

引用格式:王陆一,吴健生,李卫锋.中小城市公共自行车出行模式与驱动机制研究[J].地球信息科学学报,2019,21(1):25-35. [Wang L Y, Wu J S, Li W F. Usage patterns and driving mechanisms of public bicycle systems in small and medium-sized cities based on space-time data mining[J]. Journal of Geo-information Science, 2019,21(1):25-35. ] DOI:10.12082/dqxxkx.2019.180199

# 中小城市公共自行车出行模式与驱动机制研究

王陆一<sup>1,5</sup>, 吴健生<sup>1,2\*</sup>, 李卫锋<sup>3,4</sup>

1. 北京大学城市规划与设计学院,城市人居环境科学与技术重点实验室,深圳518055;2. 北京大学城市与环境学院,地表过程与模拟教育部重点实验室,北京100871;3. 香港大学建筑学院城市规划及设计系,香港999077;4. 香港大学深圳研究院,深圳518057;5. 滴滴出行地图事业部,北京100094

## Usage Patterns and Driving Mechanisms of Public Bicycle Systems in Small and Medium-Sized Cities based on Space-Time Data Mining

WANG Luyi<sup>1,5</sup>, WU Jiansheng<sup>1,2\*</sup>, LI Weifeng<sup>3,4</sup>

1. Key Laboratory for Urban Habitant Environment Science and Technology, School of Urban Planning and Design, Peking University, Shenzhen, Guangdong 518055, China; 2. Laboratory for Earth Surface Processes of Ministry of Education, College of Urban and Environmental Sciences, Peking University, Beijing 100871, China; 3. Department of Urban Planning and Design, Faculty of Architecture, The University of Hong Kong, Hongkong 999077, China; 4. Shenzhen Institute of Research and Innovation, The University of Hong Kong, Shenzhen 518057, China; 5. Map Division, Didi Chuxing, Beijing 100094, China

**Abstract:** Given the importance of environment-friendly cities, the development of public bicycle systems (PBSs) has become more popular in recent years around the world. The purpose of this study was to explore the usage patterns of PBSs in small and medium-sized cities in Guangdong province, China, and to infer the driving mechanisms of system attributes and the built environment. The research applied time series analysis of global activity patterns, hierarchical clustering algorithm using Dynamic Time Warping distances as features and spatial data visualization on station-based data, and then compared different systems by employing a random forest algorithm to evaluate the influencing factors. The study objective was to better understand the relationship between public bicycle usage activity and underlying built environment characteristics. In Huicheng District of Huizhou City and Shaoguan City, the public bicycle usage patterns are regular, and bicycle stations are grouped into several clusters based on usage patterns of "morning destination, night origin" "morning origin, night destination" and "steady throughout the day". The PBS in Huicheng District plays various roles by helping users commute to and from jobs and schools, and to make short distance trips. The PBS also is a complementary tool for bus transit facilities. The PBS in Shaoguan City mostly serves as a mode for commuting. The PBS is inefficiently used in Huiyang District of Huizhou City owing to the poor road conditions. This research provides a study framework that can be reproduced in other areas, and offers a way of optimizing PBSs, thereby assisting urban transportation planning and urban land use allocation.

**Key words:** public bicycle; small and medium-sized cities; spatio-temporal data mining; random forest; driving mechanisms

收稿日期:2018-04-23;修回日期:2018-09-23.

基金项目:国家自然科学基金面上项目(41471370). [ **Foundation item:** National Natural Science Foundation of China, No.41471370. ]

作者简介:王陆一(1993-),男,硕士生,研究方向为城市与区域规划,时空数据分析与挖掘. E-mail: wangluyi@pku.edu.cn

\*通讯作者:吴健生(1965-),男,博士,教授,博士生导师,研究方向为遥感与地理信息系统、数字城市与城市安全、景观生态学与土地利用. E-mail: wujs@pkusz.edu.cn

\*Corresponding author: WU Jiansheng, E-mail: wujs@pkusz.edu.cn

**摘要:**公共自行车在促进交通可持续发展、方便居民出行等方面意义重大。本文以广东省惠州市惠城区和惠阳区、韶关市区公共自行车为研究对象,利用时间序列数据分析、层次聚类、地理可视化等方法探究出行模式,利用随机森林算法分析出行行为的影响因素及其重要性程度。研究发现,惠州市惠城区和韶关市的公共自行车出行行为有规律、成规模,站点呈现生活居住、工作就业等类型;惠州市惠城区自行车使用目的多元,包括通勤、短途办事、公交接驳等,韶关市中,通勤是主要出行目的。惠州市惠阳区受到骑行道路限制,公共自行车使用率低。本研究可以提高公共自行车系统运营效率、引导慢行交通政策制定、评估城市用地布局提供参考和建议,也能为其他区域公共自行车的研究提供借鉴作用。

**关键词:**公共自行车;中小城市;时空数据挖掘;随机森林;驱动机制

## 1 引言

中国近年来城市快速扩张,交通机动化的趋势明显。为了减少机动车带来的污染排放、交通拥堵,不仅需要限制小汽车出行,同时要改善公共出行环境,鼓励低碳交通方式。其中,公共自行车作为共享经济模式在交通出行发展的产物,具有重要意义和发展空间,近年来在全世界得到重视并迅速发展。至2016年底,全球大约有1200个公共自行车系统<sup>[1]</sup>。自2008年开始,中国先后在杭州、上海、北京、广州、深圳等城市实施了公共自行车系统<sup>[2]</sup>。

随着信息通信和大数据技术的发展,许多学者利用公共自行车的历史数据,对站点时空特征、人群移动模式进行探索。在国外,Froehlich等<sup>[3]</sup>挖掘了巴塞罗那公共自行车出行的时空行为模式,利用系统树图聚类Ward's算法分析出行与地理位置、周围环境、时间的关系。纽约市的公共自行车可行性研究结合了自行车出行状况与人口密度、工作岗位密度、服务设施、文化与休闲设施和商业中心等建成环境,推断各地区的需求水平<sup>[4]</sup>。O'Brien等<sup>[5]</sup>收集了全球38个公共自行车系统的使用数据,根据系统规模、日常并发量、站点车桩的占有和空闲情况、站点间距离、空间紧凑度、自行车供给数量等指标,对不同地区的系统进行聚类,为解决站点不平衡问题提供分析基础。在国内,刘冰等<sup>[6]</sup>利用杭州市公共自行车IC卡数据,通过数据挖掘测度了站点的使用特点,揭示了杭州市“圈层式”的结构,以及不同区域的站点在使用量、潮汐性、周转率等指标的差异。郭素萍等<sup>[7]</sup>对南京公共自行车数据进行了时间统计和空间可视化分析,对不同目的(如通勤、生活等)下出行的时空特征进行了总结和对比,并建议站点规划布局应充分考虑用地功能和空间形

态。邓力凡等<sup>[8]</sup>利用北京市的摩拜单车数据分析了骑行时空特征,发现工作日存在明显的早晚高峰。在公共交通和商业、餐饮等空间要素的影响下,骑行存在不同的空间特征,其还提出了共享单车停车设施优化建议。朱玮等<sup>[9-10]</sup>以上海市闵行区公共自行车为研究对象,采用调查问卷与刷卡数据分析相结合的方式,对不同用地类型站点的使用特征进行了描述性分析。在以往的定量城市研究中,掌握公共自行车用户的时空间行为模式和特征有助于更好地理解系统内在机理和城市运行特点,通过出行大数据挖掘,可以捕捉到用户使用公共自行车的时空规律和更细致的出行行为。

为了帮助规划部门、公共自行车运营部门评估城市潜在的运行机理,促进系统运营,需要对自行车出行模式的驱动机制进行研究。Zhang等<sup>[11]</sup>建立多元线性回归和空间回归模型,探讨中山市建成环境对出行需求量和需求供给比的影响,其中的建成环境变量主要是指站点自身特征和可达性、自行车出行基础设施、公共交通以及土地利用特征。Wang等<sup>[12]</sup>建立了Log-linear和负二项回归模型,评估了居民的经济社会属性、土地利用特征(包括水体、商业区、学校)、其他交通站点等因素对于日均出行量影响的边际效应;该研究成果可以为优化公共自行车运营、新站点选址提供参考。Faghih-Imani等<sup>[13]</sup>研究了蒙特利尔土地利用和城市形态对自行车出行量的影响,建立线性回归和线性混合模型,利用log-likelihood ratio进行模型检验;主要结论表明临近的商务区、饭店、大学对出行量有显著影响,增加站点数量比增加某站点车位数对车流量的影响更加显著。已有研究表明,影响公共自行车出行的主要因素可能包括:系统本身属性、交通基础设施和公共交通换乘、土地利用情况等。

现有对于公共自行车的研究不足主要体现在

3个方面:①研究区大多集中在欧洲大中城市、北美等地,中国区域研究主要关注杭州、北京等一线城市,缺乏对于欠发达地区的关注。中小城市发展空间紧凑、土地开发强度低,居民出行特征与发达城市会有所不同;②国内对于公共自行车出行模式的研究有待大规模时空数据的进一步分析和佐证;③国内鲜有研究定量分析建成环境对于公共自行车使用行为的驱动机制。本文将对以上不足开展以下研究:选取广东省惠州市惠城区和惠阳区、韶关市区3个公共自行车系统作为研究对象,基于历史使用数据,挖掘全局使用情况,对站点使用模式进行识别并聚类,探索自行车出行的时空特征;利用随机森林算法检测站点使用模式的影响因素及其重要性程度;在研究结果的基础上,结合调研和规划文件,为城市规划和公共自行车部门提供政策依据,同时对城市用地布局和其他交通系统的规划与评估提供参考。

## 2 研究区概况和数据来源

惠州市位于广东省东南部,属于珠江三角洲城市群。惠州市管辖惠城区、惠阳区等区县。惠城区位于惠州市中部,是市政府所在地,功能分区包括住宅区、商业区、行政区和文化区,2015年底常住人口120.01万人。惠阳区位于惠州市南部,是惠州区域副中心城市,西靠深圳、东莞,南临大亚湾,2015

年常住人口59.56万人。惠州市中,惠城区和惠阳区相距大约30 km,基本不存在跨区骑行的情况,故分开讨论。韶关市被定位为“粤北区域中心城市”,是广东的老工业基地和重要生态屏障,2016年末韶关市区常住接近100万人。3个区域(图1)均建设了公共自行车系统且已运营时间较长,进行中小城市研究具有一定的典型性。系统的基本属性如表1所示。不同的区域具有不同的经济水平和功能区定位、人口分布、交通建设水平、土地利用和建成环境,各方面的差异势必造成自行车出行模式的差异。通过不同区域的比较研究,尤其针对缺乏关注度的中小城市居民出行特征的研究,有利于全面细致地探索居民出行和城市结构的交互关系。

本文采用的公共自行车站点数据来自惠民运营系统管理有限公司网站<sup>[14]</sup>,编写Python脚本实时获取各个站点的可借车辆和剩余车位数据。选取的数据时间跨度为2017-05-01至2017-09-01,时间分辨率为10 min。根据以往研究<sup>[5]</sup>,该频率获取的数据可以较精确地反映一天中出行行为和自行车可用性的变化,描绘出通勤高峰时段和其他的特征。每一条记录值包含的属性字段主要有:站点ID、名称、具体地址、经纬度、总车位数、可借车数、可还车数、日期时间和星期。

惠州市惠民自行车公共服务系统项目已作为惠州市2012年十大民生实事之一进行重点推广;另外,摩拜、ofo等无桩共享单车在研究数据时段内未



图1 研究区域

Fig. 1 Study area



表1 研究区域公共自行车系统概况

Tab. 1 Attributes of the public bicycle systems

研究区域	惠州市惠城区	惠州市惠阳区	韶关市市区
系统开始运营时间	2012年4月	2013年12月	2013年1月
自行车站点数量	97	37	31
总车位数/个	1762	714	600
自行车数/辆	760	260	270
车位/站点比	18.16	19.30	19.35
车位/自行车数比	2.32	2.75	2.22

进入市场,只在个别园区试运营。在韶关市,主要道路的自行车高峰小时交通流量不大,部分主要道路的自行车交通量在100辆/h以下<sup>[15]</sup>。市区建成范围有限,31个自行车站点覆盖了绝大部分交通要道。综上,惠民运营系统能够一定程度反映研究区公共自行车的整体情况。

其他使用到的数据包括:人口分布数据(2013年)为Oak Ridge国家实验室提供的LandScanTM<sup>[16]</sup>数据;道路数据来自众包地图平台OpenStreetMap<sup>[17]</sup>;公交站点数据和地理兴趣点(Points of Interest, POI),POI数据(电子地图上的地标、景点、处所等)来自北京城市实验室(Beijing CityLab)<sup>[18-19]</sup>。

### 3 研究方法

#### 3.1 时间序列数据分析和层次聚类

本节利用站点可借车数和空余车位的时间序列数据,探究使用行为的时空模式。首先,分析系统总体使用强度,即所有站点可借车数总和随时间和日期的变化,找出自行车的使用高峰、低谷期以及工作日和周末的差异。此外,对站点的时间序列数据进行层次聚类,解析站点使用日变化的规律,识别使用类别不同的站点,并分析不同类型的站点在城市空间上的分布。层次聚类基于不同组数据间的相似度来创建一棵有层次的嵌套聚类树,优点体现在:距离和规则的相似度容易定义,不需要预先设定聚类数目,可以发现类与类之间的层次关系。

由于系统中站点的总车位数并不是统一的,引入标准化可借车数(Normalization Available Bicycles, NAB)的概念,即可借车数的数量占总车位的比值<sup>[3,20]</sup>。NAB的范围是0(代表车位全空,无法借车)到1(代表车位占满,无法还车)。考虑到某些站点的NAB序列的日变化趋势相近,但是真实值不接近,为了提高聚类的效果,对可借车数进行二次标准化,引入NNAB(Normalized NAB)的概念。

对于某一天的数据,NNAB计算如下:

$$NNAB = NAB - \text{mean}_{\text{day}}(NAB) \quad (1)$$

式中: $\text{mean}_{\text{day}}(NAB)$ 代表某一天中NAB的平均值。通过二次标准化的方法,时间序列值反映的是趋势的变化,便于提高对站点NAB日变化序列的聚类效果。

对于时间序列数据相似度的度量,采取动态时间规整(Dynamic Time Warping, DTW)距离<sup>[21]</sup>。给定2个时间序列集Q和C,DTW距离的计算需要构建一个 $n \times m$ 的邻接矩阵,寻找最短路径。矩阵中第 $(i, j)$ 个元素为 $(q_i - c_j)^2$ ,代表匹配时间序列Q的点 $q_i$ 到时间序列C的点 $c_j$ 的距离。在本文中,由于数据每10 min收集一次,一天一个站点的数据量为 $6 \times 24 = 144$ ,所以 $n = m = 144$ 。2组时间序列的匹配可以表示为矩阵中一条弯曲的路径: $W = w_1, w_2, \dots, w_k, \dots, w_K$ ,且该路径是连续的、单调的,始于矩阵的左下角、终于矩阵的右上角。2组时间序列的最佳匹配需要该路径经过的点的匹配距离和最小,对应的最小距离和即是DTW距离<sup>[22]</sup>:

$$DTW(Q, C) = \arg \min_{W = w_1, w_2, \dots, w_k, \dots, w_K} \sqrt{\sum_{k=1, w_k=(i,j)}^K (q_i - c_j)^2} \quad (2)$$

欧氏距离要求2组时间序列长度相同,且严格计算每个时刻的对应距离,因此对噪声较敏感。相较于欧氏距离,DTW距离在度量时间序列数据相似度上,优势体现在对微小的不协调数据有较低的敏感度<sup>[22]</sup>。将站点层次聚类后,对属于各个类别的站点进行人工判别,观察其时间序列变化趋势,避免时序被过度拉伸、压缩。数据预处理和层次聚类通过R语言实现。

#### 3.2 影响因素特征生成和随机森林

根据已有的研究<sup>[11-12,23-24]</sup>可知,系统本身属性和建成环境对自行车出行模式有一定影响,其影响因素归结可为3大类:①自行车站点属性,包含自行车站点总车位数、距离最近站点的距离、缓冲区内自行车站点数;②可达性和公共交通接驳,包含缓冲区内的人口、非机动车道长度、机动车道长度、公交站数量;③城市土地利用和功能区情况,包含缓冲区内POI种类、不同类别(就业、居住、商业、公共服务、学校、休闲娱乐)POI数量。本文选取300 m欧氏距离作为缓冲区,计算站点建成环境的特征变量,因为该值在以往的研究中被证明可以较好地反



映自行车站点服务范围的效果<sup>[11,25]</sup>,且步行长度大约为10 min,取值合理。

随机森林是机器学习中由决策树理论发展而来的组合学习方法。决策树能够表示给定特征条件下类的条件概率分布(定义在特征空间的一个划分上),各叶结点(特征空间的单元)上的条件概率往往偏向某一类,决策树分类时将该结点的实例分到条件概率大的类<sup>[26]</sup>。Breiman<sup>[27]</sup>于2001年创造性地提出了随机森林:双重随机构建模型方法,即对数据集随机有放回抽样与节点处随机特征选取。由随机化产生若干的决策树结果投票得到最优的分类结果。随机森林能够在训练过程中检测到特征之间的相互影响以及特征的重要性程度。

利用随机森林识别工作日期间站点使用行为的影响因素,对建模结果和变量因子的重要性进行评价分析。本研究中,随机森林模型本质为分类问题,因变量为站点工作日使用情况的聚类结果,特征变量为前文中系统本身属性和建成环境。利用袋外数据估计错误率评估模型结果。在构建某决策树时,对训练集使用随机且有放回的抽取,没有参与树的生成的训练实例,被称为该树的“袋外数据”样本。利用袋外数据估计错误率可以在模型生成的过程中对误差建立无偏估计,而无需使用独立的测试集。

## 4 公共自行车出行模式研究

### 4.1 全局出行行为分析

分析工作日和周末,系统所有自行车使用强度的日变化,3个区域的结果如图2-图4所示。可以看出,对于惠州市惠城区和韶关市,系统的使用情况呈现出了比较明显的早晚高峰。而在惠州市惠阳区,整体的使用强度低,且没有特殊的规则模式。

惠州市惠城区的公共自行车系统颇受市民欢迎,呈现出较规律、成规模的出行行为。使用高峰从7时开始并持续到10时,在8-9时之间达到顶峰;下午高峰出现在17-19时。高峰期时被使用的车辆数量约为80辆,占到总车数约11%。而在周末,虽然也呈现出早晚高峰的趋势,但是高峰期使用量明显低于工作日,推断出该系统对于日常工作通勤的服务作用较为明显。惠州市是全国文明城市,为了倡导低碳、绿色的出行概念,缓解城市交通拥堵,注重发展慢行交通,2012年开始启动“惠民自行车”建

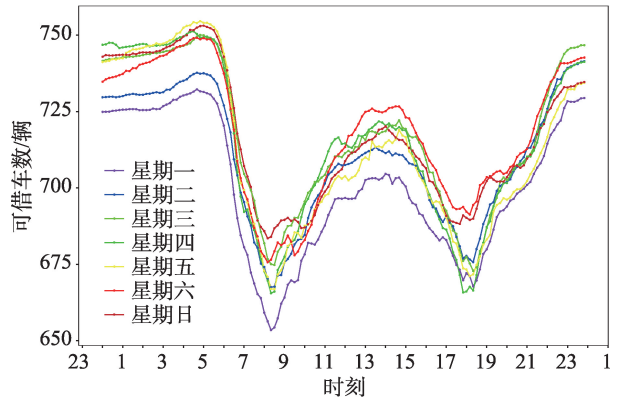


图2 惠州市惠城区站点可借车数平均值日变化

Fig. 2 Average of the total bicycles available in the stations in Huicheng, Huizhou

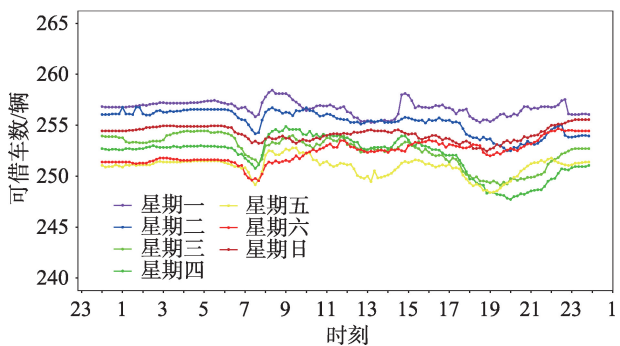


图3 惠州市惠阳区站点可借车数平均值日变化

Fig. 3 Average of the total bicycles available in the stations in Huiyang, Huizhou

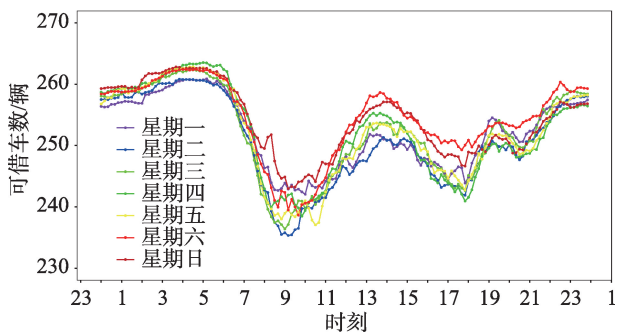


图4 韶关市站点可借车数平均值日变化

Fig. 4 Average of the total bicycles available in the stations in Shaoguan

设工程。惠城区地貌多为平原区,比较适合自行车出行。惠城区存在公交车、出租车等候时间较长的现象,公共自行车可以对其他出行方式起到一定的补充或替代作用。

惠州市惠阳区公共自行车使用量小,没有得到有效的利用。通过调研发现,该区主要的问题是骑

行道路欠缺,且有众多的摩托车骑行者,自行车与机动车距离过近,还会发生抢道混行的现象,带来了一定程度的骑行安全问题。对于公共自行车的建设推广,除了站点和车辆,还需要有通畅的路网配套,优化市区路网空间,才能更加充分地发挥自行车的出行作用。另外,市民对于公共自行车的了解程度不高,知道系统运营、使用方法的市民不多,宣传力度有待加强。

韶关市中心区的公共自行车租借行为存在较明显的“潮汐性”。系统的使用高峰出现在8-10时和17-18时,早高峰使用极值出现在9时左右。高峰期自行车使用量占到总车数的9%。周末相较于工作日来说,自行车的使用情况变化较为平缓,使用强度整体较工作日低,尤其是2个高峰时段。与惠州市惠城区整体使用情况不同的是,在20-21时,韶关市公共自行车系统出现了一个局部使用高峰,揭示出该地区部分市民夜间出行、户外休闲活动的特点。韶关市区生态环境优美,市区范围内到达目的地耗时大多不超过1 h,一定程度上有利于自行车出行。

#### 4.2 出行模式聚类分析

本节对自行车站点在工作日的使用情况进行时间序列数据聚类,发现主要的出行模式规律,并探析不同类别站点在空间上的分布特点。

惠州市惠城区的公共自行车站点可大致分为5种类别模式(图5-图6),其中类别1、4、5对应“早间出发地、晚间到达地”型,类别3对应“早间到达

地、晚间出发地”型,类别2对应“平稳”型。从空间分布来看,“早间到达地、晚间出发地”型的站点集中分布在穿城而过的东江北部惠州市政府附近以及东江南部的世贸中心、港惠新天地等商圈,日间到达热点区域呈现双中心的格局。“早间出发地、晚间到达地”型站点也呈现一定程度上的空间集聚,如沿二环路外围、东江支流西枝江两岸等。

惠州市惠阳区的公共自行车系统使用强度微弱,绝大多数站点呈现“平稳”型,没有明显的多种模式特征(图7-图8,图8见第32页)。

韶关市公共自行车站点在工作日主要呈现3种模式,分别为“早间出发地、晚间到达地”型、“早间到达地、晚间出发地”型、“平稳”型(图9-图10,图9见第32页,图10见第33页)。从站点类别的空间分布上,可以看出明显的单中心模式:日间自行车使用者向浈江和武江中部的城市中心区域集聚,附近分布有事业单位、商城等,夜间向“Y”字型河流两侧分散,目的地多为居住功能区,如鸿洲花园站、碧岛豪庭站、时代花园站等。未得到充分利用的站点主要分布在省级干线公路和城市建成区边缘地带。

综上所述,通过对于站点使用模式的识别,能够映射出城市活动中心的时空分布特点和出行运行机制。基于传感技术捕捉到公共自行车出行实时动态,利用“社会感知”和“城市计算”<sup>[28-29]</sup>挖掘市民的移动行为模式,有助于理解城市交通需求的产生机理,识别出城市核心功能区域和外围组团的人群活动强度,为完善城市综合交通体系、土地复合开发利用等提供借鉴。

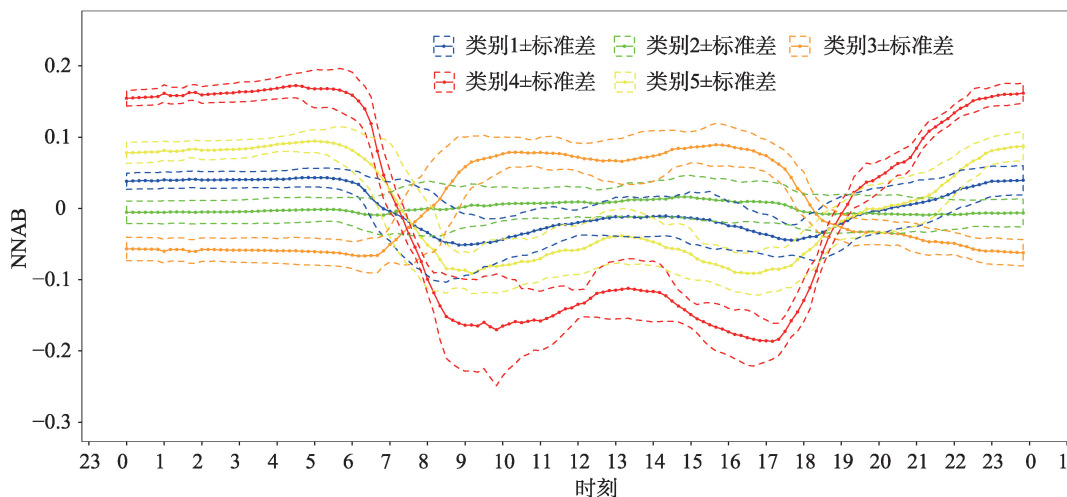


图5 惠州市惠城区站点工作日聚类结果NNAB日变化趋势

Fig. 5 Average NNAB of clusters on weekdays in Huicheng, Huizhou

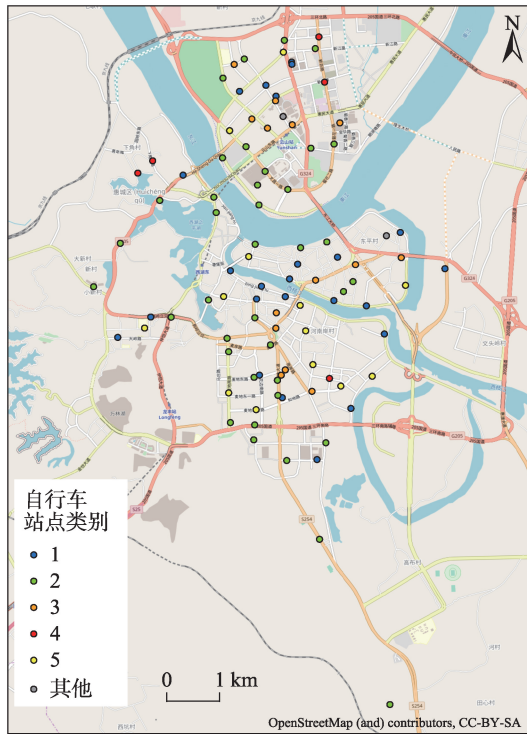


图6 惠州市惠城区站点工作日 NNAB 聚类结果空间分布

Fig. 6 Map visualization of NNAB clustering results on weekdays in Huicheng, Huizhou

## 5 出行模式驱动机制研究

本节对惠州市惠城区和韶关市中心区的公共自行车站点模式和影响因素的关系分别进行随机森林分析,得到特定出行行为的驱动机制。由于惠州市惠阳区公共自行车使用强度微弱,且没有识别出有规律的模式,故不在本文讨论范围。利用R语

言的"randomForest"工具包进行随机森林算法的实现。对于惠州市惠城区和韶关市,训练样本剔除了使用行为模式规则不明显的站点,剩余样本数分别为56和25。选取合适的树节点预选变量个数(mtry)和模型包含的决策树的数目(ntree),控制模型误差,并剔除对模型重要性程度小的特征变量,得到的因子重要性评价结果如图11-图12所示(图11见第33页,图12见第34页)。2个模型的袋外数据估计错误率分别为17.86%和24.00%,效果较好。结果中,Mean Decrease Accuracy(平均准确率降低度)和Mean Decrease Gini(平均基尼指数降低度)描述的是特征变量重要性度量<sup>[30]</sup>。Mean Decrease Accuracy描述的是当把一个变量变成随机数时,随机森林预测准确度的降低程度,该值越大表示该变量的重要性越大。Mean Decrease Gini通过基尼指数计算每个变量对决策树上每个节点的观测值的异质性的影响,该值越大表示该变量的重要性越大。

在惠州市惠城区,多种类别POI数量重要性程度较高,推断出该地市民使用自行车目的多元,包括工作日通勤和通学、短途办事、公交接驳、休闲娱乐等等,其中居住和公共服务类POI起到最重要的驱动作用;惠城中心两江贯城,具有一定的生态景观资源和休闲活动空间,也为城市骑行提供了活力。《惠州市低碳生态规划(2014-2030)》中提出,建设舒适便民的慢行交通系统,结合大型社区、交通枢纽站、公园广场、重点景区等场所设置公共自行车站点,加大公共自行车投放力度,这符合基于历史数据挖掘出的自行车出行的需求。

在东江南北两个日间活动中心的自行车站点

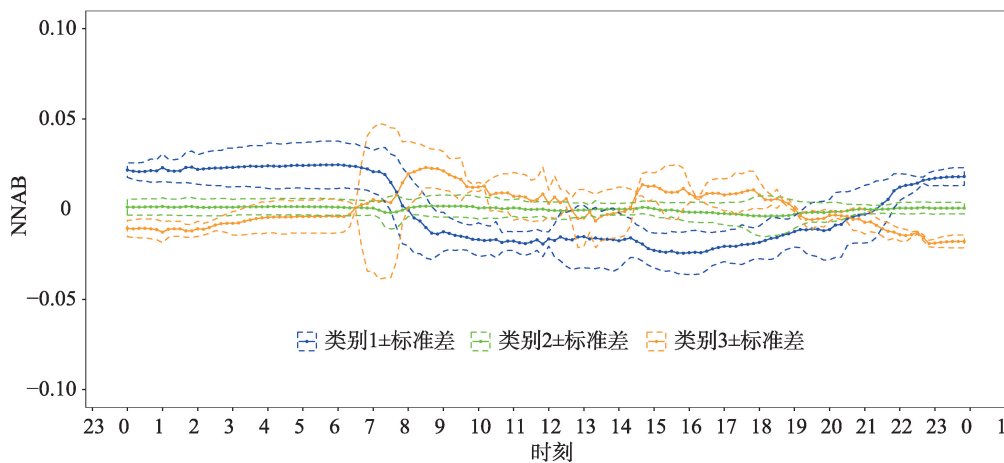


图7 惠州市惠阳区站点工作日聚类结果 NNAB 日变化趋势

Fig. 7 Average NNAB of clusters on weekdays in Huiyang, Huizhou





图8 惠州市惠阳区站点工作日 NNAB 聚类结果空间分布

Fig. 8 Map visualization of NNAB clustering results on weekdays in Huiyang, Huizhou

附近,公交站点较多、非机动车道长度较小,而居住型站点附近公交站数量相对较少,也侧面反映出公交发展水平、道路条件等的不平衡问题。目前惠州市区某些道路没有专门的非机动车道,公共骑行的条件存在一定不足。未来慢行系统的发展,应结合区域机动车和公交发展,实施区域差别化的政策指引:如接驳类站点布置在交通干道临近处;生活居住类站点更多考虑渗透到支路,并调查生活

居住区对于公交站点的诉求,统筹发展公共交通体系。另外,在核心建城区的东北和东南部、东江绿道沿线等地,自行车站点布局欠缺,为了推进“慢行+公交”为出行方式主体,还需要增加自行车站点的数量和密度。

对于韶关市,根据已有研究<sup>[15]</sup>,居民使用自行车的出行目的包括通勤通学、健身或休闲旅游、公交换乘等。根据随机森林模型,其中工作日通勤是出行的主要目的,而公交换乘所占的比例很少,这也与现状调查的结论相一致。韶关市上班高峰期存在较堵车的现象,而城市整体范围不大,公共自行车一定程度上可以缓解机动车通勤的压力,有较大的发展空间。

自行车站点服务区覆盖人口数显著影响了使用强度,居住型站点服务的人口多于平稳型的站点,后者大多位于省道等交通干线和城市建成区边缘,距离最近的自行车站点也较远。另外,韶关市存在自行车道及停车设施不足的问题,调研发现部分道路机动车和非机动车混行,有机动车违章占道,降低了自行车通行能力,带来了安全隐患。为了促进公共自行车的使用,要完善自行车出行的设施配置,增强居民的自行车出行意愿。

## 6 结论与讨论

本研究利用时空数据挖掘技术剖析了公共自行车出行模式。惠州惠城区和韶关市区的公共自行车系统,都呈现规律明显的早晚高峰现象。与北京、杭州、上海<sup>[6,8-9]</sup>等城市相比,二者的早高峰稍延后,晚

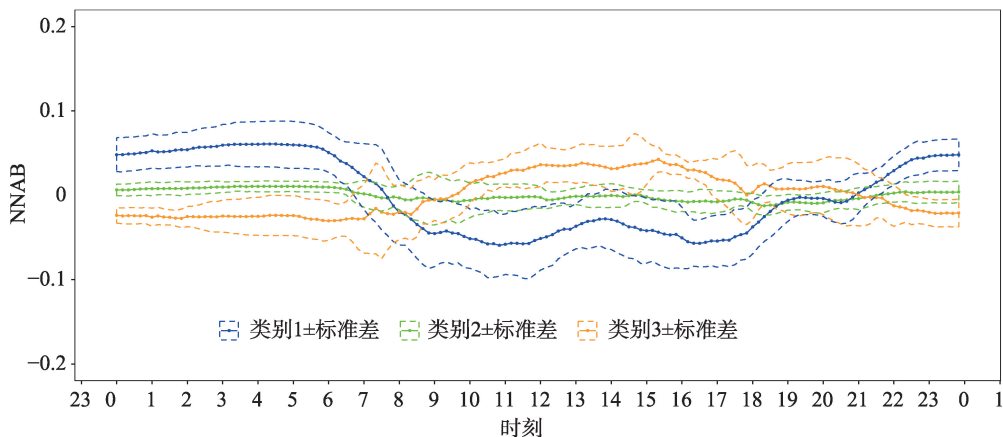


图9 韶关市站点工作日聚类结果 NNAB 日变化趋势

Fig. 9 Average NNAB of clusters on weekdays in Shaoguan

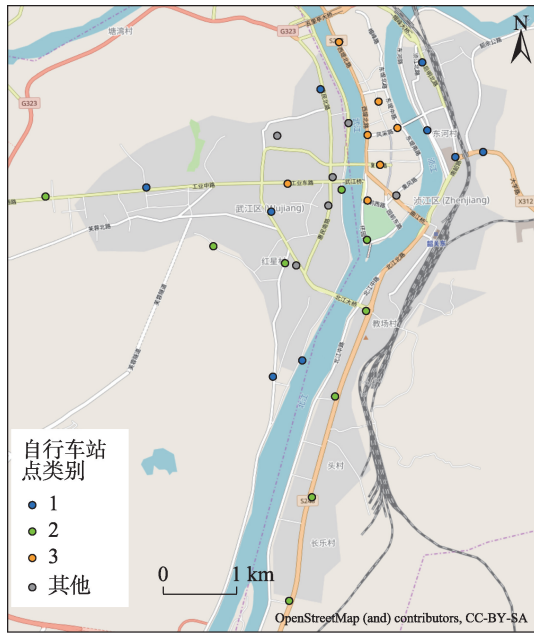


图10 韶关市站点工作日NNAB聚类结果空间分布  
Fig. 10 Map visualization of NNAB clustering results on weekdays in Shaoguan

高峰持续时间较短;且在北京中午12时出现骑行需求的次高峰,在研究区并未出现。韶关市公共自行车在20-21时呈现使用次高峰,推测原因为:该地生态环境、户外休闲场所吸引部分市民晚间骑行。

通过对站点的工作日时间序列数据进行聚类,识别出主要的出行模式,深入探讨了自行车出行的运行机制和城市空间的特点。惠州惠城区和韶关市的站点主要呈现出生活居住型、工作就业型、平稳型。相较而言,在北京、上海等地,骑行活动具有更多元的变化,如单向型、若干小波峰型等<sup>[8-9]</sup>。从站点的模式可以识别出,惠州惠城区日间活动区域

呈现双中心格局;韶关市区呈现单就业中心格局。

自行车出行模式与站点属性和建成环境息息相关。利用随机森林提取出影响站点模式类别的特征变量和重要性程度。惠州惠城区市民使用公共自行车出行目的更为多元,包括通勤、短途办事、休闲娱乐、公交接驳等。在韶关市,影响站点使用模式的主要因素为通勤和通学,而不包含公交换乘。类似的,在某些中等城市如湖南株洲市,市民使用公共自行车作为整个行程的工具,而非公交换乘<sup>[31]</sup>;而在诸如北京、杭州的大城市,公共自行车对其他公共交通系统同时起到补充和替代的作用<sup>[32-33]</sup>。在建设公共自行车系统前需要了解当地出行需求,针对不同地区差异化定位公共自行车的角色。

推进新型城镇化的进程中不能只关注大城市,我国中小城市数量众多,存在着交通低效、用地粗放等问题,未来要更多地针对中小城市居民出行特征进行探索,为中小城市的城市规划和交通精细化管理提供参考<sup>[34]</sup>。面对不同城市形态和市民的交通需求,需要针对性提出规划建议和政策引导:如惠州市定位为绿色化现代山水城市,惠城区自行车使用目的多元,扩大站点规模时应全面考虑社区、就业聚集地、公园广场、公交枢纽等场所。惠阳区更多地需要考虑市民的出行习惯,加强宣传引导的力度,并且优化路网空间。对于韶关市,主要问题体现在部分自行车道路设施不足、局部区域站点利用率低等。由于建成区范围有限,市民对于最后一公里交通接驳的需求并不大。基于历史数据,可以识别出低频使用站点,有助于改善站点的辐射效应,促进系统的均衡运营。

目前城市规划处在由增量规划向存量规划转

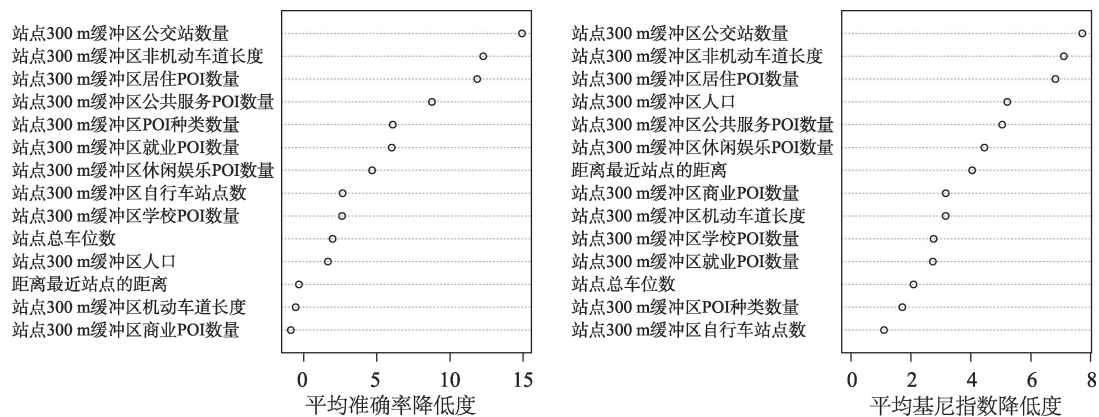


图11 惠州市惠城区站点类别影响因素重要程度

Fig. 11 Importance evaluation of factors on station cluster in Huicheng, Huizhou

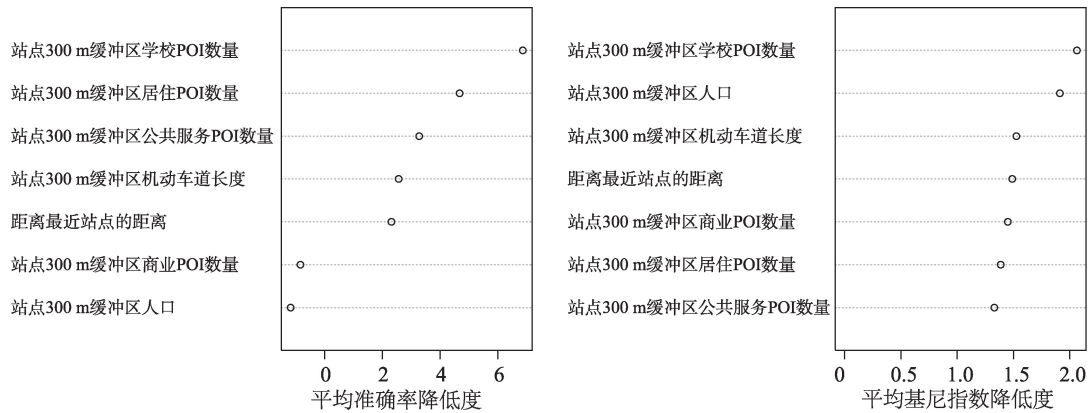


图12 韶关市站点类别影响因素重要程度

Fig. 12 Importance evaluation of factors on station cluster in Shaoguan

变的阶段,互联网大数据和时空数据挖掘算法对存量空间优化起到重要的作用。近年出现了无桩共享单车,此类单车造成的严重问题是无序的市场扩张带来车辆浪费、挤占公共空间。本研究分析框架也可以应用到无桩单车的投放量评估、停车设施空间布局,促进企业转型和城市管理。

#### 参考文献(References):

- [1] Duran A C, Anaya-boig E, Shake J D, et al. Bicycle-sharing system socio-spatial inequalities in Brazil[J]. *Journal of Transport & Health*, 2018,8:262-270.
- [2] 朱玮,庞宇琦,王德.自行车出行行为和决策研究进展[J]. *国际城市规划*,2013,28(1):54-59. [Zhu W, Pang Y Q, Wang D. Progress of research on bicycle travel behavior and decisions[J]. *Urban Planning International*, 2013,28(1):54-59.]
- [3] Froehlich J, Neumann J, Oliver N. Sensing and predicting the pulse of the city through shared bicycling[C]. Pasadena: Proceedings of the International Joint Conference on Artificial Intelligence, 2009:1420-1426.
- [4] Burden A M, Barth R. Bike-share opportunities in New York City[R]. New York: Department of City Planning, 2009:11-19.
- [5] O'brien O, Cheshire J, Batty M. Mining bicycle sharing data for generating insights into sustainable transport systems[J]. *Journal of Transport Geography*, 2014,34(219):262-273.
- [6] 刘冰,曹娟娟,周于杰,等.城市公共自行车使用活动的时空特征研究——以杭州为例[J].*城市规划学刊*,2016(3):77-84. [Liu B, Cao J J, Zhou Y J, et al. A study on the temporal-spatial features of bicycle-sharing activities: A case of Hangzhou[J]. *Urban Planning Forum*, 2016(3):77-84.]
- [7] 郭素萍,甄峰,尹秋怡.城市公共自行车出行时空特征分析[J].*规划师*,2017,33(6):112-118. [Guo S P, Zhen F, Yin Q Y. Analysis on space-temporal characters of public bicycles[J]. *Planners*, 2017,33(6):112-118.]
- [8] 邓力凡,谢永红,黄鼎曦.基于骑行时空数据的共享单车设施规划研究[J].*规划师*,2017,33(10):82-88. [Deng L F, Xie Y H, Huang D X. Bicycle-sharing facility planning base on riding spatio-temporal data[J]. *Planners*, 2017,33(10):82-88.]
- [9] 朱玮,庞宇琦,王德,等.上海市闵行区公共自行车出行特征研究[J].*上海城市规划*,2012(6):102-107. [Zhu W, Pang Y Q, Wang D, et al. Research on the travel characteristics of public bicycle in Minhang District, Shanghai [J]. *Shanghai Urban Planning Review*, 2012(6):102-107.]
- [10] 朱玮,庞宇琦,王德,等.公共自行车系统影响下居民出行的变化与机制研究——以上海闵行区为例[J].*城市规划学刊*,2012(5):76-81. [Zhu W, Pang Y Q, Wang D, et al. Travel behavior change after the introduction of public bicycle systems: A case study of Minhang District, Shanghai[J]. *Urban Planning Forum*, 2012(5):76-81.]
- [11] Zhang Y, Thomas T, Brussel M, et al. Exploring the impact of built environment factors on the use of public bikes at bike stations: Case study in Zhongshan, China[J]. *Journal of Transport Geography*, 2017,58(SupC):59-70.
- [12] Wang X, Lindsey G, Schoner J E, et al. Modeling bike share station activity: Effects of nearby businesses and jobs on trips to and from stations[J]. *Journal of Urban Planning and Development*, 2015,142(1):4015001.
- [13] Faghieh-Imani A, Eluru N, El-Geneidy A M, et al. How land-use and urban form impact bicycle flows: Evidence from the bicycle-sharing system (BIXI) in Montreal[J]. *Journal of Transport Geography*, 2014,41:306-314.
- [14] 广东惠民运营股份有限公司.2017.http://hz.2773456.com/, http://sg.2773456.com/.



- [15] 欧彦婵.韶关主城区自行车交通系统规划研究[J].建筑知识:学术刊,2014(B02):93-93. [ Ou Y C. Research on the planning of bicycle transportation system in main urban areas of Shaoguan[J]. Building Knowledge: Academic Journal, 2014(B02):93-93. ]
- [16] "2013 LandScan data for population distribution from the Oak Ridge National Laboratory in USA." [EB/OL].Oak Ridge National Laboratory. <https://landscan.ornl.gov/>.
- [17] OpenStreetMap contributors[EB/OL]. "OpenStreetMap." <https://www.openstreetmap.org>.
- [18] Long Y. Redefining Chinese city system with emerging new data[J]. Applied Geography, 2016, 75:36-48.
- [19] Jin X, Long Y, Sun W, et al. Evaluating cities' vitality and identifying ghost cities in China with emerging geographical data[J]. Cities, 2017,63:98-109.
- [20] Lathia N, Ahmed S, Capra L. Measuring the impact of opening the London shared bicycle scheme to casual users [J]. Transportation Research Part C: Emerging Technologies, 2012,22:88-102.
- [21] Berndt D J, Clifford J. Using dynamic time warping to find patterns in time series[C]. Working notes of the knowledge discovery in databases workshop, 1994:359-370.
- [22] Kate R J. Using dynamic time warping distances as features for improved time series classification[J]. Data Mining and Knowledge Discovery, 2016,30(2):283-312.
- [23] El-Assi W, Mahmoud M S, Habib K N. Effects of built environment and weather on bike sharing demand: A station level analysis of commercial bike sharing in Toronto [J]. Transportation, 2017,44(3):589-613.
- [24] Etienne C, Latifa O. Model-based count series clustering for bike sharing system usage mining: A case study with the Vélib' System of Paris[J]. Acm Transactions on Intelligent Systems and Technology, 2014,5(3):1-21.
- [25] Gutierrez J, Cardozo O D, Garciapalomares J C. Transit ridership forecasting at station level: An approach based on distance- decay weighted regression[J]. Journal of Transport Geography, 2011,19(6):1081-1092.
- [26] 李航.统计学习方法[M].北京:清华大学出版社,2012:55-56. [ Li H. Statistical learning method[M]. Beijing: Tsinghua University Press, 2012:55-56. ]
- [27] Breiman L. Random forests[J]. Machine Learning, 2001, 45(1):5-32.
- [28] Liu Y, Liu X, Gao S, et al. Social sensing: A new approach to understanding our socioeconomic environments [J]. Annals of the Association of American Geographers, 2015,105(3):512-530.
- [29] 牟乃夏,张恒才,陈洁,等.轨迹数据挖掘城市应用研究综述[J].地球信息科学学报,2015,17(10):1136-1142. [ Mou N X, Zhang H C, Chen J, et al. A review on the application research of trajectory data mining in urban cities[J]. Journal of Geo-information Science, 2015,17(10):1136-1142. ]
- [30] Han H, Guo X, Yu H. Variable selection using mean decrease accuracy and mean decrease gini based on random forest[C]. IEEE International Conference on Software Engineering and Service Science, 2017:219-224.
- [31] Zhang L, Zhang J, Duan Z Y, et al. Sustainable bike-sharing systems: Characteristics and commonalities across cases in urban China[J]. Journal of Cleaner Production, 2015,97:124-133.
- [32] Shaheen S A, Zhang H, Martin E. Hangzhou public bicycle: Understanding early adoption and behavioral response to bikesharing in Hangzhou, China[J]. Transportation Research Record, 2011,2247(2247):34-41.
- [33] Campbell A A, Cherry C R, Ryerson M S, et al. Factors influencing the choice of shared bicycles and shared electric bikes in Beijing[J]. Transportation Research Part C, 2016,67:399-414.
- [34] 吴健生,李博,黄秀兰.小城市居民出行行为时空动态及驱动机制研究[J].地球信息科学学报,2017,19(2):176-184. [ Wu J S, Li B, Huang X L. Spatio-temporal dynamics and driving mechanisms of resident trip in small cities[J]. Journal of Geo-information Science, 2017,19(2):176-184. ]