

参与式感知系统中基于社会关系的移动节点位置预测算法

于瑞云¹⁾, 夏兴有¹⁾, 李婕²⁾, 周岩¹⁾, 王兴伟³⁾

¹⁾(东北大学 软件学院, 沈阳 中国 110819)

²⁾(东北大学 计算中心, 沈阳 中国 110819)

³⁾(东北大学 信息科学与工程学院, 辽宁 沈阳 110819)

摘要 移动节点位置预测是参与式感知系统进行有效数据采集和消息转发的关键, 本文提出了一种基于社会关系的移动节点位置预测算法 (SMLP)。该算法基于位置对应用场景进行建模, 通过节点的移动规律挖掘节点之间的社会关系。SMLP 算法以 Markov 模型为基础对节点的移动性进行初步预测, 然后利用与其社会关系较强的其他节点的位置对该节点的预测结果进行修正。算法基于马尔可夫模型和加权马尔可夫模型进行了优化, 分别提出了 SMLP₁ 和 SMLP_N 两种算法实现。最后基于 UCSD WTD 数据集对算法进行仿真实验, 实验结果表明, SMLP₁ 比马尔可夫模型有更高的预测精确度; SMLP_N 与 SMLP₁ 相比有了更大程度的性能提升, 并以比 2 阶 Markov 模型小得多的算法复杂度获得了与 2 阶 Markov 模型相似的预测精度。由于加权系数的引入, SMLP₁ 和 SMLP_N 两种算法都具有良好的可扩展性。

关键词 参与式感知; 位置预测; Markov 模型; 社会关系

中图法分类号 TP393.0

Social-Relationship-Based Mobile Node Location Prediction Algorithm in Participatory Sensing Systems

YU Rui-Yun¹⁾, XIA Xing-You¹⁾, LI Jie²⁾, ZHOU Yan¹⁾, WANG Xing-Wei³⁾

¹⁾(Software College, Northeastern University, Shenyang 110819, China)

²⁾(Computing Center, Northeastern University, Shenyang 110819, China)

³⁾(College of Information Science and Engineering, Northeastern University, Shenyang 110819, China)

Abstract Mobile node location prediction is critical to efficient data acquisition and message forwarding in participatory sensing systems. This paper proposes a Social-relationship-based Mobile Node Location Prediction algorithm (SMLP). The SMLP algorithm models application scenarios based on geographic locations, and extracting social relationships of mobile nodes from nodes' mobility. The SMLP algorithm preliminarily predicts node's mobility based on the Markov model, and then amends the prediction results using location information of other nodes which have strong relationship with the node. Two algorithms, SMLP₁ and SMLP_N, are proposed based on the Markov model and the weighted Markov model, respectively. Finally, the UCSD WTD data set are exploited for simulations. Simulation results show that SMLP₁ acquires higher prediction accuracy than the

本课题得到国家自然科学基金面上项目 (No. 61272529); 国家杰出青年科学基金项目 (No. 61225012, No. 71325002); 教育部-中国移动科研基金 (No. MCM20130391); 高等学校博士学科点专项科研基金 (优先发展领域) 项目 (No. 20120042130003); 教育部中央高校基本科研业务费项目 (No. N120417002, No. N110204003, No. N120104001, No. N130817003) 资助。于瑞云, 男, 1974年1月生, 博士, 副教授, 主要研究领域为参与式感知系统、无线传感器网络、普适与移动计算、移动网络大数据分析等, E-mail: yury@mail.neu.edu.cn。夏兴有, 男, 1988年7月生, 博士研究生, 主要研究领域为参与式感知系统、移动计算等, E-mail: neuxiaxy@163.com。李婕, 女, 1982年10月生, 博士, 讲师, 主要研究领域为认知网络、移动计算等, E-mail: lijie@mail.neu.edu.cn。周岩, 男, 1988年2月生, 硕士研究生, 主要研究领域为参与式感知技术、移动计算等, E-mail: zhouy@mail.neu.edu.cn。王兴伟 (通信作者), 男, 1968年生, 博士, 教授, 博士生导师, 主要研究领域为未来互联网、云计算、网络安全和信息安全, E-mail: wangxw@mail.neu.edu.cn。

Markov model. $SMLP_N$ achieves more accuracy on prediction compared with $SMLP_1$, and obtains comparable prediction accuracy with order-2 Markov model while presents extra lower algorithm complexity. Due to the introduction of weighting coefficients, both $SMLP_1$ and $SMLP_N$ demonstrate high flexibility.

Key words Participatory sensing; Location prediction; Markov model; Social relationship

1 引言

参与式感知 (Participatory sensing) 是最近几年出现的一种感知技术, 又称以人为中心的感知 (People-centric sensing)。参与式感知利用多种数据采集设备和集成了特定传感器的智能手机对人类社会状态信息 (包括环境、交通、社会活动等) 进行交互式或自主式采集、分类、传输和分析, 进而做出智能决策, 为人类生活和社会活动提供服务。

美国加州大学洛杉矶分校的 Deborah Estrin 教授所领导的 CENS (Center for Embedded Networked Sensing) 研究中心是最早开展这方面工作的研究机构之一, 从 2009 年以来, 该中心已经将主要精力投入到参与式感知技术研究中, 在 CENS 研究中心 2011 年年度进展报告^[1]的 Research 部分, 参与式感知方面的内容占了将近一半的篇幅, 其主要介绍了 CENS 实验室在参与式感知研究领域的最新进展, 包括 PEIR、AndWellness、Mobilize、Boyle Heights 等几个典型应用项目的研究进展, 以及支撑参与式感知应用的相关技术的研究情况, 包括参与者招募、任务规划、基于手机传感器的人类活动识别、用户隐私等等; 美国达特茅斯学院的 MSG (Mobile Sensing Group) 实验室也在这方面做了很多研究, 主要项目有 BikeNet^[2]、SoundSense^[3]、CenceMe^[4,5]、MetroSense^[6]、Bubble-sensing^[7]等。

参与式感知强调感知过程中人的参与, 人们对赖以生存的城市和社会状态进行感知、记录, 一个人感知的信息或者群体感知的经过融合处理后的信息可以为其他人或群体所用。美国南加州大学的 ENL (Embedded Networks Laboratory) 实验室以此为背景开发了 Urban Tomography^[8,9]平台, 通过手机进行基于位置信息的移动影像的采集, 获取城市生活的声音影像记录, 揭示生活状态模式。LiveCompare^[10]和 Micro-Blog^[11]研究了参与式感知与社交网络的结合应用, LiveCompare 利用手机中的图像传感器获取特定商品在不同卖场的价签信息, 经过处理、分析、位置标定后上传到网站, 为用户提供实时的商品特价信息; Micro-Blog 讨论了社交网络、分布式内容共享与参与式感知相融合的体系框架, 提出了一个基于参与式感知的微博系统

原型。

利用智能手机结合多种类型的传感设备 (如加速度传感器、计步器、光线传感器、温湿度传感器、气体传感器、磁力传感器、二氧化碳传感器、PM2.5 传感器等), 可以构造多种基于参与式感知的应用系统, 比如生活模式感知、移动社交、空气污染实时监测、交通状况监测等等。美国杜克大学的 Martin Azizyan 等人在文献^[12,13]中提出用安装在手机上的光线、声音传感设备辅助 GSM、WIFI 进行相对精确的定位, 这种定位方式比 GPS 定位具有更小的开销, 并且能在 GPS 不能覆盖的室内环境下实现定位。美国麻省理工学院的 CarTel 项目^[14,15]利用安装在汽车和智能手机上的传感器采集数据, 并将数据上传到一个特定的 WEB 服务器, 然后在服务器对数据进行分析并为用户提供实时交通状况、路面状况信息。Ville Kotovirta^[28]等人通过试验对空气质量、水质和植物疾病等数据的监测, 研究了用户激励以及参与用户的隐私保护等问题。R. Rana^[29]等根据噪声污染可以影响人类的听觉和行为的特点, 利用参与式感知节点来监控噪声等级, 并建立噪声污染地图系统, 将信息提供给专家来研究噪声和人类行为之间的关系。

在上述参与式感知应用中, 如何构造一个动态的群组以保证参与式感知任务高效、高质量的完成, 是参与式感知研究领域一个有待解决的问题: 当系统中参与者数量充足时, 系统需要通过定价与博弈的方式选择更加廉价而高效的节点去执行任务; 而对于参与者数量不足的情况, 系统需要制定相关的激励机制去刺激用户的参与。

对于与位置相关的参与式感知应用而言, 系统一般会对用户的位置数据进行采集, 但在某些时刻, 系统中往往会出现感知区域用户数量覆盖不足的情况 (当前驻留在目的位置的用户数量很少, 即使全部参与, 也可能不足以支撑感知任务的完成), 而这种情况有时会直接影响感知任务顺利的完成。对于此类问题, 我们可以利用用户的移动轨迹预测, 配合着良好的用户激励机制提供出一种有效的解决方案: 如果系统在获取用户当前位置信息的同时还会预测用户将来一段时间可能到达的位置, 则可根据情况组建一个最有可能到达感知位置的用户群组, 并通过相关机制激励其前往目的地并完成任务, 以

此来增加参与者数量,提高应用的执行效率。例如,如果希望利用参与式感知系统查询某工厂附近PM2.5颗粒物分布情况(需要多个安装了PM2.5传感器的参与式感知节点完成此任务),我们首先可以通过位置预测估计将会有哪些节点在未来的一段时间内运动到这里,然后系统会向这些节点发布参与请求,提供相关奖励,促使其主动的前往相关区域,完成采集应用,这样既保证了区域的覆盖,又要避免太多的节点传回无用数据。与之相类似的还有某风景区各个不同角度的风景图片和游人数量查询(合适的游客群组通过带拍照功能的智能手机发回符合要求的图片和游客数量信息)、交通拥挤情况查询等,在这些应用中,我们并不能保证当前时刻参与节点能够将监测区域全部覆盖,因此需要通过位置预测,选择可到达的节点执行采集任务,以保证应用的顺利完成。

与传统的移动应用相比,参与式感知应用有如下特点:

(1) 将应用延伸到了自然界和物理世界,通过参与式感知设备(智能手机、传感器节点等)直接完成从物理世界中获取状态信息(声音、温度、位置、光照、物体运动状态等)的任务;

(2) 参与式感知应用可以将任务分解成多个单元,由多个参与式感知设备分别完成各自任务单元,共同完成一个整体任务(例如由多个感知半径有限的节点共同完成一个较大区域的数据采集任务,而在传统移动应用中通常多个个体完成一个相同的任务并竞争最佳方案);

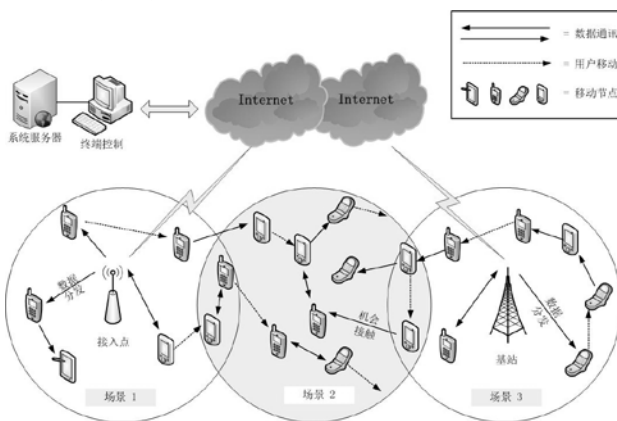


图1 参与式感知系统通信模式示意图

(3) 参与式感知系统具有机会网络特性。对于传统移动应用而言,节点一般是通过接入点或基站的方式与中心服务器进行通信(如图1中场景1和场景3所示);而在参与式感知系统中,由于节点移动、节点稀疏、射频关闭或障碍物造成信号衰减等多种原因,导致网络有时不能完全连通,移动设备

常采取一种“弱连接”的方式联网,因此某些情况下节点需要通过用户移动带来的相遇机会来完成数据的通信和传输(如图1中场景2所示)。

针对参与式感知环境中节点的机会连接问题,移动设备的位置预测不仅能够提高机会网络中数据的转发效率,减少通信时延(很多机会网络路由协议都是基于位置预测来设计的),同时还有助于充分利用系统中的一些“弱联网”节点,通过对其位置的判断,完成一些特殊的参与式感知应用^{[31][32]}。

为了解决移动预测等问题,近年来,许多科研机构 and 学者针对人类的移动模式展开了研究^[35-38]。文献[35]根据移动设备的历史位置信息研究人类的行动轨迹,作者通过6个月来对100000名参与者位置信息的调查,发现人类的移动轨迹符合一种空间概率分布,可以利用某种简单且重复的模式来描述。文献[36]对这种移动模式的可预测性进行了论证,作者通过引入“熵”的概念来判断用户移动轨迹预测的可能性,当使用所有的历史数据进行测试时,用户位置的潜在可预测性将高达93%。

基于上述理论,学者们利用人类的移动模式,针对轨迹的预测问题提出了一些不同的解决方案^{[19][20][25][33][39]}。目前应用最广泛的位置预测方法为基于马氏链(Markov chain)的方法:该方法的基本思想是建立一个 k 步Markov预测器,即是一个与一阶Markov预测器的移动概率矩阵同规模的移动概率矩阵。通过这个矩阵,就可以根据用户 k 步以前的位置来预测其下一步即将到达的位置。文献[19]利用无线校园网络移动数据^[33]来检验一些位置预测算法的性能,在文章中作者使用Wi-Fi信息定义了一些离散的位置,并据此对 k 阶Markov预测模型、基于LZ树的LZ-based预测模型等不同的方案进行了测试。经过比较发现,相对简单的2阶Markov模型拥有和其他方案相同甚至更好的预测效果,预测的平均准确率达到了75%。类似效果较好的预测方法还有文献[20]提出的扩展后的Markov预测模型,该模型考虑了到达时间和停留时间,使用延时嵌入法(time delay embedding)从时间序列中提取确定长度的位置序列,通过将上一次观察到的位置与所有嵌入的位置序列进行对比,以此来完成用户位置信息的预测。

除了马氏链外,还有一些位置预测方法如基于人工神经网络(Artificial Neural Network)的方法^[16]、基于贝叶斯网络(Bayesian Network)的方法^{[17][18]}和基于回归的方法^{[21][22][23]},这些方法从不同的角度对未来位置预测做了研究,主要针对的是单一节点的移动行为进行预测。但事实上,节点的行为状

态不仅取决于前一时间段的位置状态信息,还取决于携带移动设备的主体之间所具有的社会性关联。

综上所述,移动节点位置预测对参与式感知应用而言具有重要意义。在本文中,我们利用节点间的社会关系,对已有的位置预测算法进行优化,使其能够更好的适应参与式感知系统应用,获得了更高的预测准确率。

2 基于社会关系的节点位置预测算法

人类移动模式研究^[35]表明,人们的行动存在高度的重复性,每天人们重复访问固定的几个地点,在相对固定的时间内进行一些日常活动。根据移动节点的位置以及固定的行为模式,我们可以基于位置对场景进行建模,利用相关算法预测节点到达某一位置的概率,估计移动节点的位置。本文提出了一种基于社会关系的移动节点位置预测算法(Social-Relationship-Based Mobile Node Location Prediction Algorithm, SMLP),分别针对基于绝对分布的马尔可夫模型和加权马尔可夫模型进行了优化。算法以 Markov 模型为基础对节点的移动性进行初步预测,然后利用与其社会关系较强的其他节点位置对该节点的预测结果进行了修正,以获得更高的准确性。

2.1 基于位置的节点移动应用场景

目前大部分的机会网络研究都是针对于移动节点的,其目的是将信息通过机会连接的方式发送到目的节点上去。但由于节点的高度移动性以及接触时间的不确定性,这些研究很少在实际场景中得到应用。针对这一问题,文献[24]提出了一种基于位置的机会网络路由传输策略。Shanshan Lu 等人认为,节点在移动过程中接触时间较短,而在一些固定地点(如办公室、咖啡厅、健身场所等)接触时间较长,因此针对位置的消息传递策略相对于传统传输策略而言更加稳定,具有更高的实际应用价值。

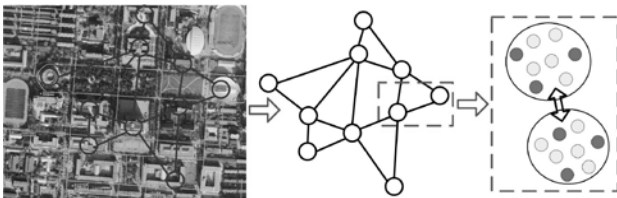


图2 基于位置的移动行为建模示意图

基于这一思想,我们提出可以基于位置对节点移动应用场景进行建模。如图2所示,每个应用场景都存在很多用户分布密集的地点,节点在这些地点进行汇聚,基于位置形成空间分散的群组。不难想象,在同一群组中的节点与处于不同群组中的节

点相比,接触更为稳定、频繁。另外,一些用户还会按照自己的日常行程从一个地点运动到另外一个地点。群组间可以利用这些移动的节点进行连接,机会性地进行通讯。

综上所述,与基于节点的模型相比,基于位置建模更有利于应用到实际场景中,主要原因有如下两点:

(1) 基于位置的接触,时间更久,建立起来的链路更加稳定,因此有利于数据的传输与交互;

(2) 基于位置的通信可以支持更大的数据量,这更有助于应用的扩展。

我们在对应用场景基于位置进行建模,在此基础上收集移动节点访问不同位置的频数信息、接触信息及时间信息,利用节点的历史移动轨迹及社会关系来预测节点到达各个位置概率,以此作为参与式感知系统机会链接特性研究的基础。

通过对CRAWDAD公开发布的实验数据集统计得出,目前人们主要通过以下三种方式来记录节点的位置信息:(1) GPS。GPS是一种较为精准的外部定位技术,其在导航系统、车载防盗系统、公交调度系统等众多项目中都得到了广泛的应用。但对于移动手持设备而言,由于耗电、数据开销大、室内定位效果欠佳等原因,通过GPS所搜集的节点移动轨迹往往存在数据缺失、数据不完整等问题。

(2) 基站(Cell Tower)。通过运营商的无线电通讯网络(如GSM网、CDMA网),人们可以获取移动终端用户的位置信息,而这些信息一般是通过跟踪记录设备所访问的基站编号而获得的(例如MIT Reality项目^[34]中的节点位置数据)。(3) WLAN。利用部署在不同区域的AP节点,通过记录带有WiFi接口的手持设备与AP节点的通讯数据,人们同样可以完成对于移动节点的位置定位工作。经典的WLAN实验项目包括美国Dartmouth学院的Wireless trace datasets^[36]和UCSD的Wireless Topology Discovery (WTD)^[37]项目等。

对于基于位置的节点移动模型而言,CPS主要记录的是节点移动轨迹的经纬度信息,但由于其收集的历史记录存在数据缺失的问题,因此不利于移动模型的建立;而通过记录Cell Tower编号来跟踪节点移动轨迹,由于基站本身覆盖面积过大,其位置记录的粒度过粗,因此建立起来的移动模型应用

^[36] Wireless trace datasets.
http://crawdad.cs.Dartmouth.edu/data.php. 2006.

^[37] UCSD Wireless Topology Discovery Project.
http://www.sysnet.ucsd.edu/wtd/wtd.html, 2013

价值不高。最后，利用 WLAN 来完成节点历史轨迹的记录，由于 AP 部署较为灵活，且 AP 通讯范围适中，因此对于校园等场景而言，使用 AP 数据来建立位置移动模型更为合适。

在大部分 WLAN 数据集中，为了方便处理，每条数据是以 (node,time,AP List) 的格式记录的，表示节点在某时间与哪些 AP 相连。为了更好的构建移动模型，我们可以利用连通分量将 AP 数据转换为位置信息，具体过程如下：

(1) 统计不同 AP 同时出现在一条记录中的次数。通过对 AP 出现次数的计算，我们可以获得 AP 间的关系矩阵 $k = [r_{ij}]$, $i \in N, j \in N$, 其中 $r_{ij} = \frac{2n_{ij}}{n_i + n_j}$, n_{ij} 表示 AP_i 与 AP_j 同时在记录中同时出现的次数, n_i 表示 AP_i 出现的总次数, n_j 表示 AP_j 出现的总次数。

(2) 将 AP 作为图中的结点，按 r_{ij} 的大小顺序依次用直线将这些结点连接起来，并标上权重，这样，就得到一个描述 AP 关系的无向图，图 3 (a) 就是这样的一个 AP 关系图。

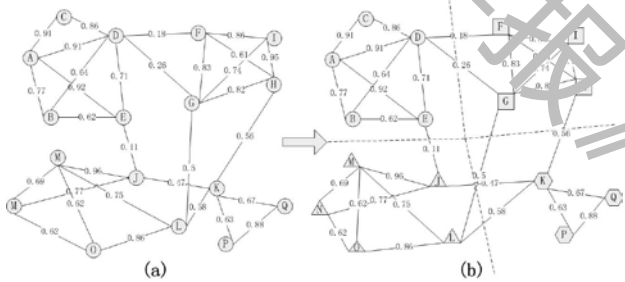


图 3 AP 位置划分

(3) 我们取定 λ ，去掉权重低于 λ 的连线，即可将结点分类，所获得的每个连通分量即为一个位置。图 3 (b) 就是当 λ 取值为 0.6 时的一个包含 4 个不同连通分量的位置示意图。

针对上述方法，阈值 λ 的选取决定了基于位置的节点移动模型位置区域粒度（每个聚类位置的覆盖区域范围）的大小：当 λ 取值较小时，每个位置包含了更多的 AP 节点，因此区域粒度较大，位置覆盖范围更加广泛；当 λ 取值较大时，每个位置包含了较少的节点个数，因此区域粒度变小，每个位置覆盖了较小的区域范围。针对不同的环境，我们可以选取不同的阈值来确定模型的位置区域粒度，以此来满足应用对位置范围的具体需求。

2.2 马尔可夫链预测模型

我们可以用马尔可夫模型来描述类似校园这样的应用场景：假设当前场景有 m 个地点，我们将地点 i 作为马尔可夫过程的第 i 个状态 X_i ，则状态空间为 $E = \{X_1, X_2, \dots, X_m\}$ 。由此，场景移动模型可

以定义为 $\{X, T\}$ ， T 为时间序列。

在参与式感知系统中，针对每个应用场景，我们可以利用马尔可夫链模型对各个节点未来的地点状态进行预测，具体的建模及预测过程如下：

(1) 准备过程

预测前的准备过程包括以下几个步骤：

① 确定状态集合

根据系统收集的用户行动轨迹，对数据中出现过的地点进行统计，记为集合 L 。由于集合 L 包含多个地点元素，所以选择访问频次较高的地点集作为系统的状态集合 $E, E \subset L$ 。

② 数据离散化处理

统计所有用户与状态集合 E 有关的数据，将每个用户的数据按照时间进行离散化，因此离散后的集合可以表示为

$$\{(t_k, X_i)\}, k = 1, 2, 3, \dots, i \in \{1, 2, 3, \dots, m\} \quad (1)$$

③ 计算一步转移概率矩阵

设 n_{ij} 是观测节点 A 由 i 位置出发访问 j 位置的次数，则节点由 i 位置出发到达 j 位置的概率为

$$p_{ij} = \frac{n_{ij}}{n} \quad (2)$$

其中 n 为节点由 i 出发访问各位置的次数总和。

因此，设集合中有 m 个地点，则可以产生一个 $m \times m$ 的转移概率矩阵

$$P = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1m} \\ p_{21} & p_{22} & \dots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \dots & p_{mm} \end{bmatrix} \quad (3)$$

(2) 基于绝对分布的马尔可夫链预测

设 $p_j^{(l)}$ 是节点初始时刻 t 处于 X_j 状态的概率，取各个状态的概率，则可获得马尔可夫链的初始分布，记为

$$P^{(l)} = (p_1^{(l)}, p_2^{(l)}, \dots, p_m^{(l)}) \quad (4)$$

例如取初始状态为 X_2 ， $m=5$ ，则初始分布 $P^{(l)} = (0, 1, 0, \dots, 0)$ ，则 $l+1$ 时刻的绝对分布

$$P^{(l+1)} = P^{(l)}P = (p_1^{(l+1)}, p_2^{(l+1)}, p_3^{(l+1)}, p_4^{(l+1)}, p_5^{(l+1)}) \quad (5)$$

可认为时刻 $l+1$ 时系统所取的状态 X_j 满足

$$X_j = \arg \max \{p_j^{(l+1)}\}。$$

(3) 加权马尔可夫链改进预测

基于绝对分布的马尔可夫链利用状态转换矩阵和初始分布进行预测，简单直观，但由于状态转移概率测定不准和系统初始状态划分不科学等原因，这种方法的预测结果往往会产生较大误差。为了减小误差，用各种步长的马尔可夫链加权和来改进预测方法^[27]。与普通的马尔可夫链预测相比，加权马尔可夫链可以更充分、合理地利用信息，并与相关分析有效结合，提高预测结果的准确率。

已知节点在 l 时刻的初始概率分布，我们可以多步预测递推 $l+t$ 时刻的概率分布：

$$P(l+t) = P(l+t-1) \cdot P = P(l)P^t = (p_1^{(l+t)}, p_2^{(l+t)}, p_3^{(l+t)}, p_4^{(l+t)}, p_5^{(l+t)}) \quad (6)$$

其中， $P(l)$ 是初始状态分布， P^t 为多步转移概率矩阵。

我们还可以将 $l, l+1, \dots, l+(t-1)$ 几个不同时刻的状态作为多步预测的初始状态，通过公式(6)，分别推算出 $l+t$ 时刻的概率分布，最后通过加权的方式获得最后的预测分布。

设初始状态集合为 $l = \{X^{(l)}, X^{(l+1)}, \dots, X^{(l+t-1)}\}$ ，

了描述方便，将 l 集合简记为 $l = \{X^{(k)}, k=1, \dots, t\}$ ，使用相关系数对各步结果进行加权，具体的加权公式如公式(7)所示。

$$P(l+t) = \sum_{k=1}^t w_k P(l+k-1)P^{t-k+1} = \{p_j^{(l+t)}\}, j=1, 2, \dots, m \quad (7)$$

其中， w_k 为相关权重， $w_k = |r_k| / \sum |r_k|$ ， $\sum w_k = 1$ 且 $w_k \geq 0$ ， r_k 为相关系数。

时刻 $l+t$ 所处状态 X_j 满足公式 (9)。

$$X_j = \arg \max \{p_j^{(l+t)}\} \quad (9)$$

举例来说，若系统用前3个状态 ($l, l+1, l+2$) 去预测下一个状态，根据式 (7) 进行加权计算，则节点位置预测过程的列表分析如表1所示，通过权重，初始分布和转换矩阵三个要素计算 $l+3$ 时刻的概率分布 $P(l+3)$ ，则节点在 $l+3$ 时刻所处状态即为

$$X_j = \arg \max \{p_j^{(l+3)}\}。$$

表1 加权马尔可夫链预测分析表

初始时刻	状态	权重	初始分布	转换矩阵	1	2	3	4	5
l	4	w_1	$P(l)$	P^3	$w_1 p_{11}^{(l)}$	$w_1 p_{12}^{(l)}$	$w_1 p_{13}^{(l)}$	$w_1 p_{14}^{(l)}$	$w_1 p_{15}^{(l)}$
$l+1$	1	w_2	$P(l+1)$	P^2	$w_2 p_{11}^{(l+1)}$	$w_2 p_{12}^{(l+1)}$	$w_2 p_{13}^{(l+1)}$	$w_2 p_{14}^{(l+1)}$	$w_2 p_{15}^{(l+1)}$
$l+2$	3	w_3	$P(l+2)$	P^1	$w_3 p_{31}^{(l+2)}$	$w_3 p_{32}^{(l+2)}$	$w_3 p_{33}^{(l+2)}$	$w_3 p_{34}^{(l+2)}$	$w_3 p_{35}^{(l+2)}$
$P(l+3)$					$p_1^{(l+3)}$	$p_2^{(l+3)}$	$p_3^{(l+3)}$	$p_4^{(l+3)}$	$p_5^{(l+3)}$

2.3 基于社会关系的预测优化

现实社会中，人与人之间具有社会关系，且社会关系有强有弱。如果用个体间的接触时间和接触次数来衡量节点间的社会关系强弱，即强关系的个体接触时间长，接触次数多；弱关系的个体接触时间短，接触次数少。在现实生活中，一个个体通常与少数个体具有较强的关系，我们将较强关系个体组成的群体称为社团 (Community)。

社团内部节点之间的连接相对紧密，而不同社团之间的连接相对稀疏。同一社团的个体，会经常地聚集在某几个地方 (下文称为社团位置)，而有些地方不经常去。一般，个体移动到不同位置的概率和停滞时间不同。个体通常移动到社团位置的概率较大，停滞时间较长；移动到非社团位置的概率相对较小，停滞时间相对较短。

根据社会节点的群居性特征，我们要对某一节点下一时刻的位置进行预测，如果在当前时刻某一

位置汇聚了很多与其社会关系亲密的节点，则该节点下一时刻移动到该位置的概率相对较大。因此，我们可以用社会关系去估计节点未来的位置，用以优化 Markov 模型的预测结果。

利用社会关系对位置信息进行预测，我们首先需要获得节点间的社会关系权值。而接触频率往往被作为衡量节点间社会关系强弱的一个标准^[30]。对于参与式感知系统而言，我们往往认为接触频率较高的节点联系更为密切，协作性更高，因此具有更强的社会关系。同时在社会关系量化过程中，接触频率的计算需要与时间关联在一起，只有在连续较长的时间产生的接触状态对系统来说才是有意义的。

定义1 移动节点A与B在统计时长 T 内的接触记录表示为 $M = \{(t_i^s, t_i^e)\}, i=1, 2, 3, \dots, n$ ，其中 n 是接触次

数, t_i^s 表示第*i*次接触开始的时间, t_i^e 表示第*i*次接触结束的时间。则节点*A*与*B*的接触概率 $W(A, B)$ 如公式 (10) 所示。

$$W(A, B) = \frac{\sum_{i=1}^n (t_i^e - t_i^s)}{T} \quad (10)$$

根据接触概率量化公式 (10), 我们可以获得系统中节点的社会关系矩阵, 根据矩阵对节点关系进行社团划分, 确定关系较为亲密的节点子群。鉴于现有复杂网络聚类算法已经比较成熟, 本文使用 Girvan-Newman 算法聚类算法^[26]作为社团划分的基本算法。

假设我们希望利用社会关系对节点*A*下一时刻到达某一位置 $i (i = 1, 2, \dots, n)$ 的概率进行计算, 已知节点*A*所处于的社团为 C , 将当前时刻该社团中处于位置*i*的节点集合记为 $S = \{S_1, \dots, S_j, \dots, S_n\}$, 其中 $S \subseteq C$, 则根据条件概率有

$$P_i(A|S_j) = \frac{P_i(A, S_j)}{P_i(S_j)}, \quad j = 1, \dots, n \quad (11)$$

其中, $P_i(A|S_j)$ 代表在节点 S_j 处于 i 位置的各种条件下节点*A*到达位置*i*的概率; $P_i(S_j)$ 代表节点 S_j 继续留在 i 地点的概率, 它可以通过Markov模型计算获得; $P_i(A, S_j)$ 代表节点*A*与节点 S_j 在 i 位置相遇的概率, 具体计算公式如式 (12) 所示。

$$P_i(A, S_j) = \frac{f_i(A, S_j)}{\sum_{i=1}^m f_i(A, S_j)} \quad (12)$$

其中, $f_i(A, S_j)$ 代表节点*A*与节点 S_j 在 i 位置的相遇次数。

设节点*A*与节点 S_j 的关系权值 $W(A, S_j) = \rho_j$, 则下一时刻节点*A*到达位置*i*的概率为

$$P_i(A) = \sum_{j=1}^n \lambda_j P_i(A|S_j), \quad \lambda_j = \frac{\rho_j}{\sum_{j=1}^n \rho_j} \quad (13)$$

其中, λ_j 为各条件概率的权值, 它通过节点*A*与节点集*S*中所有节点的关系权值进行归一化处理得到,

$$\sum_{j=1}^n \lambda_j = 1。$$

根据 C 内所有节点的位置分布情况, 我们可以获得节点*A*到达不同位置的概率, 结合之前 Markov 模型的预测结果, 通过加权公式(14)计算获得节点*A*到达所有位置的概率分布, 取访问概率最大的位置作为预测算法的输出。

$$P_i = \alpha P_i^{social} + \beta P_i^{markov} \quad (14)$$

其中, P_i^{markov} 为 Markov 模型针对状态 x_i 的位置预测概率, P_i^{social} 为基于社会关系的预测模型针对位置 i 预测概率。

令权重 $\alpha + \beta = 1$, 式 (14) 可进一步表示为公式 (15)。

$$P_i = \alpha(P_i^{social} - P_i^{markov}) + P_i^{markov} \quad (15)$$

其中, $P_i^{social} - P_i^{markov}$ 为社会关系预测模型对 Markov 模型的修正因子, α 为修正系数。

利用本节提出的基于社会关系的预测优化方法, 我们可以对基于绝对分布的 Markov 模型进行优化, 将其记为 SMLP₁, 同时对加权 Markov 模型进行优化, 将其记为 SMLP_N。对于 Markov 预测模型而言, 随着移动用户的活动, 转移概率矩阵将逐渐变得不再稀疏, 预测的准确率也会随之下降, 利用社会关系对预测结果进行优化, 有助于转移概率矩阵的稀疏化, 对于提高模型预测精度具有很好的效果。

2.4 修正系数参数估计

对于公式 (15) 中提出的修正系数 α , 我们可以通过普通最小二乘法进行参数估计。参数估计的过程为:

从数据集中随机抽出数量为 n 的连续样本集 $Y = \{y_i\}, i = 1, 2, \dots, n$, 其中 $y_i = (y_1^{(i)}, y_2^{(i)}, \dots, y_m^{(i)})$ 表示第 i 个样本 (时刻) 节点的位置分布。 $y_j^{(i)}$ 定义如公式

(16) 所示, 当节点处于 j 位置时 $y_j^{(i)}$ 取值为 1, 否则取值为 0。

$$y_j^{(i)} = \begin{cases} 0, & \text{位置} \neq j \\ 1, & \text{位置} = j \end{cases}, j = 1, 2, \dots, m \quad (16)$$

计算相应时刻的预测结果集合 $X = \{x_i\}$, $i = 1, 2, \dots, n$, 其中 $x_i = (x_1^{(i)}, x_2^{(i)}, \dots, x_m^{(i)})$ 表示第 i 个时刻通过概率判断出的节点位置分布。 $x_j^{(i)}$ 定义如公式

(17) 所示, 其中 $p_j^{(i)}$ 表示 i 个时刻节点处于 j 位置的概率, 计算公式见式 (15)。

$$x_j^{(i)} = \begin{cases} 0, & j \neq \arg \max(p_j^{(i)}) \\ 1, & j = \arg \max(p_j^{(i)}), j = 1, \dots, m \end{cases} \quad (17)$$

利用普通最小二乘法, α 从 0~1 取值, 令 $\sum_{i=1}^h (n - \sum_{j=1}^m (x_j - y_j)^2)$ 方差结果取值最小, 则可以获得最佳的修正因子值。最后的 α 计算公式如公式 (18) 所示。

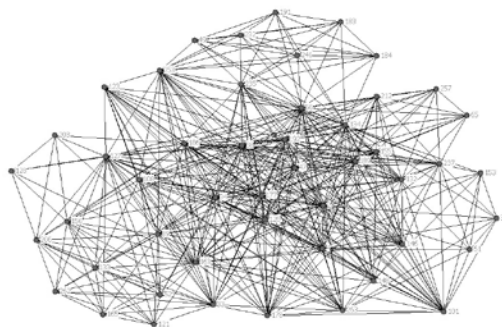
$$\alpha = \arg \min \left(\sum_{i=1}^h \left(n - \sum_{j=1}^m (x_j - y_j)^2 \right) \right) \quad (18)$$

, 其中 h 为节点的个数。

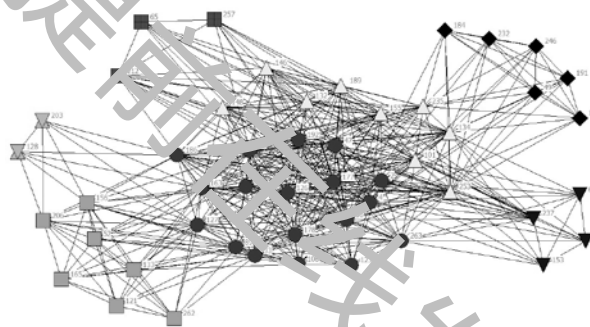
通过参数估计方法, 我们可以利用历史数据来估计修正系数 α , 使算法能够更好地适用于具体场景, 以此来进一步提高算法的准确度。

3 实验分析

3.1 仿真实验配置



(a)



(b)

图 4 WTD 数据集中的节点社会网络结构

3.3 预测精确度分析

为了对预测模型的精度进行评测, 将整理后的用户位置信息分为两部分: 按照时间从原始信息中抽取前 50% 的数据作为训练数据集用于模型的训练, 剩下的 50% 作为预测模型的测试数据集。预测精度 P_{result} 的计算公式如式 (19) 所示:

$$P_{result} = \frac{\sum_{i=1}^n accuracy_i}{n} \quad (19)$$

在进行参与式感知系统机会特征研究时, 节点通信需要通过节点移动带来的相遇机会来完成, 因此节点移动模型研究对机会连接条件下的参与式应用来说尤为重要。本文采用 Wireless Topology Discovery (WTD) 项目提供的数据集对算法进行实验。WTD 数据集记录了 UCSD 校园内 275 个携带 PDA 设备的用户与部署的 AP 间约 11 周的通信数据。由于 WTD 数据集对本文实验来说相当庞大, 因此我们选择了其中一个月 (前 5068105 条) 的数据对算法进行了仿真。在截取的数据集中共有 267 个节点, 447 个 AP, 我们取定 λ 值为 0.2, 根据 AP 位置的邻近性对其进行聚类, 共抽取出 269 个地点位置。

3.2 社会关系矩阵析取

为了利用社会关系对用户的移动轨迹进行预测, 我们首先需要了解系统中节点的社会网络结构。我们使用接触因素来对数据集中的用户关系进行量化, 量化结果如图 3 所示, 其中图 4 (a) 为 WTD 数据集的节点社会网络结构, 图 4 (b) 为使用 Girvan-Newman 算法聚类后的节点关系, 图中相同颜色的节点具有更为密切的社会关系。通过聚类后, 可以将当前时刻该社团中处于具体位置的节点集合 S 析取出来, 用于下一步的分析和计算。

其中, n 表示预测次数, $accuracy_i$ 定义如公式 (20) 所示, 表示对第 i 个预测位置是否预测正确。

$$accuracy_i = \begin{cases} 1, & \text{当预测正确} \\ 0, & \text{当预测错误} \end{cases} \quad (20)$$

首先, 我们使用测试数据集对训练后的标准 Markov 模型 (SMM) 及 SMLP₁ 进行测试。图 5 显示了两种预测模型对各用户位置的预测精度, 图 5 (a) 是编号为 1—89 节点的预测精度, 图 5 (b) 是编号为 89—179 节点的预测精度, 图 5 (c) 是编

号为 180—267 节点的预测精度。从图中可以看出，SMLP₁ 与 SMM 相比，预测精度有一定的提升。

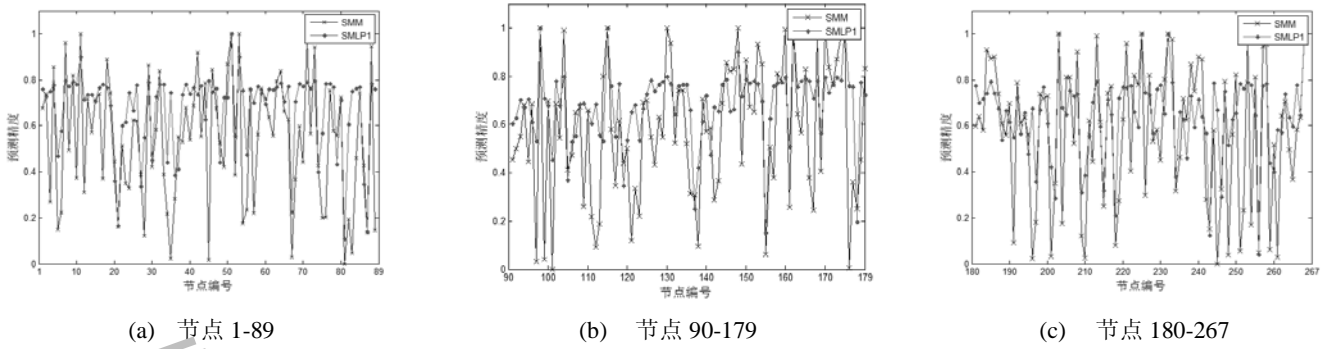


图 5 SMM、SMLP₁ 的预测精度对比

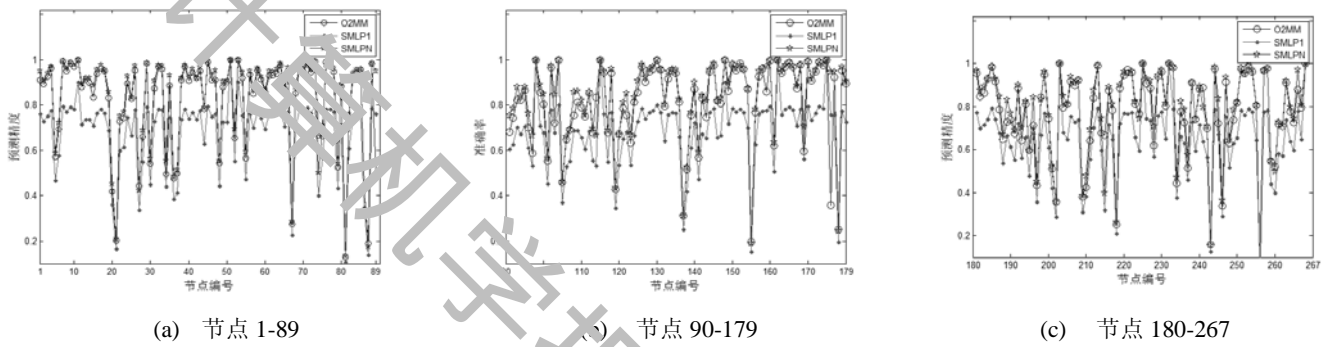


图 6 O2MM、SMLP₁ 和 SMLPN 的预测精度对比

接着，我们将 SMLP_N、SMLP₁ 及 2 阶 Markov 模型 (O2MM) 进行了对比。图 6 为三种模型预测精度的对比图，可以看出，SMLP_N 模型与 SMLP₁ 相比，预测精度有了进一步的提升，同时与公认预测效果较好的 O2MM 相比，具有相当的预测精度。图 7 描述了 SMLP_N、O2MM 及 SMLP₁ 在不同预测精度范围下的节点数量分布情况，从图中我们可以看出 SMLP 在较高的预测精度范围节点分布数量最多，O2MM 稍逊，SMLP₁ 数量最少。例如，SMLP_N 预测精度大于 90% 的节点共有 133 个，O2MM 有 117 个，而 SMLP₁ 只有 34 个。

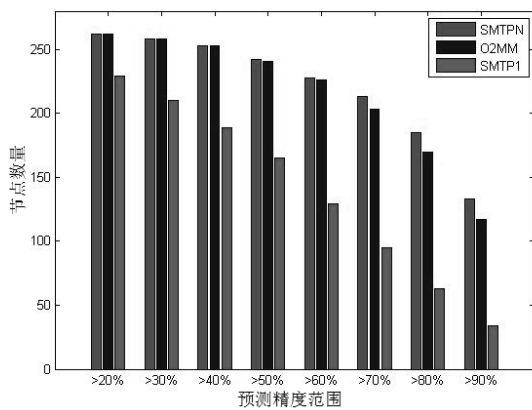


图 7 不同预测精度范围的节点数量情况

表 2 算法性能比较

	SMM	O2MM	SMLP ₁	SMLP _N
预测精度	0.5610	0.8030	0.7126	0.8349
时间复杂度	$O(N)$	$O(N^2)$	$O(N)$	$O(N)$
存储空间	$O(N^2)$	$O(N^3)$	$O(N^2)$	$O(N^2)$

最后，我们对各种模型的算法性能做了比较，比较结果如表 2 所示。从表 2 中可以看出，基于社会关系的预测模型相对于标准的 Markov 模型提高了接近 25% 的预测精度，而相对公认预测效果较好的 2 阶 Markov 模型而言精度提高不大。在空间代价比较上，基于社会关系的预测模型状态空间的复杂度为 $O(N)$ ，而 2 阶 Markov 模型为 $O(N^2)$ ，其中， N 为应用场景中的位置数目；基于社会关系的预测模型的存储空间需求为 $O(N^2)$ ，而 2 阶 Markov 模型为 $O(N^3)$ 。因此，我们可以得出结论，基于社会关系的预测模型能够以比 2 阶 Markov 模型小得多的空间代价获得与 2 阶 Markov 模型相似的预测精度，具有较高的实际应用价值。

3.4 参数估计

在进行预测时，修正系数 α 的取值可能会影响预测结果的精度。在本实验中，我们令 α 从 0.1~0.9 依次取值，观察预测准确率的变化，结果如图 8 所示。在图 8 中，当 $\alpha = 0.65$ 时，预测精确度达到了

最高值。

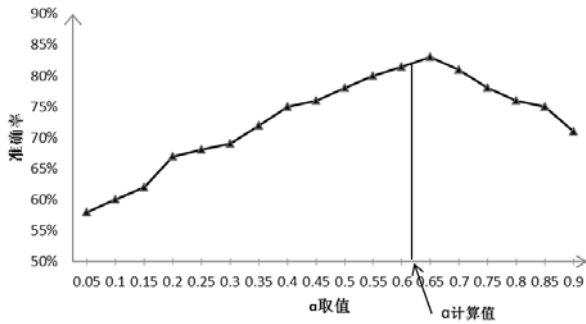


图8 修正系数预测准确率的影响

我们利用参数估计方法(公式(18))对参数进行了估计,获得了 $\alpha = 0.622$ 。这与实验得出的 α 基本一致,因此我们可以认为参数估计法能够获得较为准确的修正系数。

通过实验我们可以看出,参数 α 的取值对最后的预测精度会产生一定的影响:当 α 取值较小时,Markov预测器对最后预测结果影响较大,因此算法对于节点位置改变不频繁的情况预测更加准确;当 α 取值较大时,社会关系对最后预测结果产生较大影响,因此算法对位置经常变更的移动节点预测更加准确。

3.5 位置模型中区域粒度对预测精度的影响

在基于位置的节点移动模型中,位置区域粒度的大小将会对预测算法的准确性产生一定的影响。为了测试不同区域范围下的算法性能,我们通过调整阈值 λ 的大小来获得不同的位置区域粒度,以此来验证不同算法在不同条件下的准确性,具体的实验结果如图9所示。

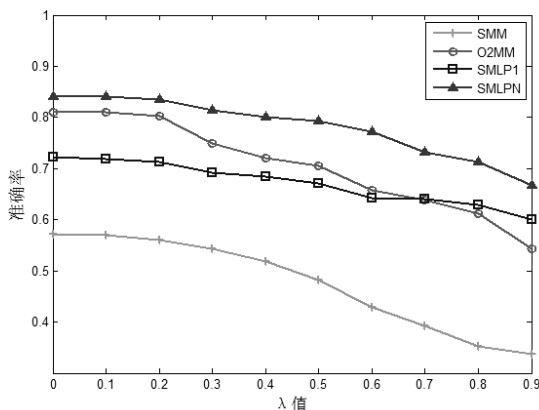


图9 位置粒度对预测准确率的影响

如图9所示,随着 λ 值的变大,模型中位置区域粒度逐渐减小,四种算法的准确性也随之降低。比较四种算法,SMM与O2MM受位置因素影响更

为明显,准确性下降了接近25%,而SMLP受位置粒度因素影响较小,下降趋势较为平缓。通过此实验证明,模型的位置粒度是影响预测精度的重要指标,随着位置覆盖范围变小,节点的移动轨迹位置变换得更加频繁,预测难度也会有所升高,而SMLP算法对于这种位置粒度的变化具有较好的适应性。

4 结论

本文介绍了参与式感知系统的概念及其相关应用,分析了移动节点位置预测对参与式感知系统的重要意义。其次,提出了一种基于社会关系的移动节点位置预测算法(SMLP)。该算法基于位置对应用场景进行建模,以Markov模型为基础,利用其它节点的位置对预测结果进行修正,通过对状态转移矩阵的稀疏化,提高了模型的预测精度。最后,使用UCSD WTD数据集对预测算法进行了实验,验证了算法的有效性。

参考文献

- [1] D. Estrin, et al. 2011 Annual Progress Report, UCLA: Center for Embedded Networked Sensing, technical report. 2011.
- [2] S. Eisenman, E. Miluzzo, N. Lane, et al. BikeNet: A Mobile Sensing System for Cyclist Experience Mapping. ACM Transactions on Sensor Networks (TOSN), 2009, 6(1): 6:1-6:39.
- [3] H. Lu, W. Pan, N. D. Lane, et al. SoundSense: Scalable Sound Sensing for People-Centric Sensing Applications on Mobile Phone. Proceedings of the 7th ACM Conference on Mobile Systems, Applications, and Services. Krakov, Poland, 2009: 165-178.
- [4] Miluzzo, Emiliano, et al. sensing meets mobile social networks: the design, implementation and evaluation of the cenceme application. Proceedings of the 6th ACM conference on Embedded network sensor systems. Raleigh, North Carolina, 2008: 337-350.
- [5] Miluzzo E, Lane N, Eisenman S, et al. Cenceme: Injecting Sensing Presence into Social Networking Applications. Smart Sensing and Context, 2007, 4793: 1-28.
- [6] S. B. Eisenman, N. D. Lane, E. Miluzzo, et al. MetroSense Project: People-Centric Sensing at Scale. Proceedings of the Workshop on World-Sensor-Web, Boulder, USA, 2006: 6-11.
- [7] H. Lu, N. D. Lane, S. B. Eisenman, et al. Bubble-sensing:

- Binding sensing tasks to the physical world. *Pervasive and Mobile Computing*, 2010, 6 (1): 58-71.
- [8] M. H. Krieger, M. Ra, J. Paek, et al. Urban Tomography. *Journal of Urban Technology*, 2010,17(2):21-36.
- [9] M. H. Krieger, R. Govindan, M. Ra, et al. Commentary: Pervasive Urban Media Documentation. *Journal of Planning Education and Research*, 2009, 29 (1): 114-116.
- [10] Deng L, Cox L P. Livecompare: grocery bargain hunting through participatory sensing//Proceedings of the 10th workshop on Mobile Computing Systems and Applications. California, USA, 2009: 4.
- [11] S. Gadepkar, J. Li, R. R. Choudhury, et al. Micro-Blog: sharing and querying content through mobile phones and social participation//Proceedings of the 6th international conference on Mobile systems, applications, and services. Colorado, USA, 2003: 174-186.
- [12] M. Azizyan, R. R. Choudhury. SurroundSense: mobile phone localization using ambient sound and light. *ACM SIGMOBILE Mobile Computing and Communications Review*, 2009, 13 (1): 69-72.
- [13] M. Azizyan, I. Constandache, R. R. Choudhury. SurroundSense: mobile phone localization via ambient fingerprinting//Proceedings of the 15th annual international conference on Mobile computing and networking, Beijing, China, 2009: 261-272.
- [14] B. Hull, V. Bychkovsky, K. Chen, et al. CarTel: A Distributed Mobile Sensor Computing System //Proceedings of the 11th ACM SenSys. Boulder, Colorado, USA, 2006: 125-138
- [15] A. Thiagarajan, L. Ravindranath, K. LaCurts, et al. VTrack: Accurate, Energy-Aware Road Traffic Delay Estimation Using Mobile Phones//Proceedings of the 14th ACM SenSys. Berkeley, CA, 2009:85-98
- [16] Mozer M C. The neural network house: An environment that adapts to its inhabitants//Proceedings of the AAAI Spring Symp. Stanford, CA, 1998: 110-114.
- [17] Karimi H A, Liu X. A predictive location model for location-based services//Proceedings of the 11th ACM international symposium on Advances in geographic information systems. New Orleans, LA, USA, 2003: 126-133.
- [18] Patterson D J, Liao L, Fox D, et al. Inferring high-level behavior from low-level sensors// Proceedings of UbiComp 2003, the 5th Annual Conference on Ubiquitous Computing. Seattle, WA, USA, 2003: 73-89.
- [19] Song L, Kotz D, Jain R, et al. Evaluating location predictors with extensive Wi-Fi mobility data//Proceedings IEEE INFOCOM 2004, the Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies. Hong Kong, China, 2004, 2: 1414-1424.
- [20] Scellato S, Musolesi M, Mascolo C, et al. Nextplace: a spatio-temporal prediction framework for pervasive systems. *Pervasive Computing*. San Francisco, USA : Springer Berlin Heidelberg, 2011: 152-169.
- [21] MA Lin-bing, Zhang Xin-chang, Research on Full-period Query Oriented Moving Objects Spatio-Temporal Data Mode. *Acta Geodaetica et Cartographica Sinica*. 2008, 37(8):32-35.
- [22] Tao Y, Faloutsos C, Papadias D, et al. Prediction and indexing of moving objects with unknown motion patterns//Proceedings of the 2004 ACM SIGMOD international conference on Management of data. Paris, France, 2004: 611-622.
- [23] He Y., Fan S. et.al. whole trajectory modeling of Moving Objects Based on MOST Model[J].*Computer Engineering*, 2008,34(16):41-43.
- [24] Lu S, Liu Y, Liu Y, et al. LOOP: A location based routing scheme for opportunistic networks//Proceedings of the 9th IEEE International Conference on Mobile Ad-hoc and Sensor Systems. Las Vegas, Nevada, USA, 2012: 118-126.
- [25] Mei A S, Andrade-Navarro M A. A novel approach for protein subcellular location prediction using amino acid exposure. *BMC bioinformatics*, 2013, 14(1): 342.
- [26] Girvan M, Newman M E J. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 2002, 99(12): 7821-7826.
- [27] Sengupta D, Maulik U, Banerjeeacharyay S. Weighted markov chain based aggregation of bio molecule orderings. *Computational Biology and Bioinformatics*, 2012, 9(3): 924-933.
- [28] Ville Kotovirta, Timo Toivanen, Renne Tergujeff, Markku Huttunen, Participatory Sensing in Environmental Monitoring-Experiences//Proceedings of the Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing. Palermo, Italy, 2012: 155-162.
- [29] R. Rana, C.T. Chou, S. Kanhere, N. Bulusu and W. Hu, Ear-Phone: An End-to-End Participatory Urban Noise

- Mapping System//Proceedings of ACM/IEEE IPSN, Stockholm, Sweden, April 2010.
- [30] An Jian, Gui Xiaoli, Zhang Wendong, et al. Social Relation Cognitive Model of Mobile Nodes in the Internet of Things. Chinese Journal of Computers, 2012, 35(6): 1164-1174 (in Chinese)
(安健, 桂小林, 张文东, 等. 物联网移动感知中的社会关系认知模型. 计算机学报, 2012, 35(6): 1164-1174)
- [31] Zhou P, Zheng Y, Li M. How long to wait?: predicting bus arrival time with mobile phone based participatory sensing//Proceedings of the 10th international conference on Mobile systems, applications, and services. Long Beach, CA, United Kingdom, 2012: 319-322.
- [32] Lane N D, Eisenman S B, Musolesi M, et al. Urban sensing systems: opportunistic or participatory?//Proceedings of the 9th workshop on Mobile computing systems and applications. Napa, CA, USA, 2008: 11-16.
- [33] Yavaş G, Katsaros D, Ulusoy Ö, et al. A data mining approach for location prediction in mobile environments. Data & Knowledge Engineering, 2005, 54(2): 121-145.
- [34] Eagle N, Pentland A S, Lazer D. Inferring friendship network structure by using mobile phone data. Proceedings of the National Academy of Sciences, 2009, 106(36): 15274-15278.
- [35] Gonzalez M C, Hidalgo C A, Barabasi A L. Understanding individual human mobility patterns. Nature, 2008, 453(7196): 779-782.
- [36] Song C, Qu Z, Blumm N, et al. Limits of predictability in human mobility. Science, 2010, 327(5968): 1018-1021.
- [37] Jiang B, Yin J, Zhao S. Characterizing the human mobility pattern in a large street network. Physical Review E, 2009, 80(2): 021136.
- [38] Qin S M, Verkasalo H, Mohtaschemi M, et al. Patterns, entropy, and predictability of human mobility and life. PloS one, 2012, 7(12): e51353.
- [39] Akoush S, Sameh A. The use of bayesian learning of neural networks for mobile user position prediction//Proceedings of the Seventh International Conference on Intelligent Systems Design and Applications. Rio de Janeiro, Brazil, 2007: 441-446.



Yu Ruiyun, born in 1974, Ph.D., associate professor. His research interests include participatory sensing systems, wireless sensor networks, pervasive and mobile computing, mobile big data analysis, etc.

Xia Xingyou, born in 1988, Ph.D. Candidate. His research interests include participatory sensing systems, mobile computing, etc.

Li Jie, born in 1982, Ph.D., Lecturer. Her research interests include cognitive networks, mobile computing, etc.

Zhou Yan, born in 1988, master. His research interests include participatory sensing systems, mobile computing, etc.

Wang Xingwei, born in 1968, Ph. D., professor, Ph. D. supervisor. His research interests include future Internet, cloud computing, network security and information security.

Background

Participatory sensing is a new sensing technology, which emphasize that people participate in the sensing process. People can sense and record environmental or social information by carrying or wearing sensing equipments or smart phones. In participatory sensing systems, mobile equipments are usually weakly connected. Due to uncertainty of connection, nodes sometimes need encounter opportunities to accomplish data communication and transmission. If the location of mobile nodes can be predicted ahead of several time slots, the service quality and efficiency of the system will be remarkably improved.

In this paper, a location prediction algorithm named the Social-relationship-based Mobile node Location Prediction algorithm (SMLP) is proposed. Social relationship between mobile nodes is exploited to optimize location prediction. It first models the network based on geographic locations, and then use Markov model to characterize evolution of the application. Two algorithm variations, SMLP1 and SMLPN, are proposed based on the Markov model and the weighted Markov model, respectively. Social-relationship between nodes is exploited for optimization and amendment of location prediction. The UCSD WTD data set is used in both the learning phase and the prediction phase. From simulation results, the algorithm performs well on prediction accuracy. A correction coefficient is introduced and makes the algorithm

more scalable.

This work is supported by the National Natural Science Foundation of China under Grant No. 61272529, the National Science Foundation for Distinguished Young Scholars of China under Grant No. 61225012 and No. 71325002; Ministry of Education-China Mobile Research Fund under Grant No. MCM20130391; the Specialized Research Fund of the Doctoral Program of Higher Education for the Priority Development Areas under Grant No. 20120042130003; the Fundamental Research Funds for the Central Universities under Grant No. N12041002, No. N110204003, No. N120104001 and No. N12031003.

One of main tasks of the abovementioned projects is to investigate mobile nodes' behavior in participatory sensing systems. Some published papers were focusing on node mobility model, opportunistic networking and data dissemination in participatory sensing systems and related applications.

This paper highly addresses the mobility model of nodes in participatory sensing systems, and proposes the mobile node location prediction algorithm based on social relationship between mobile nodes. The algorithm is used for opportunistic connection and data forwarding of mobile nodes. It is beneficial to use social relationship for improving location prediction accuracy in participatory sensing systems.