Inline Text Wrapping Picture

北京邮电大学

硕士研究生学位论文开题报告

学 号: 2013111243

姓 名: 王东升

学 院: 网络技术研究院

专业(领域): 通信与信息系统

研究方向: 信息理论与信息处理

导师姓名: 龚向阳

攻 读 学 位: 工学硕士

2014年12月2日

|  |  |  |  |
| --- | --- | --- | --- |
| 论文题目 | 基于参与式感知的数据处理及激励机制的研究与实现 | | |
| 选题来源 | 国家自然科学基金项目 | 论文类型 | 综合研究 |
| 开题日期 | 2014-12-03 | 开题地点 | 新科研楼431 |
| **一、立题依据（包括研究目的、意义、国内外研究现状和发展趋势，需结合科学研究发展趋势来论述科学意义；或结合国民经济和社会发展中迫切需要解决的关键科技问题来论述其应用前景。附主要参考文献目录）（不少于800字）**  近年来，随着通信技术的发展和硬件水平的提升，移动智能终端（如智能手机、平板电脑等）日益普及，这些智能终端拥有强大计算能力，集成了大量先进的传感器（如重力感应、方向感应、GPS等）。在这样的背景下，为了更好地完成感知任务，研究人员提出了一种通过不同个体共同完成感知任务的感知数据收集方式——移动参与式感知（Mobile Participatory Sensing）。与传统的传感网工作方式不同，移动协作感知不需要预先部署特定传感设备，它将普通手持/车载智能终端用户设备所嵌入的多种传感器作为基本感知单元，保证了信息获取的多元化与及时性、实现了感知任务（来源于任务需求方）与感知动作（来源于任务参与者）的充分互动、感知数据的有效收集和分析，进而形成具有明显交互特性的“参与式感知网络”。利用有意识或无意识的用户参与，感知系统具有了覆盖面广、数据丰富、不需要额外硬件投资等优点，能够向基于移动协作感知网络的应用提供全面、便捷、深度的数据，在完成大规模、复杂的社会感知任务方面拥有广阔的应用前景。  参与式感知技术与固定部署的传感网技术相比具有诸多优点：1 ）部署成本极低；2 ）移动终端用户的固有流动性决定了参与式感知史无前例的时空覆盖范围；3 ）服务开发部署难度低，便于大规模部署；4 ）用户不仅是参与式感知系统的使用者，同时也是内容和服务的提供者，更是最终的受益者。  但在相关系统实际部署和运行的过程中依然面临着巨大的挑战与不足。由于参与式感知系统依赖于志愿者来贡献数据，所以感知数据只会出现在志愿者所出现的时间和空间。这就会造成感知数据分布的不均匀性。同时，参与式感知系统上的应用高度依赖数据的质量，但是平台的开放性使得感知数据质量参差不齐。所以，能辨别出数据已损坏或是恶意数据，对参与式感知平台服务器来说至关重要。  由于移动节点的不确定性，在数据收集过程中会受到参与者的影响，参与者不同的文化、教育背景等对感知数据收集在精度、可信度、时延等方面有影响，并且与参与热情有关。如何通激励机制提高用户积极性、增加数据数量、提高数据质量以及评定用户的信誉成为了这一课题迫切需要解决的问题。  近年来，国外已经涌现出大量关于参与式感知技术的应用和研究。国际比较著名的参与式感知研究小组包括美国麻省理工大学SENSEable City Lab、加州大学洛杉矶分校CENS Lab、美国达特茅斯大学Mobile Sensing Group、美国哥伦比亚大学Sensor Networks Group、IBM 研究院、微软研究院等，对参与式感知进行了初步研究。  我国国民经济和社会发展十二五规划中强调要加大环境保护力度，建立健全公共服务体系等。考虑到我国城市环境中具有人口密度大和移动设备普及度高的特点，参与式协作感知技术部署成本极低，在环境保护、公共设施服务、提升全民健康水平和灾难救援等方面有着广泛的应用前景。例如在公共卫生领域，通过收集参与者的身体状况数据监控恶性传染病的传播人群；交通运输领域，通过参与者收集的实时车速信息等，监控城市道路交通状况，为交通管理部门的决策提供重要依据；环境保护方面，收集参与者上传的环境感知数据或图片，获取城市细粒度实时空气质量信息；灾难救援中，通过参与者提供的受灾情况照片及位置信息等，分析灾害在不同区域所造成的损失，为灾难救援决策的制定提供指导。因此,本课题的研究工作将具有重要的理论意义和实用价值。 | | | |

|  |
| --- |
| **二、研究内容和目标（说明课题的具体研究内容，研究目标和效果，以及拟解决的关键科学问题。此部分为重点阐述内容）（不少于2500字）** |

|  |
| --- |
| **三、研究方案设计及可行性分析（包括：研究方法，技术路线，理论分析、计算、实验方法和步骤及其可行性等）（不少于800字）**  研究方法：  本课题首先通过文献研究为基础，研究相应的数据融合方案，随后以实验研究为基础，实现数据融合方案的验证，并对该方案进行改进，同时实现参与式感知平台数据融合，结合激励机制改进数据分布，提升数据质量，校验数据融合推测的结果。  技术路线：  1.感知数据的处理、融合、预测  a)首先在本课题依托的参与式感知平台实现用户系统，然后为用户建立起信誉档案，在数据清洗阶段，如果特定区域有冗余数据，对于用户每次提交的采集数据，与特定时空的其他数据平均值比较，偏差比例超过预定阈值的数据分别判为损坏数据或者恶意数据，并更新用户信用记录。数据融合阶段，分为信号级和特征级两个层次的数据融合。信号级数据融合类似于多传感器的感知数据融合，利用多用户上传数据融合出可信性更强、代表性更全面的数据，去除冗余信息；特征级的数据融合为更高一层的数据融合，根据不同种类数据的时空相关性，用感知数据融合出我们关注的知识，在本系统中要实现的是根据其他感知数据和获取的天气、交通等信息，融合、推测出细粒度的实时的PM2.5，拟在融合、推测的过程中利用机器学习的方法，建立PM2.5数据与其他类型的数据（时间、空间两类数据）训练出相关模型，推测缺失区域的PM2.5信息，完成对感知数据的信息挖掘。  2.激励机制的设计与实现  a)激励机制的验证  i. 研究需要验证已有的竞价激励机制。  ii. 首先，设置分组：价格通知优先、数据上传优先。价格通知优先又分为：竞价分组、固定价格分组。针对所有的场景，均对用户进行区分。其次需要招募志愿者，为每组分配相同的人数。然后通过事先为对不同分组设计不同的UI和用户权限，以及数据的采集模式和积分的反馈机制，从而使得不同的激励机制得到运行。  实验运行一段时间，通过对于数据的数量、种类、覆盖范围、及时性等评估数据质量的参数对每个分组的数据进行分析，从而得出个中激励方式对于数据质量的影响。  b) 激励机制对数据的调控  为实现激励机制对数据质量的调控，数据融合部分，需为激励机制部分提供接口。使得该模块能够获得服务器对于各地点和类型的数据的需求情况，通过为需求程度顶级，并且根据需求的级别选择不同级别的激励方式，以期达到能够通过大力度的激励获取最紧急的数据，实现对数据完整性的调控  可行性分析：  随着通信、微电子、集成电路和软件等技术的飞速发展，移动智能终端集成了丰富的计算，感知和通信能力。且随着智能终端成本和价格的逐渐降低，智能移动设备在人们生活中得到了广泛的普及。因此有大量的带有移动感知功能的设备，广泛的分布在城市中，为参与式感知提供了必要条件。本课题基于参与式感知概念，针对移动感知端数据的融合，包括感知数据与专业知识之间的相似度计算，检查原始数据中存在的错误和不一致。通过对原始数据的剔除和改正，提高数据的质量。研究将利用基于实例的机器学习方法，不断修改知识库，以提高数据清洗的准确度。研究还将采用数据挖掘技术，按照一定比例组合成新的特征，把各个模型在自身模态上的输出结果按一定方式进行融合，挖掘模态之间的相关性，提高融合技术的描述能力。参与式感知是移动感知研究领域近年来兴起的一个研究热点。但基于该研究领域的前期积累和本项目提出的合理技术方案，以及对关键研究问题给出的基本解决思路，我们认为本项目在技术上是可行的，我们有信心把这项研究工作做好。 |

|  |
| --- |
| **四、本研究课题可能的创新之处（不少于500字）**  本课题实现参与式感知平台采集的多种数据的处理、有效融合，利用机器学习的方式充分发掘其潜在价值，结合有效的激励机制对数据分布进行调控，提高感知数据质量，提升融合效果。本课题的主要创新点在于：  1.将传统的多传感器感知中的数据融合技术引入到参与式感知系统中，目前这种结合所做的工作还比较少，而且参与式感知应用类型广泛、灵活，感知数据类型和用途多样，需要寻找每种应用下的最优的融合方案；将机器学习的技术用来解决PM2.5推测 问题，把这一需要特定气象领域知识的问题用机器学习的典型学习算法解决。  2. 目前已有激励机制的验证试验场景设计简单，参与者较少，对比试验也比较简单。相对复杂的激励机制，大多停留在理论研究阶段，没有能够证明其作用的试验。其中存在竞价的激励机制缺少验证。本课题通过实现对该类型的验证试验，填补这部分实践空白，并且分析其实验结果；  现有激励机制对数据质量的研究，集中在不同的激励方式，对数据数据质量、数量、及时性、覆盖范围等条件的影响研究，并没有实现与服务器进行交互的功能。没能够达到通过与服务器端的实时交互，采用不同等级的激励方式控制采样的数据类型和采样地点的目的。本课题通过实现与服务器进行实时交互，从而获得服务器对于不同数据的需求情况，根据服务器的需求，给出用户不同力度的激励，使得用户能够上传服务器缺失的信息。从而达到平衡信息来源，辅助服务器提升数据可用性的目的。 |
| **五、研究基础与工作条件（1.与本项目相关的研究工作积累基础 2.包括已具备的实验条件，尚缺少的实验条件和拟解决途径）（不少于500字）**  与本项目相关的研究工作积累基础  1)实验室已经建立起参与式感知平台的基本框架，目前可以采集图片、GPS数据、噪音值和光照强度值，通信接口与缓存队列的基本功能已经实现；  2)调研参与式感知的激励机制的发展情况，并对其进行总结。调研其他系统中的激励机制的模式，为新型激励机制提供理论支撑和原始模型。已阅读参与式感知系统中激励机制相关论文10余篇，主要参考本实验室发表过的激励机制的论文Song, Zheng, et al. "Incentive mechanism for participatory sensing under budget constraints." Wireless Communications and Networking Conference (WCNC), 2014 IEEE. IEEE, 2014.  3)调研了数据融合的基本概念与应用领域，阅读了城市计算中关于气象估计与大数据结合的相关文献；平台已实现北京范围的天气数据和PM2.5值的抓取程序，每小时更新一次数据，为数据融合和推测做了必要的准备工作。  缺少的试验条件：可采集完整数据的可运行的移动终端应用，以及大量的可挖掘数据。  尚缺少的实验条件和拟解决途径：  客户端：本课题的工程任务即完成客户端部分的实现工作，由其他负责客户端工作的同学完成；  用户数据：1.强制征集组：通过组内成员强制征集数据，试运行系统，修改bug，并通过用户的反馈进行相应的完善；  2.非强制征集组：通过激励机制的改进，鼓励用户参与，征集志愿者，参与到数据采集。  对于数据融合阶段缺失的天气、车流量等数据，采取从其他服务平台（如百度地图等）获得或网络抓取的方式获得。 |

**学位论文工作计划**

|  |  |  |
| --- | --- | --- |
| 时间 | 研究内容 | 预期效果 |
| 2014.10~ 2014.12.31 | 阅读相关文献 | 完成开题工作，确定研究方向，提交开题报告 |
| 2015.1.1 ~ 2015.1.31 | 提出激励方案，仿真方案，根据效果改进方案 | 可以推送任务，根据方案给用户激励 |
| 2015.3.1 ~ 2015.5.31 | 实现感知数据的分布分析，进行集成和插值 | 对感知数据进行集成和插值 |
| 2015.6.1 ~ 2015.7.15 | 完成模块之间的联调 | 实现本人负责的激励和融合模块的功能 |
| 2015.9.1~2015.10.31 | 撰写学术论文 | 提交一篇能够被EI/ISTP检索的学术论文 |
| 2015.11.1~2016.1 | 整理文档，完成硕士研究生毕业论文，准备答辩 | 提交项目相关文档，完成硕士学位论文 |
| 2016.1~2016.3 | 答辩完成，整理工作内容 | 归档学术成果和工程实现成果，提交论文和相关资料 |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 评  定  小  组  成  员 | 姓 名 | 职 称 | 单位名称 | 职务 |
| 龚向阳 | 教授 | 北京邮电大学 | 组长 |
| 阙喜戎 | 副教授 | 北京邮电大学 | 成员 |
| 王文东 | 教授 | 北京邮电大学 | 成员 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| 导师意见： | | | | |
| 同意开题。 | | | | |
| 导师（签名）：  日期： 年 月 日 | | | | |
| 开题报告小组意见： | | | | |
| 组长（签名）：  日期： 年 月 日 | | | | |
| 学院意见（签章）： | | | | |
| 负责人：  日期： 年 月 日 | | | | |