

# 控制与决策

Control and Decision

## 移动群智感知中基于任务质量的多任务分发参与者选择

蒋伟进, 陈君鹏, 张婉清, 陈萍萍, 孙永霞, 刘晓亮

引用本文:

蒋伟进, 陈君鹏, 张婉清, 陈萍萍, 孙永霞, 刘晓亮. 移动群智感知中基于任务质量的多任务分发参与者选择[J]. 控制与决策, 2022, 37(10): 2667–2676.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2021.0049>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 群智感知中移动用户招募的防贪婪激励机制研究

Anti-greedy incentive mechanism for mobile user recruitment in crowd sensing  
控制与决策. 2022, 37(1): 28–36 <https://doi.org/10.13195/j.kzyjc.2020.0744>

#### 基于改进生物激励神经网络算法的多移动机器人协同变电站巡检作业

Multi-mobile robot cooperative inspection operation based on improved biological excitation neural network algorithm in substation  
控制与决策. 2022, 37(6): 1453–1459 <https://doi.org/10.13195/j.kzyjc.2020.1714>

#### 基于云平台的共享制造模式定价策略

Pricing strategies for sharing manufacturing model based on cloud platform  
控制与决策. 2022, 37(4): 1056–1066 <https://doi.org/10.13195/j.kzyjc.2020.1464>

#### 一种集成最后一公里的四方物流网络设计问题启发式算法

A heuristic algorithm for fourty logistics network design problem with last-mile delivery  
控制与决策. 2022, 37(6): 1601–1608 <https://doi.org/10.13195/j.kzyjc.2020.1516>

#### 基于李雅普诺夫随机优化的车辆边缘计算资源管理

Resource management of vehicle edge computing based on Lyapunov stochastic optimization  
控制与决策. 2022, 37(3): 721–728 <https://doi.org/10.13195/j.kzyjc.2020.1211>

# 移动群智感知中基于任务质量的多任务分发参与者选择

蒋伟进<sup>1,2†</sup>, 陈君鹏<sup>1</sup>, 张婉清<sup>2</sup>, 陈萍萍<sup>1</sup>, 孙永霞<sup>1</sup>, 刘晓亮<sup>1</sup>

(1. 湖南工商大学 计算机学院, 长沙 410205; 2. 新零售与虚拟现实湖南省重点实验室, 长沙 410205)

**摘要:** 任务分发作为移动群智感知领域的关键研究问题是目前的热点和难点, 针对该领域中多任务分发条件下的参与者选择方法进行研究, 根据参与者的历史任务完成情况, 利用参与者累计信誉以及参与意愿构建参与者服务质量模型 (QoS). 在最大化 QoS 的基础上, 采用启发式贪心算法解决参与者的选择问题. 提出以任务为中心和以用户为中心两种选择方案, 在算法中引入距离约束因子、完整度约束因子以及信誉度约束因子, 在保证任务完成质量的前提下, 尽可能提升平台最终收益和参与者的收益. 对两种算法的可行性和有效性进行评估, 与现有的算法在选择出的参与者人数、移动距离以及数据质量等方面进行详细对比, 分析各种因素对选择结果的影响. 在实验过程中, 建立阶跃型数据定价模型, 对参与者上传的数据质量进行量化对比. 实验结果表明, 所提出的算法较现有的算法在服务质量方面取得了很好的效果.

**关键词:** 移动群智感知; 参与者选择; 服务质量; 多任务分发; 启发式算法; 累积信誉

中图分类号: TP373

文献标志码: A

DOI: 10.13195/j.kzyjc.2021.0049

开放科学(资源服务)标识码(OSID):



**引用格式:** 蒋伟进, 陈君鹏, 张婉清, 等. 移动群智感知中基于任务质量的多任务分发参与者选择[J]. 控制与决策, 2022, 37(10): 2667-2676.

## Multitask-oriented participant selection based on task quality in mobile crowd sensing

JIANG Wei-jin<sup>1,2†</sup>, CHEN Jun-peng<sup>1</sup>, ZHANG Wan-qing<sup>2</sup>, CHEN Ping-ping<sup>1</sup>, SUN Yong-xia<sup>1</sup>, LIU Xiao-liang<sup>1</sup>

(1. School of Computer Science, Hunan University of Technology and Business, Changsha 410205, China; 2. Key Laboratory of Hunan Province for New Retail Virtual Reality Technology, Changsha 410205, China)

**Abstract:** As a key research issue, task distribution is a hot and difficult point of current research in the field of mobile crowd sensing (MCS). The method of multi-task distribution participant selection is studied in this paper. According to the historical task completion status of participants, the cumulative reputation and willingness of participants are used to build the quality of service model (QoS). On the basis of maximizing the QoS, the greedy heuristic algorithm is used to solve the participation problem, and the two design schemes including task-centric and user-centric are proposed. The distance constraint factor, integrity constraint factor and reputation constraint factor are introduced into the proposed algorithms. The purpose is to improve the quality of perception tasks as much as possible, so as to improve the final benefits of platforms. Two algorithms are evaluated in the feasibility and effectiveness, and compared with existing algorithms in terms of the number of participants selected, moving distance, and data quality. During the experiment, a step data pricing model is established to quantitatively compare the quality of data uploaded by participants. Experimental results show that the two algorithms proposed are better than the existing algorithms in task quality.

**Keywords:** mobile crowd sensing; participant selection; service quality; multitask-oriented; heuristic algorithm; accumulated reputation

收稿日期: 2021-01-11; 录用日期: 2021-08-18.

基金项目: 国家自然科学基金项目(61772196); 湖南省自然科学基金项目(2020JJ4249); 湖南省教育厅科学研究项目(21A0374); 湖南省研究生科研创新项目(CX20211151, CX20211108).

†通讯作者. E-mail: jwj3666@163.com.

## 0 引言

作为一种新的数据采集模式,移动群智感知<sup>[1]</sup>利用现有的传感设备和通信基础设施,节省了额外搭建传感设备的开支<sup>[2]</sup>.同时,移动群智感知技术具有高度的移动性和广泛的潜在参与者等优势,特别针对突发性、不可预测性事件提供了前所未有的时空覆盖条件<sup>[3]</sup>.如今,群智感知技术在公共安全<sup>[4]</sup>、环境监测<sup>[5]</sup>、智慧交通<sup>[6]</sup>等方面都有了广泛的应用,这给日常生活带来了诸多方便的同时也提高了生活质量<sup>[7]</sup>.

移动群智感知中的参与者选择问题是指在满足一定的约束条件下,从庞大的用户群体中有效地选择合适的参与者来执行各种感知任务<sup>[8-9]</sup>.一方面,数据平台希望利用较少的开支得到期望的数据;另一方面,参与者也希望在感知任务中尽可能多地获利.不同的参与者所能收集到的数据有所差异,其向平台索要的回报也不尽相同<sup>[10]</sup>.如何实现平台与用户之间的博弈均衡,是当前学术界研究的一个重要方向<sup>[11]</sup>.除了上述两个主要目标之外,每项任务的参与人数、每个参与者分配的任务个数、参与者所提交数据的质量以及参与者完成任务所需要移动的总距离等都是选择过程中需要考虑的因素<sup>[12]</sup>.参与者选择是当前研究的重点和难点,目前的相关研究主要停留在单目标优化,例如以平台为中心寻求最大化利润以及以用户为中心实现最大化用户收益<sup>[13-14]</sup>.单目标优化往往会产生各种问题,比如最大化数据平台利益时会损害参与者收益,从而降低用户的参与意愿,最终也会影响平台收益.在现实情况中,要保证数据平台以及参与者的收益,综合考虑各种影响条件是很有必要的.因此,本文综合考虑多种因素,针对多任务条件下参与者选择方法进行深入研究.

## 1 相关工作

现阶段,已经有很多学者对移动群智感知中的参与者选择问题进行了广泛的研究与探索.

一些学者的研究侧重于设计优秀的激励机制鼓励参与者进行感知任务. Jin等<sup>[15]</sup>提出了一种新型的移动群智感知系统框架,它融合了数据聚合、激励以及扰动机制等; Hu等<sup>[16]</sup>重点研究了基于位置的MCS系统,提出了一种基于需求的动态激励机制,该机制可以按需动态地改变每个任务的奖励,以平衡任务之间的受欢迎程度; Xu等<sup>[17]</sup>以最小化社会成本为目的,提出了MCT-M和MCT-S两种模型,这两种方案利用社交网络的现实关系对参与者进行分组,使每个协作任务都可以由一组兼容的用户来完成,实验结果表

明,其提出的模型可以在减少分组时间的情况下进一步降低社会成本.

更多的研究者则是注重对于优秀参与者遴选机制的研究. Guo等<sup>[18]</sup>主要研究面向多样性任务类型中的参与者选择问题,分别从微观和宏观视觉任务类型出发,设计了UtiPay视觉群体感知框架,极大地提高了群智感知数据质量; Estrada等<sup>[19]</sup>提出了一种顺序选择参与者的框架,综合考虑了参与者声誉和支付模型,但该方法并未收集参与者的实际移动速度,同时也没有考虑用户的移动轨迹; Pu等<sup>[20]</sup>提出了另外一种顺序选择方案,对任务性质、用户工作能力以及实时性进行建模,从而预测参与者所提供的服务质量,但并没有考虑参与者的声誉以及移动距离带来的开支; 杨朔等<sup>[21]</sup>提出了有预算机制的多臂老虎机用户选择模型,解决了在无先验知识的情况下有效地学习选择每个用户的成本问题,并且在很大程度上降低了选择过程中的累计遗憾,但并未对用户环境信息的时间相关性进行深入研究; Wang等<sup>[22]</sup>将用户分为两组,分别制定不同的定价计划,提出了一个基于半马尔可夫模型来确定用户的兴趣点分布,并且提出了一种基于预测的参与者优选方法,最小化数据的上传成本; Jiang等<sup>[23]</sup>提出了基于投票机制的最优参与者决策模型,设计了一种基于逆向拍卖模型的参与者决策模型,该方法增加了有效感知任务数的参数,减少了参与者的冗余,但没有充分考虑参与者的地理位置信息; 刘琰等<sup>[24]</sup>针对多任务参与者选择分别以任务为中心以及以用户为中心提出了T-Random、T-Most以及PT-Most三种参与者选择算法,对每个任务的参与人数、每个用户分配的任务数、移动距离等多方面进行实验对比,选择出在不同指标下最合适的参与者选择策略. 但该方法仅以用户数和任务数以及移动距离为主要优化指标,并未考虑参与者的服务质量.

综上所述,目前大多数的研究方法是将选择问题表示为满足若干约束条件下的目标优化问题,在其过程中考虑参与者信誉、地理位置、移动距离以及激励机制等诸多因素,进而将多目标约束转化为单目标约束问题.但是,利用参与者历史信息来预估参与者任务质量的研究较少.一些方法在考虑优化参与者人数或者移动距离时忽略了平台对任务质量的需求;还有一些方法虽然考虑了任务质量需求,但是却忽略了参与者的实际移动速度或者参与意愿.本文兼顾参与者历史任务情况,建立参与者服务质量模型,在选择算法中,综合考虑移动距离、信誉以及任务完成程度等诸多因素,以寻求在服务质量最优的情况下最

好的选择方案. 最后, 利用真实数据集和模拟数据集对提出的两种算法进行仿真实验, 就距离弱化情况对参与者任务集合以及运行时间进行分析和研究.

## 2 基于QoS的多任务参与者选择模型

任务平台一次性会下发多个任务构成一个任务集合  $T = \{t_1, t_2, \dots, t_n\}$ . 同时, 数据平台会对任务完成时间有一定的限制, 迟到的数据往往是没有价值的, 在这里约定每个任务集合  $T$  的完成时间不超过  $h$  小时. 每项任务需要多个参与者完成, 同时每项任务  $t_i$  需要的完成人数  $s_i$  也不一定相同. 为了便于分析, 假定每项任务的完成时间为 5 min. 任务平台中有  $U = \{u_1, u_2, \dots, u_m\}$  名参与者, 每位参与者的移动速度是不相同的,  $V = \{v_1, v_2, \dots, v_m\}$ . 每个参与者所需完成的任务集合用  $TU_i = \{t_{i1}, t_{i2}, t_{i3}, \dots\}$  表示. 参与者完成任务所耗费的时间以及行程开销的时间之和要低于平台的要求时间. 其约束条件定义如下:

$$TU_i \times 5 + \frac{\text{Dist}(TU_i)}{v_i} \leq h \times 60, \quad 1 \leq i \leq m, \quad (1)$$

其中  $\text{Dist}(TU_i)$  为参与者完成该任务集合需要移动的总距离.

本文综合参与意愿以及参与者以往完成任务的信誉, 由此构建参与者服务质量模型.

某参与者的参与意愿越高, 该参与者便会越积极地收集数据, 其上传的数据质量也会越高; 若该参与者采取消极的态度收集数据, 则平台收到的数据质量就会很低. 然而, 仅仅是考虑参与者意愿对数据质量产生的影响是不够的, 参与者的客观感知能力对数据质量的影响也至关重要. 对参与意愿和数据质量分别建立评价指标, 为了体现两者具有同等的重要性, 将参与意愿  $w_i$  以及数据质量  $D_i$  都约束在  $[0, 1]$  的范围之中, 即

$$w_i \begin{cases} \min(w_i, 1), \\ \max(0, w_i); \end{cases} \quad (2)$$

$$D_i \begin{cases} \min(D_i, 1), \\ \max(0, D_i). \end{cases} \quad (3)$$

通常情况下, 参与者从接收到数据平台下发的任务到确认接受该任务所消耗的时间越长, 该参与者对于该项感知任务的参与意愿就会越低. 将该时间称为犹豫时间, 把参与意愿定义为一个与犹豫时间相关的函数, 受社会原则<sup>[25]</sup>的启发, 参与者的犹豫时间越长, 其贡献给平台的价值便越少. 将所有参与者的平均犹豫时间作为临界值, 如果某个参与者的犹豫时间与之相等, 则其参与意愿就是中性的, 其意愿值为 0.5.

基于以上讨论, 针对参与意愿和犹豫时间建模为以下函数:

$$w_i = 1 - \max \left[ 0, \min \left( \frac{1}{2} \log_{\bar{t}} t_i, 1 \right) \right]. \quad (4)$$

其中:  $w_i$  为参与者的意愿,  $t_i$  为犹豫时间,  $\bar{t}$  为平均犹豫时长. 由式(4)可知, 参与者犹豫越久, 其参与意愿就会越低.

参与者的失信情况可以从参与感知任务的历史数据中获取, 没有失信记录则表明参与者参与责任感较高, 所取得的数据质量也会较高.

为简便起见, 本文选用改进的信誉模型对参与者信誉进行刻画, 其计算公式如下:

$$R = \frac{T + 1}{T + F + 2} \times \xi, \quad 0 \leq R \leq 1, T > 0, F > 0. \quad (5)$$

其中:  $R$  值越接近于 1, 参与者的信誉值越高;  $T$  和  $F$  为参与者历史感知任务中“正确”和“错误”的记录;  $\xi$  为参与者恶意事件的权重因子, 计算公式为

$$\begin{cases} \xi = \lambda^k, & 0 \leq k < K; \\ \xi = 0, & k \geq K. \end{cases} \quad (6)$$

这里:  $k$  为某些感知参与者的恶意事件;  $K$  为数据平台设定的恶意事件阈值, 一旦超过该阈值, 参与者的信誉分数将被设定为 0. 考虑到平台对恶意事件的容忍度是很低的, 上传 3 次错误信息的参与者可以认为是恶意参与者, 在实验过程中默认  $K = 3$ .  $\lambda$  是阈值范围内的衰减因子, 设置  $\lambda$  的默认值为 0.8. 参与者的信誉可以无限接近但不等于 1, 由式(5)和(6)可知, 第一次进行感知任务的参与者的信誉值为 0.5, 随着进行感知任务数量的增加, 没有错误记录与有错误记录的参与者的信誉值的差距是很大的. 同时, 每次提交正确数据的参与者的信誉应该也是不同的, 拥有 10 次正确记录的参与者比仅有 1 次正确记录的参与者的可信程度肯定要高. 在选择参与者时, 对信誉不好的参与者往往会进行剔除. 建立参与者信誉约束条件  $\delta$ , 当  $R \in (0, \delta)$  时, 便不考虑选择此参与者进行感知任务.

综合上述的信誉模型以及参与意愿模型, 建立参与者服务质量(QoS), 其计算公式如下:

$$\text{QoS} = w \times R. \quad (7)$$

参与者携带的移动设备的配置或者品牌有所差异, 感知设备的好坏也会对数据质量产生一定的影响. 对参与者的设备的感知能力以及专业条件进行量化研究显然需要考虑多种正、负相关因素, 而参与者提交的数据与平台预期数据之间的差异程度则可以代表参与者的感知能力.

对参与者历史提交数据进行评估,利用历史提交数据完整情况刻画参与者数据完整性  $D_i$ ,其计算公式如下:

$$D_i = \begin{cases} 0.8, & N = 0; \\ \frac{1}{N} \sum_{t=1}^N P_t, & N \geq 1. \end{cases} \quad (8)$$

其中:  $P_t \in (0, 1)$  为历史时刻  $t$  时参与者上传的数据完整程度,对  $P_t$  的确定需要任务平台有一套完善的打分机制;  $N$  为参与者完成的历史任务数目;  $D$  的值越大,说明参与者上传的数据质量越高. 对于初次进行感知任务的参与者而言,平台总是愿意给他更多的机会完成任务,因而本文预计该参与者提交任务的完整性能达到 0.8. 同样的,建立完整性约束指标  $\varepsilon$ , 当  $D \in (0, \varepsilon)$  时,认为该参与者的任务完成质量难以达到当前任务的完成需求,则不考虑选择此参与者进行感知任务.

数据平台对数据质量的容忍程度往往会比对失信参与者的容忍程度要高,有些感知任务仅仅是粗略感知就能满足平台的需求,即每个任务对参与质量的要求可能是不同的,任务平台在发布任务集合  $T = \{t_1, t_2, \dots, t_n\}$  时往往会给出其对应的数据最低完整程度集合  $C = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ . 只要参与者提交的任务完整程度高于此标准,便能满足平台的需求.

本文提出的参与者选择方法的优化目标便是在保证服务质量的条件下,尽可能减少参与者的移动距离,以实现数据平台利益的最大化.

### 3 参与者选择方法

本文在 T-Most 算法基础上进行改进,以任务为中心进行参与者选择,将最大化参与者服务质量作为主要优化目标,同时兼顾最小化参与人数,要求每个参与者能在规定时间、尽可能小的移动距离条件下完成更多的任务的同时,满足数据平台对数据质量的要求.

参与者选择具体过程可做如下考虑:选择当前待完成任务中需要人数最多的任务作为初始任务,选择在该任务点一定距离范围内服务质量最高的参与者完成该任务. 接下来选择距离初始任务点最近的任务作为该参与者的第 2 个需要完成的任务. 当第 2 个任务完成之后,以此任务为初始任务节点,选择最近的任务作为第 3 个需要完成的任务. 如此反复,按照上述流程,直到选择出该参与者在  $h$  小时内完成的任务集合. 剔除上述过程中的参与者并对任务集合中其参与过的每个任务所需人数减 1,按照上述方法,

继续选择出参与者以及对应的任务集合,直到所有任务都完成. 其优化模型如下:

$$\begin{aligned} & \max(\text{QoS}); \\ & \text{s.t.} \begin{cases} \min \sum_{i=1}^m \text{Dist}(\text{TU}_i), \\ \min \text{Count}(P), \\ \text{TU}_i \times 5 + \frac{\text{Dist}(\text{TU}_i)}{v_i} \leq h \times 60. \end{cases} \quad (9) \end{aligned}$$

其中:  $\text{Dist}(\text{TU}_i)$  为第  $i$  个参与者待完成任务的总距离.

#### 算法 1

输入: 任务集合  $T$ 、用户集合  $U$ 、完整性约束集  $C$ 、距离约束  $\theta$ 、信誉约束  $\delta$ ;

输出: 参与者集合  $P$  及其完成的任务集合  $\text{TU}$ .

- 1) 计算所有候选参与者的信誉度  $R$  以及数据完整指标  $D$ ;
- 2) 删除信誉度不达标的参与者;
- 3) 选择待完成任务集中需要人数最多的任务作为初始任务  $t_{ij}(j = 1)$ ;
- 4) 计算距离初始任务  $t_{ij}$  约束距离  $\theta$  范围内符合  $D \geq \varepsilon_j$  参与者的 QoS;
- 5) 选择  $\theta$  范围内 QoS 最高的参与者  $p_i(i \geq 1)$ ;
- 6) 选择离任务  $t_{ij}$  最近的且满足  $\varepsilon_{j+1} \leq D$  的任务  $t_{i(j+1)}(j \geq 1)$  作为新的中心任务节点;
- 7) 循环执行 4)~6), 当参与者  $p_i$  完成这些任务的时间超出时间约束时, 停止循环;
- 8) 输出参与者  $p_i$  的任务集合  $\text{TU}_i = \{t_{i1}, t_{i2}, t_{i3}, \dots\}$ ;
- 9) 循环执行 3)~8), 直至所有任务都完成;
- 10) 输出参与者集合  $P = \{p_1, p_2, p_3, \dots\}$  及其完成的任务集合  $\text{TU} = \text{TU}_1 \cup \text{TU}_2 \cup \dots \cup \text{TU}_i$ ;
- 11) 结束.

本文提出的算法 1 是以任务为中心进行参与者选择,其主要目标是在最大化参与者的服务质量的条件下,尽量减少参与者的移动距离,进而提高数据平台的利益. 本文在 PT-Most 算法的基础上提出一种以参与者为中心的选择方案. 下面将提出另一种算法,与算法 1 不同,其主要优化目标是最大化参与者完成的任务数量,在满足数据平台数据质量要求的前提下,尽量提高参与者的收益.

参与者选择的具体过程如下:随机选择用户集合中的参与者作为候选参与者,选择在该候选者  $\theta$  范围内满足完整性约束的最近任务作为初始任务. 接

下来选择距离初始任务最近且候选者有资格完成的任务作为下一个任务. 如此反复, 直至选出每个参与者在约定时间内的任务集合. 选择具有最多任务的候选者作为第1个参与者. 剔除上述过程中的参与者并对任务集合中其参与过的每个任务所需人数减1, 按照上述方法, 继续选择出参与者以及对应的任务集合, 直到所有任务都完成, 其优化模型如下:

$$\begin{aligned} & \max \text{Count}(\text{TU}_j), j = 1, 2, \dots; \\ & \text{s.t.} \begin{cases} \max \text{QoS}, \\ \text{TU}_j \times 5 + \frac{\text{Dist}(\text{TU}_j)}{v_j} \leq h \times 60. \end{cases} \end{aligned} \quad (10)$$

### 算法2

输入: 任务集合  $T$ 、用户集合  $U$ 、完整性约束集  $C$ 、距离约束  $\theta$ 、信誉约束  $\delta$ ;

输出: 参与者集合  $P$  及其完成的任务集合  $\text{TU}$ .

1) 计算所有候选参与者的信誉度  $R$  以及数据完整性指标  $D$ ;

2) 删除信誉度不达标的参与者;

3) 随机选择参与者集合中的一个用户作为任务候选者  $p_{ji}$ ;

4) 选择用户  $p_{ji}$  距离约束条件内最近的满足  $\varepsilon_k \leq D_i$  的任务  $t_{ik} (k=1)$  作为初始任务;

5) 选择离任务  $t_{ik}$  距离约束下最近的满足  $\varepsilon_{k+1} \leq D_i$  的任务  $t_{i(k+1)}$  作为下一任务;

6) 循环执行5), 当参与者  $p_{ji}$  完成这些任务的时间超出时间约束时停止循环;

7) 输出候选者  $p_{ji}$  的任务集合  $\text{TU}_{ji} = \{t_{i1}, t_{i2}, t_{i3}, \dots\}$ ;

8) 循环执行3)~7), 确定每个候选者  $p_{ji}$  在时间约束内的任务合集  $\text{TU}_{ji}$ ;

9) 选择拥有最大任务合集的参与者  $p_j$  完成任务集合  $\text{TU}_j$ ;

10) 循环执行3)~9), 直至确定每个满足约束条件的参与者的任务集合;

11) 输出参与者集合  $P = \{p_1, p_2, p_3, \dots\}$  及其完成的任务集合  $\text{TU} = \text{TU}_1 \cup \text{TU}_2 \cup \dots \cup \text{TU}_i$ ;

12) 结束.

## 4 实验评估

### 4.1 数据集与实验设置

#### 4.1.1 数据集

本文选择一组真实的数据集和一组模拟数据集对多任务条件下的参与者优选方案进行评估.

真实数据集: 中国工业和应用数学学会的众包

任务分配数据集, 该数据集包含广州、深圳、佛山以及东莞4个城市的835个任务以及1877个任务参与者. 在该数据集的基础上, 为每个任务随机指定一个任务完整指标  $\varepsilon$  进行约束, 同时为每位参与者随机指定一定的速度值来代表其真实移动速度, 其值为  $[0.83, 16.6] \text{ m/s}$ . 同时, 该数据集还提供了参与者的信誉分数  $R$ , 将其映射到  $[0, 1]$  之中.

模拟数据集: 本文遵循已有的数据生成方法生成模拟数据集. 把感知任务和参与者的位置均匀分配在一个  $20 \text{ km} \times 20 \text{ km}$  的矩形平面之中. 此外, 为每个参与者设定一个  $[0, 1]$  范围内的信誉度参数  $R$ . 在实验进行的过程中, 分别考虑不同规模的任务集合和参与者集合, 从  $[100, 800]$  和  $[50, 200]$  中选择  $m$  个感知任务与  $n$  个参与者进行组合. 对模拟数据集指定与真实数据集一样的其他参数.

#### 4.1.2 评价指标设计

$\theta$  的范围越大, 选择出的参与者的 QoS 无疑会越高, 但在追求更高的 QoS 的同时, 参与者完成任务所需的时间开销也会越高, 这就增加了平台的开支压力. 将参与者的收益定义为任务收益和移动收益, 针对参与者的移动时间建立定价机制, 参与者到达某一任务所消耗的时间越长, 参与者的移动收益就会越高, 即平台的可变开支就会越大. 为简便起见, 建立线性时间开支函数表示平台的可变开支, 其函数可以设置为

$$P(u) = kt. \quad (11)$$

其中:  $P(u)$  为平台的可变开支;  $t$  为参与者到达下一任务点所需要的时间;  $k$  为单位时间的移动开销, 为简便起见, 设置  $k$  的值为1.

同时, 针对参与者收集的数据进行定价, 参与者服务质量越高, 数据平台的收益就会越大, 建立阶跃型定价模型, 其定价模型如下:

$$G(u) \begin{cases} 10\text{QoS}, & 0 \leq \text{QoS} < 0.2; \\ 3 + 10(\text{QoS} - 0.2), & 0.2 \leq \text{QoS} < 0.5; \\ 7 + 10(\text{QoS} - 0.5), & 0.5 \leq \text{QoS} < 0.8; \\ 10 + 10(\text{QoS} - 0.8), & 0.8 \leq \text{QoS} < 1. \end{cases} \quad (12)$$

其中  $G(u)$  为随 QoS 的提升平台最终能够获得的收益.

设置目标函数  $\gamma$  表示随距离变化平台的最终收益,  $\gamma$  函数表示如下:

$$\gamma = G(u) - P(u). \quad (13)$$

### 4.1.3 实验设置

由于本文选用的真实数据集已经给定任务位置以及参与者当前位置,生成的模拟数据集在参与者和任务位置上也是随机分布的,本文不再对位置信息进行其他规约.但是在现实生活中,参与者与任务之间往往不是一条直线连接的,考虑两者之间的建筑以及交通路径更符合真实情况.选择合适的路径显然需要一项专门的研究来实现,由于本文研究的内容是参与者优选方案,为实验简便,选择两点之间的欧氏距离代表需要移动的距离.

在后续的实验之中,为方便算法性能对比,在没有特殊说明的情况下,默认每个任务需要5个用户完成.同时,设置每个任务需要的完成时间为5 min,每个任务集合需要的完成时间为2 h.

## 4.2 实验结果

### 4.2.1 合适的距离约束条件 $\theta$ 的选择

本实验分别针对真实数据集和模拟数据集进行研究.针对真实数据集,通过实验对比,确定合适的 $\theta$ ;针对模拟数据集,在改变感知任务数量以及可供选择的参与者人数的基础上,进行多次实验验证,寻找不同任务密度与参与者密度情况下的 $\theta$ 变化情况对平台收益的影响,继而选择出合适的 $\theta$ 约束条件.

图1是在真实数据集上进行实验的结果.可以明显看出,随着约束条件的放宽,平台的最终收益会先增加后减少,且当约束距离大于900 m以后,平台的收益会明显降低,这与实际情况是符合的.过短的距离约束难以选择出服务质量较好的参与者,这会使得平台的收益达不到最理想的水平;而 $\theta$ 的值选择过大,虽然能选择出QoS较好的参与者,但是用户需要移动的距离也会增加,这无疑会使平台的额外开支变大.

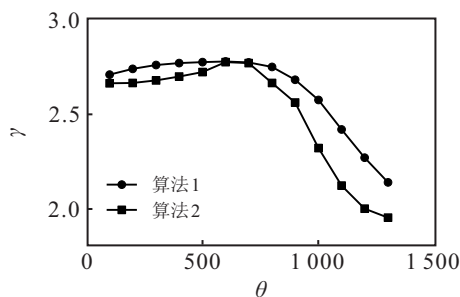
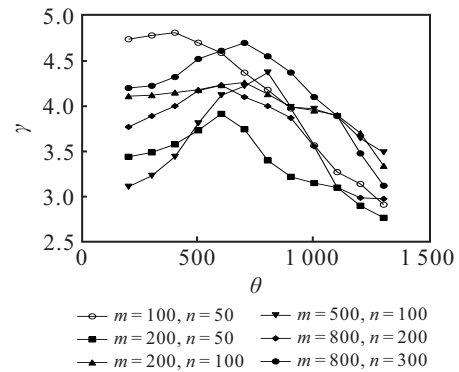


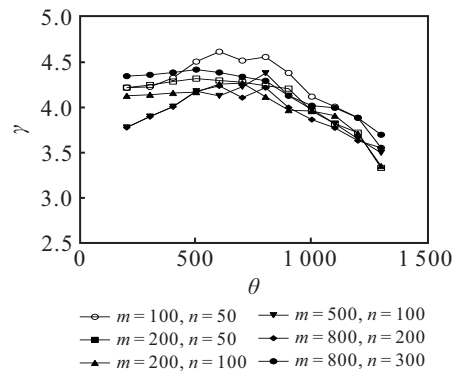
图1  $\theta$ 值的变化对真实数据集 $\gamma$ 的影响

图2是在模拟数据集上通过对任务数以及参与者人数进行动态组合的实验验证结果.可以看出,随着任务分发数目和参与者人数的变化, $\gamma$ 取得最大值的距离也有很大的变化,过高或者过低的 $\theta$ 都会使 $\gamma$ 值较低.因此针对不同的数据选择一个最为合适

的 $\theta$ 值是不容易的,但从图2可以看出,针对任务数量以及参与者人数的变化, $\theta \in [500, 800]$ 的条件下大多数数据都会取得较好的收益效果.在后续的实验,若无特殊说明,则默认选取 $\theta = 600$ 作为距离约束值.



(a)  $\theta$ 变化对算法1的影响



(b)  $\theta$ 变化对算法2的影响

图2  $\theta$ 值的变化对模拟数据集 $\gamma$ 的影响

### 4.2.2 任务个数变化对算法的影响

平台下发的任务数量并不是恒定的,本实验考虑在其他因子确定的情况下,改变每次分发任务的个数来进行参与者的选择.图3是算法1与T-Most算法的对照实验结果.随着任务个数的增加,算法1与T-Most算法选择出的参与者人数的差距并不是太大,甚至在某些任务个数的条件下,算法1会选择出更少的用户执行感知任务.对于用户平均分配到的任务数量,算法1与T-Most算法也没有太大的差距.在移动距离方面,由于算法1弱化了距离约束条件,T-Most算法中参与者的移动距离要比算法1中的少一些.T-Most算法的优化目标是 minimized distance while reducing cost, while algorithm 1 is less strict on distance requirements and focuses more on service quality. Therefore, on the platform's final revenue evaluation index  $\gamma$ , algorithm 1 achieves a significantly better result than T-Most algorithm, which is consistent with the assumptions made when proposing algorithm 1.

图4是算法2与PT-Most算法进行的对照实验结果.与图3中所能得到的结果一样,图4表明,在其他各项指标的差距并不明显的情况下,本文提出的算法

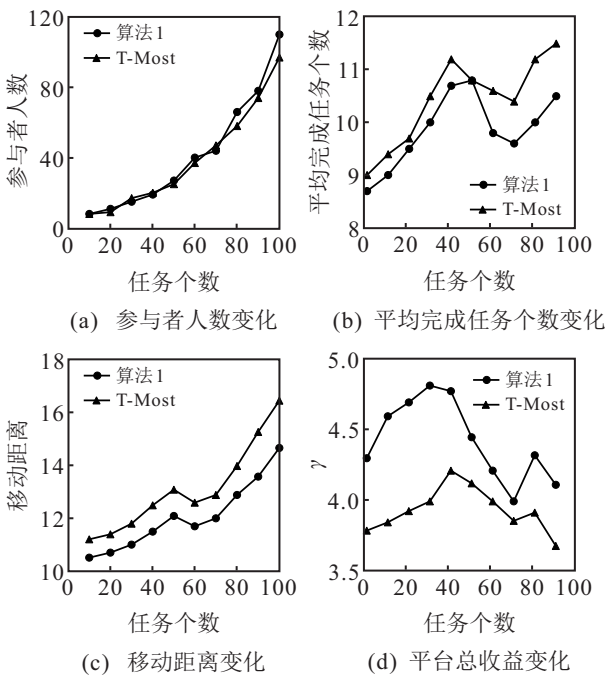


图3 任务个数变化对算法性能的影响(与T-Most对照)

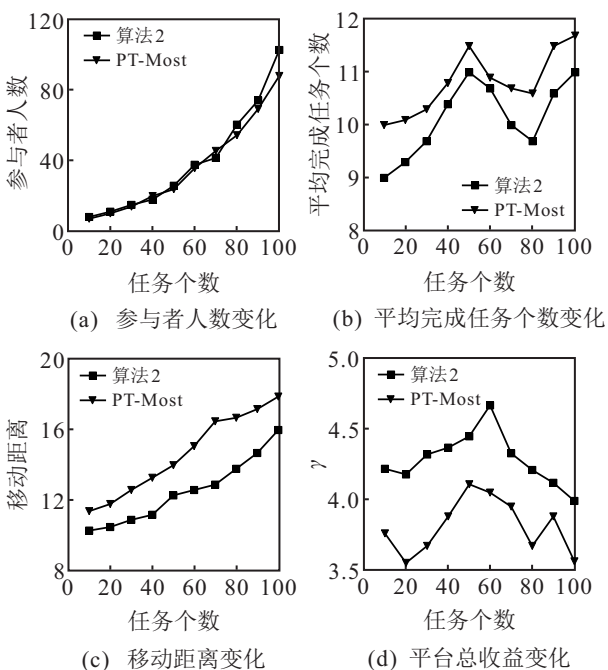


图4 任务个数变化对算法性能的影响(与PT-Most对照)

2在评价指标 $\gamma$ 上取得的效果要明显优于PT-Most算法。

### 4.2.3 候选人数量变化对算法的影响

在参与者选择的过程中,除了待完成的任务个数对选择有很大的影响之外,待选择的参与者人数也会在很大程度上对选择效果产生影响.候选人越多,选择出的参与者离任务的距离便会越近,参与者服务质量也会进一步提高.本实验考虑在待完成任务恒定的情况下,改变任务区域内的候选参与者人数进行实验.评价指标与上节实验一致.实验结果如图5和图6

所示.

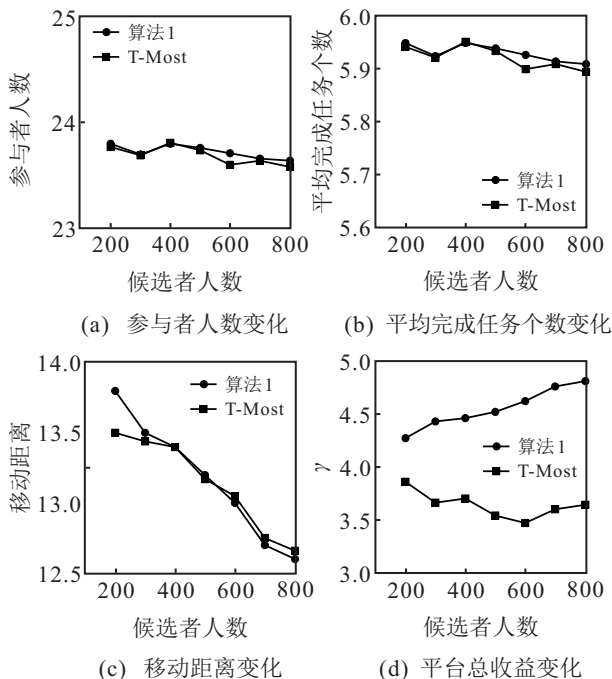


图5 候选人数量变化对算法性能的影响(与T-Most对照)

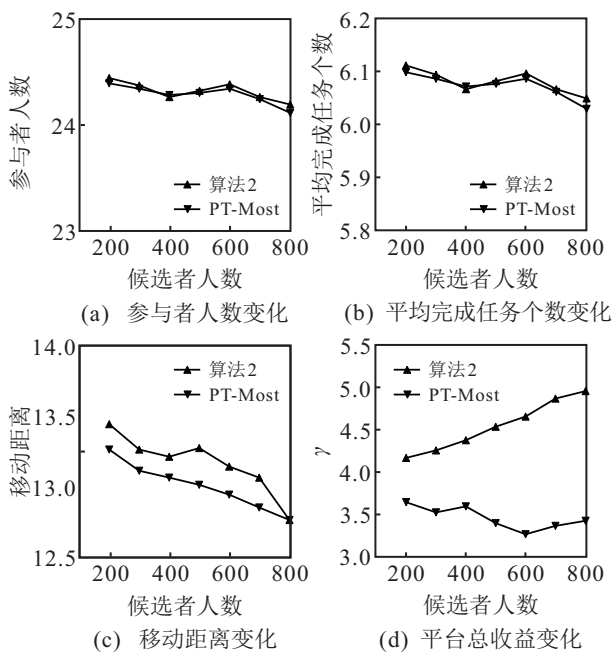


图6 候选人数量变化对算法性能的影响(与PT-Most对照)

由图5和图6可知,随着区域内候选人人数的增多,4种算法选择出的参与者人数以及每个参与者平均完成的任务数并没有太大的变化,且在数量指标上没有太大的差距.随着可选择的人数增加,选择出来的参与者会离任务越来越近,移动距离相应也会降低.很明显可以看出,在任务完成质量方面,本文所提的两种算法分别优于T-Most和PT-Most算法.而且随着候选人数的增加,本文的两种算法在 $\gamma$ 指标上的表现效果会越来越好,T-Most和PT-Most算法则没有明

显的变化.

### 4.2.4 感知时间变化

任务的规定完成时间也是影响算法性能的一个重大因素,考虑在任务数量不变的情况下,改变每一组任务的限定时间为1h、2h和3h进行实验对照,对上述实验中的5个性能指标进行分析,结果如图7所示.

由图7可知,随着限制时间的放宽,执行任务的

人数会不断降低,每个用户需要完成的任务数量相应增多,移动距离稍微增加.本文所提的两种算法与对照算法相比,在上述3种指标上的差距并不明显;从图7可以看出,大多数情况下,本文的算法与对照算法效果近似;而在 $\gamma$ 指标上,本文提出的算法要明显好于两种对照算法,且随着时间限制的放宽,在 $\gamma$ 指标上有着逐渐上升的趋势.

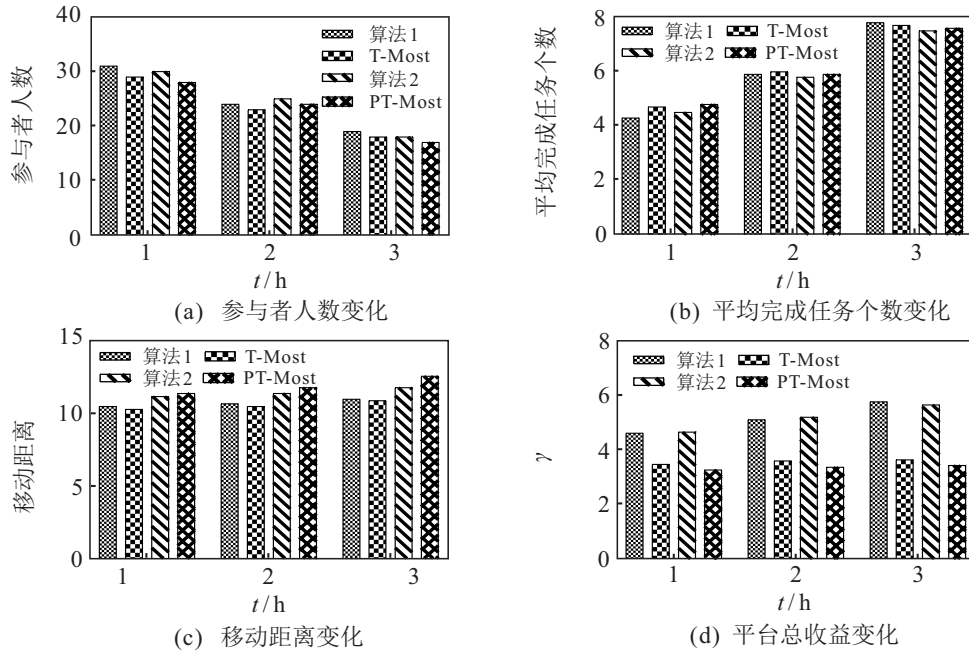


图7 感知时间变化对算法的影响

### 4.2.5 每个任务需要的人数变化

上述的实验都是限定在每个任务都需要5个参与者来完成的前提条件下进行的.但在现实生活中,因为每个任务的难易程度以及数据请求者对数据质

量的要求不同,对每个任务下达的预算有所差异,每个任务所需要的参与者人数是不尽相同的,所以本实验在任务需要相同和不同参与者的条件下,研究上述实验中的5个对比指标的变化情况.

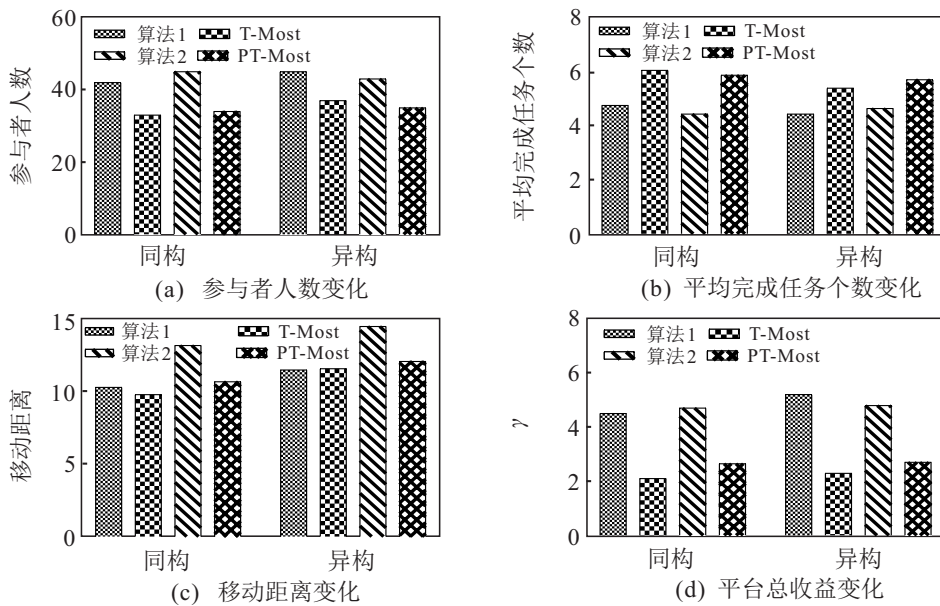


图8 同构异构条件对算法的影响

为方便实验, 设定待完成的任务个数为20, 同构条件下每个任务需要10个参与者来完成, 异构条件下分别设置任务需要人数为5、10、15, 且占总任务个数的25%、50%、25%, 即有5个任务需要5个参与者来完成, 10个任务需要10个参与者完成, 5个任务需要15个参与者来完成. 生成两个 $3\text{ km} \times 3\text{ km}$ 的矩形区域来模拟现实环境, 并让参与者和任务均匀分布在矩形平面上. 两个场景中除了任务需要的参与者人数存在差异, 其他参数设置大体一致.

由图8可以看出, 任务需要的参与者有差异的情况下与任务需要的参与者完全一致的情况下有较大的差异. 算法1、算法2与T-Most、PT-Most在前3项对比指标上有一定的弱势, 这与本文预期的情况是一致的. 但纵向对比算法1和算法2可以发现, 在参与者人数和平均完成任务数上面, 任务同构条件下, 算法1取得了较好的效果, 而在任务异构条件下, 算法2的效果较好. 在移动距离方面, 算法2在两种条件下相比算法1都移动了较长的距离, 因为算法2的主要目标是每个参与者的完成任务最大化, 因而这个结果也是符合预期的. 而在 $\gamma$ 因子上面, 本文提出的两种算法无疑还是优于现有的算法.

## 5 结论

本文针对移动群智感知中多任务情况下的参与者选择方法进行了研究, 基于参与者历史任务完成情况, 利用参与者累计信誉和参与者参与意愿建立服务质量模型, 针对以任务为中心和以参与者为中心的情况分别建立了基于服务质量的参与者优选方案. 在满足服务质量约束以及距离约束等条件下, 达到平台开支最小化以及用户收益最大化的要求. 建立了阶跃型数据质量定价函数, 对参与者上传数据进行量化评估. 在实验评估方面, 与现有的算法从多个方面进行对比, 实验结果表明, 本文提出的两种算法在参与者选择评价指标如参与者人数、平均完成任务数量以及移动总距离等方面与现有的算法没有明显的差距, 而在服务质量方面, 本文提出的算法1和算法2明显优于对比算法. 本文提出的基于服务质量的参与者选择方案能有效选择出较好的参与者进行感知任务, 从而大幅度提升平台收益.

在后续的研究中, 计划对用户的异质性以及历史移动规律进行研究, 以进一步提高服务质量.

## 参考文献(References)

- [1] 刘云浩. 群智感知计算[J]. 中国计算机学会通讯, 2012, 8(10): 38-41.  
(Liu Y H. Crowd sensing computing[J]. Communications of the CCF, 2012, 8(10): 38-41.)
- [2] 于瑞云, 王鹏飞, 白志宏, 等. 参与式感知: 以人为中心的智能感知与计算[J]. 计算机研究与发展, 2017, 54(3): 457-473.  
(Yu R Y, Wang P F, Bai Z H, et al. Participatory sensing: People-centric smart sensing and computing[J]. Journal of Computer Research and Development, 2017, 54(3): 457-473.)
- [3] 蒋伟进, 刘晓亮. 群智感知中移动用户招募的防贪婪激励机制研究[J]. 控制与决策, 2022, 37(1): 28-36.  
(Jiang W J, Liu X L. Anti-greedy incentive mechanism for mobile user recruitment in crowd sensing[J]. Control and Decision, 2022, 37(1): 28-36.)
- [4] Zhang D Q, Xiong H Y, Wang L Y, et al. CrowdRecruiter: Selecting participants for piggyback crowdsensing under probabilistic coverage constraint[C]. Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. Seattle, 2014: 703-714.
- [5] Ganti R K, Ye F, Lei H. Mobile crowdsensing: Current state and future challenges[J]. IEEE Communications Magazine, 2011, 49(11): 32-39.
- [6] Huang J Q, Kong L H, Dai H N, et al. Blockchain-based mobile crowd sensing in industrial systems[J]. IEEE Transactions on Industrial Informatics, 2020, 16(10): 6553-6563.
- [7] Restuccia F, Ghosh N, Bhattacharjee S, et al. Quality of information in mobile crowdsensing[J]. ACM Transactions on Sensor Networks, 2017, 13(4): 1-43.
- [8] Zhu X Y, Luo Y Y, Liu A F, et al. A deep learning-based mobile crowdsensing scheme by predicting vehicle mobility[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(7): 4648-4659.
- [9] 安健, 彭振龙, 桂小林, 等. 群智感知中基于公交系统的任务分发机制研究[J]. 计算机学报, 2019, 42(2): 65-78.  
(An J, Peng Z L, Gui X L, et al. Research on task distribution mechanism based on public transit system in crowd sensing[J]. Chinese Journal of Computers, 2019, 42(2): 65-78.)
- [10] 吴焱, 曾菊儒, 彭辉, 等. 群智感知激励机制研究综述[J]. 软件学报, 2016, 27(8): 2025-2047.  
(Wu Y, Zeng J R, Peng H, et al. Survey on incentive mechanisms for crowd sensing[J]. Journal of Software, 2016, 27(8): 2025-2047.)
- [11] Jin H M, Su L, Nahrstedt K. CENTURION: Incentivizing multi-requester mobile crowd sensing[C]. IEEE INFOCOM 2017-IEEE Conference on Computer Communications. Atlanta, 2017: 1-9.
- [12] Tian B Q, Yuan Y B, Zhou H Y, et al. Pavement

- management utilizing mobile crowd sensing[J]. *Advances in Civil Engineering*, 2020, 2020: 1-16.
- [13] Lee J S, Hoh B. Sell your experiences: A market mechanism based incentive for participatory sensing[C]. 2010 IEEE International Conference on Pervasive Computing and Communications (PerCom). Mannheim, 2010: 60-68.
- [14] Xiao M J, Wu J, Huang L S, et al. Multi-task assignment for crowdsensing in mobile social networks[C]. 2015 IEEE Conference on Computer Communications. Hong Kong, 2015: 2227-2235.
- [15] Jin H M, Su L, Xiao H P, et al. Incentive mechanism for privacy-aware data aggregation in mobile crowd sensing systems[J]. *IEEE/ACM Transactions on Networking*, 2018, 26(5): 2019-2032.
- [16] Hu J H, Wang Z B, Wei J, et al. Towards demand-driven dynamic incentive for mobile crowdsensing systems[J]. *IEEE Transactions on Wireless Communications*, 2020, 19(7): 4907-4918.
- [17] Xu J, Rao Z Q, Xu L J, et al. Incentive mechanism for multiple cooperative tasks with compatible users in mobile crowd sensing via online communities[J]. *IEEE Transactions on Mobile Computing*, 2020, 19(7): 1618-1633.
- [18] Guo B, Chen H H, Han Q, et al. Worker-contributed data utility measurement for visual crowdsensing systems[J]. *IEEE Transactions on Mobile Computing*, 2017, 16(8): 2379-2391.
- [19] Estrada R, Mizouni R, Otrok H, et al. A crowd-sensing framework for allocation of time-constrained and location-based tasks[J]. *IEEE Transactions on Services Computing*, 2020, 13(5): 769-785.
- [20] Pu L J, Chen X, Xu J D, et al. Crowd foraging: A QoS-oriented self-organized mobile crowdsourcing framework over opportunistic networks[J]. *IEEE Journal on Selected Areas in Communications*, 2017, 35(4): 848-862.
- [21] 杨朔, 吴帆, 陈贵海. 移动群智感知网络中信息量最大化的用户选择方法研究[J]. *计算机学报*, 2020, 43(3): 409-422.
- (Yang S, Wu F, Chen G H. On designing most informative user selection methods for mobile crowdsensing[J]. *Chinese Journal of Computers*, 2020, 43(3): 409-422.)
- [22] Wang E, Yang Y J, Wu J, et al. An efficient prediction-based user recruitment for mobile crowdsensing[J]. *IEEE Transactions on Mobile Computing*, 2018, 17(1): 16-28.
- [23] Jiang N, Xu D, Zhou J, et al. Toward optimal participant decisions with voting-based incentive model for crowd sensing[J]. *Information Sciences*, 2020, 512: 1-17.
- [24] 刘琰, 郭斌, 吴文乐, 等. 移动群智感知多任务参与者优选方法研究[J]. *计算机学报*, 2017, 40(8): 1872-1887.
- (Liu Y, Guo B, Wu W L, et al. Multitask-oriented participant selection in mobile crowd sensing[J]. *Chinese Journal of Computers*, 2017, 40(8): 1872-1887.)
- [25] Hardisty D J, Appelt K C, Weber E U. Good or bad, we want it now: Fixed-cost present bias for Gains and Losses explains magnitude asymmetries in intertemporal choice[J]. *Journal of Behavioral Decision Making*, 2013, 26(4): 348-361.

#### 作者简介

蒋伟进(1964—), 男, 教授, 博士, 从事社会计算、群智协同、分布式网络等研究, E-mail: jwj3666@163.com;

陈君鹏(1997—), 男, 硕士生, 从事群智感知、算法优化的研究, E-mail: jjpchen@163.com;

张婉清(1997—), 女, 硕士生, 从事智能决策、社交网络安全的研究, E-mail: 727149424@qq.com;

陈萍萍(1998—), 女, 硕士生, 从事群智感知、分布式网络的研究, E-mail: cpp0628@163.com;

孙永霞(1997—), 女, 硕士生, 从事机器学习、边云协同的研究, E-mail: 1552865513@qq.com;

刘晓亮(1996—), 男, 硕士生, 从事深度学习、模式识别的研究, E-mail: v1zone@163.com.

(责任编辑: 闫妍)