Accepted by IEEE Internet of Things, 2018

Task Allocation in Spatial Crowdsourcing: Current State and Future Directions

Bin Guo, *Senior Member, IEEE*, Yan Liu, Leye Wang, *Member, IEEE*, Victor O.K. Li, *Fellow, IEEE*, Jacqueline C.K. LAM, Zhiwen Yu, *Senior Member, IEEE*

***Abstract*—Spatial crowdsourcing (SC) is an emerging paradigm of crowdsourcing, which commits workers to move to some particular locations to perform spatio-temporal-relevant tasks (e.g., sensing, activity organization). Task allocation or worker selection is a significant problem that may impact the quality of completion of SC tasks. Based on a conceptual model and generic framework of SC task allocation, this paper firstly gives a review of the current state of research in this field, including single task allocation, multiple task allocation, low-cost task allocation, and quality-enhanced task allocation. We further investigate the future trends and open issues of SC task allocation, including skill-based task allocation, group recommendation and collaboration, task composition and decomposition, and privacy-preserving task allocation. Finally, we discuss the practical issues on real-world deployment as well as the challenges for large-scale user study in SC task allocation.**

***Index Terms*—spatial crowdsourcing, task allocation, data quality, optimization, grouping and collaborating.**

I. Introduction

Crowdsourcing is firstly presented by Jeff Howe in 2006, which is the combination of two words, crowd and outsourcing [1]. It defined as the "act of taking a job traditionally performed by a designated agent (usually an employee) and outsourcing it to an undefined, generally large group of people in the form of an open call" [2]. In recent years, crowdsourcing is becoming a popular way to take advantage of the collaboration of a large number of individuals to obtain available information. Unlike traditional methods which rely on dedicated staff, any human worker can play the role of data sourcing in crowdsourcing, and people who want to obtain the information need to actively ask for some workers to answer their questions. Crowdsourcing has a wide range of application areas. Classically, most of the traditional tasks in crowdsourcing are seen as participative activities based on online platforms, such as Amazon Mechanical Turk[1]. The participative online activities may be proposed by an individual or an organization, such as in natural language understanding [3], image labeling [4], speech transcription [5], software development [6], and information mining [7]. A group of individuals with varying knowledge and skills can undertake those tasks via a flexible open call with some incentives. However, sometimes online crowdsourcing may not work when the tasks contain special requirements associated to physical places.

In the real world, geographic information plays an important role in many aspects of human life, e.g., real-time traffic information, air quality at different places, etc. Therefore, the collection of geospatial data has become a focused area. In addition, with the rapid development of pervasive computing technology, mobile devices with more powerful computing and sensing capabilities have become increasingly popular. People with mobile devices (e.g., mobile phones, tablets, wearable devices) embedded with sensors have the ability to collect various types of data (e.g., pictures, videos) close to their location, such as collecting the information of air quality. These factors promote the emergence and growth of new way of crowdsourcing, named spatial crowdsourcing (SC) [8].

In general, SC is a new paradigm of crowdsourcing platforms that outsources different types of spatio-temporal tasks to the workers or participants in the real world. Due to the rapid development of mobile networks and the widespread usage of mobile devices, SC has become a promising research area. In SC, the task requesters could publish some spatial tasks in the real-world and ask for data related to a specific spot, and workers with mobile devices are invited to perform tasks by moving to target locations specified by the tasks. Different from traditional crowdsourcing, SC applies the principles of crowdsourcing to perform tasks with human involvement and powerful mobile devices. On the one hand, workers with different capabilities (e.g., human perception, knowledge, common sense) could perform various tasks. On the other hand, human mobility of workers offers unprecedented opportunities for both data collection and transmission with mobile devices. These advantages have enabled a variety of novel SC applications in different domains, such as urban dynamics mining [9], public safety [10], traffic planning [11], etc. Recently, many SC platforms have been developed to support publishing and completing of spatial tasks, like gMission [12], mCrowd [13].

Note that most SC tasks described in the literature are sensing tasks, and this new way to complete sensing tasks is known as Mobile Crowd Sensing (MCS) [14,15]. Specifically, the participants in MCS use their smart devices to perform large-scale sensing tasks, such as collecting traffic information.

This work is supported the National Basic Research Program of China (grant no.2015CB352400), and the National Natural Science Foundation of China (grant nos. 61332005, 61772428, 61725205).

Bin Guo, Yan Liu, Zhiwen Yu are with Northwestern Polytechnical University, Xi'an, 710072, China (e-mail: guob@nwpu.edu.cn).

Leye Wang is with Hongkong University of Science and Technology, Hong Kong, China (e-mail: wangleye@gmail.com).

Victor O.K. Li and Jacqueline C.K. LAM are with the University of Hong Kong, Hong Kong, China (e-mail: vli@eee.hku.hk).

[1] https://www.mturk.com

In recent years, some new spatial tasks different from sensing tasks have appeared, such as organizing activities by the community, service sharing with multiple people [16], etc. For example, Ridesharing [17] is a typical application in service sharing that allows people to share free seats in taxis with a low transport cost. Based on the concept of ridesharing, other new shared service applications have been developed, for example, taxis could deliver packages while taking passengers [18].

Typically, there are three components in SC: *tasks*, *workers* and *the server*. The task requester first submits tasks with some information (e.g., location, time) to the server, and a number of available workers are selected to complete the tasks and send data to the server. Finally, the requester obtains the results processed and integrated by the server. Note that the selected workers play the main role in spatial crowdsourcing, and will impact the efficiency and quality of the tasks performed. While SC takes advantage of huge number of participants to enable massive mobile data sensing, it also brings many new challenges. One of the key challenges is the participant recruitment problem [19], namely, how to effectively select appropriate participants from the citizen community to perform various tasks while satisfying certain constraints.

In general, there are two ways for tasks to be assigned in crowdsourcing: worker-selection (WS) and server assignment (SA) [8]. In WS, workers can choose which task(s) to complete by browsing the published tasks on the server. For example, there is a task about exploring a specific piece of the urban environment in the platform of Campaignr [20], and any person who registers on the platform can perform the task and upload the data collected via a smart phone. Different with WST, SAT collects some information of all workers (e.g., location, preference) and requirements of tasks (spatio-temporal contexts, domains), and directly assigns workers with appropriate tasks. Note that the server allocates spatial tasks to the optimal set of workers in view of different strategies, such as maximizing the quality of completing tasks, minimizing the system cost. Therefore, SAT is more popular in spatial crowdsourcing, as it could make the best use of worker resources and improve the quality of the tasks performed. However, there are some challenges in SAT task allocation. For example, a large number of location-based tasks makes the solution space of the problem very large, and various constraints of tasks and workers increase the complexity of the problem.

This paper gives an overview of task allocation and presents its future trends. In particular, we have the following contributions:

- Presenting the basic concepts of tasks, workers and the server in SC, then proposing a generic framework for task allocation.
- Reviewing the current state of research in SC task allocation, and presenting the key challenges and techniques of task allocation, including single task allocation, multiple task allocation, low-cost task allocation, and quality-enhanced task allocation.
- Investigating the future trends and open issues of SC task allocation. We first present some new forms of SC tasks, like object delivery, object tracking and complex tasks. Then some future research directions on task allocation are explored, such as skill-based task allocation, group recommendation and collaboration.
- Discussing the practical issues and challenges in SC. We identify some issues on real-world deployment, such as human participation, load balancing, etc. Based on the existing SC platforms, we illustrate the challenges for large-scale user study in SC task allocation.

The remainder of this paper is organized as follows. In Section II, we present some basic concepts in SC, and propose a generic framework for task allocation. Section III presents the key challenges and techniques of task allocation. In Section IV, we investigate the future trends and open issues of SC task allocation. The practical issues on real-world deployment and the challenges of large-scale user study in SC task allocation are discussed in Section V. Finally, we conclude the article in Section VI.

II. Characterizing SC Task Allocation

Before introducing the research challenges of SC task allocation, we first present the conceptual model.

A. The Conceptual Model

SC refers to the process that various spatial tasks are completed by moving workers. Generally, there are three objects in the process of SC: *tasks*, *workers* and *the server*. For the task, anyone can submit requests (with spatio-temporal contexts) for real-world objects as tasks. For the worker, any individual who has the ability to complete tasks can be seen as a worker, and may receive monetary reward for completing tasks. For the server, it is necessary to allocate each of the submitted tasks to an optimal set of workers to achieve the task successfully. Then, we will give the detail concepts of tasks, workers and the server. The conceptual model of SC is shown in Fig. 1.

1) Tasks

In the real world, SC tasks refer to a series of physical processes which could be performed collaboratively by a number of independent individuals. Usually, there are several important aspects that are used to model the task in SC, which are described in detail in the following.

a) Type

We define various types of SC tasks in view of the temporal and spatial information in SC.

First, there are two task types according to the temporal information: *urgent task* [21] and *normal task* [22]. For the urgent task, workers are asked to complete tasks as soon as possible to obtain timely and useful information, such as collecting traffic dynamics information, monitoring drainage status, etc. Unlike the timeliness requirements for workers in urgent tasks, normal tasks do not need to be finished immediately. For example, workers could collect information of public facility over a period of time (say one week).

Second, spatial tasks may be divided into three types based on the spatial information: *point task* [23], *region task* [24] and *complex task* [25]. Point task, such as observing the flow of people at a street intersection, is location-specific. A worker thus has to visit the specific location to complete the task. Region task is accomplished in an area instead of the precise location, like measuring the air quality in a district/region

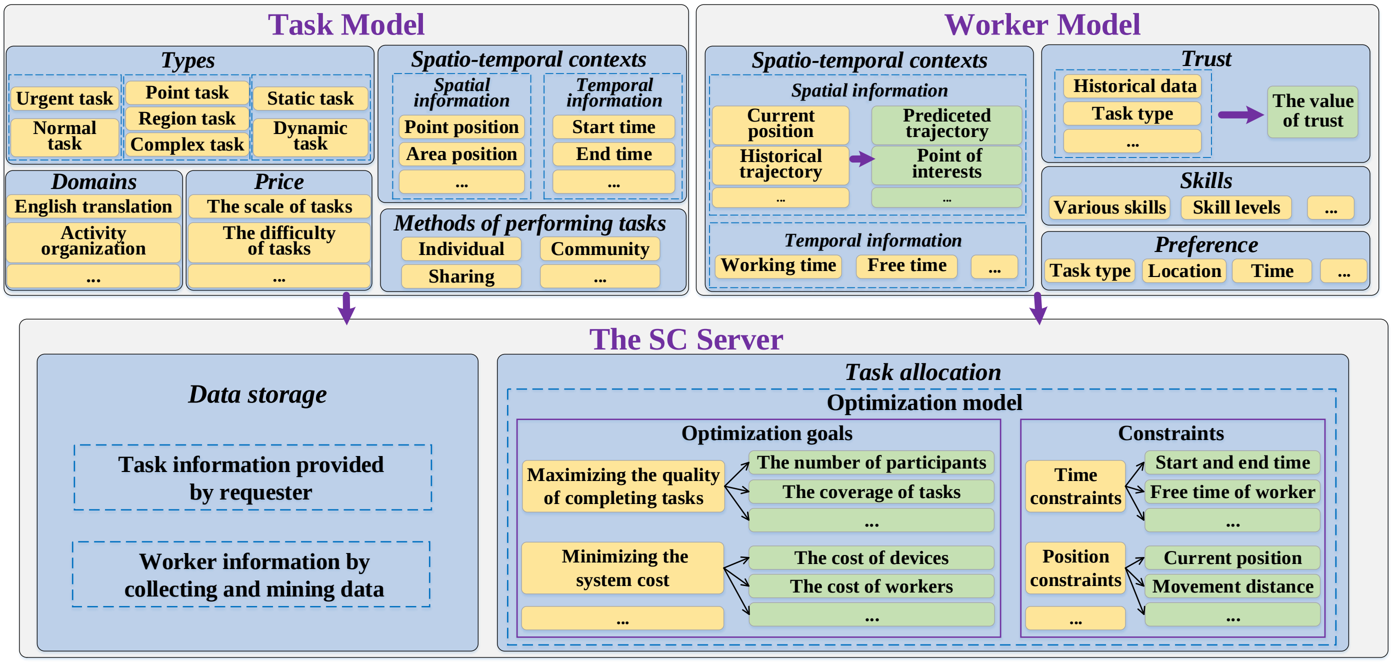


Fig. 1. The conceptual model of SC.

within a city. Note that each worker is required to perform the same action to complete point task and region task. However, complex task contains multiple different sub-tasks, which means that a task could be divided into multiple sub-tasks according to different demands for workers to accomplish. For example, given the task to obtain pictures of a large building, two sub-tasks may be performed in order to get a good quality image, one is getting a perspective of the whole building, while the other is getting the closer view for the building.

Third, spatial tasks can be classified into two task types considering both temporal and spatial information: *static task* [26] and *dynamic task* [27]. Static task consists of the spatial information with the fixed positions, like air quality monitoring, noise information mapping and so on. On the contrary, the spatial information of a dynamic task includes a sequence of uncertain locations corresponding to the temporal information, such as mobile ridesharing service, package delivery, suspicious vehicle tracking, etc.

b) Spatio-temporal Contexts

It is necessary to define the SC task with spatio-temporal contexts in the real world. Generally speaking, the requester who posts a task at the SC platform must provide the spatial and temporal information, such as road's name in traffic monitoring and restaurant's location in food delivery.

For the spatial information, point position and area position are frequently used to describe the task, such as in recording the number of passengers at the bus station, and measuring the air quality in a large park. For the temporal information, two important times that workers need to pay attention to are the start and end times. For instance, workers may be required to collect information of a concert during the performance (say 18:00-20:00).

c) *Methods of Completing Tasks*

In general, there are various methods for workers to complete tasks due to different requirements of SC tasks. For example, tasks may be performed by a number of independent individuals [28], the collaborative community [29], and via sharing some service [30]. There may be some new methods of completing tasks with the appearance of new SC tasks and pervasive technology in the future. For an individual, he/she could utilize his/her mobile device to conduct sensing tasks independently, such as obtaining traffic information by taking photos. A community formed by some workers with different abilities may be required to perform the tasks collaboratively, like developing software. It is worth noting that workers in the community could influence each other to some extent. In addition, multiple tasks could be accomplished by leveraging the same resource, e.g., multiple passengers sharing a taxi in a ridesharing service.

d) *Domains*

As in any job/task, the SC task may involve different knowledge domains. Therefore, a set of domains $D = \{d_1, d_2 \ldots d_m\}$ denotes the knowledge topics required. For example, organizing a public activity (task t) by various volunteers may require many domains such as activity planning (d_1), advertising (d_2), personal arrangement (d_3), etc.

e) *Price*

Since workers obtain revenues from completing tasks, the price of tasks plays an important role in SC, as it affects the enthusiasm of workers to participate. Generally speaking, the requester who provides the task has a right to price for the task, which takes into account the scale of the task, the difficulty of performing the task for workers, and so on.

2) Workers

To describe SC workers, we associate each worker w_i with a set of attributes in the SC platform.

a) *Spatio-temporal Contexts*

Spatio-temporal contexts of the workers refer to the information of time and space, which are used to select appropriate participants to complete tasks effectively in SC.

First, it is necessary for the platform to confirm the available time of the worker, such as, office workers can only accomplish

SC tasks after working hours. For the worker's spatial information, some localization techniques [31,32], including indoor localization and outdoor localization techniques, are used to identify the current location. Spatial information of workers not only contains the current position, but also the moving traces of workers [33]. By studying workers' historical movement patterns, more information will be mined and obtained to make efficient task allocations, such as the path of future movements, point of interests, and so on.

b) Skills

A worker may have diverse skills. A skill of the worker corresponds to the knowledge on a particular skill domain in D, and it can be quantified in a continuous scale (e.g. in a scale [0, 1]) to indicate the level of the worker's expertise for a topic. For example, a value of 0 for a skill reflects that the worker has no expertise in the corresponding domain. Note that only the worker whose skill level is not less than the minimum knowledge requirement for the task has the opportunity to complete the task.

c) Preference

In SC, the preference of workers to complete tasks generally appears in three aspects: preference for task type, location and time. For task type, some workers are willing to perform normal tasks without the additional burdens of movement, but some workers tend to complete urgent tasks to obtain more incentive. For the preference on location and time, different people have various favorite places at the corresponding times. For example, young people may prefer to obtain information at the mall in the evening, while the elderly may be happy to perform tasks in the park in the morning.

d) Trust

The trust of worker reflects the probability that the worker correctly complete a task. In general, worker's trust can be computed based on the historical data of finishing different task types. For example, worker w_i has performed the task of labelling 10 images, where 8 images are correctly labelled. So we can conclude that worker w_i has a probability of 80 percent to correctly perform a task with labelling images, and the value of trust is 0.8.

3) SC Server

The major functions of the SC server are storing data and allocating crowdsourcing tasks.

For storing data, on the one hand, the server stores the information of tasks provided by requesters and the data of tasks collected by workers. On the other hand, the server gathers and processes the data of workers, such as obtaining the worker's current location and predicting his/her next location. Then, based on the datasets in the server, the server assigns tasks to suitable workers.

For allocating tasks, the strategy of task allocation may impact the quality of the tasks, since workers have diverse qualities on different tasks. Many task allocation research efforts select the appropriate workers to perform tasks by solving the optimization problem in view of different optimization goals and constraints [34]. Here, we introduce two important optimization problems used in task allocation. The first optimization problem is *maximizing the quality of completing tasks* [35] while satisfying some constraints of tasks (e.g. time and position constraints). The second optimization problem *minimizes the system cost* [36].

In general, there are different ways to measure the quality of completing tasks. For example, the quality of completing tasks may refer to the sensing data quality in MCS, such as the quality of pictures taken by workers. To simplify the process of computing the quality of data, some parameters are frequently used to represent the quality of sensing, like the number of participants, the coverage of the sensing task, etc. It is usually assumed that the more the number of participants to perform tasks, the higher the quality of the sensing. For the community, the collaboration of workers is the major factor, which dictates the quality of completing tasks by the community. For instance, we can compute the value of the workers' collaboration in a community to accomplish tasks in view of the skills of workers. Given the task T with a set of domains $D = \{d_1, d_2, d_3, d_4\}$, and community A (n workers) with a set of skill domains in $D_A = \{d_1, d_2, d_3\}$ of task, community B (n workers) with a set of skill domains in $D_B = \{d_1, d_2, d_3, d_4\}$. Obviously, community B could finish the task T with higher quality. In addition, the quality of completing tasks can be defined as the number of accomplished tasks leveraging the existing resource in service sharing. For example, the number of passengers (tasks) in ridesharing should be maximized to improve resource utilization.

System cost usually represents the costs consumed by devices and the incentives of recruiting participants. For the cost of devices, it mainly includes the consumed energy of collecting, transmitting, processing and storing data. For the cost of workers, the requester provides some incentives to stimulate participants to complete tasks. The total incentives of performing the task depends on various factors, such as the number of selected participants, the traveling distance to accomplish the task, and so on.

B. A Generic Framework

We show the generic framework in Fig. 2.

1) Task Publishing

There are two indispensable parts in the SC platform: workers and tasks.

In the task pool are all tasks provided by task requesters. Note that not all tasks in the task pool are supposed to be published in the platform directly for task allocation. Therefore, the platform publishes some available tasks during the same time span according to the type of tasks. The platform first selects some tasks from the task pool, and then combines or decomposes the tasks. For the task combination, considering that most tasks are likely to arrive one after the other, the platform could combine some normal/non-urgent tasks, which are published in the platform within a period of time (e.g., one hour). Therefore, multiple different tasks are published in the platform during the same time span. For the task decomposition, since a complex task usually contains multiple different sub-tasks, the platform could publish some sub-tasks with different requirements to accomplish each sub-task.

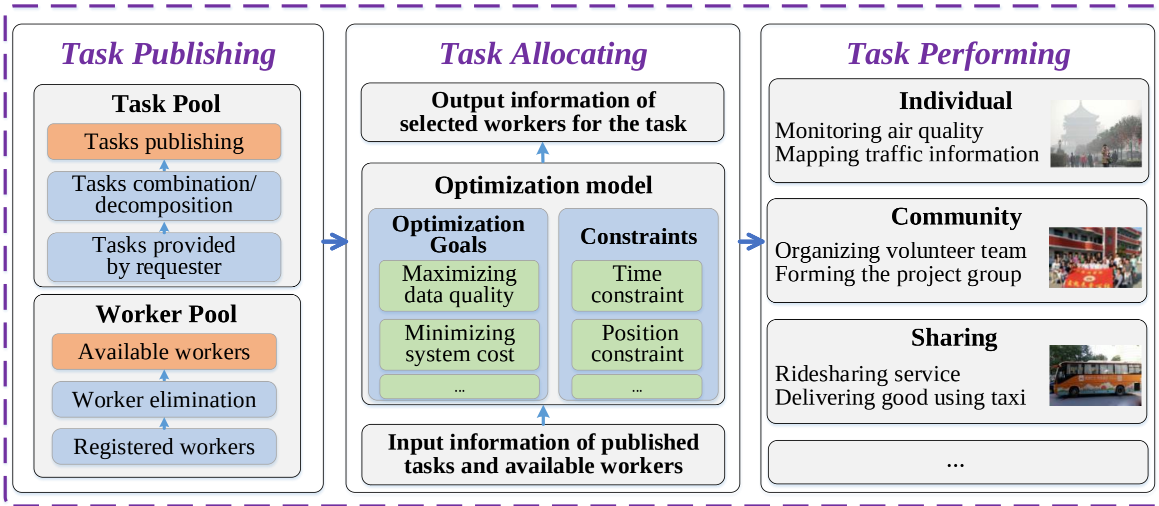


Fig. 2. The generic framework.

Some available workers in the platform are selected to perform published tasks. Owing to different demands of tasks, the platform needs to pick candidates who have the ability to perform tasks. For example, only workers with mobile devices have a chance to collect data.

2) Task Allocating

It is necessary to allocate SC tasks to some appropriate workers in order to improve the efficiency of completing tasks. Consider a crowd of available workers $W=\{w_1,w_2,w_3\ldots\}$, and a variety of tasks published by task requesters, denoted by $T=\{t_1,t_2,t_3\ldots\}$. Assuming that the optimization objective function is $f(x)$, which is usually defined as maximizing the quality of completing tasks, minimizing the system cost, etc. In addition, the selected participants are required to accomplish tasks while satisfy the constraints of tasks. For example, workers have to perform tasks at the position of tasks within the given time. Finally, the server selects the optimal set of workers S ($S \subseteq W$) corresponding to tasks via the optimization model.

3) Task Performing

Only selected workers are asked to perform SC tasks. Some common methods of completing tasks are widely used in SC, for example, tasks are performed by a number of independent individuals, the collaborative community, sharing some service.

III. Key Challenges and Techniques

Having presented the major concepts and the generic framework of SC task allocation, in this section we focus on the key challenges and techniques studied in the literature.

A. Single Task Allocation

From the time when the initial idea of SC was proposed, most task allocation studies have focused on a single task, i.e., the SC task is associated with only one kind of sensing data and one specific objective, e.g., noise or traffic sensing. Generally, such works can be seen as selecting an optimal subset of users from all the candidates to complete a task while fulfilling multiple needs. Commonly considered needs include optimizing energy consumption, reducing incentive budget cost, improving sensing quality, etc., which can either be the objectives or the constraints in the task allocation optimization. According to users' mobility patterns and task allocation timing, we can classify literature covering single-task allocation into the following categories.

1) Intentional or Unintentional Movement

SC covers two types of participants. The first type performs sensing tasks opportunistically, i.e., users will not change their routine mobility patterns and simply sense certain data unintentionally while traveling (unintentional movement). The second type of participants will intentionally travel to a certain location on purpose to finish the task (intentional movement). Traditionally, these two lines of works have been separately studied.

For unintentional movement, researchers usually need to carefully study the participants' historical movement patterns to ensure that task allocations are efficient. Ideally a task can be assigned on a participant's (near) future trajectories, and the participant can opportunistically complete the task when she moves normally. However, such task allocation is non-trivial in practice as a participant's mobility pattern is uncertain to some extent. Reddy *et al.* [37] study a recruitment framework that selects workers to maximize spatial coverage in data collection. Cardone *et al.* [38] also take user smartphones' battery levels into account when selecting appropriate workers to maximize the task completion ratio (e.g., spatio-temporal coverage). Zhang *et al.* [39] propose a worker selection framework to minimize incentive payments while satisfy the probabilistic coverage constraint under the piggyback MCS paradigm. While the above works are mostly server-centralized solutions, an autonomous and distributed task allocation strategy is proposed in [40]. Briefly, these works are based on the assumption that users' moving traces will still follow their daily routines even when they have been assigned tasks.

For intentional movement, researchers assume that participants can actively visit the task locations to complete the assigned tasks in exchange for rewards. In this case, participants' traveling distance or time becomes a critical factor to consider, since a participant usually may not be willing to travel a long distance in order to complete certain tasks. He *et al.* [41] propose an optimal task allocation scheme for location dependent MCS tasks, aiming to maximize the rewards for the MCS platform while keeping each participant's travel time to the assigned locations as a constraint. Cheung *et al.* [42] propose a distributed method for users to select their tasks, considering task time, moving costs, user reputation, etc. The proposed distributed solution achieves a comparable performance with centralized solutions.

2) Offline or Online Allocation

Regarding the time of the task allocation schemes, we can category them into two types: offline allocation and online allocation. In offline methods, the task allocation strategy is determined before the task starts. In such methods, usually the spatio-temporal information of tasks is known, while the participants' future locations are to be predicted. Under such schemes, location prediction becomes a key issue as it may affect the final task allocation performance, and commonly used algorithms are Poisson-based [39]; the task allocation performance generally increases as the participants' historical mobility traces accumulate.

Recently, online allocation methods have increasingly been examined, where the mechanism needs to be able to adapt according to the real-time micro-task and candidate participants' locations. Tong *et al.* [43] design an online task allocation mechanism to deal with the MCS scenario where tasks and participants appear dynamically. Their proposed solution provides a theoretical performance guarantee under the online random model. Han *et al.* [44] design trustful scheduling mechanisms for selecting appropriate participants under both the offline and the online settings. Pu *et al*. [45] propose an online MCS task allocation policy called CrowdLet, where service requesters can self-organize their task crowdsensing processes proactively by recruiting appropriate workers opportunistically encountered by the requesters.

B. Multiple Task Allocation

In recent years, researchers have begun to study the SC task allocation problems in a multi-task setting. While there is much less work compared to single-task ones, some interesting and important results have also been established.

Generally speaking, multiple task allocation addresses to two types of multiple tasks, *homogeneous* and *heterogeneous* tasks. As locations are usually different for different tasks, this homogeneous-heterogeneous classification depends highly on whether tasks have other specifications/requirements as well. Homogeneous multiple tasks may only have different task locations, while heterogeneous multiple tasks can have diverse specifications like temporal (different tasks need different sensing timing spans) and/or sensor (different tasks need different sensors) requirements. It is worth noting that homogeneous multiple tasks are somehow related to the single task scenario which includes multiple sub-tasks, and thus their task allocation strategies may inspire and facilitate each other under certain conditions. Despite this similarity, we separately discuss the two categories of task allocation works according to their original authors' claims for conceptual clarity.

1) Homogeneous Tasks

Xiao *et al.* [46] study the SC task allocation problem in a mobile social network (MSN), where a mobile user, called requester, may have multiple tasks that need other users in the MSN to help; however, the requester can only assign tasks and collect data when another user is within her proximity (e.g., through WiFi or Bluetooth). The authors design both offline and online task schemes for the requester to minimize the time for completing all tasks by considering other users' mobility patterns. Song *et al.* [47] consider different QoI (Quality of Information) requests of SC tasks, i.e., granularity and quantity, and propose a multi-task allocation strategy to select a minimum subset of workers to meet the QoI requirements of concurrent tasks under the total budget constraints. ActiveCrowd [48] studies the problem of multi-task worker selection under both intentional and unintentional movement situations, which are suitable for time-sensitive and delay-tolerant tasks, respectively. Liu *et al.* [21] try to address the multi-task allocation problem from another perspective. It classifies tasks into two situations: (1) MPFT (more participants, few tasks) and (2) FPMT (few participants, more tasks). In traditional MCS studies, primarily MPFT is studied, i.e., assuming the candidate participants are redundant. However, FPMT may also happen, e.g., in a heavy rain, few people are on the road, but many urgent tasks such as collecting traffic dynamics information may be published in the MCS platform.

2) Heterogeneous Tasks

Compared to homogeneous multiple tasks, heterogeneous multiple task allocation often needs to consider additional factors regarding each task's requirements. Li *et al.* [28] considers a dynamic participant recruitment problem (i.e., participants come dynamically) with heterogeneous sensing tasks (different temporal and special requirements), which aims to minimize the sensing cost while meet the coverage constraints. Three greedy algorithms are proposed to tackle this dynamic participant recruitment problem. Xiao *et al*. [49] study the deadline-sensitive worker selection problem, where each task has a deadline to satisfy. The authors propose to use a probabilistic task allocation method (i.e., a participant assigned with a task will finish it with certain probability), where multiple workers can cooperatively perform a task to meet the task deadline. In addition to temporal requirements, heterogeneous tasks may also need different types of sensors. Wang *et al.* [50] consider the sensor availability of participants by converting the multiple task allocation problem into a bipartite graph and propose an iterative greedy algorithm to address it.

C. Low-Cost Task Allocation

Reducing the costs, such as energy consumption, mobile data costs, and incentive budget, is always a key objective of SC task allocation. While the previously mentioned research studies often use these as the optimization objectives, here we explicitly summarize some recent breakthroughs in the low-cost SC task allocation literature for clarity.

1) Piggyback Crowdsensing

On a worker's smartphone, piggybacking SC task with other running mobile application can reduce the energy consumption required by the MCS task itself [36]. Thus, assigning participants tasks and letting them upload data when other mobile applications are running (e.g., making a phone call) can save the overall energy consumption of all participants significantly [51]. Based on this strategy, Xiong *et al*. [22] propose a framework of near-optimal task allocation to trade-off the participants' overall energy consumption, incentive costs, and the tasks' sensing coverage quality.

2) Compressive Crowdsensing

Recently, compressive sensing [52] has been applied in SC to infer un-sensed data based on the collected data, so as to significantly reduce the data needed in collection. [53] proposes a compressive sensing based method for various types of SC tasks, e.g., rat inspection and housing survey. In addition to missing data inference with compressive sensing, [54] identifies two other important problems to be addressed in compressive crowdsensing, i.e., cell selection (where to sense) and quality assessment (how to quantify the inference data quality in real time). It proposes to use Bayesian inference and active learning techniques to address these two issues, and also calls this paradigm as sparse MCS [55].

3) Opportunistic Encounter-based Allocation

In most SC applications, participants use 3G/4G networks to upload sensed data. However, sometimes the data size is large and thus the communication data cost is huge. To reduce such costs, short-distance wireless transmission techniques (e.g., Bluetooth, WiFi direct) are introduced into MCS task allocation. [56] studies two task allocation problems for minimizing the average makespan or largest makespan of all the tasks, respectively, considering the time that a requester needs to send tasks to the opportunistically encountered nearby participants and to receive the sensed data, as well as all the users' mobility patterns. EcoSense [57] tries to assign mobile participants with different roles in the sensed data uploading stage, so that some participants can help other participants offload their data to the server when encountered, with the objective of minimizing the overall data uploading energy consumption and communication costs.

D. Quality-Enhanced Task Allocation

Quality of sensing is a crucial issue for consideration in SC tasks. The involvement of quality needs in task allocation thus becomes an interesting and critical research direction. Quality can be characterized from various aspects, including the coverage of sensing regions, the duration of sensing time, data granularity, quantity, etc.

Since the uncertainty and uncontrollability of workers in SC, the evaluation of sensing quality should be performed online. For example, the SC platform can evaluate the quality of the photo as soon as worker uploads it when performing visual sensing tasks, and then selects other workers based on real-time results. Therefore, online algorithms are crucial in selecting proper workers who have the capacity to obtain high quality data. In view of data quality and budget in spatial crowdsensing, [54] explores the quality-guaranteed online task allocation in compressive crowdsensing. It proposes a so-called (ε, p%)-quality metric to quantify the quality of data in compressive crowdsensing, which means that the inference error of more than p% of the sensing cycles is lower than ε. A Bayesian statistical analysis method is applied to estimate whether new tasks need to be allocated further to achieve the required data-quality level. In addition, the quality of crowdsensing data cannot be fully guaranteed since workers are unprofessional. To improve sensing robustness in mobile crowdsensing, [58,59] proposes a novel framework, namely Budget LImited robuSt crowdSensing (BLISS), to tackle the problem of uncertainties about data quality, and an online learning approach is adopted to choose participants by minimizing the difference on average sense between the achieved total sensing revenue and the optimal one under a limited budget. Specifically, the online approach can acquire the statistical information about the sensing values throughout the selection process.

Some valuable quality metrics are proposed in previous research to evaluate the sensing quality in spatial crowdsourcing. [47] introduces a quality-aware metric by consideration of data granularity and quantity. A movement prediction model is proposed to estimate the quantity of data that can be collected by a worker within the sensing area. [22] defines a spatio-temporal coverage metric, called *k*-coverage, to additionally consider the fraction of covered subareas. An optimization algorithm is proposed to lower the incentive cost and meet the *k*-coverage constraints. Also quantified by spatial and temporal coverage metrics, [60] proposes a greedy approximation algorithm for worker selection in vehicular crowdsensing networks. By jointly taking into account worker ability, timeliness, and task reward, CrowdLet [45] uses the service quality metric for measuring the quality of SC task performing. A dynamic programming algorithm is proposed to maximize the expected service quality.

Data validity also impacts the quality of SC tasks. Considering that workers are not trusted equally, the challenge is how to measure the validity of the data contributed by them. To address this, [61] utilizes the reputation score to state the probability that the worker performs a task correctly. A confidence level for each spatial task is also introduced to judge whether the confidence of the results to a spatial task can be accepted or not. The quality of data is also impacted by incentives [62]. Without proper incentives, the quality of data contributed by crowd workers may be lowered. Therefore, in the future, it is also important to study the combined effects of quality measurement and incentive mechanisms in SC task allocation.

IV. Future Trends and Open Issues

There are still some limitations in recent research on SC task allocation, so we investigate the future trends and open issues in this section.

A. Toward New Forms of SC Tasks

Spatial crowdsourcing commits workers to perform spatial tasks in real-world settings. Traditional SC tasks mostly ask users to gather and share data using the integrated sensor capabilities of their mobile devices. However, not all spatial tasks are simple ones, such as taking a photo (e.g., FlierMeet [63], InstantSense [64]), reporting traffic jams (e.g., Google Waze [65]) or recording noise conditions (e.g., NoiseTube [66]). Research on spatial crowdsourcing has been evolving on multiple fronts and is now nurturing many new forms of spatial tasks, such as object delivery, object tracking, and other complex tasks.

1) Object Delivery

Spatial crowdsourcing is a part of the larger disruptive trend often referred to as the sharing economy. One of the representative SC services under this umbrella is crowdsourced object delivery, where people go about their daily lives but have the opportunity to carry individuals/objects to be delivered to specific locations. [67, 68] investigate and recommend ride sharing opportunities by analyzing human mobility patterns. CrowdDeliver [69] presents a novel passenger and package mixed transport mode which leverages the unintentional cooperation among a crowd of occupied taxis to deliver city-wide packages. They formulate it as the arriving-on-time problem to deal with the uncertainties of passengers and package requests. [18] also combines people and packages using taxis. It uses a neighborhood search method to optimize the taxi routes regarding the on-demand delivery requests defined by pickup and drop-off points.

UberEats[2] aim to improve the food delivery service using spatial crowdsourcing. Users could request food delivery from any restaurant on the service platform. The platform will then assign delivery tasks to nearby workers for picking up food packages from relevant restaurants. Compared with ride-sharing of packages, food ride-sharing problems are more challenging for several reasons. First, food delivery has more strict ‘pick-up’ (from the restaurant) and ‘arrival’ (to the food order) time constraint to ensure the quality of food and meet user dining demands/preferences. Second, the food package is usually smaller in size (by comparison with other types of packages) and to increase delivery efficiency and lower cost, a delivery-worker is often assigned multiple tasks (CrowdDeliver [69] is an SC-based package delivery system but only one package is delivered at a time by a worker). This will result in much more complicated optimization problems on SC task allocation.

2) Object Tracking

Existing SC studies mostly focus on static object sensing. We, however, can extend it to moving object sensing. Crowdsourced object tracking is an interesting research topic towards this direction. FindingNemo [70] targets the application of tracking and locating the lost child using low-power BLE peripheral via mobile crowdsensing and transparent peer collaboration. Similarly, SecureFind [71] presents a SC-based object finding system, where unique Bluetooth tags are attached to each valuable object. Though there have been initial studies in crowd tracking, special devices/tags are required. In the future, we may build crowd tracking systems by using only smartphone cameras, to form crowdsourced dynamic camera networks (as opposed to stationary-deployed camera networks [72]) for object tracking.

3) Towards Complex Tasks

SC tasks can be complex rather than simple in many cases, such as preparing for a social activity and collaborative disaster relief, which may consist of several steps and require the participation of workers with diverse skills [73]. We discuss about it in detail in the next subsection.

B. Skill-based Task Allocation

Most existing worker models assume that workers have the same expertise on different tasks. In practice, however, tasks belong to diverse domains, and workers have different expertise on different domains. It is important to select workers carefully based on the dedicated task requirements and human skills. According to WikiPedia[3], skill is defined as *“the ability to carry out a task with pre-determined results”*, and *“often divided into domain general and domain-specific skills”*. Regarding the diversity of spatial crowdsourcing tasks, human skills can be broadly linked to different abilities or attributes that a worker has, such as familiar knowledge domains, user preferences/interests, experiences, daily activity patterns, often-visited places, just name a few.

1) Skill-based Task Allocation

There have been numerous studies in skill-based crowdsourcing. For example, [74] examines user skills using knowledge base, e.g., Wikipedia and Freebase[4], to detect the domain of tasks and workers. Spatial crowdsourcing should consider more skills in real-world settings. SmartCrowd [75] assigns tasks to workers by accounting for worker expertise, wage requirements, and their availability. [73] studies skill-oriented SC task allocation by satisfying the task requirements on different skills and maximizing workers’ benefits. [22] estimates accuracies of a worker by evaluating her performance on the completed tasks, and predicting which tasks the worker is well acquainted with. [76] models tasks/workers using a hierarchical skill tree, which can map workers to tasks in a way that exploits the natural hierarchy among the skills. [77] selects workers of different daily activity patterns such that the spatio-temporal diversities of SC tasks are maximized.

2) Data-driven Skill Learning

Regarding the importance of skills in SC task allocation, another challenge is how to learn the skills of users. Explicit human inputs of skills can be one way to achieve this, but it is often time-consuming and the information obtained is often not complete. Alternatively, we can link human skills with artificial intelligence techniques and learn them from historical data. [78] finds that user skills to SC tasks are related to the attributes such as spatial proximity, the payment made, or the task theme. The logistic-regression technique is used to learn users’ individual preferences from past data. [79] identifies existing social media users who possess domain expertise (e.g., photography) and incentivize them to perform some tasks (e.g., take quality pictures). They propose a framework that extracts the potential contributors’ expertise based on their social media activity (e.g., Flickr[5]). [38] estimates worker skills by their visiting frequency to the task locations.

C. Group Recommendation and Collaboration

Complex SC tasks often need to be collaboratively completed by a team of workers. Therefore, it becomes important to recommend teams of workers to complex tasks.

1) Group Recommendation

[80, 81] study the top-k team recommendation problem in SC task allocation. The k cheapest teams that satisfy both the spatial constraints and the skill requirement of tasks will be recommended. [82] studies the formation of a group of experts to undertake a task that requires expertise in one or more domains.

2) Collaboration among Workers

Central to any group-based crowdsensing tasks is the aspect of successful collaboration among workers. Though there have

[2] https://www.ubereats.com/
[3] http://www.wikipedia.org/
[4] http://www.freebase.com/
[5] http://www.flickr.com/

been preliminary studies on group recommendation in SC task allocation, the collaboration among workers are often neglected [83]. [84, 85] investigate the notion of collaboration among workers and identify two key factors that entail successful collaboration, namely, worker-worker affinity and upper critical mass. The former represents the “comfort-level” of workers who work together on the same task; groups with low affinity often suffer from low productivity. The latter comes from organizational science and social theories, and is a constraint on the size of groups, beyond which the collaboration effectiveness diminishes. They also propose approximation or comprehensive models for collaborative crowdsourcing optimization. [86] models each worker as a selfish entity, and workers prefer to join the profitable teams where they can gain high revenue. They propose a collaborative group formation approach that allows group members to form a connected graph such that they can work together efficiently.

3) Collaboration with Online Communities

With the rapid development of Internet services and smart devices, people now work in both online and offline communities. The two types of communities are interlinked and have complementary information. It is thus important to link them and investigate the collaboration with online communities in spatial crowdsourcing [87, 88, 89]. There have recently been several studies towards this direction. For example, MoboQ [90] is a location-based question answering service that assigns spatial queries to social media (e.g., Sina Weibo[6]) users based on their online check-ins or the local intimacy learned from their posts. Similarly, [91] studies the effectiveness of employing location-based services (e.g., Foursquare[7]) for finding appropriate people to answer location-based queries. [92] presents a feature-based model to learn users’ preferences on question answering from their tweets and social connections. Expert finding is crucial to address location-based query tasks [93]. [94] formulates a top-*k* local user search problem from the tweets with geo-tags. It finds the top-*k* users who have posted tweets relevant to the desired keywords within a given query area.

D. Task Composition and Decomposition

Existing SC task allocation mechanisms mainly focus on independent tasks, while the correlations among them are omitted. In urban environments, though the spatial tasks are published by different requesters, they can be relevant from different aspects (e.g., spatial proximity, topic similarity). Thus, to improve system performance, sometimes they should be composed and completed together. On the other hand, some spatial tasks can be of high complexity and in such situations tasks should be decomposed into subtasks before being allocated to enhance efficiency.

1) Task Composition

It refers to the allocation of a sequence of tasks to a worker according to spatial properties and the association among tasks. [21] studies the multi-task allocation problem. They compose proximate micro tasks and assign them by groups to workers to maximize task throughput. Similarly, [95] optimally clusters tasks and recommends it to workers, to meet both worker’s expected wage and qualifications. [96] investigates the effect of task bundling, where workers must perform all tasks in a set to obtain payment. The results indicate that workers prefer bundled tasks, and bundled tasks have an average completion rate that is 20% higher than atomic tasks. Deng *et al*. [97] formulate task composition as a scheduling problem to maximize the number of completed tasks by each selected worker.

2) Multi-task Partitioning

To speed up task allocation when there is a large number of tasks, we need to divide the global allocation process into a set of local allocations. For example, to improve computation efficiency, [98] devises a bisection-based method which partitions tasks according to the spatial associations among them. Cheng *et al*. [77] propose the divide-and-conquer task partitioning heuristic to improve task allocation efficiency. [99] formulates task allocation as a weighted bipartite graph matching problem, and then uses a partitioning method to construct independent bipartite graphs and allocates tasks in parallel.

3) Complex Task Decomposition

To facilitate the completion of a complex task, we can split it into simple subtasks which can be processed individually [100]. The sub-task results will be combined to get the final output. [101] studies different methods for task decomposition, including the sequential implementation, the parallel implementation, and the divide and conquer implementation. Visual crowdsensing tasks [102], such as event sensing and landmark profiling, are sometimes complex as we may need pictures with spatial and temporal diversity to have a comprehensive picture of the sensing target. Regarding this, visual crowdsensing tasks can be decomposed and regrouped in task allocation. For example, a building profiling task which requests pictures about a building from different shooting directions can be split into several sub-tasks (e.g., taking pictures of the building from different directions). The sub-tasks about different buildings can be regrouped based on their spatial proximity.

E. Privacy-Preserving Task Allocation

A major concern of SC systems is the location privacy of participants. SC tasks are usually assigned to workers based on location proximity. Therefore, the location of workers should be used in task allocation. Disclosing human locations has serious privacy implications, and people may not accept SC tasks if their private information is exposed. However, the privacy issue has rarely been considered in existing SC task allocation algorithms.

Spatial cloaking is an often-used approach that allows participants to obfuscate their locations, as demonstrated in [26]. Differential privacy (DP) [103] has strong protection guarantees rooted in statistical analysis. Several works are based on this, which allows participants to obfuscate their reported locations under the guarantee of differential privacy [104, 105]. To protect user privacy, TaskMe [21] allows participants to register region-level interested areas while not precise locations for task assignment.

Besides location privacy, Li *et al*. [106] find that bids in the auctions are temporally and spatially correlated. Therefore, individual private information may be also inferred when

[6] http://weibo.com/
[7] http://www.foursquare.com/

disclosing the bids. They design a privacy-preserving task allocation method that uses Lagrange polynomial interpolation to perturb workers' bids within groups. [107] is a toolbox for interactive visualization and tuning of SC private task assignment methods, allowing us to understand the relationship between data privacy and budget setting.

F. *Incentive mechanism*

Incentive is crucial to the success of SC system because it depends on the crowd to perform a lot of tasks. Note that incentive mechanism is one of essential parts in task allocation, and it can significantly affect the quality of data contributed by workers, which is the main optimization objective in task allocation of SC [108].

There have been a lot of works that study the incentive mechanisms in SC, which can be categorized into two major types: monetary and gamification. The prior one provides workers monetary rewards, while the latter one attracts them with fun. Since workers need to consume substantial efforts and physical resources for SC tasks, monetary-based incentive mechanisms are widely used in the study of SC. MobiBee [109] is a mobile participatory game to collect fingerprints, and several incentives for participants are presented to motivate the crowd to contribute data, including financial rewards and gamification elements such as scoring system, ranking, group tasks and time pressure. The Reverse Auction (RA) [110] method is a widely used method for monetary incentives in SC, and it allows people to sell their data based on their wills and decide the payment of each worker by sealed-bidding. However, RA may result in heterogeneous payments to workers. Therefore, a more applicable method is posted pricing [111], where the unique price is announced by task requesters. Most existing commercial crowdsourcing systems (e.g., Amazon Mechanical Turk) adopt the posted pricing approach to attract user participation.

In addition to user participation, there are several other important factors that should be considered in designing incentive mechanisms, including quality of sensing [112,113], truthfulness of workers [114], cost budget [22], and so on. Note that most workers in SC are unprofessional, thus, the incentive mechanism should have the ability to ensure the truthfulness of workers and attract more workers of high reputation. [115] presents an incentive mechanism that promotes truthful reporting in crowdsourcing. [116] presents the reputation-based winner selection scheme and incentive mechanism to select workers for high-quality data contribution, which can ensure truthfulness of workers and enhance data quality. Theseus [117] is a payment mechanism that incentivizes high-effort sensing from workers, and a truth discovery algorithm is used to ensure high aggregation accuracy in MCS systems. [118] presents a truthful incentive mechanism for crowdsourcing in terms of several desirable economic properties: individual rationality, budget-balance, computational efficiency, and truthfulness.

In general, quality of sensing and payment to workers are two main factors in SC task allocation, and lower payment to workers may result in poorer data quality. At present, there have been several studies that aim to make a balance between the incentive cost and sensing quality in crowdsourcing systems. [119] presents a novel Bayesian pricing problem, and the aim of the problem is to choose an appropriate posted price and recruit a set of participants with reasonable sensing qualities to achieve sensing robustness in crowdsensing. TaskMe [120] is a novel MCS incentive mechanism, and an LBSN-powered model is leveraged for dynamic budgeting and proper worker selection. Moreover, a combination of multi-facet quality measurements and a multi-payment enhanced reverse auction scheme are proposed to improve sensing quality. Taking consideration of data quality, crowd-sensing cost, and uncertainty in offer outcomes, [121] designs a pricing mechanism that first fixes the pricing rule, and then selects users based on Unconstrained Submodular Maximization (USM). [122] proposes a constant-competitive incentive compatible mechanism, including two main objectives: maximizing the number of tasks performed under budget, and minimizing payments for a given number of tasks.

V. Real-world Deployment and Evaluation

This section discusses the practical issues on real-world deployment as well as the challenges encountered for a large-scale user study in SC task allocation.

A. *Spatial Crowdsourcing in the Wild*

When deploying spatial crowdsourcing platforms in practice, various issues have to be addressed properly, in addition to the privacy concerns and energy costs discussed above.

1) *Human Participation*

Existing work generally assumes that no rejection would occur after task allocation has been performed by the server. However, human participation should be considered a factor in practice. Workers may omit an assigned task due to the lack of time or the difficulty of the task. Therefore, how to match the right workers with the right tasks in order to maximize their acceptance rate is of great importance. One possible method is to map the human interests with the tasks during the worker selection process. For example, [123] proposes a model to measure the probability of interests for each worker-task pair, with an aim to maximize the workers' acceptance rate.

2) *Load Balancing*

As a large-scale distributed system, load balance is crucial for improving the system throughput. In SC, load balancing attempts to improve in task workload distribution across multiple participants, with an aim to minimize the task completion time, optimize resource use, and maintain long-term human participation. Not much work has addressed this issue. [124] explores the maximization of the aggregated data utilities of heterogeneous sensing tasks in task allocation, also considering the allocation of sensing task workloads among the selected workers.

3) *Uncertainty and Reliability*

Human behaviors are hard to predict and may introduce various uncertainties to the crowdsourcing systems [125]. In unintentional movement-based task allocation, we allocate tasks based on the workers' movement prediction by means of historical trajectories. However, it will introduce uncertainties as some workers may not take the same routes every day (e.g., due to interruptions in the business or the physical contexts). Also, the workers' preference or interest may evolve over time. The reliability of the SC system may also be affected by human-in-the-loop. For example, it is possible that workers can dynamically join or leave the SC system, which can impact the performance of task completion. To address this, [77] designs a grid index method that enables the dynamic update of workers/tasks to ensure system reliability.

B. Task Allocation in Existing SC Platforms

After presenting the task allocation challenges and solutions offered by the research community, this section will study the state of the art of task allocation in existing SC platforms.

Popular crowdsourcing systems, such as Amazon Mechanical Turk (AMT) [126], for performing paid tasks, and CrowdFlower[8], for collecting, cleaning, and labeling existing datasets are available in the market. Recently, with the development of spatial crowdsourcing, a number of SC platforms have emerged. They can be broadly categorized into the following types.

- *General purpose*: gMission [12], McSense [38], TaskRabbit[9], and PRISM [127] are general-purposed SC platforms which support a variety of location-based tasks.
- *Data-driven skill learning*: Gigwalk[10], EasyShift[11], and Twentify[12] are business-oriented SC platforms, which enable companies to mobilize on-demand smartphone users to collect data of retails about product performance, providing valuable insights and empowering better business decisions.
- *Dedicated use*: There are other SC platforms that provide dedicated applications, such as Postmates[13] for object delivery, MoboQ [90] for location-based query & answering, Weddar[14] for fine-grained weather information, Waze[15] for real-time traffic and road information sharing, and SeeClickFix[16] for local urban facility reporting.

These SC platforms and associated task allocation methods are summarized in Table I. Most existing platforms are based on simple task assignment schemes, i.e., location matching between task request and workers. gMission and PRISM allocates tasks based on worker location. Quite a few of them have considered users' experience/skills. For example, MoboQ assigns tasks by ranking workers in terms of locations, expertise, and availabilities; TaskRabbit considers user-specified skills in worker selection.

Other factors that should be taken into account in SC platform development include: data quality measurement, privacy-preserved task assignment, etc. In existing platforms, user skills are mostly based on human-inputs. To reduce man-made errors, data-mining methods can be considered to complement human participation. Furthermore, most platforms have focused on simple tasks; optimizations based on task decomposition to account for more complex problems should be considered. The load-balancing problem has only been considered in gMission. To tackle the issue of rejecting response, a simple strategy can be employed by assigning a task to more candidates in the neighborhood. Besides, a response deadline can often be attached when assigning a task to a participant. However, further selection schemes can be considered by incorporating factors such as user availability and interest.

TABLE I

Existing Spatial Crowdsourcing Platforms

Name	Description	Task allocation
gMission [12]	General purpose	Location-based, reputation-based, load balancing, single task
McSense [38]	General purpose	User profiling (processing power, battery-level, location) and worker ranking
PRISM [127]	General purpose	Location-based, single task
TaskRabbit[9]	General purpose	Skill and location-based task recommendation
Gigwalk[10]	Business: collect data of business performance in stores	Location-based search and recommendation
EasyShift[11]	Business: take photos of products, check prices, and review promotions.	Location-based search and recommendation
Twentify[12]	Business: enable companies to mobilize an on demand workforce of smartphone users to collect data; provide valuable insights and empower better business decisions.	Location-based search and recommendation
Weddar[14]	Weather information, invite people to report on how they feel about the weather	Self-report
SeeClickFix[16]	Allow people to report non-emergency neighborhood issues to the local governments.	Self-report
Waze[15]	Share real-time traffic and road information on their daily travels.	Self-report
Postmates[13]	On-demand local goods/food deliveries	Location-based tasks
MoboQ [90]	Location-based Q&A	Ranking (current location, location expertise, and availability)

C. Large-scale User Study

Early experimental studies on SC task allocation are mainly based on simulations, where the distribution of workers and tasks are generated based on certain models or assumptions. However, as human-in-the-loop systems, it is crucial to make experiment design closer to the real-world settings, whereas the challenge is that the recruitment of large-scale participants for real-world SC experiments is considerably high. As such, recent works on SC task allocation are mostly evaluated using synthetic datasets, i.e., integrating simulation datasets with real-life datasets to mimic the real-world environment. For example, Gowalla[17] and Brightkite[18] are two LBSNs, and we can consider LBSN users as workers, and their check-in places as task locations under evaluations. A summary of the datasets used for synthetic studies are outlined in Table II.

There are additional factors not easily simulated using employed datasets, such as workers' preference on different kinds of tasks, user-skills for performing tasks, users' task

[8] http://www.crowdflower.com/
[9] www.taskrabbit.com/
[10] http://www.gigwalk.com/
[11] http://easyshiftapp.com/
[12] http://www.twentify.com/
[13] https://postmates.com/
[14] http://www.weddar.com/
[15] http://www.waze.com/
[16] https://seeclickfix.com/
[17] http://gowalla.com/
[18] http://brightkite.com/

TABLE II
Experiments with Synthetic Datasets

Name	Dataset Description	Usage
CrowdRecruiter [39] iCrowd [22] TaskMe [21]	D4d dataset, contains 50,000 users' phone call traces from Cote d'Ivoire [129]	Historical call records as user locations; different task distributions, such as business areas and residential areas
SCAWG [128]	Check-in data from Gowalla and Yelp[130]	Initialize worker/task location
Xiao *et al*. [56]	The Cambridge Haggle Trace (Bluetooth device connections) [131] and the UMassDieselNet Trace (bus-to-bus contacts) [132]	Used for encounter-based task allocation
Song *et al*. [47]	GeoLife dataset [133], movement traces collected from 182 volunteers in Beijing	Initialize worker movements
CrowdLet [45]	Two user contact traces: Infocom06 [134] and MIT Reality Mining [135]	Used for encounter-based task allocation
CrowdPhysics [27]	The Seattle dataset: GPS-tagged tweets	Initialize worker movements
CrowdDeliver [69]	One-month taxi data of over 19,000 taxis in the city of New York, US.	Initialize the transit network and worker movements
Cheng *et al.* [77]	The POI dataset of China contains over 6 million POIs of China in 2008, while T-drive dataset includes GPS trajectories of 10,357 taxis within Beijing in 2008	Use POIs to initialize the locations of tasks. From the trajectories, they extract workers' locations, ranges of moving directions, and moving speeds.
Cheng *et al*. [73]	Meetup data set from [136], including 5,153,886 users, 5,183,840 events, and 97,587 groups from meetup.com	Use the locations of users (events) in Meetup to initialize the locations of workers (events), and use the tags of users (events) to initialize the skills of workers (the required skills of tasks)

performance reputations, etc. These, however, can only be characterized by conducting real-world user studies. Regarding its difficulty and high cost in organization and control, there are few studies that focus on real-world experiments on SC. Tasker [96] conducts a real-world user study in the SMU campus to examine the impacts of novel crowdsensing strategies. Over a two-month deployment, 30,000 tasks were performed by 900 real users. In the study conducted by FlierMeet [23], 38 recruited workers contributed more than two thousand photos during a period of two months for evaluating the crowdsourced data understanding methods. MoboQ [90] is an SC-based query answering system. During its ten-month deployment in China, 15,224 location-based query tasks performed by 35,214 registered users were proposed, with a total of 29,491 answers being collected. [77] tests their approach on a real MCS platform, namely, gMission [12]. gMission is an SC-based query answering application. Each worker in gMission is associated with his/her location and the available skills. [81] also leverage a real-world dataset collected from gMission. The dataset covers information of 11,205 workers, and each worker has an average of 5 skills. In [38], 44 university participants were recruited to collect data using the McSense mobile app, where several interesting findings about user participation are obtained. For example, users prefer automatic tasks (e.g., collecting location traces for a day) to manual tasks (e.g., taking a photo for a landmark). Though there have been several user studies about SC in the real-world setting, most systems/apps developed are dedicated and quite few of them uses a general platform (e.g., gMission) for task publishing and worker recruitment, as the usage of AMT in traditional crowdsourcing research. It indicates that the existing SC platforms are still at the early stage of development and they should be enhanced to meet different experiments and usage purposes [137].

VI. CONCLUSIONS

This paper has reviewed the current state and present future directions of task allocation in SC. First, we present the concept models of tasks, workers, and the server, which are widely used in the development process of SC. A generic framework of SC task allocation is described to select the optimal set of participants for completing spatial tasks, which includes task publishing, task allocating and task performing. Second, we summarize the current SC literature, and present the key challenges and techniques of task allocation. Particularly, this paper reviews four important task allocation studies: single task allocation, multiple task allocation, low-cost task allocation and quality-enhanced task allocation. Third, we investigate the future trends and open issues of SC task allocation in the following aspects: the new forms of SC tasks (e.g. object delivery and object tracking), skill-based task allocation, group recommendation and collaboration, task composition and decomposition, and privacy-preserving task allocation. Finally, we discuss the practical issues on real-world deployment, such as human participation and load balancing. The challenges encountered by large-scale user study in SC task allocation are also outlined.

REFERENCES

[1] J. Howe, "The rise of crowdsourcing," Wired magazine, vol.14, no. 6, pp. 1-4, 2006.

[2] J. Howe, "Crowdsourcing: A definition," Crowdsourcing: Tracking the rise of the amateur, 2006.

[3] M. Franklin, D. Kossmann, T. Kraska, S. Ramesh, and R. Xin, "CrowdDB: answering queries with crowdsourcing," in Proceedings of the ACM SIGMOD International Conference on Management of data, 2011, pp. 61-72.

[4] Y. Liu, V. Lehdonvirta, M. Kleppe, T. Alexandrova, H. Kimura, and T. Nakajima, "A crowdsourcing based mobile image translation and knowledge sharing service," in Proceedings of International Conference on Mobile and Ubiquitous Multimedia, 2010, p.6.

[5] G. Parent and M. Eskenazi, "Speaking to the crowd: looking at past achievements in using crowdsourcing for speech and predicting future challenges," in Twelfth Annual Conference of the International Speech Communication Association, 2011.

[6] M. Vukovic, "Crowdsourcing for enterprises," in Services-I, 2009 World Conference on. IEEE, pp. 686–692, 2009.

[7] X. Guo, H. Wang, Y. Song, and H. Gao, "Brief survey of crowdsourcing for data mining," Expert Systems with Applications, vol. 41, no. 17, pp. 7987-7994, 2014.

[8] L. Kazemi and C. Shahabi, "Geocrowd: enabling query answering with spatial crowdsourcing," in Proceedings of international conference on advances in geographic information systems, 2012, pp.189-198.

[9] N. Lathia, V. Pejovic, K. Rachuri, C. Mascolo, M. Musolesi, and P. Rentfrow, "Smartphones for large-scale behavior change interventions," IEEE Pervasive Computing, vol. 12, no. 3, pp. 66-73, 2013.

[10] J. Ballesteros, B. Carbunar, M. Rahman, N. Rishe, and S. Iyengar, "Towards safe cities: A mobile and social networking approach," IEEE Transactions on Parallel and Distributed Systems, vol. 25, no. 9, pp. 2451-2462, 2014.

[11] V. Coric and M. Gruteser, "Crowdsensing maps of on-street parking spaces," in Proceedings of IEEE International Conference on Distributed Computing in Sensor Systems, 2013, pp. 115-122.

[12] Z. Chen, R. Fu, Z. Zhao, Z. Liu, L. Xia, L. Chen, P. Cheng, C. Chen, Y. Tong, and C. Zhang, "gmission: A general spatial crowdsourcing

platform," in Proceedings of the VLDB Endowment, 2014, pp. 1629-1632.
[13] T. Yan, M. Marzilli, R. Holmes, D. Ganesan, and M. Corner, "mCrowd: a platform for mobile crowdsourcing," in Proceedings of ACM Conference on Embedded Networked Sensor Systems, 2009, pp. 347-348.
[14] R. Ganti, F. Ye, and H. Lei, "Mobile crowdsensing: current state and future challenges," IEEE Communications Magazine, vol. 49, no. 11, 2011.
[15] H. Ma, D. Zhao, and P. Yuan, "Opportunities in mobile crowd sensing," IEEE Communications Magazine, vol. 52, no. 8, pp. 29-35, 2014.
[16] B. Cohen and J. Kietzmann, "Ride on! Mobility business models for the sharing economy," Organization & Environment, vol. 27, no. 3, pp. 279-296, 2014.
[17] M. Furuhata, M. Dessouky, F. Ordóñez, M. Brunet, X. Wang, and S. Koenig, "Ridesharing: The state-of-the-art and future directions," Transportation Research Part B: Methodological, vol.57, pp.28-46, 2013.
[18] B. Li, D. Krushinsky, H. Reijers, and T. Woensel, "The share-a-ride problem: People and parcels sharing taxis," European Journal of Operational Research, pp. 31-40, 2014.
[19] M. Karaliopoulos, O. Telelis, and I. Koutsopoulos, "User recruitment for mobile crowdsensing over opportunistic networks," in Proceedings of IEEE INFOCOM'15, 2015, pp. 2254-2262.
[20] A. Joki, J. Burke, and D. Estrin, "Campaignr: a framework for participatory data collection on mobile phones," 2007.
[21] Y. Liu, B. Guo, Y. Wang, W. Wu, Z. Yu, and D. Zhang, "Taskme: multi-task allocation in mobile crowd sensing," in Proceedings of the UbiComp'16, 2016, pp. 403-414.
[22] H. Xiong, D. Zhang, G. Chen, L. Wang, V. Gauthier, and L. Barnes, "iCrowd: Near-optimal task allocation for piggyback crowdsensing," IEEE Transactions on Mobile Computing, vol.15, no.8, pp.2010-2022, 2016.
[23] B. Guo, H. Chen, Z. Yu, X. Xie, S. Huangfu, and D. Zhang, "FlierMeet: a mobile crowdsensing system for cross-space public information reposting, tagging, and sharing," IEEE Transactions on Mobile Computing, vol.14, no.10, pp.2020-2033, 2015.
[24] H. To, L. Fan, L. Tran, and C.s Shahabi, "Real-time task assignment in hyperlocal spatial crowdsourcing under budget constraints," in Proceedings of PerCom'16, 2016, pp.1-8, 2016.
[25] M. Saini, R. Gadde, S. Yan, and W. Tsang, "MoViMash: online mobile video mashup," in Proceedings of ACM international conference on Multimedia, 2012, pp.139-148.
[26] L. Pournajaf, L. Xiong, V. Sunderam, and S. Goryczka, "Spatial task assignment for crowd sensing with cloaked locations," in Proceedings of IEEE International Conference on Mobile Data Management, 2014, pp.73-82.
[27] A. Sadilek, J. Krumm, and E. Horvitz, "Crowdphysics: Planned and opportunistic crowdsourcing for physical tasks," 2013.
[28] H. Li, T. Li, and Y. Wang, "Dynamic participant recruitment of mobile crowd sensing for heterogeneous sensing tasks," in Proceedings of Mobile IEEE International Conference on Ad Hoc and Sensor Systems, 2015, pp. 136-144.
[29] Z. Yu, R. Du, B. Guo, H. Xu, T. Gu, Z. Wang, and D. Zhang, "Who should i invite for my party?: Combining user preference and influence maximization for social events," in Proceedings of UbiComp'15, 2015, pp.879-883.
[30] B. Li, D. Krushinsky, T. Woensel, and H. Reijers, "The Share-a-Ride problem with stochastic travel times and stochastic delivery locations," Transportation Research Part C: Emerging Technologies, vol.67, pp.95-108, 2016.
[31] X. Guo, L. Chu, and X. Sun, "Accurate localization of multiple sources using semidefinite programming based on incomplete range matrix," IEEE Sensors Journal, vol.16, no. 13, pp.5319-5324, 2016.
[32] G. Wang, S. Cai, Y. Li, and N. Ansari, "A bias-reduced nonlinear WLS method for TDOA/FDOA-based source localization," IEEE Transactions on Vehicular Technology, vol. 65, no.10, pp.8603-8615, 2016.
[33] S. Gambs, M. Killijian, and M. Cortez, "Next place prediction using mobility markov chains," in Proceedings of the First Workshop on Measurement, Privacy, and Mobility, 2012, p.3.
[34] I. Boutsis and V. Kalogeraki, "On task assignment for real-time reliable crowdsourcing," in Proceedings of IEEE International Conference on Distributed Computing Systems, pp.1-10, 2014.
[35] H. Xiong, D. Zhang, G. Chen, L. Wang, and V. Gauthier, "Crowdtasker: Maximizing coverage quality in piggyback crowdsensing under budget constraint," in Proceedings of PerCom'15, 2015, pp.55-62.
[36] N. Lane, Y. Chon, L. Zhou, Y. Zhang, F. Li, D. Kim, G. Ding, F. Zhao, and H. Cha, "Piggyback CrowdSensing: energy efficient crowdsourcing of mobile sensor data by exploiting smartphone app opportunities," in Proceedings of ACM Conference on Embedded Networked Sensor Systems, 2013, p.7.
[37] S. Reddy, D. Estrin, and M. Srivastava, "Recruitment framework for participatory sensing data collections," in International Conference on Pervasive Computing. Springer, pp.138-155, 2010.
[38] G. Cardone, L. Foschini, P. Bellavista, A. Corradi, C. Borcea, M. Talasila, and R. Curtmola, "Fostering participaction in smart cities: a geo-social crowdsensing platform," IEEE Communications Magazine, vol.51, no.6, pp.112-119, 2013.
[39] D. Zhang, H. Xiong, L. Wang, and G. Chen, "CrowdRecruiter: selecting participants for piggyback crowdsensing under probabilistic coverage constraint," in Proceedings of UbiComp'14, 2014, pp.703-714.
[40] G. Tuncay, G. Benincasa, and A. Helmy, "Autonomous and distributed recruitment and data collection framework for opportunistic sensing," ACM SIGMOBILE Mobile Computing and Communications Review, vol.16, no.4, pp.50-53, 2013.
[41] S. He, D. Shin, J. Zhang, and J. Chen, "Toward optimal allocation of location dependent tasks in crowdsensing," in Proceedings of IEEE INFOCOM'14, 2014, pp.745-753.
[42] M. Cheung, R. Southwell, F. Hou, and J. Huang, "Distributed time-sensitive task selection in mobile crowdsensing," in Proceedings of ACM International Symposium on Mobile Ad Hoc Networking and Computing, 2015, pp.157-166.
[43] Y. Tong, J. She, B. Ding, L. Wang, and L. Chen, "Online mobile micro-task allocation in spatial crowdsourcing," in Proceedings of IEEE International Conference on Data Engineering, 2016, pp.49-60.
[44] K. Han, C. Zhang, J. Luo, M. Hu, and B. Veeravalli, "Truthful scheduling mechanisms for powering mobile crowdsensing," IEEE Trans. Computing, vol.65, no.1, pp.294-307, 2016.
[45] L. Pu, X. Chen, J. Xu, and X. Fu, "Crowdlet: Optimal worker recruitment for self-organized mobile crowdsourcing," in Proceedings of IEEE INFOCOM'16, 2016, pp.1-9.
[46] M. Xiao, J. Wu, L. Huang, Y. Wang, and C. Liu, "Multi-task assignment for crowdsensing in mobile social networks," in Proceedings of IEEE INFOCOM'15, 2015, pp.2227-2235.
[47] Z. Song, C. Liu, J. Wu, J. Ma, and W. Wang, "Qoi-aware multitask-oriented dynamic participant selection with budget constraints," IEEE Transactions on Vehicular Technology, vol.63, no.9, pp.4618-4632, 2014.
[48] B. Guo, Y. Liu, W. Wu, Z. Yu, and Q. Han. "Activecrowd: A framework for optimized multitask allocation in mobile crowdsensing systems," IEEE Transactions on Human-Machine Systems, vol.47, no.3, pp.392-403, 2017.
[49] M. Xiao, J. Wu, H. Huang, L. Huang, and C. Hu, "Deadline-sensitive User Recruitment for mobile crowdsensing with probabilistic collaboration," in Proceedings of IEEE International Conference on Network Protocols, 2016, pp.1-10.
[50] J. Wang, Y. Wang, D. Zhang, F. Wang, Y. He, and L. Ma, "PSAllocator: multi-task allocation for participatory sensing with sensing capability constraints," in Proceedings of ACM Conference on Computer Supported Cooperative Work and Social Computing, 2017, pp.1139-1151.
[51] H. Xiong, D. Zhang, L. Wang, J. Gibson, and J. Zhu, "EEMC: Enabling energy-efficient mobile crowdsensing with anonymous participants," ACM Transactions on Intelligent Systems and Technology, vol.6, no.3 pp.39, 2015.
[52] D. Donoho, "Compressed sensing," IEEE Transactions on information theory, vol.52, no.4, pp.1289-1306, 2006.
[53] L. Xu, X. Hao, N. Lane, X. Liu, and T. Moscibroda, "More with less: Lowering user burden in mobile crowdsourcing through compressive sensing," in Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2015, pp.659-670.
[54] L. Wang, D. Zhang, A. Pathak, C. Chen, H. Xiong, D. Yang, and Y. Wang, "CCS-TA: Quality-guaranteed online task allocation in compressive crowdsensing," in Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2015, pp. 683-694.
[55] L. Wang, D. Zhang, Y. Wang, C. Chen, X. Han, and A. Mhamed, "Sparse mobile crowdsensing: challenges and opportunities," IEEE Communications Magazine, vol. 54, no.7, pp.161-167, 2016.
[56] M. Xiao, J. Wu, L. Huang, R. Cheng, and Y. Wang, "Online task assignment for crowdsensing in predictable mobile social networks," IEEE Transactions on Mobile Computing, vol.16, no.8, pp.2306-2320, 2016.

[57] L. Wang, D. Zhang, H. Xiong, J. Gibson, C. Chen, and B. Xie, "ecoSense: Minimize Participants' Total 3G Data Cost in Mobile Crowdsensing Using Opportunistic Relays," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol.47, no.6, pp.965-978, 2017.

[58] K. Han, C. Zhang, and J. Luo, "BLISS: Budget limited robust crowdsensing through online learning," in Proceedings of IEEE International Conference on SECON'14, 2014, pp. 555-563.

[59] K. Han, C. Zhang, and J. Luo, "Taming the uncertainty: Budget limited robust crowdsensing through online learning," IEEE/ACM Transactions on Networking, vol. 24, no. 3, pp.1462-1475, 2016.

[60] Z. He, J. Cao, and X. Liu, "High quality participant recruitment in vehicle-based crowdsourcing using predictable mobility," in Proceedings of IEEE INFOCOM'15, 2015, pp.2542-2550.

[61] L. Kazemi, C. Shahabi, and L. Chen, "Geotrucrowd: trustworthy query answering with spatial crowdsourcing," in Proceedings of ACM sigspatial international conference on advances in geographic information systems, 2013, pp.314-323.

[62] R. Kawajiri, M. Shimosaka, and H. Kashima, "Steered crowdsensing: incentive design towards quality-oriented place-centric crowdsensing," in Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2014, pp.691-701.

[63] B. Guo, H. Chen, Z. Yu, X. Xie, S. Huangfu, and Z. Wang, "FlierMeet: Cross-space public information reposting with mobile crowd sensing,"in Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, 2014, pp.59-62.

[64] H. Chen, B. Guo, Z. Yu, and Q. Han, "Toward real-time and cooperative mobile visual sensing and sharing," in Proceedings of IEEE INFOCOM'16, 2016, pp.1-9.

[65] G. Chatzimilioudis, A. Konstantinidis, C. Laoudias, and D. Zeinalipour-Yazti, "Crowdsourcing with smartphones," IEEE Internet Computing, vol.16, no.5, pp.36-44, 2012.

[66] N. Maisonneuve, M. Stevens, M. Niessen, and L. Steels, "NoiseTube: Measuring and mapping noise pollution with mobile phones," Information technologies in environmental engineering, pp. 215-228, 2009.

[67] N. Bicocchi and M. Mamei, "Investigating ride sharing opportunities through mobility data analysis," Pervasive and Mobile Computing, pp.83-94, 2014.

[68] D. Santos and E. Xavier, "Taxi and ride sharing: A dynamic dial-a-ride problem with money as an incentive," Expert Systems with Applications, vol.42, no.19, pp.6728-6737, 2015.

[69] C. Chen, D. Zhang, X. Ma, B. Guo, L. Wang, Y. Wang, and E. Sha, "Crowddeliver: planning city-wide package delivery paths leveraging the crowd of taxis," IEEE Transactions on Intelligent Transportation Systems, vol. 18, no. 6, pp. 1478-1496, 2016.

[70] K. Liu and X. Li, "Finding Nemo: Finding Your Lost Child in Crowds via Mobile Crowd Sensing," in Proceedings of IEEE International Conference on Mobile Ad Hoc and Sensor Systems, 2014, pp.1-9.

[71] J. Sun, R. Zhang, X. Jin, and Y. Zhang, "Securefind: Secure and privacy-preserving object finding via mobile crowdsourcing," IEEE Transactions on Wireless Communications, vol.15, no.3, pp.1716-1728, 2016.

[72] W. Limprasert, A. Wallace, and G. Michaelson, "Real-time people tracking in a camera network," IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol.3, no.2, pp.263-271, 2013.

[73] P. Cheng, X. Lian, L. Chen, J. Han, and J. Zhao, "Task assignment on multi-skill oriented spatial crowdsourcing," IEEE Transactions on Knowledge and Data Engineering, vol.28, no.8, pp.2201-2215, 2016.

[74] Y. Zheng, G. Li, and R. Cheng, "Docs: a domain-aware crowdsourcing system using knowledge bases," in Proceedings of the VLDB Endowment, vol.10, no.4, pp.361-372, 2016.

[75] S. Roy, I. Lykourentzou, S. Thirumuruganathan, S. Amer-Yahia, and G. Das, "Task assignment optimization in knowledge-intensive crowdsourcing," The VLDB Journal, vol.24, no.4, pp.467-491, 2015.

[76] P. Mavridis, D. Gross-Amblard, and Z. Miklós, "Using hierarchical skills for optimized task assignment in knowledge-intensive crowdsourcing," in Proceedings of International Conference on World Wide Web, 2016, pp.843-853.

[77] P. Cheng, X. Lian, Z. Chen, R. Fu, L. Chen, J. Han, and J. Zhao, "Reliable diversity-based spatial crowdsourcing by moving workers," in Proceedings of the VLDB Endowment, vol.8, no.10, pp.1022-1033, 2015.

[78] M. Karaliopoulos, I. Koutsopoulos, and M. Titsias, "First learn then earn: Optimizing mobile crowdsensing campaigns through data-driven user profiling," in Proceedings of ACM International Symposium on Mobile Ad Hoc Networking and Computing, 2016, pp.271-280.

[79] P. Micholia, M. Karaliopoulos, I. Koutsopoulos, L. Aiello, G. Morales, and D. Quercia, "Incentivizing social media users for mobile crowdsourcing," International Journal of Human-Computer Studies, pp.4-13, 2017.

[80] D. Gao, Y. Tong, J. She, T. Song, L. Chen, and K. Xu, "Top-k team recommendation in spatial crowdsourcing," in International Conference on Web-Age Information Management, Springer, pp.191-204, 2016.

[81] D. Gao, Y. Tong, J. She, T. Song, L. Chen, and K. Xu, "Top-k Team Recommendation and Its Variants in Spatial Crowdsourcing," Data Science and Engineering, pp.1-15, 2017.

[82] H. Rahman, S. Thirumuruganathan, S. Roy, S. Amer-Yahia, and G. Das, "Worker skill estimation in team-based tasks," in Proceedings of the VLDB Endowment, vol.8, no.11, pp.1142-1153, 2015.

[83] T. Ludwig, C. Kotthaus, C. Reuter, S. Dongen, and V. Pipek, "Situated crowdsourcing during disasters: Managing the tasks of spontaneous volunteers through public displays," International Journal of Human-Computer Studies, pp. 103-121, 2017.

[84] H. Rahman, S. Roy, S. Thirumuruganathan, S. Amer-Yahia, and G. Das. 2015, "Task assignment optimization in collaborative crowdsourcing," in Proceedings of IEEE International Conference on Data Mining, 2015, pp.949-954.

[85] R. Fishel, "The whole is greater than the sum of its parts," Critical care medicine, vol.35, no.10, pp.2431-2432, 2007.

[86] W. Wang, J. Jiang, B. An, Y. Jiang, and B. Chen, "Toward efficient team formation for crowdsourcing in noncooperative social networks," IEEE transactions on cybernetics, vol. 47, no. 12, pp. 4208-4222, 2016.

[87] B. Guo, Z. Yu, D. Zhang, and X. Zhou, "Cross-community sensing and mining," IEEE Communications Magazine, vol.52, no.8, pp.144-152, 2014.

[88] B. Guo, C. Chen, D. Zhang, Z. Yu, and A. Chin, "Mobile crowd sensing and computing: when participatory sensing meets participatory social media," IEEE Communications Magazine, vol.54, no.2, pp.131-137, 2016.

[89] M. Avvenuti, S. Bellomo, S. Cresci, M. Polla, and M. Tesconi, "Hybrid crowdsensing: A novel paradigm to combine the strengths of opportunistic and participatory crowdsensing," in Proceedings of International Conference on World Wide Web Companion, 2017, pp.1413-1421.

[90] Y. Liu, T. Alexandrova, and T. Nakajima, "Using stranger as sensors: temporal and geo-sensitive question answering via social media," in Proceedings of international conference on World Wide Web, 2013, pp.803-814.

[91] M. Bulut, Y. Yilmaz, and M. Demirbas, "Crowdsourcing location-based queries," in Proceedings of IEEE International Conference Pervasive Computing and Communications Workshops, 2011, pp.513-518.

[92] J. Mahmud, M. Zhou, N. Megiddo, J. Nichols, and C. Drews, "Optimizing the selection of strangers to answer questions in social media," arXiv preprint arXiv:1404, 2014.

[93] P. Shankar, Y. Huang, P. Castro, B. Nath, and L. Iftode, "Crowds replace experts: Building better location-based services using mobile social network interactions," in Proceedings of PerCom'12, 2012, pp.20-29.

[94] J. Jiang, H. Lu, B. Yang, and B. Cui, "Finding top-k local users in geo-tagged social media data," In Proceedings of IEEE International Conference on Data Engineering, 2015, pp.267-278.

[95] S. Amer-Yahia, E. Gaussier, V. Leroy, J. Pilourdault, R. Borromeo, and M.Toyama, "Task composition in crowdsourcing," in Proceedings of IEEE International Conference on Data Science and Advanced Analytics, 2016, pp.194-203.

[96] T. Kandappu, N. Jaiman, R. Tandriansyah, A. Misra, S. Cheng, C. Chen, H. Lau, D. Chander, and K. Dasgupta, "Tasker: Behavioral insights via campus-based experimental mobile crowd-sourcing," in Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2016, pp.392-402.

[97] D. Deng, C. Shahabi, and U. Demiryurek, "Maximizing the number of worker's self-selected tasks in spatial crowdsourcing," in Proceedings of ACM sigspatial international conference on advances in geographic information systems, 2013, pp.324-333.

[98] D. Deng, C. Shahabi, and L. Zhu, "Task matching and scheduling for multiple workers in spatial crowdsourcing," in Proceedings of SIGSPATIAL International Conference on Advances in Geographic Information Systems, 2013, p.21.

[99] A. Alfarrarjeh, T. Emrich, and C. Shahabi, "Scalable spatial crowdsourcing: A study of distributed algorithms," in Proceedings of IEEE International Conference on Mobile Data Management, 2015, pp.134-144.

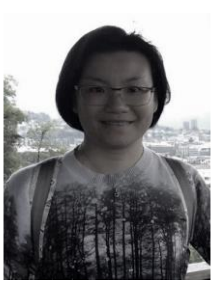
[100] H. Dang, T. Nguyen, and H. To, “Maximum complex task assignment: Towards tasks correlation in spatial crowdsourcing,” in Proceedings of International Conference on Information Integration and Web-based Applications & Services, 2013, pp.77.
[101] A. Chittilappilly, L. Chen, and S. Amer-Yahia, “A survey of general-purpose crowdsourcing techniques,” IEEE Transactions on Knowledge and Data Engineering, vol.28, no.9, pp.2246-2266, 2016.
[102] B. Guo, H. Chen, Q. Han, Z. Yu, D. Zhang, and Y. Wang, “Worker-contributed data utility measurement for visual crowdsensing systems,” IEEE Transactions on Mobile Computing, vol. 16, no. 8, pp. 2379-2391, 2017.
[103] C. Dwork, “Differential privacy: A survey of results,” in International Conference on Theory and Applications of Models of Computation, Springer, pp.1-19, 2008.
[104] H. To, G. Ghinita, and C. Shahabi, “A framework for protecting worker location privacy in spatial crowdsourcing,” in Proceedings of the VLDB Endowment, vol.7, no.10, pp.919-930, 2014.
[105] L. Wang, D. Yang, X. Han, T. Wang, D. Zhang, and X. Ma, “Location Privacy-Preserving Task Allocation for Mobile Crowdsensing with Differential Geo-Obfuscation,” in Proceedings of International Conference on World Wide Web, 2017, pp.627-636.
[106] T. Li, T. Jung, H. Li, L. Cao, W. Wang, X. Li, and Y. Wang, “Scalable privacy preserving participant selection in mobile crowd sensing,” in Proceedings of PerCom’17, 2017, pp.59-68.
[107] H. To, G. Ghinita, and C. Shahabi, “PrivGeoCrowd: A toolbox for studying private spatial crowdsourcing,” in Proceedings of IEEE International Conference on Data Engineering, 2015, pp.1404-1407.
[108] D. Zhao, H. Ma, and L. Liu, “Frugal online incentive mechanisms for mobile crowd sensing,” IEEE Transactions on Vehicular Technology, vol. 66, no. 4, pp.3319-3330, 2017.
[109] Q. Xu, and R. Zheng, “Mobibee: A mobile treasure hunt game for location-dependent fingerprint collection,” in Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct, 2016, pp. 1472-1477.
[110] H. Jin, L. Su, D. Chen, K. Nahrstedt, and J. Xu, “Quality of information aware incentive mechanisms for mobile crowd sensing systems,” in Proceedings of ACM International Symposium on Mobile Ad Hoc Networking and Computing, 2016, pp. 167-176.
[111] A. Singla, and A. Krause, “Truthful incentives in crowdsourcing tasks using regret minimization mechanisms,” in Proceedings of international conference on World Wide Web, 2013, pp. 1167-1178.
[112] L. Liu, W. Wei, D. Zhao, and H. Ma, “Urban resolution: New metric for measuring the quality of urban sensing,” IEEE Transactions on Mobile Computing, vol. 14, no. 12, pp.2560-2575, 2015.
[113] M. Pouryazdan, B. Kantarci, T. Soyata, L. Foschini, and H. Song, “Quantifying user reputation scores, data trustworthiness, and user incentives in mobile crowd-sensing,” IEEE Access, vol. 5, pp. 1382-97, 2017
[114] D. Zhao, X.Y. Li, and H. Ma, “Budget-feasible online incentive mechanisms for crowdsourcing tasks truthfully,” IEEE/ACM Transactions on Networking, vol. 24, no. 2, pp.647-661, 2016.
[115] E. Kamar, and E. Horvitz, “Incentives for truthful reporting in crowdsourcing,” in Proceedings of international conference on autonomous agents and multiagent systems-volume 3, pp. 1329-1330, 2012.
[116] B. Guo, W. Nan, Z. Yu, X. Xie, H. Chen, and X. Zhou, “TaskMe: a cross-community, quality-enhanced incentive mechanism for mobile crowd sensing,” in Adjunct Proceedings of International Joint Conference on Pervasive and Ubiquitous Computing , 2015, pp. 49-52.
[117] H. Jin, L. Su, and K. Nahrstedt, “Theseus: Incentivizing truth discovery in mobile crowd sensing systems,” in Proceedings of ACM International Symposium on Mobile Ad Hoc Networking and Computing, 2017, p. 1.
[118] X. Zhang, G. Xue, R. Yu, D. Yang, and J.Tang, “Truthful incentive mechanisms for crowdsourcing,” in Proceedings of IEEE INFOCOM’15, 2015, pp. 2830-2838.
[119] K. Han, H. Huang, and J. Luo, “Posted pricing for robust crowdsensing,” in Proceedings of ACM International Symposium on Mobile Ad Hoc Networking and Computing, 2016, pp. 261-270.
[120] B. Guo, H. Chen, Z. Yu, W. Nan, X. Xie, D. Zhang, and X. Zhou, “Taskme: Toward a dynamic and quality-enhanced incentive mechanism for mobile crowd sensing,” International Journal of Human-Computer Studies, vol. 102, pp.14-26, 2017.
[121] X. Ying, S. Roy, and R. Poovendran, “Pricing mechanisms for crowd-sensed spatial-statistics-based radio mapping,” IEEE Transactions on Cognitive Communications and Networking, vol. 3, no.2, pp.242-254, 2017.
[122] Y. Singer, and M. Mittal, “Pricing mechanisms for crowdsourcing markets,” in Proceedings of International conference on World Wide Web, 2013, pp. 1157-1166.
[123] L. Zheng, and L. Chen, “Maximizing acceptance in rejection-aware spatial crowdsourcing,” IEEE Transactions on Knowledge and Data Engineering, vol. 29, no. 9, pp.1943-1956, 2017.
[124] W. Sun, Y. Zhu, L. Ni, and B. Li, “Crowdsourcing Sensing Workloads of Heterogeneous Tasks: A Distributed Fairness-Aware Approach,” in Proceedings of International Conference on Parallel Processing, 2015, pp.580-589.
[125] S. Cheng, C. Chen and A. HoongChuinLau, “Towards city-scale mobile crowdsourcing: Task recommendations under trajectory uncertainties,” 2015.
[126] G. Paolacci, J. Chandler, and P. Ipeirotis, “Running experiments on amazon mechanical turk,” 2010.
[127] A. Hinton, M. Kwiatkowska, G. Norman, and D. Parker, “PRISM: A tool for automatic verification of probabilistic systems,” in Proceedings of International Conference on Tools and Algorithms for the Construction and Analysis of Systems. Springer, 2006, pp.441-444.
[128] H. To, M. Asghari, D. Deng, and C. Shahabi, “SCAWG: A toolbox for generating synthetic workload for spatial crowdsourcing,” in Proceedings of Pervasive Computing and Communication Workshops, 2016, pp.1-6.
[129] V. Blondel, M. Esch, C. Chan, F. Clérot, P. Deville, E. Huens, F. Morlot, Z. Smoreda, and C. Ziemlicki, “Data for development: the d4d challenge on mobile phone data,” arXiv preprint arXiv:1210.0137, 2012.
[130] S. Scellato, C. Mascolo, M. Musolesi, and V. Latora, “Distance Matters: Geo-social Metrics for Online Social Networks,” in WOSN, 2010
[131] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau, “CRAWDAD dataset cambridge/haggle,” CRAWDAD wireless network data archive, 2006.
[132] J. Burgess, B. Levine, R. Mahajan, J. Zahorjan, A. Balasubramanian, A. Venkataramani, Y. Zhou, B. Croft, N. Banerjee, M. Corner, *et al*, “CRAWDAD trace umass/diesel/transfer/spring,” 2006 (v. 2007-12-02). Downloaded (2007).
[133] Y. Zheng, Q. Li, Y. Chen, X. Xie, and W. Ma, “Understanding mobility based on GPS data,” in Proceedings of international conference on Ubiquitous computing, 2008, pp.312-321.
[134] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau, “CRAWDAD dataset cambridge/haggle” 2009 (v. 2009-05-29), http://crawdad. org/cambridge/haggle/20090529/imote traceset: imote.
[135] N. Eagle, A. Pentland, and D. Lazer, “Infering social network structure using mobile phone data,” in Proceedings of National Academy of Sciences, 2006.
[136] X. Liu, Q. He, Y. Tian, W. Lee, J. McPherson, and J. Han, “Event-based social networks: linking the online and offline social worlds,” in Proceedings of ACM SIGKDD international conference on Knowledge discovery and data mining, 2012, pp.1032-1040.
[137] B. Guo, Q. Han, H. Chen, L. Shangguan, Z. Zhou, Z. Yu. "The Emergence of Visual Crowdsensing: Challenges and Opportunities", IEEE Communications Surveys and Tutorials, vol. 19, no. 4, 2017, pp. 2526-2543

Bin Guo is a professor from Northwestern Polytechnical University, China. He received his Ph.D. degree in computer science from Keio University, Japan in 2009 and then was a post-doc researcher at Institut TELECOM SudParis in France. His research interests include ubiquitous computing, mobile crowd sensing, and HCI.

Yan Liu received the BE degree from Northwestern Polytechnical University, Xi’an, China. She is currently a PhD student with the School of Computer Science, Northwestern Polytechnical University. Her research interests include mobile crowd sensing and spatial crowdsouring.

Leye Wang is currently a postdoctoral fellow in Hong Kong University of Science and Technology. He obtained his Ph.D. from Mines-Télécom/Télécom SudParis and Universite Pierre et Marie Curie, France, in May 2016. He received his M.Sc. and B.Sc. in computer science from Peking University, China. His research interests include mobile crowdsensing, social networks, and intelligent transportation systems.

Victor O.K. Li is Chair of Information Engineering, Cheng Yu-Tung Professor in Sustainable Development, and Head of the Department of Electrical and Electronic Engineering, the University of Hong Kong. He received SB, SM, EE and ScD degrees in Electrical Engineering and Computer Science from MIT in 1977, 1979, 1980, and 1981, respectively. Previously he was Professor of Electrical Engineering at the University of Southern California (USC) and Director of the USC Communication Sciences Institute. His research is in the technologies and applications of information technology, including clean energy and environment, social networks, wireless networks, and optimization techniques. He is a Fellow of the Hong Kong Academy of Engineering Sciences, the IEEE, and the HKIE.



Jacqueline C.K. LAM is currently an associate professor at the Department of Electrical and Electronic Engineering, the University of Hong Kong. She received her PhD from Centre of Urban Planning and Environment Management, Faculty of Architecture, the University of Hong Kong, Hong Kong. She is also a Visiting Senior Research Fellow and Associate Researcher in Energy Policy Research Group, Judge Business School, the University of Cambridge. Her research interests cover clean energy and environment interdisciplinary study.

Zhiwen Yu is currently a professor at the School of Computer Science, Northwestern Polytechnical University, China. He has worked as an Alexander Von Humboldt Fellow at Mannheim University, Germany from Nov. 2009 to Oct. 2010, a research fellow at Kyoto University, Japan from Feb. 2007 to Jan. 2009. His research interests cover ubiquitous computing and HCI.