

基于移动性预测的群智感知混合式任务分配

蒋伟进^{1,2,3}, 聂彩燕^{1,2*}, 刘 茜^{1,2}, 杜熙晨^{1,2}, 杨 璇^{2,3}, 蒋意容⁴

(1. 湖南工商大学计算机学院, 湖南长沙 410205; 2. 湘江实验室, 湖南长沙 410205;
3. 湖南工商大学前沿交叉学院, 湖南长沙 410205; 4. 湖南信息学院, 湖南长沙 410151)

摘 要: 移动群智感知(Mobile Crowd Sensing, MCS)通过用户随身携带的感知设备来收集数据,是一种大规模数据感知范式,其中任务分配是主要挑战之一. 文章研究具有异构质量延迟敏感任务的混合用户任务分配问题,设计目标是在机会式用户和参与式用户共享总预算下,最大限度地提高任务完成质量. 针对现有预测方法预测精度不足问题,文章提出一种基于迁移学习的移动性预测模型,通过将轨迹丰富的旧参与者的数据转移给新参与者,解决历史数据稀缺导致的预测误差. 基于此预测模型,设计了一个混合用户任务分配算法,该算法利用移动性预测模型为机会式用户分配任务;此外,将剩余任务聚类到不同区域,构造二部图匹配问题使参与式用户和任务区域绑定;之后提出一种基于行程距离平衡的蚁群优化算法(Ant Colony Optimization algorithm based on Travel Distance Balance, ACOTDB),在用户行程距离预算下实现最优路径规划. 在真实数据集上的大量仿真实验表明,与现有算法进行比较,本文算法在任务完成质量和任务分配效率方面具有显著的优势,验证了其有效性.

关键词: 移动群智感知;任务分配;迁移学习;移动性预测模型;蚁群优化算法;质量异构

基金项目: 国家自然科学基金(No.61772196);湖南省自然科学基金(No.2020JJ4249);湖南省教育厅科学研究重点项目(No.21A0374)

中图分类号: TP274 文献标识码: A 文章编号: 0372-2112(2025)06-1979-17

电子学报 URL: <http://www.ejournal.org.cn> DOI: 10.12263/DZXB.20240632

Hybrid Task Assignment for Crowd Sensing Based on Mobility Prediction

JIANG Wei-jin^{1,2,3}, NIE Cai-yan^{1,2*}, LIU Qian^{1,2}, DU Xi-chen^{1,2}, YANG Xuan^{2,3}, JIANG Yi-rong⁴

(1. School of Computer Science, Hunan University of Technology and Business, Changsha, Hunan 410205, China;

2. Xiangjiang Laboratory, Changsha, Hunan 410205, China;

3. College of Frontier Intersection, Hunan University of Technology and Business, Changsha, Hunan 410205, China;

4. Hunan University of Information Technology, Changsha, Hunan 410151, China)

Abstract: Mobile crowd sensing (MCS) collects data through the sensing devices carried by users and is a large-scale data sensing paradigm, where task allocation is one of the main challenges. This paper studies the task allocation problem of mixed users with heterogeneous quality delay-sensitive tasks. The design objective is to maximize the quality of task completion under the shared total budget of opportunistic users and participatory users. In response to the problem of insufficient prediction accuracy of existing prediction methods, this paper proposes a mobility prediction model based on transfer learning. By transferring the data of old participants with rich trajectories to new participants, it solves the prediction errors caused by the scarcity of historical data. Based on this prediction model, a mixed user task allocation algorithm is designed. The algorithm uses the mobility prediction model to allocate tasks to opportunistic users. In addition, the remaining tasks are clustered into different areas, and a bipartite graph matching problem is constructed to bind participatory users and task areas. Subsequently, an ant colony optimization algorithm based on travel distance balance (ACOTD) is proposed to achieve optimal path planning under the user's travel distance budget. Through a large number of simulation experiments on real datasets, this paper compares with existing algorithms. The results show that the algorithm has significant advantages in task completion quality and task allocation efficiency, verifying its effectiveness.

Key words: mobile crowd sensing; task allocation; transfer learning; mobility prediction model; ant colony optimization algorithm; mass heterogeneity

Foundation Item(s): National Natural Science Foundation of China (No.61772196); Natural Science Foundation of Hunan Province (No.2020JJ4249); Key Project of Hunan Provincial Department of Education for Science and Technology Research (No.21A0374)

1 引言

随着无线网络技术的快速发展和移动设备的普及,移动群智感知(Mobile Crowd Sensing, MCS)^[1,2]已成为一种新兴的感知范式.典型的MCS以大量移动用户及其携带的智能设备作为基础感知单元,通过其广泛存在性、灵活移动性和机会连接性来完成大规模且细粒度的感知任务^[3],该技术已经广泛应用于智慧交通^[4]、环境监测^[5]、噪声采集^[6]、公共安全^[7]等领域.

MCS系统由3个部分组成:MCS平台、任务请求者和移动用户.MCS平台从请求者处收集传感任务,然后利用大量具有移动设备(例如配备多功能传感器的智能手机和可穿戴设备)的移动用户来完成传感任务.然而,移动用户需要物理地移动到感兴趣的位置来完成任务,这从本质上影响了系统的总体任务完成效率.

任务分配是MCS中的主要挑战之一^[8],对感知任务的质量和MCS系统的效率有重大影响.现有研究根据用户的运动模式和参与机制可将移动用户划分为机会式感知用户和参与式感知用户.在机会式感知中,移动用户在不改变轨迹的日常移动路线上,通过无意识地收集传感数据执行任务^[9].在参与式感知中,用户根据平台的需求移动到指定位置去完成传感任务,使参与式感知比机会式感知更具灵活性和效率,但会产生除了固有感知成本之外的额外移动成本.

尽管上述2种传感模式在任务分配方面取得了很大进展,但仍有2个至关重要的问题需要研究.

对于机会式感知模式,尽管已经提出不同的轨迹预测算法^[10,11],并在一定程度上证明了其预测能力,但是大多数移动性预测模型^[12,13]都是在基于丰富的历史轨迹数据基础之上,对于只有少量历史轨迹记录的新参与者来说,预测结果并不理想.因此,在历史轨迹数据稀缺的情况下,设计一个准确的移动性预测模型进行有效的任务分配是本文的第1个挑战.

混合式感知通过结合参与式传感和机会式传感的方式有效提升了系统的整体性能,现有混合式感知研究已经取得了一定的进展.然而,现有研究普遍假设所有任务的执行时间是固定不变的,这忽视了任务执行时机对数据质量的影响,因此现有研究并不适合具有延迟敏感的任务.综上,如何针对质量异构的延迟敏感任务实现高效的分配算法是本文的第2个挑战.

受上述挑战启发,本文研究了质量异构延迟敏感

任务的混合用户任务分配,综合利用不同类型用户的特性来提高延迟敏感任务的执行性能,以实现感知成本和数据质量之间的权衡.针对问题1,面对参与者历史轨迹数据稀缺的挑战,本文提出一种基于迁移学习的移动性预测模型,有效降低了参与者由于历史数据稀缺而导致的预测误差.针对问题2,本文的目标是在给定预算约束的情况下,结合2种模式的优势,使2种类型的用户任务完成质量最大化.因此,本文提出一种基于混合用户感知的任务分配算法(Hybrid User sensing based Task Allocation Algorithm, HUTAA),首先,在预算约束下利用移动性预测模型对机会式用户进行任务分配;然后,将剩余的任务聚类到不同的区域,并且将用户和区域绑定;最后,执行基于行程距离平衡的蚁群优化算法(Ant Colony Optimization algorithm based on Travel Distance Balance, ACOTDB),用于规划参与式用户路径.大量的仿真结果表明了该算法的有效性,本文的主要贡献如下:

(1)提出一个基于混合用户的任务分配框架,该框架集成了移动性预测模型和混合用户任务分配,利用机会式用户和参与式用户高效地执行延迟敏感任务,实现感知成本和感知质量之间的良好平衡.

(2)提出基于迁移学习的移动性预测模型,该模型提高了用户在历史数据稀缺情况下的轨迹预测精度,增加后续任务分配算法的准确性.

(3)使用2个真实的轨迹数据集评估HUTAA,实验结果证明了HUTAA在提高感知任务质量方面的有效性.

2 相关工作

MCS中高效任务分配算法的设计受到了广泛的关注,并开展了大量的研究.根据用户的参与程度,现有工作可以分为3种类型:机会式任务分配、参与式任务分配和混合式任务分配.

2.1 机会式任务分配

在机会式感知中的任务分配,要求参与者在日常移动路线覆盖的单元格中无意识地完成感知任务.关键问题是如何在预算约束下从机会式用户集合中招募一组合适的用户执行任务,同时使任务完成质量最大化.

Zhang等人^[10]提出基于概率和控制模糊系统的移动预测方法来估计工人的移动行为,根据移动预测结

果优化总体任务完成率. Nasser 等人^[13]提出一种用于连续感知任务的新型招募系统,该系统不依赖历史移动轨迹来招募工人,转而采用机器学习实时预测用户在感知期间的位置. 为提高用户轨迹预测准确率,在文献[14]和文献[15]中,通过建立基于马尔可夫链的移动性模型,预测用户空间分布以进行用户选择,从而提高感知覆盖质量. Song 等人^[16]则从另一个角度出发,提出 TaskMat 框架,该框架利用工人的任务偏好来实现覆盖率保证的任务分配,将合格的工人迁移到不太受欢迎的任务以增加任务覆盖率. Guo 等人^[17]则将 MCS 技术引入海洋观测领域,提出基于 Transformer 的轨迹预测算法,再结合离散粒子群优化算法,提出一种海洋 MCS 的动态参与者选择算法,旨在最小化系统成本.

2.2 参与式任务分配

在参与式感知的任务分配问题中,参与式用户根据平台的需求移动到指定位置去完成任务,关键问题是如何在给定的资源约束下,为参与者设计移动路线分配任务位置,从而最大化整个感知质量或最大化感知平台利润.

Peng 等人^[18]考虑每个任务都需要一定执行时间的要求,提出基于二分匹配的策略进行任务分配,在满足每个任务时间要求的同时最大化 MCS 平台的总利润. 蒋伟进等人^[19]根据用户与任务的适应程度,定义空间匹配度和技能匹配度,提出一种结合非线性递减收敛因子、最优局部抖动及动态位置更新 3 种操作的改进鲸鱼优化算法,在保证感知任务基本完成质量的前提下降低任务分配的成本. 与文献[18]类似, Yang 等人^[20]在每个任务都有执行时间的约束下,提出一种模糊时间序列分析方法来预测特定时空内每个任务可用的参与者数量. 吕翊等人^[21]从全新的角度出发,关注将经验不足的工人培养为优质工人,并令其执行这些感知任务,实现工人的长期复用,提高感知数据质量和长期平台效用. Ma 等人^[22]提出 MCS 中的多异构任务和异构用户招募问题,目标是 minimized 平台总支付并最大化任务覆盖率. Wang 等人^[23]利用异构任务之间的隐含时空相关性提出并解决 MCS 系统中重要异构多任务分配问题,最大限度地提高数据质量并最小化总激励预算. Shen 等人^[24]提出一种基于小组协作模式的异构多项目多任务分配模型,在技能匹配和完成时间的约束下,最大化感知质量并最小化平台成本.

2.3 混合式任务分配

上述工作主要考虑单一的基于机会的感知或基于参与的感知模式任务分配,文献[25~29]联合考虑了 2 种感知模式.

Wang 等人^[25]提出一个两阶段的混合任务分配框架,目标是在预算约束下最大化完成任务的数量. 该框

架通过在离线阶段利用半马尔可夫模型预测机会式参与者的移动性选择用户,从在线阶段选择参与式用户,但没有考虑在历史移动数据不足的情况下机会主义用户的预测精度问题. Wei 等人^[26]提出一种基于稀疏 MCS 中的混合感知模式的任务分配框架,在总成本约束下同时招募机会主义和参与式参与者在重要单元中执行任务,同时考虑了他们对传感地图推断的贡献. Lv 等人^[27]在机会用户选择阶段,设计一个更新的迭代算法来选择低成本、高覆盖率的机会用户集;在参与用户选择阶段,设计一种结合分组和匹配的算法来解决任务位置受限的问题. Wang 等人^[28]提出一种基于部分迁移学习的两阶段参与者选择方法,在第 1 阶段通过迁移源任务的部分用户资源到特征空间相似的目标子任务中,以确保高效准确的参与者选择. 在第 2 阶段针对未覆盖的子任务,选择机会式用户进行任务分配. Zhang 等人^[29]提出一种基于区块链的混合用户任务分配方案,该方案用区块链取代传统的中央服务器,从而确保系统的安全性.

在 MCS 领域中,任务分配算法的设计和优化始终是研究的热点. 尽管如此,现有混合感知算法通常基于一个假设,即所有任务都是长寿且具有不变的同等质量,但这种假设忽略了现实世界中任务的动态性和质量异构性.

本文特别关注任务的延迟敏感性,即任务的完成质量会随着时间的推移而衰减. 在此基础上,本文设计一种基于迁移学习的移动性预测模型. 传统的移动性预测模型往往依赖轨迹数据丰富的参与者,而本文模型考虑了新参与者轨迹稀疏性问题. 在移动性预测的基础上,提出一种混合式感知用户任务分配算法. 该算法特别考虑了任务的延迟敏感性,为参与式参与者规划了高效的任务完成路径. 本文提出的分配方法,能够在给定的时间和预算约束下,最大化任务的整体完成质量.

3 系统模型

3.1 模型概述

本文的设计目标是在给定的预算约束下最大限度地提高任务完成质量. 本文在 MCS 框架下正式定义了混合用户任务分配问题,表 1 列出了主要符号说明,系统建模如下.

3.1.1 用户模型

在 MCS 的任务分配问题中,本文假设用户是真实的,并且所有用户都具有相同的感知能力,其中 MCS 系统中有 n_1 个机会式用户和 n_2 个参与式用户,分别建模如下:

机会式用户 (Opportunistic Users): 令 $OU = \{ou_1, ou_2, \dots, ou_{n_1}\}$ 表示机会式用户集.

参与式用户 (Participatory Users): 令 $PU =$

表1 符号说明

符号	说明
ou_i, OU	机会式用户 ou_i 和机会式用户集合
pu_i, PU	参与式用户 pu_i 和参与式用户集合
B	平台招募用户的总预算
TB_i	用户 pu_i 的行程距离预算
$T = \{t_1, t_2, t_3, \dots, t_{ T }\}$	任务集合
$TE(a_x^i)$	参与者 ou_i 的第 x 个停留区域的时间熵
$A_{q_j}^i$	任务 t_j 的完成质量
$time_{ij}$	机会式用户 $ou_i \in OU$ 或参与式用户 $pu_i \in PU$ 从当前位置到达任务 t_j 位置的时间
α	时间衰减因子
$Dist(\text{tail}, t_j)$	用户已经走过的路径尾部到新任务 t_j 位置的距离
ov_i, pv_i	机会式用户 $ou_i \in OU$ 或参与式用户 $pu_i \in PU$ 的行走速度
C_{pu_i}, C_{ou_i}	参与式用户和机会式用户完成任务的奖励
θ_i, δ_i	参与式用户距离的移动成本和速度的单位成本
$Dist_i^k$	参与者 pu_i 在其分配集合执行 k 个任务的总距离
x_{ij}, y_{ij}	表示机会式用户 ou_i 是否被分配到任务 t_j , 参与式用户 pu_i 是否被分配到任务 t_j
ST^n, ST^o	表示新参与者和旧参与者的最大语义轨迹模式
$LCS(ST^n, ST^o)$	新旧参与者的最长公共序列
$A^n = \{a_1^n, a_2^n, \dots, a_l^n\}$	新参与者 ou_{new} 的停留区域
$A^o = \{a_1^o, a_2^o, \dots, a_l^o\}$	旧参与者 ou_{old} 的停留区域
$RA = \{RA_1, \dots, RA_{n_2}\}$	剩余区域
$RA_{ik} = \{RA_{i_1k}, \dots, RA_{i_{m_2}k}\}$	区域 RA_k 的任务集合
$Num(RA_{ik})$	区域 RA_{ik} 的任务数量
$\tau_{\text{tail}, j}$	连接当前路径尾部和任务 t_j 弧上的信息素浓度
ρ	信息素蒸发率
τ_0	初始信息素浓度
ΔA_{t_j}	完成任务 t_j 之后的质量增量
$Dist(t_j, \text{des}_i)$	任务 t_j 位置到用户 pu_i 目的位置的距离

$\{pu_1, pu_2, \dots, pu_{n_2}\}$ 表示参与式用户集合。

与机会式用户不同,服务平台可以选择将招募的参与式用户派遣到任务位置完成任务,每个用户 pu_i 都有一个行程距离预算 TB_i ,该预算限制了用户所走总路径的长度。

3.1.2 任务模型

本文将目标感知区域根据感知任务划分为 L 个大

小相等的单元格,记作 $L = \{l_1, l_2, \dots, l_L\}$,每个区域的数据采集代表一个感知任务,将所有任务集记 $T = \{t_1, t_2, t_3, \dots, t_{|T|}\}$ 。

在本文中,工人对任务的感知时间与工人的行走时间相比往往可以忽略不计。一个原因是本文关注的感知任务执行时间与工人的出行时间相比非常短。另一个原因是本文考虑感知任务的细粒度组成,那么在每个位置通常只有一个任务,而且用户不必在一个地方停留很长时间来完成任务,因此任务执行时间可以忽略不计,用户的到达时间则视为任务完成时间。此外,假设需要上传的任务量很小,那么其上传时间和成本都可以忽略不计。因此,用户最早到达任务位置的时间决定了任务质量。任务完成质量随着时间的推移而衰减,计算方法如下:

$$A_{q_j}^i = q_j \times e^{-\text{time}_{ij} \times \alpha} \quad (1)$$

其中, q_j 为任务 t_j 的最大质量, α 为时间衰减因子, time_{ij} 为机会式用户 $ou_i \in OU$ 或参与式用户 $pu_i \in PU$ 从初始位置到达任务位置的时间。 time_{ij} 的计算方法如下:

$$\text{time}_{ij} = \begin{cases} \frac{Dist(\text{tail}, t_j)}{ov_i}, & \text{if } ou_i \in OU \\ \frac{Dist(\text{tail}, t_j)}{pv_i}, & \text{if } pu_i \in PU \end{cases} \quad (2)$$

其中, $Dist(\text{tail}, t_j)$ 表示用户已经走过的路径尾部到新任务 t_j 位置的距离, ov_i 和 pv_i 分别表示机会式用户 $ou_i \in OU$ 或参与式用户 $pu_i \in PU$ 的行走速度。

3.1.3 激励模式

激励是鼓励员工参与感知活动的重要手段。一般来说,员工可以在完成分配任务后领取相应报酬,该报酬主要取决于完成指定任务所花费的成本,因此可以表示为成本的函数。激励机制设计是 MCS 的另一个具有挑战性的领域,与任务分配密切相关。在这项工作中,重点研究了多用户任务分配,因此简单地假设支付给员工的报酬等于其成本,以减少激励对任务分配结果的影响。

为了激励用户完成任务,平台派遣执行任务的用户根据其出行距离和参与方式获得一定的奖励。根据用户的参与模式,本文有 2 种不同的激励模式。

对于参与式用户来说,更愿意改变他们的路线来完成传感任务,每个人都可以获得与实际在线行走距离成比例的奖励。另一方面,行走速度是影响工人成本的另一个关键因素。因此综合考虑移动距离和行走速度,参与式用户完成任务的奖励为

$$C_{pu_i} = Dist(\text{tail}, t_j) \times \theta_i + pv_i \times \delta_i \quad (3)$$

其中,比例系数 θ_i 表示参与式用户 pu_i 单位距离的移动

成本, $\text{Dist}(\text{tail}, t_j)$ 表示用户从当前任务尾部完成分配的任务 t_j 的最短路径长度, p_{v_i} 表示用户 pu_i 的行走速度, δ_i 表示参与式用户行走速度的单位成本, 参与者 pu_i 在其分配集合执行 k 个任务的总距离可以计算为

$$\text{Dist}_i^k = \text{Dist}(\text{pu}_i, t_1) + \sum_{i=1}^{k-1} \text{Dist}(t_i, t_{i+1}) \quad (4)$$

对于机会式用户来说, 在日常移动轨迹中无意识地完成感知任务, 平台需要向每个人支付固定的、相等的激励奖励 C_{ou} .

3.2 问题定义

在本文中, 混合式任务分配的目标是在给定的预算约束下, 为不同类型的用户分配任务和规划路径, 以最大限度提高任务的总完成质量. 为了表述这个问题, 本文定义一组变量 x_{ij} 表示用户 ou_i 是否被分配到任务 t_j , 变量 y_{ij} 表示用户 pu_i 是否被分配到任务 t_j , 则以任务总完成质量最大化为目标的混合式任务分配问题可表述为

$$\max \left(\sum_{i=1}^{n_1} \sum_{j=1}^{|T|} x_{ij} A_{q_i}^j + \sum_{i=1}^{n_2} \sum_{j=1}^{|T|} y_{ij} A_{q_i}^j \right) \quad (5)$$

$$\text{s.t.} \quad \sum_{i=1}^{n_1} \sum_{j=1}^{|T|} x_{ij} C_{\text{ou}} + \sum_{i=1}^{n_2} \sum_{j=1}^{|T|} y_{ij} C_{\text{pu}_i} \leq B \quad (6)$$

$$\sum_{i=1}^{n_1} x_{ij} + \sum_{i=1}^{n_2} y_{ij} \leq 1, t_j \in T \quad (7)$$

$$\text{Dist}_i^k \leq \text{TB}_i, \forall \text{pu}_i \in \text{PU} \quad (8)$$

$$\text{var. } x_{ij} \in \{0, 1\}, y_{ij} \in \{0, 1\}, \quad (9)$$

$$\forall \text{ou}_i \in \text{OU}, \text{pu}_i \in \text{PU}, t_j \in T$$

其中, 目标函数式(5)表示使任务总完成质量最大化, 约束式(6)表示保证所有用户获得的总奖励不超过平台总预算, 约束式(7)表示保证每个任务只能完成1次, 约束条件式(8)表示保证不违反每个用户的出行距离预算.

4 基于移动性预测的机会式任务分配

本文提出基于迁移学习的移动性预测模型, 旨在解决目标轨迹数据稀缺时的预测精度不足问题. 迁移学习允许一个预训练的模型被调整或迁移到一个不同但相关的任务上, 通过迁移学习, 该模型从轨迹丰富的参与者向数据稀缺的参与者转移合适的轨迹数据. 基于移动性预测的机会式任务分配方案如图1所示. 第1步, 称为停留位置提取, 将每个用户的轨迹转换为停留位置序列; 第2步, 称为最大语义轨迹模式挖掘, 提取用户的语义行为, 并根据用户的语义行为相似度得到用户聚类; 第3步, 称为生活模式分类, 根据时间熵将所有参与者归类为不同的生活模式; 第4步, 称为新老用户相似性度量, 综合语义相似性和时间相似性比较新老用户之间的相似性程度; 第5步, 称为基于贪心算法的任务分配, 贪婪地选择效用最大的机会式用户.

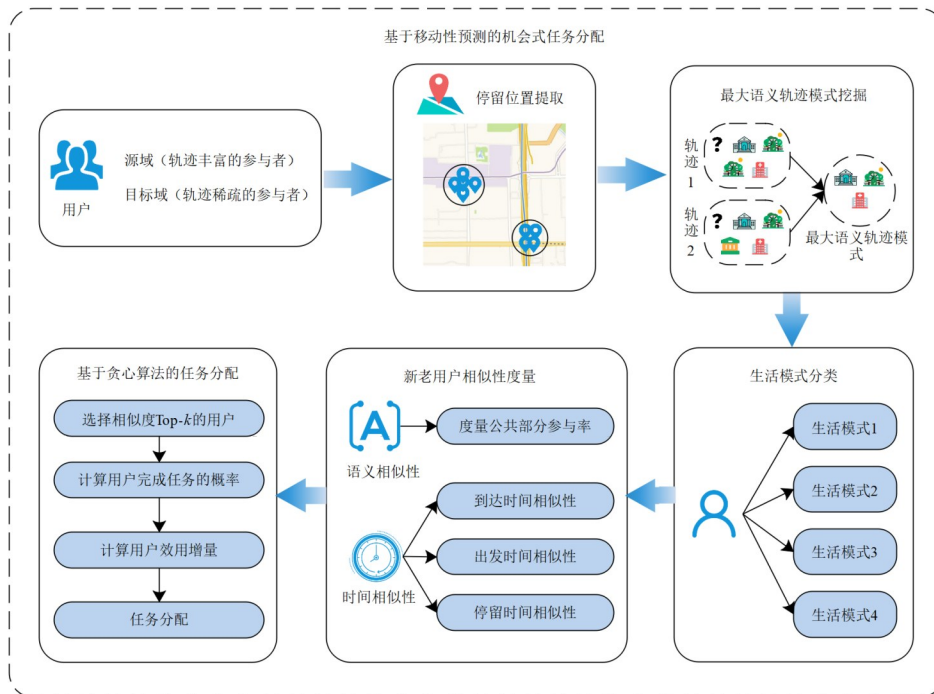


图1 机会式任务分配总体框架图

4.1 停留位置提取

地理信息通常由一系列地理点组成, 仅仅使用地

理信息不能反映用户的运动行为^[30]. 语义轨迹由一系列用语义标签标记的位置组成, 语义轨迹反映了用户

的轨迹频繁模式. 因此, 本文利用语义轨迹来衡量用户相似性. 首先, 采用基于时间和空间阈值法从地理轨迹提取停留点^[31], 停留区域是移动用户停留时间超过时间阈值 T_o 的地理区域, 区域的中心就是停留点. 然后, 基于阈值的停留点提取可能会产生大量的冗余数据. 本文采用基于密度的聚类算法^[32]将距离较近且具有相似语义标签的停留点聚成一个停留区域.

然后, 本文使用地理信息数据库为发现的停留区域分配语义词. 将语义词分配给停留区域后, 停留区域序列可以转化为地标序列, 称之为语义轨迹.

4.2 最大语义轨迹模式挖掘

将每个地理轨迹转化为语义轨迹之后, 每个用户的地理轨迹集就被转换为一个语义轨迹数据集. 用户的运动随时间变化将会导致语义轨迹的多样化, 但是用户的主要行为可能是固定的. 因此, 为了识别用户频繁的移动行为, 本文在每个用户的语义轨迹数据集上执行 Prefix-Span 序列模式挖掘算法来挖掘频繁的语义轨迹. 如图2所示, 给定一个移动用户的轨迹1和轨迹2, 轨迹1=<停留区域0, 停留区域1, 停留区域2, 停留区域3>, 轨迹2=<停留区域0, 停留区域1, 停留区域4, 停留区域3>, 其中停留区域0没有与任何地标重叠, 因此被赋予了“未知”的语义标签, 而停留区域1与“学校”和“公园”2个地标重叠, 因此被赋予了<学校, 公园>的语义标签. 移动用户的轨迹记录被转换成语义轨迹如下: 轨迹1=<{未知}, {学校, 公园}, {公园}, {医院}>, 轨迹2=<{未知}, {学校, 公园}, {银行}, {医院}>. 在此过程中, 只保留<{学校, 公园}, {医院}>, 被称为最大语义轨迹模式, 也就是覆盖范围最广、频率最高的语义轨迹模式, 用来表示用户频繁的移动行为.

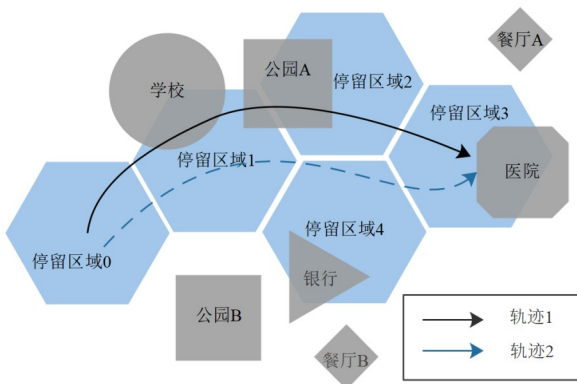


图2 语义轨迹转换图

4.3 生活模式分类

在 MCS 平台上, 有大量的旧参与者和新参与者, 需要从旧参与者中找到与新参与者行为相似的个体, 生活模式描述了参与者在一段时间内的行为模式, 如果新参与者和旧参与者频繁停留区域的数量大致相同,

则认为他们有相似的生活模式.

本文根据时间熵筛选出具有相似生活模式的旧参与者, 来判断他们的生活模式是否相似. 接下来采用时间熵来检测停留区域是否为经常停留区域. 参与者 ou_i 的第 x 个停留区域的时间熵表示为

$$TE(a_x^i) = - \sum_{k=1}^d \frac{|P_k^i|}{\sum_{k=1}^d |P_k^i|} \log \frac{|P_k^i|}{\sum_{k=1}^d |P_k^i|} \quad (10)$$

其中, a_x^i 由 ou_i 的 d 日轨迹内的停留点组成, P_k^i 是第 k 天的停留点集. 时间熵用以衡量到达停留区域的频率, 值越高表示越频繁.

4.4 新老用户相似性度量

最大语义轨迹模式代表了用户在真实世界中的语义行为之一, 本文考虑2个用户之间的语义相似性, 即它们的最大语义轨迹模式的相似性. 本文使用最大语义轨迹模式相似度来衡量新旧参与者之间语义轨迹模式相似度. 给定2个最大语义轨迹模式, 当它们有更多的共同部分时, 意味着它们更相似. 因此, 本文使用这2种模式的最长公共序列 (Longest Common Sequence, LCS) 来表示它们的最长公共部分. 本文假设新参与者的最大语义轨迹模式为 $ST^n = \langle \{School\}, \{Hospital\}, \{Park, Bank\}, \{Restaurant\} \rangle$, 旧参与者的最大语义轨迹模式为 $ST^o = \langle \{School, Market\}, \{Park\}, \{Restaurant\} \rangle$, 则新旧参与者的最长公共序列为 $LCS(ST^n, ST^o) = \langle \{School\}, \{Park\}, \{Restaurant\} \rangle$, 新用户参与者公共部分的参与率定义如下:

$$\text{ratio}(LCS(ST^n, ST^o), ST^n) = \frac{\sum_{i=1}^{|ST^n|} \sum_{j=1}^{|LCS(ST^n, ST^o)|} M(ST_i^n, LCS_j)}{|ST^n|} \quad (11)$$

where $M(ST_i^n, LCS_j)$

$$= \begin{cases} \frac{|ST_i^n \cap LCS_j|}{|ST_i^n|}, & \text{if } LCS_j \text{ is matching to } ST_i^n \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

接下来, 通过衡量新参与者 ST^n 和旧参与者 ST^o 的公共部分计算平均参与比例, 使用均等平均法直接计算2个比值的平均值, 即新旧参与者之间的语义相似度定义为

$$\text{Sim}(ST^n, ST^o) = \frac{\text{ratio}(LCS(ST^n, ST^o), ST^n)}{2} + \frac{\text{ratio}(LCS(ST^n, ST^o), ST^o)}{2} \quad (13)$$

然而, 仅从语义信息的角度计算参与者的相似性, 可能无法准确描述参与者的行为模式. 因此, 本文在计算相似性时考虑了时间因素. 停留区域的时间相似性包括到达时间相似性、出发时间相似性、停留时间相似

性. 假设新参与者 ou_{new} 的停留区域表示为 $A^n = \{a_1^n, a_2^n, \dots, a_l^n\}$, 而旧参与者 ou_{old} 的停留区域具有丰富的历史轨迹数据, 表示为 $A^o = \{a_1^o, a_2^o, \dots, a_l^o\}$, 则到达时间相似性可以定义为

$$\begin{aligned} \text{Sim}_t^A(a_i^n, a_j^o) &= \sum_{k=t_i^n - \Delta t}^{t_i^n + \Delta t} w(k) \cdot P_A(a_j^o, k) \\ &= \sum_{k=t_i^n - \Delta t}^{t_i^n + \Delta t} \frac{1}{\sigma_a \sqrt{2\pi}} e^{-\frac{(k - \mu_a)^2}{2\sigma_a^2}} \cdot P_A(a_j^o, k) \end{aligned} \quad (14)$$

其中, t_i^n 表示 ou_{new} 到达 a_i^n 的时间, $P_A(a_j^o, k)$ 表示 ou_{old} 在其历史轨迹数据中 k 时间出现在 a_j^o 的概率, 通过高斯(正态)分布的概率密度函数估计权重 $w(k)$. 用同样的方法计算出发时间相似度 $\text{Sim}_t^D(a_i^n, a_j^o)$ 和停留时间相似度 $\text{Sim}_t^S(a_i^n, a_j^o)$. 因此, a_i^n 和 a_j^o 的时间相似性表示为

$$\text{Sim}_t(a_i^n, a_j^o) = \text{Sim}_t^A(a_i^n, a_j^o) \cdot \text{Sim}_t^D(a_i^n, a_j^o) \cdot \text{Sim}_t^S(a_i^n, a_j^o) \quad (15)$$

最后, 根据新旧参与者之间的语义相似性和时间相似性可得出 ou_{new} 和 ou_{old} 的相似性定义如下:

$$\begin{aligned} \text{Sim}(ou_{new}, ou_{old}) &= \frac{\max \left(\prod_{a_i^n \in A^n} \text{Sim}(ST^n, ST^o) \cdot \text{Sim}_t(a_i^n, a_j^o) \right)}{Z} \quad (16) \\ &= \frac{\max_{a_j^o \in A^o} \left(\prod_{a_i^n \in A^n} \text{Sim}(ST^n, ST^o) \cdot \text{Sim}_t(a_i^n, a_j^o) \right)}{Z} \end{aligned}$$

其中, $\text{Sim}(ST^n, ST^o)$ 和 $\text{Sim}_t(a_i^n, a_j^o)$ 分别表示停留区域 a_i^n 和 a_j^o 的语义和时间相似性. Z 是规范化因子, 可表示为 $Z = \sum \left(\prod_{a_i^n \in A^n} \text{Sim}(ST^n, ST^o) \cdot \text{Sim}_t(a_i^n, a_j^o) \right)$, 用于确保相似度的值在 0~1.

4.5 基于迁移学习的移动性预测模型

根据相似性度量, 可以找到与新参与者轨迹模式最相似的前 k 名参与者. 平台可以使用这些旧参与者的数据作为新参与者的辅助数据.

本文采用半马尔可夫模型来捕捉参与者在不同单元格间的移动性, 并将这一过程视为状态间的转换. 该模型能够根据用户的轨迹数据, 准确预测用户在特定时间单元内到达特定区域的概率. 本文将用户能否按时完成感知任务的问题转化为一个概率问题, 从而提供一种量化方式来评估任务完成的可能性. 因此, 利用半马尔可夫模型来预测参与者 ou_i 在时间 S 内从当前区域 l_i 移动到 l_j 的概率表示为

$$\begin{aligned} Z_{ou_i}(l_i, l_j, S) &= P(L_{ou_i}^{n+1} = l_j, s_{ou_i}^{n+1} - s_{ou_i}^n \leq S | s_{ou_i}^0, \dots, s_{ou_i}^n; s_{ou_i}^0, \dots, s_{ou_i}^n) \quad (17) \\ &= P(L_{ou_i}^{n+1} = l_j, s_{ou_i}^{n+1} - s_{ou_i}^n \leq S | L_{ou_i}^n = l_i) \end{aligned}$$

其中, $L_{ou_i}^{n+1}$ 是参与者可移动区域序列, s_{ou_i} 是相应到达时间. 对于旧参与者, 可以根据历史轨迹数据计算 $Z_{ou_i}(l_i, l_j, S)$. 对于新参与者, 使用源数据和辅助数据来获得 $Z_{ou_i}(l_i, l_j, S)$. 则 ou_i 在 S 时从 l_i 移动到 l_j 的概率为

$$Q_{ou_i}(l_i, l_j, S) = \begin{cases} \sum_{l_k=1}^L \sum_{s=1}^S (Z_{ou_i}(l_i, l_k, s) - Z_{ou_i}(l_i, l_k, s-1)) \\ \cdot Q_{ou_i}(l_k, l_j, S-s), l_i \neq l_j \\ 1 - \sum_{l_k=1, l_k \neq l_i}^L Z_{ou_i}(l_i, l_k, S) \\ + \sum_{l_k=1, l_k \neq l_i, s=1}^L \sum_{s=1}^S (Z_{ou_i}(l_i, l_k, s) - Z_{ou_i}(l_i, l_k, s-1)) \\ Q_{ou_i}(l_k, l_j, S-s), l_i = l_j \end{cases} \quad (18)$$

用户 ou_i 在时间 S 内完成区域 l_j 的感知任务 t_j 的概率为

$$P(ou_i, t_j) = 1 - \prod_{s=1}^S (1 - Q_{ou_i}(l_i, l_j, s) \cdot p_{ou_i}) \quad (19)$$

本文定义 p_{ou_i} 为用户 ou_i 成功执行任务的概率, 为了简化问题, 定义 $p_{ou_i} = 1$. 也就是说, 如果选择用户中的一个在适当时间到达相应的位置, 那么任务将成功执行. 被选中的用户集合 OU 完成区域 l_j 感知任务 t_j 的期望概率为

$$E(t_j, OU) = 1 - \prod_{ou_i \in OU} (1 - P(ou_i, t_j)) \quad (20)$$

则增加一个候选机会主义工人的效用增加量为

$$\begin{aligned} \Delta \text{Utility}(OU \cup \{ou_j\}) &= \sum_{t_i \in T} E(t_i, OU \cup \{ou_j\}) - \sum_{t_i \in T} E(t_i, OU) \quad (21) \end{aligned}$$

4.6 基于移动性预测的贪心算法

在本文中, 提出一个贪心算法, 旨在迭代地识别最具效益的机会主义工作者. 该过程的伪代码如算法 1 所示. 算法 1 说明机会式任务分配的过程. 首先初始化招募的机会式用户 OU_p 集合、分配给机会用户的任务集合 OT' 和剩余预算 B_R . 所有参与者的频繁停留区域 LA 集合、生活模式 LM 集合. 第 3~5 行从所有参与者的历史轨迹数据中提取停留点. 然后利用基于密度的聚类算法将停留点聚类为停留区域. 第 6~11 行选择每个参与者的时间熵超过阈值 T_θ 的停留区域设置为频繁停留区域, 并将具有相同数量频繁停留区域的参与者划分为一种生活模式. 第 12~16 行提取新参与者的停留点和停留区域, 将停留区域转化为语义标签, 并且通过停留区域找到生活模式分类. 第 17~23 行, 遍历具有相同生活模式的旧参与者, 比较相似度, 并挑选出前 k 名相似的参与者. 第 24~27 行根据映射关系将 U_{top-k} 的数据传输到新的参与者中. 第 28~33 行迭代选择增量贡献最大的用户

算法 1 基于移动性预测的机会式用户任务分配

输入: 一组新机会式用户 OU_{new} 和旧机会式用户 OU_{old} ; 参与者相似度阈值 S_θ ; 时间熵阈值 T_θ ; 总预算 B ; 概率阈值 P_θ

输出: 机会式用户集合 OU ; 剩余预算 B_R ; 分配给机会用户的任务集合 OT' ; 剩余任务集合 T_R

1. 初始化 $OU_p, OT' = \emptyset; B_R = B; LA = \{LA_1, LA_2, LA_3, \dots\}; LM = \{LM_1, LM_2, LM_3, \dots\}$
2. FOR $n = 1$ TO $|OU_{old}|$ DO
3. 提取参与者 OU_{old}^n 的停留点 P_{old}^n
4. 利用聚类发现停留区域 A_{old}^n 区域 A_{old}^n
5. 将停留区域集转化为语义标签集 ST^o
6. FOREACH $a \in A_{old}^n$ DO
7. IF $TE(a) \geq T_\theta$ THEN
8. $LA_n = LA_n \cup \{a\}$
9. END FOR
10. $LM_{|LA_n|} = LM_{|LA_n|} \cup \{ou_{old}^n\}$
11. END FOR
12. FOR $n = 1$ TO $|OU_{new}|$ DO
13. 提取参与者 ou_{new}^n 的停留点 P_{new}^n
14. 利用聚类发现停留区域 A_{new}^n
15. 找到 ou_{new}^n 的生活模式类别 LM_x
16. 将停留区域集 A_{new}^n 转化为语义标签集 ST^n
17. FOREACH $OU_{old} \in LM_x$ DO
18. IF $Sim(ou_{new}^n, ou_{old}) \geq S_\theta$ THEN
19. $U_s = U_s \cup \{op_{old}\}$
20. 记录停留区域 ou_{new}^n 和 ou_{old} 之间的映射关系
21. END IF
22. $U_{top-k} = \{u \in U_s | u \text{ 是相似度最高的 } k \text{ 个元素}\}$
23. END FOR
24. FOREACH $u \in U_{top-k}$ DO
25. 将 u 的轨迹数据迁移到 ou_{new}^n 中
26. END FOR
27. END FOR
28. WHILE $B_R > 0$ DO
29. $OU' = \emptyset$
30. FOREACH $ou_j \in (OU_{new} + OU_{old}) \setminus OU'$ DO
31. $OU_{max} = \arg \max(\Delta Utility(OU' \cup \{ou_j\}))$
32. END FOR
33. $OU' = OU' \cup \{OU_{max}\}$
34. $OT = \emptyset$
35. FOREACH $t_j \in T$ DO
36. IF $E(t_j, OU') \geq P_\theta$ THEN
37. $OT = OT \cup \{t_j\}$
38. END IF
39. END FOR
40. $OU_p = OU_p \cup OU'; B_R = B - |OU_p| * C_{ou}$
41. $OT' = OT' \cup OT; T_R = T - OT'$
42. END WHILE
43. RETURN OU_p, B_R, T_R, OT'

OU_{max} , 第34~39行计算 OU_{max} 完成的任务. 第40~41行计算当前剩余预算 B_R 和剩余任务集合 T_R , 并且记录招募的

机会式用户集合 OU_p 和分配给机会用户的任务集合 OT' .

5 基于蚁群优化算法的参与式任务分配

本文在机会式用户任务分配 (Opportunistic User Task Assignment, OUTA) 的基础上将剩余的任务分配给参与式用户, 并且在满足预算限制的情况下最大限度地提高总体任务质量.

5.1 任务区域和用户绑定

为了让参与式用户完成任务时最大限度地减少行程距离, 本文提出的任务分配算法综合考虑任务的地理位置和任务质量. K -means++ 聚类算法^[33] 能够有效地将大量数据点划分为预定数量的簇, 使得同一簇内的数据点尽可能相似, 因此本文使用 K -means++ 聚类算法将剩余任务根据其距离和任务与用户之间的距离, 将整个感知地图划分为和参与式用户数量相等的区域, 设 $RA = \{RA_1, RA_2, \dots, RA_{n_2}\}$ 表示剩余区域, RA_{ik} ($1 \leq k \leq n_2$) 表示区域 RA_k 中的任务集合, n_2 是待招募参与式用户的数量. 为了将参与式用户与任务区域关联起来, 定义一个新的绑定效用函数确保每个用户都能被分配到最符合资源的任务, 从而最大化整个系统的运行效率. 绑定效用函数定义如下:

$$B(pu_i, RA_k) = \frac{\text{Priority}(RA_k)}{\text{Dist}(pu_i, RA_k)}, \quad \forall 1 \leq k \leq n_2 \quad (22)$$

其中, $\text{Dist}(pu_i, RA_k)$ 为用户 pu_i 当前位置到区域 RA_k 中心的距离, $\text{Priority}(RA_k)$ 为待选区域 RA_k 的优先级. 可表示为

$$\begin{aligned} \text{Priority}(RA_k) &= \frac{\text{Num}(RA_{ik})}{\pi \left(\frac{\sum_{t_j \in RA_{ik}} \text{Dist}(t_j, RA_k)}{\text{Num}(RA_{ik})} \right)^2} \cdot \frac{\sum_{t_j \in RA_{ik}} q_j}{\text{Num}(RA_{ik})} \\ &= \frac{\sum_{t_j \in RA_{ik}} q_j}{\pi \left(\frac{\sum_{t_j \in RA_{ik}} \text{Dist}(t_j, RA_k)}{\text{Num}(RA_{ik})} \right)^2} \end{aligned} \quad (23)$$

其中, $\text{Num}(RA_{ik})$ 为区域 RA_k 的任务数量,

$\frac{\sum_{t_j \in RA_{ik}} \text{Dist}(t_j, RA_k)}{\text{Num}(RA_{ik})}$ 为区域 RA_k 的半径, q_j 为任务完成质量.

用 b_{ij} 表示区域 RA_j 和用户 pu_i 之间的绑定关系, 如果区域 RA_j 和用户 pu_i 绑定则 $b_{ij} = 1$, 否则 $b_{ij} = 0$. 本文目的是将参与式用户和区域绑定起来, 以使效用总和最

大化. 最大化总效用的用户-区域绑定问题表述为

$$\max \sum_{i=1}^{n_2} \sum_{j=1}^{n_2} b_{ij} \times B(\text{pu}_i, \text{RA}_k) \quad (24)$$

$$\text{s.t.} \sum_{i=1}^{n_2} b_{ij} = \sum_{j=1}^{n_2} b_{ij} = 1 \quad (25)$$

$$\text{var. } b_{ij} \in \{0, 1\}, \forall \text{pu}_i \in \text{PU}, \text{RA}_j \in \text{RA} \quad (26)$$

目标函数式(24)通过用户和区域之间的绑定,实现总效用的最大化. 约束式(25)确保将一个用户和一个区域绑定在一起. 在解决这个优化问题时,将用户区域绑定问题转化为一个二部图匹配问题. 通过创建一个二部图,每个用户和每个区域都被视为一个顶点. 每对用户和区域通过一条边相连,边的权重基于效用值式(22)计算.

算法2给出了用户和任务区域绑定的过程. 第1行初始化剩余区域. 第2行根据任务位置的远近程度对剩余任务进行聚类并更新剩余区域. 第3~4行使用二部图匹配进行用户和区域绑定,并计算用户和区域之间每条边的权重. 第5~11行更新决策变量 $b_{ij}, \forall \text{pp}_i \in \text{PP}, A_j \in A$.

算法2 基于聚类的用户和任务区域绑定算法

输入:一组参与式参与者PU; 剩余任务 T_R

输出: $b_{ij}, \text{pu}_i \in \text{PU}, \text{RA}_j \in \text{RA}$

1. $\text{RA} = \emptyset$
2. 利用K-means++算法对剩余任务 T_R 进行聚类,其中聚类数量等于参与式用户数量,并据此更新 RA
3. 构建一个二部图 G , 顶点包括参与式用户 $\text{pu}_i \in \text{PU}$ 和区域 $\text{RA}_j \in \text{RA}$, 并根据式(22)计算每条边的权重,表示为相应的效用值
4. 利用Kuhn-Munkres算法匹配 $\text{pu}_i \in \text{PU}$ 和区域 $\text{RA}_j \in \text{RA}, \forall 1 \leq i, j \leq n_2$
5. FOREACH $\text{pu}_i \in \text{PU}$ DO
6. IF 用户 pu_i 被绑定到区域 RA_j
7. $b_{ij} = 1$
8. ELSE
9. $b_{ij} = 0$
10. END IF
11. END FOR
12. RETURN $b_{ij}, \forall 1 \leq i, j \leq n_2$

5.2 基于行程距离平衡的蚁群优化算法

用户和区域绑定策略有效地将参与式用户分配到了特定的任务区域. 然而,由于本文研究的是时间敏感任务,这些任务需要以最快速度完成,因此在特定区域内高效地规划任务完成路径成为一个关键问题. 为了

应对这一问题,本文采用ACOTDB^[34],该算法通过模拟蚂蚁寻找食物的行为来寻找最短或最优路径,能够自适应地调整搜索策略,有效避免局部最优解,在蚁群优化算法的基础上本文引入了行程距离平衡度量^[35],提出基于行程距离平衡的ACOTDB. 该算法考虑了任务对用户整体行程距离的影响,优化用户的总行程和分配区域内的任务执行路径,确保能够以最小的行程成本高效完成任务.

ACOTDB是一种模拟蚂蚁觅食行为的启发式算法,已成功应用于路径搜索问题. 起初蚂蚁是随机移动的,一组蚂蚁通过遵循特定规则获得到达目的地的路径,一旦蚂蚁完成其路径,就会沿着路径留下信息素痕迹. 由于信息素的存在,蚂蚁可以协作地找到旅行路径的合理解决方案,提高迭代解的性能.

信息素是蚂蚁在觅食过程中沿途留下的化学信号,用于指引其他蚂蚁找到食物来源的路径. 信息素更新如下:

$$\tau_{\text{tail},j} = (1 - \rho) \cdot \tau_{\text{tail},j} + \rho \cdot \tau_0 \quad (27)$$

其中, $\tau_{\text{tail},j}$ 为连接当前路径尾部和任务 t_j 弧上的信息素浓度, ρ 为蒸发率 ($0 \leq \rho \leq 1$), τ_0 为初始信息素浓度. 启发式信息 η_{ijk} 是一种优先权度量,作为蚂蚁从当前位置访问候选任务 t_j 吸引力的指标. 启发信息如下:

$$\eta_{\text{tail},j} = \frac{\Delta A_{t_j}}{\text{Dist}(\text{tail}, t_j)} \quad (28)$$

其中, ΔA_{t_j} 表示完成任务 t_j 后的质量增量,偏好函数用于计算蚂蚁选择特定路径或任务的偏好程度,是一个基于启发式信息和信息素浓度的数学表达式,用于评估每个选择的吸引力,一般使用的偏好函数如下:

$$\text{Pre}(t_j) = \tau_{\text{tail},j}^\epsilon \cdot \left(\frac{\Delta A_{t_j}}{\text{Dist}(\text{tail}, t_j)} \right)^\beta \quad (29)$$

其中, ϵ 和 β 表示加权参数. 但是上述偏好函数只通过考虑任务质量或者任务密度,从而选择高效益和低成本的任务,却都没有考虑出行距离平衡对任务的影响. 如图3所示,有2个候选任务1和任务2需要考虑. 这2种候选算法具有相同的任务质量增量和相同的额外行程距离,但是其实候选任务1到用户目的地的距离比候选任务2要长得多. 因此如果选择候选任务1,则用户必须花费更远的行程距离才能最终到达用户目的地,那么访问其他任务位置的行程距离预算将大大减少. 综上,在类似的情况下应该给候选任务2更高的优先级.

因此,在上述偏好函数的基础上引入行程距离平衡对任务的影响,在本文中偏好函数定义如下:

$$U(t_j) = \tau_{\text{tail},j}^\epsilon \cdot \frac{\Delta A_{t_j} \cdot \left(\text{TB}_i - \text{Dist}_i^k - \text{Dist}(\text{tail}, j) - \text{Dist}(t_j, \text{des}_i) \right)^\beta}{\text{Dist}(\text{tail}, j)^\beta} \quad (30)$$

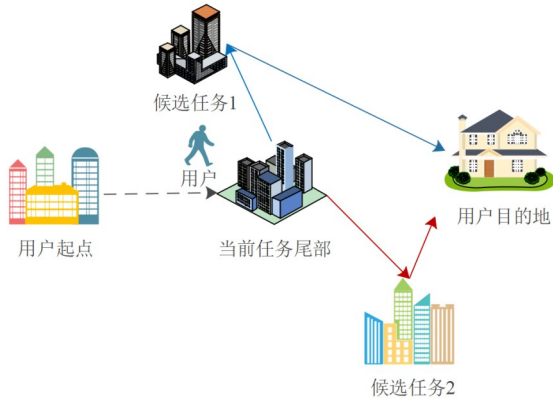


图3 行程距离示例图

其中, ΔA_i 表示完成任务 t_j 后的质量增量, TB_i 表示用户 pu_i 的出行预算. $Dist_i^t$ 表示用户 pu_i 当前完成任务所行走的路径距离. $Dist(t_j, des_i)$ 表示从任务 t_j 位置到用户 pu_i 目的地的距离. $TB_i - Dist_i^t - Dist(tail, j) - Dist(t_j, des_i)$ 表示如果用户在完成当前任务 t_j 后直接前往最终目的地, 到目前为止的剩余行程距离. η 和 γ 是 2 个权重参数, 较大的 η 参数值鼓励选择会导致较大行程距离预算的任务, 较大的 γ 会更加倾向于选择距离当前位置较近的任务, 以减少行程成本和时间.

不同的候选任务有不同的概率被选为下一个任务. 偏好值最高的任务被选中的概率为 q , 其他候选任务被选中的概率为 $1 - q$, 当选择任务 t_j 作为下一个任务时, 当前路径尾部与任务 t_j 之间边缘的局部信息素 $\tau_{tail, j}$ 随蒸发速率 λ 而变化, 使得 $\tau_{tail, j} = (1 - \rho) \cdot \tau_{tail, j} + \rho \cdot \tau_0$.

算法 3 描述了 ACOTDB 的过程. 第 1 行初始化全局信息素. 第 2 行将这一轮的局部信息素初始化为全局信息素. 第 3 行开始搜索过程, 第 4 行确保每个蚂蚁都能构建一个完整的路径. 第 5~10 行让每个蚂蚁选择它的下一个任务, 并且更新当前路径尾部和所选任务之间边缘的局部信息素. 第 11 行在一轮结束时更新全局信息素. 第 12 行将完成的路径添加到用户路径集合.

6 实验

本文实验的目的是比较本文提出方法和其他基线方法在不同情况下的性能.

本文使用 GeoLife 数据集^[36]和 San Francisco 数据集^[37]评估所提出算法的性能. 首先介绍了数据集和基线方法. 然后, 对提出的算法及其各自的基线进行全面的评估, 重点关注 2 个关键方面: 移动性预测和任务分配. 通过与基线方法的比较分析, 全面了解了所提出算法的相对优势和劣势. 此外, 使用任务完成质量和任务完成数量等指标来评估任务分配的性能. 最后, 使用

算法 3 基于行程距离平衡的蚁群优化算法

输入: 一组参与式参与者 PU ; 剩余预算 B_R ; 剩余区域 RA ; 区域 RA_k 内的任务集合 RA_{ik}

输出: 用户 pu_i 的行程路径 P_{pu_i}

1. 初始化全局信息素 τ
2. 将局部信息素设置为全局信息素 τ
3. FOR $pu_i \in PU$ DO
4. WHILE 用户 pu_i 未构建完整路径且剩余预算 $B_R > 0$ DO
5. 更新 pu_i 的可选任务集
6. 根据式(30)计算用户对区域内每个可选任务 RA_{ik} 的偏好值 $Pre(RA_{ik})$, 偏好值最高的任务为最佳候选任务
7. 以概率 q 选择最佳候选任务
8. 以概率 $1 - q$ 选择其他候选任务
9. 选择下一个任务并且更新当前路径到新路径的局部信息素
10. END WHILE
11. 更新用户 pu_i 构建的路径上的全局信息素
12. 将用户 pu_i 构建的路径添加到路径集合 $Path_{pu}$ 中
13. END FOR
14. RETURN $Path_{pu}$

Python 3.6 在处理器为 Intel i5-8265U、内存 16 GB 和 windows10 64 bit 的 PC 机上进行实验评估.

6.1 数据集

本文使用的第 1 个数据集是 GeoLife 轨迹数据集, 该数据集采集 182 名用户在日常生活中的轨迹, 包括时间戳、纬度、经度和海拔信息, 共包含 17 621 条轨迹, 覆盖大约 120 万 km 和 50 176 h 的持续时间. 本文从中选择了 30 个用户 365 d 的轨迹, 并且将选取的参与者轨迹分为 12 个子集, 每个子集包括 1 个月的数据. 首先, 本文使用每个月的最后 1 天来评估移动性预测算法, 同时使用该月前几天的轨迹来训练移动性预测模型. 其次, 引入 San Francisco 数据集, 该数据集包含 536 辆出租车在 1 个月内的行程轨迹, 每条记录包含车辆 ID、纬度、经度、占用情况和时间情况等. 本文从中选择 30 辆出租车 1 个月的轨迹, 与 GeoLife 数据集相似, 使用该月前几天的轨迹数据来训练移动性预测模型, 并使用最后 1 天的数据来评估移动性预测算法.

6.2 参数设置

本文分别设置了任务和用户的参数, 汇总如表 2 所示. 对于用户数量, 默认用户数量为 30, 并选择在轨迹迁移中排名前 4 位相似的老用户, 之后对不同数量的参与者进行算法性能评估. 对于每个轨迹数据集, 本文采用文献[38]的设置, 将机会式参与者的比例设置为 0.5, 使得有足够的机会式参与者和参与式参与者可供选择. 由于新机会式参与者的数量远少于旧机会式参与者, 因此从每个机会式参与者集合中随机选择 20% 的参与者作为新机会式参与者. 默认的任务总预算是

800,之后再改变总预算从500~1 000不等去评估成本对算法性能的约束. 本文将每个被招募的机会式参与者的成本设置为5,每个被招募的参与式参与者的单位距离成本设置为5. 默认的任务数量是200,任务的最大质量均匀且随机地分布在1~100,参与式参与者的速度均匀且随机地分布在1~4,默认的时间衰减因子 α 为0.02. 用户的行程预算比定义为用户的行程距离预算与用户从起点到目的地的最短距离的比值,行程预算比(≥ 1)服从正态分布,默认均值为1.5,标准差为0.6. γ 和 η 是2个权重参数,分别设置为3和1. 默认时间熵阈值 T_θ 、相似性阈值 S_θ 和概率阈值 P_θ 分别设置为0.6、0.5和0.8. 加权参数 ϵ 、蒸发速率 ρ 和蚂蚁选择其局部最优任务作为下一个任务概率 q 的默认设置分别为1、0.9和0.1.

表2 实验参数设置

参数	设置
任务数量	140,160,180,200,220,240
参与者数量	10,20,30,40,50,60
机会式参与者占比	0.5
新机会式参与者占比	0.2
机会式参与者任务奖励	5
参与式参与者单位距离成本	1
总预算	500,600,700,800,900,1 000
任务的最大完成质量	U(1,100)
用户的速度	U(1,4)
时间衰减因子 α	0.02
行程预算比	N(1.5,0.6)
γ 和 η	3,1
T_θ 、 S_θ 和 P_θ	0.6,0.5,0.8
ϵ 、 ρ 和 q	1,0.9,0.1

6.3 基线

为了验证本文提出的基于迁移学习的移动性预测模型的预测性能,与2种移动性预测模型比较如下.

(1)基于半马尔可夫的移动性预测方法(Mobility Prediction method based on Semi-Markov, MPSM). 与文献[39]类似,该模型采用半马尔可夫模型进行预测,但是新参与者仅根据自身的历史轨迹数据预测未来轨迹,并没有考虑机会式参与者的历史轨迹数据稀缺情况.

(2)基于迁移学习的移动性预测方法(Mobility Prediction method based on Transfer Learning, MPTL). 与本文方法不同,MPTL将新参与者的轨迹与所有旧参与者的轨迹依次进行比较,但是不进行生活模式分类.

在移动性预测模型的基础上,为证明本文提出的基于混合用户的任务分配效率,与如下算法进行比较.

(1)OUTA. 与文献[40]类似,该算法在预算约束的情况下迭代选择效用增量最大的用户,该基线方法用于比较混合式任务分配和单一的机会模式的有效性.

(2)参与式用户任务分配(Participatory User Task Assignment, PUTA). 与文献[35]类似,该算法仅招募参与式用户,直到总预算用完. 该基线方法用于比较混合式任务分配和单一的参与感知模式的性能.

(3)混合式感知用户任务分配(HUTAA-NO). 与本文提出的HUTAA类似,区别在于HUTAA-NO算法在机会式任务分配阶段不使用本文提出的迁移学习方法,在此之后,算法将遵循HUTAA完全相同的步骤进行参与式用户招募.

(4)基于两阶段的混合式用户任务分配(Hybrid User task allocation based on Two-Stage, HUTS). 与文献[38]类似,在离线阶段,采用机会式参与者招募算法;在线阶段采用基于感知子区聚类的参与式参与者招募算法. 该方案在有限的预算约束下,通过混合用户选择实现传感覆盖的最大化.

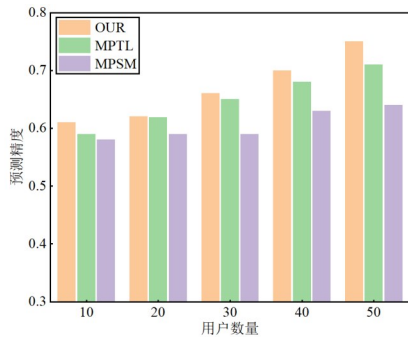
6.4 移动性预测模型性能

图4展示了3种移动性预测模型在Geolife和San Francisco数据集上的预测精度. 如图4所示,在不同用户数量的情况下,本文提出的移动性预测模型在2个数据集上的预测表现均优于对比方法. 此外,随着机会式用户数量的增加,本文提出的方法和MPTL的预测精度显著增加,是因为这2种方法考虑了历史轨迹数据稀疏的新参与者,参与者数量增加意味着有足够多的旧参与者充当数据源从而迁移轨迹数据,显著提升了对历史轨迹数据不足的新参与者的预测准确性.

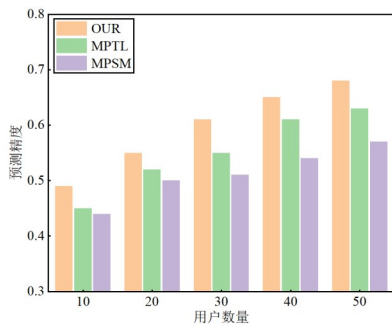
图5展示了在Geolife和San Francisco数据集上,本文提出的移动性预测模型与MPSM方法在新参与者拥有不同的历史轨迹数据天数时对预测精度的影响. 如图5所示,随着历史轨迹天数的增加,2种算法均展现出预测精度的增加趋势. 此外,当新参与者有 ≤ 7 d的历史轨迹天数时,本文方法比MPSM具有更准确的预测结果. 这表明在新参与者历史数据稀缺的情况下,本文提出的基于迁移学习的移动性预测模型能够有效提高新机会式参与者的预测准确性.

图6展示了Geolife和San Francisco数据集中参与者的生活模式分布. 本文将参与者按频繁停留区域数量分为4类:生活模式1~生活模式3分别代表参与者具有1~3个频繁停留区域,生活模式4表示有4个或更多的频繁停留区域. 在Geolife数据集中,生活模式1和生活模式2的参与者较多,这表明该数据集中的参与者倾向于在较少的区域频繁活动. 而在San Francisco数据集中,生活模式3和生活模式4的参与者占据了较大比例,是因为San Francisco数据集主要采集经常需要前往多个目的地的司机,因此他们的活动模式更为复杂,涉及更多的停留区域.

图7展示了本文提出的预测方法在不同时间熵阈值下的性能比较. 在Geolife和San Francisco数据集上,

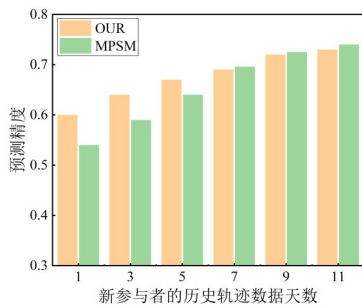


(a) Geolife

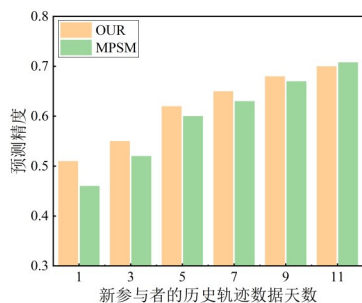


(b) San Francisco

图4 用户数量对预测精度的影响



(a) Geolife



(b) San Francisco

图5 新参与者的历史轨迹数据天数对预测精度的影响

可以观察到随着时间熵阈值的增大,预测精度呈现先上升后下降的趋势. 当时间熵阈值设置为0.6时,模型达到了最佳的预测精度.

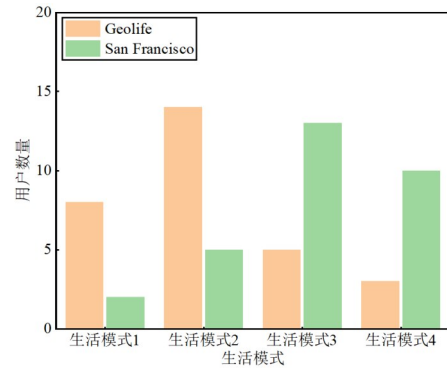


图6 生活模式分类图

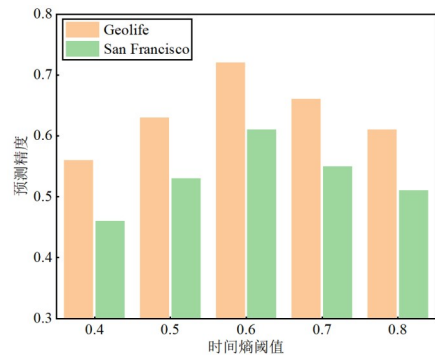


图7 时间熵阈值对预测精度的影响

图8展示了本文方法和MPTL方法在Geolife和San Francisco数据集上搜索时间的比较,即挑选出与新参与者最相似的Top-k个旧参与者所需的时间. 在2种算法中,随着旧参与者数量的增加,搜索时间均呈现增长趋势. 但是本文方法相较于MPTL方法显著地缩短了搜索时间.

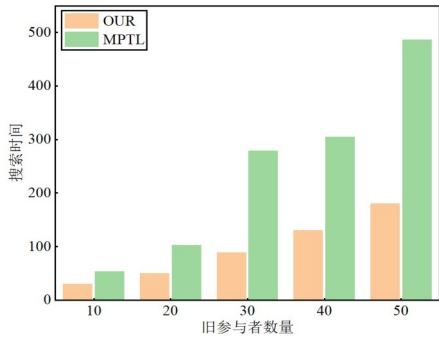
图9展示了在不同k值下移动性预测模型的预测精度. 在Geolife和San Francisco数据集上,预测精度呈现先上升后下降的趋势. 当k=4时,预测精度达到了峰值;当k>4时,额外转移的轨迹参与者与新参与者之间的相似度较低. 因此,过多的数据传输反而会降低移动性预测的性能,从而导致任务分配算法的不准确性.

6.5 任务分配性能

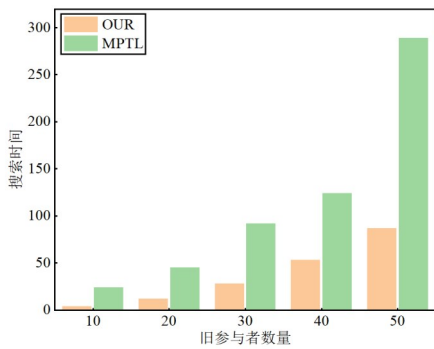
首先用不同的参数设置评估本文提出算法的性能,然后将其与不同的算法进行比较以进行性能评估.

6.5.1 权重参数η和γ的影响

图10展示了在Geolife和San Francisco数据集上,权重参数η和γ对任务完成质量的影响,这2个参数分别用于衡量行程剩余距离和路径末端到候选任务的距离. 如图10(a)所示,当η从0增加到1时,任务完成质量迅速提高,并在η=1时达到最优状态. η值的进一步增加会导致任务完成质量的轻微下降. 如图10(b)所示,随着γ的增加,任务完成质量呈现上升趋势,并在γ=3时达到峰值. 因此将参数η设置为1,参数γ设置为3更为合适,能够达到任务完成质量的最优平衡.



(a) Geolife



(b) San Francisco

图8 搜索时间图

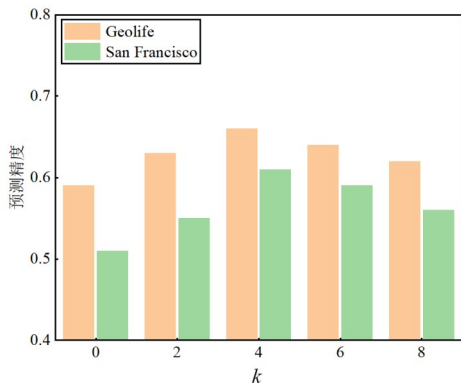
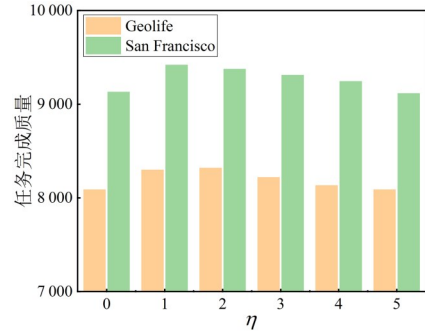


图9 Top-k值对预测精度的影响

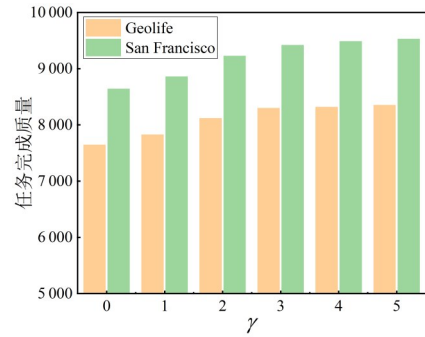
6.5.2 时间衰减因子 α 的影响

图 11 和图 12 分别展示了在 Geolife 和 San Francisco 数据集上,不同时间衰减因子 α 对各种算法在任务完成质量和完成数量方面的性能. 如图 11 所示,任务完成质量随 α 的增大而降低. 这是因为较大的衰减因子意味着任务完成质量随时间的下降速度加快,然而,图 11(b) 相较于图 11(a) 的任务完成质量下降趋势相对平缓,可能是因为图 11(b) 的数据集来源于出租车,其较快的速度缩短了到达任务位置的时间,从而减缓了质量的衰减. 图 12 表明, HUTAA 在任务完成数量方面也表现出最佳性能. 综上, HUTAA 在任务完成的质量和数量上均优于

其他模拟算法.

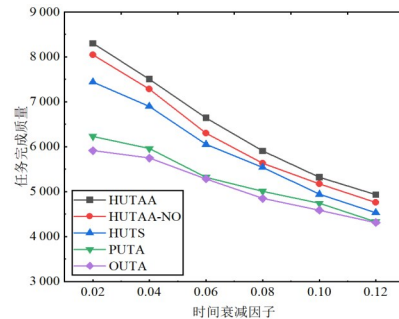


(a) 权重参数为 η

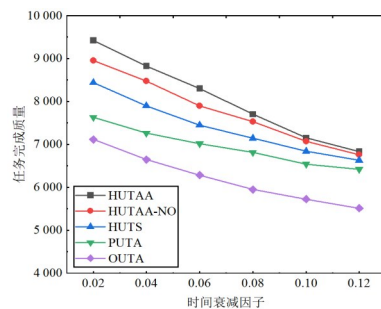


(b) 权重参数为 γ

图10 权重参数 η 和 γ 对任务完成质量的影响

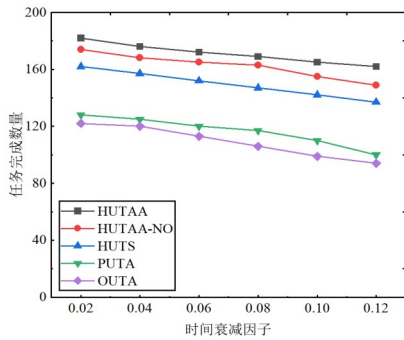


(a) Geolife

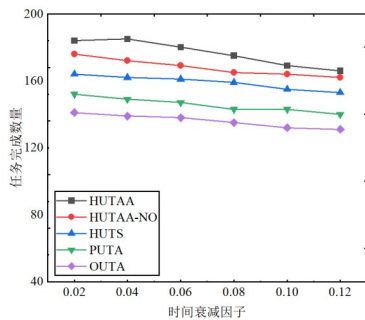


(b) San Francisco

图11 不同时间衰减因子下的任务完成质量

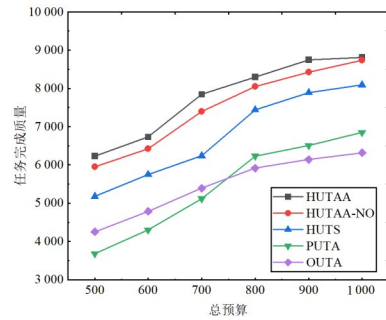


(a) Geolife

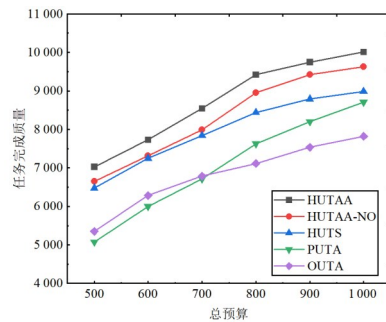


(b) San Francisco

图 12 不同时间衰减因子下的任务完成数量



(a) Geolife



(b) San Francisco

图 13 不同总预算下的任务完成质量

6.5.3 总预算的影响

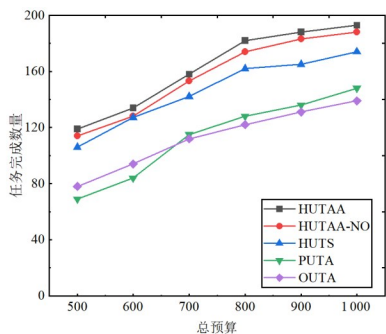
图 13 和图 14 分别展示了在不同数据集上,不同预算对算法在任务完成质量和任务完成数量方面的影响. 如图 13 和图 14 所示,随着总预算的增加,所有算法的任务完成质量和任务完成数量都在增加,表明更高的预算有助于提高任务完成质量. 综上,HUTAA 在不同预算下,无论是任务完成质量还是任务完成数量,都表现出最优的性能. 因此,在所有的仿真算法中 HUTAA 的性能是最好的.

6.5.4 用户数量的影响

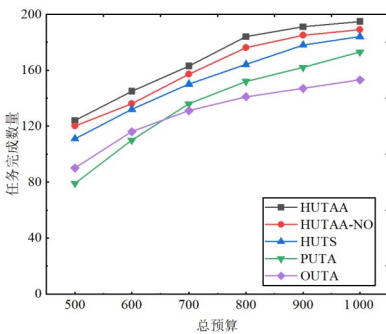
图 15 和图 16 分别展示了在不同数据集上用户数量对算法性能的影响. 控制其他变量,确保参与式和机会式用户的数量保持一致. 如图 15 所示,随着用户数量的增加,任务完成质量持续提高. 同样,在所有算法中 HUTAA 表现最好. 如图 16 所示,随着用户数量的增加,HUTAA 在任务完成数量方面也表现最佳. 最终,所有算法的性能曲线相对趋于稳定,是因为尽管用户数量增加,但总预算和任务数量保持不变,从而限制了任务完成数量.

6.5.5 任务数量的影响

图 17 和图 18 分别在不同数据集上,比较了不同任务数量对不同算法任务完成质量和任务完成数量的影响. 如图 17 所示,随着任务数量的增加,各算法任务完成质量均有所提高. 如图 18 所示,随着任务数量的增加,所有算法任务完成数量都在增加. 这是因为更多的任务提供了更多的机会给用户完成任务. 但是由于总预算的限制,所有算法曲线在最后趋向于平缓,表明任务完成率随着任务数量的增加而略有下降. 综上,可以得出 HUTAA 在所有任务数量设置下都优于其他对比算法.



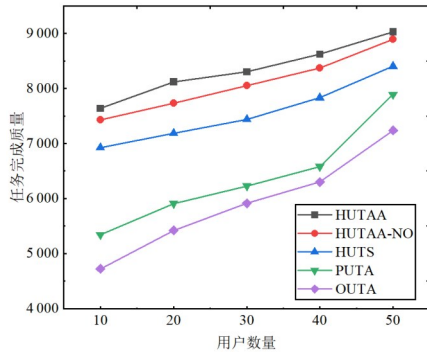
(a) Geolife



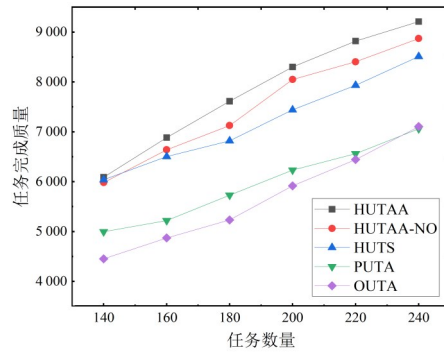
(b) San Francisco

图 14 不同总预算下的任务完成数量

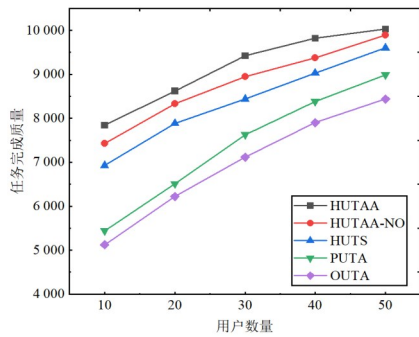
制,所有算法曲线在最后趋向于平缓,表明任务完成率随着任务数量的增加而略有下降. 综上,可以得出 HUTAA 在所有任务数量设置下都优于其他对比算法.



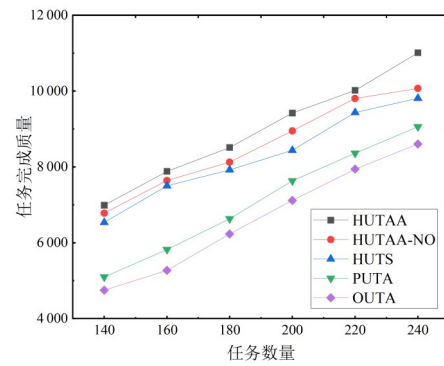
(a) Geolife



(a) Geolife



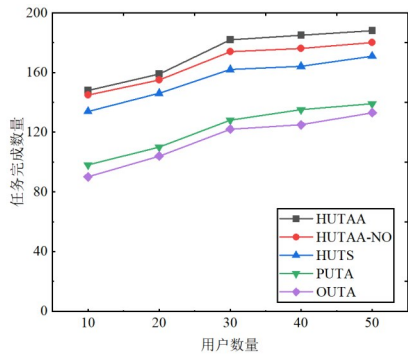
(b) San Francisco



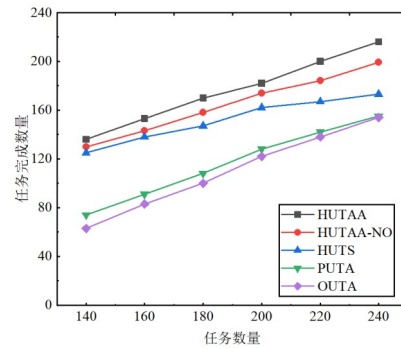
(b) San Francisco

图 15 不同用户数量下的任务完成质量

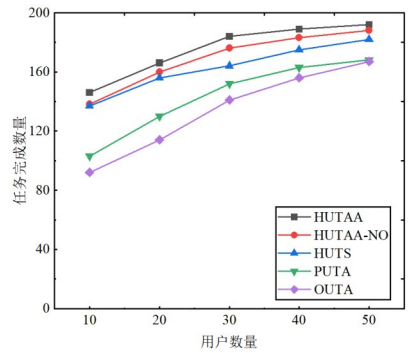
图 17 不同任务数量下的任务完成质量



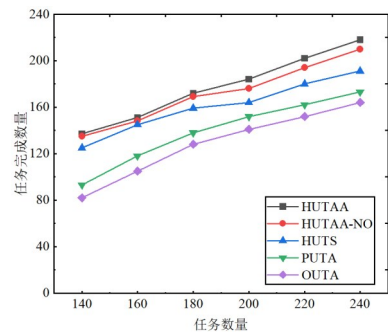
(a) Geolife



(a) Geolife



(b) San Francisco



(b) San Francisco

图 16 不同用户数量下的任务完成数量

图 18 不同任务数量下的任务完成数量

7 结论

针对异构质量的延迟敏感任务分配问题,本文在共享总预算的条件下提出一种混合机会式用户和参与式用户的任务分配算法,旨在最大限度地提高任务完成质量. 本文要求机会式用户在日常工作中完成 MCS 任务,因此设计了一种基于迁移学习的移动性预测模型,此外,本文在行程距离预算下使用 ACOTDB 规划参与式用户的完成任务路径. 然而,迁移学习在特殊情况下也可能存在一定的局限性,尤其是当用户行为模式存在显著差异时,可能会引发负迁移现象,即源数据的知识可能对目标数据的学习产生负面效果. 此外,如果目标数据集的质量不佳,也可能影响移动性预测模型的性能,使其无法达到预期水平. 针对这些问题,本文计划采取进一步的措施来改进迁移学习方法,以确保该模型在各种情况下都能提供准确的预测,并有效地支持任务分配决策.

参考文献

- [1] GANTI R K, YE F, LEI H. Mobile crowdsensing: Current state and future challenges[J]. *IEEE Communications Magazine*, 2011, 49(11): 32-39.
- [2] LIU Y. Crowdsensing computing[J]. *Communications of the Chinese Computer Society*, 2012, 8(10): 38-41.
- [3] 刘文彬, 杨永健, 王恩. 合作性移动群智感知中具有一般效用和成本的用户招募方法[J]. *计算机学报*, 2022, 45(12): 2576-2591.
LIU W B, YANG Y J, WANG E. User recruitment with general utility and cost for cooperative mobile CrowdSensing[J]. *Chinese Journal of Computers*, 2022, 45(12): 2576-2591. (in Chinese)
- [4] ZHANG Y Q, LI C, LI K H, et al. High-resolution large-scale urban traffic speed estimation with multi-source crowd sensing data[J]. *IEEE Transactions on Vehicular Technology*, 2024, 73(9): 12345-12357.
- [5] EL HAFYANI H, ABBOUD M, ZUO J W, et al. Learning the micro-environment from rich trajectories in the context of mobile crowd sensing[J]. *GeoInformatica*, 2024, 28(2): 177-220.
- [6] GUO X Y, XING W W, FANG J, et al. Noise-aware optimization for mobile crowdsensing-based travel time estimation[J]. *IEEE Transactions on Vehicular Technology*, 2024, 73(3): 4067-4080.
- [7] KIM M, KIM Y. Multi-blockchain structure for a crowdsensing-based smart parking system[J]. *Future Internet*, 2020, 12(5): 90.
- [8] ZHU W P, GUO W Z, YU Z Y. Social-aware task allocation in mobile crowd sensing[J]. *Wireless Communications and Mobile Computing*, 2020, 2020(1): 8822251.
- [9] 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. (in Chinese)
- [10] ZHANG J Y, ZHANG X L. Multi-task allocation in mobile crowd sensing with mobility prediction[J]. *IEEE Transactions on Mobile Computing*, 2023, 22(2): 1081-1094.
- [11] YANG Y J, LIU W B, WANG E, et al. A prediction-based user selection framework for heterogeneous mobile CrowdSensing[J]. *IEEE Transactions on Mobile Computing*, 2019, 18(11): 2460-2473.
- [12] LIU W B, YANG Y J, WANG E, et al. Dynamic user recruitment with truthful pricing for mobile CrowdSensing[C]//*IEEE INFOCOM 2020-IEEE Conference on Computer Communications*. Piscataway: IEEE, 2020: 1113-1122.
- [13] NASSER R, ABOULHOSN Z, MIZOUNI R, et al. A machine learning-based framework for user recruitment in continuous mobile crowdsensing[J]. *Ad Hoc Networks*, 2023, 145: 103175.
- [14] WU Y F, SUO Y N, YU F X, et al. A utility-based subcontract method for sensing task in mobile crowd sensing[J]. *IEEE Transactions on Industrial Informatics*, 2022, 18(2): 1210-1219.
- [15] DING Y, ZHANG L C, GUO L J. Dynamic delayed-decision task assignment under spatial-temporal constraints in mobile crowdsensing[J]. *IEEE Transactions on Network Science and Engineering*, 2022, 9(4): 2418-2431.
- [16] SONG S W, LIU Z D, LI Z J, et al. Coverage-oriented task assignment for mobile crowdsensing[J]. *IEEE Internet of Things Journal*, 2020, 7(8): 7407-7418.
- [17] GUO S, XIA M L, XUE H Q, et al. OceanCrowd: Vessel trajectory data-based participant selection for mobile crowd sensing in ocean observation[J]. *IEEE Transactions on Sustainable Computing*, 2024, 9(6): 889-901.
- [18] PENG S, LIU K, WANG S J, et al. Time window-based online task assignment in mobile crowdsensing: Problems and algorithms[J]. *Peer-to-Peer Networking and Applications*, 2023, 16(2): 1069-1087.
- [19] 蒋伟进, 张婉清, 陈萍萍, 等. 基于IWOA群智感知中数量敏感的任务分配方法[J]. *电子学报*, 2022, 50(10): 2489-2502.
JIANG W J, ZHANG W Q, CHEN P P, et al. Quantity sensitive task allocation method based on IWOA in group intelligence perception[J]. *Acta Electronica Sinica*, 2022, 50(10): 2489-2502. (in Chinese)
- [20] YANG G S, GUO D S, WANG B Y, et al. Participant-quantity-aware online task allocation in mobile crowdsensing[J]. *IEEE Internet of Things Journal*, 2023, 10(24): 22650-22663.
- [21] 吕翊, 王燕, 崔亚平, 等. 考虑工人培养的移动群智感知任务分配机制[J]. *电子与信息学报*, 2023, 45(4): 1505-1513.
LÜ Y, WANG Y, CUI Y P, et al. Worker development-aware task allocation strategy in mobile crowd sensing[J]. *Journal of Electronics & Information Technology*, 2023,

45(4): 1505-1513. (in Chinese)

- [22] MA G Q, CHEN H L, HUANG Y, et al. Utility-based heterogeneous user recruitment of multitask in mobile crowdsensing[J]. *IEEE Internet of Things Journal*, 2023, 10(11): 9796-9808.
- [23] WANG L, YU Z W, ZHANG D Q, et al. Heterogeneous multi-task assignment in mobile crowdsensing using spatiotemporal correlation[J]. *IEEE Transactions on Mobile Computing*, 2019, 18(1): 84-97.
- [24] SHEN X N, XU D, SONG L Y, et al. Heterogeneous multi-project multi-task allocation in mobile crowdsensing using an ensemble fireworks algorithm[J]. *Applied Soft Computing*, 2023, 145: 110571.
- [25] WANG J T, WANG F, WANG Y S, et al. HyTasker: Hybrid task allocation in mobile crowd sensing[J]. *IEEE Transactions on Mobile Computing*, 2020, 19(3): 598-611.
- [26] WEI X H, LI Z J, REN C H, et al. HSM-SMCS: Task assignment based on hybrid sensing modes in sparse mobile crowdsensing[J]. *IEEE Internet of Things Journal*, 2023, 10(5): 4034-4048.
- [27] LV Y, CHEN X, HE P, et al. Hybrid worker selection for task coverage maximization in mobile crowdsensing[C]// 2023 IEEE Wireless Communications and Networking Conference (WCNC). Piscataway: IEEE, 2023: 1-6.
- [28] WANG J, LIU J X, ZHAO G S. Two-phased participant selection method based on partial transfer learning in mobile crowdsensing[J]. *ACM Transactions on Sensor Networks*, 2023, 19(2): 1-17.
- [29] ZHANG S W, LI Z X, LIANG W, et al. Blockchain-based hybrid reliable user selection scheme for task allocation in mobile crowd sensing[J]. *IEEE Transactions on Network Science and Engineering*, 2024, 11(6): 6494-6510.
- [30] YING J J, LEE W C, WENG T C, et al. Semantic trajectory mining for location prediction[C]//Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. New York: ACM, 2011: 34-43.
- [31] CAO X, CONG G, JENSEN C S. Mining significant semantic locations from GPS data[J]. *Proceedings of the VLDB Endowment*, 2010, 3(1/2): 1009-1020.
- [32] ZHENG Y, ZHANG L Z, MA Z X, et al. Recommending friends and locations based on individual location history[J]. *ACM Transactions on the Web*, 2011, 5(1): 1-44.
- [33] ARTHUR D, VASSILVITSKII S. K-means++: The advantages of careful seeding[J]. *Soda*. 2007, 7: 1027-1035.
- [34] DORIGO M, GAMBARDELLA L M. Ant colony system: A cooperative learning approach to the traveling salesman problem[J]. *IEEE Transactions on Evolutionary Computation*, 1997, 1(1): 53-66.
- [35] GONG W, ZHANG B X, LI C. Location-based online task assignment and path planning for mobile crowdsensing[J]. *IEEE Transactions on Vehicular Technology*, 2019, 68(2): 1772-1783.
- [36] ZHENG Y, ZHANG L Z, XIE X, et al. Mining interesting locations and travel sequences from GPS trajectories[C]// Proceedings of the 18th International Conference on World Wide Web. New York: ACM, 2009: 791-800.
- [37] MICHAL P, NATASA S D, MATTHIAS G. Crowdad dataset epfl/mobility[EB/OL]. (2009-02-24) [2024-07-04]. <https://crowdad.org/epfl/mobility/20090224>.
- [38] LU A Q, ZHU J H. Worker recruitment with cost and time constraints in mobile CrowdSensing[J]. *Future Generation Computer Systems*, 2020, 112: 819-831.
- [39] LIU W B, YANG Y J, WANG E, et al. Dynamic online user recruitment with (non-) submodular utility in mobile CrowdSensing[J]. *IEEE/ACM Transactions on Networking*, 2021, 29(5): 2156-2169.
- [40] LIU Y N, LIU X, LI X, et al. Participants recruitment for coverage maximization by mobility predicting in mobile CrowdSensing[J]. *China Communications*, 2023, 20(8): 163-176.

作者简介



蒋伟进 男,1964年7月出生于湖南省益阳市. 现为湖南工商大学计算机学院二级教授、硕士生导师. 主要研究方向为群智感知、联邦学习、边缘计算.

E-mail: jwjnudt@163.com



聂彩燕 女,2001年5月出生于湖南省株洲市. 现为湖南工商大学计算机学院硕士研究生. 主要研究方向为群智感知、隐私保护.

E-mail: 516278890@qq.com