Chen YY, Lv P, Guo DK *et al.* A survey on task and participant matching in mobile crowd sensing. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 33(4): 768–791 July 2018. DOI 10.1007/s11390-018-1855-y

A Survey on Task and Participant Matching in Mobile Crowd Sensing

Yue-Yue Chen[1], Pin Lv[2,3,*], *Member, CCF, ACM, IEEE*
De-Ke Guo[4,5], *Distinguished Member, CCF, Senior Member, IEEE, Member, ACM*, Tong-Qing Zhou[1]
and Ming Xu[1], *Member, CCF, ACM, IEEE*

[1] *College of Computer, National University of Defense Technology, Changsha 410073, China*

[2] *School of Computer Electronics and Information, Guangxi University, Nanning 530004, China*

[3] *Guangxi Key Laboratory of Multimedia Communications and Network Technology*
Guangxi University, Nanning 530004, China

[4] *College of System Engineering, National University of Defense Technology, Changsha 410073, China*

[5] *School of Computer Science and Technology, Tianjin University, Tianjin 300072, China*

E-mail: yueyuechen@nudt.edu.cn; lvpin@gxu.edu.cn; {dekeguo, zhoutongqing, xuming}@nudt.edu.cn

Received August 24, 2017; revised May 12, 2018.

Abstract Mobile crowd sensing is an innovative paradigm which leverages the crowd, i.e., a large group of people with their mobile devices, to sense various information in the physical world. With the help of sensed information, many tasks can be fulfilled in an efficient manner, such as environment monitoring, traffic prediction, and indoor localization. Task and participant matching is an important issue in mobile crowd sensing, because it determines the quality and efficiency of a mobile crowd sensing task. Hence, numerous matching strategies have been proposed in recent research work. This survey aims to provide an up-to-date view on this topic. We propose a research framework for the matching problem in this paper, including participant model, task model, and solution design. The participant model is made up of three kinds of participant characters, i.e., attributes, requirements, and supplements. The task models are separated according to application backgrounds and objective functions. Offline and online solutions in recent literatures are both discussed. Some open issues are introduced, including matching strategy for heterogeneous tasks, context-aware matching, online strategy, and leveraging historical data to finish new tasks.

Keywords mobile crowd sensing, participant selection, task allocation, task and participant matching

1 Introduction

Mobile crowd sensing (MCS)[1-2] has attracted substantial attentions these years in both research literatures and commercial applications. With the increase in the availability of smart devices, including mobile phones, smart vehicles, wearable devices and so on, MCS has become more and more popular. The key feature of MCS is to recruit common participants and their smart devices to execute large-scale tasks, such as city monitoring[3-4], smart transportation[5-6], and emergency alarming[7]. A traditional way to finish such tasks is to employ dedicated staffs and devices to patrol around the city and collect required information; hence it costs much time and money. Currently, smart devices are equipped with a richer set of sensors, including GPS, cameras, accelerometers, microphones, gyroscopes, etc.[8] These sensors enhance the sensing capabilities of smart devices, and make it possible to get abundant sensing data from ordinary smart devices.

Survey

This work was partially supported by the National Natural Science Foundation for Outstanding Excellent Young Scholars of China under Grant No. 61422214, the National Natural Science Foundation of China under Grant Nos. 61402513, 61379144, and 61772544, the National Basic Research 973 Program of China under Grant No. 2014CB347800, the Hunan Provincial Natural Science Fund for Distinguished Young Scholars of China under Grant No. 2016JJ1002, the Natural Science Foundation of Guangxi Zhuang Autonomous Region of China under Grant No. 2016GXNSFBA380182, the Guangxi Cooperative Innovation Center of Cloud Computing and Big Data under Grant Nos. YD16507 and YD17X11, and the Scientific Research Foundation of Guangxi University under Grant Nos. XGZ150322 and XGZ141182.

*Corresponding Author

Moreover, the computing capability of portable equipments has become more and more powerful. With the help of massive widespread participants and their smart devices, an MCS system can collect information more economically and timely.

As illustrated in Fig.1, an MCS framework consists of three parts, i.e., the MCS server in the cloud, task publishers, and participants with smart devices. The task publishers upload tasks and requirements to the MCS server, and the MCS server is responsible of providing results to them. After receiving tasks from multiple task publishers, the server does some task management work before publishing them to participants, such as large task decomposing, similar tasks fusion and so on. The participants register their information to the server, if they are interested in some tasks. The server selects some participants as participants to execute related tasks. With abundant embedded sensors, participants are able to sense various data and help to finish the tasks. The MCS server is in charge of data collection, processing and providing service for the participants. Numerous participants can be either data providers or data consumers. That means the participants can provide data for the MCS server and request data service from it.

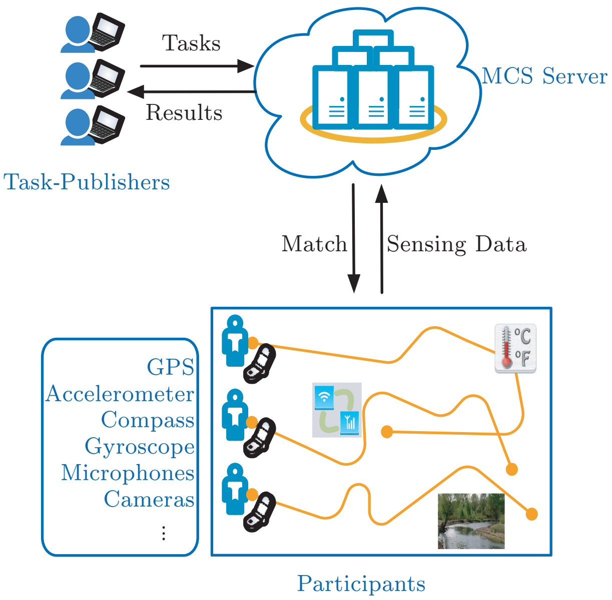

Fig.1. MCS framework.

From the MCS server's perspective, many tasks from multiple publishers need to be finished, while lots of participants wait to execute tasks at the same time. Therefore, it is important to match tasks and participants properly. However, it is also challenging to finish the matching due to multiple reasons. Firstly, both tasks and participants have many different attributes and requirements, and it is difficult to match proper participants and tasks. Secondly, mobility is an important feature of participants in the MCS framework, and it makes the matching problem more difficult. Thirdly, both tasks and participants usually appear dynamically, which brings up new challenges for matching in real time. To tackle these problems, many matching methods have been proposed and a great number of papers have been published in these years.

In this paper, we survey the up-to-date research issues about the task and participant matching problem, so as to plot the mainstream and emerging area of the matching problem. To the best of our knowledge, there is no previous survey paper about the matching problem in MCS. We propose a novel research framework for the matching problem in this survey, including participant model, task models, and solutions design. The related research work is surveyed according to the three parts in the framework as follows.

• For the participant model, the literatures are classified into three types based on different participant characters, including participant attributes, requirements, and supplements.

• For the task model, the literatures are classified into two types based on the sensing scope, i.e., area task and point task.

• For the solution design, the literatures are divided into online algorithms, and offline algorithms based on the input data availability from the start.

The remainder of the paper is organized as follows. In Section 2, the process details of MCS and the research framework of matching problem are introduced. In Section 3, multiple characters of the participants are discussed. Different task models based on task background and objective functions are listed in Section 4. The solutions are compared in Section 5. After reviewing lots of literatures, some future research directions are introduced in Section 6. Finally, Section 7 concludes this paper.

2 Preliminaries and Research Framework

In this section, we first introduce the preliminaries for the process of MCS and locate the matching problem among this process, and then illustrate the research approach in this paper and the main contents of the matching problem.

2.1 MCS Process

According to the MCS framework, the MCS process is divided into three steps, as shown in Fig.2, including participant recruitment, task execution, and data processing.

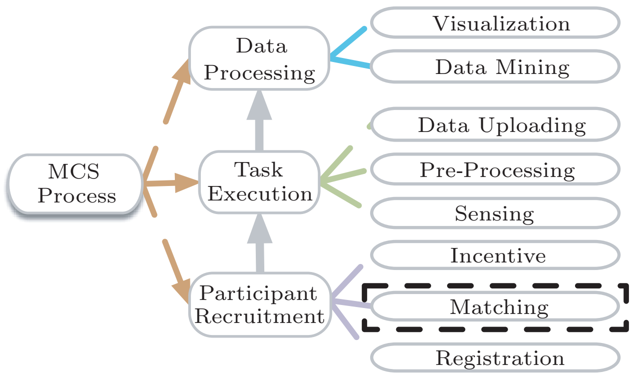

Fig.2. Three steps and related domains in the MCS process.

In the first step, to get sufficient high-quality sensing data, the MCS server needs to recruit proper participants for different tasks. Three substeps are involved in this step, i.e., registration, matching, and incentive. Incentive strategies focus on balancing benefit between the MCS server and participants. On one hand, participants are usually getting some payments for their task-execution cost. On the other hand, the budget of the MCS server is limited. Hence, there is a natural contradiction that both the server and participants want to maximize their own benefit. Some reviews about this domain[9-12] have been published. Task description is included in the registration step, and participants would sign in interested tasks. The matching between participants and tasks, marked with a dotted rectangle in Fig.2, is discussed in this paper. The matching substep focuses on selecting proper participants for proper tasks (or vice versa), and attributes of both participants and tasks should be considered in matching strategies. Hence, the matching substep is very fundamental in the whole MCS process.

The second step is task execution, which also includes three substeps. Substep 1 is information sensing with mobile sensors according to task requirements to get needed information. Substep 2 is data pre-processing (such as noise filtering, data quality enhancement) by smart devices to reduce uploading cost. Substep 3 is data uploading by different methods to the MCS server. The constraints considered in this step include energy[13], cost and hardware restraint[14].

The third step is data processing. Data mining based on information uploaded by participants is needed in this step. Moreover, most of the MCS systems may visualize their service and release APPs. Many applications have been designed in the MCS systems. For example, one kind of MCS applications is road-related. Koukoumidis *et al.* introduced a traffic schedule service based on traffic signals detecting and predicting[5]. Guo *et al.* introduced a shop profiling system through crowd sensing WiFi heat map and machine learning algorithms[15]. Zhou *et al.* predicted bus arrival time based on bus passengers' crowd sensing[16]. Morishta *et al.* aimed to find the flowering cherries along roads using mobile crowd data[6]. Some other kinds of applications are also introduced in recent papers. Cherian *et al.* designed a method to gauge the occupancy of in-door parking garages, by detecting driving status through cellphones[3]. Ludwig *et al.* proposed an emergency discovery system, which relies on crowd participants[7]. Guo *et al.* proposed to transfer the community bulletin boards to online information, which is meaningful for information sharing and propagation[4]. The constraints considered in this step include data reliability, error detection, and implementation effort.

In particular, the privacy[17-18] and the security[19] problems accompany the whole MCS process. Based on the introduction of the MCS process, we can see that, task execution is the main work of participants, and data processing is mainly the server's job, while the participant recruitment needs the efforts of both participants and the server.

2.2 General Framework for Solving the Matching Problem

As discussed before, the matching substep determines which data to select and how to collect data. Hence, the quality of sensing information collected by the MCS server heavily depends on the matching process. There are two main components in this substep, i.e., participants and tasks. Both participants and tasks have different characteristics and requirements. Therefore, to execute the matching process successfully, these characteristics and requirements have to be considered and discussed. In this survey, we propose a framework for solving the matching problem in this subsection. This framework depicts the basic research procedures for matching problem in MCS. The framework is illustrated in Fig.3. We summarize the research procedure

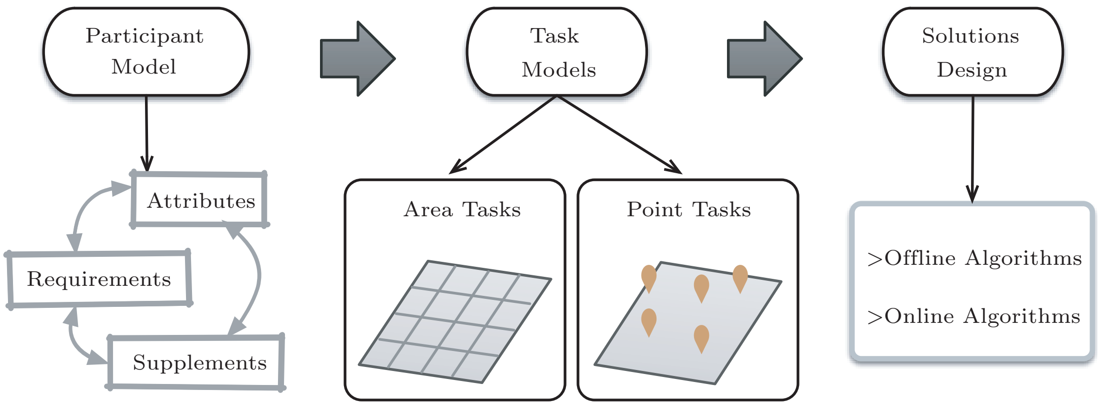

Fig.3. Research framework for participant and task matching problem in MCS.

into three parts, participant model, task model, and solution design.

In the first part, we propose a general participant model for the MCS process. We believe that most of the participant characters studied till now can be depicted by our participant model. The participant model consists of three elements, i.e., the participant attributes, the participant requirements, and supplements of the participants. The participant attributes depict the inherent characters of participants while the matching process happens. The requirements from participants are sent to the server, and they express the participants' interests in the matching process. The supplements of participants indicate the extrinsic factors, which may influence the selection results in the matching process. The combination of the three elements illustrates the status of participants in the matching process. More detailed discussion about the participant model is introduced in Section 3.

In the second part, we introduce several task models appeared in related literatures till now. We divide these task models into two main categories based on task sensing range, i.e., area tasks and point tasks. The reason is that the objective functions and requirements of these two kinds of tasks are always different. If the sensing scope of a task is an area, the task is called area task. For example, the task of sensing the PM 2.5 values in a city, and the task of finding traffic congestion in a downtown are all area tasks. A point task means that the task sensing scope is a point. For instance, the task of taking a picture of a church, and the task of collecting the noise value at the doorway of a bar, are point tasks. In the area tasks, the coverage is the most concerned problem, while the completion rate in the point tasks is discussed widely. In Section 4, we will introduce more find-grained task models within each main category.

Furthermore, solutions design with the given participant model and task models is crucial for the matching problem. We divide the existing solutions into two types, i.e., offline matching and online matching. If the participants and tasks are known in advance, the matching problem can be formulated into an offline matching problem. If the participants or tasks come to the MCS server dynamically, the matching problem is an online matching problem. Related algorithms about the two types of solutions used in recent papers are introduced and discussed respectively in Section 5.

3 Participant Model

Many participant characters have been discussed in recent literatures. To contain different characters of participants under different situations and to illustrate a participant comprehensively, we propose a general participant model, as shown in Fig.4. The participant model consists of three components, i.e., participant attributes, requirements, and supplements. The participant attributes depict the inherent characters of the participants, such as reputation and mobility. The attributes of a participant are determined by the participant itself. The participant requirements are the interests from participants to the MCS server or task publishers. For example, some participants may require to protect privacy, and some participants may require to select tasks by themselves (initiative) instead of being assigned tasks by the MCS server. The participants supplements indicate the extrinsic characters about participants, which are the complements besides the inherent characters and participant requirements. The supplements include social features of participants, the environmental context of them, etc. Actually, there are far more participant characters than those we list. However, most of the appeared characters belong to one of the three elements. In this section, we will introduce the three elements and related characters, respectively.

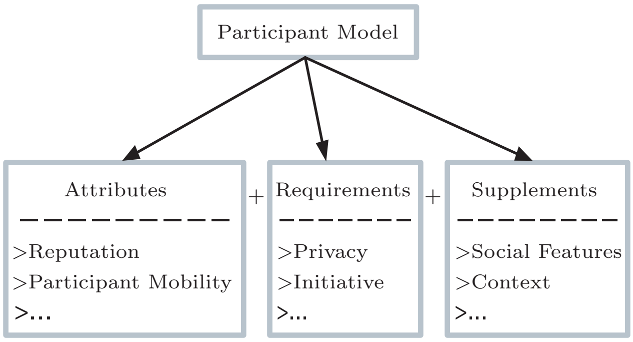

Fig.4. General participant model for MCS.

3.1 Participant Attributes

The participant attributes indicate the inherent characters of participants. These characters are determined once the participants come to the tasks. Two typical participant attributes are discussed in this subsection. The first one is the reputation, which impacts the quality of uploaded data. The second one is the participant mobility. The concerned factors include that, the needed participants are mobile or static, and the mobile trajectories of participants are alterable or not.

3.1.1 Participant Reputation

If the reputation of participants can be gotten in public, recruiters are able to pick the participants according to their reputation ranking. However, no uniform reputation record exists in mobile crowd sensing, even though some reputation records have been available in economic area. That is because evaluating the reputation of a participant is challenging, and computing methods of reputation can be various with different tasks and different participants.

Many reputation frameworks have been proposed. Huang *et al.* proposed a reputation system for evaluating the trustworthiness of participants, utilizing the Gompertz function to compute data quality uploaded by a device over a period of time[20]. A higher reputation score represents more reliable data in the past. Truskinger *et al.* concluded that the past performance of participants, the opinions of other participants, or a combination of both was used to calculate reputation[21]. A reputation framework, combining an initial score based on direct or indirect data sources and a performance score based on performance of the participant in current task, is proposed in [21]. Christin *et al.* considered the privacy of participants while establishing their reputation files, utilizing periodic pseudonyms and the reputation transfer[22]. Ren *et al.* introduced a bid price into the reputation system to evaluate the cost performance ratio of participants[23]. Mousa *et al.* reviewed the reputation frameworks clearly in mobile participatory sensing[24].

3.1.2 Participant Mobility

Participants in an MCS system can be either static or mobile. In a static situation, the selection strategies are easy to design. For example, if a task queries the temperature of an area at a certain time, the participants can upload their possible locations to the MCS server. However, in a mobile situation, how to select sensing data is challenging, because the trajectories of participants are sprawling. Usually, there are two ways to select sensing data from cluttered trajectories. The first way is to select participants and their whole trajectories, which may incur redundancy, since different trajectories may overlap in some segments. The second way is to select segments instead of the whole trajectories to avoid redundancy.

Multiple researches are based on the first way. Zhao *et al.* removed redundant participants by analyzing historical trajectories and calculating the coverage ratio of target area[25]. The constraint is the limited budget. He *et al.* selected participants based on predictable trajectories[26], and the selection goal is to maximize the spatial and the temporal coverage of the target area. The trajectory of each participant is supposed to be predicted, which can be achieved in many ways, such as navigation, periodic movement[27], or just uploaded by participants themselves. Gao *et al.* selected bus routes to satisfy coverage requirements[28]. The bus routes can be known beforehand since the bus routes of a city are usually changeless. All these papers select the whole trajectories of the participants to get sensing data. They all use the greedy selection strategy as one of the matching algorithms.

As mentioned before, selecting the whole trajectories is an intuitive method, while the drawback is obvious. The redundancy exists when the selected trajectories are overlapping, and the payment for the redundancy is a waste. As illustrated in Fig.5, suppose we want to monitor the environment condition of a city. After publishing the area scope to the crowd, three participants apply to contribute to the task. The trajectories of three participants are shown in Fig.5. If we need to recruit two of them due to the limited budget, data redundancy cannot be avoided no matter which two are chosen, because the first half of their trajectories are almost the same. Therefore, several studies are based on the second way, i.e., the segment selection. Trajectories can be divided into multiple segments, according to spatial distance or temporal distance. Based on spatial distance, segments belong to different squares. Boutsis and Kalogeraki proposed that when there were multiple segments within a same square, the MCS server assigned different costs per segment based on the availability of the segments[30]. Hence, the selection strategy can be designed to achieve a trade-off between the quality of sensing data and the cost. Hamind *et al.* discussed the segment selection problem in traffic conditions, where trajectories were along with roads, and the trajectory segments were determined by road segments[31]. Chen *et al.* proposed both offline and online segment selection mechanisms for predicted trajectories, and backward greedy algorithm was used in [29]. Based on temporal distance, the length and the pattern of different segments are different, and hence the selection problem is more challenging. However, to the best of our knowledge, few literatures focus on this issue. Zheng provided an overview for more methods to segment trajectories[32].

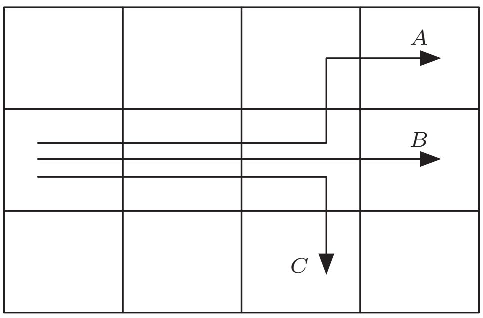

Fig.5. Example of trajectory redundancy[29].

3.2 Participant Requirements

The participant requirements indicate the need of the participants sent to the MCS server or task requesters. We discuss two of requirements, since these requirements have been repeatedly mentioned in recent literatures. The first one is the participant privacy, which should be protected in MCS applications. The second one is the participant initiative, which means the tasks are selected by participants themselves or assigned by the MCS server.

3.2.1 Participant Privacy

Smart devices are widely used nowadays. On one hand, the smart devices give us more and more convenience for our lives and work. On the other hand, the risk of our privacy leakage is getting increasingly higher. One important obstacle on the road of implementing mobile crowd sensing is that people do not want to share their information or sensing data. Therefore, protecting participant privacy when matching participants and tasks is the obligation of the MCS server.

Pournajaf *et al.* examined the problem of spatial task assignment in crowd sensing, when participants utilize spatial cloaking to obfuscate their locations[33]. Cloaking is a popular obfuscation way to protect the location privacy of the participants. It means that a coarse-grained location scope is used to represent a fine-grained location. The MCS server can only receive a scope instead of a precise location. They proposed a novel two-stage optimization approach which consists of a global optimization using cloaked locations and a local optimization using participants' precise locations without breaching privacy. Wang *et al.* adopted differential-privacy in sparse MCS to provide a theoretical guarantee for participants' location privacy[34]. Sparse MCS is a concept proposed by Wang *et al.*[35], which aims at selecting as few as possible participants to finish tasks, while guaranteeing the information quality. Differential privacy is another obfuscation mechanism to protect privacy, and it is widely used in security area. Wang *et al.*[34] argued that the effectiveness of cloaking was greatly impaired if the adversary had prior knowledge about the target participants' location distribution. Differential privacy means that an obfuscated location is mapped to an actual location, and the obfuscated location is uploaded to the MCS server instead of the actual location. As shown in Fig.6, the solid people icons indicate the actual locations, while the dotted people icons indicate the obfuscated locations, and the stars indicate the task allocation results. Data quality loss is discussed to balance privacy protection and data quality guarantee. Xiao *et al.* proposed a participant recruitment protocol to select the minimum amount of participants to guarantee the task quality, while protecting the privacy of the participants[36]. The protocol is based on a greedy strategy. The approximation ratio is analyzed, and the security of the protocol is proven.

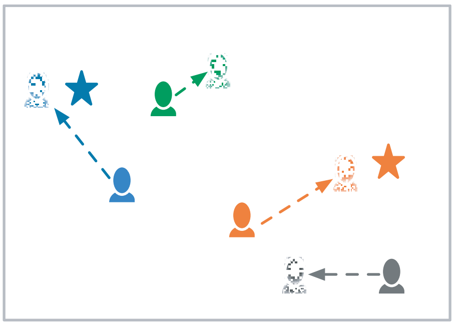

Fig.6. Location obfuscated method and task allocation in [34].

Celis *et al.* pointed out that participant privacy had attracted a lot of attentions, while the task privacy was ignored seriously[37]. Hence, to protect task privacy, it was proposed to split a whole task into multiple small components. Information loss functions were introduced to formally measure the amount of private information leaked as a function of the task assignment. Multiple assignment situations are also discussed.

3.2.2 Participant Initiative

In most literatures about the matching problem, the MCS server is responsible for selecting and assigning tasks to each participant. However, the participants can be active enough to select tasks themselves in some situations, instead of waiting to be selected as participants by the MCS server. The MCS server releases many points of interest (PoIs) to participants, and participants can choose interested ones according to their plans and upload their options. In this situation, participants do not need to expose their private information, but have to negotiate with the MCS server about their options. Cheung *et al.* proposed an asynchronous and distributed task selection strategy to help the participants choose PoIs on their own, and the selected tasks form their trajectories[38]. The MCS server releases tasks with locations, deadlines and payoffs. The participants need to upload their starting location and ending location. The participants select tasks which can be finished before the deadline. A non-cooperative game based selection process was proposed. Zhang *et al.* proposed a novel taxi orders dispatch model based on a popular ride-sharing system called DiDi[39]. The passenger orders are dispatched to multiple drivers, and each driver determines whether or not to accept the dispatch. Many factors may influence the decision of drivers, and a machine learning method is used to model the decision results.

Another way is to neutralize the initiative of both the participants and the MCS server. Celis *et al.* proposed an intermediate approach called Tug Of War (TOW), which balanced flexibility for both the participants and the MCS server[37]. The tasks are supposed to be split into multiple components, and each component can be independently completed by a user. The authors[37] presented selection strategies for all three selection methods, and analyzed the tradeoffs.

3.3 Participant Supplement

The supplements of participants indicate the extrinsic characters related to participants, including the social features and contexts. The social features can help to describe the participants, since the social activities are important components of people's lives. The contexts of participants describe the surround features around participants or their sensors. The contexts can help to match participants more precisely.

3.3.1 Social Features of Participants

With the rapid development of mobile social network, human life is becoming more and more closely linked to their social behavior. Cho *et al.* pointed out that users' long-distance movements were determined by their social networks, while short-distance trajectories were daily repeated[27]. Furthermore, participants can be distinguished by their social tags, such as location stamps and interests. Therefore, selection strategies can be designed according to their social information. Two kinds of frameworks for matching problem based on social network are introduced in the following.

One way is to select participants based on their social tags. In this situation, the MCS server is supposed to have the needed information of participants, which can be uploaded by participants themselves or mined from public data. Cardone *et al.* proposed a geo-social crowd sensing platform, profiling participants with time, location, social and other information[40]. The participants are selected for different tasks according to their profiles. Ren *et al.* proposed a social-aware selection framework, considering the social tags, task delay and reputation in crowd sensing[23]. Multiple social tags, such as sporting (interest tag), Toronto (location tag), and so on, are related to participants. Different tasks can select corresponding tags according to requirements, and hence select related participants. Similarly, some selection mechanisms were proposed based on the expertise or the speciality of the participants.

The expertise and the speciality indicate that a participant may only have expertise or speciality in some certain domains. Pu *et al.* selected participants according to their speciality tags, indicated by task attributes of categories and keywords[41]. Mavridis *et al.* modeled the tasks and the participants using a skill tree, and proposed a task assignment mechanism using hierarchical skills for the model[42]. Zhang *et al.* formulated an expertise-aware task allocation problem[43]. The expertise of a task or a participant is obtained by semantic analysis, including semantic information extraction and dynamic hierarchical clustering. The random observations of participants for a task are assumed to follow a normal distribution. Hence, the formulated task allocation problem can be solved by the EM (expectation-maximization) algorithm.

The other way is to identify participants through multi-hop friendship relations. It means that the tasks can be finished by friends and friends of the friends. Amintoosi and Kanhere proposed a trust-based selection framework based on social network, and a route selection strategy was also put forward[44]. Chang and Wu pointed out that the absence of a relay worker might disrupt the information flow in a task finishing process, and proposed several data collection strategies based on two special social structures[45]. One social structure is triangle relation among any three consecutive participants, and the other is quadrilateral relation between two intersected workflows. In such a selection framework, the route information and the route selection method are key research points.

3.3.2 Contexts of Participants

Context includes the participant status and the situations of surrounding environment. Participant status includes motion status (running, walking, etc.), movement status (location, direction, etc.), and so on. Surrounding status includes the number of companies, the environment categories (park, gym, home, etc.), and so on.

To select participants according to their context, the context information should be detected successfully. With the help of the rich set of embedded sensors in smart devices, context detection is becoming more and more accurate. Nath proposed a user status sensing framework based on a heterogeneous architecture[46]. By combing various sensor data, user status (e.g., IsDriving, IsWalking, AtHome) can be obtained. Furthermore, by mining relationships among various context attributes, more elaborated context can be deduced. Much energy can be saved due to the deducing process, because sensing, specially GPS sensing, is energy-consuming. More context detection methods can be seen in [8, 50].

Reddy *et al.* assumed that user trajectories are context annotated, and designed a data collector selection strategy based on the annotated context[47]. Tamilin *et al.* presented an integrated realtime civic awareness and engagement platform, and pointed out that user context was an essential element to determine whether the user was in conditions relevant to tasks[51]. However, the user context is represented by sensing capacities of participants, which is incomplete. Zhang *et al.* discussed the context-aware participant selection mechanism, and indicated that task utility was dynamic because participants can log in or out at any time[48]. The authors assumed that the user context can be represented by logging out probability, and the selection strategy was designed aiming to maximize the long-term task utility. Liu *et al.* proposed to select participants according to their context[49]. The ground truth of the sensing data is difficult to get, yet the context of the participants can be detected. The data quality is related to the context. Hence, the context can be a measurement to help matching. For example, the data quality is different when a participant is sitting, walking, or running. A context-quality classifier is trained to capture the relation between context information and the data quality, and the classifier is applied to guide participant recruitment.

Context is an important factor in the MCS system, and behavior recognition has been studied in many literatures. Nevertheless, selection strategies based on user context have not been researched deeply. There are various reasons. Firstly, the context changes rapidly for each user in each time slice, and continual behavior detection is energy-consuming. Secondly, surrounding situation is hard to be determined just by device sensors. For example, it is difficult to detect whether a user is in a park or on a road. Thirdly, too many context categories exist in human daily life, and no uniform taxonomy is defined. With all those obstacles, more efforts are still needed in context-aware selection framework.

3.4 Summary of the Participant Model

The summary of literatures about the participant model can be seen in Table 1. Several conclusions are drew based on Table 1 as follows. Firstly, most of the related literatures focus on one of the participant char-

Table 1. Summary of Different Characters in Participant Model

Paper	Attribute		Requirement		Supplement	
	Reputation	Mobile	Privacy	Initiative	Social	Context
Huang *et al.*[20]	√					
Truskinger *et al.*[21]	√					
Christin *et al.*[22]	√		√			
Ren *et al.*[23]	√				√	
Zhao *et al.*[25]		√				
He *et al.*[26]		√				
Gao *et al.*[28]		√				
Boutsis and Kalogeraki[30]		√				
Hamid *et al.*[31]		√				
Chen *et al.*[29]		√				
Pournajaf *et al.*[33]		√	√			
Wang *et al.*[34]		√	√			
Xiao *et al.*[36]		√	√			
Cheung *et al.*[38]		√		√		
Zhang *et al.*[39]		√		√		√
Celis *et al.*[37]		√		√		
Cardone *et al.*[40]					√	
Pu *et al.*[41]					√	
Mavridis *et al.*[42]					√	
Zhang *et al.*[43]					√	
Amintoosi and Kanhere[44]					√	
Chang and Wu[45]					√	
Nath [46]						√
Reddy *et al.*[47]		√				√
Zhang *et al.*[48]		√				√
Liu *et al.*[49]						√

acters, since it is difficult to consider all the characters. Only [39] combines the three kinds of participant characters, i.e., attributes, requirements, and supplements. Secondly, participant mobility is considered in most related papers. The privacy protection strategies and the initiative strategies are always associated with the participant mobility. The participant mobility provides more opportunities for ingenious privacy protection and initiative strategies. Thirdly, the social characters of participants are always considered alone instead of being integrated into other characters. Therefore, combinational optimization based on multiple participant characters is a future research direction.

4 Task Models

As introduced in the research framework in Section 2, we divide the task models into two main categories, i.e., the area tasks and the point tasks. The two categories have been discussed repeatedly in recent literatures. The accordance of the classification is the task's spatial scope, which is an area or a point. In this section, the fine-grained task models that belong to each category are discussed. We classify the task models under each category based on different objective functions, and the classification result can be seen in Fig.7. For the area tasks, the coverage of the target area is the most important problem. Baseline coverage functions include basic spatial-temporal coverage and weighted coverage. Other definitions about the area quality other than coverage are also discussed. For point tasks, the objective functions can be divided into three kinds, i.e., completion rate maximization (CRM), CRM with uncertain participants and participant-task utility optimization. Specially, there are some other kinds of task models, which have not been discussed so much in recent literatures. We will introduce them briefly at the end of this section.

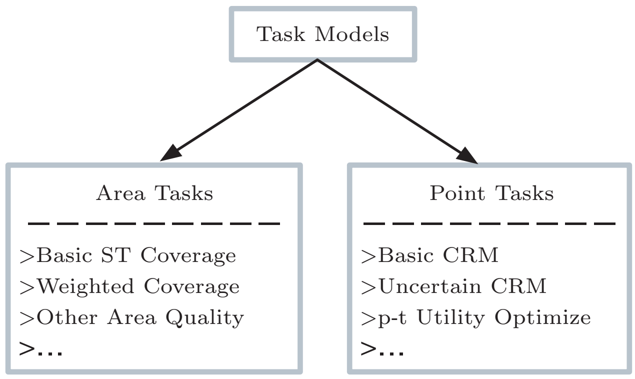

Fig.7. Task model classification in MCS.

4.1 Area Task

An area task always concerns the whole spatial area during a certain period. However, it is difficult to collect all data of the whole target area. One reason is the limited budget. It is costly to recruit enough participants to cover the whole task area. Another reason is regional inequality. Sometimes there are no participants at all in some remote areas. Hence, to meet different tradeoffs between the task requirements and the participant characters, it is important to decide different matching objective functions.

After reviewing recent literatures about the matching problems in area tasks, we divide the related objective functions into three categories, i.e., basic spatial-temporal coverage, weighted coverage, and other definitions of area quality. The concern of our classification basement about the coverage objectives is whether the sub-areas are treated equally. In the basic spatial-temporal coverage, all the divided areas in the whole task area are treated equally. In the weighted coverage, different divided areas need different kinds of participants. Moreover, some novel definitions of area quality, besides the coverage objectives, are introduced in the last category.

4.1.1 Basic Spatial-Temporal Coverage

One basic spatial-temporal coverage standard is the coverage ratio, which means the ratio of sensed area to the whole area. In order to maximize the spatial-temporal coverage of participants within a given budget, a common method can be illustrated as follows. The target area is divided into multiple squares, and the time period is divided into multiple time slices. If the amount of given budget cannot afford all the participants, a subset of them should be selected which can maximize the coverage ratio. He *et al.* formulated the participant trajectories into a spatial-temporal matrix, and defined the matching problem as maximizing spatial coverage or temporal coverage separately by selecting proper participants based on predicted trajectories[26].

Another basic spatial-temporal coverage standard is uniform coverage, which means the collected data should be uniformly distributed in the whole area. However, measuring the uniform degree of participant distribution is challenging. Ji *et al.* proposed a hierarchical entropy-based objective function to address this challenge[52]. Coverage ratio can be different according to different partition granularities, as shown in Fig.8. The distributions in Fig.8(a) and Fig.8(c) are identical, while the partition granularities are different. Fig.8(b) and Fig.8(d) also have the same distribution and different partition granularities. Fig.8(a) and Fig.8(b) are in the fine-grained partition granularity, and their coverage ratios are the same, which is 1/4, although their distributions are different. Fig.8(c) and Fig.8(d) are in the coarse-grained partition. The coverage ratio is 1/4 in Fig.8(c), while it is 1 in Fig.8(d). The entropy mean of multiple granularities is selected as the uniform degree measurement. Furthermore, a parameter is used to tune up coverage uniformity and covered square number. To maximize the proposed measurement function, a graph-based task selection strategy is designed according to starting and ending locations of the participants, and an iterative participant selection process is defined. However, the entropy only represents the amount of information inside the target area, rather than spatial location relationship between participants. Hence, a more accurate uniform measurement function is needed. Zhao *et al.* further divided time slices into multiple sampling periods, and put forward that one square should only be sampled once (or predefined times) in a time slice[25]. To reduce sampling redundancy, a greedy strategy is proposed based on historical trajectory matrix. However, if the number of candidates is so small that many squares cannot be sensed even once during a sensing time slice, this selection method cannot work any more.

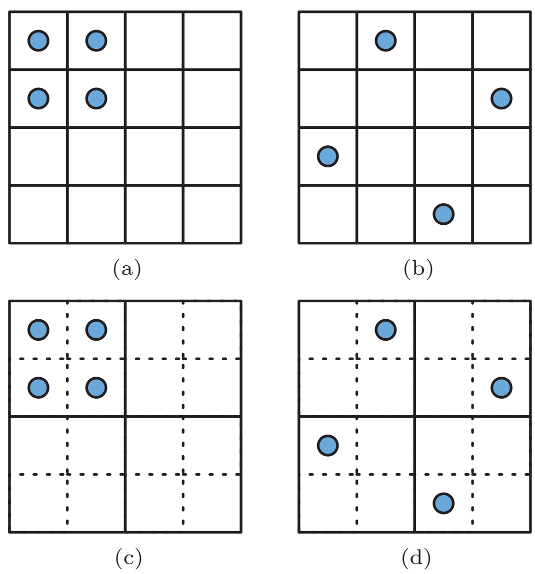

Fig.8. Same distribution showing different uniform degrees in different partition granularities[52].

4.1.2 Weighted Coverage

Sometimes the matching strategy should be designed based on area attributes, such as population, region ranking, and density requirement, instead of treating all the areas equally. For example, more people an area contains, more precise the monitoring should be. Hence, more participants should be selected within the area.

Jaimes *et al.* proposed a geometric coverage model where the weight of a sensor was the sensor number within its coverage[53], as illustrated in Fig.9. The authors aimed to find a subset of sensors whose union covers all the sensors. A combination of two kinds of greedy strategies is designed to tackle this problem. However, this measurement function may incur that participants gather in densely-populated areas, and too few or even no participants are selected in the area if the participants are distributed sparsely. Therefore, the weighted coverage should be combined with uniform coverage to illustrate a more reasonable coverage measurement function. Mendez and Labrador argued that the size of selected participant set for each area should be based on the variability of the interested factor in that area[54]. If a factor is fairly constant in a region, which can be constructed by historical data, a low number of participants should be selected, and vice versa. The weight is the temporal variety of interested parameters, which is always ignored by most other papers. Hence, besides the spatial variety, the temporal variety, such as the time slice length, temporal uniform coverage, or temporal weighted coverage, should be discussed more in the future. Song *et al.* proposed that the quality of information (QoI) of a square in a time slice can be represented by the number of sensors needed for each task[55]. The question is how to select the minimum set of participants to satisfy the QoI requirements of the whole target area within a given budget. The question can be transformed into a nonlinear knapsack problem. A mobility formulation based on probability and dynamic greedy selection strategy is designed to solve this problem.

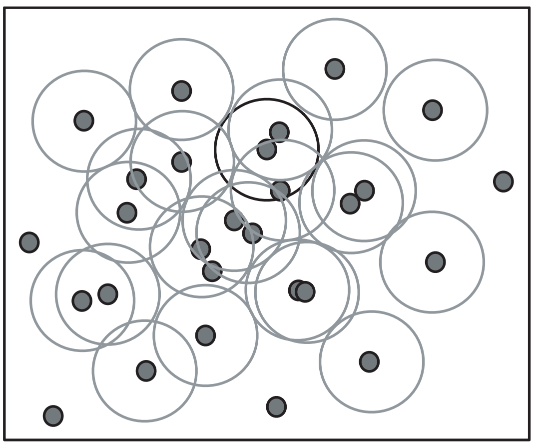

Fig.9. Geometric coverage model proposed in [53].

4.1.3 Other Definitions of Area Quality

Liu *et al.* proposed a new metric to measure the sensing quality of a target area, which was called urban resolution[56]. Like the concept of image resolution, the urban resolution is the number of gridded sensors. Statistic methods are used to explore the relationship between the number of sensors and the number of gridded sensors. Kang *et al.* further enhanced the sensing quality via data correlation[57], where a tensor decomposition was used to analyze and rebuild the sensing data.

Wang *et al.* tried to minimize the sensing squares because of the limited budget, while guaranteeing the information quality of the whole task area simultaneously[60]. The compressing sensing strategy (STCS)[58] is used to deduce unsensed cells. Static participants are supposed to be massive enough to cover all the squares. One square is selected to be sensed at one sensing cycle, until the information quality reaches the defined threshold. Leave-one-out strategy is used to measure the error rate of the reconstruction strategy, and a Bayesian inference is used to compute the posterior probability distribution of sensing errors.

Wu *et al.* leveraged the metadata of photos, including location, orientation, field of view, and range of a camera, to define photo utility[59]. The utility is used to measure how well a target area is covered by a set of photos. Based on the defined utility, a photo selection algorithm was proposed to achieve constant coverage ratio under a resource budget.

4.1.4 Summary of Matching in Area Tasks

The comparison of different matching strategies in area tasks is illustrated in Table 2. As discussed before, the basic spatial-temporal coverage metrics include coverage ratio[26] and uniform coverage[25,52]. In all the three papers, the participant information is assumed to be known before matching. However, accurate trajectory prediction is difficult to achieve. Hence, coverage ensuring with uncertain participants should be discussed in the future. In the weighted coverage, several distinctions between squares have been discussed in

Table 2. Comparison of Different Matching Strategies in Area Tasks

Paper	Summary	Assumption	Objective	Constraint	Mathematical Tool(s)
He *et al.*[26]	Participant recruitment strategy for vehicle-based crowdsourcing	Known trajectories	Max. spatial or temporal coverage	Budget	NP-complete proven; greedy approximation; genetic algorithm
Ji *et al.*[52]	Crowd-based urban sensing framework	Known departure/ destination locations, and time of trajectories	Uniform coverage with minimal participants	Budget	Dynamic programming
Zhao *et al.*[25]	Cooperative sensing and data forwarding framework	One square sampled once in a time slice	Reduce sampling redundancy	Budget	Greedy selection
Jaimes *et al.*[53]	Incentive mechanism in participatory sensing system	Static participants	Min. participant number whose union covers all participants	Budget	Greedy selection
Mendez and Labrador[54]	Method to determine participant distribution	Static participants	Improve the area estimations with a fixed participant number	N/A	N/A
Song *et al.*[55]	QoI aware dynamic participant selection	Mobile participants	Min. participants to satisfy QoIs of all squares	Budget	Nonlinear knapsack; dynamic greedy selection
Liu *et al.*[56]	A new metric to measure the sensing quality of a target area	Known distribution of sensors	N/A	N/A	Pareto distribution; Monte Carlo simulation
Wang *et al.*[60]	Leverage data reconstruction to reduce the number of sensing cells while ensuring the data quality	Static and unchanged participants	Minimize the number of sensing cells	Quality threshold	Compressive sensing; Bayesian inference; leave-one-out re-sampling
Wu *et al.*[59]	Method to crowdsource photos that best cover the target area	Photo metadata which can be obtained	Max. defined coverage utility of selected photos	Resource budget	Greedy selection

Note: Max. means maximize; min. means minimize.

recent literatures, including the participant density[53], the factor variability[54], and the quality requirement of each square[55]. More differences should be concerned in the future. Other definitions of area quality include urban resolution[56], reconstructed data quality[60], and visual coverage[59]. All three definitions are novel and interesting, and more follow-on work should appear.

4.2 Point Task

In the mobile crowd sensing system, each point task is always related to a specific location, and the participants should arrive at the location to finish the task. In most literatures about the matching strategies of point tasks, multiple tasks instead of single task are always discussed together. As different point tasks are related to different locations and participants can be static or mobile, the matching strategies are variable.

After surveying the related literatures, it is found that two important indicators are often used as optimization objectives, i.e., the completion rate maximization (CRM) and the participant-task utilities optimization (PTUO). The former one aims to complete as many tasks as possible, and the latter one aims to optimize a global participant-task utility defined according to applications. Based on this discovery, we divide the task model into three sub-models, i.e., basic CRM, CRM with uncertain participants, and PTUO. In the sub-model of basic CRM, some basic objective functions with certain participants are introduced. In the sub-model of CRM with uncertain participants, variable formulations of participant uncertainty are introduced. In the sub-model of PTUO, we focus on the different definitions of participant-task utilities. At last in this subsection, we discuss the research status about the tasks models in point tasks.

4.2.1 *Basic Completion Rate Maximization*

The completion rates of all the target point tasks are discussed widely in recent literatures. The participants in this matching problem are always supposed to be mobile.

If the trajectories of participants can be known or predicted, the matching problem is an optimization problem. Liu *et al.* analyzed two typical situations: few participants more tasks (FPMT) and few tasks more participants (MPFT)[61]. They proposed several matching strategies for both situations. In the former situation, the objective function is to maximize the total number of accomplished tasks. In the latter situation, the goal is to minimize the total number of selected participants. Similarly, Guo *et al.* proposed two kinds of situations, where all tasks should be finished[62]. In the first situation, the tasks are delay-tolerant, and the participant trajectories are predicted. The delay-tolerant task means that the deadline of the task is loose, and there is no need to upload sensing data in real time. The objective function under this situation is to minimize the total number of selected participants. In the second situation, the tasks are time-sensitive, and the participant movements are intentional. Time-sensitive means the sensing data should be uploaded to the MCS server in real time. Intentional movement means participants can move arbitrarily during the task execution. The objective function under this situation is to minimize the total distance between participants and tasks.

Gao *et al.* aimed to monitor PM 2.5 of several important locations within a city[28], such as schools and hospitals. The monitoring nodes are deployed on buses, since the bus routes are distributed in the whole city and are relatively stable. However, it is expensive to install monitoring nodes on every bus. To design bus selection strategy, the city area is divided into multiple squares, and each square has an importance value based on the distances between them and those important locations. Moreover, each square can be covered by several bus routes or several time slices; hence there is a coverage degree for each square. As illustrated in Fig.10, the number in each square indicates the importance, and the color in each square indicates the coverage degree. The weight of each square is the product of importance and coverage degree, and the objective function is to maximize the weight sum of the selected important locations (PoIs).

If the trajectories of participants cannot be predicted, the matching problem becomes more complex. In most related literatures, some of restraints are set for participants. For instance, the original location, the destination, the start time, and the end time of participants are known. The MCS server can plan trajectories for participants under their permits. With this precondition, Kang *et al.* proposed that each task must be executed multiple times to ensure the reliability of the sensing result[63]. The objective function of this problem is to maximize the number of finished tasks with a certain reliability by assigning appropriate sets of tasks to participants. The trajectories of participants in this problem are assumed to be alterable. The allocated tasks for a participant determine the trajectory of this participant. Offline and online matching algorithms are discussed, respectively, and the competitive ratio of proposed algorithms are calculated.

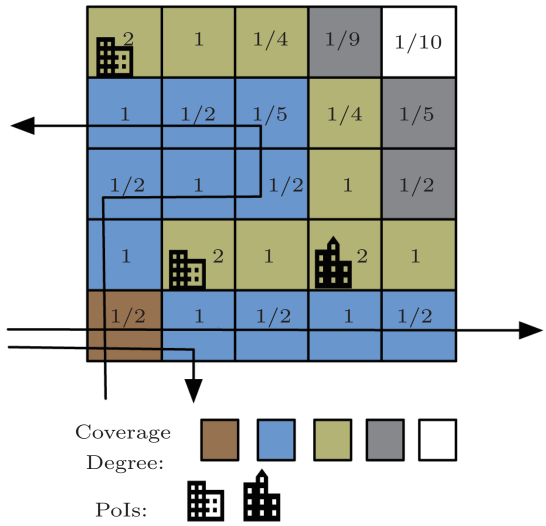

Fig.10. Weight calculation of PoIs in [28].

In some other literatures, only static participants are considered. Kazemi and Shahabi defined the completion rate maximization problem, and three alternative solutions were proposed to address this problem, including a greedy strategy, a least location entropy priority strategy, and a nearest neighbor priority strategy[64].

4.2.2 *Completion with Uncertain Participants*

In many situations, it is uncertain about whether or not a sensor (or a participant) covers a task. Hence, the probability of a task being finished by uncertain participants should be considered. We introduce the completion rate maximization problem with uncertain participants.

Piggyback Crowdsensing (PCS) was proposed by Lane *et al.*[65], which is an energy-efficient model leveraging smartphone application opportunities to perform sensing and data uploading. Based on PCS, many matching strategies aiming to maximize probabilistic matching were proposed. It was proposed that area coverage constraints can be replaced by cell tower coverage constraint, since cell towers are distributed among cities[66-68]. A high covering percentage of cell towers in a given region ensures that most part of the given area is covered. The call sequence of a user is assumed to be an inhomogeneous Poisson process, and the intensity can be estimated as the mean number of related historical data. After that, the probability of a cell tower being covered by a user can be computed. The matching problem is formulated as maximizing the coverage rate of cell towers at each time slice. Furthermore, Xiong *et al.* aimed to achieve the full cell tower coverage, and presented several participant selection methods based on the calculated coverage probabilities[69]. The coverage probability of adjacent squares can be computed in some methods, and more complicated models based on this precondition were studied.

Wang *et al.* also computed the coverage probability based on historical data from a telecom operator, and then further converted the matching problem into the representation of a bipartite graph[70]. An iterative greedy process is employed to optimize the matching problem. Similarly, based on the piggyback model, Li *et al.* formulated the coverage probability of a square by a user to be the probability of making calls by the user in this square[71]. The total coverage rate of square is the probability sum of selected participants. The probability of making calls by a user in a square can be computed according to historical data. Li *et al.* introduced a new MCS architecture which leverages the cached sensing data to fulfill partial sensing tasks in order to reduce the size of selected participant set[72].

Chen *et al.* studied the stochastic task recommendation problem[73]. Each participant is associated with several predicted trajectories, and the probability of each trajectory can be calculated. The goal is to maximize the expected utility of all tasks achieved by selected participants. Furthermore, Xiao *et al.* proposed a deadline-sensitive task allocation framework with uncertain participants[74]. Each participant is associated with multiple probabilistic trajectories. Several participants might be recruited to cooperatively perform one task, to ensure that the task can be finished before the deadline. The objective function of this paper is to maximize the finish probability of all tasks within each deadline. The problem was proven to be NP-hard, and several approximation algorithms were proposed in this paper. The approximation ratios were also calculated.

4.2.3 Participant-Task Utilities Optimization

Besides the completion rates of tasks introduced above, some participant-task utilities have been discussed in recent literatures as the objective functions. The definitions of participant-task utilities are always associated with application requirements. The utilities combine both the participant characters and the task objectives.

Tong *et al.* proposed to minimize the distance between the participants and the tasks[75]. The bipartite graph was used to formulate the matching problem between the participants and the tasks. The two-sided online situation was considered, where both participants and tasks come to the MCS server dynamically. Two approximation algorithms were discussed. Lee *et al.* proposed to minimize the total driving time from drivers to passengers in the online taxi-hailing systems[76]. Real-time traffic conditions were considered. The previous dispatch strategies in DiDi① are to maximize the passenger-driver scores[39]. The score is calculated through a learning-to-rank based method. Yu *et al.* proposed a participant selection problem for offline event marketing[77]. Three important factors were considered, i.e., distances between participants and event locations, overlapping social influence of participants, and item coverage of participants. The participant selection problem was transformed into a combinatorial optimization problem with the objective function of the marketing effect maximization.

Cheng *et al.* pointed out that previous studies on the matching strategies that maximize the participant-task scores were only based on the present information[78]. Thus, a prediction-based matching problem was proposed to maximize a global matching score, including the present matching and the future matching. A grid-based prediction method was designed, and a heuristic method was proposed to tackle the matching problem.

4.2.4 Summary of Matching in Point Tasks

The comparison of different matching strategies in point tasks is illustrated in Table 3. CRM is a traditional problem in point tasks, and we introduce different studies on this topic, including CRM with static participants[64], with mobile participants[61-62], and with mobile participants and heterogeneous tasks[63], respectively. To support these researches, trajectory prediction algorithms are needed. In the sub-model of CRM with uncertain participants, we introduce several kinds of formulations. The first kind is to formulate the uncertainty of participants as the probabilities of calls[66]. The second kind is to assume that there are several possible trajectories of each participant[73-74]. The third kind is to assume that the coverage of each sensor follows a Gaussian distribution[28]. In the sub-model of PTUO, the participant-task utilities are determined by applications. Several utili-

① http://www.didichuxing.com, Apr. 2018

Table 3. Comparison of Different Matching Strategies in Point Tasks

Paper	Summary	Assumption	Objective	Constraint	Mathematical Tool (s)
Liu *et al.*[61]	Multi-task allocation framework	Known mobile participants	Max. the number of finished tasks; min. total payments	Movement distance	Minimum cost maximum flow; multi-objective optimization
Guo *et al.*[62]	Multi-task allocation framework	Known mobile participants	Min. movement distance; min. the number of participants	Time sensitive or not	Greedy-enhanced genetic selection
Kang *et al.*[63]	Quality-aware online task allocation	A task must be executed multiple times; alterable trajectories	Max. the number of tasks exceeding quality threshold	Worker capacity; cost budget	Greedy selection; branch and bound
Kazemi and Shahabi[64]	Completion ratio maximization	Static participants	Completion ratio maximization	Budget	Greedy selection; entropy; nearest neighbor
Zhang *et al.*[66]	Participant selection framework for piggyback crowdsensing	Piggyback; known historical records of participant calls	Min. participant number to meet coverage requirement	Resource budget	Poisson process; greedy selection
Xiong *et al.*[69]	Participant selection framework for piggyback crowdsensing	Piggyback; known historical records of participant calls	Min. participant number to full cover the task area	Resource budget	Poisson process
Wang *et al.*[70]	Participant selection framework on a multi-task sensing system	Piggyback; known historical records of participant calls	Max. defined utility of participant and task matching	Resource budget	Poisson process; bipartite graph; greedy selection
Li *et al.*[71]	Offline and online algorithms for piggyback crowdsensing	Piggyback; known historical records of participant calls	Max. coverage probability	Resource budget	Poisson process; greedy selection
Chen *et al.*[73]	Task allocation with uncertain trajectories	N/A	Max. the expected utility	Time budget	Lagrangian relaxation
Xiao *et al.*[74]	Deadline sensitive task allocation with uncertain trajectories	N/A	Max. the finish probability of all tasks within each deadline	Task deadline	Non-trivial set cover problem; greedy selection
Gao *et al.*[28]	Novel air quality monitoring system by deploying monitoring nodes on buses	N/A	Maximize the weight sum of the selected locations	Budget	Greedy selection
Tong *et al.*[75]	Matching in two-side online situation	Participants and tasks come to the server dynamically	Min. the distance sum between participants and tasks	Budget	Weighted bipartite graph; Hungary algorithm; greedy selection
Yu *et al.*[77]	Participant selection framework for offline event marketing	Known participant information	Max. the marketing effect	Budget	Greedy selection
Cheng *et al.*[78]	Prediction-based matching	Static participants	Max. the present and future matching scores	Budget	Greedy selection; divide-and-conquer

Note: Max. means maximize; min. means minimize.

ties are introduced in this paper, such as the total distance minimization[75], the total driving time minimization[76], and the total scores considering different factors maximization[39,77]. Moreover, the future utilities optimization is also discussed[78].

4.3 Other Kinds of Tasks

4.3.1 Joint of Area Tasks and Point Tasks

Zhang *et al.* pointed out that the optimization goal of a point task was usually to maximize the completion rate, while that of an area task was usually to maximize the coverage[79]. The authors proposed to combine these two goals by leveraging the two kinds of tasks. The coverage scope of participant trajectories formed by the participants' traveling to the point tasks can be used to improve the quality of area tasks. A task management framework was designed to efficiently match participants to the combined tasks. More researches about the joint of area tasks and point tasks should be made in the future, since the task fusion is an efficient way to improve the matching quality.

4.3.2 Line Tasks

In vehicle-based MCS applications, the coverage of paths is significant, because road-based services, such as congestion warning and accident alarming, can be improved through coverage information analysis. Hamid *et al.* proposed a greedy vehicle-trajectory selection strategy aiming to maximize the coverage ratio of a given path[31]. Since a vehicle may not stick to the trajectory it announces, some applications require redundancy at some part of the given path (for higher accuracy). They introduced a leaving probability for the former situation, and a coverage degree for the latter situation, to improve the selection strategies separately. More detailed trajectory related formulations and computations are discussed in [32].

4.3.3 Visual Crowdsensing

Visual crowdsensing is an important paradigm of crowdsensing. It leverages cameras of smart devices to attain the information of targets. A target can be a person, a building, a game, and so on. Guo *et al.* surveyed the challenges and opportunities in the visual crowdsensing[80]. In visual crowdsensing, a task is usually to take photos, which incurs new research issues such as data redundancy identification and elimination, high data processing cost, complex data structure, and high data transmission cost. Several papers[81-82] have been published to tackle the related problems in the visual crowdsensing.

5 Solutions Design

Based on the arrival time of participants, the selection strategies are classified into offline selection and online selection. For online selection, the participants arrive dynamically, and the decision on whether or not to choose the participants is made on the occasions of their arrivals. For offline selection, all the participants are ready, and the selection is made in advance based on historical or predictable information.

As depicted in Fig.11, there are multiple kinds of algorithms in both offline and online situations. In the offline situations, the algorithms can be divided into three types, i.e., greedy-based algorithms, machine learning (ML) based algorithms, and graph-based algorithms. In the online situations, we discuss the one-side online situations and multi-side online situations separately. One-side online situations mean that only tasks or only participants are coming dynamically, while the other component is known in advance. Multi-side online situations refer to that more than one of the components in the MCS system are coming to the MCS server dynamically. In the former situations, we introduce three kinds of algorithms, i.e., greedy-based algorithms, game theory-based algorithms, and graph-based algorithms. In the latter situations, we discuss two kinds of algorithms, i.e., greedy-based algorithms and graph-based algorithms.

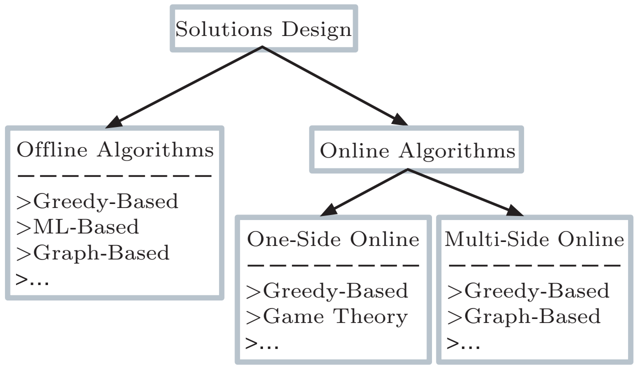

Fig.11. Solutions design in MCS.

5.1 Offline Matching Algorithm

If the information of the tasks and the participants is all known in advance, the matching problem is seen as offline matching. One important strategy used in many papers is the greedy selection strategy. In most cases, the matching problem is proven to be an NP-complete problem, and the greedy selection strategy is an efficient way to solve such a problem. For example, in [26], one of the objectives is to maximize the spatial coverage ratio. The relevant strategy is to select a user who can maximize the total coverage into the participant set in each step until the budget is exhausted. In [83], a reliable task assignment problem is studied. The authors formulated two optimization goals, maximum reliability assignment (MRA) under a recruitment budget, and minimum cost assignment (MCA) under a task reliability requirement. These two problems were proven to be NP-hard, and greedy algorithms were designed to obtain approximate solutions. The key feature of greedy selection strategy is to choose the best option in each step until the objective function is satisfied or the constraint is reached. Hence, it is important to quantify the objective function for one step. The greedy selection strategy is so widely used that most of the papers mentioned in our survey adopt this strategy or an improved version. To improve the matching efficiency, some other heuristic methods, such as genetic algorithm, simulation annealing algorithm, and ant colony algorithm, are also used in some papers.

Machine learning based algorithms are also popular in related work. Hsieh *et al.* proposed to estimate the air quality in unsensed areas by training a multilayer perceptron, because the sensing results of different sensing areas were supposed to be inter-related[84]. Based on the perceptron, a greedy selection algorithm combined with a sorting algorithm was used to select proper areas as the sensing tasks. Wang *et al.* utilized a Bayesian inference to estimate the air quality in unsensed areas[60]. The prior knowledge is gotten through cross-validation results of the sensing areas. Based on the estimated results, a participant selection algorithm was designed. The main idea is to deduce the data of unsensed areas through different methods, and to select the one with maximum difference as the sensing area. The Bayesian inference plays a key role.

Some matching mechanisms are based on graph theory. Ji *et al.* selected PoIs for each participant to optimize their trajectories, and achieved a maximum uniform coverage[52]. Proper locations are selected firstly for each participant, and possible trajectories are listed in a location graph. The goal of the matching algorithm is to maximize the utility (uniform coverage) of selected PoIs. A dynamic programming strategy was used to find an approximate optimal path. Liu *et al.* proposed to solve a matching problem based on the minimum cost and maximum flow (MCMF) algorithm in graph theory[61]. The participants are assumed to be few, and tasks are many (FPMT). The representation graph can be seen in Fig.12. Each participant must finish q tasks, and each task can be executed at most p times. The task set in the third level is produced by selecting q tasks randomly, and the weight between the second level (participants) and the third level is the executing cost. A greedy algorithm was used to select the minimum cost flow in the graph. Actually, by reducing to the maximum flow problem, any algorithm that computes the maximum flow in the network can be used to solve the matching problem. However, the computation of the third level is time-consuming.

5.2 Online Matching Algorithm

The online matching problem can be divided into one-side online matching and multi-side online matching. One-side online matching means that only the tasks arrive dynamically, while the participants are known in advance; or only the participants arrive dynamically, and the tasks are determined. In most related work, the tasks are supposed as known at first, and the participants dynamically come to the MCS server. Zhao *et al.* designed an online selection strategy combining with an incentive mechanism[85]. Li *et al.* designed three progressive strategies: offline selection strategy based on comprehensive information, online selection strategy with dynamic participants and static tasks, and online selection strategy with dynamic participants and dynamic tasks[71]. Hu *et al.* proposed a framework consisting of an inference model and an online task assigner[86]. The inference model was used to get reliable label results, which is implemented by the EM (Expectation Maximization) algorithm based on the reputation of participants, the distance between participants and PoIs, and the influence of PoIs. The online task assignment is based on greedy algorithm. First, the quality improvement of each task is computed if it is assigned to one or more participants. Then, the best tasks for each participant are selected to maximize

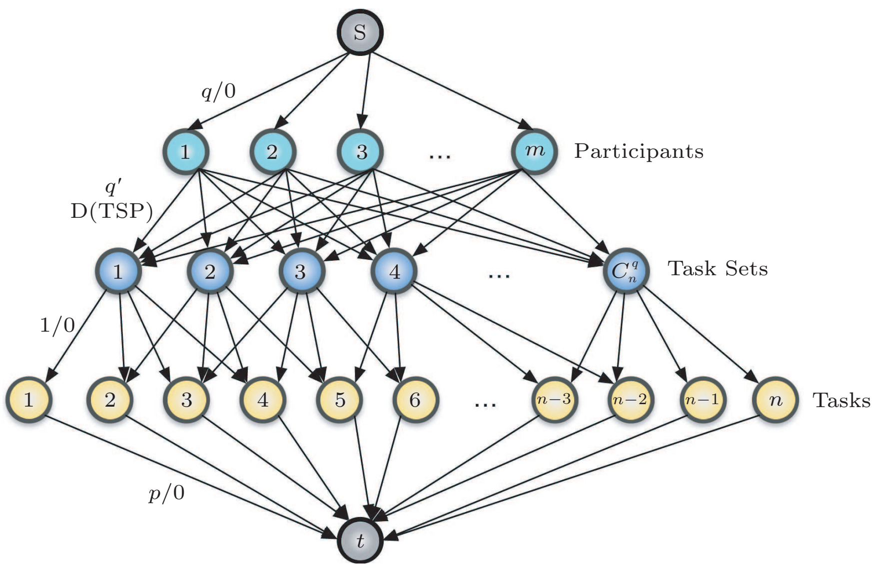

Fig.12. MCMF for the FPMT problem[61].

the total improvement. Zhang *et al.* formulated the online task assignment problem, and proposed three incentive mechanisms to solve this problem based on online reverse auction[87]. Furthermore, the authors designed two online incentive mechanisms motivated by a sampling-accepting process and weighted maximum matching to assign most valuable tasks to the selected participants in [88].

Other kinds of strategies used in online matching problems are applications of game theory. Han *et al.* formulated the dynamic participant selection problem as a multi-armed bandit (MAB)[89]. A feasible sensing engagement (FSE) is selected in each time slot, with cost and revenue as attributes of each FSE. The selection series of FSE make a robust sensing policy (RSP). Hence the problem can be abstracted as defining an RSP which can get the maximum revenue. The cost of each FSE is dynamic to make sure the robustness. The authors further proposed a task scheduling mechanism in mobile crowd sensing system in [90]. Similarly, She *et al.* formulated the task-participant arrangement problem into the MAB problem, and two matching algorithms based on traditional MAB solutions were proposed and compared[91]. The tasks, with capacity and conflicting event pairs, are regarded as arms. The feedback (accepting or rejecting the arranged tasks) from the participants is regarded as the reward. One of the proposed matching algorithms is based on Thompson sampling (TS), and the other one is based on upper confidence bound (UCB). The experimental results indicate the performance of UCB is better than that of TS. Cheung *et al.* formulated the task selection process in a pull situation based on a non-cooperative game[38]. In every negotiation round, each participant selects tasks which can maximize the utility. The utility of each participant equals the payoff of selected tasks minus the moving cost, which is the function of the moving distance. It is proved that the algorithm converges to Nash equilibrium within a polynomial time. To compare the efficiency of the proposed algorithm, a centralized greedy selection algorithm and a distributed greedy selection algorithm are introduced respectively. The experimental results show that the social surplus and task coverage of the proposed algorithm and the centralized greedy selection algorithm are similar. However, the payoff fairness between participants of the former one is much better than that of the latter one. Pu *et al.* proposed that the arrivals of the participants to a PoI follow a Poisson process[41]. They then formulated the matching problem as a dynamic programming problem, and designed a matching strategy based on Bellman equation.

Multi-side online matching means that more than one part of the MCS system is unknown beforehand, and a real-time decision on whether or not to select should be made. Particularly, two-side online matching indicates that the tasks and the participants both arrive dynamically. Tong *et al.* identified a new online micro-task allocation problem, and called it the global online micro-task allocation (GOMA) problem[75]. Two-side online situations are considered in this paper, including dynamic tasks and dynamic participants. The allocation problem is formulated into the maximum bipartite matching problem, as shown in Fig.13. A two-phase framework based on a Hungary algorithm, which is a traditional solution to bipartite matching problem, was proposed to balance the competitive ratio and the efficiency. However, the experimental results show that the performance of the greedy matching algorithm is similar to that of the proposed algorithm, while the former one is less complex than the latter one. Moreover, Song *et al.* proposed a new three-side online matching problem, where the participants, the tasks, and the workplaces should be selected in real time, to maximize the utility function[92]. Three matching algorithms were proposed, including a basic greedy algorithm, a threshold greedy algorithm, and an adaptive threshold greedy algorithm. The basic greedy algorithm aims to select fresh triple which can maximize the utility function. Considering the high time complexity of the basic greedy algorithm, the threshold greedy algorithm was proposed, and its main idea is to select triple upper the predefined threshold. Furthermore, the adaptive threshold greedy algorithm aims to adapt the threshold to a proper value through analyzing the matching results in each cycle.

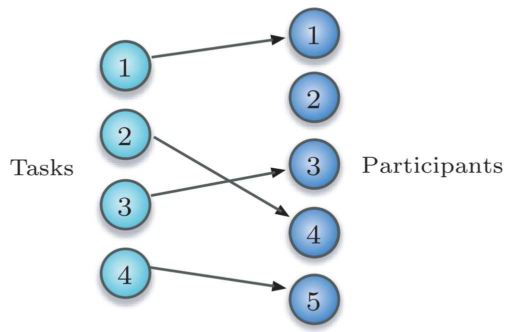

Fig.13. Bipartite matching model for two-side online situations[75].

5.3 Summary of Solutions

After surveying the proposed solutions by recent literatures, we can find that the greedy-based algorithms are used in both offline and online situations. The reason is that in most of the discussed cases, the global optimization method to solve the matching problem in MCS is time-consuming. The greedy-based algorithms can be proved as an efficient approximation of the global optimization. Another popular direction to solve the matching problem is the graph-based method. The reason is that the matching problem can be naturally formulated through the bipartite graph, and traditional algorithms for the bipartite matching can be used in the matching problem in MCS. The game theory based algorithms are popular in one-side online situations. One of the reasons is that the related game theories focus on the trade-off between exploiting and exploring, which is suitable to the online situations. The matching algorithms in multi-side online situations are less than those in the one-side online situations, and we believe that more game theory based algorithms for multi-side online situations will be discussed in the future. The machine learning based algorithms are popular in offline algorithms, because they are always time-consuming and proper features are needed.

6 Future Research Directions

Based on our review of the matching strategies, we propose several directions for future research. These directions include matching strategy for heterogeneous tasks, context-aware matching, online selection, and leveraging historical data to finish new tasks.

6.1 Matching Strategy for Heterogeneous Tasks

As described above, the task publishers upload the tasks to the MCS server, and the MCS server assigns the tasks to the participants. Hence, multiple task publishers may upload tasks simultaneously, and those tasks can be very different. It is a challenge for the MCS server to deal with all the different tasks, especially when there are connections between tasks. For example, if one task is about traffic accident sensing and collection, another task is about traffic jam detection and navigation. It is reasonable to merge these two kinds of tasks, and to select the same participants for both tasks. By introducing the heterogeneous tasks, task fusion and task decomposition become research issues. On one hand, a large task can be divided into multiple sub-tasks according to different steps, and parallel assignment can improve the task efficiency. On the other hand, related tasks uploaded by different publishers can be merged as one task, or accomplished by the same set of participants. Coordinating these two kinds of situations is important to optimize the productivity of the MCS server.

6.2 Context-Aware Matching

Context includes the environmental information surrounding the participants as well as the behavioral information of the participants. Examples of the environmental information include whether a participant is in a park, how many companions are around a participant, how much sensing battery is left. Examples of the behavioral information include whether a participant is on a driving, in a meeting, or walking. By analyzing the context, the MCS server can assign tasks more wisely. If a participant is on a driving, tasks such as temperature or PM 2.5 information collection cannot be assigned, since the sensing results are within the car and are not representative. Therefore, context-aware selection mechanisms are needed. However, it is challenging to collect the contextual information. There are few papers till now focusing on the context-aware sensing. Participant selection based on the context is rare. There are some reasons. First, the needed context for a task is hard to be abstracted and defined. Second, the context change in different time slices makes the context sensing more complicated. Third, context information is always associated with participant privacy, and people usually are not willing to share.

6.3 Online Strategy

Existing papers usually assume that the number and the distribution of the participants are known at first, and the selection strategies are based on the rigorous precondition. However, in a real MCS framework, the participants usually appear randomly. Hence, it is more useful to design online selection strategies than offline ones. Nevertheless, it is challenging to determine whether to select a participant or not for a task as soon as the participant signs up. Machine learning based methods may help to design online strategies, since machine learning can excavate patterns according to historical information.

6.4 Leveraging Historical Data to Finish New Tasks

Big data has been a hot research point in recent years. Large amount of historical data is stored in both the MCS server and terminal devices. Most existing researches on participant selection in the MCS framework focus on designing selection strategies according to instant information. However, they usually ignore making use of the abundant historical data. Actually, some applications are related to historical data tightly. For instance, how to recruit witnesses or records of an accident is a useful application. However, it is difficult to search useful information among tremendous amount of data. Therefore, it is meaningful to find a way to leverage proper historical data to finish new tasks.

7 Conclusions

Task and participant matching is an important problem in mobile crowd sensing. In this survey, we provided the reader with an extensive review of the state-of-the-art researches on this topic. The contents of mobile crowd sensing and the importance of the matching problem were introduced first. The researching framework was then presented, including the participant model, task models, and matching solutions. By analyzing the limitations and weaknesses of existing work, possible directions of novel future research in this field are discussed.

The extensive literature analysis showed that most matching solutions are designed based on the participant situations and the task requirements. In addition, it is found that more and more factors are considered in the participant model to select the proper participants efficiently. Concerning the task models, it is found that technologies and algorithms applicable at either area tasks or point tasks have evolved significantly, yet the matching problems with other task types need more researches. For the solutions, multiple kinds of mathematical tools can be used to design new matching solutions. The task and participant matching is a significant issue in mobile crowd sensing, which still needs great effort to be studied in the future.

References

[1] Guo B, Wang Z, Yu Z W, Wang Y, Yen N Y, Huang R H, Zhou X S. Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm. *ACM Computing Surveys*, 2015, 48(1): 1-31.

[2] Guo B, Chen C, Zhang D Q, Yu Z W, Chin A. Mobile crowd sensing and computing: When participatory sensing meets participatory social media. *IEEE Communications Magazine*, 2016, 54(2): 131-137.

[3] Cherian J, Luo J, Guo H L, Ho S S, Wisbrun R. Parkgauge: Gauging the occupancy of parking garages with crowdsensed parking characteristics. In *Proc. the 17th IEEE Int. Conf. Mobile Data Management* (*MDM*), July 2016, pp.92-101.

[4] Guo B, Chen H H, Yu Z W, Xie X, Huangfu S L, Zhang D Q. FlierMeet: A mobile crowdsensing system for cross-space public information reposting, tagging, and sharing. *IEEE Trans. Mobile Computing*, 2015, 14(10): 2020-2033.

[5] Koukoumidis E, Peh L S, Martonosi M R. SignalGuru: Leveraging mobile phones for collaborative traffic signal schedule advisory. In *Proc. the 9th ACM Annu. Conf. Mobile Systems, Applications, and Services* (*MobiSys*), June 2011, pp.127-140.

[6] Morishita S, Maenaka S, Nagata D, Tamai M, Yasumoto K, Fukukura T, Sato K. SakuraSensor: Quasi-realtime cherry-lined roads detection through participatory video sensing by cars. In *Proc. ACM Int. Joint Conf. Pervasive and Ubiquitous Computing* (*UbiComp*), September 2015, pp.695-705.

[7] Ludwig T, Reuter C, Pipek V. What you see is what I need: Mobile reporting practices in emergencies. In *Proc. the 13th European Conf. Computer Supported Cooperative Work* (*ECSCW*), September 2013, pp.181-206.

[8] Lane D D, Miluzzo E, Lu H, Peebles D, Choudhury T, Campbell A T. A survey of mobile phone sensing. *IEEE Communications Magazine*, 2010, 48(9): 140-150.

[9] Gao H, Liu C H, Wang W D, Zhao J X, Song S, Su X, Crowcroft J, Leung K. A survey of incentive mechanisms for participatory sensing. *IEEE Communications Surveys & Tutorials*, 2015, 17(2): 918-943.

[10] Jaimes L G, Vergara-Laurens I J, Raij A. A survey of incentive techniques for mobile crowd sensing. *IEEE Internet of Things Journal*, 2015, 2(5): 370-380.

[11] Zhang X L, Yang Z, Sun W, Liu Y H, Tang S H, Xing K, Mao X F. Incentives for mobile crowd sensing: A survey. *IEEE Communications Surveys & Tutorials*, 2016, 18(1): 54-67.

[12] Guo B, Chen H H, Yu Z W, Nan W Q, Xie X, Zhang D Q, Zhou X S. TaskMe: Toward a dynamic and quality-enhanced incentive mechanism for mobile crowd sensing. *International Journal of Human Computer Studies*, 2017, 102: 14-26.

[13] Rafael P T, César T H, Hiram G Z. Power management techniques in smartphone-based mobility sensing systems: A survey. *Pervasive and Mobile Computing*, 2016, 31: 1-21.

[14] Liaqat D, Jingoi S, Lara E D, Goel A, To W, Lee K, Garcia I D, Saldana M. Sidewinder: An energy efficient and developer friendly heterogeneous architecture for continuous mobile sensing. In *Proc. the 21st Annu. Conf. Architectural Support for Programming Languages and Operating Systems* (*ASPLOS*), April 2016, pp.205-215.

[15] Guo X N, Chan E, Liu C, Wu K S, Liu S Y, Ni L M. ShopProfiler: Profiling shops with crowdsourcing data. In *Proc. the 33th IEEE Annu. Conf. Computer Communications* (*INFOCOM*), April 2014, pp.1240-1248.

[16] Zhou P F, Zheng Y Q, Li M. How long to wait? Predicting bus arrival time with mobile phone based participatory sensing. *IEEE Trans. Mobile Computing*, 2014, 13(6): 1228-1241.

[17] Pournajaf L, Garcia-Ulloa D A, Xiong L, Sunderam V. Participant privacy in mobile crowd sensing task management: A survey of methods and challenges. *ACM SIGMOD Record*, 2016, 44(4): 23-34.

[18] Christin D. Privacy in mobile participatory sensing: Current trends and future challenges. *Journal of Systems and Software*, 2016, 116: 57-68.

[19] Yang K, Zhang K, Ren J, Shen X. Security and privacy in mobile crowdsourcing networks: Challenges and opportunities. *IEEE Communications Magazine*, 2015, 53(8): 75-81.

[20] Huang K L, Kanhere S S, Hu W. Are you contributing trustworthy data? The case for a reputation system in participatory sensing. In *Proc. the 13th ACM Annu. Conf. Modeling, Analysis, and Simulation of Wireless and Mobile Systems* (*MSWiM*), October 2010, pp.14-22.

[21] Truskinger A, Yang H F, Wimmer J, Zhang J L, Williamson I, Roe P. Large scale participatory acoustic sensor data analysis: Tools and reputation models to enhance effectiveness. In *Proc. the 7th IEEE Annu. Conf. eScience*, December 2011, pp.150-157.

[22] Christin D, Roßkopf C, Hollick M, Martucci L A, Kanhere S S. IncogniSense: An anonymity-preserving reputation framework for participatory sensing applications. *Pervasive and Mobile Computing*, 2013, 9(3): 353-371.

[23] Ren J, Zhang Y X, Zhang K, Shen X M. SACRM: Social aware crowdsourcing with reputation management in mobile sensing. *Computer Communications*, 2015, 65: 55-65.

[24] Mousa H, Mokhtar S B, Hasan O, Younes O, Hadhoud M, Brunie L. Trust management and reputation systems in mobile participatory sensing applications: A survey. *Computer Networks*, 2015, 90: 49-73.

[25] Zhao D, Ma H D, Tang S J, Li X Y. COUPON: A cooperative framework for building sensing maps in mobile opportunistic networks. *IEEE Trans. Parallel and Distributed Systems* (*TPDS*), 2015, 26(2): 392-402.

[26] He Z J, Cao J N, Liu X F. High quality participant recruitment in vehicle-based crowdsourcing using predictable mobility. In *Proc. the 34th IEEE Annu. Conf. Computer Communications* (*INFOCOM*), April 2015, pp.2542-2550.

[27] Cho E, Myers S A, Leskovec J. Friendship and mobility: User movement in location-based social networks. In *Proc. the 17th ACM Annu. Conf. Knowledge Discovery and Data Mining* (*SIGKDD*), August 2011, pp.1082-1090.

[28] Gao Y, Dong W, Guo K, Liu X, Chen Y, Liu X J, Bu J J, Chen C. Mosaic: A low-cost mobile sensing system for urban air quality monitoring. In *Proc. the 35th Annual IEEE Annu. Conf. Computer Communications* (*INFOCOM*), April 2016.

[29] Chen Y Y, Lv P, Guo D K, Zhou T Q, Xu M. Trajectory segment selection with limited budget in mobile crowd sensing. *Pervasive and Mobile Computing*, 2017, 40(9): 123-138.

[30] Boutsis I, Kalogeraki V. Mobile stream sampling under time constraints. In *Proc. the 14th IEEE Annu. Conf. Mobile Data Management* (*MDM*), June 2013, pp.227-236.

[31] Hamid S A, Takahara G, Hassanein H S. On the recruitment of smart vehicles for urban sensing. In *Proc. the 2013 IEEE Global Telecommunications Conference* (*GLOBECOM*), December 2013, pp.36-41.

[32] Zheng Y. Trajectory data mining: An overview. *ACM Trans. Intelligent Systems and Technology* (*TIST*), 2015, 6(3): 1-41.

[33] Pournajaf L, Xiong L, Sunderam V, Goryczka S. Spatial task assignment for crowd sensing with cloaked locations. In *Proc. the 15th IEEE Annu. Conf. Mobile Data Management* (*MDM*), July 2014, pp.73-82.

[34] Wang L Y, Zhang D Q, Yang D Q, Lim B Y, Ma X J. Differential location privacy for sparse mobile crowdsensing. In *Proc. the 16th IEEE Annu. Conf. Data Mining* (*ICDM*), December 2016, pp.1257-1262.

[35] Wang L Y, Zhang D Q, Wang Y S, Chen C, Han X, M'Hamed A. Sparse mobile crowdsensing: Challenges and opportunities. *IEEE Communications Magazine*, 2016, 54(7): 161-167.

[36] Xiao M J, Wu J, Zhang S, Yu J P. Secret-sharing-based secure user recruitment protocol for mobile crowdsensing. In *Proc. the 36th IEEE Annu. Conf. Computer Communications (INFOCOM)*, April 2017.

[37] Celis L E, Reddy S P, Singh I P, Vaya S. Assignment techniques for crowdsourcing sensitive tasks. In *Proc. the 19th ACM Conf. Computer-Supported Cooperative Work and Social Computing* (*CSCW*), February 2016, pp.836-847.

[38] Cheung M H, Southwell R, Hou F, Huang J W. Distributed time-sensitive task selection in mobile crowdsensing. In *Proc. the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing* (*MobiHoc*), June 2015, pp.157-166.

[39] Zhang L Y, Hu T, Min Y, Wu G B, Zhang J Y, Feng P C, Gong P H, Ye J P. A taxi order dispatch model based on combinatorial optimization. In *Proc. the 23rd ACM Annu. Conf. Knowledge Discovery and Data Mining* (*SIGKDD*), August 2017, pp.2151-2159.

[40] Cardone G, Foschini L, Bellavista P, Corradi A, Borcea C, Talasila M, Curtmola R. Fostering participaction in smart cities: A geo-social crowdsensing platform. *IEEE Communications Magazine*, 2013, 51(6): 112-119.

[41] Pu L J, Chen X, Xu J D, Fu X M. Crowdlet: Optimal worker recruitment for self-organized mobile crowdsourcing. In *Proc. the 35th Annual IEEE Annu. Conf. Computer Communications* (*INFOCOM*), April 2016.

[42] Mavridis P, Gross-Amblard D, Miklós Z. Using hierarchical skills for optimized task assignment in knowledge-intensive crowdsourcing. In *Proc. the 25th Annu. Conf. World Wide Web* (*WWW*), April 2016, pp.843-853.

[43] Zhang X M, Wu Y B, Huang L F, Ji H, Cao G H. Expertise-aware truth analysis and task allocation in mobile crowdsourcing. In *Proc. the 37th IEEE Annu. Conf. Distributed Computing Systems* (*ICDCS*), June 2017, pp.922-932.

[44] Amintoosi H, Kanhere S S. A trust-based recruitment framework for multi-hop social participatory sensing. In *Proc. the 2013 IEEE Annu. Conf. Distributed Computing in Sensor Systems* (*DCoSS*), May 2013, pp.266-273.

[45] Chang W, Wu J. Reliability enhanced social crowdsourcing. In *Proc. the 2015 IEEE Global Communications Conf.* (*GLOBECOM*), December 2015.

[46] Nath S. ACE: Exploiting correlation for energy-efficient and continuous context sensing. *IEEE Trans. Mobile Computing*, 2013, 12 (8): 1472-1486.

[47] Reddy S, Shilton K, Burke J, Estrin D, Hansen M, Srivastava M. Using context annotated mobility profiles to recruit data collectors in participatory sensing. In *Proc. Int. Sym. Location- and Context-Awareness* (*LoCA*), May 2009, pp.52-69.

[48] Zhang H L, Xu Z K, Du X J, Zhou Z G, Shi J T. CAPR: Context-aware participant recruitment mechanism in mobile crowdsourcing. *Wireless Communications and Mobile Computing*, 2016, 16(15): 2179-2193.

[49] Liu S Z, Zheng Z Z, Wu F, Tang S J, Chen G H. Context-aware data quality estimation in mobile crowdsensing. In *Proc. the 36th IEEE Annu. Conf. Computer Communications* (*INFOCOM*), April 2017, pp.802-810.

[50] Xia S H, Gao L, Lai Y K, Yuan M Z, Chai J X. A survey on human performance capture and animation. *Journal of Computer Science and Technology* (*JCST*), 2017, 32(3): 536-554.

[51] Tamilin A, Carreras I, Ssebaggala E, Opira A, Conci N. Context-aware mobile crowdsourcing. In *Proc. ACM Int. Joint Conf. Pervasive and Ubiquitous Computing* (*UbiComp*), September 2012, pp.717-720.

[52] Ji S G, Zheng Y, Li T R. Urban sensing based on human mobility. In *Proc. ACM Int. Joint Conf. Pervasive and Ubiquitous Computing* (*UbiComp*), September 2016, pp.1040-1051.

[53] Jaimes L G, Vergara-Laurens I, Labrador M A. A location-based incentive mechanism for participatory sensing systems with budget constraints. In *Proc. IEEE Annu. Conf. Pervasive Computing and Communications* (*PerCom*), March 2012, pp.103-108.

[54] Mendez D, Labrador M A. Density maps: Determining where to sample in participatory sensing systems. In *Proc. the 3rd FTRA Annu. Conf. Mobile, Ubiquitous, and Intelligent Computing* (*MUSIC*), June 2012, pp.35-40.

[55] Song Z, Liu C H, Wu J, Ma J, Wang W D. QoI-aware multitask-oriented dynamic participant selection with budget constraints. *IEEE Trans. Vehicular Technology*, 2014, 63(9): 4618-4632.

[56] Liu L, Wei W Y, Zhao D, Ma H D. Urban resolution: New metric for measuring the quality of urban sensing. *IEEE Trans. Mobile Computing*, 2015, 14(12): 2560-2575.

[57] Kang X, Liu L, Ma H D. Enhance the quality of crowdsensing for fine-grained urban environment monitoring via data correlation. *Sensors*, 2017, 17(1): Article No. 88.

[58] Zhang Y, Roughan M, Willinger W, Qiu L L. Spatio-temporal compressive sensing and Internet traffic matrices. *Network*, 2009, 20(3): 267-278.

[59] Wu Y B, Wang Y, Cao G H. Photo crowdsourcing for area coverage in resource constrained environments. In *Proc. the 36th IEEE Annu. Conf. Computer Communications* (*INFOCOM*), April 2017.

[60] Wang L Y, Zhang D Q, Pathak A, Chen C, Xiong H Y, Yang D Q, Wang Y S. CCS-TA: Quality-guaranteed online task allocation in compressive crowdsensing. In *Proc. ACM International Joint Conf. Pervasive and Ubiquitous Computing* (*UbiComp*), September 2015, pp.683-694.

[61] Liu Y, Guo B, Wang Y, Wu W L, Yu Z W, Zhang D Q. TaskMe: Multi-task allocation in mobile crowd sensing. In *Proc. ACM Int. Joint Conf. Pervasive and Ubiquitous Computing* (*UbiComp*), September 2016, pp.403-414.

[62] Guo B, Liu Y, Wu W L, Yu Z W, Han Q. ActiveCrowd: A framework for optimized multitask allocation in mobile crowdsensing systems. *IEEE Trans. Human-Machine Systems*, 2017, 47(3): 392-403.

[63] Kang Y R, Miao X, Liu K B, Chen L, Liu Y H. Quality-aware online task assignment in mobile crowdsourcing. In *Proc. the 12th IEEE Annu. Conf. Mobile Ad Hoc and Sensor Systems* (*MASS*), October 2015, pp.127-135.

[64] Kazemi L, Shahabi C. GeoCrowd: Enabling query answering with spatial crowdsourcing. In *Proc. the 20th Annu. Conf. Advances in Geographic Information Systems*, November 2012, pp.189-198.

[65] Lane N D, Chon Y H, Zhou L, Zhang Y Z, Li F, Kim D W, Ding G Z, Zhao F, Cha H J. Piggyback crowdsensing (PCS): Energy efficient crowdsourcing of mobile sensor data by exploiting smartphone app opportunities. In *Proc. the 11th ACM Conf. Embedded Networked Sensor Systems* (*SenSys*), November 2013, pp.7:1-7:14.

[66] Zhang D Q, Xiong H Y, Wang L Y, Chen G L. CrowdRecruiter: Selecting participants for piggyback crowdsensing under probabilistic coverage constraint. In *Proc. ACM Int. Joint Conf. Pervasive and Ubiquitous Computing* (*UbiComp*), September 2014, pp.703-714.

[67] Xiong H Y, Zhang D Q, Chen G L, Wang L Y, Gauthier V. CrowdTasker: Maximizing coverage quality in piggyback crowdsensing under budget constraint. In *Proc. IEEE Annu. Conf. Pervasive Computing and Communications* (*PerCom*), March 2015, pp.55-62.

[68] Xiong H Y, Zhang D Q, Chen G L, Wang L Y, Gauthier V, Barnes L E. ICrowd: Near-optimal task allocation for piggyback crowdsensing. *IEEE Trans. Mobile Computing*, 2016, 15(8): 2010-2022.

[69] Xiong H Y, Zhang D Q, Wang L Y, Chaouchi H. EMC3: Energy-efficient data transfer in mobile crowdsensing under full coverage constraint. *IEEE Trans. Mobile Computing*, 2015, 14(7): 1355-1368.

[70] Wang J T, Wang Y S, Zhang D Q, Wang F, He Y D, Ma L T. PSAllocator: Multi-task allocation for participatory sensing with sensing capability constraints. In *Proc. ACM Conf. Computer Supported Cooperative Work and Social Computing* (*CSCW*), February 2017, pp.1139-1151.

[71] Li H S, Li T, Wang Y. Dynamic participant recruitment of mobile crowd sensing for heterogeneous sensing tasks. In *Proc. the 12th IEEE Annu. Conf. Mobile Ad Hoc and Sensor Systems* (*MASS*), October 2015, pp.136-144.

[72] Li H S, Li T, Li F, Wang W C, Wang Y. Enhancing participant selection through caching in mobile crowd sensing. In *Proc. the 24th IEEE/ACM International Symposium on Quality of Service* (*IWQoS*), June 2016.

[73] Chen C, Cheng S F, Misra A, Lau H C. Multi-agent task assignment for mobile crowdsourcing under trajectory uncertainties. In *Proc. Annu. Conf. Autonomous Agents and Multiagent Systems*, May 2015, pp.1715-1716.

[74] Xiao M J, Wu J, Huang H, Huang L S, Hu C. Deadline-sensitive user recruitment for probabilistically collaborative mobile crowdsensing. In *Proc. the Annu. Conf. Distributed Computing Systems* (*ICDCS*), August 2016, pp.721-722.

[75] Tong Y X, She J Y, Ding B L, Wang L B, Chen L. Online mobile micro-task allocation in spatial crowdsourcing. In *Proc. the 32nd IEEE Annu. Conf. Data Engineering* (*ICDE*), May 2016, pp.49-60.

[76] Lee D H, Wang H, Cheu R, Teo S. A taxi dispatch system based on current demands and real-time traffic conditions. *Journal of the Transportation Research Board*, 2004, 1882(1): 193-200.

[77] Yu Z Y, Zhang D Q, Yu Z W, Yang D Q. Participant selection for offline event marketing leveraging location-based social networks. *IEEE Trans. Systems, Man, and Cybernetics*: *Systems*, 2015, 45(6): 853-864.

[78] Cheng P, Lian X, Chen L, Shahabi C. Prediction-based task assignment in spatial crowdsourcing. In *Proc. the 33rd IEEE Annu. Conf. Data Engineering* (*ICDE*), April 2017, pp.997-1008.

[79] Zhang X L, Yang Z, Gong Y J, Liu Y H, Tang S H. SpatialRecruiter: Maximizing sensing coverage in selecting workers for spatial crowdsourcing. *IEEE Trans. Vehicular Technology*, 2017, 66(6): 5229-5240.

[80] Guo B, Han Q, Chen H H, Shangguan L F, Zhou Z M, Yu Z W. The emergence of visual crowdsensing: Challenges and opportunities. *IEEE Communications Surveys* & *Tutorials*, 2017, 19(4): 2526-2543.

[81] Guo B, Chen H H, Han Q, Yu Z W, Zhang D Q, Wang Y. Worker-contributed data utility measurement for visual crowdsensing systems. *IEEE Trans. Mobile Computing*, 2017, 16(8): 2379-2391.

[82] Chen H H, Guo B, Yu Z W, Chen L M, Ma X J. A generic framework for constraint-driven data selection in mobile crowd photographing. *IEEE Internet of Things Journal*, 2017, 4(1): 284-296.

[83] Zhang X L, Yang Z, Liu Y H, Tang S H. On reliable task assignment for spatial crowdsourcing. *IEEE Trans. Emerging Topics in Computing* (*Early Access*), 2017. DOI: 10.1109/TETC.2016.2614383.

[84] Hsieh H P, Lin S D, Zheng Y. Inferring air quality for station location recommendation based on urban big data. In *Proc. the 21st ACM Annu. Conf. Knowledge Discovery and Data Mining* (*SIGKDD*), August 2015, pp.437-446.

[85] Zhao D, Li X Y, Ma H D. Budget-feasible online incentive mechanisms for crowdsourcing tasks truthfully. *IEEE/ACM Trans. Networking*, 2016, 24(2): 647-661.

[86] Hu H Q, Zheng Y D, Bao Z F, Li G L, Feng J H, Cheng R. Crowdsourced POI labelling: Location-aware result inference and task assignment. In *Proc. the 32nd IEEE Annu. Conf. Data Engineering* (*ICDE*), May 2016, pp.61-72.

[87] Zhang X L, Yang Z, Zhou Z M, Cai H B, Chen L, Li X Y. Free market of crowdsourcing: Incentive mechanism design for mobile sensing. *IEEE Trans. Parallel and Distributed Systems* (*TPDS*), 2014, 25(12): 3190-3200.

[88] Zhang X L, Yang Z, Liu Y H, Li J Q, Ming Z. Toward efficient mechanisms for mobile crowdsensing. *IEEE Trans. Vehicular Technology*, 2017, 66(2): 1760-1771.

[89] Han K, Zhang C, Luo J. BLISS: Budget limited robust crowdsensing through online learning. In *Proc. the 11th Annual IEEE Annu. Conf. Sensing, Communication, and Networking* (*SECON*), June 2014, pp.555-563.

[90] Han K, Zhang C, Luo J, Hu M L, Veeravalli B. Truthful scheduling mechanisms for powering mobile crowdsensing. *IEEE Trans. Computers*, 2016, 65(1): 294-307.

[91] She J Y, Tong Y X, Chen L, Song T S. Feedback-aware social event-participant arrangement. In *Proc. ACM Annu. Conf. Management of Data* (*SIGMOD*), May 2017, pp.851-865.

[92] Song T S, Tong Y X, Wang L B, She J Y, Yao B, Chen L, Xu K. Trichromatic online matching in real-time spatial crowdsourcing. In *Proc. the 33rd IEEE Annu. Conf. Data Engineering* (*ICDE*), April 2017, pp.1009-1020.

Yue-Yue Chen received her B.S. and M.S. degrees in computer science and technology from National University of Defense Technology (NUDT), Changsha, in 2013 and 2015, respectively. She is currently a Ph.D. candidate in College of Computer, NUDT, Changsha. Her main research interests include mobile crowd sensing, task assignment, etc.

Pin Lv received his B.S. degree in software engineering from Northeastern University, Shenyang, in 2006. He received his Ph.D. degree in computer science and technology from National University of Defense Technology, Changsha, in 2012. He is currently with the School of Computer Electronics and Information, Guangxi University, Nanning, and also with Guangxi Key Laboratory of Multimedia Communications and Network Technology, Nanning. His research interests include wireless networks, mobile computing, Internet of Things, etc. He is a member of CCF, ACM, and IEEE.

De-Ke Guo received his B.S. degree in industry engineering from Beijing University of Aeronautics and Astronautics, Beijing, in 2001, and his Ph.D. degree in management science and engineering from the National University of Defense Technology (NUDT), Changsha, in 2008. He is currently a professor with the College of System Engineering, NUDT, Changsha, and a professor with the School of Computer Science and Technology, Tianjin University, Tianjin. His research interests include distributed systems, software-defined networking, data center networking, wireless and mobile systems, and interconnection networks. He is a distinguished member of CCF, a senior member of IEEE, and a member of ACM.

Tong-Qing Zhou received his B.S. and M.S. degrees in computer science and technology from National University of Defense Technology (NUDT), Changsha, in 2012 and 2014, respectively. He is currently a Ph.D. candidate in College of Computer, NUDT, Changsha. His main research interests include wireless networks, mobile sensing, and data security.

Ming Xu is a professor in Department of Network Engineering, College of Computer, National University of Defense Technology, Changsha. He is an editor of the Journal of Communications and the International Journal of Pervasive Computing, and a technical program member of several international conferences such as IEEE PerCom. His current research interests include ad-hoc networks, vehicular networks, and wireless mesh networks. He is a member of CCF, ACM, and IEEE.