed in the setting area. In the simulations, we randomly allocate the weight to each POI, as well as the cost to each mobile user, according to a uniform distribution. We obtain the simulation results by averaging the evaluation results in 50 different instances.

We use the random selection algorithm (denoted by Random) and the method proposed by Kuo et al (denoted by Kuo's) [13], for comparison. In the random selection algorithm, the mobile user are randomly selected into the target user set. In Kuo's scheme, due to the connection constraint, we set the $r_c = 70m$ (a typical communication range for commonly mobile devices) as the communication range of each mobile user in the simulation scenario. Two mobile user are connected as long as they come into each other's communication range. When performing our algorithm (denoted by Proposed) in the simulation, the parameter k (defined in Algorithm 1) is set to be 3, for enhancing the efficiency of the algorithm.

2) Simulation Results: The primary concern we need to evaluate is the impact of the budget constraint B on the total coverage quality. Intuitively, the bigger the given budget, the larger the total coverage quality. We note that the budget is normalized for better presentation. As Fig. 3 shows, both our algorithm and the random selection achieve more coverage

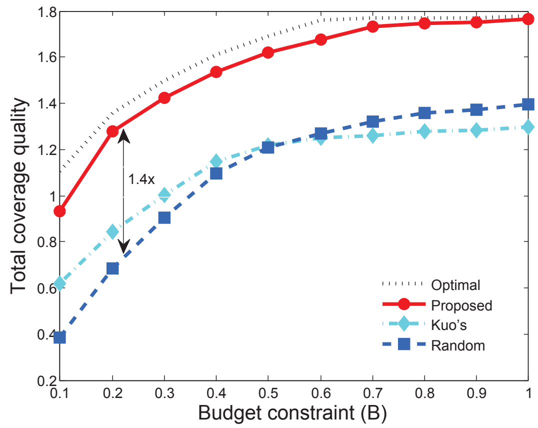


Fig. 3. Impact of the budget, B, on the total coverage quality. ($n = 10$ and $m = 50$)

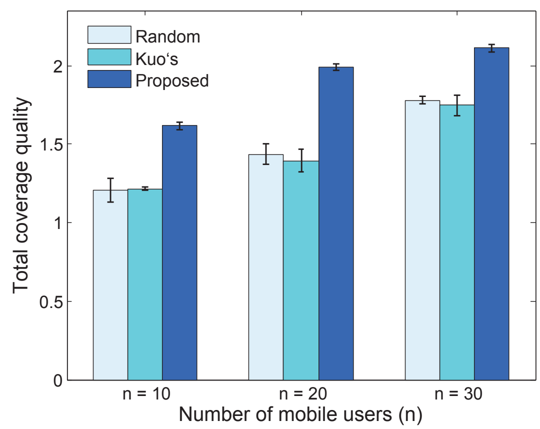


Fig. 4. Impact of the number of mobile users, n, on the total coverage quality. ($B = 0.5$ and $m = 50$)

quality as B increases. For better comparison, we also show the results of brute force selection (denoted by Optimal) as a upper bound of the total coverage quality. The difference between the total coverage quality achieved by ours and the best achievable level is small, say, 4.1% on average. The performance improvement achieved by our algorithm is significant compared to the random selection and Kuo's method. That is, our algorithm outperforms the random selection and Kuo's method in terms of total coverage quality by at most 140.4% and 50.2%, and by averagely 40.6% and 38.9%, respectively. The connection constraint in Kuo's method limits the amount of selectable mobile users, resulting in the decrease of total coverage quality, compared to our scheme.

We then evaluate the impacts of the number of mobile users n on the total coverage quality. While other factors including budget ($B = 0.5$ and $m = 50$) remain unchanged, the total coverage quality increases with increasing the number of mobile users, as shown in Fig. 4. However, the increase gap is becoming small as the number of mobile users increases. Again, our algorithm outperforms the random selection and Kuo's scheme by at most 38.9% and 60.7%, and by 27.4% and 33.2% on average, respectively. Besides, Fig. 5 shows how the

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/TVT.2015.2490679, IEEE Transactions on Vehicular Technology

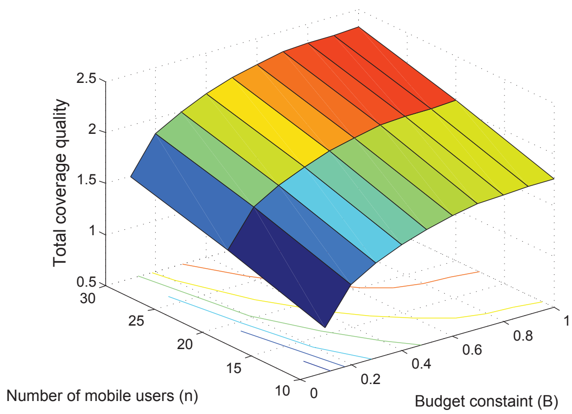


Fig. 5. Joint impact of both the budget B and the number of mobile users n on the total coverage quality. ($m = 50$)

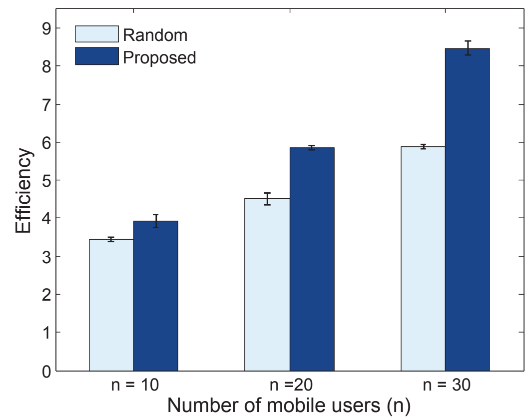


Fig. 7. Impact of the number of mobile users, m, on the efficiency. ($B = 0.5$ and $m = 50$)

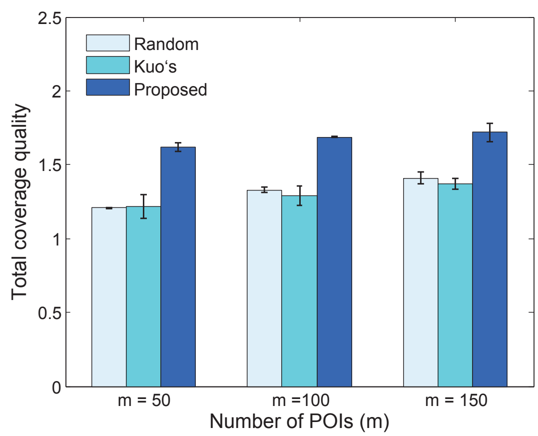


Fig. 6. Impact of the number of POIs, m, on the total coverage quality. ($B = 0.5$ and $n = 10$)

total coverage quality changes as the budget and the number of mobile users jointly increase. The total coverage quality escalates to the maxima with increasing both the budget and the number of mobile users.

Moreover, the impacts of the number of POIs m on the total coverage quality is evaluated. We set the budget $B = 0.5$ and the number of mobile users $n = 10$. From Fig. 6, we observe that the total coverage quality increases slightly as the number of POIS increases, using the three comparative methods. Still, our algorithm achieves better performance than both the random selection and Kuo's method.

Finally, we introduce a metric to measure the efficiency of different selection schemes. Given the set of selected mobile users S and POIs P, the efficiency η is defined as the ratio between the coverage quality $U_S(P)$ achieved by S and the sum of cost $b(S)$ for recruiting S, *i.e.*, $\eta = \frac{U_S(P)}{b(S)}$. As shown in Fig. 7, for both our algorithm and the random selection, the efficiency increases with increasing the number of mobile users. Our algorithm achieves bigger improvement of the efficiency than the random selection, by at most 44.1% and 32.4% on average. The rationale lies in that, recalling Algorithm 1 (specially, line 5), it greedily selects the mobile user whose efficiency is maximized into the target user set, and hence the total efficiency is maximized.

B. Real-world Trace-driven Simulation

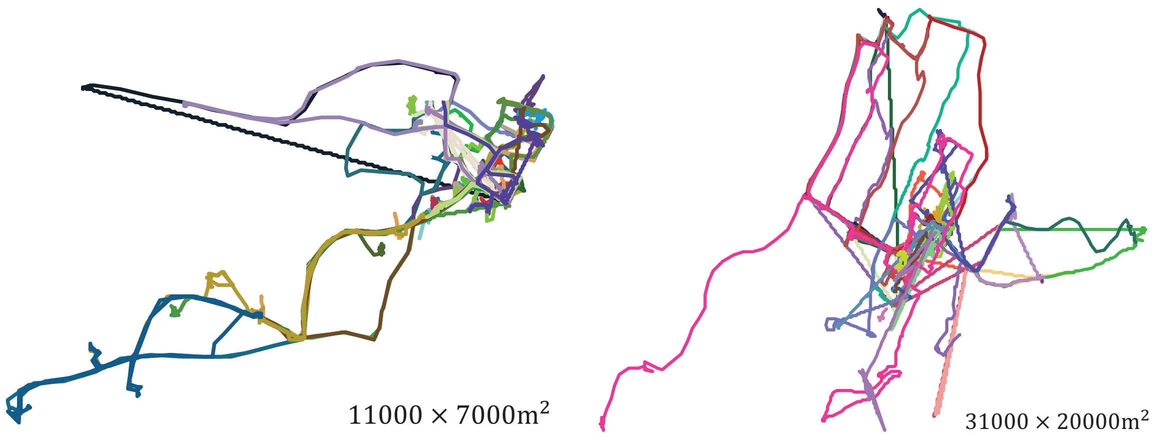
1) Simulation Settings: To further evaluate our proposed user selection method, we conduct real-world trace-driven simulations based on two mobile datasets collected from NCSU and New York city (denoted by NCSU and NewYork respectively) [22] [23]. In NCSU and NewYork datasets, there are respectively 35 and 39 daily trajectories that are collected by recruited participants who are equipped with GPS hand-held receivers. In the simulations, we map these two mobility traces into two dimensional regions, as Fig. 8.

To perform our algorithm in trace-driven simulations, we capture POIs from the Google Map based on the constructed regions and the real sites of NCSU campus and New York city (concretely, Manhattan and its vicinity). In particular, to measure the impact of distinct POI, we allocate the weight for each POI according to two different distributions, namely, uniform and power-law distributions. We note that [6] have revealed that the relation between the places of interest and the number of mobile users followed a power law distribution. Furthermore, we compare our algorithm with the random selection scheme. The simulation results are obtained by averaging the evaluation results in 50 different instances.

2) Simulation Results: We first measure the impact of the budget, B, on the total coverage quality. The number of POIs are $m = 50$ and $m = 100$ in NCSU and NewYork dataset, respectively. As shown in Fig.9, the total coverage quality increases with increasing budget for these two approaches. Besides, our algorithm outperforms the random selection in terms of the total coverage quality. When the budget is low (< 0.5), the total coverage quality is bigger when using power law distribution than that when using uniform distribution, especially for our algorithm. The reasons are that the majority of weight is dominated by few POIs when using power law distribution, and our algorithm tends to select the mobile users who visit these POIs.

We then evaluate the impact of the number of POIs, m, on the total coverage quality, where the budget B and the number

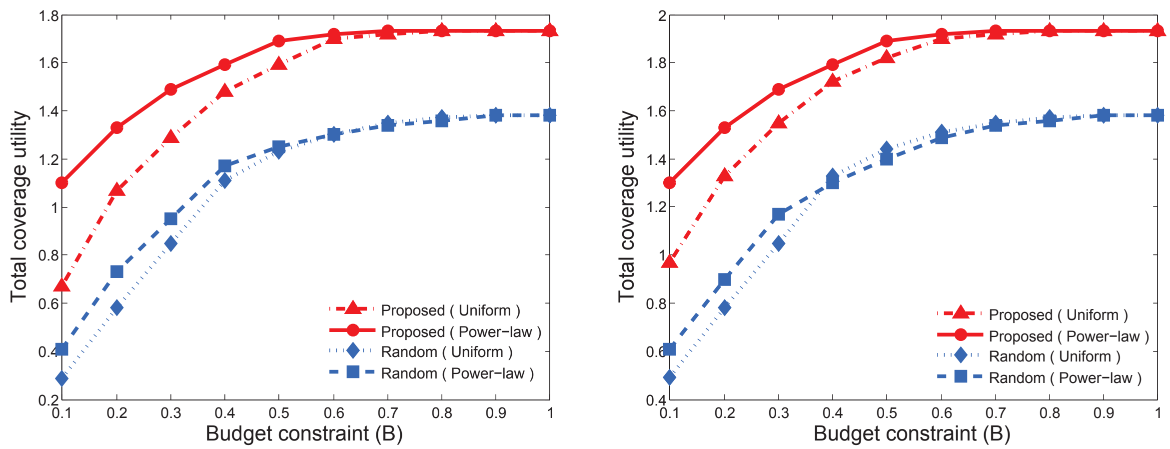
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2490679, IEEE Transactions on Vehicular Technology



(a) 35 trajectories extracted from NCSU dataset

(b) 39 trajectories extracted from NewYork dataset

Fig. 8. Mapping trajectories extracted from mobile datasets NCSU and NewYork into a two dimensional region.



(a) Using NCSU dataset(Number of POIs $m = 50$)

(b) Using NewYork dataset(Number of POIs $m = 100$)

Fig. 9. Impact of the budget, B, on the total coverage quality using mobile datasets NCSU and NewYork respectively.

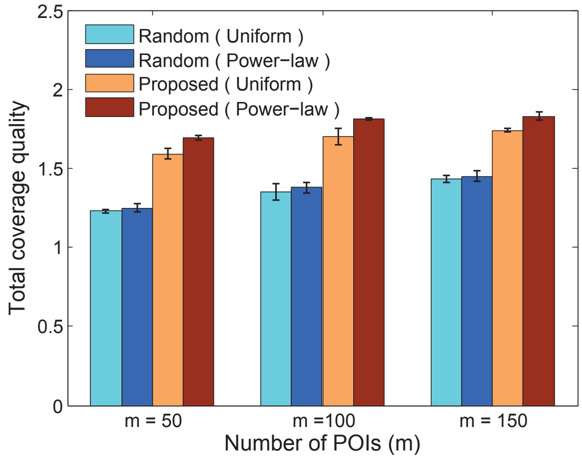
of mobile users n are set to be 0.5 and 10. The results in Fig.10 are similar to that in Random-walk based simulations. The proposed algorithm achieves better performance than the random selection for using both NCSU and NewYork datasets.

V. Related Work

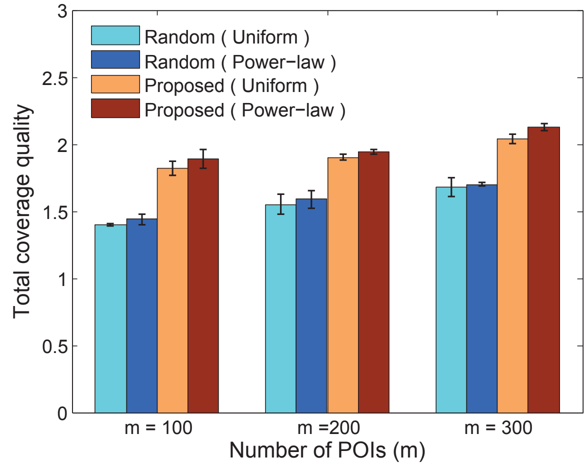
Mobile crowdsensing [1], [2], [24]–[27] is a promising frontier, where individuals with sensing and computing devices (*e.g.*, smartphone, tablet and in-vehicle sensor) collectively share data and extract information to measure phenomenon of common interest. A growing class of mobile crowdsensing systems are place-centric, designed to provide place-related information. Meanwhile, previous localization techniques [28], [29] help us target the places that provide fruitful information. A number of prototypes have been developed, for example, Jigsaw has been designed for reconstructing floor plan [3], and SecureFind is a privacy-preserving object finding system via mobile crowdsensing [30]. Another instances include an autonomous place naming system [4] and a crowd counting application [5].

One fundamental problem in mobile place-centric crowdsensing is how to select the qualified users to achieve considerable coverage of places of interest. Also, coverage is a key issue associated with the quality of sensing [31]. Coverage is also one of the most import issues (*e.g.*, routing [32] and data delivery [33]) in traditional wireless sensor networks. Despite numerous researches on traditional coverage (such as area coverage [10], [12], [34]–[36] and barrier coverage [11], [37], [38]) and user selection [13] [39], very little effort has been devoted to this new field. One systematic study of the coverage properties of place-centric crowdsensing has been conducted by Chon *et al.* [6]. Based on the deployment experiences and analysis of the collected dataset, they revealed the relationship between the user population and coverage of places of interest, *i.e.*, the number of place covered by participants follows a power-law distribution. Their work inspires us to study the maximum coverage utility and user selection problem for mobile place-centric crowdsensing. In terms of the problem of coverage, a relatively related work is SmartPhoto [14], in which the coverage of photos obtained via crowdsourcing was investigated, and three optimization problems including

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/TVT.2015.2490679, IEEE Transactions on Vehicular Technology



(a) 35 trajectories extracted from NCSU dataset



(b) 39 trajectories extracted from NewYork dataset

Fig. 10. Impact of the number of POIs, m, on the total coverage quality using mobile datasets NCSU and NewYork respectively(the budget $B = 0.5$).

maximum photo utility, online maximum photo utility, and minimum selection problem have been studied. However, our scheme focuses on selecting qualified sensing data with budget limitations. Specifically, this work proposes a quality-aware design, where the covered areas are evaluated with weights. Zhang *et al.* [40] have proposed a participants selection framework named CrowdRecruiter for mobile crowdsensing. CrowdRecruiter aims to minimize incentive payment while satisfying probabilistic coverage constraint. In contrast, we aim to maximize the place-centric sensing coverage with a predefined budget constraint. Zhao *et al.* [41], recently have proposed an energy-efficient opportunistic coverage scheme for people-centric urban sensing [42]. In their work, the human mobility features were exploited to construct an opportunistic coverage model and design an offline user selection scheme. While their work cares more about the trade-off between energy consumption and coverage, ours explores both the diversity of places of interest and the spatial property of human involved.

VI. Conclusion

In this paper, we investigated the coverage problem in the mobile crowdsensing network. In particular, we studied how to select a set of mobile users so that we could maximize the total coverage quality with budget constraint. In tackling the problem, we have proved the submodularity of the set function of coverage quality with respect to the weights of POIs covered by mobile users. Then, we presented an $(1 - \frac{1}{e})$ approximation algorithm for the problem. Finally, we conducted trace-driven simulations with synthetic random-walk-based mobility traces to evaluate the performance of our algorithm. The simulation results validated our algorithms, and showed that it outperforms the random selection algorithm and the state-of-art.

In the future, we plan to study how to leverage the temporal-spatial properties of both the mobile users and the POIs to improve the coverage of mobile crowdsensing networks. Moreover, we would like to conduct the systematic implementation using our algorithm.

References

[1] R. K. Ganti, F. Ye, and H. Lei, "Mobile crowdsensing: Current state and future challenges," *IEEE Commun. Mag.*, vol. 49, no. 11, pp. 32–39, 2011.

[2] M.-R. Ra, B. Liu, T. F. La Porta, and R. Govindan, "Medusa: A programming framework for crowd-sensing applications," in *Proc. ACM MobiSys*, pp. 337–350, 2012.

[3] R. Gao, M. Zhao, T. Ye, F. Ye, Y. Wang, K. Bian, T. Wang, and X. Li, "Jigsaw: indoor floor plan reconstruction via mobile crowdsensing," in *Proc. ACM MobiCom*, pp. 249–260, 2014.

[4] Y. Chon, Y. Kim, and H. Cha, "Autonomous place naming system using opportunistic crowdsensing and knowledge from crowdsourcing," in *Proc. ACM IPSN*, pp. 19–30, 2013.

[5] P. G. Kannan, S. P. Venkatagiri, M. C. Chan, A. L. Ananda, and L.-S. Peh, "Low cost crowd counting using audio tones," in *Proc. ACM SenSys*, pp. 155–168, 2012.

[6] Y. Chon, N. D. Lane, Y. Kim, F. Zhao, and H. Cha, "Understanding the coverage and scalability of place-centric crowdsensing," in *Proc. ACM Ubicomp*, pp. 3–12, 2013.

[7] L. Deng and L. P. Cox, "Livecompare: grocery bargain hunting through participatory sensing," in *Proc. ACM HotMobile*, pp. 4–9, 2009.

[8] B. Han and A. Srinivasan, "Your friends have more friends than you do: identifying influential mobile users through random walks," in *Proc. ACM MobiHoc*, pp. 5–14, 2012.

[9] D. Yang, G. Xue, X. Fang, and J. Tang, "Crowdsourcing to smartphones: incentive mechanism design for mobile phone sensing," in *Proc. ACM MobiCom*, pp. 173–184, 2012.

[10] R. Tan, G. Xing, B. Liu, J. Wang, and X. Jia, "Exploiting data fusion to improve the coverage of wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 20, no. 2, pp. 450–462, 2012.

[11] Y. Wang and G. Cao, "Barrier coverage in camera sensor networks," in *Proc. ACM MobiHoc*, pp. 12–21, 2011.

[12] B. Liu, O. Dousse, P. Nain, and D. Towsley, "Dynamic coverage of mobile sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 2, pp. 301–311, 2013.

[13] T.-W. Kuo, K. C.-J. Lin, and M.-J. Tsai, "Maximizing submodular set function with connectivity constraint: Theory and application to networks," in *Proc. IEEE INFOCOM*, pp. 1977–1985, 2013.

[14] Y. Wang, W. Hu, Y. Wu, and G. Cao, "Smartphoto: a resource-aware crowdsourcing approach for image sensing with smartphones," in *Proc. ACM MobiHoc*, pp. 113–122, 2014.

[15] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Elsevier Comput. netw.*, vol. 52, no. 12, pp. 2292–2330, 2008.

[16] Y. Liu, Y. He, M. Li, J. Wang, K. Liu, and X. Li, "Does wireless sensor network scale? a measurement study on greenorbs," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 10, pp. 1983–1993, 2013.

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/TVT.2015.2490679, IEEE Transactions on Vehicular Technology

[17] J. Bao, Y. Zheng, and M. F. Mokbel, "Location-based and preference-aware recommendation using sparse geo-social networking data," in *Proc. ACM SIGSPATIAL*, pp. 199–208, 2012.

[18] S. Khuller, A. Moss, and J. S. Naor, "The budgeted maximum coverage problem," *Elsevier Inform. Process. Lett.*, vol. 70, no. 1, pp. 39–45, 1999.

[19] A. Krause and C. Guestrin, "A note on the budgeted maximization of submodular functions," *Tech. Rep. CMU-CALD-05-103*, 2005.

[20] L. Lovász, "Submodular functions and convexity," in *Mathematical Programming The State of the Art*, pp. 235–257, Springer Berlin Heidelberg, 1983.

[21] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functionsłi," *Springer Math. Program.*, vol. 14, no. 1, pp. 265–294, 1978.

[22] I. Rhee, M. Shin, S. Hong, K. Lee, S. J. Kim, and S. Chong, "On the levy-walk nature of human mobility," *IEEE/ACM Trans. Netw.*, vol. 19, no. 3, pp. 630–643, 2011.

[23] I. Rhee, M. Shin, S. Hong, K. Lee, S. Kim, and S. Chong, "CRAWDAD data set ncsu/mobilitymodels (v. 2009-07-23)." Downloaded from http://crawdad.org/ncsu/mobilitymodels/, July 2009.

[24] B. Guo, Z. Wang, Z. Yu, Y. Wang, N. Y. Yen, R. Huang, and X. Zhou, "Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm," *ACM Comput. Surv.*, vol. 48, no. 1, p. 7, 2015.

[25] 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 *Proc. ACM MobiSys*, pp. 55–68, 2009.

[26] R. K. Rana, C. T. Chou, S. S. Kanhere, N. Bulusu, and W. Hu, "Ear-phone: an end-to-end participatory urban noise mapping system," in *Proc. ACM IPSN*, pp. 105–116, 2010.

[27] H. Ma, D. Zhao, and P. Yuan, "Opportunities in mobile crowd sensing," *IEEE Commun. Mag.*, vol. 52, no. 8, pp. 29–35, 2014.

[28] J. Wang, Q. Gao, H. Wang, P. Cheng, and K. Xin, "Device-free localization with multidimensional wireless link information," *IEEE Trans. Veh. Technol.*, vol. 64, no. 1, pp. 356–366, 2015.

[29] J. Wang, Q. Gao, P. Cheng, Y. Yu, K. Xin, and H. Wang, "Lightweight robust device-free localization in wireless networks," *IEEE Trans. Ind. Electron.*, vol. 61, no. 10, pp. 5681–5689, 2014.

[30] J. Sun, R. Zhang, X. Jin, and Y. Zhang, "Securefind: Secure and privacy-preserving object finding via mobile crowdsourcing," *arXiv preprint arXiv:1503.07932*, 2015.

[31] S. He, J. Chen, P. Cheng, Y. J. Gu, T. He, and Y. Sun, "Maintaining quality of sensing with actors in wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 9, pp. 1657–1667, 2012.

[32] J. Shen, H. Tan, J. Wang, J. Wang, and S. Lee, "A novel routing protocol providing good transmission reliability in underwater sensor networks," *Journal of Internet Technol.*, vol. 16, no. 1, pp. 171–178, 2015.

[33] S. Xie and Y. Wang, "Construction of tree network with limited delivery latency in homogeneous wireless sensor networks," *Springer Wireless Pers. Commun.*, vol. 78, no. 1, pp. 231–246, 2014.

[34] G. S. Kasbekar, Y. Bejerano, and S. Sarkar, "Lifetime and coverage guarantees through distributed coordinate-free sensor activation," *IEEE/ACM Trans. Netw.*, vol. 19, no. 2, pp. 470–483, 2011.

[35] W. Wang, V. Srinivasan, B. Wang, and K.-C. Chua, "Coverage for target localization in wireless sensor networks," *IEEE Wireless Commun.*, vol. 7, no. 2, pp. 667–676, 2008.

[36] X. Wang, S. Han, Y. Wu, and X. Wang, "Coverage and energy consumption control in mobile heterogeneous wireless sensor networks," *IEEE Trans. Autom. Control*, vol. 58, no. 4, pp. 975–988, 2013.

[37] A. Chen, S. Kumar, and T. H. Lai, "Local barrier coverage in wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 9, no. 4, pp. 491–504, 2010.

[38] Z. Wang, J. Liao, Q. Cao, H. Qi, and Z. Wang, "Achieving k-barrier coverage in hybrid directional sensor networks," *IEEE Trans. Mobile Compt.*, vol. 13, no. 7, pp. 1443–1455, 2014.

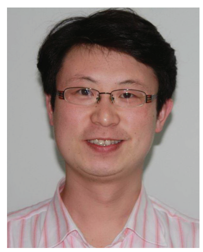
[39] B. Han, P. Hui, V. A. Kumar, M. V. Marathe, J. Shao, and A. Srinivasan, "Mobile data offloading through opportunistic communications and social participation," *IEEE Trans. Mobile Comput.*, vol. 11, no. 5, pp. 821–834, 2012.

[40] D. Zhang, H. Xiong, L. Wang, and G. Chen, "Crowdrecruiter: selecting participants for piggyback crowdsensing under probabilistic coverage constraint," in *Proc. ACM Ubicomp*, pp. 703–714, 2014.

[41] D. Zhao, H. Ma, and L. Liu, "Energy-efficient opportunistic coverage for people-centric urban sensing," *Springer Wireless Netw.*, vol. 20, no. 6, pp. 1461–1476, 2014.

[42] A. T. Campbell, S. B. Eisenman, N. D. Lane, E. Miluzzo, R. A. Peterson, H. Lu, X. Zheng, M. Musolesi, K. Fodor, and G.-S. Ahn, "The rise of people-centric sensing," *IEEE Internet Comput.*, vol. 12, no. 4, pp. 12–21, 2008.

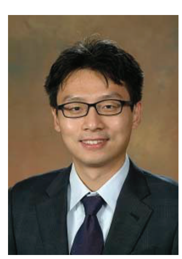
Maotian Zhang (S'12) received the B.S. degree in communication engineering from Tianjin University, China, in 2011. He is currently pursuing the Ph.D. degree in computer science and engineering at PLA University of Science and Technology, China. His current research interests include mobile computing, mobile crowdsensing networks and mobile social networks.



Panlong Yang (M'02) received his B.S. degree, M.S. degree, and Ph.D. degree in communication and information system from Nanjing Institute of Communication Engineering, China, in 1999, 2002, and 2005 respectively. Dr. Yang is now a professor in the College of Computer Science and Technology, University of Science and Technology of China. His research interests include wireless mesh networks, wireless sensor networks and cognitive radio networks. He is a member of the IEEE Computer Society and ACM SIGMOBILE Society.

Chang Tian (M'02) received his B.S. degree in communication and information system, M.S. degree and Ph.D. degree in computer science and engineering from Institute of Communications Engineering, Nanjing, China, in 1984, 1989 and 2001, respectively. During September 1995 to September 1996, he was a visiting scholar in Universit di Pisa of Italy. He is currently a professor at the PLA University of Science and Technology, China.

He has published widely in the areas of distributed system, wireless ad-hoc network, signal processing for communications, information theory. His current research interests include wireless sensor network, network coding, etc.



Shaojie Tang is currently an assistant professor of Naveen Jindal School of Management at University of Texas at Dallas. He received his PhD in computer science from Illinois Institute of Technology in 2012.

His research interest includes social networks, mobile commerce, game theory, e-business and optimization. He received the Best Paper Awards in ACM MobiHoc 2014 and IEEE MASS 2013. He also received the ACM SIGMobile service award in 2014. Tang served in various positions (as chairs and TPC members) at numerous conferences, including ACM MobiHoc and IEEE ICNP. He is an editor for Elsevier Information Processing in the Agriculture and International Journal of Distributed Sensor Networks.

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/TVT.2015.2490679, IEEE Transactions on Vehicular Technology

Xiaofeng Gao (M'09) received the B.S. degree in information and computational science from Nankai University, Tianjin, China, in 2004; the M.S. degree in operations research and control theory from Tsinghua University, Beijing, China, in 2006; and the Ph.D. degree in computer science from The University of Texas at Dallas, Richardson, TX, USA, in 2010.

She is an Associate Professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China. Her research interests include data engineering, wireless communications, and combinatorial optimizations.

Baowei Wang received his B.S. and Ph.D. degrees in Computer Science from Hunan University in 2005 and 2011, respectively. He is currently working as a lecturer in School of Computer and Software, Nanjing University of Information Science and Technology. His research interests include steganography, wireless networks and securing ad hoc networks.

Fu Xiao received Ph.D in Computer Science and Technology from Nanjing University of Science and Technology, China. He is currently professor and PhD supervisor in School of Computer, Nanjing University of Posts and Telecommunications, China. His main research interests are Sensor Networks.