

感知质量优化的移动群智感知任务在线分发算法

张伟^{1,2}, 李卓^{1,2}, 陈昕²

(1. 网络文化与数字传播北京市重点实验室, 北京 100101; 2. 北京信息科技大学计算机学院, 北京 100101)

摘要: 感知质量优化和用户招募是移动群智感知的两个重要问题, 随着数据量的大幅度增加, 感知内容出现冗余, 存在感知质量降低的风险。提出了一种感知质量优化的任务分发机制, 在保证覆盖率的情况下, 提高群体的感知质量。利用聚类算法评估任务真值, 量化用户数据质量; 基于汤普森抽样算法和贪婪算法设计并实现了一种用户招募策略, 在保证任务空间覆盖率的基础上优化感知质量。针对 TSUR (Thompson based user recruit) 算法的性能进行仿真分析, 并与已有的 BBTA (bandit-based task assignment) 算法和 BUR (basic user recruitment) 算法作比较。实验表明, 在同一区域进行任务感知, 与 BBTA 算法和 BUR 算法相比, 累计感知质量分别提高了 16% 和 20%, 空间覆盖率分别提高了 30% 和 22%。

关键词: 移动群智感知; 任务分发; 数据质量; 在线学习

中图分类号: TP393.01

文献标识码: A

doi: 10.11959/j.issn.2096-3750.2020.00185

Data quality optimized online task allocation method for mobile crowdsensing

ZHANG Wei^{1,2}, LI Zhuo^{1,2}, CHEN Xin²

1. Beijing Key Laboratory of Internet Culture and Digital Dissemination Research, Beijing 100101, China

2. Beijing Information Science and Technology University, Beijing 100101, China

Abstract: Optimization of the perceived quality and the recruitment of user are two important issues of mobile crowdsensing. As the amount of data increases rapidly, perceived data becomes redundant, and perceived quality is at risk of decreasing. A mechanism of task assignment based on the perceptive quality optimization was proposed to improve the perceived quality under the condition of full coverage. The clustering algorithm was used to evaluate the truth value of the task and quantify the quality of the user data. Based on Thompson sampling algorithm and greedy algorithm, a user recruitment strategy was designed and implemented to optimize the perceived quality on the basis of ensuring the spatial coverage of the task. The performance of Thompson based user recruit (TSUR) algorithm was simulated and analyzed that compared with the existing algorithms of BBTA and basic user recruitment (BUR). Experiments show that in the same area, compared with bandit-based task assignment (BBTA) algorithm and BUR algorithm, the quality of the cumulative sensing data was improved by 16% and 20%, and the spatial coverage was improved by 30% and 22%.

Key words: mobile crowdsensing, task assignment, data quality, online learning

收稿日期: 2019-10-21; 修回日期: 2020-02-18

通信作者: 李卓, lizhuo@bistu.edu.cn

基金项目: 国家自然科学基金资助项目 (No.61872044, No.61502040); 北京市青年拔尖人才项目、北京市属高校高水平教师队伍建设支持计划青年拔尖人才培养计划资助项目 (No.CIT&TCD201804055); 网络文化与数字传播北京市重点实验室开放课题资助项目 (No.ICDDXN001); 北京信息科技大学“勤信人才”培育计划资助项目

Foundation Items: The National Natural Science Foundation of China (No.61872044, No.61502040), The Beijing Municipal Program for Top Talent, Beijing Municipal Program for Top Talent Cultivation (No.CIT&TCD201804055), The Open Program of Beijing Key Laboratory of Internet Culture and Digital Dissemination Research (No.ICDDXN001), The Qinxin Talent Program of Beijing Information Science and Technology University

1 引言

数据驱动的移动群智感知模式利用群体的移动设备来完成复杂的感知任务,包括感知任务的发起、执行、提交以及反馈。大量的感知数据从边缘节点聚合到云服务中心,通过一系列数据处理形成有价值的数据库资源。同时,大数据、人工智能等新兴技术也将为其所用,使得移动群智感知得到更快的发展。在此模式下,感知任务的分发问题、参与者的激励问题以及数据的隐私保护问题等成为研究重点,随着感知数据量的倍增和任务规模的扩大,有效的质量控制显得尤为重要^[1]。移动群智感知模式的感知质量包含时空覆盖率和数据质量两个层面,前者关注是否能采集足够全面的数据,后者关注数据是否足够准确和可信^[2],用户的移动性和复杂性使感知质量更难以控制。因此,针对感知质量提出了一种考虑覆盖率和数据质量的动态任务分发机制,实时监测感知质量,保证感知模式的高实用性。

如何量化用户数据层面的感知质量,是解决任务分发问题的难点之一。本文利用聚类算法的性质来评估任务真值和用户数据质量,将簇心数据判定为真值,利用真值与感知数据的差异来量化用户数据质量。用户数据质量随着不同群体数据的变化而变化,不同群体表现出不同的数据质量水平。为了进一步提高感知质量,引入贪婪策略保障覆盖率层面的感知质量,实时控制区域的感知覆盖情况,减少了不必要的重复感知和数据冗余。在此基础上,采用汤普森抽样算法招募具有高数据质量的用户,汤普森抽样算法平衡了任务分发问题的两个阶段,即利用现有信息最大化用户数据质量阶段和通过探索用户未知信息获取未来收益阶段,也就是在最短的时间内充分探索用户感知数据质量规律,从而实现群体数据质量的最大化。

在仿真实验中,模拟了感知任务的分发、执行、反馈等过程。在预设区域内发起多次感知任务,从参与者数据质量和任务覆盖率两个方面进行实时的参与者选择,在对比实验中,与 BBTA^[3]在线招募算法和 BUR^[4]贪婪招募算法相比,本文的 TSUR 任务分发算法的参与者数据质量较两者分别提高了 16%和 20%,在用户资源充足的条件下,任务覆盖率可达 100%。

2 相关工作

移动群智感知是一种新的感知范式,它利用人们在移动设备中产生的感知数据,通过云中的数据

聚集和融合,凝聚人类的智慧,用于提供以人为中心的服务^[5]。已有的工作研究主要应用于智能交通^[6]、环境监测^[7]、城市管理^[8]和公共安全^[9]等领域,随着各类应用的开展,有效的质量控制对提升数据的可用性具有重要作用。

在移动群智感知任务分发问题上,学者们已经提出了众多分发方案。然而,随着任务规模的扩大,感知质量呈下降趋势,为了解决这个问题,文献^[10]提出了多任务分发方案,该方案利用最小限度的数据质量阈值达到最优的全局效用。在感知覆盖层面,文献^[11]提出了一种可以保证覆盖效果和节能的参与者选择策略并实现了能量高效传输方式。关于移动轨迹的招募策略^[12],提出了一种利用移动群智感知的数据属性进行用户选择的方法,该方法不仅考虑了用户的时空轨迹概率,还考虑了感知任务的属性,并利用计算任务完成概率进行用户招募。这种方法关注用户执行任务的时效性,但缺少对任务执行结果的评估。文献^[13]建立了用户信用体系来评估用户的历史任务完成情况,基于上下文的分发策略把感知环境、个人和社会信息作为招募用户的考虑因素^[14]。本文的任务分发算法不仅考虑用户长期的感知能力,还对用户群体执行感知任务的覆盖情况进行优化。对用户感知能力做概率统计分析,减少由用户能力考查产生的大量开销,根据用户的地理位置做出实时性的决策,从而提高分发性能。

感知任务的数据质量一般和任务真值相关。一些质量控制的群智感知系统分发同一感知任务给多个用户,通过投票等统计方法对调查结果进行汇总,从而确定群体意见一致的答案,并将其作为任务真值^[15]。文献^[16]利用混合高斯模型来描述感知数据,然后提出了一种期望最大化的算法来评估感知任务真值。文献^[17]利用距离函数定义用户数据质量,每轮任务下的数据总质量相同,然而在现实中,不同群体的数据质量一般能分出不同层次,因此,本文数据质量定义以群体的差异化为基础,量化个体的数据质量。随着不同群体的数据质量不同,用户数据质量也能计算出不同差异程度,这表明用户数据质量随着群体数据质量的变化而变化,不同群体间也可以分辨出不同层次的数据质量。截至目前,感知数据已经拓展到图片、音频、视频等形式^[16,18-19]。文献^[18]提出了视觉群智感知用户贡献度量框架,减少了由冗余数据产生的用户贡献。文献^[19]对视频数据质量进行研究和推进,提出了主观质量评估方法。

3 移动群智感知任务分发模型

移动群智感知任务分发过程包括任务分发模型和感知地图模型。任务分发模型记录群体数据质量和招募人员; 感知地图模型将目标区域划分成若干感知单元, 然后统计单元格上的覆盖情况和感知质量。针对不同目标区域用户群体感知质量不同的情况, 提出感知质量最大化的优化问题。

3.1 群智感知任务分发基本架构

移动群智感知任务分发基本架构如图1所示。移动感知云平台发布感知任务, 任务分发模型确定感知任务适用区域和感知群体, 对招募的用户分发任务。用户接收感知任务后进行感知作业, 将数据上传至系统, 系统根据感知地图模型评估任务质量, 更新参数传送至移动感知云平台。

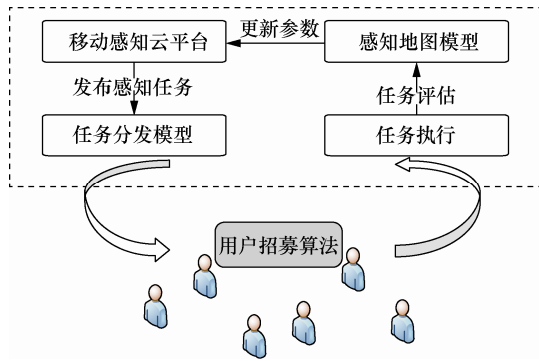


图1 移动群智感知任务分发基本架构

3.2 任务分发模型

在感知目标区域内, 将不同时间段的感知任务记为 $T = \{t_1, t_2, \dots, t_m\}$, 每个任务的 用户集为 $U = \{u_1, \dots, u_n\}$, 招募若干个用户构成集合 $U' \subseteq U$, U' 进行任务感知, 其提交的感知数据集为 $x = \{x_1, x_2, \dots, x_k\}$ 。其数据质量记为 $Q = \{q_1, \dots, q_k\}$, $q_i \in Q$ 且 $0 \leq q_i \leq 1$ 。存在数据质量阈值 μ , 当 $q_i \geq \mu$ 时, 用户 u_i 的数据质量达到标准。

用 $\text{dis}(x_i, x_j)$ 表示数据 x_i 与数据 x_j 的差异; D_i 表示数据 x_i 与其他用户数据的差距, $D_i = \sum_{1 \leq j \leq k, j \neq i} \text{dis}(x_i, x_j)$, 从簇心到其他用户的数据距离为 $D_r = \min_{1 \leq r \leq k} \{ \sum_{1 \leq j \leq k, j \neq r} \text{dis}(x_r, x_j) \}$, 簇心数据即为任务真值 x_r [17]; 在任务真值为 x_r 的条件下, 用户 u_i 的数据质量为 $q_i = \frac{1}{\text{dis}(x_r, x_i) + 1}$ 。

3.3 感知地图模型

本文利用感知地图模型评估区域的数据质量, 同时记录任务的空间覆盖情况。将感知地图 M 划分为若干个单元, $M = \{m_1, \dots, m_d\}$ 。用户 u_i 的感知范围为 $S_i = \{m_p, \dots, m_q\}, m_p \subseteq M, m_q \subseteq M$ 。 $C = \{c_1, \dots, c_d\}$ 记录感知地图子单元的覆盖情况, $c_i = 0$ 表示感知单元 m_i 没有用户进行覆盖; $c_i = 1$ 表示 m_i 已有用户完成感知任务。子单元上感知质量为 $P = \{p_1, \dots, p_d\}$, p_i 表示 m_i 上的感知质量, $p_i = \max \left\{ \frac{q_j}{|S_j|} \right\}$, $m_i \in S_j, j \in 1, \dots, k$, 其中, $|S_j|$ 表示 S_j 集合中所包含的感知单元个数。感知单元 m_i 可能被多个用户重复感知, 不重复累加用户感知质量, 整个感知地图的感知质量为 $\sum_{i=1}^d p_i$ 。

3.4 移动群智感知任务分发问题定义

本文在保证感知质量的条件下进行任务分发, 在给定的感知地图 $M = \{m_1, \dots, m_d\}$ 中, 招募一定数量用户, 最大化区域 M 中的数据质量。假设招募的最多用户数 $|U'|$ 与区域的大小有关, 设置最多招募用户数为 k 人, 同时保证区域 M 中感知单元的全覆盖。

$$\max \sum_{i=1}^d p_i \quad (1)$$

$$\text{s.t. } |U'| \leq k \quad (2)$$

每个感知单元都有用户进行感知覆盖, 可表示为

$$\sum_{i=0}^d c_i = d \quad (3)$$

将各个感知单元的感知质量 p_i 累加构成整个群体的感知质量。 $c_i = \{0, 1\}$, 记录 m_i 区域的覆盖情况。将区域感知质量作为优化目标, 不仅保证了移动群智感知的高实用性, 还考虑了优化感知开销。

4 系统算法

群智感知任务分发过程中包括分发算法和数据质量评估算法, 分发算法通过汤普森抽样算法对用户历史数据质量进行预测, 并结合贪婪策略覆盖目标区域。数据质量评估算法以群体间的差异性为基础, 量化个体的数据质量。

4.1 任务分发算法

4.1.1 汤普森抽样算法

汤普森抽样算法是一种在线决策问题的算法，其他在线算法包括探索阶段和利用阶段，汤普森抽样算法可以在一个阶段完成在线问题的两个关键点，一方面利用现有的信息最大化目标优化，另一方面积累未知信息以获得更多收益。通过这种随机策略招募了数据质量较优的用户，同时也给信息量较少的用户尝试感知任务的机会。在群智感知平台中，集合 U 中的 u_1, \dots, u_n 序列就是系统的决策变量。每个决策变量 $i \in \{1, \dots, n\}$ 都有相应的评估概率 θ_i ，表示用户达到数据质量阈值的概率， $1 - \theta_i$ 表示未达到数据质量阈值的概率。云平台通过历史数据得到用户数据质量达到标准的概率为 $(\theta_1, \dots, \theta_n)$ 。 $u^* = \max_i u_i$ 表示每次任务中的最优决策。在每轮任务中，感知任务的用户集合为有限集。系统随机产生条件概率为 $H_\theta(q_i | u_i)$ ，并选择决策变量为 u_i ，之后获取 u_i 的感知数据质量 q_i ，结果模型计算 u_i 的奖励 $r_i = r(q_i)$ 。首先，系统对 θ_i 未知，需要利用先验概率分布进行选择。汤普森抽样算法在分布 H_θ 中随机抽取样例，产生 θ^\wedge 。在有限结果集中， u_i 的奖励期望值为

$$E_{H_{\theta^\wedge}} [r(q_i) | u_i = u] = \sum_q H_{\theta^\wedge}(q | u) r(q) \quad (4)$$

通过观察可知，更新 P 的概率分布，如果 θ^\wedge 在有限集合 $(\theta_1, \dots, \theta_n)$ 中产生，那么条件概率密度函数 P 为

$$P_{p,H}(\theta = u | q_i, u_i) = \frac{P(u) H_u(q_i | u_i)}{\sum_v p(v) H_v(q_i | u_i)} \quad (5)$$

本文将汤普森抽样算法的 Beta 伯努利分布应用在任务分发问题中。在用户招募阶段，执行决策变量 u_i ，其奖励函数为 $r_i = \{1, -1\}$ ，并产生用户数据质量达到标准的概率为 $P(r_i = 1 | u_i, \theta) = \theta_i$ 。在用户招募初期，平台从独立的先验分布中获取 θ_i ，其分布包含参数 $\alpha = \{\alpha_1, \dots, \alpha_k\}$ 和 $\beta = \{\beta_1, \dots, \beta_k\}$ 。 θ_i 的先验概率密度函数 $p(\theta_i)$ 为^[20]

$$p(\theta_i) = \frac{\Gamma(\alpha_i + \beta_i)}{\Gamma(\alpha_i) \Gamma(\beta_i)} \theta_i^{\alpha_i - 1} (1 - \theta_i)^{\beta_i - 1} \quad (6)$$

其中， Γ 为 Gamma 函数。当得到用户数据质量后，根据贝叶斯准则更新用户后验概率分布为

$$\text{Beta}(\alpha_k, \beta_k) =$$

$$\begin{cases} \text{Beta}(\alpha_k + r(q_i), \beta_k), r(q_i) = 1 \text{ 且 } q_i \geq \mu \\ \text{Beta}(\alpha_k, \beta_k - r(q_i)), r(q_i) = -1 \text{ 且 } q_i < \mu \end{cases} \quad (7)$$

其中， μ 表示数据质量阈值。由此通过历史记录更新用户质量达到标准的后验概率，后续对其抽样作为用户招募指标。

4.1.2 区域覆盖策略

本文采用贪婪策略来解决区域全覆盖问题，考虑当前感知地图中的覆盖情况和用户抽样结果来选择最佳用户执行感知任务。在每个任务下的招募时序 $\langle u_1, \dots, u_n \rangle$ 会影响整个区域的覆盖情况，招募集合中的用户 u_i 则是招募时序 $\langle u_1, \dots, u_{i-1} \rangle$ 后的最佳招募人员。根据用户 u_i 当前的感知范围计算感知地图上新增的覆盖区域，即 $\sum_{m_j \in S_i, c_j=0} m_j, t \in 1, \dots, d$ 。同时，利用上述用户数据质量达标的后验概率来估计用户当前任务的数据质量，从而进行综合评估。选出优胜的用户后，更新感知地图上的感知质量及新增覆盖的单元格，TSUR 算法如下。

算法 1 TSUR 算法

输入 用户集为 U ，感知范围为 C

输出 招募的用户集为 U'

抽样模型

for $i=1$ to k do:

 抽样 $\theta_i \sim \text{Beta}(\alpha_i, \beta_i)$

end for

招募模型

while $|U'| < k$:

$$u_i \leftarrow \arg \max_i \left(\theta_i^* \frac{1}{|S_i|} * \left| \sum_{m_j \in S_i, c_j=0} m_j \right| \right)$$

$$U' = U' \cup u_i$$

for $m_j \subseteq S_i$:

$$p_j = \max \left\{ \frac{q_j}{|S_i|}, p_j \right\}, c_j = 1$$

end for

end while

更新概率分布

if $q_i \geq \mu: (\alpha_i, \beta_i) \leftarrow (\alpha_i + 1, \beta_i)$

if $q_i < \mu: (\alpha_i, \beta_i) \leftarrow (\alpha_i, \beta_i + 1)$

return U'

根据算法 1, 基于汤普森抽样算法和贪婪策略进行用户招募。首先, 从用户数据质量达到标准的先验概率分布中采样, 之后可替换为更新的后验分布概率。其次, 根据采样结果和用户覆盖范围实时评估感知地图中新增的感知质量, 从中挑选区域感知质量最佳的用户, 并将其加入招募用户集合。随即执行感知任务, 将感知结果提交至感知云平台, 由感知地图模型评估任务质量。最后, 更新平台反馈结果及相应参数, 如式(7)所示。

4.1.3 任务分发算法时间复杂度分析

假设在一次任务分发过程中, 用户集 U 存在 n 个用户。用户 u_i 的感知范围为 S_i , TSUR 算法的时间复杂度为 $O\left(\frac{n(3+n)}{2} + \max_i |S_i|^{n+1}\right)$ 。

$$\begin{aligned} \text{证明} \quad & n + [|S_1| + \dots + |S_n| + n] + [|S_1| + \dots + |S_{n-1}| + \\ & n - 1] + [|S_1| + 1] \leq n + \frac{n(1+n)}{2} + \max_i |S_i|^n + \max_i |S_i|^{n-1} \\ & + \dots + \max_i |S_i| \leq n + \frac{n(1+n)}{2} + \frac{\max_i (|S_i|^{n+1} - 1)}{\max_i (|S_i| - 1)} = \\ & \frac{n(3+n)}{2} + \frac{\max_i (|S_i|^{n+1} - 1)}{\max_i (|S_i| - 1)} = O\left(\frac{n(3+n)}{2} + \max_i |S_i|^{n+1}\right) \end{aligned}$$

在文献[21]中的用户招募算法, 其目标函数为使用户的招募花销最小化, 并保证用户覆盖感知任务区域。设所需覆盖的区域为 l_i , 在相同约束下, 每个区域 l_i 内存在 n_i 个用户, 其招募的时间为 $\sum_i n_i + n_i - 1 + \dots + 1$, 在用户资源分布均匀的条件

下, 时间复杂度为 $O\left(\frac{n(|l_i| + n)}{2|l_i|}\right)$, 与本文招募算法时间复杂度都在 $O(n^2)$ 数量级上。

4.2 数据质量评估算法

当用户招募结束后, 将感知数据提交至群智感知云平台, 系统根据用户的感知数据计算其数据质量。数据质量评估算法如算法 2 所示, 在算法 2 中, $\text{dis}(x, y)$ 函数可以采用欧氏距离、余弦距离、标准化距离等公式表示, 数据质量评估算法首先计算用户数据与其他所有用户数据的距离总和, 然后将最小距离和的用户数据定义为任务真值, 最后计算每个用户的数据质量, 其值与任务真值距离成反比, 具体算法如下。

算法 2 数据质量评估算法

输入 感知数据 X

输出 用户 U' 的感知质量 Q

x_i for 1 to k :

$$D_i = \sum_{j=1}^k \text{dis}(x_i, x_j)$$

end for

$$D_r = \min_i \{D_i\}$$

x_i for 1 to k :

$$q_i = \frac{1}{\text{dis}(x_i, x_r) + 1}$$

end for

return Q

对于可量化的感知数据, 根据算法 2 可以得到其数据质量。不同群体执行感知任务呈现不同的群体差异性, 差异性较小的群体达成意见一致, 表现出较好的数据质量。任务分发算法将逐渐提高这类群体的选择概率。

5 仿真实验

本文基于 Python 进行任务分发的仿真实验。与 BBTA 算法和 BUR 算法进行对比, 考查算法性能。BBTA 算法基于信心上界 (UCB, the upper confidence bound) 算法进行用户招募, 其算法通过积累用户的答题错误次数来降低用户的招募概率。BUR 是基于贪婪算法的用户招募策略, 将数据质量作为招募指标, 每次选择数据质量最好的用户。本文算法不仅考虑了用户长期以来的数据质量, 还考虑了用户实时的地理位置及感知范围进行任务分发。

对比实验的设置如下, 将感知任务发布到 15×15 的感知单元区域上, 每个用户随机出现在不同位置, 其附近若干个感知单元作为用户覆盖的感知区域。用户群体人数为 100~500 人, 不同任务设置不同招募人数进行测试实验。

本文目标优化为最大化区域数据质量。区域覆盖良好、用户数据质量优良的感知群体胜于过分集中、质量超群的群体。区域覆盖率如图 2 所示, 该场景由 500 个用户组成, 用户招募人数为 100 人, 展示了 100 个不同时刻的感知任务的覆盖情况。在用户资源充足的条件下, TSUR 算法的区域覆盖率可达 100%, 而 BBTA 算法和 BUR 算法的区域覆盖率约为 80%, TSUR 的平均区域覆盖率高出其余者约 20%。其中, 在 7 次感知任务中, 由于用户资源匮乏, 部分区域始终没有用户进行感知行为, 感知覆盖率未达 100%, 但同样高于 BBTA 算法和 BUR 算法的区域覆盖率。

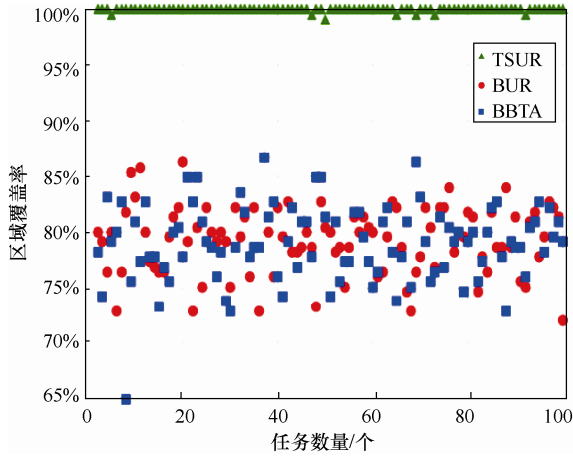


图2 区域覆盖率

区域累计感知质量如图3所示，图3展示了本文优化目标的50次任务招募对比实验，不改变场景参数设置。随着任务数量的增加，TSUR算法的累计区域感知质量增速大于其他两个算法的，不仅因为TSUR算法有效地招募了优良的群体，还因为它避免了大量重复感知的用户进行感知任务。

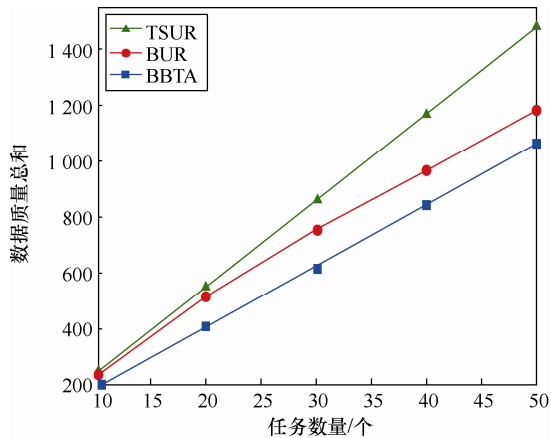


图3 区域累计感知质量

在本实验的基础上，改变招募人数，不同招募人数下的感知质量如图4所示，4个柱形的招募人数分别为50、100、150、200。TSUR算法的区域感知质量明显优于其余两者的，当招募100人以后，TSUR算法的区域数据质量变化不大。在此场景下，500人的群体招募100个用户就可以保证区域数据质量最佳。不同区域范围的感知质量如图5所示，图5展示了在基础设置上改变感知单元个数的区域感知质量。分别为10×10、15×15、20×20、25×25的感知单元，随着区域面积的增大，TSUR算法和BBTA算法的区域感知质量明显增长，而BUR算法的感知质量出现急剧下滑，主要因为其算法不能避免

由用户群体过于密集而导致的重复感知、质量失效的情况。在10×10的感知单元区域内，TSUR算法区域质量略低于其他两者，原因在于感知初期TSUR算法探索用户质量信息，而减少了对高质量用户的招募，随着感知区域的扩大，TSUR算法掌握了用户较稳定的概率分布，从而达到较好的区域感知质量。

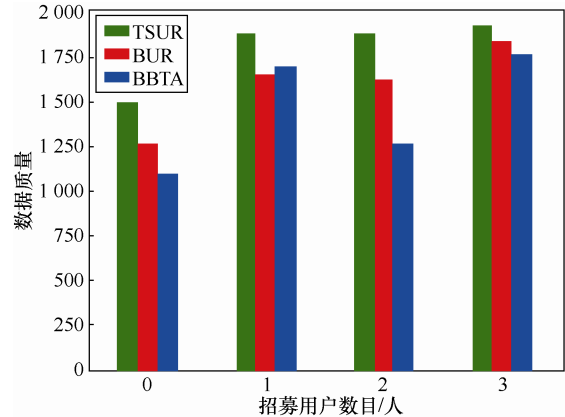


图4 不同招募人数下的感知质量

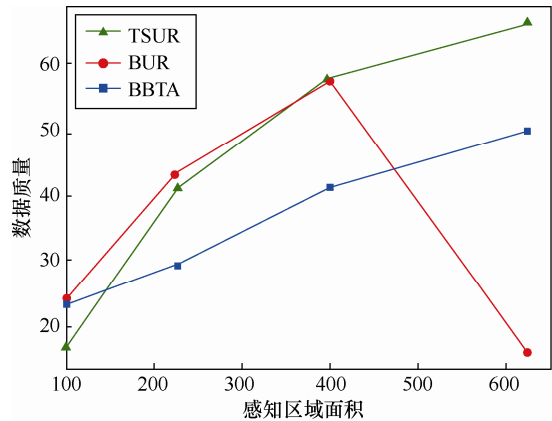


图5 不同区域范围的感知质量

在本文实验部分，通过区域覆盖率和用户的数据质量来保证群智感知模式的可靠性。在上述4组任务分发实验中，TSUR算法的区域数据质量高于同期BUR算法和BBTA算法的区域数据质量，同时TSUR算法的感知任务覆盖率高于BUR算法和BBTA算法的感知任务覆盖率约20%。TSUR算法的高性能表现在对用户招募指标的准确评估和感知覆盖的实时贪婪策略方面。

6 结束语

本文提出一种感知质量优化的群智感知任务分发算法，是一种基于汤普森抽样的在线学习算法，该算法将感知范围和数据质量作为指标，结合贪

婪算法, 实时地招募感知覆盖性能强、数据质量优的用户进行感知任务。对于无标准真值类型的感知任务提出了一种解决方案, 利用聚类性质评估用户数据质量和任务真值。在保障任务覆盖率的基础上, 优化群体感知质量。经过充分的实验表明, TSUR 算法在用户资源充足的条件下的覆盖率达 100%, 与已有的 BUR 算法和 BBTA 算法相比, 覆盖率分别提高了 30% 和 22%, 累计区域质量分别提高了 16% 和 20%。

参考文献:

- [1] LEASE M. On quality control and machine learning in crowdsourcing[C]//The 11th AAAI Workshop on Human Computation. AAAI Press, 2011: 97-102.
- [2] 赵东, 马华东, 刘亮. 移动群智感知质量度量与保障[J]. 中兴通讯技术, 2015, 21(6): 2-5.
ZHAO D, MA H D, LIU L. Quality measuring and assurance for mobile crowd sensing[J]. ZTE Technology Journal, 2015, 21(6): 2-5.
- [3] ZHANG H, MA Y, SUGIYAMA M. Bandit-based task assignment for heterogeneous crowdsourcing[J]. Neural Computation, 2015, 27(11): 2447-2475.
- [4] XIAO M J, WU J, ZHANG S, et al. Secret-sharing-based secure user recruitment protocol for mobile crowdsensing[C]//IEEE INFOCOM 2017 IEEE Conference on Computer Communications. IEEE, 2017.
- [5] GUO B, YU Z W, ZHANG D Q, et al. From participatory sensing to mobile crowd sensing[C]//2014 IEEE International Conference on Pervasive Computing and Communication Workshops (PERCOM WORKSHOPS). IEEE, 2014: 593-598.
- [6] ZHAO Y R, ZHANG Y, YU T, et al. CityDrive: a map-generating and speed-optimizing driving system[C]//IEEE INFOCOM 2014 IEEE Conference on Computer Communications. IEEE, 2014: 1986-1994.
- [7] KIM S, ROBSON C, ZIMMERMAN T, et al. Creek watch: pairing usefulness and usability for successful citizen science[C]//SIGCHI Conference on Human Factors in Computing Systems. ACM, 2011: 2125-2134.
- [8] MATHUR S, JIN T, KASTURIRANGAN N, et al. ParkNet: drive-by sensing of road-side parking statistics[C]//The 8th International Conference on Mobile Systems, Applications, and Services. ACM, 2010: 123-136.
- [9] SIMOENS P, XIAO Y, PILLAI P, et al. Scalable crowd-sourcing of video from mobile devices[C]//The 11th Annual International Conference on Mobile Systems, Applications, and Services. ACM, 2013: 139-152.
- [10] WANG J T, WANG Y S, ZHANG D Q, et al. Multi-task allocation in mobile crowd sensing with individual task quality assurance[J]. IEEE Transactions on Mobile Computing, 2018, 17(9): 2101-2113.
- [11] KO H, PACK S, LEUNG V C M. Coverage-guaranteed and energy-efficient participant selection strategy in mobile crowdsensing[J]. IEEE Internet of Things Journal, 2018: 1.
- [12] WANG E, YANG Y J, LOU K H, et al. User selection utilizing data properties in mobile crowdsensing[J]. Information Sciences, 2019, 490: 210-226.
- [13] MAARRY K E, BALKE W T, CHO H, et al. Skill ontology-based model for quality assurance in crowdsourcing[C]//International Conference on Database Systems for Advanced Applications. Springer, 2014.
- [14] HU X, LI X, NGAI E C H, et al. Multidimensional context-aware social network architecture for mobile crowdsensing[J]. IEEE Communications Magazine, 2014, 52(6): 78-87.
- [15] ALLAHBAKHS M, BENATALLAH B, IGNJATOVIC A, et al. Quality control in crowdsourcing systems: issues and directions[J]. IEEE Internet Computing, 2013, 17(2): 76-81.
- [16] LIU S Z, ZHENG Z Z, WU F, et al. Context-aware data quality estimation in mobile crowdsensing[C]//IEEE INFOCOM 2017 IEEE Conference on Computer Communications. IEEE, 2017.
- [17] YANG S, WU F, TANG S J, et al. On designing data quality-aware truth estimation and surplus sharing method for mobile crowdsensing[J]. IEEE Journal on Selected Areas in Communications, 2017: 832-847.
- [18] GUO B, CHEN H H, QI H, et al. Worker-contributed data utility measurement for visual crowdsensing systems[J]. IEEE Transactions on Mobile Computing, 2017, 16(8): 2379-2391.
- [19] SALAS O F, ADZIC V, KALVA H. Subjective quality evaluations using crowdsourcing[C]//2013 Picture Coding Symposium(PCS). IEEE, 2013.
- [20] RUSSO D, VAN R B, KAZEROUNI A, et al. A tutorial on Thompson sampling[J]. Foundations and Trends in Machine Learning, 2017, 11(1): 112.
- [21] KARALIOPOULOS M, TELELIS O, KOUTSOPOULOS I. User recruitment for mobile crowdsensing over opportunistic networks[C]//Computer Communications. IEEE, 2015: 2254-2262.

[作者简介]



张伟 (1996-), 女, 北京人, 北京信息科技大学硕士生, 主要研究方向为移动群智感知。



李卓 (1983-), 男, 河南南阳人, 博士, 北京信息科技大学副教授, 主要研究方向为移动无线网络、分布式计算。



陈昕 (1965-), 男, 江西南昌人, 博士, 北京信息科技大学教授, 主要研究方向为网络性能评价、网络安全。