

利用位置的虚拟社交网络地理骨干网提取

常晓猛^{1,2} 乐 阳² 李清泉^{2,1} 陈碧宇¹ 萧世伦^{3,1} 涂 伟²

1 武汉大学测绘遥感信息工程国家重点实验室,湖北 武汉,430079

2 深圳大学空间信息智能感知与服务深圳市重点实验室,广东 深圳,518060

3 美国田纳西大学地理系,田纳西州诺克斯维尔,37996-0925

摘要:基于海量的中国大型社交网络新浪微博人际社交关系数据,利用用户的地理位置信息对人-人(人际)社交关系进行虚拟网络空间到物理空间的映射,形成地-地关系。以城市为尺度,对个体间的社交关系链进行聚合,构建城市-城市(城际)社交关系网。根据网络的全局异构和局部异构等拓扑特征以及城市空间的相互作用,提出了一种融合重力模型和信息熵技术的地理骨干网提取方法。该研究有助于揭示虚拟网络环境下的城市体系结构、城市辐射力、城市吸引力和开放程度等问题。

关键词:社交网络;空间结构;地理骨干网;香农熵;重力模型

中图法分类号:P208

文献标志码:A

城市体系的空间联系研究是城市地理学的传统课题。长期以来,主要运用人流、物流、交通流、技术流、金融流等进行数据的收集、分析及挖掘^[1-3]。近年来,以 Facebook、Twitter、LinkedIn、微博、微信等为代表的社交网络及社会媒体迅速兴起,社交网络正通过其巨大的影响力对人类社会的各个方面产生革命性的影响。人类活动从现实物理空间延伸到虚拟网络空间^[4],为地理学者研究虚拟社交网络中城市间信息流的空间结构、城市空间联系体系及空间相互作用机制、城市时空行为等研究提供新的方向,对于重构和丰富城市学科理论具有重要意义^[5-6],也有助于社会学家揭示虚拟网络空间与线下现实空间的映射关系、人类社会新的组织结构和活动模式^[7]。

然而,社交网络中节点和社交关系数目、网络行为及活动数据常常是结构复杂、数量惊人的,是典型的大数据和复杂网络;另外,85%的数据具有非结构化的特点,使得人们从直观上把握该网络的整体结构和性质变得不现实。这对利用此类数据研究城市空间关系和体系结构是一个巨大的挑战。通过技术手段对社交网络进行科学抽样、抽稀,抽取地理骨干网,从而降低信息理解难度,有助于揭示网络中的“流”的走向,把握城市体系的

整体空间联系特征,挖掘复杂现象背后规律的能力。

传统的对非均匀复杂网络骨干的提取方法主要包括基于无权网络及基于带权网络骨干的提取^[8-13]。前者主要考虑网络节点的拓扑性质如节点度、中介中心度等筛选重要的节点,进而剔除次要节点之间的边;后者主要从边入手,考虑复杂网络全局及局部异构特征保留重要的链接,进而剔除没有链接的节点。然而,这些方法共同的缺陷都是针对非空间网络,没有考虑地理空间网络的空间分布特征及空间相互作用力。本文充分考虑城际社交网络的空间特征和拓扑异构特征,提出了一种融合重力模型和信息熵理论的地理骨干网提取方法。

1 城际空间社交关系网络

1.1 城际空间社交关系网络构建

基于图论,复杂社交网络中的个体及个体间的社会关系通常可以被抽象为图中的节点和边,用 $G_p = (V_p, E_p)$ 表示,其中点集 V_p 和边集 E_p 分别代表网络节点集合以及节点间的社交关系集,节点数记为 $N_p = |V_p|$,边数记为 $M_p = |E_p|$ 。

收稿日期:2014-02-15

项目来源:国家自然科学基金资助项目(41371377);深圳市科技研发资金资助项目(ZDSY20121019111146499);深圳市战略性新兴产业发展专项资金资助项目(JCYJ20121019111128765);CCF-腾讯犀牛鸟科研基金资助项目(CCF-TencentARG20130115)。

第一作者:常晓猛,博士生。主要从事地理社交网络、时空可视分析、ITS等方面的研究。E-mail: changxiaomeng@gmail.com

通讯作者:李清泉,博士,教授。E-mail: liqq@szu.edu.cn

由于典型的在线社交网络服务中(如微博等)是基于关注和被关注的单向关系,因此, G_p 是一个有向无权网络。

根据社交网络用户的地理位置信息,将虚拟空间的人-人社交关系映射到地理空间,形成地-地的关系,进一步将 $G_p = (V_p, E_p)$ 拓展为地-地的空间网络,用 $G_g = (V_g, E_g)$ 表示,其中点集 V_g 和边集 E_g 分别代表地理节点集以及地理节点间的社会链接集。对 G_g 网络中起讫节点分别属于相同地理区域的地-地社会关系以城市为尺度进行空间聚合,构建了城际地理社交关系网 G_c , 城市节点数记为 N_c , 城际社交连接边数记为 M_c 。

1.2 城际空间社交关系网络的交通量定义

$\forall e_{ij} \in E_c$, 赋予边 e_{ij} 权重 σ_{ij} , 在本研究中,将 σ_{ij} 定义为城市 i 到城市 j 的出向社交通量;同理, σ_{ji} 定义为城市 i 从城市 j 获得的入向社交通量;而 $\sigma(i, j)$ 或 $\sigma(j, i)$ 被定义为城际总社交通量,为城市 i 和城市 j 的双向社交通量之和,它在数值上直观地反映两个城市之间的联系强度,即

$$\sigma(i, j) = \sigma(j, i) = \sigma_{ij} + \sigma_{ji} \quad (1)$$

为了刻画在城际社交网络中各城市节点的地位,本研究定义了城市总入向社交通量、总出向社交通量及总社交通量等参数。

城市 i 总出向社交通量定义为城市 i 到网络中其他所有城市的城际出向社交通量之和,用 S_i^{out} 表示,即

$$S_i^{\text{out}} = \sum_{j=1, j \neq i}^{N_g} \sigma_{ij} \quad (2)$$

城市 i 总入向社交通量定义为网络中其他所有城市到城市 i 的城际入向社交通量之和,用 S_i^{in} 表示,即

$$S_i^{\text{in}} = \sum_{j=1, j \neq i}^{N_g} \sigma_{ji} \quad (3)$$

城市 i 总社交通量定义为城市 i 与网络中其他所有城市的城际总社交通量之和,用 S_i^{total} 表示,即

$$S_i^{\text{total}} = \sum_{j=1, j \neq i}^{N_g} \sigma(i, j) \quad (4)$$

2 城际空间社交关系地理骨干网提取

2.1 方法概述

大型社交网络的社交关系都是人际间的拓扑社交联系。为了获取虚拟社交网络地理骨干网结构,根据在线社交网络中的用户节点、用户社交关

系以及用户的地理位置信息,对该数据进行三步操作:① 空间映射。根据社交网络用户的地理位置信息,对人-人社交关系进行一次虚拟网络空间与物理空间的映射,形成地-地关系。② 空间聚合。以城市为尺度,对个体间的社交关系链进行聚合,构建城市-城市(城际)社交关系网。③ 结构简化。研究城际社交关系网的网络性质,根据网络的全局异构和局部异构特征,提出了一种基于信息熵技术的滤波方法,抽取了城市社交关系网的地理骨架。

中国城际社交网络在统计学上被证明是一个严重不均衡的网络,即网络中少数的城市节点(Hub)集中了网络中大部分的社交关注与被关注度,少数的 Hub 城际链接集中了网络中大部分信息流的交换^[6]。这一现象显示中国城际社交网络具有非常鲜明的全局和局部异构性以及层次结构。在提取城际社交骨干网的过程中,首先必须考虑该非均匀网络的异构和层次的拓扑特征,即通过选择有效的参数和滤波手段,既要考虑骨干网络城市节点的完备性,保留全局网络中权重较大的城市链接;又要充分考虑单个城市节点的局部异构性和对外联系的多样性,保留影响局部城市重要的城市链接。此外,传统的复杂网络骨架提取方法往往仅考虑网络的拓扑性质,而忽视网络的地理空间特征和空间相互作用。本研究将从网络的空间特征和拓扑特征入手,综合考虑这两个因素,设计一体化的骨干网提取滤波函数。

2.2 空间紧密性

网络紧密度是反映不同区域间联系强弱的指标,最早被用在交通运输领域反映区域间旅客的交流强弱^[14]。本研究基于城际间的社交通量与城市的总社交通量两个参数,引用了网络连接紧密度的定义作为城市间空间联系强弱的指标,即

$$W_{ij} = \frac{\sigma_{ij}}{S_i^{\text{total}}} + \frac{\sigma_{ji}}{S_i^{\text{total}}} + \frac{\sigma_{ij}}{S_j^{\text{total}}} + \frac{\sigma_{ji}}{S_j^{\text{total}}} \quad (5)$$

在空间网络中,具有同样权重的连接由于“空间阻尼”的不同,其存在的价值也就不同。大量研究显示,随着空间距离的增加,地理区域间的联系强度逐步衰减,计量地理学常常用重力模型来模拟距离衰减效应和空间相互作用,即

$$T_{ij} = K \cdot \frac{P_i \cdot P_j}{f(d_{ij})} \quad (6)$$

其中, P_i 和 P_j 分别为 i 和 j 的相关总量指标; $f(d_{ij})$ 为空间距离阻尼因子; K 为模型比例常数。通常情况下, $f(d_{ij})$ 被定义为关于空间距离 d_{ij} 的指数模型,即 $f(d_{ij}) = d_{ij}^\gamma$, 其中 γ 为空间距离阻

尼常数。这一理论在研究中国城市体系的空间网络结构、交通网络、社交网络等领域广泛应用^[1,15-16],在移动通信网络,空间距离阻尼常数被海量数据证实约等于0.5^[17]。本研究中,具有相同网络连接紧密度的两组城际联系,其中城市间的距离越远的那组城市联系在骨干网提取过程中将会得到优先保留。为了削弱城际联系空间阻尼的影响,在网络连接紧密度指数的基础上定义了网络连接空间紧密度指数,即

$$W'_{ij} = W_{ij} \cdot d_{ij}^{-1/2} \quad (7)$$

2.3 拓扑异质性

在信息论中,熵被用来衡量一个随机变量出现的期望值。熵越高,则能传输越多的信息;熵越低,则意味着传输的信息越少。香农多样性指数(Shannon's diversity index, SHDI)是一种基于信息理论的测量指数,即

$$H_i = - \sum_{k=1}^{m_i} p_{ik} \ln p_{ik} \quad (8)$$

其中, p_{ik} 为种间个体的分担率。式(8)包含两个成份:①种数 m_i ;②各种间个体分配的均匀性,该指标在生态学中的应用很广泛,能反映景观异质性与生物多样性,特别对景观中各拼块类型以及生物种类非均衡分布的状况较为敏感,即强调稀有类型对信息的贡献^[17]。本文将该指数引用到城际社交网络,选择SHDI值作为衡量网络城市节点对外联系异质性和多样性的指标,其中

$$p_{ik} = W'_{ik} / \sum_{k=1}^{m_i} W'_{ik} \quad (9)$$

如果在城际社交网络中,城市节点与网络中其他城市的联系越丰富、异质,破碎化程度越高,则该城市对外联系不确定性的信息含量也越大,则SHDI值也就越高。

2.4 骨干网提取算法

在综合考虑城际社交网络空间相互作用和拓扑特征的基础上,本文提出了基于重力模型(gravity model, GM)和香农多样性指数地理骨干网提取算法(SHDI-GM),制定了骨干网滤波函数,即

$$f(i, \text{top}-k) = - \sum_{k=1}^{\text{top}-k} p_{ik} \ln p_{ik} \geq H^* \quad (10)$$

其中, H^* 为全局社交多样性指数阈值。该滤波函数从单个城市节点 i 入手,对该城市的所有城际链接按空间紧密度进行排序,通过迭代增加 k ,筛选出最少数量的 $\text{top}-k$ 条与当前城市具有较高空间紧密度的城市链接,计算当前条件下描述该城市社交关系的香农多样性指数,即 $f(i, \text{top}-k)$ 值,直至 $k=m_i$ 或者 $f(i, \text{top}-k) \geq H^*$ ($k \leq m_i - 1$),并保留该 $\text{top}-k$ 条边,剔除剩余的城市链接。对于具体城市来说,如果在以该城市为中心的局部空间社交网络中,对外联系越平均,则该城市的社交异质性越低,SHDI值越小,因而在骨干网提取过程中需要保留更多的边。通过动态控制 H^* ,从而实现在总体上控制城际社交骨干网的尺度和规模。具体算法流程如下:

算法:SHDI-GM
 输入:(V, E):原始城市网络; $|\sigma|$:城际社交流量矩阵; $|d|$:城际空间距离矩阵; H^* :全局社交多样性指数阈值
 输出:骨干网络(V', E')
 BEGIN
 Step 1: 计算城际空间网络紧密度矩阵 $|W'|$
 Step 2: for each 城市 C_i in V , 对 $|W'_{ij}|$ 中的元素进行排序,调用 $\text{filter}(|W'_{ij}|, H^*)$
 END
 Procedure $\text{filter}(C_i, |W'_{ij}|, H^*)$
 While ($f(i, \text{top}-k) < H^*$ and $k \leq m_i$) {
 $k++$;
 选择 $\text{top}-k$ 条较大空间网络紧密度的链接 $E_i^{\text{top}-k}$
 }
 Output $E_i^{\text{top}-(k-1)}$
 }

算法:SHDI-GM

输入:(V, E):原始城市网络; $|\sigma|$:城际社交流量矩阵; $|d|$:城际空间距离矩阵; H^* :全局社交多样性指数阈值

输出:骨干网络(V', E')

BEGIN

Step 1: 计算城际空间网络紧密度矩阵 $|W'|$

Step 2: for each 城市 C_i in V , 对 $|W'_{ij}|$ 中的元素进行排序,调用 $\text{filter}(|W'_{ij}|, H^*)$

END

Procedure $\text{filter}(C_i, |W'_{ij}|, H^*)$

While ($f(i, \text{top}-k) < H^*$ and $k \leq m_i$) {
 $k++$;

选择 $\text{top}-k$ 条较大空间网络紧密度的链接 $E_i^{\text{top}-k}$

}

Output $E_i^{\text{top}-(k-1)}$

}

3 实验与分析

3.1 实验数据

本文通过新浪微博开放平台随机选择了网络中的1000个种子用户,基于Snow-ball的数据抓取方法,共采集了7286310个用户的基本信息,包括用户地理信息和用户之间的社交关系链,去除境外用户和不活跃用户,共采集到5845329个有效样本用户,占原始样本的80.2%,有效样本用户之间的有效社会关系对为367044740条,数据采集时间从2011-12-01至2012-06-25。

3.2 结果分析

本实验以地级市为空间尺度(台湾地区以县为尺度),共计选择了365个城市。对所有有效用户及用户关系进行地理映射,并对用户关系进行空间聚合,生成了 365×365 的空间矩阵,共计132860条城市链接(不包含城市自身)。通过实验,本文选择三个指标来验证SHDI-GM方法的有效性:①城市节点保留率($\%N_T$);②城市链接保留率($\%E_T$);③城市社交流量保留率($\%W_T$)。图1(a)为原始城际社交关系网络,图1(b)为全局

社交多样性指数阈值 $H^* = 0.2$ 条件下的城际社交关系地理骨干网。该城际社交骨干网络与原始网络相比,城市间的连接数降低了 98.1%, 1.90% 的边保留了原始网络社交流量总和的 67.24%。从图 1(b) 可以发现, 中国城际社交关系网呈现以 5 个中心为主导的“五角星座型”主体空间结构, 这 5 个中心分别为京津、长三角、珠三角、成渝及台北地区, 显示在虚拟社交网络, 中国的城市仍然呈现了层次鲜明的等级结构。

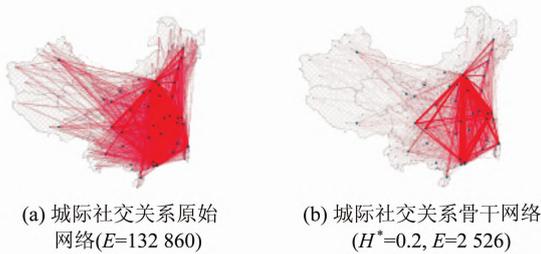


图 1 城际社交原始网络和骨干网络对比图
Fig. 1 Comparison of the Origin and the Backbone of the Inter-city Social Network

表 1 显示的是选取不同全局社交多样性指数阈值 H^* 下骨干网的基本特征。可以发现, 随着 H^* 的降低, 网络的节点数保持不变, 保证了网络的连通性和结构的完备性; 网络的链接规模得到较好的控制, 当 $H^* = 0.8$ 时, 边被剔除了超过 40%, 而网络总社交流量仅仅减少了约 3.5%。随着 H^* 的降低, 网络的链接数量得到了大幅的减少并趋于平衡, 使得网络的总城际社交流量得到最大程度的保留, 同时也使得网络中有全局影响力的城市链接以及在局部社交网络重要的城市链接在全局社交多样性指数阈值变化的过程中都得到了自适应的保留。该方法既保证了骨干网络的节点完备性和网络连通性, 又能最大程度地保留城际社交网络关系的异构特征, 并且通过调节单变量, 能够变尺度控制地理骨干网络的整体规

表 1 不同全局社交多样性指数阈值下的骨干网特征
Tab. 1 Sizes of the Disparity Backbones in Terms of the Percent-age of Total Weight ($\%W_T$), Nodes ($\%N_T$), and Edges ($\%E_T$) for Different Values of the Significance Level H^*

H^*	$\%W_T$	$\%N_T$	$\%E_T$
0.8	96.51	100	57.49
0.5	88.64	100	12.57
0.3	77.60	100	3.84
0.2	67.24	100	1.90
0.05	43.09	100	0.63
0.02	34.29	100	0.58
0.01	30.87	100	0.57

模和细节, 有助于揭示城市社交网络的多尺度的空间结构特征。

4 结 语

利用社交网络的城际关系可以研究城市功能、城市空间关系及体系结构等。但是社交网络中的节点和社交关系复杂, 且为典型的大数据, 为了从复杂的非结构化的社交网络数据研究中国城市体系的空间结构, 本研究根据海量虚拟社交网络服务的用户节点、用户社交关系以及用户的地理位置等信息, 以城市为尺度, 对个体间的社交关系链进行聚合, 构建城市-城市社交关系网; 并在此基础上综合考虑该网络的空间和拓扑特征, 提出了一种融合重力模型和信息熵理论的地理骨干网提取方法。实验验证, 在保持网络节点完备度的前提下, 本文的方法通过调节全局社交多样性指数阈值, 能够变尺度地控制网络规模, 并能保证网络节点的连通性, 在不同尺度下都能最大程度地保留网络中关键的城际链接。该方法能有效减少复杂网络信息的冗余, 有利于降低对大数据信息的理解难度, 这对挖掘虚拟网络环境下中国城市空间结构体系以及相互作用关系有重要的作用。后期将会在此基础上进一步研究虚拟社交网络环境下中国城市体系的组织结构和层次关系, 并探讨城市及城市群辐射范围、城市吸引力和开放性等问题。

参 考 文 献

[1] Gu Chaolin, Pang Haifeng. Study on Spatial Relations of Chinese Urban System: Gravity Model Approach[J]. *Geographical Research*, 2008, 27(1): 1-12 (顾朝林, 庞海峰. 基于重力模型的中国城市体系空间联系与层域划分[J]. *地理研究*, 2008, 27(1): 1-12)

[2] Leng Bingrong, Yang Yongchun, Li Yingjie, et al. Spatial Characteristics and Complex Analysis: A Perspective from Basic Activities of Urban Networks in China[J]. *Journal of Geographical Science*, 2011, 66(2): 199-211 (冷炳荣, 杨永春, 李英杰, 等. 中国城市经济网络结构空间特征及其复杂性分析[J]. *地理学报*, 2011, 66(2): 199-211)

[3] Zhou Yixing, Hu Zhiyong. Looking into the Network Structure of Chinese Urban System from the Perspective of Air Transportation[J]. *Geographical Research*, 2002, 21(3): 276-286 (周一星, 胡智勇. 从航空运输看中国城市体系的空间网络结构[J]. *地理研究*, 2002, 21(3): 276-286)

- [4] Kwan M P. Mobile Communications, Social Networks, and Urban Travel: Hypertext as a New Metaphor for Conceptualizing Spatial Interaction [J]. *The Professional Geographer*, 2007, 59(4): 434-446
- [5] Qin Xiao, Zhen Feng, Xiong Lifang, et al. Methods in Urban Temporal and Spatial Behavior Research in the Big Data Era[J]. *Progress in Geography*, 2013, 32(9): 1 352-1 361 (秦萧, 甄峰, 熊丽芳, 等. 大数据时代城市时空行为研究方法[J]. 地理科学进展, 2013, 32(9): 1 352-1 361)
- [6] Li Qingquan, Chang Xiaomeng, Shaw S L, et al. Characteristics of Micro-blog Inter-city Social Interactions in China[J]. *Journal of Shenzhen University Science and Engineering*, 2013, 30(5): 441-449 (李清泉, 常晓猛, 萧世伦, 等. 中国城际社交关系网络特征分析[J]. 深圳大学学报(理工版), 2013, 30(5): 441-449)
- [7] Yu H B, Shaw S L. Exploring Potential Human Activities in Physical and Virtual Spaces: A Spatiotemporal GIS Approach[J]. *International Journal of Geographical Information Science*, 2008, 22(4): 409-430
- [8] Glattfelder J B, Battiston S. Backbone of Complex Networks of Corporations: The Flow of Control [J]. *Physical Review E*, 2009, 80(3): 036104
- [9] Radicchi F, Ramasco J J, Fortunato S. Information Filtering in Complex Weighted Networks[J]. *Physical Review E*, 2011, 83(4): 046101
- [10] Kim J R, Kim J, Kwon Y K, et al. Reduction of Complex Signaling Networks to a Representative Kernel[J]. *Sci Signal*, 2011, 4(175): ra35
- [11] Serrano M N, Boguna M N, Vespignani A. Extracting the Multiscale Backbone of Complex Weighted Networks[J]. *The National Academy of Sciences*, 2009, 106(16): 6 483-6 488
- [12] Chen Jiangfeng. Extract Backbones of Large-Scale Networks Using Data Field Theory[J]. *Journal of Harbin Institute of Technology (New Series)*, 2010, 17(1): 119-122
- [13] Zhang Q M, Zeng A, Shang M S. Extracting the Information Backbone in Online System[J]. *PLoS ONE*, 2013, 8(5): e62624. doi:10.1371/journal.pone.0062624
- [14] Wang Jiao'e, Mo Huihui, Jin Fengjun. Spatial Structural Characteristics of Chinese Aviation Network Based on Complex Network Theory[J]. *Journal of Geographical Science*, 2009, 64(8): 899-910 (王皎娥, 莫辉辉, 金凤君. 中国航空网络空间结构的复杂性[J]. 地理学报, 2009, 64(8): 899-910)
- [15] Leng Bingrong, Yang Yongchun, Hung Xing, et al. Analysis on Complex Characteristics in Chinese Urban Networks Structure Based on Gravity Model Since 1990s[J]. *Areal Research and Development*, 2011, 30(5): 66-71 (冷炳荣, 杨永春, 黄幸, 等. 1990年代以来基于重力法的中国城市网络结构复杂性分析[J]. 地域研究与开发, 2011, 30(5): 66-71)
- [16] Kang C G, Zhang Y, Ma X J, et al. Inferring Properties and Revealing Geographical Impacts of Inter-city Mobile Communication Network of China Using a Subnet Data Set[J]. *International Journal of Geographical Information Science*, 2013, 27(3): 431-438
- [17] Keylock C J. Simpson Diversity and the Shannon-Wiener Index as Special Cases of a Generalized Entropy[J]. *Oikos*, 2005, 109(1): 203-207

Extracting the Geographic Backbone of Location-based Social Network

CHANG Xiaomeng^{3,1} YUE Yang² LI Qingquan^{2,1} CHEN Biyu¹
SHAW Shihlung^{3,1} TU Wei²

1 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

2 Shenzhen Key Laboratory of Spatial Smart Sensing and Services, Shenzhen University, Shenzhen 518060, China

3 Department of Geography, University of Tennessee, Knoxville 37996-0925, USA

Abstract: Based on the massive amount of users social relationship and location information data collected from the largest social network website in China-SINA micro-blog, an inter-city geographic social network was built from the aggregated inter-person linkages mapped from virtual space onto the physical space. Considering about the topological characteristics such as the global and local heteroge-

(下转第 718 页)

- [37] Crandall DJ, Backstrom L, Cosley D. Inferring Social Ties from Geographic Coincidences [J]. *The National Academy of Sciences*, 2010, 107(52): 22 436-22 441
- [38] Pham H, Hu L, Shahabi C. GEOSO—A Geo-social Model: From Real-world Co-occurrences to Social Connections[J]. *Databases in Networked Information Systems*, 2011, 7 108:203-222
- [39] Zheng Y, Zhang L, Xie X. Recommending Friends and Locations Based on Individual Location History [J]. *ACM Transaction on the Web*, 2011,5(1):60-67

An Integrated Data Model for Spatial-Temporal Trajectories and the Social Relationships of Moving Objects

ZHANG Hengcai¹ LU Feng¹ CHEN Jie¹

1 State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

Abstract: In the pan geographic information age, with the development of the SoLoMo (social, local and mobile) modes in the mobile internet industry, how to effectively manage massive location-based data, social network and geospatial data has become a major challenge in the moving object databases field of research. In this paper, a spatial-temporal integrated data model for moving object trajectories and social relationships, called the Geo-Social-Moving (GSM) model, is proposed to manage both moving object geographical locations, trajectories and social relationships in an integrated framework. We define a series of common operations including interception, extension, projection and multi-dimensional projection. Then, the GSM model is implemented based on a graph database system. We conducted detailed experiments with two free location-based social network datasets from Brightkite and Gowalla to verified the effectiveness of the GSM model. Results show that the proposed GSM model can avoid query inefficiency caused by relational database table joins with the efficiency and stability to provide support for location-based social network applications.

Key words: moving objects; data model; trajectory; social network

First author: ZHANG Hengcai, PhD. He focus on moving objects database and spatial-temporal data mining. E-mail: zhanghc@lreis. ac. cn

Corresponding author: LU Feng, PhD, researcher, PhD supervisor. E-mail: luf@lreis. ac. cn

Foundation support: The National High Technology Research and Development Program of China(863 Program), Nos. 2012AA12A211, 2013AA120305; the National Natural Science Foundation of China, Nos. 41271408, 41101149.

(上接第 710 页)

neity, and the spatial interactions among cities, a hybrid geographic backbone extraction approach is proposed which is drawn from the gravity model and information entropy theories. This research can increase understanding of the structure of urban systems, radiant ability, attractiveness and openness of cities in the virtual web society.

Key words: social network; spatial structure; geographic backbone; Shannon entropy; gravity model

First author: CHANG Xiaomeng, PhD candidate, specializes in geographic social network, spatio-temporal data visual analysis, and ITS. E-mail: changxiaomeng@gmail. com

Corresponding author: LI Qingquan, PhD, professor. E-mail: liqq@szu. edu. cn

Foundation support: The National Natural Science Foundation of China, No. 41371377; Shenzhen Scientific Research and Development Funding Program, No. ZDSY20121019111146499; Shenzhen Dedicated Funding of Strategic Emerging Industry Development Program, No. JCYJ20121019111128765; CCF-TencentARG20130115.