

基于时空感知的用户角色推理

胡亚慧^{①②} 李石君^{*①} 余伟^① 杨莎^① 方其庆^②

^①(武汉大学计算机学院 武汉 430079)

^②(空军预警学院 武汉 430019)

摘要: 随着大数据和信息技术的发展,更好地理解用户的行为轨迹在个性化推荐、广告推荐等方面越来越重要。该文依据大数据环境下的城市计算理论,提出一种基于情境感知的用户角色推理模型。通过用户的行为轨迹分析其行为的时空特性;结合时间、语义分析等构造识别用户角色概率推理模型;通过算法克服识别用户角色的主观性、动态适应性差等问题。实验结果证明了该文所提模型的可行性、精确性和预测准确性。

关键词: 大数据; 城市计算; 情境感知; 用户角色

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2016)03-0517-06

DOI: 10.11999/JEIT150700

Inferring Social Roles with Spatio-temporal Awareness

HU Yahui^{①②} LI Shijun^① YU Wei^① YANG Sha^① FANG Qiqing^②

^①(Computer School, Wuhan University, Wuhan 430079, China)

^②(Air Force Early Warning Academy, Wuhan 430019, China)

Abstract: With the development of big data and information technology, a better understanding of users' trajectories is of great importance for the design of many applications, such as personalized recommendation, behavioral targeting and computational advertising. In this paper, with the theory of urban computing based on big data, a model of recognizing information veracity of users on the social media networks is proposed. The behavior characteristics of users' trajectories based on context awareness are analyzed. The model of recognizing the truth of social roles is formalized and built. The subjectivity of recognizing users' roles is overcome. Furthermore, experiments are conducted with large-scale and real-world datasets. The results show that the proposed model offers a powerful ability for recognition of truth social roles.

Key words: Big data; Urban computing; Context awareness; Social roles

1 引言

随着城市信息化、大数据分析技术和数据科学技术的不断发展,基于位置的服务(Location Based Service, LBS)作为新一代的信息网络服务方式,能够收集到带有明确地理位置坐标的数据,将地图与兴趣点(Point Of Interest, POI)有效地结合,不仅可以识别出用户的兴趣点 POI、常出入的功能区域,

还可以识别出所在城市的交通状况、发展状态等信息^[1-4]。

目前,从用户的行为轨迹数据中发现其社会角色,为不同社会角色的用户提供有针对性的服务是基于位置服务研究的一个热点^[5-8],其主要思想是通过轨迹聚类分析,将用户的角色按照不同的轨迹进行划分。文献[8]使用图挖掘的方法将图中具有相似结构的行为轨迹提取出来,对节点进行无监督聚类,从而实现角色提取。文献[9]从轨迹时空关联语义和时态熵角度发现移动对象社会角色。文献[10]首次提出社交媒体中的用户具有物以类聚人以群分的特性即同质性,包括性别、年龄、宗教信仰、教育程度、职业、社会地位、网络位置、行为等。而造成同质性的原因主要有地理位置、家庭关系、所在机构。可见人的社会角色分析离不开以上这些因素。文献[11]发现具有朋友关系的用户具有相同的属

收稿日期: 2015-06-08; 改回日期: 2015-09-16; 网络出版: 2015-11-19

*通信作者: 李石君 shjli@whu.edu.cn

基金项目: 国家自然科学基金(61272109), 中央高校基本科研业务费专项资金(2042014kf0057), 湖北省自然科学基金(2014CFB289), 空军预警学院青年创新基金(2013ZDJC0101)

Foundation Items: The National Natural Science Foundation of China (61272109), The Fundamental Research Funds for Central Universities (2042014kf0057), The National Natural Science Foundation of Hubei Province of China (2014CFB289), Air Force Early Warning Academy (2013ZDJC0101)

性,例如同一所大学的同学或者同一公司的职员具有相同的职业。但是,在进一步的研究中发现即使同一公司的职员,其分工也不同,承担的角色也就有所不同,例如设计师、工程师、销售员、经理、测试员等等。文献[12]用因子图模型实现了社交媒体的时间感知分析。文献[13]对异构网络建立了一个用户角色推理模型。文献[14]分析了与用户角色和地位相关的社会因素和原理,并进行量化分析。

考虑到用户的角色识别,其行为轨迹离不开所在城市的功能区域划分。对于城市功能区域的识别,文献[15]通过人的行为轨迹和兴趣点 POI 能够识别出城市的功能区域。紧接着,该研究团队改进了方法,利用潜在轨迹发现功能区域^[16]。

因此,通过用户的行为轨迹中识别出用户角色,为不同的社会角色提供有针对性的信息服务,从而提高个性化服务的质量,这也是本文研究的目的。本文的主要研究工作包括:

(1)从时间序列角度分析出用户的行为特征,包括轨迹的规律性和周期性、访问频次、区域跳变性、区域游离性等特征;

(2)考虑到不用角色的用户在某个特定时间出现在某个位置存在着一定的规律性,结合马尔科夫模型对用户的移动轨迹建模,从上下文情境感知角度推理用户角色;

(3)在真实数据上进行实验分析,表明本文算法在准确率上与现有方法相比有较大的提高。

本文内容安排如下:第1节介绍国内外相关研究工作;第2节描述基于时空感知的用户行为特征;第3节构建基于情境感知的用户角色识别模型和算法实现过程;第4节通过实验验证分析了本文方法;第5节总结全文。

2 用户行为特征分析

首先从用户的行为轨迹中发现用户的出行规律,对其行为特征进行分析,从而为判别其角色做基础。如图1所示,从用户的签到数据中,将用户频繁去的位置,即经纬度信息结合地图,将其逐一标注出来,即可得出用户常去的区域有哪些。如图2所示,5个用户(横坐标)在某天的24小时(纵坐标)的频繁功能区域。

定义1 时间序列 用 $T=\{t_1, t_2, \dots, t_i, \dots, t_n\}$ 表示。其中, Δt 表示任何两个相邻时间差,即 $\Delta t=t_i-t_{i+1}$ 。时间序列对应用户的每个签到时间。

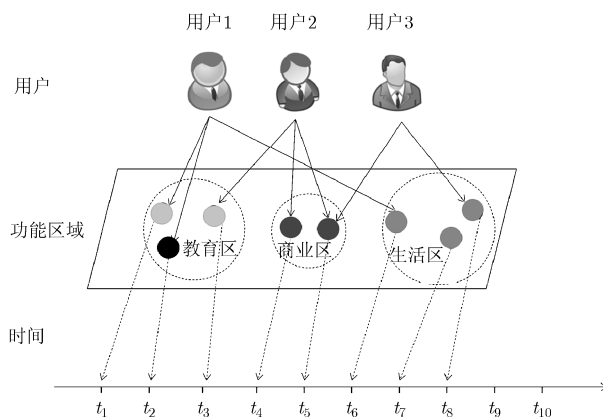


图1 用户行为轨迹示例

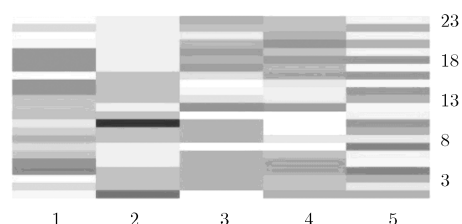


图2 用户一天24小时的轨迹记录示例

定义2 功能区域和区域轨迹序列 将用户的签到数据与百度地图 API^[17]对应获取其地理标签,从地理标签中提取诸如“学校”、“医院”、“饭店”等表示区域类型的关键词。结合文献[9,16]提出的基于语义的功能区域划分方法,功能区域集合 $R=\{r_1, r_2, \dots, r_n, \dots\}=\{\text{学校, 医院, } \dots, \text{生活区, } \dots\}$, 表示用户所在的不同区域。用户 u_i 的区域轨迹序列 $R_i=\{r_i(1), r_i(2), \dots, r_i(l), \dots\}$, 其中 $r_i(l)$ 表示用户 u_i 的第 l 个位置, Δt 表示从位置 $r_i(l-1)$ 到位置 $r_i(l)$ 的时间间隔。随着时间的推移,用户所在的位置是在不断变化的。任意时刻,用户都有一个物理位置与之对应。

定义3 区域访问频次 设用户 u_i 在访问区域所有区域驻留的总时间为 $T_i^{R_i}$ 和所有区域访问的次数为 $F_i^{R_i}$, 如果用户 u_i 在区域 $r_i(j)$ 中驻留的时间和访问该区域的次数分别记为 $T_i^{r_i(j)}$ 和 $F_i^{r_i(j)}$, 用户 u_i 访问区域 $r_i(j)$ 的频次为

$$\text{Tim}(u_i, r_i(j)) = \log \left(\frac{T_i^{r_i(j)}}{T_i^{R_i}} \times \frac{F_i^{r_i(j)}}{F_i^{R_i}} \right) \quad (1)$$

当用户 u_i 访问区域 $r_i(j)$ 的频次 $\text{Tim}(u_i, r_i(j)) < \theta$ 时视为噪声数据, θ 是访问频次阈值。

定义4 区域周期性 特定角色的用户其行为轨迹在时间序列上存在着规律性和周期性。例如公司职员、企业员工等上班族的早出晚归的生活作息,在时间和空间上的概率分布近似为高斯分布^[18]:

$$P(t_i = t | r) = N_r(\mu_t, \sigma_t^2) \quad (2)$$

以功能区域：家(H)和办公场所(W)为例，用户在一天内的空间分布存在两个2维正态分布，且两种分布的峰值正好在家(H)和办公场所(W)^[17]。

$$P(r_i = r | c_u(t_i)) = \begin{cases} \sim N_H(\mu H, \sum H), c_u(t_i)=H \\ \sim N_W(\mu W, \sum W), c_u(t_i)=W \end{cases} \quad (3)$$

其中， $c_u(t_i)$ 表示时刻 t_i 用户所处的状态属于 H 或 W 分布。

定义5 区域游离度 某种角色的用户常常频繁往返于一个或多个区域。如学生常常往返于教室、宿舍和食堂，上班族在工作日常常往返于单位和家庭住所，即从式(3)中找出 $P > \gamma$ ，即用户频繁游离的区域。其中， γ 是概率阈值。

定义6 区域跳变性 在某一时间段内，用户平均访问区域的个数。例如出租车司机的区域跳变性很高，而医生或教师每天的区域跳变性就很低，即

$$B = \frac{\sum_{t=1}^n \text{Tim}(u_i, r_i(j))}{\sum_{t=1}^n R_i} \quad (4)$$

当 $B > \xi$ 时视为噪声数据， ξ 是区域跳变阈值。

定义7 用户角色 用户的角色有多种，如医生，学生，教师，职员，销售员等。刻画一个用户的角色，可以从区域访问频次，区域亲密度，区域跳变性等方面进行逻辑推理，即 $A = \{a_1, a_2, a_3, \dots\} = \{\text{教师, 医生, 职工, } \dots\}$ 。

3 基于情境感知的用户角色识别

3.1 模型建立

为了有效地识别用户角色，对于每一个用户的行为轨迹分别从区域，时间和用户3维向量空间 \mathbf{S} 进行分析，如图3所示，即 $\mathbf{S} = (\mathbf{R}, \mathbf{T}, \mathbf{U})$ ，其中 \mathbf{R} 表示区域集合， \mathbf{T} 表示时间集合， \mathbf{U} 表示用户集合。图中，黑色方形表示用户在某个时刻出现在某功能区域，白色方形表示用户在相应时刻没有出现在某功能区域。

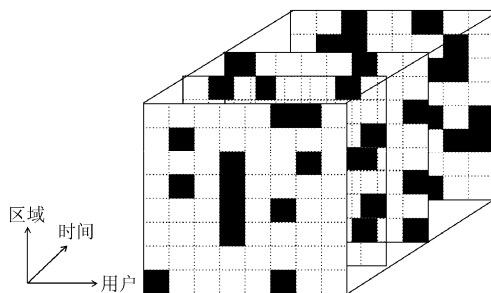


图3 基于情境感知的用户角色识别模型

假定总共有 l 个用户， m 个区域，用户 u_i 在3维空间向量 \mathbf{S} 中的记录所占比例为

$$P(u_i | \mathbf{U}) = \frac{\sum_{j=1}^m \alpha(u_i, r_i(j))}{\sum_{i=1}^l \left(\sum_{j=1}^m \alpha(u_i, r_i(j)) \right)} \quad (5)$$

其中，函数 $\alpha(u_i, r_i(j)) = \begin{cases} 1, \text{用户 } u_i \text{ 访问了区域 } r_i(j) \\ 0, \text{用户 } u_i \text{ 没有访问区域 } r_i(j) \end{cases}$ ；

分母表示所有用户访问所有区域的总次数。

在 t 时间内，用户 u_i 访问区域 $r_i(j)$ 的概率为

$$P(u_i, r_i(j)) = \frac{\sum_{t=1}^n \alpha(u_i, r_i(j))}{P(u_i | \mathbf{U})} \quad (6)$$

分子表示用户 u_i 访问区域 $r_i(j)$ 的总数。

结合式(2)和式(3)，用户 u_i 访问区域 $r_i(j)$ 的时间概率分布为

$$P(t | r_i(j)) = N_{r_i(j)}(\mu_t, \sigma_t^2) \quad (7)$$

假设用户 u_i 的角色是教师 a_1 ，其访问学校区域的频率应该比其他功能区域高，且教师属于上班族，有早出晚归的工作习惯，在时间上满足式(2)和式(3)的概率分布。

如果用户 u_i 的角色是出租车司机，其访问各种功能区域的随机性很大，且在每个位置停留的时刻很短，其轨迹序列很难找出规律，需要通过区域访问频次、区域游离度和区域跳变性设定的阈值来衡量。但是如果用户 u_k 的角色是公交车司机，不仅有规律的作息时间，其轨迹也具有规律性和周期性。因此很容易区别出租车司机和公交车司机。

根据 N 阶Markov空间，用户 u_i 从当前区域 $r_i(j)$ 转到下一个区域 $r_i(j+1)$ 的概率为

$$P(r_i(j+1) | r_i(j)) = \frac{P(u_i, r_i(j)) \times P(u_i, r_i(j+1))}{\sum_{j=2}^m P(u_i, r_i(j-1)) \times P(u_i, r_i(j))} \quad (8)$$

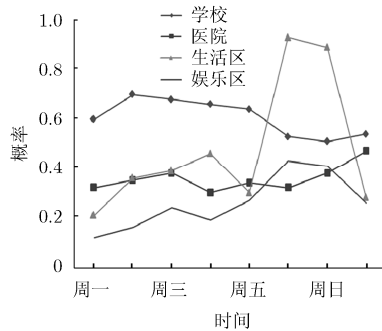
其中，分子表示用户 u_i 在 t 时刻位于区域 $r_i(j)$ ，而在 $t+1$ 时刻位于区域 $r_i(j+1)$ 总数。

推理一个用户是某种角色，需要从时间概率分布上和空间区域分布上来衡量，即

$$P(a_k, u_i) = \left[\sum_{u_i \in \mathbf{U}, r_i(j) \in \mathbf{R}, a_k \in \mathbf{A}} (\varphi P(t | r_i(j)) + \tau P(u_i, r_i(j)) + (1 - \varphi - \tau) P(r_i(j+1) | r_i(j))) \right] / P(u_i | \mathbf{U}) \quad (9)$$

$\varphi, \tau \in (0, 1)$ 是权重。

如图4所示，用户 u_i 一周内在学校、医院、生活区、商业区4个区域内访问的概率。

图4 用户 u_i 一周内访问不同区域的概率

3.2 算法

(1)数据预处理阶段: 为了提高数据质量需要去掉数据中的噪声数据, 在实验前先将数据集进行预处理, 处理规则如下:

根据式(1), 将用户 u_i 访问区域 $r_i(j)$ 的频次 $\text{Tim}(u_i, r_i(j)) < \theta$ 的数据删除。根据文献[9,15,16,19]以及实验发现, 用户访问某区域频率低于5次, 并且在该区域停留时间低于30 min的数据会严重降低实验的准确率, 即频次阈值 $\theta = 0.1$ 。例如, 一个销售员轨迹频次太低时, 他的行为轨迹全部集中在某个区域(如学校), 推断模型会将该用户误判为教育人员(教师和学生), 从而会训练出一个错误的推断模型。根据式(3), 概率至少要大于20%, 即概率阈值 $\gamma = 0.2$; 根据式(4), 跳变阈值 $\xi = 5$ 。

(2)算法实现过程: 对任意用户进行角色识别算法如表1所示。

表1 用户角色识别算法

算法1 用户角色识别算法

输入: 用户 check in 数据, 城市功能区域集合

输出: 用户角色集合

① Map 阶段

输入: $\langle \text{key}=u_i; \text{value}=\langle \text{region1}, t1 \rangle, \langle \text{region2}, t2 \rangle, \dots \rangle$

处理: 以实体的 id 为键, 将 value 与 ts 进行映射;

② Reduce 阶段

输入: $\text{key}=\langle u_i \rangle; \text{value}=\langle r_i, ts_i \rangle$

处理:

for each item of D do

 Compute $P(u_i | U)$ according to 式(5);

 Compute $P(u_i, r_i(j))$ according to 式(6);

 Compute $P(t | r_i(j))$ according to 式(7);

 Compute $P(r_i(j+1) | r_i(j))$ according to 式(8);

 Compute $P(a_k, u_i)$ according to 式(9);

end for

4 实验结果与分析

本实验利用虚拟化技术, 通过开源的虚拟化软件 Oracle Virtualbox 将高性能 NF8560M2 服务器虚拟出30个主机节点, 并以此为底层的分布式硬件环境。每个节点虚拟出一颗 XENO E7-4807 的 CPU 和 4 G 的内存, 主机采用的是 Windows Server 2008 R2 的操作系统, 节点采用的是 Ubuntu 12 操作系统的 Hadoop 0.2 平台。

4.1 实验数据集

为了更充分地体现算法的效果, 实验采用真实的数据集, 数据集信息如表2所示。在 Gowalla 提供的数据集中抽取了 4206641 个 Gowalla 上的签到记录, 这些签到记录由 25653 个用户在 2009 年 1 月到 2010 年 11 月期间生成。每个签到记录包含用户 ID, POI ID, POI 地理位置坐标, 签到时间。Foursquare 数据集包括 17, 908 个用户在 2010 年 1 月到 2011 年 1 月期间生成的 2, 127, 401 个签到记录。将实验数据集按照 2: 8 的比例分为训练集和测试集, 以验证实验的准确性和合理性。

表2 数据集信息表

	GoWalla	Foursquare
Duration	01/2009-11/2010	01/2010-01/2011
Users	25653	17908
Check-ins	4206641	2127401
Unique locations	694617	43509
Links	203502	258131
Test check-ins	3017602	1834735

4.2 评价指标

为了判断本实验的效果, 从算法的准确性方面进行检验。

本文采用平均准确率(Precision, P)、平均召回率(Recall, R)和平均综合评价指标(F -measure)衡量实验的准确性^[6,7]。设候选的数据集空间 D , 其中实际不一致的数据集为 IND, 用本文算法自动发现的不一数据集为 TIND, 则整个算法的平均准确率:

$$P = \frac{|\text{IND} \cap \text{TIND}|}{|\text{TIND}|} \quad (9)$$

整个算法的平均召回率:

$$R = \frac{|\text{IND} \cap \text{TIND}|}{|\text{IND}|} \quad (10)$$

整个算法的平均综合评价指标:

$$F_1 = \frac{2 \times P \times R}{P + R} \quad (11)$$

4.3 实验结果与分析

(1)参数设置 本实验采用 2 种人工标注的方式进行训练。一种是抽验选择数据集，将数据进行人工提取，构造精确的数据集进行初始训练；另一种是根据初始训练参数进行识别，对结果进行抽验和标注，根据标注结果进行参数调整，求解满足最佳效果的参数值。训练得到的各阈值参数如表 3 所示。

(2)准确率比较 根据轨迹语义识别用户角色特征的典型算法 MSTP^[20]与本文算法进行比较(见图 5)。说明本文算法准确率明显优于 MSTP。

(3)运行时间比较 在时间效率上，使用相同规模的数据，调整分析次数，两种方法的表现结果如图 6 所示。在固定的数据规模情况下，两种方法的运行时间呈线性增长。当数据规模增大时，MSTP 的时间复杂度为多项式，而本文算法仍为线性时间复杂度。

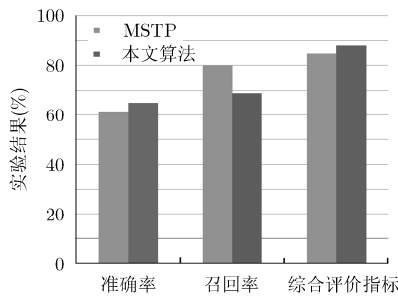


图 5 实验结果对比

5 结束语

本文将社交媒体中用户的行为轨迹和城市功能区域相结合，从而识别出用户的角色信息。依据大数据环境下的城市计算理论，提出一种基于情境感知的用户角色识别模型：(1)通过用户的行为轨迹分析其行为的时空特性；(2)综合考虑影响用户角色的多种要素。结合时间、语义分析等构造识别用户角色模型；(3)通过算法克服识别用户角色的主观性、动态适应性差等问题。实验结果表明，该模型具有更好的动态适应性和预测准确性。

表 3 本文所采用的阈值参数

参数名	γ	ξ	θ	φ	τ
参数值	0.2	5	0.1	0.3	0.35

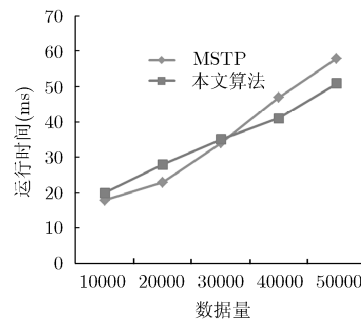


图 6 性能分析曲线

参考文献

[1] 王静远, 李超, 熊璋, 等. 以数据为中心的智慧城市研究综述[J]. 计算机研究与发展, 2014, 51(2): 239-259.
WANG Jingyuan, LI Chao, XIONG Zhang, et al. Survey of data-centric smart city[J]. *Journal of Computer Research and Development*, 2014, 51(2): 239-259.

[2] 李鹏, 于晓洋, 孙渤禹, 等. 基于用户群组行为分析的视频推荐方法研究[J]. 电子与信息学报, 2014, 36(6): 1485-1491. doi: 10.3724/SP.J.1146.2013.01225.
LI Peng, YU Xiaoyang, SUN Boyu, et al. Video recommendation method based on group user behavior analysis[J]. *Journal of Electronics & Information Technology*, 2014, 36(6): 1485-1491. doi: 10.3724/SP.J.1146.2013.01225.

[3] 徐风苓, 孟祥武, 王立才. 基于移动用户上下文相似度的协同过滤推荐算法[J]. 电子与信息学报, 2011, 33(11): 2785-2789. doi: 10.3724/SP.J.1146.2011.00384.
XU Fengling, MENG Xiangwu, and WANG Licai. A collaborative filtering recommendation algorithm based on context similarity for mobile users[J]. *Journal of Electronics & Information Technology*, 2011, 33(11): 2785-2789. doi: 10.3724/SP.J.1146.2011.00384.

[4] PAN Gang, QI Guande, ZHANG Wangsheng, et al. Trace analysis and mining for smart cities: Issues, methods, and applications[J]. *IEEE Communications Magazine*, 2013, 51(6): 120-126.

[5] 王玉斌, 孟祥武, 胡勋. 一种基于信息老化的协同过滤推荐算法[J]. 电子与信息学报, 2013, 35(10): 2391-2396. doi: 10.3724/SP.J.1146.2012.01473.
WANG Yubin, MENG Xiangwu, and HU Xun. Information aging-based collaborative filtering recommendation algorithm[J]. *Journal of Electronics & Information Technology*, 2013, 35(10): 2391-2396. doi: 10.3724/SP.J.1146.2012.01473.

[6] YAN Zhixian, CHAKRABOTRY D, PARENT C, et al. Semantic trajectories: Mobility data computation and annotation[J]. *ACM Transactions on Intelligent Systems and Technology*, 2012, 9(4): 39: 1-39: 34.

[7] YING J J C, LEE W C, and TSENG V S. Mining

- geographic-temporal-semantic patterns in trajectories for location prediction[J]. *ACM Transactions on Intelligent Systems and Technology*, 2013, 5(1): 2:1–2:34.
- [8] HENDERSON K, GALLAGHER B, ELIASSI-RAD T, *et al.* RolX: Structural role extraction & mining in large graphs[C]. Proceeding of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Beijing, 2012: 1231–1239.
- [9] 马宇驰, 杨宁, 谢琳, 等. 基于轨迹时空关联语义和时态熵的移动对象社会角色发现[J]. *计算机研究与发展*, 2012, 49(10): 2153–2160.
- MA Yuchi, YANG Ning, XIE Lin, *et al.* Social roles discovery of moving objects based on spatial temporal associated semantics and temporal entropy of trajectories[J]. *Journal of Computer Research and Development*, 2012, 49(10): 2153–2160.
- [10] MCPHERSON M, SMITH-LOVIN L, and COOK J M. Birds of a feather: homophily in social networks [J]. *Annual Review of Sociology*, 2001(27): 415–444.
- [11] MISLOVE A, VISWANATH B, KRISHNA P G, *et al.* You are who you know: Inferring user profiles in online social networks[C]. Proceedings of the 3rd International Conference on Web Search and Web Data Mining, New York, 2010: 251–260.
- [12] TAN Chenhao, LEE Lililan, TANG Jie, *et al.* User-level sentiment analysis incorporating social networks[C]. Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, 2011: 1397–1405.
- [13] TANG J, LOU T, and KLEINBERG J. Inferring social ties across heterogeneous networks[C]. Proceedings of the 5th International Conference on Web Search and Web Data Mining, Seattle, 2012: 743–752.
- [14] ZHAO Yuchen, WANG Guan, YU Philip S, *et al.* Inferring social roles and statuses in social networks[C]. Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Chicago, 2013: 695–703.
- [15] YUAN Jing, ZHENG Yu, and XIE Xing. Discovering regions of different functions in a city using human mobility and pois[C]. Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Beijing, 2012: 186–194.
- [16] YUAN Jing, ZHENG Yu, XIE Xing, *et al.* Discovering urban functional zones using latent activity trajectories[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2015, 27(3): 712–725.
- [17] BAIDU. Map[OL]. <http://map.baidu.com/>, 2012.
- [18] CHO E, MYERS S A, and 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, 2011: 1082–1090.
- [19] 马武彬, 刘明星, 黄宏斌, 等. 依据历史轨迹构建城市出租车移动概率模型[J]. *国防科技大学学报*, 2014, 36(6): 129–134.
- MA Wubin, LIU Mingxing, HUANG Hongbin, *et al.* The moving probability model of urban cabs based on history trajectory[J]. *Journal of National University of Defense Technology*, 2014, 36(6): 129–134.
- [20] YING Jiaching, LU Chan, LEE Wangchien, *et al.* Mining user similarity from semantic trajectories[C]. Proceedings of the 2010 International Workshop on Location Based Social Networks, San Jose, 2010: 19–26
- 胡亚慧: 女, 1980年生, 博士生, 研究方向为数据挖掘、城市计算、大数据.
- 李石君: 男, 1964年生, 教授, 研究方向为数据挖掘、数据库、大数据.
- 余伟: 男, 1987年生, 博士, 研究方向为数据挖掘、城市计算、大数据.
- 杨莎: 女, 1980年生, 博士生, 研究方向为数据挖掘、电子商务、大数据.
- 方其庆: 男, 1979年生, 博士, 研究方向为数据挖掘、大数据.