

# 基于众包工人移动轨迹的任务推荐模型



胡颖 王莹洁 童向荣

烟台大学计算机与控制工程学院 山东烟台 264005

(sd\_huying@163.com)

**摘要** 随着移动众包的发展,越来越多的任务被发布在众包平台上。然而,由于移动众包系统中有大量的任务,当众包工人根据自己的兴趣主动选择任务时,选择出适合众包工人执行的任务会花费大量的时间。此外,由于众包工人不能完全了解众包系统中存在的所有任务的信息,他们很难选出最适合自己的任务。移动众包系统中的任务具有时间和空间特性,需要众包工人在指定的时间区间内移动到指定的区域完成任务。而众包工人有自己的工作和生活,为了适应众包工人的日常移动规律,提出了一种移动预测模型来预测众包工人的移动行为。基于预测结果和众包工人的需求,提出了一种基于众包工人移动轨迹的任务推荐模型。文章在两个真实数据集上进行了大量的仿真,结果证明所提模型具有较高的准确性和良好的适应性。

**关键词:** 移动众包; 移动预测; 众包工人; 任务推荐; 移动轨迹

**中图法分类号** TP391

## Task Recommendation Model Based on Crowd Worker's Movement Trajectory

HU Ying, WANG Ying-jie and TONG Xiang-rong

School of Computer and Control Engineering, Yantai University, Yantai, Shandong 264005, China

**Abstract** With the development of mobile crowdsourcing, more and more tasks are published on crowdsourcing platforms. However, crowd workers choose tasks suitable for them will take a lot of time according to their interests, because there are a large number of tasks in the mobile crowdsourcing system. In addition, it is difficult for them to select the tasks that are most suitable for their own execution, because the crowd workers have no knowledge of the information of all tasks existing in the crowdsourcing system. The tasks in the mobile crowdsourcing system have the spatio-temporal characteristic, which requires crowd workers to move to the specified region to complete the task within the specified time interval. However, crowd workers have their own works and life, in order to adapt to their daily movement, a mobile prediction model is proposed to predict the movement behavior of them. Based on the prediction results and the needs of crowd workers, a task recommendation model based on the movement trajectory of crowd workers is proposed to recommend tasks for crowd workers. Finally, a lot of simulations are carried out on two real data sets. The results prove that the proposed model has high accuracy and good adaptability.

**Keywords** Mobile crowdsourcing, Mobile prediction, Crowd workers, Task recommendation, Movement trajectory

## 1 引言

群智感知<sup>[1]</sup>是数据获取的一种新模式,它利用人携带的各种移动设备,依靠其内置传感器对周围环境进行感知,进而达到收集感知数据的目的。移动众包是实现群智感知的一种方式,它利用群体力量来解决问题。由于只需要携带智能设备就能执行任务,每个智能设备的携带者都可以成为众包工人。随着科学技术的发展以及智能设备的普及,几乎有人在的地方就有智能设备用户,因此移动众包有大量的参与者以及潜在参与者。目前,移动设备配备的传感器种类越来

越多,可以实现的功能也越来越多,因此可以执行多种多样的任务。目前国内比较流行的众包 APP 有美团众包、蜂鸟众包、京东众包、阿里众包、蚂蚁众包等。各种众包 APP 中有各种异质任务,包括语音、图片收集、外卖、快递配送等。通过对移动众包的多方面研究<sup>[2-5]</sup>,可以激励用户积极地参与众包任务的执行,也可以实现任务的合理分配以及对众包工人任务执行路径的合理规划,最大化平台和众包工人的效用。

移动众包可以为人们提供诸多便利,但是移动众包应用的正常运行需要足够多的参与者使用其随身携带的智能设备(如智能手机、IPad、智能手环等)来执行众包任务。为了能够

收到稿日期:2020-04-30 返修日期:2020-08-09 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家自然科学基金(61822602,61772207,61802331,61702439,61773331);中国博士后基金(2019T120732,2017M622691);烟台大学研究生科技创新基金资助(YDZD1908)

This work was supported by the National Natural Science Foundation of China (61822602,61772207,61802331,61702439,61773331), China Postdoctoral Science Foundation (2019T120732,2017M622691) and Graduate Innovation Foundation of Yantai University(YDZD1908).

通信作者:王莹洁(towangyingjie@163.com)

吸引更多的用户参与众包任务,需要众包应用提供令人满意的服务,例如:保护众包工人的隐私<sup>[6-9]</sup>、节约能源<sup>[10]</sup>、最大化众包工人的效用并设计公平的机制进行任务分配。目前,众包系统中的任务分配方式分为系统主动分配和众包工人自主选择<sup>[11-12]</sup>两种。系统主动分配的方式会忽略众包工人的兴趣,有可能导致众包工人被分配到自己不愿意执行的任务,从而影响任务的完成质量,因此一些关于众包任务分配的研究通过考虑任务和众包工人的需求来进行任务分配。在众包工人自主选择任务的研究中,大多数都是设定众包工人从一些任务中选择感兴趣的任務提交到众包平台<sup>[13]</sup>,由于众包工人不能完全了解众包系统中所有任务的信息,因此会导致以下情况:

1) 众包工人选择了自己感兴趣的任務,但是该众包工人并不是最适合执行该任务的众包工人,从而可能导致任务完成质量不高。

2) 众包工人选择了自己感兴趣的任務,但是该任务要求高,众包工人没有竞聘成功,长期的竞聘失败导致众包工人丧失信心退出众包系统。

3) 众包工人选择的任務要求很低,但是众包工人的能力完全可以做价值更高的任务。但是由于系统中发布的任务太多,而众包工人很难发现与自己能力匹配的任务。这样导致了众包工人才能的浪费,且一些对能力要求高的任务也会因找不到合适的众包工人来执行而长时间等待。

目前,为了提高服务质量,各种基于服务的研究<sup>[14-15]</sup>中都进行了关于推荐的研究。为了解决上述问题,移动众包研究领域的一些研究者对任务推荐进行了研究。但现有的任务推荐研究中,大多数侧重于考虑众包工人的兴趣,通过任务和用户的相似度计算来进行任务的推荐,很少有研究考虑众包工人的时间和地理位置因素的影响。与传统众包只要在线上就可以完成任务不同,移动众包需要众包工人在线上竞拍任务,然后根据任务的要求,在任务截止时间前移动到指定的区域去完成任务。众包工人有自己的日常生活,执行众包任务对大多数的众包工人来说是兼职工作,因此众包工人的时间和地理位置相比其他因素更能影响众包工人的任务执行意愿。为了给众包工人推荐更适合的任务,在任务推荐时除了要考虑众包工人的兴趣之外,还需要考虑他们的时间和位置<sup>[16-17]</sup>。因此我们设计了移动预测模型,在任务推荐时先通过众包工人的移动轨迹信息确定任务推荐的区域。本文具体贡献如下:

1) 设计了一个移动预测模型,通过研究众包工人的移动规律对众包工人的移动行为进行预测。考虑到众包工人的变化性(如学生毕业、上班族换工作、搬家),在轨迹预测时加入时间衰减因子来使模型忘记众包工人过去的移动行为,增强模型的适应性。

2) 通过考虑众包工人的兴趣、历史任务执行情况、技能、信誉来计算任务对众包工人的吸引力,然后通过计算请求者对众包工人的技能、信誉等要求,来计算请求者对众包工人的满意度。综合考虑这两方面计算任务和众包工人的匹配成功率,最终为众包工人推荐最适合执行的 Top- $n$  个任务。

3) 实验结果证明,本文提出的移动预测模型能够比较精

确地预测众包工人下一步要去的区域,并且个性化任务推荐方法可以为众包工人推荐更适合的任务。

## 2 相关工作

随着众包工人和任务数量的增长,如何快速地为众包工人选择适合执行的任務以减少众包工人的浏览时间,以及如何为任务选择适合的众包工人以保证数据质量,成为当前众包系统的研究热点。基于提高众包平台的服务质量、任务完成率以及保证众包工人和任务请求者的利益的目的,同时为了吸引更多的参与者参与任务的执行,一些研究者通过考虑各方需求来实现任务和众包工人的最优匹配,进行了有关任务分配和推荐的研究。主要方法包括概率矩阵分解、基于历史信息计算众包工人和任务相似度,以及其他考虑隐私<sup>[18]</sup>、可信度<sup>[19]</sup>和偏好的推荐方法。

Yuen 等<sup>[20]</sup>针对先前基于分类的任务推荐没有考虑众包系统中动态出现新员工和新任务的情况,提出了一种基于统一概率矩阵因子分解的任务推荐(TaskRec)框架,旨在为动态场景中的众包工人推荐任务。该方法不需要工人提供任务评分,评分可从工人的互动行为中推断出来。之后,他们针对之前一些研究没有考虑到真实众包场景中不同类别的任务需要具有不同技能的工人,且推荐中没有考虑到新工人在没有收到推荐的首选任务列表前是不想执行大量任务的情况,在 TaskRec 推荐框架的基础上提出了基于主动学习的概率矩阵分解方法 ActivePMFv2<sup>[21]</sup>。该模型在推荐任务时考虑了工人的偏好以及员工的历史绩效,主动为最可靠的工人选择最不确定的任务来再训练分类模型。

Baba<sup>[22]</sup>为了避免众包参与者因为竞争任务的获胜率太低而离开众包平台,在为工人推荐任务时通过分析历史任务的获胜信息、工人与任务之间的吸引力和工人的能力来预测工人竞争任务时的获胜概率,为工人推荐获胜概率高的任务;并且采用了迁移学习和基于特征的矩阵因子分解的方法来解决数据稀疏的问题。Safran 等<sup>[23]</sup>通过分析众包系统任务推荐和传统商品推荐的区别,分别从众包工人的角度和请求者的角度出发,为众包系统设计了实时的推荐算法 TOP-K-T 和 TOP-K-W,该算法通过综合计算工人的文件信息与任务类别的相似度、任务接受率和类别偏好的匹配分数,利用独特的数据结构进行映射,来为工人和请求者进行任务或工人的推荐。但该推荐算法由于考虑到耗时问题,不会在完成一项任务后实时更新匹配分数,从而不能满足工人希望推荐与其最近选择和完成任务相似的任务的需要。此后,Safran 等通过定义工人类别适宜性评分和工人任务吸引力评分来支持个性化的任务和工人推荐,并分别提出了 Top- $n$  任务推荐算法和 Top- $n$  工人推荐算法来为工人和任务进行推荐建议<sup>[24]</sup>。Wang 等<sup>[25]</sup>通过考虑用户相似度和任务相似度来推测任务推荐概率,通过用户的停留时间来判断用户的感兴趣程度,以此进行任务潜在推荐概率的预测,最后通过考虑任务的可信度来保证推荐的有效性。Pan 等<sup>[26]</sup>基于 Word2vec 深度学习计算词向量的相似度,并建立语义标签相似矩阵数据库,通过语义标签相似矩阵计算任务与工人之间的相似度,最终实现了众包任务的个性化推荐。

Guo 等<sup>[27]</sup>提出了一种细粒度的推荐机制 IntexCrowd, 首先为每个任务分配一个主题, 然后根据任务的历史记录计算用户对于特定主题的兴趣程度和专业水平, 综合考虑这两点因素为特定主题任务推荐适合的用户。该方法适用于传统的基于问答平台的众包系统, 对于需要考虑地理位置以及时间的移动众包任务则不适用。考虑到执行众包任务会导致用户隐私泄露, 降低了用户参与众包任务的积极性, Liu 等<sup>[28]</sup>和 Gong 等<sup>[29]</sup>在任务推荐中考虑到用户的隐私偏好, 提出了基于隐私的推荐机制, 为用户推荐符合他们隐私要求的任务。Zhang 等<sup>[30]</sup>提出了一种基于二元模糊语言(2-tuple fuzzy linguistic method)的任务推荐方法, 该方法通过考虑工人执行任务的能力(capability-to-complete, CTC)和接受任务推荐的可能性(possibility-to-accept, PTA)来为用户进行任务推荐, 但是随着任务和工人的数量增加, 计算工人的 CTC 和 PTA 变得困难。

以上关于众包任务推荐的研究大多适用于传统的线上完成任务的众包系统, 对于目前比较流行的需要众包工人在线上竞拍任务, 在线下到指定地点完成的移动众包任务, 由于没有考虑到任务的时空特性, 往往不能取得很好的效果。要实现大规模的群智感知, 需要大量众包参与者携带智能手机来执行任务, 因此本文的研究是面对大量普通智能手机用户而不是专门的以此为生计的众包工人(如滴滴打车专车司机、美团外卖专职外卖员), 这些参与者有自己的日常生活, 不太会改变自己的行程去任务指定的地点执行任务, 而是在不改变或者基本不改变自己行程的情况下顺便执行任务。为了实现这种情况下的任务推荐, 本文在为众包工人推荐任务时考虑了众包工人的日常移动轨迹信息。

### 3 移动预测模型

移动众包系统中众包工人是很重要的角色, 他们是一些持有智能设备的群体, 每天有自己的工作和其他的事情, 除此之外, 还有些空余时间可以执行众包系统中的任务。这些众包工人倾向于在不影响其正常生活计划的区域内执行任务。例如, 众包工人( $K, G, A, W, D, T_{net}, \lambda_{max}, B, \mu, c$ )每天都会去某个工厂上班, 如果有一个检测工厂附近区域噪声水平的任务, 则该任务对( $X, Y$ )来说就非常方便执行。为了激励智能设备持有者加入众包系统, 本文设计了一个移动预测模型。该模型通过挖掘众包工人的历史轨迹信息来预测众包工人的移动, 然后众包平台为众包工人推荐最适合的任务。

在移动众包系统中, 众包工人  $W = \{\omega_1, \omega_2, \dots, \omega_i, \dots, \omega_n\}$  携带智能设备进行日常活动, 在空闲时会登录众包系统查看是否有适合自己执行的任务。每个众包系统中的众包工人都有各自的日常活动轨迹, 由经验可知, 由于人们的生产、生活的规律性使得大多数众包工人在一天之内经过的区域和经过某一区域的时间会长期保持一定的规律。本文将众包工人  $\omega_i$  一天内经过的区域、到达时间及停留时间组成的记录称为众包工人的日常轨迹信息记录, 并将其表示为  $TR_i = \langle L_1, T_1, D_1 \rangle, \dots, \langle L_d, T_d, D_d \rangle$ , 其中,  $\langle L_d, T_d, D_d \rangle$  表示第  $d$  天的轨迹信息,  $L_d = \{l_1, l_2, \dots, l_n\}$  表示第  $d$  天到达的区域集合,  $T_d = \{t_1, t_2, \dots, t_l\}$  表示第  $d$  天到达对应区域的时间集合,

$D_d = \{dw_1, dw_2, \dots, dw_l\}$  表示第  $d$  天在对应区域的停留时间集合。

为了分析众包工人轨迹信息的规律, 将时间窗划分为一天 24 段, 时间窗的大小可根据具体应用调整。在本文中,  $tw_0$  表示 0:00-1:00 的时间区间, 以此类推。由众包工人的轨迹信息可以推测众包工人在  $tw$  区间内在区域  $l$  的概率<sup>[31]</sup>。概率用  $p_{tw,l}$  表示并由式(1)得出:

$$p_{tw,l} = \frac{\sum_{d=1,2,\dots,|TR_i|} \sum_t num_l^{tw,d} \cdot \beta_d}{\sum_{d=1,2,\dots,|TR_i|} \sum_{l \in L_d^{tw}} num_l^{tw,d} \cdot \beta_d} \quad (1)$$

其中,  $L_d^{tw}$  表示众包工人在第  $d$  天的时间窗  $tw$  期间经过的区域集合,  $num_l^{tw,d}$  表示众包工人在第  $d$  天的时间窗  $tw$  期间经过区域  $l$  的次数。  $\beta_d$  为时间衰减因子, 因为众包工人的移动行为会随时间改变, 在计算概率时加入时间衰减因子可以加强最近的轨迹记录并减弱距当前时间较久的轨迹记录的影响。时间衰减因子可通过式(2)计算:

$$\beta_d = \begin{cases} a \geq 1, & d = |TR_i| \\ \frac{1}{(|TR_i| - d + 1)}, & 1 \leq d < |TR_i| \end{cases} \quad (2)$$

然而, 只考虑众包工人的访问频率有可能增大一些众包工人只是经常访问的区域的预测概率。但是, 众包工人可能在这些区域停留很短的时间, 并且停留时间可能不足以完成任务。因此, 计算概率时还应该考虑众包工人在区域的停留时间。众包工人在该时间区间停留在该区域的时间越长则概率越大, 且众包工人越适合在此区域执行任务, 因为他在此区域内的停留时间足够完成任务。考虑停留时间的概率表示为:

$$p_{tw,l}^{dw} = \frac{\sum_{d=1,2,\dots,|TR_i|} \sum_t dw_l^{tw,d} \cdot \beta_d}{\sum_{d=1,2,\dots,|TR_i|} \sum_{l \in L_d^{tw}} dw_l^{tw,d} \cdot \beta_d} \quad (3)$$

其中,  $dw_l^{tw,d}$  表示众包工人在第  $d$  天的时间窗  $tw$  内经过区域  $l$  的次数。

由式(1)和式(3), 众包工人在时间窗  $tw$  内到区域  $l$  的概率可通过式(4)计算:

$$p_{tw,l}^* = \rho \cdot p_{tw,l} + (1 - \rho) \cdot p_{tw,l}^{dw} \quad (4)$$

其中, 参数  $\rho$  是权值且  $\rho \in (0, 1)$ 。

然而, 上述概率计算中没有考虑众包工人在不同区域间转移的关联性。众包工人当前所在的区域有可能离下一时间段常在的区域很远, 众包工人不会从当前区域转移过去。因此, 本文的概率计算方法考虑了众包工人在不同区域间的转移概率。

本文将众包工人  $\omega_i$  的移动建模成离散时间的时间齐次隐马尔可夫模型( $l_n^i, t_n^i$ )<sup>[32]</sup>,  $l_n^i$  表示众包工人  $\omega_i$  的第  $n$  个状态, 随机变量  $t_n^i$  表示众包工人  $\omega_i$  到达第  $n$  个状态的开始时间。由于  $\omega_i$  从状态  $l_n^i$  转移到状态  $l_{n+1}^i$  是完全独立于状态  $l_{n-1}^i$  的, 则过程  $l_n^i$  是一个标准离散时间马尔可夫链。随机变量  $dw_n^i$  表示在状态  $l_n^i$  的停留时间。这些随机变量独立同分布且不会随时间改变。

用  $M_{xy}^i(t)$  表示  $\omega_i$  在最多  $t$  个时间单位内从状态  $x$  转移到状态  $y$  的概率(开始离开状态  $x$ , 还没到状态  $y$ ), 其公式为:

$$\begin{aligned} M_{xy}^i(t) &= P(l_{n+1}^i = y, dw_{x,remain}^i \leq t | l_0^i \dots l_n^i, t_0^i \dots t_n^i) \\ &= P(l_{n+1}^i = y, dw_{x,remain}^i \leq t | l_n^i = x) \end{aligned} \quad (5)$$

其中,  $d\omega_{x,remain}^i$  表示  $\omega_i$  在状态  $x$  的剩余停留时间,且  $d\omega_{x,remain}^i = d\omega_x^i - (t_c - t_x)$ ,  $d\omega_x^i$  表示  $\omega_i$  在状态  $x$  的停留时间,该值可从众包工人历史记录中获得。 $t_c$  代表当前时间,  $t_c - t_x$  表示  $\omega_i$  在状态  $x$  已经停留的时间。

用  $p_{x,y}$  表示众包工人从状态  $x$  转移到状态  $y$  的转移概率。基于马尔可夫链,又考虑到时间因素的影响,  $p_{x,y}$  可由式(6)计算:

$$p_{x,y} = P(l_{n+1}^i = y | l_n^i = x) = \frac{\sum_{d=1,2,\dots,|TR_i^d|} \sum_{l \in L_{x,next}^d} num_y^d \cdot \beta_d}{\sum_{d=1,2,\dots,|TR_i^d|} \sum_{l \in L_{x,next}^d} num_l^d \beta_d} \quad (6)$$

其中,  $num_y^d$  表示众包工人在第  $d$  天从状态  $x$  转移到下一状态是  $y$  的次数;  $L_{x,next}^d$  表示众包工人在第  $d$  天内从状态  $x$  直接转移到其他状态的集合。结合众包工人在  $tw$  期间转移到区域  $l$  的概率,可得众包工人在离开状态  $x$  后在  $tw$  期间转移到下一状态是  $y$  的概率为:

$$p_{x,y}^{tw} = \omega \cdot p_{x,y} + (1-\omega) \cdot p_{tw,t}^* \quad (7)$$

其中,  $\omega$  为权重且  $\omega \in (0,1)$ 。由此可得  $\omega_i$  最有可能转移到下一状态为:

$$y = \operatorname{argmax}_{y \in L_{x,next}} (p_{x,y}^{tw}) \quad (8)$$

其中,  $L_{x,next}$  表示历史记录中  $\omega_i$  从状态  $x$  转移到的所有状态的集合。则  $\omega_i$  将在最多  $t$  个时间单位后离开状态  $x$  的概率为:

$$W_x^i(t) = P(d\omega_{x,remain}^i \leq t | l_n^i = x) = \sum_{y=1, y \neq x} M_{x,y}(t) = \frac{num_{d\omega_{x,remain}^i \leq t}}{num_{x,t}^i} \quad (9)$$

其中,  $num_{d\omega_{x,remain}^i \leq t}$  表示历史记录中剩余时间小于或等于  $t$  的次数,  $num_{x,t}^i$  表示  $\omega_i$  在状态  $x$  的停留次数。因此,  $\omega_i$  在状态  $x$  的最大剩余停留时间为:

$$t_{leave}^* = \operatorname{argmax}(W_x^i(t)) \quad (10)$$

用  $Q_{x,y}^i(t)$  表示  $\omega_i$  当前状态是  $x$ ,但在最少  $t$  个时间单位转移到了  $y$  的概率。若只考虑区域间的一次转移,则  $Q_{x,y}^i(t)$  可由式(11)得出:

$$Q_{x,y}^i(t) = P(t_{x,y}^i \geq t | l_n^i = x, l_{n+1}^i = y) = \sum_{k=t}^{t_{m,y}} P(t_{x,y}^i = k | l_n^i = x, l_{n+1}^i = y) = \frac{num_{t_{x,y}^i \geq t}}{num_{x,y}^i} \quad (11)$$

其中,  $t_{x,y}^i$  表示  $\omega_i$  从状态  $x$  到达状态  $y$  需要的时间,  $num_{t_{x,y}^i \geq t}$  表示转移时间大于  $t$  的次数,  $num_{x,y}^i$  表示  $\omega_i$  从状态  $x$  转移到状态  $y$  的次数。

## 4 任务推荐模型

本节介绍所提出的任务推荐模型。在移动众包系统中,携带智能设备的众包工人  $W = \{\omega_1, \omega_2, \dots, \omega_n\}$  随机到达,系统中有请求者发布的一些任务,用  $\Gamma = \{\tau_1, \tau_2, \dots, \tau_j, \dots, \tau_m\}$  表示。任务  $\tau_j$  的信息用  $Re_j = \{d_j, t_j, rp_j, skill_j, l_j\}$  表示。其中,  $d_j$  表示任务的截止时间,众包工人必须在截止时间前完成任务并提交任务数据;  $t_j$  表示执行该任务大概需要的时间;  $rp_j$  表示任务对执行者的信誉要求,不满足信誉要求的众包工人将不具备执行任务的资格;  $skill_j$  是任务对众包工人的技能要求,如会拍照、会开车等;  $l_j$  代表任务对众包工人的位置

要求。这些任务分为不同的类别,本文用  $C = \{C_1, C_2, \dots, C_o, \dots, C_k\}$  来表示任务类别的集合,其中,  $C_o$  代表任务类别  $o$ ,每个任务属于至少一种类别。为了给众包工人推荐适合的任务,本文提出了时空众包任务推荐模型,其运行流程如图1所示。

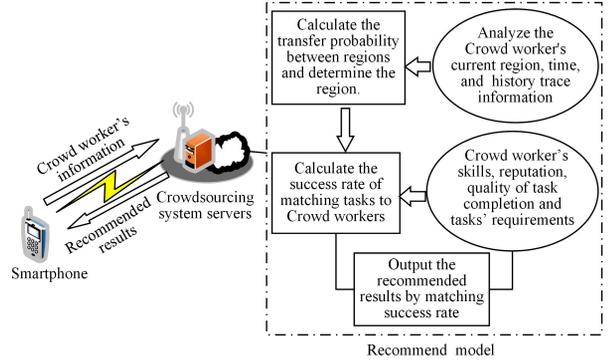


图1 移动众包系统任务推荐的基本流程

Fig.1 Process of task recommendation in mobile crowdsourcing system

考虑到时空众包任务的特性,所提出的时空众包系统的任务推荐模型有两个主要环节:

1)为任务推荐确定区域。根据众包工人日常移动记录的规律性,确定适合众包工人的任务执行区域,使众包工人在不影响日常生活的情况下完成任务。

2)众包工人和任务的匹配成功率。匹配成功代表众包工人选择了推荐的任务,且该任务的请求者选择了这个众包工人来执行任务。任务推荐的目的是给众包工人推荐其感兴趣的,还要尽可能保证所推荐的任务最终能与众包工人匹配成功。因此,为了避免屡次竞争任务失败的众包工人退出众包系统的情况,在为众包工人推荐任务时还要考虑他被成功选上执行该任务的概率。

### 4.1 为任务推荐确定区域

本文基于第3节提出的移动预测模型来确定任务推荐区域。算法1根据众包工人的当前所在区域、当前时间及其历史轨迹信息计算当前区域与其他区域之间的转移概率。

#### 算法1 确定区域算法

输入:众包工人历史轨迹记录  $TR_i = \langle \langle L_1, T_1, D_1 \rangle, \dots, \langle L_d, T_d, D_d \rangle \rangle$   
输出:适合为众包工人推荐任务的区域  $l^*$

1. 获取众包工人当前所在区域  $l_c$  和当前时间  $t_c$
2. 计算  $t_{leave}^* = \operatorname{argmax}(W_x^i(t))$
3. if  $\Gamma_c \neq \emptyset$  /\* 当前区域有任务 \*/
4.  $l^* = l_c$
5. else:
6.  $t_{leave} = t_c + t_{leave}^*$  /\* 计算大概离开当前区域的时间 \*/
7. 根据离开时间  $t_{leave}$  计算所在时间区间  $tw$
8.  $y^* = \operatorname{argmax}_{y \in L_{candidate}} (p_{x,y}^{tw})$
9. if  $\Gamma_{y^*} = \emptyset$
10. 将  $y^*$  从候选区域集合中删除
11. 返回第8行
12. else
13.  $l^* = y^*$
14. end if
15. end if

算法 1 描述了众包系统确定推荐任务所在区域的过程。当众包工人进入系统,系统获得众包工人当前所在区域和当前时间(第 1 行),然后计算众包工人可能在多长时间后会离开当前区域(第 2 行)。如果有任务在当前区域(第 3 行),将为众包工人推荐在当前区域的任务(第 4 行);否则,系统计算当众包工人离开当前区域时所在的时间窗(第 6,7 行)并预测众包工人将要去的下一区域(第 8 行)。如果没有任务在这个区域(第 9 行),将该区域从预测区域候选集中删除并选择一个新的区域(第 10,11 行);否则,系统将此区域内的任务推荐给众包工人(第 13 行)。

#### 4.2 众包工人和任务的匹配成功率

任务和众包工人匹配成功,表示众包工人选择该任务且该任务分配给了众包工人去执行。本文考虑了请求者发布的一般要求,包括众包工人的信誉、技能、过往执行此类任务的综合质量。影响众包工人选择任务的因素主要包括任务的类型、技能要求、报酬、时间适宜性(通过考虑众包工人的状态和任务截止时间来分析众包工人在时间上是否能够执行这个任务)和任务所在的区域。

众包工人只能完成符合自身技能的任务。任务的报酬是否符合众包工人的期望可以根据众包工人过往执行同类型任务的报酬来判断。

众包工人具备的技能与任务要求技能的符合度定义为:

$$s_{ij} = \begin{cases} 1, & skill_j \subseteq skill_i \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

其中,  $skill_i$  表示  $w_i$  的技能集合。式(12)表示当  $w_i$  具备执行任务  $\tau_j$  的所有技能时,则为完全匹配,否则为不匹配。

众包工人有各自的偏好<sup>[33]</sup>,  $w_i$  对类别为  $C_o$  的任务的偏好程度定义为:

$$pre_{C_o}^i = \frac{|\Gamma_{C_o}^i|}{|\Gamma^i|} \quad (13)$$

$w_i$  过往执行类别为  $C_o$  的任务的平均报酬为:

$$avg\_reward_{C_o}^i = \frac{\sum_{\tau_j \in C_o} reward_{ij}}{num_{C_o}} \quad (14)$$

众包工人的时间适合度为众包工人在留有执行任务的时间的情况下,在任务截止时间前能到达任务执行地点的概率,用公式表示为:

$$\begin{aligned} t\_s_{ij} &= P(t_{x,y}^i < t | l_n^i = x, l_{n+1}^i = y, t = d_j - t_c - t_j) \\ &= 1 - P(t_{x,y}^i \geq t | l_n^i = x, l_{n+1}^i = y, t = d_j - t_c - t_j) \\ &= 1 - Q_{xy}^i(t) = 1 - \frac{num_{t_{x,y}^i \geq t}}{num_{x,y}^i} \end{aligned} \quad (15)$$

即时间适合度表示众包工人在  $t$  个时间单位内从当前区域移动到指定区域的概率。本文使用加权的方法来综合评定任务对众包工人的吸引程度,由式(13)一式(15)可得任务对众包工人的吸引力为:

$$attraction_{ij} = \frac{1}{2} \alpha_1 \cdot e^{\frac{reward_j - avg\_reward_{C_o}^i}{avg\_reward_{C_o}^i}} + \alpha_2 \cdot pre_{C_o}^i + \alpha_3 \cdot t\_s_{ij} \quad (16)$$

其中,  $\alpha_1, \alpha_2, \alpha_3$  为控制报酬、偏好类别、时间适合度对吸引程度的影响的权重,且  $\alpha_1 + \alpha_2 + \alpha_3 = 1$ 。  $reward_j$  为任务  $\tau_j$  的报酬。式(16)表示众包工人倾向于选择自己感兴趣的,且报酬和时间适合度高的任务。

众包工人过往执行某类任务的综合质量由请求者对众包工人的反馈信息得出,用  $q_{ij}$  表示  $w_i$  执行任务  $\tau_j$  的质量,则  $w_i$  执行  $C_o$  类任务的综合质量为:

$$q_{C_o}^i = \frac{\sum_{\tau_j \in r_{C_o}^i} q_{ij}}{|\Gamma_{C_o}^i|} \quad (17)$$

请求者在众包系统中发布任务时会描述想要招募的众包工人的条件。设请求者对众包工人执行此类任务的综合质量要求为  $q_j$ , 信誉要求为  $rp_j$ , 且  $q_j, q_{C_o}^i, rp_j, rp_i \in (0, 1]$ 。则请求者对众包工人的满意程度可定义为:

$$satisfaction_{ij} = \begin{cases} 1, & q_i \geq q_j \text{ and } rp_i \geq rp_j \\ 0, & \text{otherwise} \end{cases} \quad (18)$$

式(18)表示当众包工人执行任务的综合质量和信誉达到请求者要求时,请求者对众包工人的满意程度为 1, 否则众包工人不满足条件且满意度为 0。

任务与众包工人的匹配成功率定义为任务对众包工人的吸引力和请求者对众包工人满意度的加权重,表示为:

$$match_{ij} = \lambda \cdot attraction_{ij} + (1 - \lambda) satisfaction_{ij} \quad (19)$$

其中,  $\lambda$  是用于平衡任务对众包工人的吸引力和请求者对众包工人的满意度对匹配成功率影响的权重系数。

#### 算法 2 任务推荐算法

输入: 区域  $l^*$ , 该区域中的任务集  $\Gamma_1^* = \{\tau_1, \tau_2, \dots, \tau_j, \dots, \tau_m\}$ , 任务要求  $Re_j = \{d_j, rp_j, skill_j, l_j\}$

输出: 推荐任务集  $Task_{rec}$

1. for  $\tau_j$  in  $\Gamma_1^*$
2. if  $s_{ij} = 1$
3.  $\Gamma_{candidate}^i \leftarrow \Gamma_{candidate}^i \cup \tau_j / *$  将满足要求的任务放入任务候选集中  $*$  /
4. end if
5. end for
6. for  $\tau_j$  in  $\Gamma_{candidate}^i$
7. Calculate  $match_{ij}$
8. end for
9. 将任务候选集  $\Gamma_{candidate}^i$  以匹配分数  $match_{ij}$  降序排列
10. while  $j < n$  do
11.  $Task_{rec} \leftarrow Task_{rec} \cup \tau_j$
12. end while

算法 2 描述了确定区域后的任务推荐过程。首先,对该区域内的所有任务,检测众包工人所具备的技能和任务对技能的要求是否匹配,将技能匹配的任务加入到众包工人的任务候选集中(第 1-5 行)。然后计算每个候选集中的任务与众包工人的匹配成功率(第 6,7 行),将匹配成功率按值的大小降序排列(第 9 行)。最后将排名前  $n$  的任务添加到推荐任务集(第 10,11 行)。

## 5 实验与结果

本节利用两个真实数据集来评估所提出的任务推荐算法的有效性和适应性。实验在 Windows 7 操作系统上运行,使用 Intel(R) Core(TM) i5-6500 CPU @ 3.20 GHz, 8.00 GB 内存, PyCharm 2017.1.4 仿真平台。

### 5.1 数据集

为了在实际场景下验证本文所提出的移动预测模型的准确性,我们使用了捷克博士生 Michal Ficek 的一个 142 天的

地面真实移动记录的数据集  $ctu^{1)}$  来模拟移动众包场景中的众包工人的轨迹信息。 $ctu$  数据集包含 Michal Ficek 142 天的信息记录时间、所在经纬度、位置区号、Cell-ID、移动国家代码(MCC)和移动网络代码(MNC)字段。我们将  $ctu$  数据集的位置区号字段映射为本文实验所需要的区域。为了获取区域的停留时间,将数据集中同一位置区号的连续时间段的数据合为一条数据,根据起止时间计算在该位置区的停留时间。根据数据集中的日期加入星期字段,并根据时间信息加入时间窗字段。实验使用一个包含了 12 294 个动画信息和 73 516 个用户的用户偏好数据集  $anime^{2)}$  来模拟众包系统中的众包工人和任务,选取其中有效数据最多的用户的数据作为实验数据,以此来验证本文所提出的推荐方法的准确性。其他实验参数的默认值如表 1 所列。在本实验中,为了增强当天轨迹的影响,将时间衰减函数的参数设为 5。基于人们对利益的追求,将  $\alpha_1$  设为 0.5,  $\alpha_2$  和  $\alpha_3$  分别设为 0.3 和 0.2,这是因为本文认为众包工人的偏好比时间适宜性对众包工人选择任务的影响更大。权重  $\rho$  设置为 0.6,以增强访问频率对在  $tw$  期间到达区域  $l$  的概率的影响。将  $\omega$  设置为 0.5,以平衡状态转移概率和访问概率的影响。为了使任务对众包工人的吸引力和请求者对众包工人的满意度在推荐计算时有同样的影响力,将  $\lambda$  设置为 0.5。此外,每轮推荐中的任务数设置为 10。本节使用相同的实验参数和实验环境进行对比实验,保证实验的真实性和公平性。

表 1 参数设置

$\alpha$	$\alpha_1, \alpha_2, \alpha_3$	$\rho$	$\omega$	$\lambda$	$n$
5	0.5, 0.3, 0.2	0.6	0.5	0.5	10

## 5.2 对比算法

由于目前关于移动众包的任务推荐中很少有研究考虑众包工人的历史移动轨迹来对众包工人进行任务的推荐,为了验证本文所提出的移动预测模型的准确性,实验选取了一些文献中提到的基线模型和用户移动预测模型,具体如下。

最频繁位置模型(Most Frequented Location Model, MF)<sup>[34]</sup>:该模型将用户一天中在给定时间区间  $h_c$  在位置  $x_c$  上签到的概率指定为以前在时间区间  $h_c$  内在位置  $x_c$  上签到的比例。更准确地说,令  $C_u$  为用户  $u$  的所有签到地点集合,则最频繁位置模型为:

$$P_M[x_u(t) = x | t \in h] = \frac{|\{c | c \in C_u, x_c = x, h_c = h\}|}{|\{c | c \in C_u, h_c = h\}|} \quad (20)$$

此模型选取指定时间区间  $h_c$  中签到比例最大的点为预测结果。虽然该模型比较简单,但是由于人的活动与时间的关联具有规律性,使得此模型有较好的效果。

工人移动模型(Worker Mobility Model, WMM)<sup>[35]</sup>:基于用户的历史位置记录,该模型将用户  $w$  在时间区间  $tw$  将要通过指定区域  $l$  的预测概率定义为:

$$\rho(w, tw, l) = \frac{|TD_{(w, tw, l)}|}{|TD|} \quad (21)$$

其中,  $|TD_{(w, tw, l)}|$  表示历史记录中用户  $w$  在时间区间  $tw$  通

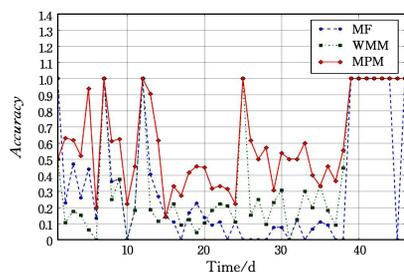
过指定区域  $l$  的天数,  $|TD|$  表示记录中总的天数。

## 5.3 实验结果

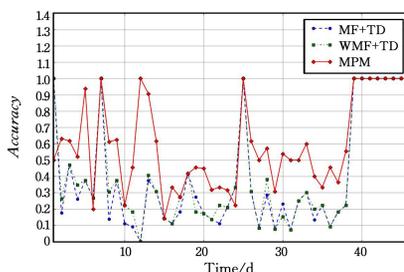
本文根据历史轨迹信息记录计算概率。由于众包工人在相同时间区间经常去的区域并不是单一的,通常会有很多个,因此,本文分别选取了概率最高的一个区域和概率排名前三的区域验证所提出的移动预测模型的预测准确性,当选取概率排名前三的区域作为预测结果时,认为实际结果在预测集中即为预测正确。实验中将上述数据集中 80% 的数据作为众包工人的历史轨迹数据,20% 作为测试数据。并且为了使实验更符合真实场景,每个测试完的数据会作为众包工人的历史轨迹信息加入训练集,作为下一次区域预测的参考数据。

(1) 每次选择概率最大的一个区域作为预测结果

图 2 为当选取概率最大的一个区域为预测结果时每一天的预测准确性。 $x$  轴表示测试数据集中的第几天,  $y$  轴表示预测准确性。从图 2(a) 中可以明显看出, MF 和 WMM 的预测准确性大部分都低于本文所提出的方法。其中 WMM 方法大部分的预测准确性在 10%~30%。与 WMM 和 MPM 不同, MF 的预测准确性出现分段情况,前 15 天左右的预测准确性大部分落在 20%~50% 之间,第 15~38 天左右的准确性大部分落在 0%~20% 之间,且有很多 0 值。然而, MPM 的预测准确性大部分在 30%~70%。图 2(b) 为增加了时间衰减参数后, MF, WMM 与 MPM 的对比结果。可以看出,增加时间衰减参数后, MF 和 WMM 的预测准确性均有所提高。其原因是 MF 和 WMM 在预测时只考虑了轨迹信息,没有考虑时间因素,以至于众包工人换了新的活动区域时算法不能及时地调整而出现预测失灵的现象。本文所提出的算法既考虑了时间因素又考虑了区域转移关联性,因此相比两个基线方法有较好的效果。



(a)



(b)

图 2 每天预测准确性的对比结果

Fig. 2 Experimental results on daily prediction accuracy

<sup>1)</sup> <https://crawdad.org/ctu/personal/20120315>

<sup>2)</sup> <https://www.kaggle.com/CooperUnion/anime-recommendations-database>

图3为随着测试数据增多,整体预测准确性的变化。 $x$ 轴表示测试数据数, $y$ 轴表示预测准确性。从图3(a)可以看出,随着数据量的增多,MF和MPM算法整体的预测准确性略有下降,但都高于WMM,且MPM算法的整体预测准确性比MF算法高出约32%,比WMM算法高出约35%。图3(b)为增加时间衰减因子后,MF,WMM与MPM的对比结果。可以看出,增加时间衰减因子后,WMM算法的准确性提升了约10%,MF的准确性提升了约5%,这是因为数据集中该工人有一段时间出差去了别的地方,原始的方法在预测时没有考虑时间因素,依然用过时的数据预测众包工人经常去区域,而那些区域当时离众包工人很远,因此其准确性很低。同样因为这个原因,MPM的预测也需要时间调节,因此准确性也受到了影响,整体的准确性略有下降但仍然比其他两个算法高出25%左右。

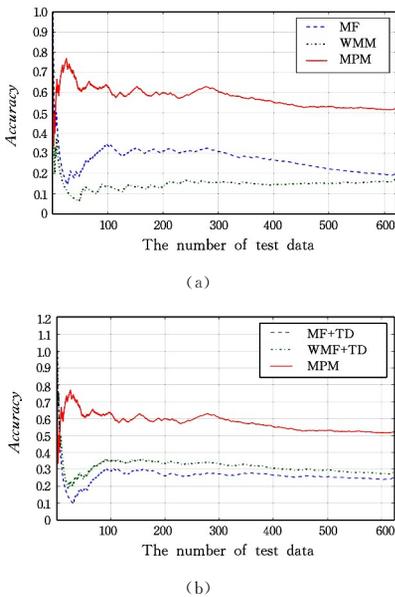


图3 随数据量增多后整体预测准确性的对比结果  
Fig. 3 Experimental results on overall prediction accuracy with increase of data

## (2) 每次选择概率值排前三的区域作为预测结果

图4为选取概率值前三的区域为预测结果时每一天的预测准确性。从图4(a)可以很明显看出,MF和WMM几乎每天的预测准确性都低于本文所提出的方法。其中WMM方法大部分的预测准确性在20%~60%,相比于只选择一个概率最大的区域作为预测结果的准确性提高了约20%,这说明有至少20%的区域是众包工人经常去但不是最频繁访问的区域。MF的预测准确性也在20%~60%,相比图2(a),前15天提升了约10%,15~38天提升了约30%。本文所提出的MPM算法的准确性大部分在60%~100%,提升了约30%。这充分表明本文提出的方法可以将众包工人最可能访问的区域从候选集中选出来。图4(b)为增加时间衰减参数后,MF,WMM与MPM的对比结果。对比图4(a)和图4(b)可以看出,增加时间衰减参数后,MF和WMM的准确性有很大的提高。这是因为增加时间衰减因子使得这两个算法可以迅速提高众包工人最近经常访问的区域的预测概率。

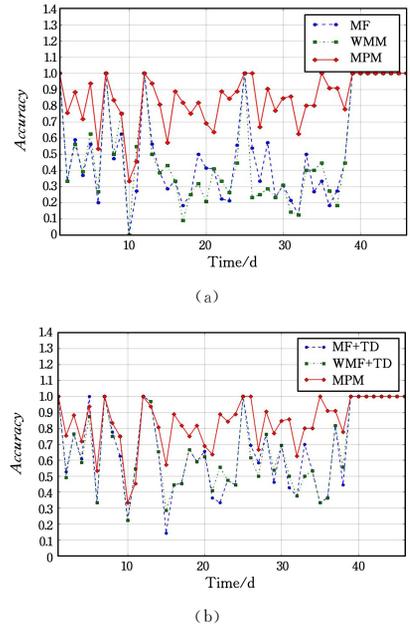


图4 选取3个区域为预测结果时每天预测准确性的对比结果  
Fig. 4 Experimental results of daily prediction accuracy when selecting three regions as the prediction result

图5为随着测试数据增多,整体预测准确性的变化。从图5(a)中可以看出,MPM算法的整体准确性比其他两个算法高40%,而MF与WMM算法的准确性几乎相同,这是因为这两种算法都是使用一段时间的访问频率来计算概率,当选取3个地点为预测结果时得出的结果相差不大。曲线在开始阶段有较大的波动,这是因为数据量少,有少量值的变化就会导致整体准确性有很大变化。图5(b)为增加了时间衰减因子后,MF,WMM与MPM的对比结果。

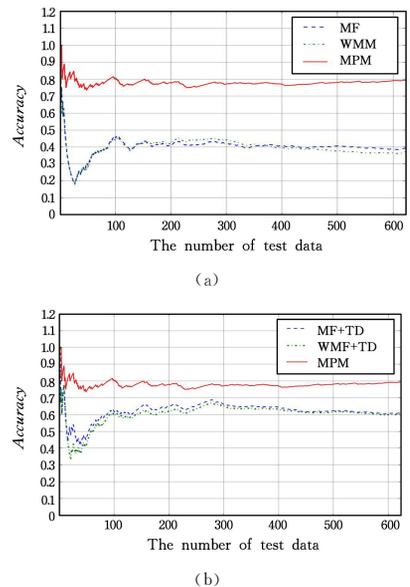


图5 选取3个区域为预测结果时随数据量增多后整体预测准确性的对比结果  
Fig. 5 Experimental results on overall prediction accuracy with increase of test data when selecting three regions as prediction result

对比图5(a)和图5(b)可以看出,增加时间衰减因子后,

WMM算法和MF算法的预测准确性均提升了约20%,相比于图3(a)提升了40%以上。而MPM算法相比于图3(a)提升了30%以上,这也说明众包工人过往30%的轨迹数据中访问了他访问概率较大的几个区域中的一个,而不是概率最大的区域。

图6为MPM算法预测准确性随着 $\alpha$ 值变化的变化情况, $x$ 轴表示 $\alpha$ 的值, $y$ 轴表示预测准确性。 $\alpha$ 值越大,说明当天访问区域对预测的影响越受重视。可以看出,当 $\alpha=3$ 时预测效果最好,这是因为 $\alpha$ 值太小不能突出当天访问区域的影响效果,而当 $\alpha$ 值太大则过于强调最近的访问区域,容易忽视历史访问记录的作用。

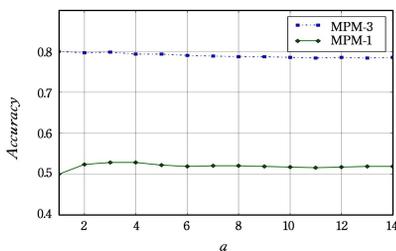


图6 随着 $\alpha$ 值增加整体准确性的变化

Fig. 6 Changes of overall prediction accuracy with the increase of  $\alpha$

图7表示权重参数 $\lambda$ 的值对推荐准确性的影响。横坐标表示推荐轮次,纵坐标代表推荐准确率。本文设定当任务所在区域为众包工人下一步将要去的区域且任务被众包工人选中则为推荐成功,否则认为推荐的任务未被众包工人选择。由图7可以看出,当 $\lambda=0$ 时推荐准确性最差,而其他取值下的推荐准确性相同。这是因为当 $\lambda=0$ 时,推荐任务时没有考虑任务对众包工人的吸引力,而请求者对众包工人的满意度只有两个取值,即0或者1,这样只会把众包工人胜任的任务筛选出来,并不能选出众包工人满意的任务,因此准确性很低;同时,由于任务对众包工人的满意度只有两个取值,当众包工人的质量和信誉满足大多数任务的要求时,决定任务推荐结果的就是任务对众包工人吸引力的大小,因此只要这个值不为0,得出的结果都是一样的。

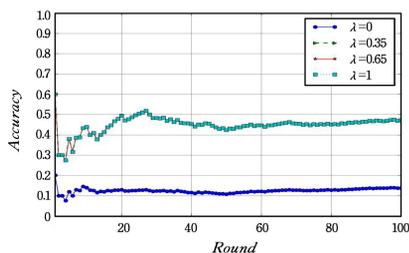


图7  $\lambda$ 值对推荐准确率的影响

Fig. 7 Influence of  $\lambda$  value on recommendation accuracy

本节对移动预测模型和任务推荐模型进行了实验,由于之前移动众包系统中的任务推荐没有考虑众包工人的移动规律,因此没有进行推荐方法的对比。若将没有考虑移动规律的方法与本文方法进行对比,由于其推荐任务时不考虑位置信息,推荐的任务所在区域是众包工人不会去的,准确性将会很低。预测准确性的实验结果显示,当预测区域设为概率值排前三的地点时,整体准确性会达到80%。为了保证区域预测的准确性,以此来保证任务推荐的准确性,本文为众包工人

按照区域预测概率分配每个区域的推荐任务数。例如每次为众包工人推荐10个任务,区域1的概率为0.3,区域2的概率为0.2,区域3的概率为0.1,则推荐 $(0.3/0.6) \times 10 = 5$ 个区域1中的任务,3个区域2中的任务和2个区域3中的任务。

**结束语** 本文提出了一种基于众包工人移动轨迹的任务推荐模型。首先,通过综合考虑停留时间、访问频率和转移概率建立移动预测模型,利用众包工人历史轨迹信息来对众包工人的移动行为进行预测。为了增强移动预测模型的适应性,在进行概率计算时加入了时间衰减因子,使得模型忘记距当前较久的轨迹信息并加强当天访问信息对预测的影响。然后,根据当前时间和位置预测众包工人下一步要去的区域来确定任务推荐区域,根据众包工人的属性和任务的要求对众包工人和任务计算匹配概率,以此进行任务推荐。最后,通过实验证明了本文所提出的移动预测模型和任务推荐模型的准确性。

本文采用的方法仅仅考虑了单一众包工人的历史轨迹信息,未来,将针对众包工人的社会关系对众包工人移动行为和任务选择的影响对模型加以改进,使推荐的结果更加精准、有效。

## 参考文献

- [1] WU Y, ZENG J R, PENG H, et al. Survey on incentive mechanisms for crowd sensing[J]. Ruan Jian Xue Bao/Journal of Software, 2016, 27(8): 2025-2047.
- [2] WANG Y J, CAI Z P, TONG X R, et al. Truthful incentive mechanism with location privacy-preserving for mobile crowdsourcing systems[J]. Computer Networks, 2018, 135: 32-43.
- [3] 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.
- [4] CAI Z P, ZHENG X. A Private and Efficient Mechanism for Data Uploading in Smart Cyber-Physical Systems[J]. IEEE Transactions on Network Science and Engineering, 2020, 7(2): 766-775.
- [5] LIU T E, WANG Y J, LI Y S, et al. Privacy protection based on stream cipher for spatio-temporal data in IoT[J/OL]. IEEE Internet of Things Journal, 2020. <https://ieeexplore.ieee.org/abstract/document/9079473>.
- [6] WANG Y J, CAI Z P, YIN G S, et al. An Incentive Mechanism with Privacy Protection in Mobile Crowdsourcing Systems[J]. Computer Networks, 2016, 102: 157-171.
- [7] QI LY, ZHANG X Y, DOU W C, et al. A Two-stage Locality-Sensitive Hashing Based Approach for Privacy-Preserving Mobile Service Recommendation in Cross-Platform Edge Environment[J]. Future Generation Computer Systems, 2018, 88: 636-643.
- [8] BI M N, WANG Y J, LI Y S, et al. A privacy-preserving mechanism based on local differential privacy in edge computing[J]. China Communications, 2020, 17(9): 1-17.
- [9] CAI Z P, ZHENG X, YU J G. A Differential-Private Framework for Urban Traffic Flows Estimation via Taxi Companies[J]. IEEE Transactions on Industrial Informatics, 2019, 15(12): 6492-6499.

- [10] QI L Y, CHEN Y, YUAN Y, et al. A QoS-Aware Virtual Machine Scheduling Method for Energy Conservation in Cloud-based Cyber-Physical Systems [J]. *World Wide Web Journal*, 2020, 23:1275-1297.
- [11] DUAN Z J, LI W, CAI Z P. Distributed Auctions for Task Assignment and Scheduling in Mobile Crowdsensing Systems[C]// 2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS). Atlanta, GA: IEEE, 2017: 635-644.
- [12] WANG Y J, CAI Z P, ZHAN Z H, et al. An Optimization and Auction-Based Incentive Mechanism to Maximize Social Welfare for Mobile Crowdsourcing [J]. *IEEE Transactions on Computational Social Systems*, 2019, 6(3): 414-429.
- [13] WANG Y J, GAO Y, LI Y S, et al. A Worker-selection Incentive Mechanism for Optimizing Platform-centric Mobile Crowdsourcing Systems[J]. *Computer Networks*, 2020, 171: 107144.
- [14] LIU H W, KOU H Z, YAN C, et al. Link prediction in Paper Citation Network to Construct Paper Correlated Graph[J]. *EURASIP Journal on Wireless Communications and Networking*, 2019, 2019(1).
- [15] GONG W W, QI L Y, XU Y W. Privacy-aware Multidimensional Mobile Service Quality Prediction and Recommendation in Distributed Fog Environment [J]. *Wireless Communications and Mobile Computing*, 2018, 2018: 1-8.
- [16] LI J, CAI Z P, YAN M Y, et al. Using Crowdsourced Data in Location-based Social Networks to Explore Influence Maximization[C]// The 35th Annual IEEE International Conference on Computer Communications. San Francisco, CA: IEEE, 2016: 1-9.
- [17] LI J, CAI Z P, WANG J B, et al. Truthful Incentive Mechanisms for Geographical Position Conflicting Mobile Crowdsensing Systems[J]. *IEEE Transactions on Computational Social Systems*, 2018, 5(2): 324-334.
- [18] QI L Y, ZHANG X Y, DOU W C, et al. A Distributed Locality-Sensitive Hashing based Approach for Cloud Service Recommendation from Multi-Source Data[J]. *IEEE Journal on Selected Areas in Communications*, 2017, 35(11): 2616-2624.
- [19] WANG Y J, YIN G S, CAI Z P, et al. A Trust-based Probabilistic Recommendation Model for Social Networks[J]. *Journal of Network and Computer Applications*, 2015, 55: 59-67.
- [20] YUEN M C, KING I, LEUNG K S. TaskRec: A Task Recommendation Framework in Crowdsourcing Systems[J]. *Neural processing letters*, 2015, 41(2): 223-238.
- [21] YUEN M C, KING I, LEUNG K S. An online-updating algorithm on probabilistic matrix factorization with active learning for task recommendation in crowdsourcing systems[J]. *Big Data Analytics*, 2016, 1(1): 14.
- [22] BABA Y, KINOSHITA K, KASHIMA H. Participation recommendation system for crowdsourcing contests[J]. *Expert Systems with Application*, 2016, 58(Oct. ): 174-183.
- [23] SAFRAN M, CHE D. Real-time recommendation algorithms for crowdsourcing systems[J]. *Applied Computing and Informatics*, 2017, 13(1): 47-56.
- [24] SAFRAN M, CHE D. Efficient Learning-Based Recommendation Algorithms for Top-N Tasks and Top-N Workers in Large-Scale Crowdsourcing Systems[J]. *ACM Transactions on Information Systems*, 2019, 37(1): 2. 1-2. 46.
- [25] WANG Y J, TONG X R, WANG K, et al. A Novel Task Recommendation Model for Mobile Crowdsourcing Systems[J]. *International Journal of Sensor Networks*, 2018, 28(3): 139-148.
- [26] PAN Q X, DONG H B, WANG Y J, et al. Recommendation of Crowdsourcing Tasks Based on Word2vec Semantic Tags [J/OL]. *Wireless Communications and Mobile Computing*, 2019. <https://doi.org/10.1155/2019/2121850>.
- [27] GUO Z W, TANG C W, NIU W J, et al. Fine-Grained Recommendation Mechanism to Curb Astrourfing in Crowdsourcing Systems[J]. *IEEE Access*, 2017, 5: 15529-15541.
- [28] LIU R, LIANG J B, GAO W Y, et al. Privacy-based recommendation mechanism in mobile participatory sensing systems using crowdsourced users' preferences[J]. *Future Generation Computer Systems*, 2018, 80(3): 76-88.
- [29] GONG Y M, FANG Y G, GUO Y X. Optimal Task Recommendation for Mobile Crowdsourcing with Privacy Control[J]. *IEEE Internet of Things Journal*, 2017, 3(5): 745-756.
- [30] ZHANG X F, SU J F. An approach to task recommendation in crowdsourcing based on 2-tuple fuzzy linguistic method[J]. *Kybernetes the International Journal of Systems & Cybernetics*, 2017, 47(8): 1623-1641.
- [31] LAN R S, ZHOU Y C, LIU Z B, et al. Prior Knowledge-Based Probabilistic Collaborative Representation for Visual Recognition[J]. *IEEE Transactions on Cybernetics*, 2020, 50(4): 1498-1508.
- [32] 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.
- [33] DUAN Z J, LI W, ZHENG X, et al. Mutual-Preference Driven Truthful Auction Mechanism in Mobile Crowdsensing[C]// The 39th IEEE International Conference on Distributed Computing Systems. ICDCS, 2019.
- [34] CHO E, MYERS S A, LESKOVEC J. Friendship and Mobility: User Movement In Location-Based Social Networks[C]// Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Diego, CA, USA: ACM, 2011: 1082-1090.
- [35] 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.



**HU Ying**, born in 1993, postgraduate. Her main research interests include mobile crowdsourcing and so on.



**WANG Ying-jie**, born in 1986, Ph.D., associate professor, is a member of China Computer Federation. Her main research interests include mobile crowdsourcing, privacy protection and trust computing.