

大数据时代的交通工程

王璞¹, 黄智仁², 龚航¹

(1. 中南大学交通运输工程学院 长沙 410075; 2. 中南大学软件学院 长沙 410075)

【摘要】面向当代严重的交通拥堵, 交通规划与组织水平受到人们的广泛关注, 交通需求作为交通规划与组织的重要基础数据, 其获取方法一直以来都是非常重要的研究课题。进入大数据时代, 复杂网络与人类出行研究领域快速发展, 为交通需求预测与交通网络优化提供了崭新途径。该文综述人类出行与交通网络研究方向的重要成果, 展望这些成果为交通工程带来的契机, 并介绍两个方向的最新交叉研究。

关键词 大数据; 人类出行特性与建模; 交通网络; 交通工程

中图分类号 N94

文献标志码 A

doi:10.3969/j.issn.1001-0548.2013.06.002

Transportation Engineering in the Big Data Era

WANG Pu¹, HUANG Zhi-ren², and GONG Hang¹

(1. School of Traffic and Transportation Engineering, Central South University Changsha 410075;

2. School of Software, Central South University Changsha 410075)

Abstract Faced with severe traffic congestions, the level of traffic planning and organization has been paid much attention. As a crucial and fundamental data for traffic planning and traffic organization, travel demand has been an important research topic for a long time. Stepping into the big data era, the fast developments in the research fields of human mobility and complex network offer new methodologies for predicting travel demand and improving transportation networks. In this paper, we review the important works in the areas of human mobility and transportation network, discuss the potential opportunities which bring to transportation engineering, and finally introduce a new interdisciplinary study of the two research areas.

Key words big data; human mobility laws and modeling; transportation network; transportation engineering

交通工程是一门交叉学科, 从其诞生伊始便大量运用数学、运筹学、物理学和心理学的研究方法。交通工程理论与技术的革新和很多研究领域的发展密切相关, 上世纪末智能交通系统的发展依靠的就是成熟的信息、通讯、传感、控制和计算机技术。进入21世纪, 计算机的存储和计算能力进一步提高, 存储和计算成本大幅下降, 大量宝贵数据得以保存与分析。在这个数据爆发的时代, 一方面, 复杂网络科学兴起, 为探索各种复杂系统的拓扑结构、演化机理提供了崭新的理论与方法; 另一方面, 基于各种大数据, 学术界对人类出行特性有了更深的认识, 建立了更精确、更高效的人类出行预测模型。那么, 大数据时代复杂网络和人类出行研究的蓬勃发展能为交通工程领域带来哪些新的契机呢? 本文通过综述两个方向的重要科研成果对此进行展望。

1 人类出行特性分析与建模

大数据驱动下的人类出行统计特性研究始于科学家对疫情传播模型的深层次思考。2004年, 美国科学院院刊刊发了Borckmann与其合作者的关于SARS病毒传播预测的研究成果^[1]。文献[1]建立了基于航空旅客运输网络的SARS疫情传播预测模型, 并得到了与实际统计数据相吻合的预测结果。由于仅考虑了航空出行导致的疫情传播, 在更小的空间尺度上, 如相邻城市之间, 该模型无法预测疫情传播。此后, Brockmann一直在寻找记录人类中短程出行的数据。终于在一次访友中, 他得知wheresgeorge.com拥有大量美元钞票流通的时间位置数据^[2]。钞票的流通间接地记录了人类的出行, 文献[2]发现钞票的位移分布服从幂律衰减, 如图1所示, 具有类似Lévy

收稿日期: 2013-10-15; 修回日期: 2013-11-01

基金项目: 国家自然科学基金(51208520)

作者简介: 王璞(1983-), 男, 教授, 博士生导师, 主要从事交通运输规划与管理、统计物理学、复杂网络和数据挖掘等方面的研究。

飞行的特征: 存在大量短距离移动, 同时也有一些长距离移动。文献[2]还发现钞票的扩散速度远低于具有相同幂指数的Lévy飞行的扩散速度, 分析其原因, 发现钞票在两次移动之间的停留时间也服从幂律分布, 如图1所示。因此, 作者采用同时具有幂律步长分布和停留时间分布的连续时间随机游走(CTRW)模型来模拟钞票的空间移动, 很好地再现了钞票低速扩散的特征。

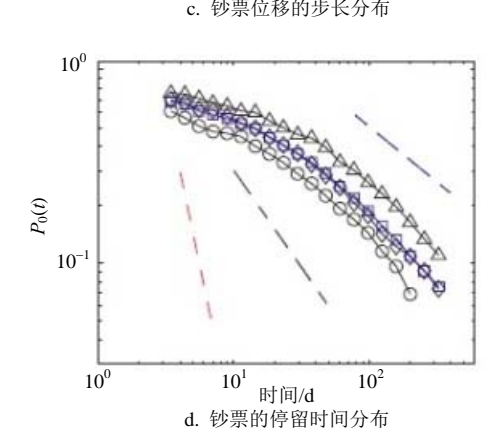
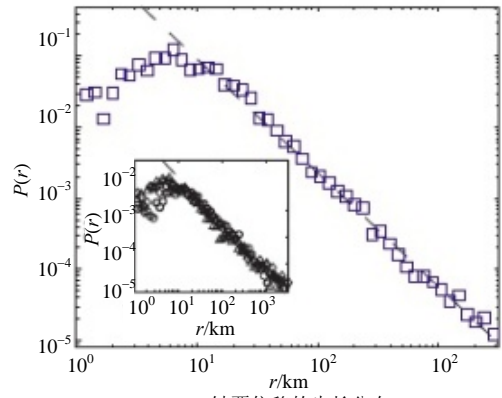
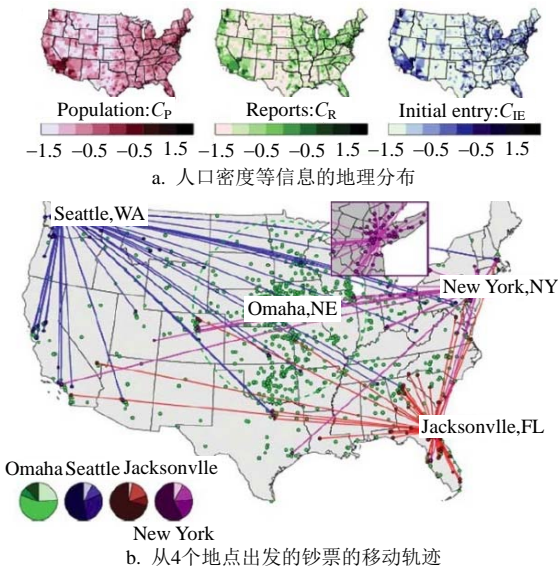


图1 美元钞票位移的步长分布与停留时间分布^[2]

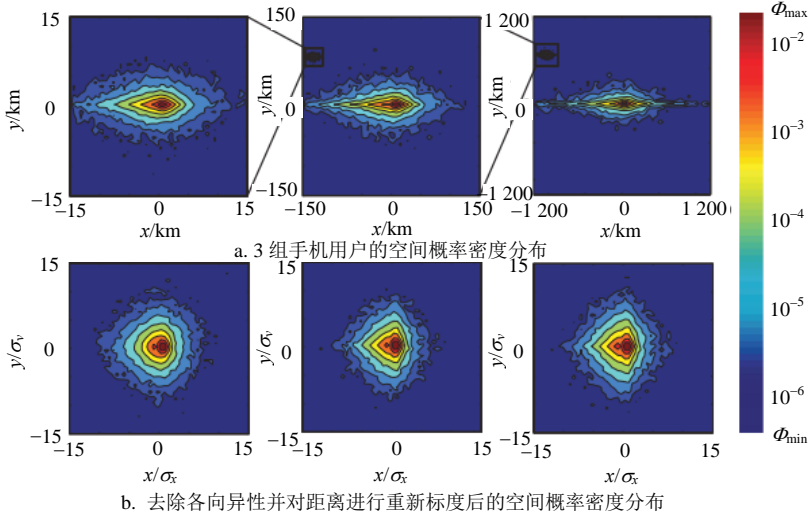


图2 不同出行距离出行者的位置概率密度分布^[3]

一张钞票的流通并不是由同一个人完成, 使用钞票流通数据研究人类出行的步长分布规律是可行的, 但是对于个体出行特征的更深层次研究, 该数据是不能胜任的。每当人们使用手机时, 时间和服务通讯的基站会被记录, 所以手机数据中蕴含着大量的时间、位置信息。文献[3]基于600万手机用户长达6个月的账单记录深入研究了手机用户的出行特

征, 发现个体位移的步长服从带有指数截断的幂律分布, 和文献[2]得到结果非常相近。在个体出行特征方面, 发现个体平均出行距离差异显著, 个体对地点的访问频率也服从幂律分布, 不同平均出行距离的个体在空间运动位置分布上具有很高的相似性。图2展示了3组手机用户的空间概率密度分布, 尽管3组用户的日活动范围差异很大, 当对个体出行

方向主轴重新定向并对距离进行重新标度后，他们的空间概率密度分布非常相似。这说明人类可能遵循着某种普适的出行模式，该发现为建立个体位置预测模型提供了理论基础和保障。

这两篇开创性工作掀起了学术界探索人类出行特性的热潮。基于手机数据，文献[4]研究了个体和群体的人类行为动力学，测量了不同时段手机用户的移动概率；文献[5]分析了纽约附近的一个郊区城市的人员流动，指出其研究成果能够帮助城市规划者了解城市动态；文献[6]研究了实时通信和人类出行在紧急情况(炸弹、袭击、地震)和非紧急情况(音乐会、体育赛事)下的反应模式。基于GPS数据，文献[7]跟踪了车辆的详细轨迹，研究了城市居民出行的频繁模式、交通密集区的预测方法、交通拥堵的描述方法；文献[8]利用GPS数据再次发现人类出行距离服从幂律分布。基于出行调查数据，文献[9]研究了国内两个经济发展程度不同的城市的居民出行。作者发现在经济发达的深圳地区，居民的出行模式类似于发达国家居民的出行模式。而在欠发达的安徽农村地区，人们的出行距离非常局限；文献[10]比较分析了北京和纽约的小汽车保有量、小汽车出行出发时间分布、出行距离、出行目的和车载人数分布的差异。从能源消耗上角度，文献[11]发现在不同的出行方式下，人们的出行距离服从一个普遍分布规律。文献[12]分析了瑞士弗劳恩菲尔德市230名志愿者填写的日常出行日志，作者发现群体的出行距离近似服从一个带有指数截断的幂律分布，与基于手机数据得到的实证研究结果非常接近，但绝大多数个体的出行距离并不符合幂律分布。如图3所示，学生和职员更容易在固定的地点之间频繁移动，出行距离呈现有峰值的分布。该研究更深入地理解了人类出行位移步长的幂律分布。

人类出行的统计特性是出行建模的基础。通过分析人类出行特征，研究人员建立了多种出行模型。文献[13]中，研究人员为100名麻省理工学院的学生与职员提供了智能手机，这些智能手机采集了实验人员长达400 000 h的位置数据。基于该数据作者提出了行为空间的概念，并利用特征向量计算特征行为，预测实验人员各小时的位置，准确率可达95%，如图4所示。该模型的预测准确度在不同的群体中有所差异，高年级学生的位置较容易预测，而商学院学生的位置较难预测。通过获取实验人员一天中前12 h的位置数据，该模型可以以平均79%的准确率预测该人员后12 h的位置。

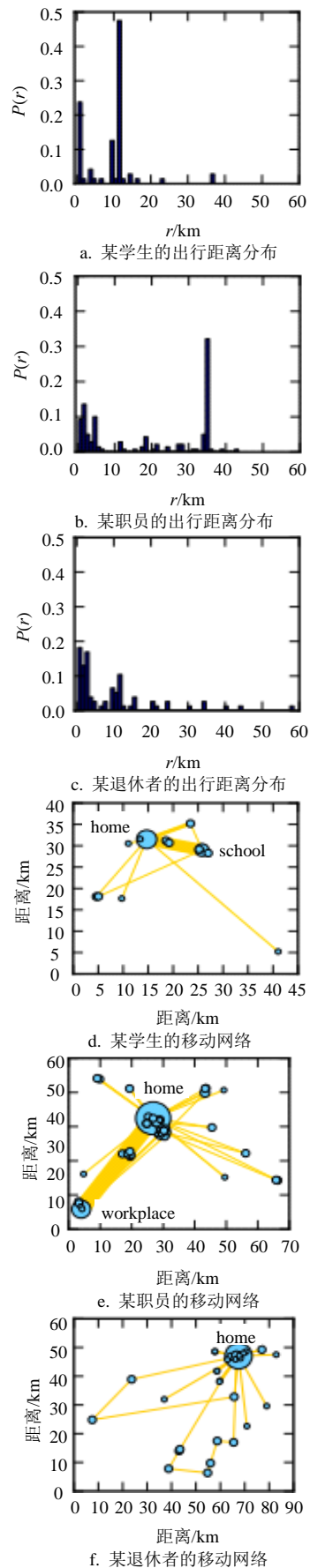


图3 三种不同职业个体的出行位移分布^[12]

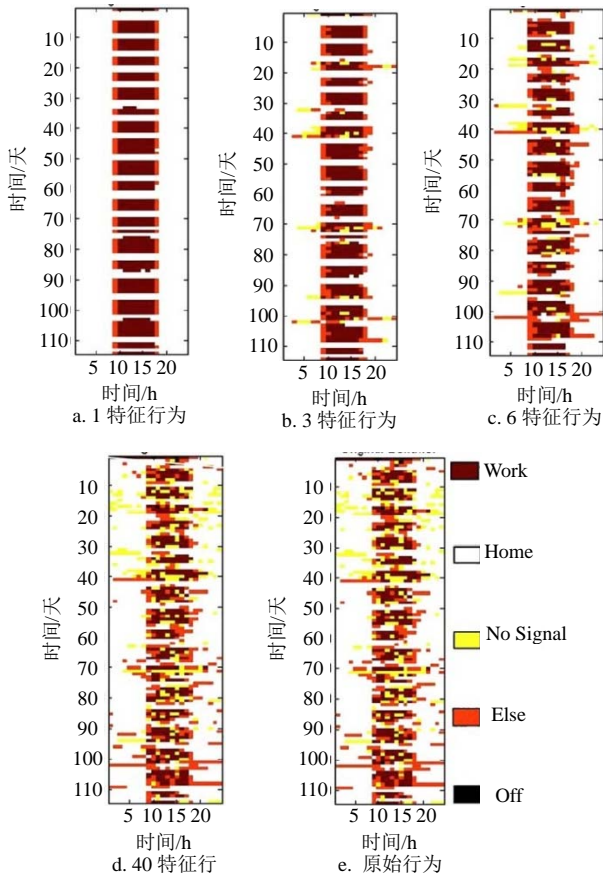
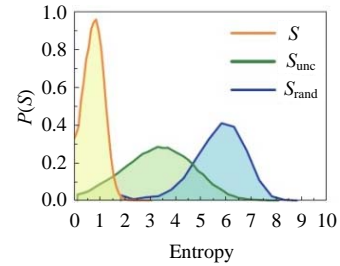


图4 基于不同数量特征行为估计实验者115天的行为^[13]

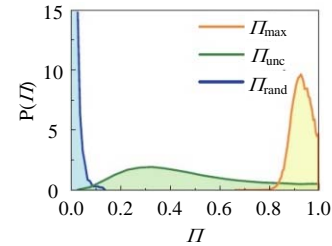
文献[13]非常准确地预测了100名实验人员的小时位置。那么, 人类空间位置到底有多大的可预测性? 文献[14]回答了这个问题, 作者从1 000万手机用户中选取了5万具有高频率位置记录的手机用户(平均2 h至少有一个记录), 并建立了这些手机用户的出行网络。文献[14]基于信息论中的Fano不等式计算了每个用户的访问地点能够被准确预测的概率上限 Π_{max} , 发现 Π_{max} 的分布在0.93处有峰值, 且方差很小, 如图5所示。即平均而言人类位置最高有93%的预测准确率, 并且可预测性与年龄、性别、出行距离等参数关联不大。该研究作为人类出行建模奠定了坚实的理论基础。

文献[15]利用手机数据分析了用户空间位移的标度律, 发现: 1) 随着时间的增长, 个体访问新地点数量服从 $S(t) = t^{-\mu}$, $\mu \sim 0.6$; 2) 个体对第 k 个最可能访问地点的访问频率服从 $f_k = k^{-\xi}$, $\xi \sim 1.2$; 3) 个体均方位移(MSD)超缓慢增长。作者指出传统CTRW模型不足以描述人类运动的统计特性, 并建立探索回归模型再现人类运动的统计特性。探索回归模型的主要思想是: 个体以 $\rho S^{-\gamma}$ 的概率探索新地点, 以 $1 - \rho S^{-\gamma}$ 的概率返回以前访问的地点, 访问

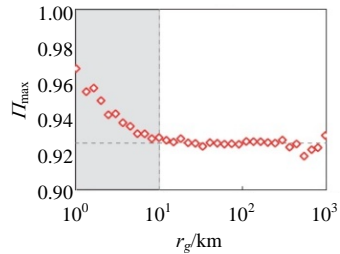
某地点的概率与个体在该地点被发现的概率成正比, 如图6所示。



a. 手机用户3种熵的分布图



b. 3种熵对应的可预测性最大值分布



c. Π_{max} 与回转半径 r_g 的关系

图5 人类位置的高可预测性^[14]

文献[16]基于手机用户的时变移动概率、出行位置分布和出行方向主轴, 预测了600万手机用户一个月的小时位置, 建立了一个能预测大规模人群位移轨迹的高效模型。文献[17]提出了人员流动的多层网络模型, 并用于预测猪流感的大范围地域传播。文献[18]提出了一个基于个体行为和集体行为的个体位置预测模型, 并利用波士顿地区的手机数据验证。文献[19]指出交通系统的层次性影响着人类出行行为标度性特征的形成, 作者采用了一套基于概率流平衡的解析方法对此模型进行了解析研究。文献[20]采用统计方法估计个体时空路径, 并通过定义位置稳定指数衡量个体出行的规律性和稳定性。文献[21]基于手机位置数据、浮动车数据在地理信息系统平台上对城市居民出行进行仿真。

重力模型是预测区域间人口流动、货物运输、城际通讯与国家贸易的常用模型, 在交通工程“四阶段法”中的交通分布阶段中有着广泛的应用。然而, 重力模型仍然需根据区域特征的不同动态调整参数。为了解决这个问题, 文献[22]建立了“辐射模型”, 该模型预测区域间人口流动的精度比传统重力

模型高，且不需要进行大规模的交通调查即可实施

预测。其具体形式是：

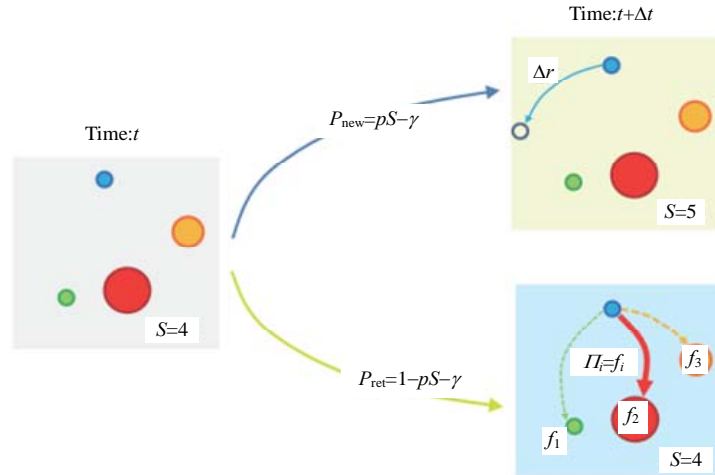


图6 探索回归模型示意图^[15]

$$\langle T_{ij} \rangle = T_i \frac{m_i n_j}{(m_i + S_{ij})(n_j + S_{ij})}$$

式中， T_i 是从*i*区域出发的通勤总数； m_i 是出发地*i*的人口； n_j 是工作地*j*的人口； S_{ij} 是*i*与*j*间圆弧内的人口，如图7所示。辐射模型具有无参数特性，可在缺乏先前测量数据的情况下，只需要易于采集的人口信息便可使用。

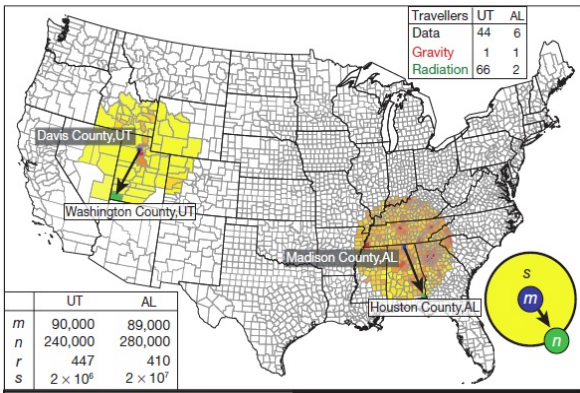


图7 辐射模型示意图^[22]

以上简要介绍了人类出行特性与建模研究方向的发展历程及标志性成果，关于该方向更详细的综述请见文献[23]，下面展望人类出行研究领域的研究成果可以为交通工程的发展带来哪些崭新途径。交通工程领域的两类基本模型是集计模型和非集计模型。集计模型的典型代表是“四阶段法”，自20世纪60年代建立以来一直在工业界广泛应用。“四阶段”包括交通生成、交通分布、交通方式划分、交通分配。前三个阶段估计某种出行方式的OD矩阵，第四阶段将OD矩阵中的出行分配在交通网络上。20世纪70年代，针对“四阶段法”的理论欠缺，McFadden、

Ben-Akiva等著名科学家提出非集计模型，其基本假设是当出行者面临选择时，对某种选择的偏好可以用被选择对象的“效用值”来描述。非集计模型在交通方式划分和交通分配阶段有着十分广泛的应用。两类模型分别从宏观和微观角度估计交通需求和交通状态。在宏观方面，“四阶段法”中的交通需求预测多通过采集实际数据，建立统计学模型实现，而辐射模型^[22]的理论推导不仅为交通分布提供了理论基础，同时该模型还具有无参数、预测精度高的特点，可在缺乏先前测量数据的情况下使用。因此，辐射模型可以革新“四阶段法”中交通分布模块的建模框架。在微观方面，传统非集计模型主要应用于交通方式划分方面，在交通生成与交通分布预测中应用很少。文献[13]提出的人类特征行为模型，文献[15]提出的探索回归模型以及文献[16]建立的大规模人群出行预测模型都属于建立在实证数据上的非集计模型，在交通生成与分布阶段可以有很广泛的潜在应用。尽管当前人类出行领域的研究成果在疫情传播预测及控制领域得到了广泛的应用，但在城市规划、交通预测组织等领域的研究还比较欠缺，将人类出行建模与网络交通流分配结合的交叉研究非常少；而交通工程领域也更注重传统方法的优化，似乎两个领域都没有把对方的经典方法融入到本领域的研究之中，这可能源于两个领域学者研究背景和技术手段存在较大的差异。另外，人类出行研究领域在出行方式识别方面的工作还非常欠缺，亟待发展。

2 交通网络分析与优化

复杂网络科学的兴起不仅为信息科学、交通科

学、生命科学、社会科学等学科领域提供了技术支撑, 而且为理解复杂系统的性质和功能提供了理论保障。早在复杂网络科学兴起之初, 就有大批学者把工作重点放在交通网络的分析与优化方面。本文综述的交通网络研究限于道路、铁路等陆路交通。航空、海运、管道等交通的研究方法与陆路交通的研究方法存在较大差异, 故不做讨论, 有兴趣的读者可以参考相关文献。

交通网络拓扑结构分析是网络优化的基础, 因此, 本文首先介绍交通网络拓扑结构分析方向的一些研究成果和方法。文献[24-25]研究了瑞士Gävle市的道路网络, 发现了路网的分级结构。Gävle市的道路网络中80%的道路是普通道路, 20%的道路是重要道路, 这20%的重要道路承载了城市80%的交通流量。该研究丰富了人们对城市道路组织形式的认识。拓扑结构分析中的一个重要测量就是中心度, 文献[26]研究了世界18个城市的路网结构, 测量了路网节点的4种中心度: 紧密中心度 C^C , 介数中心度 C^B , 平直中心度 C^S , 信息中心度 C^I , 如图8所示。通过分析中心度的空间分布和统计分布, 作者指出中心度测量可以用来拓展城市结构和特性的描述。为了探索交通网络的演化机制, 文献[27]提出了一个简单的基于局部优化和叶脉结构的模型。这个模型的统计特性与所观察到的道路分布模式吻合, 作者指出在不受宏观控制的情况下, 许多不同交通网络的演化遵循一个简单的普遍机制。

文献[28]以一种崭新的信息观点研究了城市路网的拓扑结构, 将每条道路映射为信息城市网络中的节点, 其中拥有较多交叉口的长道路成为网络的中心节点, 一般支路成为普通节点, 如图9所示。文献[28]发现信息城市网络的度分布范围很广。为了量化一个城市导航的难易程度, 定义了搜寻信息 S , 较高的 S 代表新的访问者需要更多信息发现路径。文献[28]研究了曼哈顿路网, 瑞典三个城市的路网及机场网络的平均信息搜寻难度, 发现各网络的导航难度都要大于其相应的随机网络。定义 $\delta = (S - S_r) / \log_2(N)$ 衡量一定路网规模的导航效率, 发现机场与曼哈顿的网络搜寻信息 S 与其等规模的随机网络的搜寻信息 S_r 接近, 而传统城市道路的层次性较强, 存在较偏僻的道路, S 与 S_r 的差距较大。

文献[29]研究了由物理基础设施层和交通流量层构成的多层交通网络, 提出了一个普适的多层模型, 分析了三个尺度的交通网络(华沙公交网、瑞士

铁路网、中欧铁路网)。该研究测量了交通网络边的度、介数、实际负载。通过测量Pearson相关系数, 发现几者之间的相关性是很差的, 度与边介数的相关性为0.5, 介数与实际负载的相关性仅为0.26, 如图10所示。文献[29]指出交通网络拓扑结构研究与交通需求研究的分离, 两个方向的研究都存在着忽视对方研究成果和方法的情况。

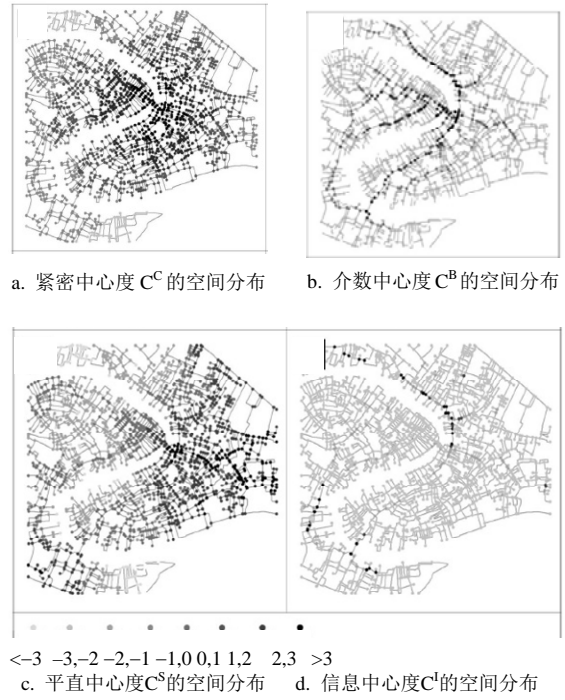


图8 紧密中心度 C^C 、介数中心度 C^B 、平直中心度 C^S 、信息中心度 C^I 的空间分布^[26]

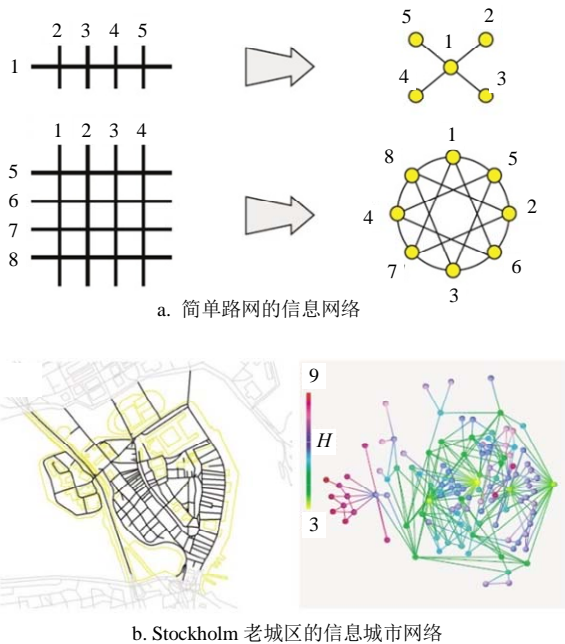
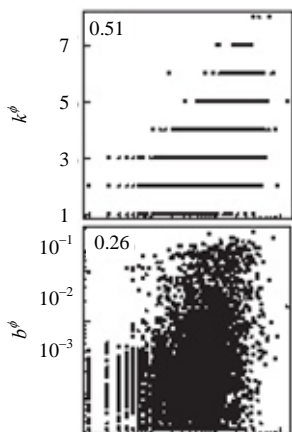


图9 城市道路网络的信息表示^[28]



c. 铁路网络节点的实际负载 l



d. 实际负载与度和介数的相关性

图10 铁路网络中边的度、介数、实际负载的分布^[29]

通过网络拓扑分析，可以定位网络中的关键部分以提高网络的效率。文献[30]构建了SF网络、ER网络和晶格网络，并在这些网络的边上均匀赋以0~1之间的权重以模拟交通网络。作者发现这些构造网络的最小生成树(MST)分为两个截然不同的部分，具有显著不同的运输性质。其中一部分是超级高速

公路(IIC superhighway)，另一部分是普通公路，改善小部分超级高速公路可以大幅提升整个道路网络的运输性能，如图11所示。作者测量了MST中IIC的使用率，随着路网规模的增大，IIC中的边在整个路网中占有的比例会下降，但其使用率却不随之下降，趋近于常数。作者通过两个实例(随机电阻网络，计算机科学中的最大流问题)验证了提高IIC的容量要比提高相同数量最高介数的边在提升网络运输能力方面具有更好的效果。

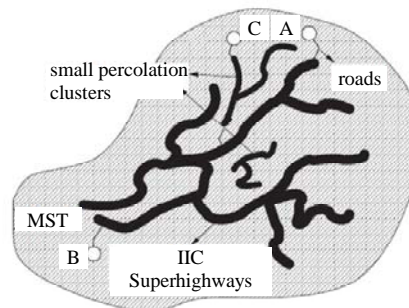


图11 最小生成树中的超级高速公路和普通公路^[30]

近年来，晶格网络成为研究网络导航优化的重要载体，一个重要的问题是：如何搭建长距离链接优化晶格网络的导航效率，如图12所示。文献[31]在两点*i*和*j*之间以概率 $P = r_{ij}^{-\alpha}$ 搭建长距离连接(r_{ij} 为*i*与*j*之间的曼哈顿距离)，发现在仅使用网络局部信息的情况下，当 $\alpha = 2$ 时平均信息传递时间最短。文献[32]把搭建长距离链接的成本约束条件引入，当长距离链接的总长度达到*N*时停止搭建长距离链接(*N*为晶格网络的节点总数)。当 α 较大的时，长距离链接长度短但数量多，当 α 较小的时，长距离链接长度长但数量少。文献[32]发现无论利用网络局部信息还是全局信息，当 $\alpha = 3$ 时，网络的平均最短距离最小，如图12所示。导航优化问题至今仍是学术界的一个热点研究问题，更多的研究工作请见文献[33-40]。在晶格网络上，长距离链接的运输费用*C*多被假定为一个单位长度($C = 1$)，这在社会网络上可行的，但在空间网络上是不合理的。例如，道路网络中的高速公路可以被看作长距离链接，汽车使用高速公路(长距离链接)的时间大概是使用普通公路(短距离链接)的一半($C \sim 0.5r_{ij}$)，并不是很多理论工作中过分低估的交通运输费用($C = 1$)。另外，当前导航优化研究多假设交通需求均匀分布，这与实际附加在交通网络上的交通需求有很大的出入，文献[29, 53]等都指出了交通需求在网络节点间的不均匀分布。在经典导航优化模型中考虑更加真实的边

属性和交通需求分布是评价和优化实际交通网络的必要条件。

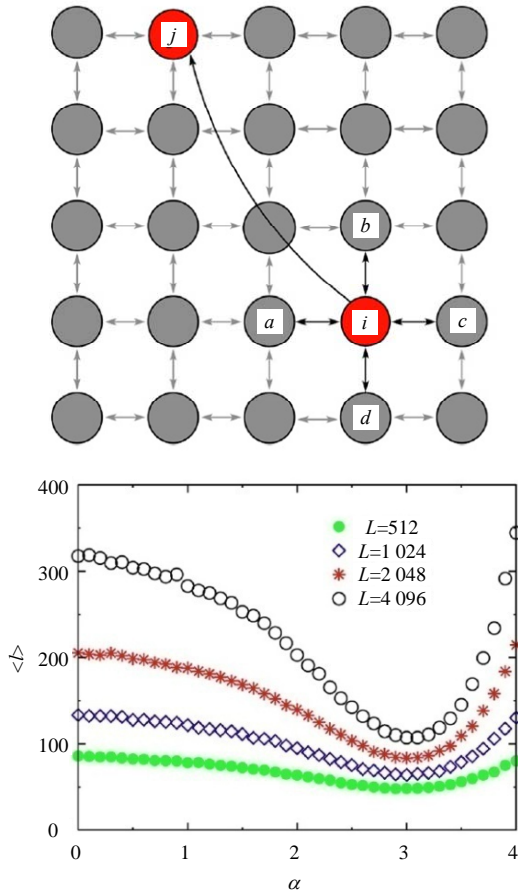
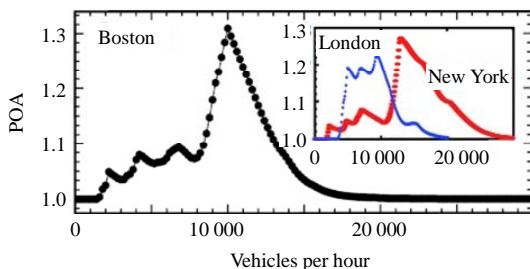
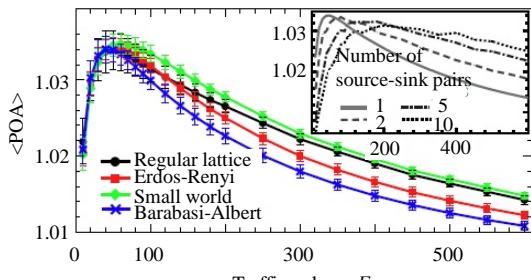


图12 晶格网络中的长距离链接与最优导航参数^[32]

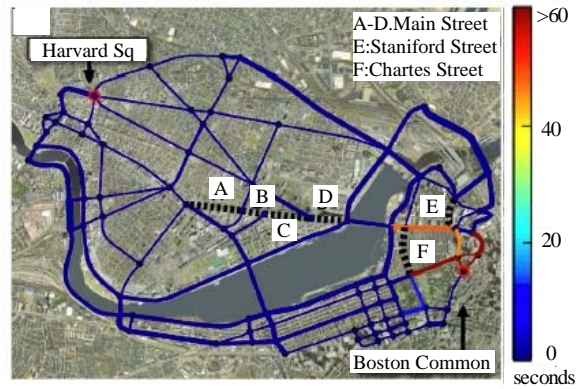


a. 3 城市不同交通流量下的 POA

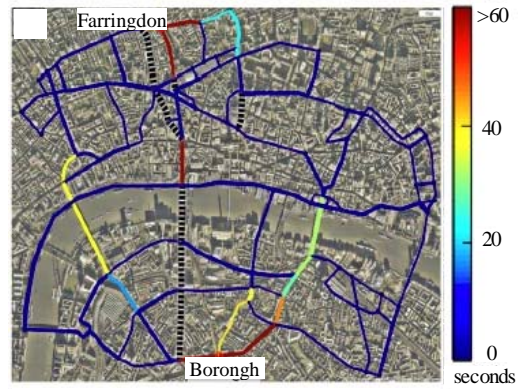


b. 理论网络的 POA

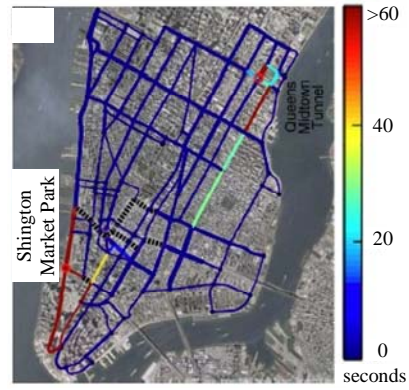
图13 交通网络与理论网络中混乱的代价^[41]



a. 关闭波士顿某路段引起的平均行驶时间延迟



b. 关闭伦敦某路段引起的平均行驶时间延迟



c. 关闭纽约某路段引起的总行驶时间延迟

图14 交通网络中布雷斯悖论的验证^[41]

近年来,更实际的交通网络得到了更多的关注,文献[41]就是一个很有趣的例子。作者构建了波士顿、伦敦、纽约的简单路网,研究了个体最优和系统最优的差异以及交通网络结构与交通效率的关系。作者阐述了两个概念:用户平衡(user equilibrium)与系统最优(social optimal)。用户平衡是个体追求自身利益最大化所导致的平衡状态,而系统最优是通过牺牲某些个体的利益达到系统效率的最优状态。文献[41]中定义混乱的代价(price of anarchy, POA)为用户平衡情况下的总行驶时间和系统最优情况下的总行驶时间之比,用以度量个体追求自身利益对

系统整体造成的效率损失。作者发现POA在一定的出行量达到峰值,说明当出行量较少或较大时,POA都比较小,当出行量达到某个特定值时,由于人们追求自身利益最大化而造成的系统效率损失达到极值,如图13所示。另外作者还发现:删去路网中的某些路段会使出行时间大幅增长,而删去某些路段反而会使出行时间降低,如图14所示,为交通领域中的布雷斯悖论,这说明调整网络结构可以提高网络的效率。

3 大数据与交通

大数据是复杂网络科学、人类动力学蓬勃发展的基础,而复杂网络科学和人类动力学的发展又会为交通工程提供崭新的建模途径。实际上,大数据挖掘技术已经被广泛应用到交通工程研究,特别是在交通需求预测方面,由于传统的OD调查分析方法耗费巨大的人力物力,利用各种信息资源进行数据挖掘,分析获取OD流成为了新的研究方向。

基于路段车流量数据,文献[42-44]研究了基于道路交通计数的OD矩阵估计算法。基于车载GPS位置数据,文献[45]估计了高速公路和少量主干公路的动态OD矩阵。文献[46]研究了基于车牌号数据的OD调查方法。基于手机数据,文献[47]对波士顿地区的每一个手机用户进行居住地点和工作地点定位,估计通勤OD矩阵;文献[48]基于手机用户的移动轨迹估计日OD矩阵,并根据用户提供的交通问卷数据(年龄、家庭收入等社会经济因素)对该方法进行了拓展。基于手机数据的交通流预测研究方向,文献[49]指出当前数据采集技术不能满足动态交通分配模型的所有需求,而手机数据可以同时提供流量,行驶时间和动态交通需求。文中作者详细阐述了利用手机数据预测网络动态交通流的模型。文献[50]指出手机数据可即时使用的优点,提出了两个基于贝叶斯框架和手机数据的交通估计模型。文献[51]基于手机定位技术、分层结构、手机分类聚类 and 地图匹配技术研究了估计交通信息的方法,证明手机数据在智能交通中的广泛应用前景。文献[52]开发了一个基于手机数据和GIS的城市交通显示系统,此系统可以显示实时交通状况,并具有最佳路径选择、OD数据总结以及行程统计的功能。

文献[53]是近期一项将大数据挖掘、复杂网络、人类出行建模有机融合的交通研究工作。文献[53]从大规模手机数据中提取出行分布信息,利用机动车使用率数据、单位时段交通量数据生成一天四个

时段的出行OD矩阵,提出了利用手机数据等多数据源预测交通需求的建模框架。作者通过经典的“四阶段法”预测交通流,并以文献[13-14]发现的人类位置的高预测性为理论保障定位手机用户住址小区(车源)。基于二分网络思想,作者关联交通网络与车源小区,建立“道路使用网络”,不仅预测道路交通流,而且定位道路上的车辆来源,如图15所示。文献[53]发现不同交通小区的“拥堵贡献”存在很大差异,导致拥堵的大部分车流仅出自少量交通小区。由于手机定位信息便于获取,采集成本低,这项成果可为优化交通组织方案、有效缓解交通拥堵提供重要科学依据。该工作把“溯源”的想法应用到道路网络的拥堵缓解,可以广泛地推广到各种交通网络(地铁网络、数据传输网络等)。

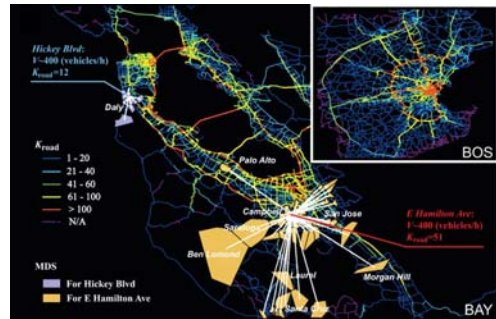


图15 旧金山湾区路网和波士顿路网中路段的主要车源^[53]

通过介绍大数据在交通中的一些最新应用,可以看出大数据在交通领域受到了很大的重视,多源信息挖掘已成为交通研究中的重要方法。不同特性的数据被用来研究不同的问题,例如浮动车数据多被用来估计路段车速而不被用来估计OD矩阵,而手机数据由于纪录频率低不能被用在估计路段车速。根据不同的研究问题和研究目的,数据挖掘的方法可能是很不相同的。合理、充分、巧妙地挖掘大数据中的交通信息将是未来的一个重要研究课题。

4 总结

人类出行和复杂网络领域的丰硕成果可以为“四阶段法”提供崭新的建模元素;大数据挖掘技术与人类出行建模可以为交通需求预测提供底层支持;复杂网络分析可以为交通网络优化提供重要工具。怎样通过大数据挖掘,利用复杂网络与人类出行建模工具为交通工程带来新的研究思维是一个很值得思考的问题。本文综述了人类出行特性与建模及交通网络研究方向的重要科研进展,指出了两个研究方向当前研究的欠缺,展望了其研究成果在交通工程领域的应用前景。最后,本文介绍了一项把

大数据挖掘、复杂网络理论、人类出行建模技术融合的研究工作。

参 考 文 献

- [1] HUFNAGEL L, BROCKMANN D, GEISEL T. Forecast and control of epidemics in a globalized world[J]. PNAS, 2004(101): 15124-15129.
- [2] BROCKMANN D, HUFNAGEL L, GEISEL T. The scaling laws of human travel[J]. Nature, 2006(439): 462-465.
- [3] GONZÁLEZ M C, HIDALGO C A, BARABÁSI A L. Understanding individual human mobility patterns[J]. Nature, 2008(435): 779-782.
- [4] CANDIA J, GONZÁLEZ M C, WANG P, et al. Uncovering individual and collective human dynamics from mobile phone records[J]. Journal of Physics A: Mathematical and Theoretical, 2008(41): 224015.
- [5] BECKER R A, CACERES R, HANSON K, et al. A tale of one city: Using cellular network data for urban planning[J]. IEEE Pervasive Computing, 2011(10): 18-26.
- [6] BAGROW J P, WANG D, BARABÁSI A L. Collective response of human populations to large-scale emergencies[J]. PLoS One, 2011(6):1-8.
- [7] GIANNOTTI F, NANNI M, PEDRESCHI D, et al. Unveiling the complexity of human mobility by querying and mining massive trajectory data[J]. VLDB JOURNAL, 2011(20): 695-719.
- [8] JIANG B, YIN J, ZHAO S. Charactering the human mobility pattern in a large street network[J]. Physical Review E, 2009(80): 021136.
- [9] GARSKE T, YU H, PENG Z, et al. Travel patterns in china[J]. PloS One, 2011(6): e16364.
- [10] 刘明君, 郭继孚, 高利平, 等. 私人小汽车出行行为特征分析与建模[J]. 吉林大学学报(工学版), 2009, 39(2): 25-30.
LIU Ming-jun, GUO Ji-fu, GAO Li-ping, et al. Analysis and modeling of private car usage through travel behavior[J]. Journal of Jilin University (Engineering and Technology Edition), 2009, 39(2): 25-30.
- [11] KÖLBL R, HELBING D. Energy laws in human travel behavior[J]. New Journal of Physics, 2003(5): 1-12.
- [12] YAN X, HAN X, WANG B, et al. Diversity of individual mobility patterns and emergence of aggregated scaling laws[J]. Scientific Reports, 2013(3): 2678.
- [13] EAGLE N, PENTLAND A. Eigenbehaviors: identifying structure in routine[J]. Behavioral Ecology and Sociobiology, 2009, (63): 1057-1066.
- [14] SONG C, QU Z, BLUMM N, et al. Limits of predictability in human mobility[J]. Science, 2010(327): 1018-1021.
- [15] SONG C, KOREN T, WANG P, et al. Modelling the scaling properties of human mobility[J]. Nature Physics, 2010(10): 818-823.
- [16] WANG P, GONZÁLEZ M C, HIDALGO C A, et al. Understanding the spreading patterns of mobile phones viruses[J]. Science, 2009(324): 1071-1076.
- [17] BALCAN D, COLIZZA V, GONCALVES B, et al. Multiscale mobility networks and the spatial spreading of infectious diseases[J]. Proceedings of the National Academy of Sciences, 2009(106): 21484-21489.
- [18] CALABRESE F, DILORENZO G, RATTI C. Human mobility prediction based on individual and collective geographical preferences[C]//IEEE International Conference on Intelligent Transportation Systems. Piscataway, NJ, USA: IEEE, 2010: 312-317.
- [19] HAN X, HAO Q, WANG B, et al. Origin of the scaling law in human mobility: hierarchy of traffic systems[J]. Physical Review E, 2011(83): 036117.
- [20] TIAN H, MA X, WANG H, et al. A novel approach to estimate human space-time path based on mobile phone call records[C]//18th International Conference on Geoinformatics. Piscataway, NJ, USA: IEEE, 2010: 1-6.
- [21] HU J, CAO W, LUO L, et al. Dynamic modeling of urban population travel behavior based on data fusion of mobile phone positioning data and FCD[C]//18th International Conference on Geoinformatics. Piscataway, NJ, USA: IEEE, 2010: 1-5.
- [22] SIMINI F, GONZALEZ M C, MARITAN A, et al. A universal model for mobility and migration patterns[J]. Nature, 2012(484): 96-100.
- [23] 周涛, 韩筱璞, 闫小勇, 等. 人类行为时空特性的统计力学[J]. 电子科技大学学报, 2013, 42(4): 481-540.
ZHOU Tao, HAN Xiao-pu, YAN Xiao-yong, et al. Statistical mechanics on temporal and spatial activities of human[J]. Journal of University of Electronic Science and Technology of China, 2013, 42(4): 481-540.
- [24] JIANG B. Street hierarchies: a minority of streets account for a majority of traffic flow[J]. International Journal of Geographical Information Science, 2009(23): 1033-1048.
- [25] JIANG B, ZHAO S, YIN J. Self-organized natural roads for predicting traffic flow: a sensitivity study[J]. J Stat Mech, 2008(07): P07008.
- [26] CRUCITTI P, LATORA V, PORTA S. Centrality measures in spatial networks of urban streets[J]. Phys Rev E, 2006(73): 036125.
- [27] BARTHÉLEMY M, FLAMMINI A. Modeling urban streets patterns[J]. Phys Rev Lett, 2008(100): 138702.
- [28] ROSVALL M, TRUSINA A, MINNHAGEN P, et al. Networks and cities: an information perspective[J]. Phys Rev Lett, 2005(94): 028701.
- [29] KURANT M, THIRAN P. Layered complex networks[J]. Phys Rev Lett, 2006i(96): 138701.
- [30] WU Z, BRAUNSTEIN L A, HAVLIN S, et al. Transport in weighted networks: partition into superhighways and roads[J]. Phys Rev Lett, 2006(96): 148702.
- [31] KLEINBERG J M. Navigation in a small world[J]. Nature, 2000(406): 845.
- [32] Li G, REIS S D S, MOREIRA A A, et al. Towards design principles for optimal transport networks[J]. Phys Rev Lett, 2010(104): 018701.
- [33] CARMÍ S, CARTER S, SUN J, et al. Asymptotic behavior of the kleinberg model[J]. Phys Rev Lett, 2009(102): 238702.
- [34] CARTOZO C C, RIOS P D L. Extended navigability of small world networks: exact results and new insights[J].

- Phys Rev Lett, 2009(102): 238703.
- [35] ROBERSON M R, BEN-AVRAHAM D. Kleinberg navigation in fractal small-world networks[J]. Phys Rev E, 2006(74): 017101.
- [36] BOGUNA M, KRIOUKOV D. Navigating ultrasmall worlds in ultrashort time[J]. Phys Rev Lett, 2009(102): 058701.
- [37] KOSMIDIS K, HAVLIN S, BUNDE A. Structural properties of spatially embedded networks[J]. Europhys Lett, 2008(82): 48005.
- [38] MOUKARZEL C F, DE-MENEZES M A. Shortest paths on systems with power-law distributed long-range connections[J]. Phys Rev E, 2002(65): 056709.
- [39] GUIMERA R, DIAZ-GUILERA A, VEGA-REDONDO F, et al. Optimal network topologies for local search with congestion[J]. Phys Rev Lett, 2002(89): 248701.
- [40] GASTNER M T, NEWMAN M E J. Optimal design of spatial distribution networks[J]. Phys Rev E, 2006(74): 016117.
- [41] YOUNG H, GASTNER M T, JEONG H. Price of anarchy in transportation networks: efficiency and optimality control[J]. Phys Rev Lett, 2008(101): 128701.
- [42] GAO Z, SI B. A new algorithm for estimation of origin-destination demands from link traffic flows[J]. Journal of Transportation Systems Engineering and Information, 2002, 2(1): 366-375.
- [43] BRENNINGER-GÖTHE M, RNSTEN K J, LUNDGREN J T. Estimation of origin-destination matrices from traffic counts using multi objective programming formulations[J]. Transportation Research Part B, 1989(238): 257-269.
- [44] HAZELTON M L. Estimation of origin-destination matrices from link flows on uncongested networks[J]. Transportation Research Part B, 2000(34): 549-566.
- [45] HERRERA J C, AMIN S, BAYEN A, et al. Dynamic estimation of OD matrices for freeways and arterials[R]. USA California Berkeley: Institute for Transportation Studies, UC Berkeley, 2007.
- [46] 陈满堂. 关于车牌号OD调查方法的探讨[J]. 公路, 2004(9): 86-88.
- CHEN Man-tang. Discussion about method of OD (origin-destination) investigation of vehicle license plate number[J]. Highway, 2004(9): 86-88.
- [47] CALABRESE F, LORENZO D, LIU G, et al. Estimating origin-destination flows using mobile phone location[J]. IEEE Pervasive Computing, 2011(10): 36-44.
- [48] ZHANG Y, QIN X, DONG S, et al. Daily OD matrix estimation using cellular probe data[Z]. Transportation Research Part B Annual Meeting CD-ROM, 2010.
- [49] DONG S, QIN X, ZHANG Y, et al. Dynamic network flow modeling based on cell probe data[C]//2010 IEEE Intelligent Vehicles Symposium, Piscataway, NJ, USA: IEEE, 2010.
- [50] CHENG P, QIU Z, RAN B. Particle filter based traffic state estimation using cell phone network data[C]//Proceedings of the IEEE ITSC 2006. Piscataway, NJ, USA: IEEE, 2006.
- [51] GUO L, LUO D. Apply cellular wireless location technologies to traffic information gathering[C]//ICICTA 09 Proceedings of the 2009 Second International Conference on Intelligent Computation Technology and Automation 3. Piscataway, NJ, USA: IEEE, 2009.
- [52] LU B, HUANG M. Urban traffic information system based on the integration of gis and mobile phone message[C]// Proceedings of the 10th International Conference of Chinese Transportation Professionals. Reston, VA, USA: American Society of Civil Engineers, 2010.
- [53] WANG P, HUNTER T, BAYEN A M, et al. Understanding road usage patterns in urban areas[J]. Scientific Reports, 2012(2): 01001.

编辑 蒋晓