

# 基于跨空间多元交互的群智感知动态激励模型

南文倩 郭 斌 陈荟慧 於志文 吴文乐 周兴社

(西北工业大学计算机学院 西安 710072)

**摘 要** 随着群智感知系统的大量出现,用户参与度成为影响系统服务质量的关键因素.为了提高用户参与感知应用的积极性,该文提出了基于跨空间多元交互的群智感知动态激励模型 CSII,利用跨空间数据分析并通过系统、任务发布者和参与者三者之间的多元交互来激励参与者高质量完成感知任务.具体来说,CSII 利用时空情景数据首先对感知任务进行评估得到任务代价并反馈给任务发布者和参与者,然后通过多元动态交互确定任务预算和工作集合,最后系统基于工作者历史信誉和出价使用逆向拍卖计算并支付报酬.该模型能够利用社交网络的时空情景数据完成任务的代价评估和工作者能力评估,与传统方法相比,CSII 能够提高工作者接受任务的意愿,并且促进线下任务的完成率.通过基于真实签到数据集实施的模拟实验表明,使用 CSII 激励机制后感知任务的完成率、用户的参与度和感知数据质量都得到了提高.

**关键词** 激励机制;跨空间;多元交互;移动群智感知;时空情境

**中图法分类号** TP301 **DOI号** 10.11897/SP.J.1016.2015.02412

## A Cross-Space, Multi-Interaction-Based Dynamic Incentive Mechanism for Mobile Crowd Sensing

NAN Wen-Qian GUO Bin CHEN Hui-Hui YU Zhi-Wen WU Wen-Le ZHOU Xing-She

(School of Computer Science, Northwestern Polytechnical University, Xi'an 710072)

**Abstract** With the surge of varied crowdsensing systems, active user participation becomes a crucial factor that determines whether a crowdsensing system can provide good service quality. To encourage user participation in mobile crowd sensing, we propose a novel incentive mechanism called CSII—a Cross-Space, multi-Interaction-based dynamic Incentive mechanism. CSII can estimate the value of a task based on the sensing context and historical data. It then has multiple interactions with both the task requester and potential contributors to dynamically adjust the budget and select suitable people to form the worker group. Finally, the requester pays the workers reward that they deserved based on their reputation and bids. Both online and offline data are leveraged to estimate task value and user quality for a particular task. Experiments show that the incentive mechanism can achieve good performance in terms of task completion rate, enthusiasm and data quality, and so on.

**Keywords** incentive scheme; cross-space; multi-interaction; mobile crowd sensing; spatio-temporal contexts

收稿日期:2014-07-01;最终修改稿收到日期:2015-05-26. 本课题得到国家“九七三”重点基础研究发展规划项目基金(2015CB352400)、国家自然科学基金(61332005,61373119,61222209)、教育部“新世纪优秀人才支持计划”(NCET-12-0466)和陕西省青年科技新星(2014KJXX-39)资助.南文倩,女,1989年生,硕士,研究方向为普适计算. E-mail: nwq\_xian@163.com.郭斌,男,1980年生,博士,教授,中国计算机学会(CCF)高级会员,主要研究领域为普适计算与移动群智感知.陈荟慧,女,1978年生,博士研究生,研究方向为普适计算与移动群智感知.於志文,男,1977年生,博士,教授,中国计算机学会(CCF)高级会员,主要研究领域为普适计算、社会感知计算.吴文乐,男,1990年生,硕士,研究方向为普适计算.周兴社,男,1955年生,教授,中国计算机学会(CCF)理事,主要研究领域为嵌入式计算和普适计算.

## 1 引言

随着智能手机的普及和移动感知技术的发展,大量基于智能手机的群智感知系统(Mobile Crowd Sensing Systems)被人们重视并应用于现实生活中<sup>[1-2]</sup>。例如,交通信息采集系统 VTrack<sup>[3]</sup>和 Nericell<sup>[4]</sup>,噪声采集系统 NoiseTube<sup>[5]</sup>和 CSNS<sup>[6]</sup>以及研究个人行为对环境影响的 PEIR<sup>[7]</sup>,都是由大量参与者完成数据采集任务的群智感知系统。

用户的参与度是决定群智感知系统能否获取高质量数据的关键因素,但是感知过程中的各种消耗会影响到用户参与和感知数据的质量。例如,群智感知系统通常需要采集一些内置传感器(如 GPS、麦克风、摄像头等)的数据,这可能会造成参与者隐私的泄露和手机电量的快速消耗;有些应用需要在移动客户端对数据进行预处理并上传到服务器,其中数据预处理会占用手机大量的计算资源,而上传数据会耗费参与者大量的通信流量。总之,隐私泄露、电量消耗和流量消耗等都会降低参与者采集数据的积极性,如果没有合理激励策略,感知任务很可能无法按时按需完成。

为了吸引更多的参与者加入并且提高数据质量,研究人员提出了许多激励模型<sup>[8-17]</sup>。但是,现有的激励模型并没有“以人为中心”考虑参与者的动态变化和时空特征,并且固定的预算机制不能适应感知环境和感知任务的动态变化。为了解决这些问题,本文提出了基于跨空间多元交互的群智感知动态激励模型 CSII(Cross-Space, multi-Interaction-based dynamic Incentive scheme)。

相比已有工作,CSII进行了两方面的创新:(1)在任务代价评估、任务分配、任务预算计算、报酬计算等环节,CSII促使群智任务管理系统、任务发布者和任务参与者之间进行多元交互,形成情境自适应和以人为本的动态激励模型,进而提高任务完成质量,促进用户参与;(2)通过跨空间数据分析进行群智任务组织和实施,例如 CSII结合 LBSN的线上签到数据和线下采集历史数据进行分析,完成具有时空特性的任务代价评估,以及具有时空特性和行为规律性的参与者能力评估。

## 2 相关工作

目前,群智感知中的激励模型分别针对在线激

励和离线激励两种情况进行了研究。

在线激励模型的特点是工作者按时间次序不重复认领任务,当某一工作者认领任务时,系统需要根据之前认领任务的其他工作者信息做出是否给该工作者分配任务的决定,且该决定不可更改。由于缺乏所有工作者的完整信息,文献[8-11]将部分工作者作为训练数据学习得到一个阈值,该阈值作为判断是否给空闲工作者分配任务的依据。由于该方案具有前期的训练过程,可能会造成任务分配和执行的延后;另外,前期工作者的信息只作为训练数据并不参与任务分配,因此该方案会损害前期工作者的利益。

离线激励模型的特点是系统给工作者分配任务前已经掌握了所有工作者的信息,包括任务成本和工作者出价等。下面介绍离线激励方面的代表性工作。

文献[12]为基于移动智能手机的群智感知任务设计了两种不同的激励模型:以系统为中心的激励模型和以用户为中心的激励模型。以系统为中心的激励模型采用斯坦伯格博弈(Stackelberg Game)寻求系统和用户利益最大化的均衡,文献[13]则采用贝叶斯博弈(Bayesian Game)实现此目的。以用户为中心的 LSB(Local Search-Based) auction 和 MSensing auction 是基于逆向拍卖的激励模型<sup>[12]</sup>,目的在于寻求最优的工作者集合以达到系统效用最大化的目的。

文献[14]提出了第二价格密封拍卖模型(Sealed-bid Second-price auction),工作者以密封的形式独立出价,发布者选择出价最低的工作者为胜出者,但是支付给胜出者的是所有出价中的第二低价。由于固定价格的拍卖模型<sup>[12,14]</sup>不能适应动态变化的数据采集环境,文献[15]提出了基于动态价格的逆向拍卖(RADP-VPC)来保证预算的可控性和工作者数量的稳定性。

文献[12,15]仅考虑了发布者提供的任务预算和工作者提供的出价,因此,文献[16]在 RADP-VPC的基础上提出了基于位置的激励模型,在固定预算约束下选择区域覆盖率最大的工作者集合,但是以用户为中心的区域覆盖并不能适应用户位置的动态可变性和多样性。虽然文献[16]在文献[15]的基础上考虑到了工作者位置信息,但忽略了工作者的能力特性和时间特性。文献[17]为不同工作能力的用户分配不同难度的任务,虽然考虑了用户的能力特性,但是没有考虑激励模型对用户参与积极性的影响。

综上所述,现有研究大多没有考虑参与者在激励模型中与发布者和系统之间的动态交互需求,也

没有考虑任务完成代价随时空特征变化的特点和参与者能力的个性化差异. 为了综合考虑预算约束、工作者能力特性、任务时空特性等在激励模型中的作用, 本文提出了基于跨空间多元交互的群智感知动态激励模型 CSII. CSII 通过多元动态交互实现感知任务的动态评估与激励优化, 通过跨空间数据分析评估任务代价和参与者能力, 以提高任务的完成质量和用户参与积极性.

## 3 问题定义与建模

### 3.1 问题定义

为了提高群智感知系统中用户参与感知任务的积极性, 并提高用户提交数据的质量, 本文提出了跨空间多元交互的激励机制 CSII. 在 CSII 激励机制中发布任务的用户称为发布者 (task provider), 将所有潜在执行任务的用户称为参与者 (participant), 将被群智任务管理系统 (后简称“系统”) 选择并推送任务的参与者称为目标参与者 (objective participant), 将接受并执行任务的目标参与者称为工作者 (worker), 将系统选中提交数据并获得报酬的工作者称为胜出者 (winner).

本文将任务表示为五元组的形式  $task = \langle m, b, o, s, d \rangle$ .  $m$  指完成该任务需要的期望工作者人数,  $b$  指任务的成本预算,  $o$  表示任务所要求采集数据的矩形区域,  $s$  指发布者发布任务的开始时间,  $d$  指该工作者将感知数据提交给系统的截止时间. 从开始时间  $s$  到截止时间  $d$  的时间段称为任务有效时间. 本文将任务执行的所有目标地理位置定义为矩形区域, 将整个矩形区域划分为很多小的等大矩形区域块. 此处定义的任务均为简单任务, 即任务的执行区域均为其中一个小矩形区域块, 复杂任务可由多个简单任务组合而成, 比如, 当发布者需要采集多个区域的交通情况时, 可以根据任务执行的多个矩形区域块将该任务划分为多个只在一个区域块内执行的简单任务.

群智感知系统中发布者发布感知任务, 系统根据任务需求进行工作者选择和任务推送. 传统方式下, 发布者一般凭借经验给出任务成本预算  $b$ , 缺少参考依据且没有可控性, 感知任务不一定能在合理的开销下被高质量地完成. 为此, CSII 提出了任务评估模型以通过分析 LBSN 信息空间数据来估算任务的代价, 为发布者的任务成本预算提供依据. 同时, 为了提高用户的参与度, 需要有目的地选择执行

任务的工作者, CSII 提出了参与者评估模型, 根据任务的需求并融合信息空间和物理空间数据选择执行任务最适合的工作者集合. 数据质量也影响着群智感知系统的性能, 如果工作者提供低质量的数据来赚取报酬, 则发布者的利益会受到损害. 因此, 在 CSII 激励机制中采用了信誉模型, 以保证收集到的数据的质量.

### 3.2 任务评估模型

由于发布者在很多情况下不容易估计任务代价, 所以设定的任务预算很可能不能满足任务的需求. 针对该问题, CSII 利用 LBSN 的线上签到数据对感知任务进行评估得到一个参考任务代价, 以辅助发布者设置合适的任务预算.

在任务评估模块中, 系统对 LBSN 的线上签到数据进行数据分析, 再根据分析结果对具有时间地点属性的任务进行评估得到执行任务的代价. 任务评估模块有两个影响因素: 区域热度和时间热度. 区域热度是任务所在区域中用户访问频度的度量, 时间热度是任务有效时间内用户访问频度的度量. 这两个因素根据信息空间中的用户签到数据和物理空间中的感知任务的时空属性计算得到的.

#### 3.2.1 区域热度

CSII 通过分析 LBSN 线上数据中的参与者及其对应的签到地点计算任务的区域热度. 区域热度随着任务覆盖区域的参与者数量和签到次数的增加而增加.

用  $H(o)$  表示发布者指定的任务区域  $o$  的区域热度, CSII 模型将区域热度归一化到  $0 \sim 1$  之间, 取值越大说明该区域的访问人数越多、签到频率越高. 此处定义的区域热度是单位区域内签到次数和签到用户多样性的体现, 式(1)用来计算区域热度  $H(o)$ .  $UT_{x,j}$  表示参与者  $x$  在地点  $j$  的签到次数,  $\Gamma$  表示所有签到地点的集合,  $U$  表示所有参与者集合. 式(1)前半部分通过用户的签到信息来刻画该区域的用户访问频率, 后半部分  $En(X)$  表示区域  $o$  中访问用户的多样性, 该多样性通过计算区域信息熵<sup>[18]</sup> 得到, 其中  $X$  表示在区域  $o$  签到的用户集合, 即  $X = \{x_1, x_2, \dots, x_n\}$ . 信息熵是信息论中用于度量信息量的概念, 是系统有序化程度的一个度量, 如果指标的信息熵越大, 该指标提供的信息量就越大. 区域信息熵通过式(2)得到, 其中  $p(x_i)$  表示用户  $x_i$  在区域  $o$  签到的概率, 即用户  $x_i$  在区域  $o$  签到次数占在所有区域签到次数的比例. 如果区域信息熵越大, 则在该区域签到的用户提供的信息量越大, 即该区域签到

的用户多样性程度越高.

$$H(o) = \frac{\sum_{x \in U, j \in o} UT_{x,j}}{\sum_{x \in U, j \in \Gamma} UT_{x,j}} \times En(X) \quad (1)$$

$$En(X) = - \sum_{x_i \in X} (p(x_i) \times \log_2 p(x_i)) \quad (2)$$

### 3.2.2 时间热度

CSII 通过分析 LBSN 线上数据中参与者的签到时间分布计算任务的时间热度.

由于参与者签到时间具有一定的弹性,为了任务评估的准确性,CSII 以一个小时为单位将一天分为 24 个时间段.  $VR_{x,z,j}$  表示参与者  $x$  在第  $z$  个时间段内访问地点  $j$  的次数,那么在第  $z$  个时间段内所有参与者访问区域  $o$  的概率表示为  $VP_z$ ,  $VP_z$  的计算公式如式(3)所示,其中  $AT$  表示 24 个时间段,  $AT = \{1, 2, \dots, 24\}$ . 根据任务描述信息中的开始时间  $s$  和持续时间  $(d-s)$ ,可以得到该任务覆盖的时间段集合  $ZT = \{z_1, z_2, \dots, z_{d-s}\}$ ,  $d-s \geq 1$ ,  $ZT$  为  $AT$  的子集. 时间热度定义为  $F_{ZT}$ , 取值在  $0 \sim 1$  之间, 计算如式(4)所示.

$$VP_z = \frac{\sum_{x \in U, j \in o} VR_{x,z,j}}{\sum_{x \in U, k \in AT, j \in o} VR_{x,k,j}} \quad (3)$$

$$F_{ZT} = \sum_{z \in ZT} VP_z \quad (4)$$

### 3.2.3 任务代价评估

任务评估模型利用 LBSN 线上签到数据分析得到任务的区域热度和时间热度,然后根据这两个属性评估任务代价.

$LWA$  表示通过式(1)得到的参与者签到数据中所有区域矩形块的区域热度平均值,  $TWA$  表示通过式(3)、(4)计算得到的所有参与者签到的时间热度的平均值. 定义当区域热度为  $LWA$ , 时间热度为  $TWA$  时任务代价为基准值 1, 该基准值是任务的单位代价. 对于任意一个任务,其代价  $v$  的计算如式(5)所示,其中,该式表示任务代价与所在区域的空间热度和时间热度成负相关.

$$v = \frac{LWA}{H(o)} \times \frac{TWA}{F_{ZT}} \quad (5)$$

### 3.3 参与者评估模型

在传统的激励模型中,通过所有的参与者执行任务,系统可以得到大量的感知数据,虽然大量数据会在一定程度上提高感知数据的质量,但会因数据冗余造成系统负担. 因此,CSII 通过评估参与者并与目标参与者交互选择理想的工作者集合. 每个参

与者都具有一定的时空活动规律,因此具有不同属性的参与者执行同一任务的能力不同. 参与者评估的目的是客观评价工作者是否适合执行某任务,从而根据其执行该任务的能力来决定是否向其推送任务. 该过程可提高参与者执行任务的完成率.

系统生成任务后,需要向最有可能完成该任务的参与者推送任务,并确定工作者集合  $WS$ . 系统选择在任务区域和有效时间内在 LBSN 中签到的参与者,并对这些参与者进行评估得到最适合执行任务的参与者,向其推送任务. 如果参与者接受任务,则将其加入到工作者集合  $WS$ ; 如果参与者拒绝任务,则继续在剩余的参与者中选择合适的参与者推送任务,直到  $WS$  中人数达到需求人数或候选参与者集合已全部和系统有过交互记录.  $WS$  集合中的工作者接受任务后根据任务要求感知数据. 参与者评估模型的影响因素为参与者的行为时空特性,即时空影响度.

将参与者在任务的有效时间和矩形区域  $o$  的签到次数作为参与者时空影响度的评判标准. 在任务区域  $o$  内签到的参与者作为候选参与者集合  $CS$ , 集合  $CS$  初始为空集. 若参与者  $x \in o$ ,  $CS = CS \cup \{x; x \in o\}$ .  $wt_x$  表示在任务的有效时间内  $CS$  集合中的参与者  $x$  在区域  $o$  内签到的次数,那么可通过式(6)将参与者签到次数进行归一化处理,计算得到时空特性对参与者  $x$  的影响度  $V_x$ ,  $V_x$  表示在任务区域和任务有效时间内参与者  $x$  的签到次数占任务区域内所有参与者签到总次数的比例,通过该比例可得到参与者  $x$  和感知任务在时空特性上的相关度,  $V_x$  取值范围在  $0 \sim 1$  之间. 该值越高说明参与者执行任务的能力越高.

$$V_x = \frac{wt_x}{\sum_{k \in CS} wt_k} \quad (6)$$

### 3.4 信誉模型

#### 3.4.1 数据质量评价指标

为保证工作者提交数据的质量,CSII 在向工作者支付任务报酬时将其信誉考虑在内. 本文假设信誉值高的工作者,其感知数据的质量相对较高. 本文通过两方面对工作者的信誉值进行计算:基于行为分析的客观评价和基于数据质量的主观评价<sup>[19-20]</sup>.

客观评价指参与者采集数据的时间地点和任务要求的时间地点之间的相符程度,主观评价指发布者根据数据内容与任务要求采集的数据之间的相关程度对参与者进行打分评价. 具体来讲,CSII 从采集数据的及时性、覆盖率和相关性<sup>[21]</sup> 3 个方面对数

据质量进行评价,其中前两点为客观评价,后一点为主观评价.本文采用模糊逻辑技术将及时性、覆盖率、相关性进行融合得到数据的综合质量.

(1)及时性.假设数据采集时间为  $t$ ,任务开始时间为  $s$ ,截止时间为  $d$ .如果参与者在任务有效时间内采集并提交数据,即  $s \leq t \leq d$ ,那么影响其信誉的时间因子为 1;如果在截止时间之后提交数据,即  $t > d$ ,随着用户提交数据的延后,时间因子对数据质量造成的负影响越大.式(7)中的  $g(t-d)$ 表示时间对数据质量的影响因子,其中  $f(t-d)$ 为 sigmoid 函数,  $sgn(t-d)$ 为符号函数,函数表达式分别为式(8)和式(9).图 1 显示了时间因子随时间变量(式(7)中的变量  $(t-d)$ 即为图 1 中的变量  $x$ )变化的函数模型.

$$g(t-d) = 2 \times sgn(t-d) \times f(d-t) + sgn(d-t) \tag{7}$$

sigmoid 函数  $f(x)$ 的表达式为

$$f(x) = \frac{1}{1 + e^{-x}} \tag{8}$$

符号函数  $sgn(x)$ 的表达式为

$$sgn(x) = \begin{cases} 1, & x > 0 \\ \frac{1}{2}, & x = 0 \\ 0, & x < 0 \end{cases} \tag{9}$$

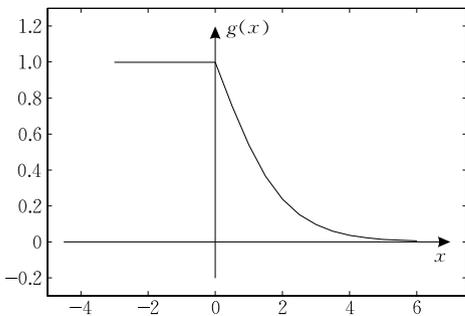


图 1 时间因子和位置因子的函数模型

(2)覆盖率.随着采集地点距目标采集地点间距的增大,地点因子对数据质量的负影响也变大.该规律同样可以用 sigmoid 函数来刻画.本文中任务执行地点为区域,即图 2 中的格子为一个区域,以步长为单位进行衡量(一个步长为一个区域块),步长数表示采集数据的区域距要求区域最短路径的区域数,例如图 2 中灰色区域为目标采集地点,三角形区域的步长为 1,五角星区域的步长为 2.假设步长数为  $r$ ,由于地点因子和时间因子对数据质量具有相同的影响规律,因此地点因子  $g(r)$ 的计算方法可通过式(10)表示,该规律符合图 1 所示的函数模型.

$$g(r) = 2 \times sgn(r) \times f(-r) + sgn(-r) \tag{10}$$

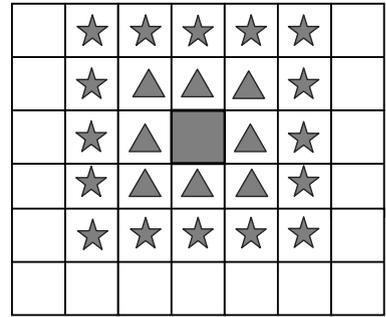


图 2 覆盖率评测模型

(3)相关性. CSII 使用评分制反映发布者对工作者提交数据的满意程度,即提交的数据和任务需求之间的相关程度,定义评价分数  $s$  的取值区间为  $[0, 1]$ .

### 3.4.2 模糊推理系统

在上述 3 个参数定义基础上,进一步采用模糊逻辑技术将时间因子、地点因子和相关性分数进行数据融合得到数据质量.模糊系统如图 3 所示,模糊系统由知识库、模糊化模块、模糊推理模块、反模糊化模块 4 部分组成.在该系统中,首先将输入变量按照图 4(a)的隶属度函数进行模糊化得到语言变量<sup>[22]</sup>(即模糊变量),然后根据模糊规则进行推理得到数据质量的语言变量值,最后将语言变量表示的数据质量按照图 4(b)的隶属度函数进行反模糊化得到数据质量  $q$ .

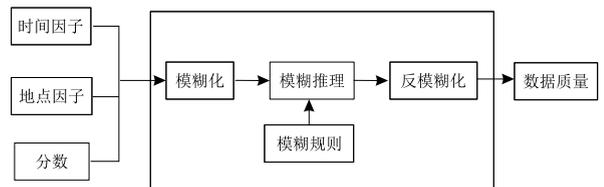
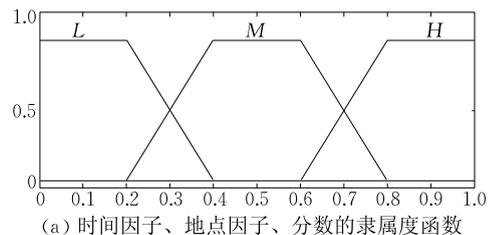
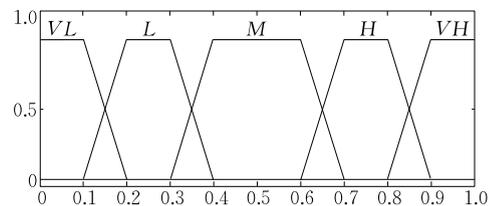


图 3 模糊逻辑模型结构



(a) 时间因子、地点因子、分数的隶属度函数



(b) 数据质量的隶属度函数

图 4 输入输出变量的隶属度函数

知识库. 模糊集合和模糊规则一起构成了推理系统的知识库. 一个变量可取的语言值组成该变量的模糊集合  $A$ . 本文规定输入变量“时间因子”、“地点因子”和“分数”的模糊集合由“高”、“中”、“低”组成, 并分别用符号  $H$ 、 $M$ 、 $L$  表示; 输出变量“数据质量”的模糊集合由“非常高”、“高”、“中”、“低”、“非常低”组成, 并分别用符号  $VH$ 、 $H$ 、 $M$ 、 $L$ 、 $VL$  表示. 则输入输出参数的模糊集合分别表示为

$$A(\text{时间因子}) = A(\text{地点因子}) = A(\text{分数}) \\ = \{L, M, H\},$$

$$A(\text{数据质量}) = \{VL, L, M, H, VH\}.$$

数值型变量  $x$  到语言值变量  $A(x)$  (隶属度) 的映射构成隶属度函数. 图 3 模糊系统中的输入输出变量的隶属度函数如图 4 所示. 若隶属度为 1 表示  $x$  完全属于模糊集合  $A$  中的对应变量; 若为 0 表示完全不属于  $A$  中的相应变量; 若位于区间  $(0, 1)$  则表示属于模糊集合中相应变量的概率.

模糊规则描述了模糊系统的动态行为, 这些规则基于专家知识或经验. 模糊规则一般为 IF-THEN 的形式. 本文定义的模糊规则如表 1 所示.

表 1 根据时间因子、地点因子和评价分数确定数据质量的模糊规则

规则编号	IF 时间因子	AND 地点因子	AND 分数	THEN 数据质量
1	L	L	L	VL
2	L	L	M	VL
3	L	L	H	L
4	L	M	L	VL
5	L	M	M	VL
6	L	M	H	M
7	L	H	L	L
8	L	H	M	M
9	L	H	H	H
10	M	L	L	VL
11	M	L	M	M
12	M	L	H	M
13	M	M	L	L
14	M	M	M	M
15	M	M	H	H
16	M	H	L	M
17	M	H	M	M
18	M	H	H	VH
19	H	L	L	L
20	H	L	M	M
21	H	L	H	H
22	H	M	L	M
23	H	M	M	M
24	H	M	H	VH
25	H	H	L	M
26	H	H	M	VH
27	H	H	H	VH

模糊化. 该模块是将模糊系统的数值型输入变量转换成模糊变量的过程. 该过程通过输入变量相

关联的隶属度函数来实现, 即为模糊系统中的输入值赋予了相对应的隶属度.

模糊推理. 模糊推理模块将模糊化的输入映射到表 1 所示的 IF-THEN 规则库, 并为之生成一个模糊化输出的过程. 本文采用 Mamdani<sup>[23]</sup> 推理机实现输入变量的模糊推理过程.

反模糊化. 该模块是模糊化模块的逆操作, 即将模糊推理得到的语言变量根据隶属度函数转换成数值变量的过程.

### 3.4.3 信誉计算模型

CSII 根据参与者历史采集数据的质量来计算参与者的信誉, 定义时间衰减因子来表示历史数据质量对信誉的影响权重. 历史数据对信誉的影响随时间变化, 早期数据对信誉的影响力变弱, 即近期采集数据的质量更大程度上影响下一次的采集行为. 因此, 为了准确预测参与者在当前时间之后采集数据的质量, CSII 考虑了参与者采集数据的行为模式随时间的动态变化, 以准确反映参与者的信誉.

CSII 采用时间窗的方法计算时间衰减因子, 将参与者签到的时间分为若干个时间窗, 每个时间窗的时间间隔为  $\theta$ , 参与者最早接受任务的时间段为第一个时间窗  $T_1$ . 指数函数  $e^{(T_i - T_h)}$  表示时间衰减因子, 时间衰减因子表明距当前时间较近的数据对参与者信誉影响较高, 而距当前时间较远的数据对参与者信誉影响较低. 假设当前时间为  $t$ , 其所在的时间窗为  $T_h$ , 在  $T_h$  个时间窗内参与者  $x$  的信誉  $R_x(T_h)$  的计算方法如式(11)所示, 该式表示数据质量按照时间衰减的规律决定参与者的信誉. 式中  $q_{x,t,k}$  表示参与者  $x$  在第  $T_t$  个时间窗内执行第  $k$  个任务提交数据的质量, 该值由 3.4.2 节的模糊推理系统得到,  $n_t$  表示参与者  $x$  在第  $T_t$  个时间窗内执行任务的个数.

$$R_x(T_h) = \sum_{t=1}^h \left( e^{(T_t - T_h)} \times \frac{1}{n_t} \sum_{k=1}^{n_t} q_{x,t,k} \right) \quad (11)$$

## 4 动态激励方法设计

用户在执行群智感知任务的过程中会因为电量、流量、计算资源等原因产生一定的花销, 在不提供激励的情况下这些因素会降低用户参与感知任务的积极性和采集数据的质量. 本文提出了跨空间多元交互的激励机制 CSII, CSII 融合信息空间人的活动信息和物理空间的数据采集行为以及历史信息构建了任务评估模型、参与者评估模型和信誉模型(已

在第 3 节介绍),在这 3 个模型的基础上通过发布者、系统、工作者三者之间的交互实现向发布者提供任务预算建议,提高了用户参与度和数据质量等。

#### 4.1 整体设计

CSII 包括 3 个交互主体:任务发布者、群智任务管理系统和参与者。激励过程如图 5 所示:首先发布者通过系统发布群智感知任务,系统根据任务评估模型得到任务代价并反馈给发布者和参与者,然后通过系统、发布者和参与者之间的多元动态交互确定任务预算,并通过参与者评估模型确定工作者集合,最后系统根据工作者的信誉和任务出价按照逆向拍卖的原则选出胜出者,并分别向胜出者和拍卖失败者支付任务报酬和任务补偿。

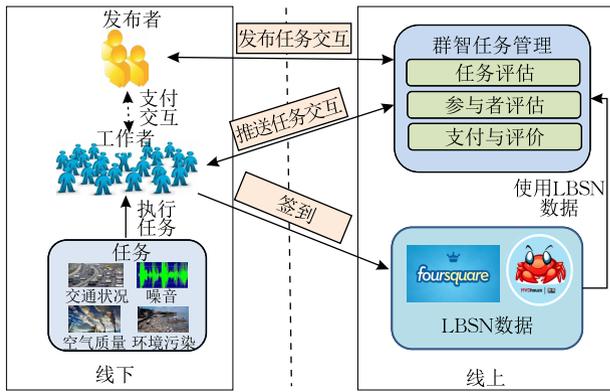


图 5 基于跨空间多元交互的群智感知动态激励模型

CSII 通过多元交互实现感知任务的分配与激励优化,从交互主体角度出发,本文将 CSII 激励过程分为 4 个阶段:任务发布阶段、工作者选择及任务分配阶段、胜出者选择阶段、数据提交和报酬支付阶段。

**任务发布阶段.**发布者和系统进行交互,发布者向系统提交任务,系统评估任务后将任务代价和预算建议反馈给发布者,发布者确定预算并发送给系统。

**工作者选择及任务分配阶段.**参与者和系统进行交互,系统评估参与者执行任务的能力选择理想目标参与者推送任务,被推送任务的参与者决定是否接受该任务,并反馈给系统,系统选出工作者集合。

**胜出者选择阶段.**工作者和系统进行交互,工作者完成任务后向系统提交出价,系统根据逆向拍卖的原则,综合考虑工作者的出价和信誉选择胜出者并通知工作者。

**数据提交和报酬支付阶段.**发布者和工作者进行交互,胜出者向发布者提交感知数据,发布者根据数据质量对胜出者进行评价更新胜出者的信誉,并向胜出者和其他工作者支付报酬和任务补偿。

#### 4.2 激励机制流程

**任务生成.**发布者向系统提交任务信息后,系统利用 3.2 节的任务评估模型得到任务代价,根据代价调整发布者的预算,并将任务代价和建议预算反馈给发布者,发布者得到这些信息后对初始设定的预算做出调整发送给系统,如果发布者的初始预算高于任务代价,则建议发布者将预算降低到任务代价,以降低发布者的任务开销;如果发布者的初始预算低于任务代价,那么建议发布者将预算提高到任务代价,以提高任务的完成质量和任务完成率。系统根据发布者调整结果重新生成任务信息。

**选择工作者.**任务发布后,系统利用 3.3 节的参与者评估模型选择工作者集合。通过该模型得到参与者的时空影响度  $V_x$ ,按照  $V_x$  从大到小对参与者进行排序,然后向前  $m$  个用户推送任务,如果用户  $x$  接受任务,则将其加入到工作者集合  $WS$  中,即  $WS = WS \cup \{x\}$ ;如果用户  $x$  拒绝任务,则向第  $m+1$  个工作者推送,依次类推得到工作者集合  $WS$ 。 $WS$  中的工作者按照任务要求执行任务。

**逆向拍卖.**系统确定工作者集合后,工作者根据任务描述执行任务感知数据,并在规定的截止日期前提交任务出价。系统综合考虑工作者的出价和信誉并按照逆向拍卖模型选择胜出者,并将选择结果发送给所有工作者。为了避免数据冗余,CSII 只选择一个工作者作为胜出者,胜出者向发布者提交感知数据,发布者评价胜出者的工作,然后向胜出者和其他工作者支付相应的报酬和任务补偿。

本文在逆向拍卖过程中考虑到工作者感知数据的质量,认为工作者的信誉是对其感知数据质量的直接反映,所以综合考虑工作者的信誉和任务出价,通过式(12)计算得到工作者  $x$  的新出价  $\omega_x$ ,假设该时间位于第  $T_h$  个时间窗,工作者信誉为  $R_x(T_h)$ 。系统选择新出价  $\omega_x$  最小的工作者作为胜出者。式中  $bid_x$  表示工作者  $x$  拍卖其数据时的出价,本文假设所有工作者都是理性的,为了赢得拍卖成为胜出者,其出价小于任务预算。同时工作者有心理底价  $bp$  (basic price),其出价不小于该心理底价,即工作者  $x$  的出价满足条件:  $bp_x \leq bid_x \leq b$ 。

$$\omega_x = bid_x \times (1 - R_x(T_h)) \quad (12)$$

**报酬支付.**胜出者向发布者提交感知数据,发布者评价胜出者的工作,系统根据 3.4 节的信誉模型更新胜出者的信誉,最后发布者根据式(13)分别向拍卖胜出者和失败者支付相应的报酬  $p_x$ 。为了提高工作者的参与度,CSII 对拍卖失败者支付一定的报

酬作为任务补偿. 发布者向胜出者支付其在逆向拍卖中的出价  $bid_x$ , 该出价表示工作者提交数据所希望得到的报酬. 如果有剩余预算, 则将这些剩余预算平均分发给拍卖失败者作为任务补偿.

$$p_x = \begin{cases} bid_x, & x \text{ 为胜出者} \\ \frac{b - bid_{winner}}{m - 1}, & x \text{ 为失败者} \end{cases} \quad (13)$$

## 5 实验验证

本文采用实验仿真结合真实采集数据的形式对 CSII 模型进行验证. 下面将分别从数据准备、实验目标及测评指标、实验结果等几个方面进行说明.

### 5.1 数据准备

本文采用结合在线 LBSN 数据的方法进行任务代价评估和工作者选择, 实验中采用来自于国内的 LBSN 服务——街旁<sup>①</sup>的数据进行测试. 实验以用户较为密集的上海市为背景, 共采集上海市 54148 个用户的 966 814 条签到记录. 每个用户都有签到地点和签到时间两个属性, 签到地点覆盖的区域为 (121. 2138, 31. 0179) 至 (121. 8059, 31. 3422). 在进行任务评估和参与者评估时, 参与者的签到时间以小时为单位, 966 814 条记录的签到时间涵盖了一天中的 24 小时. 将数据集标识的区域划分为多个边长为  $0.01^\circ$  的矩形区域.

### 5.2 实验目标及测评指标

#### 5.2.1 实验目标

CSII 中的 3 个交互主体共同影响着群智感知系统提供服务的质量: 参与者的参与积极性、发布者发布任务的积极性. 对此, 本文主要针对这两个方面进行仿真实验以验证 CSII 在群智感知系统中所起到的促进作用.

参与者的参与积极性. 提高参与者的参与积极性是本文设计激励机制的主要目的, 对此本文提出了影响参与者参与积极性的实验评测指标: 用户参与度. 较高的用户参与度能够保证群智感知系统向任务发布者提供可靠稳定的群智感知服务.

发布者发布任务的积极性. 群智感知系统中任务发布者对整个系统的运作起着决定性的作用. 然而, 以下一些因素会降低发布者发布任务的积极性, 进而降低 CSII 的现实可用性: 参与者在认领任务后提供较低质量的数据; 参与者认领任务后由于一些客观或主观因素造成较低的任务完成率; 由于发布者缺少对任务代价客观准确的评估而导致任务预算

不可控, 不可控的任务预算引起较高任务开销. 因此, 本文中的仿真实验需要对胜出者信誉、任务完成率以及任务开销 3 个方面进行验证.

#### 5.2.2 测评指标

根据 5.2.1 节的实验目的, 本文的实验测评指标为用户参与度、胜出者信誉、任务完成率、任务开销.

用户参与度. 用户参与度指用户参与感知任务的意愿程度. 由于多数工作者受到发布者的支付激励而参与感知任务, 在群智系统中用户收益是影响用户参与的一个主要因素. 而发布者所提供的任务预算是用户收益的一个决定性因素, 如果任务预算低于执行任务时电池电量、通信流量和计算资源的花销, 由于较低的任务预算导致的低收益阻碍了参与者的参与意愿. 因此, 可以研究工作者在多次接受任务并提交感知数据后得到的总收益以及发布者的任务预算来评价用户参与积极性.

(1) 任务接受率. 当发布者的预算低于任务代价时, 会因目标参与者拒绝执行系统推送的任务而导致任务不能完成或完成质量下降. 由于当预算小于任务代价时参与者的任务接受率随着预算的降低而降低, 那么当目标工作者集合人数为  $OM$  时, 接受任务成为工作者的人数  $WM$  可通过式(14)得到. 式中  $b$  表示发布者的任务预算,  $v$  表示根据 3.2 节的任务评估模型得到的任务代价,  $\lambda$  为调整参数. 该式表示当任务预算低于任务代价时, 任务接受率随着任务预算成比例降低.

$$WM = \lambda \times \frac{b}{v} \times OM \quad (14)$$

(2) 用户参与度. 由于用户参与度受用户收益以及发布者的任务预算两方面因素的影响, 因此, 可根据式(15)得到用户参与度的表达式,  $c$  表示发布者得到数据服务的实际花销,  $\alpha$  和  $\beta$  为两个因素的影响权重.

$$D_x = \begin{cases} \alpha \times \frac{bid_x}{bp_x} + \beta \times \frac{c}{b}, & x \text{ 为胜出者} \\ \alpha \times \frac{b - bid_{winner}}{(m - 1) \times bp_x} + \beta \times \frac{c}{b}, & x \text{ 为失败者} \end{cases} \quad (15)$$

胜出者信誉. 本文通过考虑参与者的信誉来选择胜出者. 由于信誉是对参与者历史任务完成情况的客观评价, 信誉值高的参与者其提交的数据质量也相对高. 因此, 选择信誉值高的参与者执行任务可以提高其提交数据的质量. 此外, 文献[24]指出竞争

① <http://jiepang.com/>

环境下的微支付更有利于高质量数据的采集,因此,系统采用逆向拍卖模型有利于提高数据的质量.

**任务完成率.** 任务完成率指工作者完成任务的总数占发布者发布任务总数的比例. 本文采用 LBSN 线上签到数据和线下历史数据相结合的方式实现激励模型中的参与者评估和工作者选择. 选择工作者时,根据所有用户在街旁的签到地点和时间有目标地向理想用户推送任务. 在传统的逆向拍卖激励模型中,由于没有考虑到用户的时空特性,在选择工作者集合时具有一定的盲目性,从而造成任务有可能不能完成,从而降低了任务完成率.

**任务开销.** 任务开销指发布者对所有工作者支付的报酬总和,即发布者对任务的总支出. 发布者发布任务的同时向系统提交了任务预算. 如果发布者预算高于系统对任务的估算代价,那么根据任务评估模型发布者可以将预算降低到任务代价. 本文假设所有参与者都是理性的,当预算不低于任务代价时所有参与者都会接受任务,当预算低于任务代价时参与者的任务接受率随着预算的降低而降低. 那么在所有用户都接受任务的情况下,发布者将预算降低到评估的任务代价时可以降低任务开销. 式(16)中,  $ED$  表示任务开销,从公式中可以看出,发布者通过调整预算可降低任务开销.

$$ED = \begin{cases} v, & v < b, \text{将预算调整为任务代价} \\ b, & v < b, \text{保持原预算不变} \end{cases} \quad (16)$$

### 5.3 实验结果

本文先对用户的签到数据进行预处理,选择签到次数最多的 7 个区域进行实验,这 7 个区域的签到次数分别为 26380, 5838, 4264, 1469, 1453, 1263, 995 次,在这 7 个区域中一共有 830 个用户签到,总签到记录数为 110306 个. 在这 7 个区域随机生成 1000 个任务,根据任务评估模型评估得到任务代价,并根据参与者评估模型选择 30 个工作者执行任务.

#### 5.3.1 用户参与度

用户参与度受任务接受率和工作者收益两个因素影响,因此本文通过研究任务接受率和工作者收益这两个指标来研究工作者参与感知任务的积极性.

(1) 任务接受率. 参与者接受任务的意愿程度受任务预算的影响,因此本节将研究任务预算对参与者任务接受率的影响. 实验任务选取的区域块的签到次数为 4264 次,任务执行时间为 11 点到 18 点,需求人数为 15 人. 当候选目标参与者集合增大时,可供系统选择的工作者数量增大,因此在预算和任

务代价确定的情况下,选择到的工作者人数与候选目标参与者人数成一定比例关系. 图 6 显示了利用式(14)计算的在不同预算和候选集合下能够选择的工作者人数,即任务接受率. 当候选目标参与者从 30 变化到 60 时,工作者人数随着候选目标参与者的增多而增多,直到达到需求人数(15 人)时不再变化. 系统根据第 3.2 节任务评估模型中的式(1)~(4)估算的任务代价为  $v$ ,从图中可看到当预算  $b$  分别取  $(0.1 \times v)$ ,  $(0.2 \times v)$ ,  $(0.3 \times v)$ ,  $(0.5 \times v)$ ,  $v$  时工作者人数的变化. 在实验中,当预算高于  $(0.5 \times v)$ , 候选参与者人数为需求工作者人数 2 倍以上时,参与者的任务接受率为 100%;当预算为  $(0.3 \times v)$  时,只有当候选目标参与者人数达到一定数量(实验中为 50 人以上)时系统选择的工作者人数才能达到发布者期望的人数;而当预算小于  $(0.3 \times v)$  时,即使候选者人数为需求人数的 4 倍,接受任务的人数也不能达到任务所需人数,对任务的完成质量甚至任务完成率会造成一定的影响. 因此从实验中可以得到,候选目标参与者集合大小和需求者的预算都会影响参与者的任务接受率,因此在足够的候选目标参与者的情况下,如果发布者的预算低于系统的任务代价,发布者可以根据系统建议将预算提高到任务代价,从而用户的任务接受率将会提高到 100% (理想情况下).

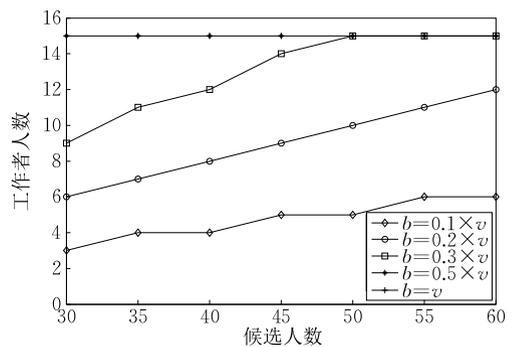


图 6 任务接受率(工作者人数预定=15 人)

(2) 工作者收益. Lee 等人<sup>[15]</sup>提出的动态价格逆向拍卖(Reverse Auction based Dynamic Pricing incentive mechanism, RADP)、固定价格随机选择激励机制<sup>[15]</sup>(Random Selection based Fixed Pricing incentive mechanism, RSFP)中只有胜出者才可以得到报酬,而在 CSII 中,胜出者和失败者都可以得到相应的报酬,这种支付方案能够提高用户的总体收益,从而提高用户参与群智感知任务的积极性. CSII 中工作者收益的计算方法如式(13)所示, RADP 按照胜出者的出价支付报酬, RSFP 向胜出者支付

固定价格的报酬, 实验中设置 RSFP 激励机制支付的固定价格为预算的 0.3 倍. 图 7 显示了 CSII 和 RADP、RSFP 的参与者收益对比, CSII 的全支付策略相对 RADP 和 RSFP 使工作者有较高的收益. 实验中, 以均匀分布随机生成工作者的心理底价, 工作者出价均匀分布于心理底价和预算之间.

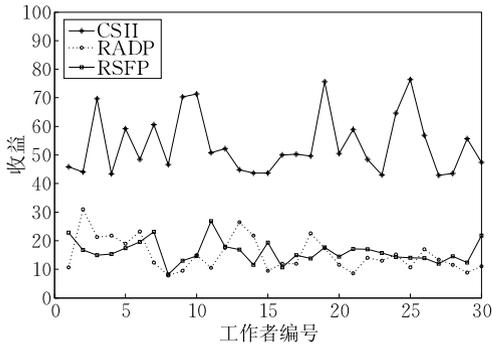


图 7 工作者收益对比

实验设定式(15)中的影响参数  $\alpha$  和  $\beta$  的取值分别为 0.6 和 0.4, 那么可以得到在不同激励模型下的用户参与度. 如图 8 所示, 与 RADP、RSFP 相比, CSII 具有较高的用户参与度. 在较高的任务预算下, 工作者即使没有在逆向拍卖中获胜, 也会得到相对较高的任务补偿, 而在 RADP 和 RSPF 的逆向拍卖中工作者会因为拍卖失败而没有任务补偿. 因此 CSII 相对这两种模型给予工作者更高的收益, 提高了工作者接受任务的意愿程度.

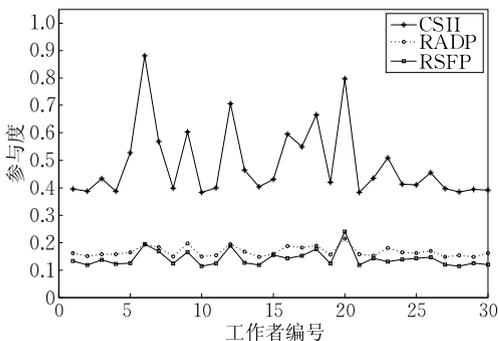


图 8 用户参与度

### 5.3.2 胜出者信誉

CSII 在逆向拍卖选择胜出者时, 综合考虑了工作者的出价和信誉, 有助于发布者得到高质量的数据. 而 RADP 只考虑发布者的出价, RSFP 则是随机选择工作者成为胜出者, 这两种方案都忽略了工作者信誉这一重要因素, 因此, 在图 9 所示的 1000 次实验中, 在 RADP 和 RSFP 激励机制下大约 90% 的胜出者的信誉处于 0.1~0.5 之间, 而在 CSII 下有 50% 以上的胜出者信誉大于 0.5. 由此可得出, CSII

激励机制下的逆向拍卖模型相对其他两种激励机制能够得到较高的数据质量.

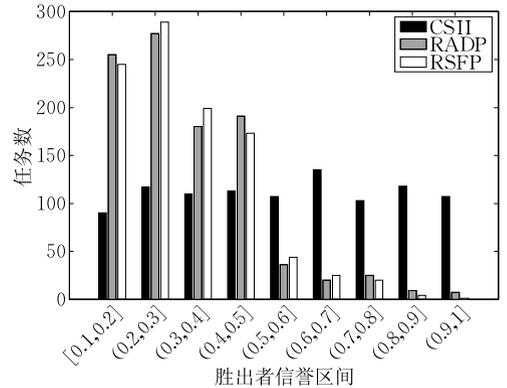


图 9 胜出者信誉

### 5.3.3 任务完成率

本文从 110306 条签到记录中随机选出 70306 条作为 LBSN 的训练数据集, 剩余的 30000 条作为测试数据集, 在训练数据集的基础上进行群智感知任务分发并进行 CSII 激励. 本文假设, 如果选择的工作者在测试数据集中的签到记录符合任务需求, 那么认为该工作者能够完成任务, 否则认为该工作者在任务发布之后没有在任务需求的时间地点进行签到, 从而不能完成任务. Baseline 则是在训练数据集中随机选择工作者执行任务, 并在测试数据集中进行测试. 图 10 显示了 1000 次任务中 CSII 和 Baseline 两种机制下工作者任务完成率的情况统计, CSII 中 90.3% 的任务其完成率介于 0.8~1 之间, 而 Baseline 中 61.9% 的任务其完成率小于 0.6. 因此, 采用 3.3 节的参与者评估模型选择工作者相对随机选择工作者能够提高任务的完成率.

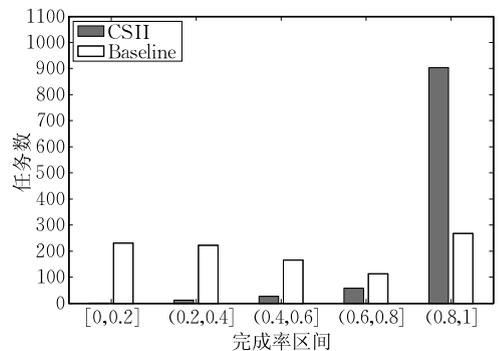


图 10 任务完成率

### 5.3.4 发布者的任务开销

根据式(16)可计算 CSII 下的任务开销, CSII 和 RADP、RSFP 的任务开销如图 11 所示. 从图中可以看出, 1000 次任务中, RADP 和 RSFP 任务开销大部分处于 0.1~0.8 之间, 而 CSII 中超过 350 个

任务的开销大于 1.5. 在 CSII 任务预算的建议下, 发布者的任务开销为任务评估模型得到的任务代价, RADP 和 RSFP 分别为胜出者的出价和固定价格. 和 RADP、RSFP 两种支付策略相比, CSII 的全支付策略具有较高的任务开销. 但是如式(16)所示, CSII 的支付方案并没有超出发布者的预算, 仍然在可控的支付范围内. 同时, 虽然相对 RADP 和 RSFP 发布者在 CSII 激励模型下需要支付更多的任务报酬, 但根据 5.3.1 节和 5.3.2 节的分析, 该支付方案具有更高的工作者收益和用户参与度, 而且相比 RADP、RSFP 两种方案能够得到更高的数据质量.

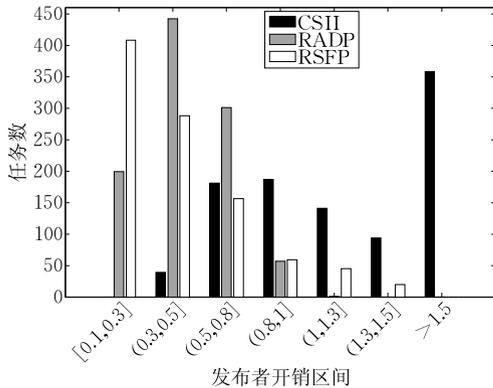
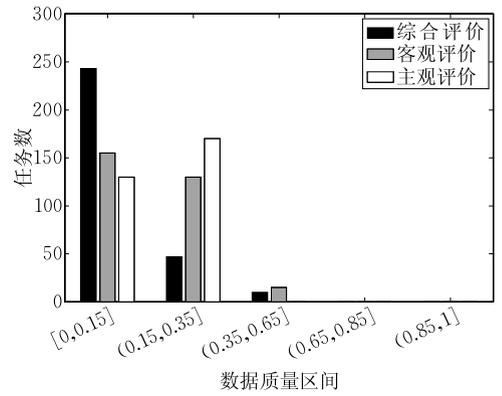


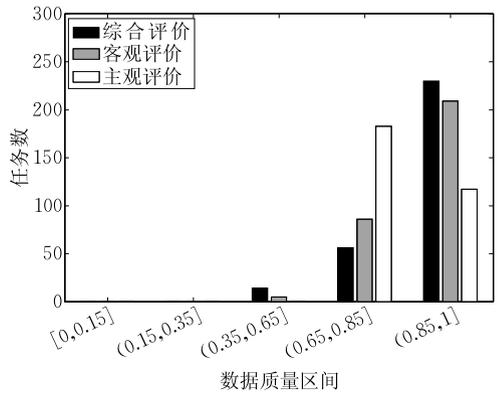
图 11 发布者开销

5.3.5 模糊逻辑性能

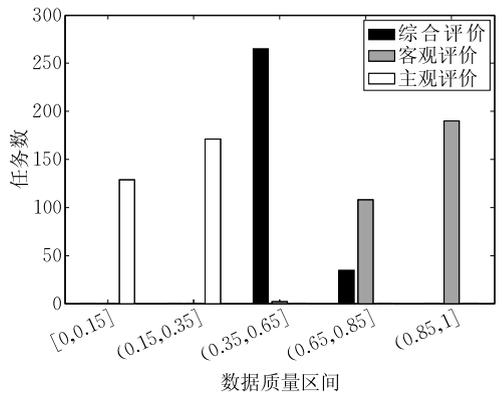
本文采用 3.4.1 节的数据质量评估系统从任务的及时性、覆盖率、相关性 3 个方面对工作者提供的任务质量进行了评价. 实验采用了 3 种方法进行对比, 综合评价指利用 3.4.1 节的模糊推理系统从及时性、覆盖率、内容相关性 3 个方面对数据进行综合分析; 客观评价指利用模糊逻辑系统从及时性、覆盖率两方面进行客观分析; 主观评价指分析数据的内容相关性. 本文按照 3.4.2 节定义的输入、输出变量的模糊集合进行区间划分, 图 12 分别显示了及时性、覆盖率、相关性取不同区间时 3 种评价方法下的数据质量结果. 实验中对这 3 个参数进行了不同组合, 即时间因子、地点因子以及分数的语言值变量分别取“L”和“H”两种情况下的组合, 由于时间因子和地点因子同时影响客观评价的结果, 因此时间因子和地点因子同时变化, 那么根据时间因子、地点因子和分数的取值不同可得到 4 种组合方式:  $\{L, L, L\}$ ,  $\{L, L, H\}$ ,  $\{H, H, L\}$ ,  $\{H, H, H\}$ . 当时间因子、地点因子的语言值为“L”时, 式(7)中时间变量  $(t-d)$  和式(10)中的地点变量  $r$  的取值均匀分布于  $[1.5, 3]$ ; 当语言变量取“H”时,  $(t-d)$  均匀分布于  $[-0.5, 0.75]$ ,  $r$  均匀分布于  $[0, 0.75]$ . 分数的语言变量取



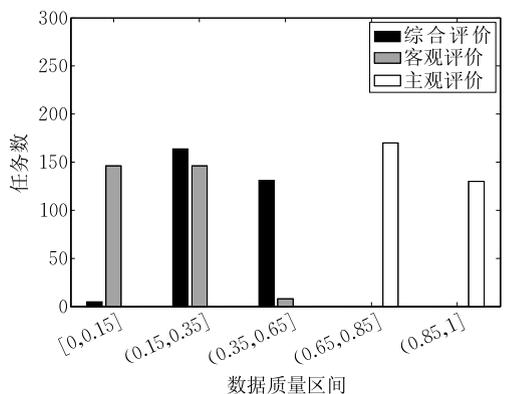
(a) 变量取值为  $\{L, L, L\}$



(b) 变量取值为  $\{H, H, H\}$



(c) 变量取值为  $\{H, H, L\}$



(d) 变量取值为  $\{L, L, H\}$

图 12 时间因子、地点因子、分数不同组合下的数据质量

“L”时其数值型变量取值均匀分布于 $[0, 0.35]$ , 取“H”时数值型变量取值均匀分布于 $[0.65, 1]$ . 从图中可以看出, 当3个参数都为“L”时的综合评价比主观评价和客观评价得到的数据质量低(图12(a)), 都为“H”时综合评价比客观评价和主观评价得到的数据质量高(图12(b)). 在图12(c)和(d)两种组合下综合评价得到的数据质量的语言值为“M”. 因此主客观评价相融合的方式从数据采集的时间地点和任务内容相关程度方面对数据进行评价, 更加客观地描述了数据的质量, 增加了数据的可信度.

#### 5.4 多元交互激励机制讨论

CSII 通过跨空间数据融合和多元交互提高了群智感知系统中用户参与度和数据质量. 发布者、系统、工作者这3个交互元在各个阶段的交互过程中相互协作共同完成群智感知应用中任务的发布、执行等过程. 但由于处于研究的初级阶段, CSII 模型仍然存在一些不完善的方面, 目前提出的方案仅作为面向群智感知系统提高用户参与度和数据质量的参考方案, 仍然需要在现有基础上对激励机制模型进行更加深入的探索研究.

(1) 任务评估. 在任务评估模型中, CSII 只考虑了任务的时空特性, 从任务的区域热度和时间热度两方面进行了任务代价评估, 该模型目前只针对简单的单任务, 而且在任务评估时, 忽略了任务类型(例如图像、音频、视频、文本等)对任务代价的影响. 在将来的工作中, 将考虑任务类型在任务评估模型中所起到的作用.

(2) 参与者评估. 在物理世界中, 群智感知系统中的激励机制模型由多个因素决定, 例如时空特性、LBSN 线上参与者之间的好友关系、天气等环境因素以及参与者的情绪影响等, 将来研究会结合参与者的历史数据采集行为、参与者的采集行为相似度、好友关系、参与者采集的数据和历史数据的相似度、任务的复杂程度、任务的采样数量等因素进行改进.

(3) 拍卖支付. 参与者在逆向拍卖中会因为没有任何任务补偿而在下次的任务分配中拒绝参与任务. 为了防止用户流失, CSII 在报酬支付时采用全支付的策略, 即所有参与任务采集的工作者都给予一定的任务奖励. 虽然这种支付方案一定程度上能够提高工作者参与任务的积极性, 但可能会造成一部分工作者采集较低质量的数据, 甚至采集不相关的数据. 针对该问题, 在将来的工作中会根据工作者采集

数据的质量进行数据优选, 并且只支付一部分的工作者, 即采集高质量数据的工作者, 对于采集低质量数据的工作者, 或接受任务但没有执行任务的工作者采用惩罚机制以提高数据的质量.

(4) 任务类型支持. 目前 CSII 针对的是不可分割的简单任务, 只面向局部粗粒度的任务, 并没有全覆盖所有的任务类型, 因此将来 CSII 模型将考虑分布复杂的多类型任务, 例如需求者需要采集某时间段内多个区域的噪声情况, 那么对于这种需要多人协作才能完成的复杂任务, CSII 需要将其分解成简单任务进行处理. 然而, 将简单任务融合成复杂任务的过程中, 系统需要将多个工作者的数据从多角度多方位进行融合, 从而得到全面完整的数据集合. 在数据选择和融合后, CSII 需要根据不同工作者的任务量、任务难度等因素支付给工作者相应的任务报酬. 但目前 CSII 的支付阶段只选择一个胜出者, 这对于需要工作者相互协作才能完成的复杂任务来讲, 需要进行改进.

(5) 工作者选择. 在任务评估和参与者评估模型中, CSII 利用参与者的历史 LBSN 签到数据而非实时的用户位置信息, 这种评估模型必然存在一定的误差, 因此, 在将来的工作中需要采用一定的算法降低历史信息 and 实时信息间的误差, 提高任务推送的准确率.

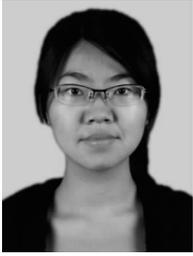
## 6 总结和展望

本文提出了基于跨空间的多元动态交互激励模型 CSII, CSII 在系统、发布者、参与者三者之间的动态交互基础上实现感知任务的分配和激励过程, 同时通过对跨空间的 LBSN 线上数据的分析实现任务评估和工作者评估. 在群智感知系统中, 利用 CSII 对参与者进行激励能够提高用户的参与度, 利用信誉模型提高发布者获取数据的质量. 同时, 该模型还能够有效地降低任务开销.

CSII 只研究了系统发布简单任务的情况, 认为复杂任务可由多个简单任务组合而成, 这适用于任务可组合的情况, 但对于更加复杂的不可组合的任务则没有考虑. 同时, CSII 模型没有考虑不同任务类型(例如噪声、图像、周围环境信息等)对任务难度的影响. 在将来的工作中, 针对以上问题将对 CSII 做出改进, 使其能够普遍应用于不同类型的群智感知系统.

## 参 考 文 献

- [1] Guo B, Yu Z, Zhou X, et al. From participatory sensing to mobile crowd sensing//Proceedings of the Pervasive Computing and Communications Workshops. Budapest, Hungary, 2014: 593-598
- [2] Liu Yun-Hao. Crowdsourcing computing. Communications of the China Computer Federation, 2012, 8(10): 38-41(in Chinese)  
(刘云浩. 群智感知计算. 中国计算机学会通讯, 2012, 8(10): 38-41)
- [3] Thiagarajan A, Ravindranath L, LaCurts K, et al. VTrack: Accurate, energy-aware road traffic delay estimation using mobile phones//Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems. Berkeley, USA, 2009: 85-98
- [4] Mohan P, Padmanabhan V N, Ramjee R. Nericell: Rich monitoring of road and traffic conditions using mobile smartphones//Proceedings of the 6th ACM conference on Embedded network sensor systems. Raleigh, USA, 2008: 323-336
- [5] Stevens M, D'Hondt E. Crowdsourcing of pollution data using smartphones//Proceedings of the Workshop on Ubiquitous Crowdsourcing. Copenhagen, Denmark, 2010: 1-4
- [6] Wu Wen-Le, Guo Bin, Yu Zhi-Wen. Crowd sensing based urban noise map and temporal-spatial feature analysis. Journal of Computer-Aided Design & Computer Graphics, 2014, 26(4): 638-643(in Chinese)  
(吴文乐, 郭斌, 於志文. 基于群智感知的城市噪音检测与时空规律分析. 计算机辅助设计与图形学学报, 2014, 26(4): 638-643)
- [7] Mun M, Reddy S, Shilton K, et al. PEIR, the personal environmental impact report, as a platform for participatory sensing systems research//Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services. Kraków, Poland, 2009: 55-68
- [8] Badanidiyuru A, Kleinberg R, Singer Y. Learning on a budget: Posted price mechanisms for online procurement//Proceedings of the 13th ACM Conference on Electronic Commerce. Valencia, Spain, 2012: 128-145
- [9] Zhao D, Li X Y, Ma H. OMG: How much should I pay Bob in truthful online mobile crowdsourced sensing? Correll University Library, New York, USA; Technical Report 1306.5677, 2013
- [10] Singer Y, Mittal M. Pricing mechanisms for crowdsourcing markets//Proceedings of the 22nd International Conference on World Wide Web. Seoul, Korea, 2013: 1157-1166
- [11] Zhang X, Yang Z, Zhou Z, et al. Free market of crowdsourcing: Incentive mechanism design for mobile sensing. IEEE Transactions on Parallel and Distributed Systems, 2014, 25(12): 3190-3200
- [12] Yang D, Xue G, Fang X, et al. Crowdsourcing to smartphones: Incentive mechanism design for mobile phone sensing//Proceedings of the 18th Annual International Conference on Mobile Computing and Networking. Istanbul, Turkey, 2012: 173-184
- [13] Koutsopoulos I. Optimal incentive-driven design of participatory sensing systems//Proceedings of the IEEE Conference on Computer communications. Turin, Italy, 2013: 1402-1410
- [14] Danezis G, Lewis S, Anderson R J. How much is location privacy worth?//Proceedings of the 4th Workshop on the Economics of Information Security. Berlin, Germany, 2005: 1-13
- [15] Lee J S, Hoh B. Dynamic pricing incentive for participatory sensing. Pervasive and Mobile Computing, 2010, 6(6): 693-708
- [16] Jaimes L G, Vergara-Laurens I, Labrador M A. A location-based incentive mechanism for participatory sensing systems with budget constraints//Proceedings of the IEEE International Conference on Pervasive Computing and Communications. Lugano, Switzerland, 2012: 103-108
- [17] Ho C J, Vaughan J W. Online task assignment in crowdsourcing markets//Proceedings of the Association for the Advancement of Artificial Intelligence. Toronto, Canada, 2012: 45-51
- [18] Shannon C E. A mathematical theory of communication. ACM SIGMOBILE Mobile Computing and Communications Review, 2001, 5(1): 3-55
- [19] Yang Hao-Fang, Zhang Jing-Lan, Roe Paul. Using reputation management in participatory sensing for data classification. Procedia Computer Science, 2011, 5: 190-197
- [20] Amintoosi H, Kanhere S S. A reputation framework for social participatory sensing systems. Mobile Networks and Applications, 2014, 19(1): 88-100
- [21] Reddy S, Burke J, Estrin D, et al. A framework for data quality and feedback in participatory sensing//Proceedings of the 5th International Conference on Embedded Networked Sensor Systems. Sydney, Australia, 2007: 417-418
- [22] Mamdani E H, Assilian S. An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Studies, 1975, 7(1): 1-13
- [23] Zheng Ya-Lin, Huang De-Long, Guo Jian. Mamdani algorithm of Fuzzy reasoning. Journal of Baoji College of Arts and Science (Natural Science), 2001, 21(3): 168-173 (in Chinese)  
(郑亚林, 黄德隆, 郭健. Fuzzy推理的Mamdani算法. 宝鸡文理学院学报: 自然科学版, 2001, 21(3): 168-173)
- [24] Reddy S, Estrin D, Hansen M, et al. Examining micro-payments for participatory sensing data collections//Proceedings of the 12th ACM International Conference on Ubiquitous Computing. Copenhagen, Denmark, 2010: 33-36



**NAN Wen-Qian**, born in 1989, M. S. Her research interest is pervasive computing.

**GUO Bin**, born in 1980, Ph. D., professor. His research interests include pervasive computing and mobile crowd sensing.

## Background

This research is partly supported by the National Basic Research Program (973 Program) of China under Grant No. 2015CB352400, the National Natural Science Foundation of China under Grant Nos. 61332005, 61373119, 61222209, the Program for New Century Excellent Talents in University under Grant No. NCET-12-0466, and the Scientific and Technology New Star of Shaanxi Province under Grant No. 2014KJXX-39.

The emergence of crowd sensing service presents us huge opportunities to publish tasks which can be performed by large amount of participants. Whether there are adequate users to participate is a critical influence on the quality of the crowdsensing system. However, some factors like privacy threat, smartphone power and 3G dataflow would hinder the large-scale deployment of the crowdsensing system. To overcome the problem, incentive mechanism is proposed to motivate participating, with which the participants would be highly interesting in participating in crowdsensing systems.

**CHEN Hui-Hui**, born in 1978, Ph. D. candidate. Her research interests include pervasive computing and mobile crowd sensing.

**YU Zhi-Wen**, born in 1977, Ph. D., professor. His research interests include pervasive computing and social sensing computing.

**WU Wen-Le**, born in 1990, M. S. His research interest is pervasive computing.

**ZHOU Xing-She**, born in 1955, professor. His research interests include embedded computing and pervasive computing.

Many works have been done on incentive mechanism for crowdsensing systems. These works have well solved this problem by proposing different algorithms. However, these works do not consider the characteristics of crowd sensing: spatio-temporal and user-centric. Given the characteristics, we propose a novel incentive mechanism that is Cross-Space multi-Interaction-based dynamic Incentive mechanism (CSII). Two contributions about the CSII system are as follows. Firstly, cross-space data about online and offline is merged to select suitable workers to perform the sensing task. Secondly, interactions between the entities in the system are considered to improve the quality of sensing data. What's more, in CSII system, the sensing task is evaluated to get the value of the task and the participants are evaluated in selecting suitable workers with cross-space data. Given the quality of the sensing data, the reputation of the workers is also involved. Hence, CSII system can provide an effective incentive mechanism to motivate participating and ensure the quality of sensing data.