

引文格式: 吴华意, 黄蕊, 游兰, 等. 出租车轨迹数据挖掘进展[J]. 测绘学报, 2019, 48(11): 1341-1356. DOI: 10.11947/j.AGCS.2019.20190210.  
WU Huayi, HUANG Rui, YOU Lan, et al. Recent progress in taxi trajectory data mining[J]. Acta Geodaetica et Cartographica Sinica, 2019, 48(11): 1341-1356. DOI: 10.11947/j.AGCS.2019.20190210.

## 出租车轨迹数据挖掘进展

吴华意<sup>1</sup>, 黄蕊<sup>1</sup>, 游兰<sup>2</sup>, 向隆刚<sup>1</sup>

1. 武汉大学测绘遥感信息工程国家重点实验室, 湖北 武汉 430072; 2. 湖北大学计算机与信息工程学院, 湖北 武汉 430062

## Recent progress in taxi trajectory data mining

WU Huayi<sup>1</sup>, HUANG Rui<sup>1</sup>, YOU Lan<sup>2</sup>, XIANG Longgang<sup>1</sup>

1. State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430072, China; 2. School of Computer Science and Information Engineering, Hubei University, Wuhan 430062, China

**Abstract:** The development of big data technology, internet of thing and precise positioning has promoted the progress of city perception. The increasing taxi trajectory data not only records the pathway of taxis, but also implies the real-time traffic status, the information of urban dwellers' travel rule, urban structure and potential social problems. It is of great significance to mine and analyze the taxi trajectory data for smart transportation, urban planning etc. This paper reviews the field of taxi trajectory data analysis and applications in the past ten years. From the perspective of research methodology, four categories are identified: spatial statistical, time series analysis, graph and network analysis, and machine learning. Each category is reviewed with its current research situation, advantages disadvantages. Later on, applications, hot topics and future trends of taxi trajectory analysis are summarized to four areas including traffic management, resources and environmental protection, city planning, and human mobility. Finally, the current challenges and the future research directions in the field of taxi trajectory data mining are proposed.

**Key words:** trajectory data; data mining; taxi trajectory; review

**Foundation support:** The National Natural Science Foundation of China (No. 41771474)

**摘 要:** 大数据、物联网与精密定位技术的发展推动了城市感知的进步。随着社会活动的与日俱增, 出租车轨迹数据不仅记录了出租车的行车轨迹, 还蕴藏着道路交通状态、城市居民出行规律、城市结构及其他社会问题。通过各种数据分析与挖掘手段对出租车轨迹数据进行深入探究, 对于智能交通、城市规划等有着重要意义。本文综述了近十年国内外基于出租车轨迹大数据的相关研究, 按照空间统计方法、时间序列方法、图论与网络方法及机器学习方法等 4 类, 详细阐述各类方法的研究现状。随后, 本文分析了现有研究的应用领域、热点主题和发展趋势。最后, 本文指出了出租车轨迹数据挖掘研究领域面临的挑战和未来研究方向。

**关键词:** 轨迹数据; 数据挖掘; 出租车轨迹; 综述

**中图分类号:** P208

**文献标识码:** A

**文章编号:** 1001-1595(2019)11-1341-16

**基金项目:** 国家自然科学基金(41771474)

近年来随着物联网感知与高性能计算技术的不断发展, 大规模城市感知数据受到了各个领域的广泛关注。作为最重要的一个数据类型, 轨迹数据隐含了丰富的城市信息, 往往规模巨大且来

源广泛。轨迹数据主要包括人类活动轨迹数据、交通轨迹数据、动物活动轨迹数据和自然现象轨迹数据<sup>[1]</sup>。这些轨迹数据的时空尺度与粒度各有不同, 分别适用于不同研究。其中人类活动轨迹

数据与交通轨迹数据是与人类日常生活联系最紧密的两类轨迹数据,因此有大量研究围绕他们展开。

人类活动轨迹数据,包括微博签到数据、Flickr 照片数据、手机基站定位数据、信用卡消费数据、手机 GNSS 定位数据等,按照时间顺序跟踪这些位置记录就形成关联人在一段时间内的移动轨迹。这类轨迹数据精度不高,通常在 200 m 左右<sup>[2]</sup>,而且大多涉及个人隐私,并没有得到广泛的应用,一般在研究中作为辅助数据用于丰富语义信息。

交通轨迹数据包括两类:一类是通过固定设备采集得到的数据,如城市道路卡口照片、视频监控数据、地铁刷卡数据等;另一类是通过车载 GNSS 设备采样得到的交通工具移动轨迹,如公交车轨迹数据、出租车轨迹数据、船只轨迹数据等。这其中,固定设备数据只能粗粒度地描述人群或车辆在不同固定位置之间的移动,覆盖范围有限;公交车轨迹数据也只能记录运营时间内公交车的固定路线轨迹。出租车不受线路和时间的约束,是最灵活、覆盖范围最广的轨迹数据,且精度较高、较少涉及隐私问题,常被作为轨迹数据研究和应用的主要数据集。

出租车轨迹数据不仅反映城市道路的交通状态,还折射出隐含的城市问题与挑战。通过轨迹数据挖掘可以帮助乘客了解出行信息<sup>[3-6]</sup>、为司机推荐导航路线<sup>[7-8]</sup>、改善出租车的运营管理<sup>[9-12]</sup>。同时,轨迹数据与其他社会、经济、人口数据的关联分析,能发现城市人口流动模式<sup>[13-16]</sup>、社会活动动态<sup>[17-18]</sup>、能源消耗分布<sup>[19-20]</sup>及环境污染状况<sup>[21-22]</sup>等,帮助提高城市管理决策水平。

近年来,围绕出租车轨迹数据挖掘展开了大量研究,并取得了丰硕的成果。同时,随着滴滴出行、Uber 等新型出行方式的兴起,以及神经网络与深度学习方法的热潮,出租车轨迹数据挖掘领域正面临着重大机遇和挑战,有必要对该领域的研究现状进行梳理总结。为此,本文综合分析近十年来出租车轨迹数据挖掘方向的研究成果,从研究方法与研究应用两个维度对出租车轨迹数据挖掘研究成果进行剖析,试图描绘出该领域的发展历程与研究前景,以为为轨迹数据挖掘等相关领域学者提供参考。

本文首先从空间统计、时间序列、图论与复杂网络及机器学习 4 类研究方法阐述出租车轨迹数

据挖掘领域的研究现状;其次从智能交通、环境与资源保护、城市规划及社会感知 4 个应用方面来归纳分析该领域的最新进展;最后讨论该领域目前面临的主要问题和有待探索的研究方向。

## 1 出租车轨迹数据

### 1.1 数据内容

原始的出租车轨迹数据集本质上都是出租车轨迹点集,由多行采样记录构成,每条记录代表一个轨迹点,包括出租车编号、时间戳、经纬度坐标、速度和方向等基本行驶数据,部分数据集还会记录载客状态、车辆类型等信息。从数据集中抽取以车为单位的轨迹点序列,就形成轨迹。此外,出租车轨迹在一些特殊场合,也通过二维曲线或关键位置序列如路段编号字符串、交通格网编号序列等来描述。

具有不同属性特征的轨迹适用于不同研究主题,如载客轨迹与空载轨迹常被用于人群移动<sup>[14,17]</sup>、出行模式<sup>[23-24]</sup>和运营策略<sup>[25]</sup>等研究;低速行驶轨迹可以用于发现交通拥堵现象<sup>[26]</sup>;异常轨迹能作为评估司机驾驶行为<sup>[27]</sup>或感知交通突发事件<sup>[28]</sup>的依据。

### 1.2 数据特点

(1) 覆盖范围广。出租车轨迹数据在时间和空间尺度上的覆盖范围比其他交通轨迹数据更广。出租车运营时间可达全天 24 h,而且出租车行驶在城市交通路网中,不受线路制约。如图 1 所示,武汉市一天的出租车轨迹数据可覆盖城市中心路网的 80% 以上。



图 1 武汉市一天的出租车轨迹数据覆盖范围

Fig. 1 Road network coverage of one-day taxi trajectories in Wuhan

(2) 采样密度高。出租车轨迹数据的采样间隔一般在 1 min 以内,部分数据集可达到 3 s 以内,能完整地记录出租车的行驶路径,具有时空序列性和连续性。

(3) 位置精度高。出租车轨迹数据是通过车载 GNSS 采集获得的出租车位置数据,其精度较高,一般为 5~20 m<sup>[2]</sup>。同时由于出租车始终行驶在城市路网中,因此可以通过地图匹配等方法进一步提高数据的位置精度<sup>[29]</sup>。

(4) 数据规模大。由于城市出租车数量多、运营时间长、采样密度高,因此出租车轨迹数据集往往规模庞大,中心城市如武汉等一天就可产生 200 多万条出租车轨迹数据记录。而滴滴出行等商业平台每日新增轨迹数据就超过 106 TB,每日处理数据更是达到 4875 TB<sup>[30]</sup>。

(5) 蕴含信息丰富。出租车轨迹数据中记录的最直接信息是出租车在每个时间点的位置坐标,这些连续坐标点刻画了出租车在一段时间内的移动轨迹。移动轨迹体现出出租车司机的运营特征,包括寻客策略<sup>[11]</sup>、路径选择偏好<sup>[31]</sup>、载客范围<sup>[10]</sup>或欺诈绕路现象<sup>[27]</sup>等。

出租车在城市交通中所占比重较大,达到总交通流的 20% 左右,在一些关键区域甚至可以达到 50%<sup>[32]</sup>。国家交通运输部发表的《2018 年交通运输行业发展统计公报》表明,2018 年出租车客运量达到 351.67 亿人。因此,出租车速度和密度等属性一定程度上能够反映城市交通流和人群移动的整体情况,从而用于估计或预测通行时间和交通流量、监测交通拥堵情况<sup>[4-6]</sup>,并结合 POI 数据进行城市规划结构、用地分类<sup>[33-35]</sup>等的进一步探测。

### 1.3 数据预处理

原始的出租车轨迹数据多存在异常点、噪音点、轨迹点漂移等问题,因此在对数据进行挖掘分析之前通常要对数据预处理。出租车轨迹数据预处理过程包括数据清洗、地图匹配、数据轨迹化与轨迹划分、质量评价 4 个部分。

数据清洗主要是为了剔除数据中的异常点和噪音点,其中卡尔曼滤波<sup>[36]</sup>、粒子滤波<sup>[37]</sup>等都是常用的数据清洗方法。而地图匹配的目的是要将出租车轨迹点准确地匹配到城市路网上。现有的地图匹配算法可分为确定性地图匹配算法与不确定性地图匹配算法两大类<sup>[29]</sup>,具体有投影算法<sup>[38]</sup>、概率统计算法<sup>[39]</sup>、模糊逻辑算法<sup>[40]</sup>、相关

性分析算法<sup>[41]</sup>等。数据轨迹化则是将原始的出租车轨迹点数据转化为线数据的方法。基本思想是将每辆出租车的连续 GNSS 采样点按照时间顺序先后连接起来,映射到地图上即得到一条与道路地图相匹配的有向曲线<sup>[42]</sup>。出租车轨迹数据质量评价方面可分为两部分,一是对出租车轨迹数据定位精度进行评定;二是出租车与真实的交通流之间存在密度与速度<sup>[43]</sup>上的差异,这些差异会导致最终交通流计算或预测结果中存在误差<sup>[44]</sup>,因此还需要对出租车轨迹数据的计算精度进行评定与校正。

### 1.4 开放数据集

目前通过互联网可免费获取到许多公开出租车轨迹数据集。如微软 T-Drive 项目提供的 2008 年北京出租车一周内的轨迹数据<sup>[45]</sup>;纽约 Taxi & Limousine Commission(TLC)官方网站发布的纽约市出租车轨迹数据集<sup>[46]</sup>,时间区间为 2009—2018 年,可以月为单位下载所需的数据;CRAWDAD 是达特茅斯学院的一个无线数据资源网站<sup>[47]</sup>,包含了大量出租车轨迹数据集,如旧金山海湾地区 500 辆出租车 30 d 内的轨迹数据、上海 4000 辆出租车在 2007 年 2 月 20 日这一天内 24 h 的轨迹数据、罗马市区 320 辆出租车在 2014 年 2 月 1 日至 3 月 2 日一个月内的行驶数据等;滴滴出行通过盖亚数据开放计划向研究者免费开放了部分网约车数据<sup>[48]</sup>,目前已提供西安市和成都市 2016 年的局部轨迹。

## 2 研究趋势分析

出租车轨迹数据挖掘相关研究的历史最早可追溯至 1999 年,第 3 代蜂窝通信网络与基于位置的服务(LBS)在这一时期开始兴起,如图 2 所示。移动特性的知识对规划、设计和运行通信网络具有重要意义,因此研究者们利用全球定位系统(GNSS)测定一段时间内出租车的位置数据,来评估物体的移动特性<sup>[49-50]</sup>。之后这些出租车轨迹数据开始被尝试用于估计路段速度<sup>[51]</sup>、进行出租车调度<sup>[52]</sup>和监测交通排放量<sup>[53]</sup>等智能交通与环境保护方面。

在 2011 年第十三届普适计算国际会议上,文献<sup>[54]</sup>提出利用出租车轨迹数据检测城市规划中的缺陷。此后,利用出租车轨迹数据发现并评价城市规划结构的研究成果开始大量涌现,如发现城市中的功能区域<sup>[55]</sup>、识别城市土地利用分



类<sup>[34,56]</sup>、评估城市交通系统应对突发事件的能力<sup>[10]</sup>等。基于出租车轨迹数据进行城市动态研究的文献在同一时期开始出现,如文献[57]以北京1万多辆出租车为研究对象,对人类移动行为建模分析;文献[13]基于上海158万条出租车轨迹数据,识别城市居民出行模式;文献[58]从葡萄牙里斯本5个月的出租车轨迹数据中发现城市流动规律等。

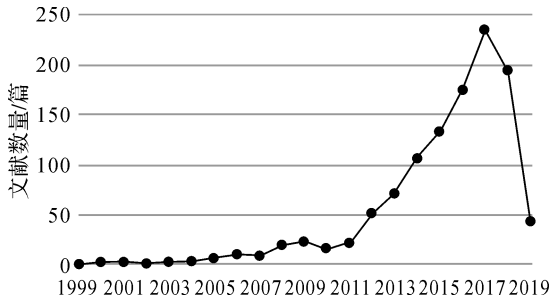


图2 1999—2019年出租车轨迹数据挖掘领域文献数量  
Fig.2 Numbers of publications on taxi trajectory data mining from 1999 to 2019

总的来说,出租车轨迹数据挖掘领域在2011—2012年间发生了重大变化,出现了两个新的研究方向:一是城市规划,指利用出租车轨迹数据发现、识别并评估静态的城市规划和城市结构;二是社会感知,指基于出租车轨迹数据对城市内人群活动的动态变化、移动模式进行分析和监测。这时期出现的一系列重要成果彻底改变了之前出租车轨迹数据只能在交通领域内得到应用的情况,出租车轨迹数据挖掘领域进入新的发展阶段。

智能交通、资源与环境保护两个应用方面贯穿出租车轨迹数据挖掘领域始终,尤其智能交通方面,是出租车轨迹数据的直接应用领域,直到现在也是每年发表研究成果数量最多的方向,如图3所示。城市规划和社会感知虽然都是于2011年前后开始起步,但这两个方面的联系较为紧密,城市规划结构可以看作是社会感知研究的基础,因此城市规划方面相对来说得到了更多的关注和发展,但近年来社会感知研究也开始出现上升的趋势。

空间统计、时间序列、图与复杂网络等传统的分析挖掘方法是该领域早期的主要研究方法,如图4所示。空间统计方法适用范围广,能解决多种类型的问题,如热点区域发现、通行时间估计、预测等。时间序列方法主要用于出租车轨迹数据

的频繁模式挖掘与相似性度量。图与复杂网络方法则是对基于出租车轨迹数据抽取出的网络结构如道路网、市民出行网等进行网络相关特性的挖掘分析。值得注意的是,从2014年开始,由于计算机处理能力和硬件设备的提升,机器学习中的神经网络与深度学习方法重新吸引了各领域科学家的目光,也开始被用于出租车轨迹数据挖掘领域中预测通行时间、打车需求等问题。

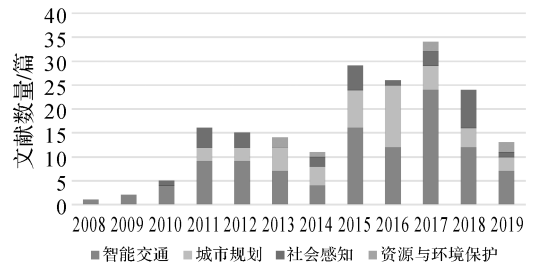


图3 2008—2019年出租车轨迹数据挖掘研究数量的应用领域分布

Fig.3 Application distribution of researches on taxi trajectory data mining from 2008 to 2019

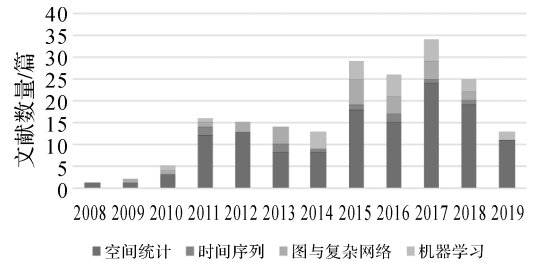


图4 2008—2019年出租车轨迹数据挖掘研究数量的研究方法分布

Fig.4 Methodologies distribution of researches on taxi trajectory data mining from 2008 to 2019

2014—2015年可看作是出租车轨迹数据挖掘领域的第2个分界点。从2014年至今,出租车轨迹数据挖掘领域都依然处于转型过渡阶段。这一时期,除了引入神经网络与深度学习方法外,更重要的是出现了滴滴出行、Uber、共享单车等新型出行方式。《中国共享经济发展年度报告(2019)》表明,网约出租车客运量占总出租车客运量的比重达到36.3%,这对传统出租车行业造成了一定冲击。许多城市的出租车公司与滴滴出行合作,传统出租车也能通过在线接单来寻客,因此出租车行业整体的运营方式也改变了。出租车行业的巨大转变与前沿技术的发展,对出租车轨迹数据挖掘领域,既是机遇也是挑战。

一方面,基于滴滴平台获得的网约出租车轨迹数据比传统出租车轨迹数据的采样密度更高,达到 1~3 s 一个轨迹点;采集精度更高,由于滴滴依据行驶轨迹计费,GNSS 数据与实际行驶路径匹配准确率可达到 100%;覆盖范围更广,以北京为例,62%的滴滴网约车起点或终点位于城市边缘公共交通覆盖不足地区,且包含大量跨城轨迹<sup>[59]</sup>。这为出租车轨迹数据挖掘研究提供了质量更高、更具代表性的数据源。

另一方面,滴滴网约车与传统出租车之间因其行业背景不同而存在差异。例如,主要寻客方式由过去的巡游寻客变为在线派单,滴滴网约车本身存在专车、快车、出租车和顺风车等多种运营形式,这些因素影响行车经验挖掘、载客点推荐等问题的结果。此外,神经网络与深度学习方法的应用尚未深入。在解决不同问题时应该怎样选择网络?在使用不同网络时又应该怎样选择出租车轨迹数据的输入形式?神经网络与深度学习对轨迹数据来说是否真的适用?这些问题都还未得到完整解答。从图 2 可看到,2017 年后出租车轨迹数据挖掘研究的文献数量有所回落,一定程度上表示出租车轨迹数据挖掘领域渐渐开始进入新的瓶颈期,需要进一步探索新型前沿技术在该领域的应用,以及与多源数据的联合分析。

近年来,部分学者发表了出租车轨迹数据挖掘综述成果,为本文方法和应用综述研究奠定了一定基础。文献[60]从社会动态、交通动态和行为动态 3 方面来对出租车轨迹研究进行分析总结;文献[61]以城市计算为出发点,综述出租车轨迹数据在城市感知和规划方面的应用;文献[62]总结了出租车轨迹数据在交通状态分析、运营管理及支持和路径规划及预测等智能交通方面的研究。这些研究综述的角度有所不同,各有侧重。本文综述研究的重点是出租车轨迹数据挖掘与分析的理论和方法,以及总结该领域研究主题变化趋势和挑战。在出租车轨迹数据领域面临转变的背景下,本文旨在梳理出租车轨迹数据挖掘研究的现状,同时思考未来可能的发展趋势,为广大研究人员掌握现有研究状况、定位和规划未来研究方向提供参考。

### 3 出租车轨迹数据挖掘方法

多年来,各个领域的学者都不断试图从出租车轨迹数据中挖掘出新的知识与经验,所涉及的

理论与方法覆盖多个学科,采用的主要方法可以归纳为 4 类:空间统计、时间序列方法、图论与复杂网络及机器学习。

#### 3.1 空间统计

空间统计方法能直观地帮助研究者获得数据集的整体分布特征,从而选择相应的统计模型对轨迹数据进行后续分析。因此,空间统计方法是应用最多最广泛的基本数据挖掘技术,它可以作为独立的方法对轨迹数据进行分析,也可以作为前期的数据处理方法为后续的研究提供依据。

##### 3.1.1 探索性空间分析

探索性空间数据分析方法的特点是对数据集的总体不作假设,而是使用统计图表、图形和统计概括等探索性、描述性的方法对数据的特征进行分析和描述<sup>[63]</sup>。在面对出租车轨迹数据时,数据的整体特征对研究者来说往往是未知的。采用该类方法可以获得数据集的整体统计特征,计算得到载客里程、载客时间、载客数、上下客点数、行驶速度、换乘时间间隔等统计量,帮助发现时空分布规律,进而建立统计模型解决相应问题。

例如,司机收入这一统计量通常被作为衡量司机寻客策略有效性的标准<sup>[64-67]</sup>。在此基础上,通过跟踪高收入司机的行车轨迹,可以为其他出租车司机推荐载客点和提供行车指导<sup>[11,25,68-69]</sup>,从而提高出租车服务质量和司机收入水平。

除了对出租车行业本身的探索与评价之外,部分研究综合考虑公交站点分布<sup>[70]</sup>、网约车服务<sup>[71-72]</sup>、天气<sup>[9]</sup>等外界因素从而进一步发掘出租车行为规律。在统计量的基础上进行简单计算能粗略估计城市道路交通状况和交通容纳量<sup>[3,73-75]</sup>,提取居民出行与移动的规律<sup>[23,57,76]</sup>。基于与道路网和社会经济数据的联合分析,可以探测城市功能区域和土地利用布局<sup>[56,77]</sup>、估计城市汽油消耗和尾气排放总量<sup>[19,20,78]</sup>。

探索性空间分析方法能帮助简单、直观地把握整个数据集的特征,但结果都是比较概括性的结论,容易受到数据处理过程中的误差影响,不能充分挖掘轨迹数据的价值。

##### 3.1.2 空间聚类

空间聚类方法是数据挖掘领域的关键技术之一。根据聚类对象的不同,出租车轨迹数据的空间聚类方法可分为点聚类方法和轨迹聚类方法。

点聚类方法主要是对出租车轨迹数据中的点数据进行空间聚类,包括出租车的 GNSS 采样

点、停留点或者是上下客点。常用的聚类算法有 k-means 算法和 DBSCAN 算法。聚类分析的结果直观简明,能直接从结果中发现特定的出租车行为规律。例如,通过对出租车上下客点进行点聚类能获得热门上下客区域,从而向出租车司机推荐最佳载客点<sup>[79-81]</sup>。下客点聚类结果反映了市民出行偏好,可辅助城市规划布局<sup>[16,34,82]</sup>。道路网中大量低速聚集状态的出租车集群是道路拥挤现象的表现,因此,通过道路低速点聚类能对城市道路拥堵模式进行估计和分析<sup>[83-84]</sup>。文献<sup>[85]</sup>还基于空间聚类提出一种城市交叉口自动识别方法,为城市交通路网局部结构的探测提供了新途径。

点聚类得到的结果通常是要素聚集的面状区域,但在数据量较大时,难以快速识别出不规则形状的聚类簇。同时,如何确定具有多重属性的点数据相似系数也是目前难点之一。如文献<sup>[86]</sup>针对现有算法在出租车载客热点区域提取结果的不足,提出一种顾及路网约束的改进 DBSCAN 算法,将道路拓扑关系与路段长度数据加入聚类算法的相似性度量中。

轨迹聚类是对出租车行驶轨迹进行聚类,得到的结果是具有相似形态且聚集的轨迹簇。例如,对下客热点到上客热点之间的轨迹进行聚类,得到最具有寻客潜力的最优路径<sup>[87]</sup>。或者,跟踪特定区域间的下客轨迹,从而识别出城市交通流向规律<sup>[88]</sup>。此外,文献<sup>[26]</sup>基于相似轨迹聚类的思想提出拥堵同伴的概念和发现算法,筛选出可能发生拥堵的浮动车数据,进而对拥堵区域的变化趋势进行预测。

轨迹聚类的难点是如何定义轨迹在时空维度上的相似性度量,基于整体的轨迹聚类会忽略子轨迹的细节信息,而基于分段的轨迹聚类会分割轨迹,使一条轨迹分属于多个聚类簇。从聚类准确度出发,多数研究者认为基于分段的轨迹聚类方法粒度更细、准确度更高。

### 3.1.3 空间回归分析

空间回归分析是定量描述空间数据之间关系的常用方法。出租车轨迹数据记录了出租车的行驶路线,在空间分布上具有明显的空间异质性,并与社会经济因素和城市规划布局紧密相关;同时,也受城市居民的出行规律影响,在时间上表现出明显的周期性。因此,空间回归分析方法常被用来拟合并预测出租车客流量,并量化分析乘客需

求与社会经济因素、环境质量之间的关系。

例如,各种回归模型被用来拟合出租车上客点的数量分布<sup>[89]</sup>,再对热点地区乘客的时空变化和等待时间进行预测<sup>[10,90-92]</sup>。或者,利用地理加权回归模型来刻画出租车的空间异质性<sup>[93]</sup>,并将客流量与社会人口和建筑环境变量关联起来。

空间回归分析能准确地刻画因子之间的相关性和拟合程度,但其应用范围会受到回归模型假设条件的限制,且模型的构建通常是一个探索性的过程,因此结果依赖于研究者对因子的选择和表达。

### 3.1.4 密度分析

密度分析是用来计算整个研究区域内数据聚集情况的方法,利用离散的点(线)生成连续的曲面,从而发现要素较为集中的区域,包括普通的点(线)密度分析方法和核密度分析方法。

普通的点(线)密度分析对落入搜索区域的点或线进行求和,计算区域单位面积内的元素数量。出租车停留点密度能作为有效估计区域内打车需求与出租车充电需求<sup>[94-95]</sup>的指标、提取城市热点区域和人们出行规律<sup>[35]</sup>的依据;同时,跟踪出租车停留点密度变化过程可以探测城市中上下客事件的冷热点集群变化规律<sup>[17,96]</sup>。这种方法的局限在于会导致网格边缘出现过大落差,也就是“断崖”现象。

核密度分析方法中,落入搜索区的点具有不同的权重,分析的结果是平滑、连续的表面,避免了普通的点(线)密度分析中“断崖”的出现。通过核密度分析方法能提取出租车集聚的热点区域<sup>[97]</sup>、分析出租车行为的时空分布模式<sup>[98]</sup>。与普通密度分析方法相比,核密度分析方法提取的区域边界更平滑、更合理,但由于带宽的设置会直接影响结果的好坏,因此根据数据集的特点选择合适的带宽是该方法的关键。

### 3.1.5 其他方法

除了上述方法之外,还有大量研究使用基于空间统计理论的其他方法来挖掘出租车轨迹数据,如利用 OD 矩阵分析<sup>[15,99]</sup>、隐马尔可夫模型<sup>[100]</sup>、离散选择模型<sup>[31]</sup>等对出租车上下客行为进行建模预测,通过图像处理从出租车轨迹数据中提取路网结构<sup>[101]</sup>等。另外统计检验方法<sup>[102]</sup>、基于证据理论的方法<sup>[103]</sup>、主成分分析方法<sup>[28]</sup>等也逐渐受到人们的关注。总之,空间统计方法种类繁多,所应用的研究领域也涵盖甚广,但



使用时多受前提假设的限制,是一类基础而又需谨慎使用的方法。

### 3.2 时间序列方法

时间序列数据反映了一类事物或现象随时间的变化状态或程度。出租车轨迹数据可以看作是按照一定时间间隔通过定位设备采样得到的时间序列数据,因此部分研究利用时间序列方法的思想对出租车轨迹数据进行分析。

例如,文献[104]使用 PrefixSpan 的思想搜索轨迹的频繁模式子序列,并构建一个频繁轨迹图模型,该模型能计算得到最佳打车推荐结果。文献[105]对 GSP 算法进行改进以应对大规模轨迹数据,挖掘出租车频繁轨迹特征<sup>[106]</sup>,从而分析出租车司机的驾驶经验与行为模式,并为其他司机提供路径规划。

相比轨迹数据,时间序列数据的相似性度量要简单得多,因此有部分研究首先将轨迹数据转换为时间序列数据,再对其进行轨迹聚类,从而探测十字路口<sup>[107]</sup>和道路网变化<sup>[108]</sup>、挖掘司机行为规律<sup>[109]</sup>等。但其缺点在于转换过程中会损失信息,且时间序列方法是用于分析一维数据的方法,在应用于二维数据的过程中存在许多限制。

### 3.3 图论与复杂网络

出租车轨迹数据可以看作是城市道路网的映射,能够从中提取出城市道路网的主干结构。同时,由于出租车轨迹数据记录了居民出行路径,将出行热点区域作为节点,区域间的交通轨迹作为边,可以构建乘客出行网络图。利用从轨迹数据中提取出的图和网络特性,通过图论与复杂网络相关方法能解决更多特定问题。

#### 3.3.1 路径搜索方法

路径搜索方法是解决在图或网络中怎样按照特定的规则从一个顶点到达另一个顶点问题的一类方法,实际应用时通常先构建交通道路网,再遵循长度最短或时间最短的条件在道路网中计算最佳路径,最后将计算的结果推荐给司机,为司机提供路径规划和导航<sup>[110-113]</sup>。例如文献[7]基于 k-最短路径算法提出了一种具有负载均衡分配的时空轨迹模型,该模型为出租车司机提供推荐点间的最短路线。最短路径搜索可以提供理想情况下的路径参考,然而在现实情况下,距离最短的路径不一定是司机第一选择的路径,还受到许多环境因素的影响如道路拥堵、交通限速等,同时也依赖于道路网构建合理与否。

#### 3.3.2 可达性衡量方法

在城市道路网络中,可达性指的是道路网中从一个区域到达某个特定区域的容易程度,是评价城市区域和道路网规划是否满足市民需求的指标<sup>[114-115]</sup>。如文献[116]构建了两个矩阵模型来分别表示每条轨迹的起讫网格和经过的网格坐标,并采用基于位置的可达性衡量法对研究区域的城市规划进行评价。这类方法要求研究者对城市规划和交通规划有比较深刻的认识,充分理解可达性的内涵与类型,同时考虑各个指标的假设条件限制。

#### 3.3.3 社区发现方法

社区发现方法是用于发现复杂网络中的社区结构的一种算法,社区是一个子图,而整个网络可以看作是由多个社区构成的。其中社区内节点与节点之间的连接很紧密,而社区与社区之间的连接比较稀疏,因此社区发现方法与聚类方法有一定的相似性。社区发现方法可以探测道路网、市民出行 OD 网络的子结构,并分析评价城市结构与空间的相互作用,跟踪其动态演化过程<sup>[33,117]</sup>。如文献[118]基于出租车轨迹数据,把空间单元看作节点,人类运动流看作边,建立了空间嵌入式网络模拟城市内部空间互动。目前社区发现方法所研究的对象已不局限于静态单一的网络,如何在多模式、多维度网络当中解决不同模式及维度下的信息融合、共享及动态演化是该方法的最大挑战。

#### 3.3.4 中心性分析方法

中心性(centrality)是量化判定网络中节点重要性的指标,道路网的中心性评价结果可以用以估计城市交通流量并分析城市路网的结构特征<sup>[4,119]</sup>。常见的中心性指标有度中心性、介数中心性和 PageRank 中心等,如文献[120]首先评价了道路网的中心性对交通量的预测能力,然后提出了基于现有中心性测度的扩展方法。但现有的大部分中心性指标都是针对静态网络的,自然界中的复杂系统本质是不断变化,面向动态复杂网络的中心性度量还有待进一步研究。

### 3.4 机器学习

随着近年来人工智能技术的发展,机器学习受到了各个领域的广泛关注。在出租车轨迹数据挖掘中,机器学习方法自发、主动的特征学习过程满足了轨迹数据挖掘的需求,同时,轨迹数据其规模巨大、来源广泛的特点为机器学习方法提供了

足够的样本。在出租车轨迹数据挖掘中常用的机器学习方法有神经网络、贝叶斯模型、支持向量机及遗传算法等。

### 3.4.1 神经网络

神经网络是利用计算机模拟人的神经系统结构,并通过它从大规模数据中学习规律和知识的一种过程<sup>[121]</sup>。神经网络模型的一个重要应用是预测城市中不同区域的打车需求与寻客潜力,并将其推荐给出租车司机<sup>[8,122-123]</sup>。文献[12]以路段集群为预测单位,采用基于排名的 ELM 回归模型(极限学习机)预测未标记集群的寻客潜力。神经网络模型同样能分析道路交通状况,文献[124]采用一种深度限制的玻尔兹曼机和递归神经网络结构模型基于出租车轨迹数据对交通拥堵的演化过程进行建模和预测,能够快速识别出交通拥堵区域。神经网络学习的能力非常强大,但由于是直接从数据出发得到规律和经验,其结果往往难以解释,而且前期需要大量数据集进行训练。

### 3.4.2 贝叶斯分类

对于分类问题来说,贝叶斯理论考虑的是在所有相关概率都已知的理想情形下,如何基于这些概率和误判损失来选择最优的类别标记。不同于其他的分类预测方法,贝叶斯分类不仅利用模型信息和数据信息,还会充分利用先验信息。贝叶斯分类方法能够基于历史数据预测交通流量<sup>[22]</sup>、空载出租车数量<sup>[125]</sup>、乘客出行目的<sup>[126]</sup>及司机决策行为<sup>[127]</sup>。如文献[22]在估计每个路段的交通状况基础之上,利用一个基于贝叶斯网络的非监督图形模型 TVI 来学习交通速度、流量和密度之间的关系,并计算得到交通流量。该方法的优势在于分类效率稳定、容易解释,而且所需参数少,适用于大规模数据。但贝叶斯分类中的属性独立性假设往往在实际中不成立,且分类之前需要知道先验概率,存在分类决策错误率。

### 3.4.3 支持向量机

支持向量机(support vector machine, SVM)是一类二分类模型,其基本模型定义是特征空间上间隔最大的线性分类器,其学习策略就是间隔最大化。同时作为一种预测算法,通常能对有效客源、交通速度以及出租车排队等待时间等<sup>[128-130]</sup>进行预测。如文献[130]基于支持向量机模型建立了一个短期交通速度预测模型,并与人工神经网络、k-最近邻模型、基于历史数据的模

型、基于移动平均的模型等方法进行比较,其中带有时空参数的支持向量机模型展现了良好的性能。支持向量机方法对小集群分类效果比较好而且可以避免局部极小点问题,但其局限是难以应用于大规模数据。

### 3.4.4 遗传算法

遗传算法是一种最优化方法,通过模拟遗传学中生物进化过程来搜索最优解,常应用于解决城市交通领域的各类优化问题,如对电动出租车充电站选址进行优化<sup>[131]</sup>、对出租车应急调度方案进行优化<sup>[81,132]</sup>等。文献[80,132]利用多种群遗传算法进行最短路径计算,实现出租车应急调度模型,同时为司机推荐最佳载客路线。遗传算法采用的是启发式搜索,易于并行化处理,但可能出现早熟现象,并且处理大规模数据时效率不高。

## 4 出租车轨迹数据挖掘应用

目前出租车轨迹数据挖掘领域研究的应用很广泛,其中在智能交通、资源与环境保护、城市规划、社会感知等方面的应用,是目前的应用热点。

### 4.1 智能交通

(1) 交通状态分析。从出租车轨迹数据中提取城市道路交通相关的信息,并利用这些信息来识别交通拥堵现象<sup>[26,75,124]</sup>与特殊社会事件<sup>[28,133]</sup>、估计并预测交通流量<sup>[4,134]</sup>和行程时间<sup>[6]</sup>等的一类研究,这类研究的意义是辅助管理者及时处理道路交通中的突发事件。

(2) 出租车运营管理和支持。将出租车轨迹数据挖掘结果用于改善出租车的运营管理、提高出租车司机平均收入是该领域应用最为广泛的一个方面。如将出租车轨迹数据中提取得到的上下客点、载客里程、空载率等数据作为出租车行为分析<sup>[31,70]</sup>、打车需求预测<sup>[89,92,100]</sup>的依据,从而为出租车的寻客策略<sup>[8,11]</sup>和调度方针<sup>[132]</sup>提供指导。

这一应用领域相对广泛和深入,一方面是因为出租车轨迹数据本身是通过安装在出租车上的 GNSS 定位设备采样而来,轨迹数据呈现出的时空规律实际上就是出租车行为的映射,因此与出租车行业经营与管理紧密相关;另一方面,智能交通领域的研究成果能直接改善乘客与司机信息不对等的现状,并切实解决现实中打车难、司机收入低等问题,具有现实意义性,而且利用数据作为基础支撑,直观且具有说服力。

(3) 路径规划与异常检测。路径规划主要指



为出租车司机规划特定两个位置之间的最短路径,在路径规划时多会结合从出租车轨迹数据中挖掘到的高收入司机的择路经验<sup>[7,112-113]</sup>。异常检测则指对出租车行驶路径进行异常检测,如将待检测路径与从轨迹数据中提取的常规路径进行比较,从而判断路径是否存在绕路、超速等异常现象<sup>[135-136]</sup>。该类研究目的是预防司机的欺骗行为及监测突发事件的发生。

#### 4.2 资源与环境保护

车辆在行驶过程中的燃油消耗、尾气排放等活动是自然资源与环境保护方面关注的一个重要部分。将出租车作为城市行驶车辆的一个样本,计算出出租车行驶里程、行驶速度并提取加油行为<sup>[19]</sup>可估计城市内整体油耗量<sup>[78]</sup>和尾气排放量<sup>[20]</sup>,并可作为相关政策制定的辅助依据。

#### 4.3 城市规划

(1) 城市规划优化。这里的城市规划主要指城市功能区域规划及道路交通规划,通过出租车轨迹数据挖掘而获得的城市功能单元<sup>[35]</sup>、区域可达性指标<sup>[116]</sup>和道路容纳量<sup>[74]</sup>等,能作为评价现有城市规划的指标,并为城市规划优化<sup>[95,98]</sup>提供依据和支撑。

(2) 路网更新。从出租车轨迹数据中能提取出较为清晰的交通路网结构,通过比对历史路网数据可以识别出路网变化,实现路网更新<sup>[101,108,137]</sup>。同时,结合出租车速度与方向信息对道路交叉口进行分析能够识别复杂的交叉口结构<sup>[85,138]</sup>、学习交叉口交通规则<sup>[107]</sup>。

#### 4.4 社会感知

社会感知是指利用各类地理空间大数据研究人类时空间行为的一类研究。出租车轨迹数据是地理空间大数据中的一类重要数据,它所反映的人类时空间行为主要体现在城市人群出行规律与人类移动特性两个方面。

(1) 城市人群出行规律。城市道路网中或城市各区域间交通流方向与流量的变化可以反映市民的日常出行规律变化。这类研究通过挖掘交通热点、跟踪热点区域间的交通轨迹来探究人群的移动规律<sup>[17-18,117]</sup>与市民社会活动与娱乐生活的偏好规律<sup>[16]</sup>,其关注的是城市尺度上人群的集体流动模式。

(2) 人类移动特性。人类移动特性研究是利用特定的统计模型,在个体层面上对人类出行的步长、时间与频率等量进行建模分析<sup>[23-24]</sup>。与前

一类应用不同的是其不局限于特定的出行目的或时间段,而是更多地关注个体的随机漫步模式<sup>[139]</sup>,通过数理统计方法对人类移动进行分析。

值得注意的是近年来在城市规划、社会感知方面出现了许多有价值的工作,出租车轨迹数据挖掘研究已不仅应用于智能交通领域与解决打车问题,而是逐渐开始扩展应用领域外沿,寻求与其他专业领域相结合的综合应用,如提取城市功能结构、挖掘社会动态等。现如今,城市计算相关主题已渐渐成为出租车轨迹数据挖掘领域的热点话题,如何利用城市中产生的大规模数据来辅助决策、应对城市化进程中出现的各种挑战是未来值得继续深入探索的方向。

## 5 研究展望

出租车轨迹数据是城市交通状态、城市结构规划及人口流动规律的数据映射,通过挖掘出租车轨迹数据可以改善城市交通服务、优化城市规划、提高市民生活质量。但是出租车轨迹数据挖掘研究的相关理论与技术仍然亟待深入。基于现有研究,本文提出还需进一步探索的几个方向:

### 5.1 支持轨迹大数据实时清洗、管理、挖掘的高性能计算

轨迹数据的规模越来越庞大,与此并存的还有数据缺失、精度低、语义信息单一等一系列问题,如何提高数据质量并开发出更高效的轨迹数据实时挖掘方法是一直以来面临的难题。一方面,急需结合大数据、云计算<sup>[140]</sup>等新兴技术,在充分挖掘轨迹数据的前提下,提高挖掘方法的精度和效率。另一方面,轨迹数据压缩技术能在保证轨迹数据的结构和语义完整性的前提下,尽可能地减小轨迹数据量。结合分布式计算、量子计算等高性能技术,设计高效的轨迹数据实时压缩算法与管理查询机制是未来的研究方向。

运用出租车轨迹数据进行司机行为分析或市民出行目的预测等研究时会涉及个人隐私问题,在不暴露用户敏感信息的前提下进行有效的数据挖掘,既能挖掘出规律,又能隐私保护,并严格遵守法律法规。

### 5.2 基于泛在信息融合的跨领域研究

随着传感网、物联网、射频识别技术与智能嵌入技术的发展,信息社会正逐步转变为泛在信息社会,人与物、物与物之间交流所产生的泛在信息关联分析将带来大量潜在新型应用。出租车轨迹

数据本身语义信息有限,与其他泛在信息如共享单车数据、夜光数据<sup>[141]</sup>、POI数据、微博签到数据、交通刷卡数据、道路卡口数据等进行联合分析能扩展应用领域范围,进行多领域多学科的综合应用研究,有效解决城市问题。其难点在于泛在信息标准不一、结构多样,需要提出成熟可行的数据度量与融合方法。

### 5.3 基于轨迹大数据的城市和城市群建设与管理的新应用

让健康的人生活在健康的环境里,让生活在城市中的人都有安全感是城市规划与设计的一个重要目标。出租车轨迹大数据所反映的正是城市居民的移动行为和出行偏好,能为城市规划优化、改善城市环境提供依据和支撑。如基于出租车轨迹大数据探测交通网络中车道设计与实际承载交通流量不相匹配的区域,从而对交通网络进行重新规划与建设,在一定程度上缓解城市拥堵现象,从而实现城市的健康与安全。

同时城市本身也处于不断发展的过程中,城乡一体化和城市群的发育发展等,也提出了一系列新的课题,为出租车轨迹数据的挖掘提供了新的空间。

### 5.4 结合轨迹数据与新兴信息技术的创新研究

从出租车轨迹大数据中能挖掘出新的规律和知识,结合知识图谱与知识库等技术对这些获取到的知识进行组织、关联和管理能进一步发挥其应用价值。目前已有部分研究尝试将神经网络、深度学习等智能方法应用于出租车轨迹数据挖掘,但大多并未深入。大规模的历史轨迹数据集为机器学习提供了足够的训练样本,机器学习方法主动的特征学习过程能充分挖掘轨迹数据的价值,在未来这仍是一个前景广阔的方向。此外,将出租车轨迹数据作为一类辅助数据源用于语义理解、自然语言识别等领域也是值得尝试与探索的一个方向。

## 参考文献:

[1] 高强,张凤荔,王瑞锦,等. 轨迹大数据: 数据处理关键技术研究综述[J]. 软件学报, 2017, 28(4): 959-992.  
GAO Qiang, ZHANG Fengli, WANG Ruijin, et al. Trajectory big data: a review of key technologies in data processing[J]. Journal of Software, 2017, 28(4): 959-992.

[2] 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.

[3] 孙超,张红军,陈小鸿. 基于多源浮动车数据融合的道路交通运行评估[J]. 同济大学学报(自然科学版), 2018, 46(01): 46-52.  
SUN Chao, ZHANG Hongjun, CHEN Xiaohong. Road traffic operation assessment based on multi-source floating car data fusion [J]. Journal of Tongji University (Natural Science), 2018, 46(01): 46-52.

[4] ZHAO Shuangming, ZHAO Pengxiang, CUI Yunfan. A network centrality measure framework for analyzing urban traffic flow: a case study of Wuhan, China [J]. Physica A: Statistical Mechanics and Its Applications, 2017(478): 143-157.

[5] DUAN Zongtao, YANG Yun, ZHANG Kai, et al. Improved deep hybrid networks for urban traffic flow prediction using trajectory data[J]. IEEE Access, 2018(6): 31820-31827.

[6] TANG Kun, CHEN Shuyan, LIU Zhiyuan. Citywide spatial-temporal travel time estimation using big and sparse trajectories[J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 19(12): 4023-4034.

[7] WU Liang, HU Sheng, YIN Li, et al. Optimizing cruising routes for taxi drivers using a spatio-temporal trajectory model [J]. ISPRS International Journal of Geo-Information, 2017, 6(11): 373.

[8] HUANG Zhenhua, ZHAO Zhenqi, E Shijia, et al. PRACE: a taxi recommender for finding passengers with deep learning approaches[C]//Proceedings of the 13th International Conference on Intelligent Computing. Liverpool: Springer, 2017: 759-770.

[9] CHEN Dandan, ZHANG Yong, GAO Liangpeng, et al. The impact of rainfall on the temporal and spatial distribution of taxi passengers [J]. PLoS One, 2017, 12(9): e0183574.

[10] KONG Xiangjie, XIA Feng, WANG Jinzhong, et al. Time-location-relationship combined service recommendation based on taxi trajectory data[J]. IEEE Transactions on Industrial Informatics, 2017, 13(3): 1202-1212.

[11] RONG Huigui, WANG Zepeng, ZHENG Hui, et al. Mining efficient taxi operation strategies from large scale geo-location data[J]. IEEE Access, 2017(5): 25623-25634.

[12] WANG Ran, CHOW C Y, LYU Y, et al. TaxiRec: recommending road clusters to taxi drivers using ranking-based extreme learning machines [J]. IEEE Transactions on Knowledge and Data Engineering, 2018, 30(3): 585-598.

[13] PENG Chengbin, JIN Xiaogang, WONG K C, et al. Collective human mobility pattern from taxi trips in urban area[J]. PLoS One, 2012, 7(4): e34487.

[14] ZHANG Wangsheng, LI Shijian, PAN Gang. Mining the semantics of origin-destination flows using taxi traces[C]//Proceedings of 2012 ACM Conference on Ubiquitous Computing. Pittsburgh: ACM, 2012.

- [15] TANG Jinjun, LIU Fang, WANG Yin Hai, et al. Uncovering urban human mobility from large scale taxi GPS data[J]. *Physica A: Statistical Mechanics and Its Applications*, 2015(438): 140-153.
- [16] 付鑫, 孙茂棚, 孙皓. 基于 GPS 数据的出租车通勤识别及时空特征分析[J]. *中国公路学报*, 2017, 30(7): 134-143. FU Xin, SUN Maopeng, SUN Hao. Taxi commute recognition and temporal-spatial characteristics analysis based on GPS data[J]. *China Journal of Highway and Transport*, 2017, 30(7): 134-143.
- [17] SCHOLZ R W, LU Yongmei. Detection of dynamic activity patterns at a collective level from large-volume trajectory data[J]. *International Journal of Geographical Information Science*, 2014, 28(5): 946-963.
- [18] MOREIRA-MATIAS L, GAMA J, FERREIRA M, et al. Time-evolving O-D matrix estimation using high-speed GPS data streams[J]. *Expert Systems with Applications*, 2016(44): 275-288.
- [19] ZHANG Fuzheng, YUAN N J, WILKIE D, et al. Sensing the pulse of urban refueling behavior: a perspective from taxi mobility [J]. *ACM Transactions on Intelligent Systems and Technology*, 2015, 6(3): 37.
- [20] LUO Xiao, DONG Liang, DOU Yi, et al. Analysis on spatial-temporal features of taxis' emissions from big data informed travel patterns: a case of Shanghai, China[J]. *Journal of Cleaner Production*, 2017(142): 926-935.
- [21] 赵旻玥. 基于出租车轨迹数据分析交叉口尾气排放的时空特征[D]. 上海: 华东师范大学, 2013. ZHAO Minyue. Analyze spatial and temporal characteristics of exhaust emissions around intersections based on taxi trajectory data[D]. Shanghai: East China Normal University, 2013.
- [22] SHANG Jingbo, ZHENG Yu, TONG Wenzhu, et al. Inferring gas consumption and pollution emission of vehicles throughout a city [C] // *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM, 2014.
- [23] ZHAO Kai, MUSOLESI M, HUI Pan, et al. Explaining the power-law distribution of human mobility through transportation modality decomposition [J]. *Scientific Reports*, 2015(5): 9136.
- [24] ZHENG Zhong, RASOULI S, TIMMERMANS H. Two-regime pattern in human mobility: evidence from GPS taxi trajectory data[J]. *Geographical Analysis*, 2016, 48(2): 157-175.
- [25] TANG Luliang, SUN Fei, KAN Zihan, et al. Uncovering distribution patterns of high performance taxis from big trace data[J]. *ISPRS International Journal of Geo-Information*, 2017, 6(5): 134.
- [26] 吴佩莉, 刘奎恩, 郝身刚, 等. 基于浮动车数据的快速交通拥堵监控[J]. *计算机研究与发展*, 2014, 51(1): 189-198. WU Peili, LIU Kuien, HAO Shengang, et al. Rapid traffic congestion monitoring based on floating car data[J]. *Journal of Computer Research and Development*, 2014, 51(1): 189-198.
- [27] GE Yong, XIONG Hui, LIU Chuanren, et al. A taxi driving fraud detection system[C] // *Proceedings of the 2011 IEEE 11th International Conference on Data Mining*. Vancouver: IEEE, 2011.
- [28] KUANG Weiming, AN Shi, JIANG Huifu. Detecting traffic anomalies in urban areas using taxi GPS data[J]. *Mathematical Problems in Engineering*, 2015(2015): 809582.
- [29] 刘兴权, 金美含. 地图匹配算法综述[J]. *科技信息*, 2014(4): 64-65. LIU Xingquan, JIN Meihan. A review on map matching algorithm[J]. *Science & Technology Information*, 2014(4): 64-65.
- [30] ZHANG Lingyu, HU Tao, MIN Yue, et al. A taxi order dispatch model based on combinatorial optimization[C] // *the 23rd ACM SIGKDD International Conference*. ACM, 2017.
- [31] TANG Jinjun, JIANG Han, LI Zhibin, et al. A two-layer model for taxi customer searching behaviors using GPS trajectory data [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2016, 17(11): 3318-3324.
- [32] YANG Hai, YE Min, TANG W H, et al. Regulating taxi services in the presence of congestion externality [J]. *Transportation Research Part A: Policy and Practice*, 2005, 39(1): 17-40.
- [33] ZHAO Kai, CHINNASAMY M P, TARKOMA S. Automatic city region analysis for urban routing[C] // *Proceedings of 2015 IEEE International Conference on Data Mining Workshop*. Atlantic City: IEEE, 2015.
- [34] PAN Gang, QI Guande, WU Zhaohui, et al. Land-use classification using taxi GPS traces[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2013, 14(1): 113-123.
- [35] ZHANG Daqiang, WAN Jiafu, HE Zongjian, et al. Identifying region-wide functions using urban taxicab trajectories[J]. *ACM Transactions on Embedded Computing Systems*, 2016, 15(2): 36.
- [36] LEE W C, KRUMM J. Trajectory preprocessing[M] // ZHENG Yu, ZHOU Xiaofang. *Computing with Spatial Trajectories*. New York: Springer, 2011.
- [37] GUSTAFSSON F. Particle filter theory and practice with positioning applications [J]. *IEEE Aerospace and Electronic Systems Magazine*, 2010, 25(7): 53-82.
- [38] QUDDUS M A, OCHIENG W Y, ZHAO LIN, et al. A general map matching algorithm for transport telematics applications[J]. *GPS Solutions*, 2003, 7(3): 157-167.
- [39] 许志海, 张昭云, 陈应东. 车辆导航系统中定位数据处理和地图匹配技术[J]. *测绘学院学报*, 2003, 20(4): 247-250. XU Zhihai, ZHANG Zhaoyun, CHEN Yingdong. Technique of positioning data preceding and map matching in naviga-



- tion system for vehicle [J]. *Journal of Institute of Surveying and Mapping*, 2003, 20(4): 247-250.
- [40] QUDDUS M A. High integrity map matching algorithms for advanced transport telematics applications [D]. London: Imperial College London, 2006.
- [41] 常菲, 浦争艳, 李明禄, 等. 综合地图匹配定位技术研究 [J]. *计算机工程与应用*, 2004, 40(19): 200-202.  
CHANG Fei, PU Zhengyan, LI Minglu, et al. Study on integrated map matching technology [J]. *Computer Engineering and Applications*, 2004, 40(19): 200-202.
- [42] 安实, 匡伟明. 出租车 GPS 数据轨迹化方法研究 [J]. *科学技术与工程*, 2015, 15(11): 125-130.  
AN Shi, KUANG Weiming. Research on taxi GPS trajectory composing method [J]. *Science Technology and Engineering*, 2015, 15(11): 125-130.
- [43] 陈皓, 张磊. 基于大数据的出租车与普通小汽车行驶速度差异研究 [J]. *公路*, 2017, 62(2): 147-152.  
CHEN Hao, ZHANG Lei. Research on the difference of taxi and car speed based on big data [J]. *Highway*, 2017, 62(2): 147-152.
- [44] 王辉. 考虑出租车影响的城市道路交通流建模与仿真 [D]. 兰州: 兰州交通大学, 2018.  
WANG Hui. Modeling and simulation of the influence of taxi on urban road traffic flow [D]. Lanzhou: Lanzhou Jiaotong University, 2018.
- [45] YUAN JING, ZHENG Yu, ZHANG Chengyang, et al. T-drive: driving directions based on taxi trajectories [C] // Proceedings of the 18th Sigsatial International Conference on Advances in Geographic Information Systems. San Jose: ACM, 2010.
- [46] NYC. Taxi & limousine commission (TLC) [EB/OL]. [2019-05-23]. <https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page>.
- [47] KOTZ D, HENDERSON T, CRAWDAD; A community resource for archiving wireless data at dartmouth [J]. *IEEE Pervasive Computing*, 2005, 4(4):12-14.
- [48] DANIEL S, ZHANG Kaisheng, SHEN Suwan. Analyzing spatiotemporal traffic line source emissions based on massive didi online car-hailing service data [J]. *Transportation Research Part D Transport and Environment*, 2018, 62: 699-714.
- [49] KOBAYASHI T, SHINAGAWA N, WATANABE Y. Vehicle mobility characterization based on measurements and its application to cellular communication systems [J]. *IEICE Transactions on Communications*, 1999, E82-B(12): 2055-2060.
- [50] SAITOH K, HIDAKA H, SHINAGAWA N, et al. Vehicle motion in large and small cities and teletraffic characterization in cellular communication systems [J]. *IEICE Transactions on Communications*, 2001, E84-B(4): 805-813.
- [51] CHEU R L, XIE Chi, LEE D H. Probe vehicle population and sample size for arterial speed estimation [J]. *Computer-Aided Civil and Infrastructure Engineering*, 2002, 17(1): 53-60.
- [52] LIAO Z. Taxi dispatching via global positioning systems [J]. *IEEE Transactions on Engineering Management*, 2001, 48(3): 342-347.
- [53] KETZEL M, WÄHLIN P, BERKOWICZ R, et al. Particle and trace gas emission factors under urban driving conditions in Copenhagen based on street and roof-level observations [J]. *Atmospheric Environment*, 2003, 37(20): 2735-2749.
- [54] ZHENG Yu, LIU Yanchi, YUAN Jing, et al. Urban computing with taxicabs [C] // Proceedings of the 13th International Conference on Ubiquitous Computing. Beijing: ACM, 2011.
- [55] YUAN Jing, ZHENG Yu, 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: ACM, 2012.
- [56] LIU Yu, WANG Fahui, XIAO Yu, et al. Urban land uses and traffic "source-sink areas": evidence from GPS-enabled taxi data in Shanghai [J]. *Landscape and Urban Planning*, 2012, 106(1): 73-87.
- [57] LIANG Xiao, ZHENG Xudong, LV Weifeng, et al. The scaling of human mobility by taxis is exponential [J]. *Physica A: Statistical Mechanics and its Applications*, 2012, 391(5): 2135-2144.
- [58] VELOSO M, PHITHAKKITNUKON S, BENTO C, et al. Exploratory study of urban flow using taxi traces [C] // Proceedings of the 1st Workshop on Pervasive Urban Applications (PURBA) in Conjunction with the Ninth International Conference on Pervasive Computing 2011. San Francisco: [s.n.], 2011.
- [59] 朱金清. 基于滴滴轨迹大数据的智慧交通控制 [EB/OL]. (2017-10-25) [2019-05-23]. <http://www.caai.cn/index.php?s=/home/article/detail/id/397.html>.  
ZHU Jinqing. Intelligent traffic control based on DiDi trajectory big data [EB/OL]. (2017-10-25) [2019-05-23]. <http://www.caai.cn/index.php?s=/home/article/detail/id/397.html>.
- [60] CASTRO P S, ZHANG Daqing, CHEN Chao, et al. From taxi GPS traces to social and community dynamics: a survey [J]. *ACM Computing Surveys*, 2013, 46(2): 17.
- [61] ZHENG Yu, CAPRA L, WOLFSON O, et al. Urban computing: concepts, methodologies, and applications [J]. *ACM Transactions on Intelligent Systems and Technology*, 2014, 5(3): 38.
- [62] 张红, 王晓明, 过秀成, 等. 出租车 GPS 轨迹大数据在智能交通中的应用 [J]. *兰州理工大学学报*, 2016, 42(1): 109-114.  
ZHANG Hong, WANG Xiaoming, GUO Xiucheng, et al. Application of taxi GPS big trajectory data in intelligent traffic system [J]. *Journal of Lanzhou University of Tech-*

- nology, 2016, 42(1): 109-114.
- [63] 王远飞, 何洪林. 空间数据分析方法[M]. 北京: 科学出版社, 2007.
- WANG Yuanfei, HE Honglin. Spatial data analysis method[M]. Beijing: Science Press, 2007.
- [64] LIU Liang, ANDRIS C, RATTI C. Uncovering cabdrivers' behavior patterns from their digital traces[J]. Computers, Environment and Urban Systems, 2010, 34(6): 541-548.
- [65] WANG Wenjun, PAN Lin, YUAN Ning, et al. A comparative analysis of intra-city human mobility by taxi[J]. Physica A: Statistical Mechanics and its Applications, 2015(420): 134-147.
- [66] ZHANG Daqing, SUN Lin, LI Bin, et al. Understanding taxi service strategies from taxi GPS traces[J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(1): 123-135.
- [67] SU Rongxiang, FANG Zhixiang, LUO Ningxin, et al. Understanding the dynamics of the pick-up and drop-off locations of taxicabs in the context of a subsidy war among e-hailing apps[J]. Sustainability, 2018, 10(4): 1256.
- [68] HWANG R H, HSUEH Y L, CHEN Yuting. An effective taxi recommender system based on a spatio-temporal factor analysis model [J]. Information Sciences, 2015 (314): 28-40.
- [69] ZHANG Sihai, WANG Zhiyang. Inferring passenger denial behavior of taxi drivers from large-scale taxi traces[J]. PLoS One, 2017, 12(2): e0171876.
- [70] WANG Hongwei, PENG Zhongren, LU Qingchang, et al. Assessing effects of bus service quality on passengers' taxi-hiring behavior [J]. Transport, 2018, 33 (4): 1030-1044.
- [71] NIE Yu. How can the taxi industry survive the tide of ridesourcing? Evidence from Shenzhen, China[J]. Transportation Research Part C: Emerging Technologies, 2017(79): 242-256.
- [72] JIANG Weiwei, ZHANG Lin. The impact of the transportation network companies on the taxi industry: evidence from Beijing's GPS taxi trajectory data[J]. IEEE Access, 2018(6): 12438-12450.
- [73] SHI W, LIU Y. Real-time urban traffic monitoring with global positioning system-equipped vehicles[J]. IET Intelligent Transport Systems, 2010, 4(2): 113-120.
- [74] OMER M, MOSTASHARI A, NILCHIANI R. Assessing resilience in a regional road-based transportation network [J]. International Journal of Industrial and Systems Engineering, 2013, 13(4): 389.
- [75] KONG Xiangfu, YANG Jiawen, YANG Zhongyu. Measuring traffic congestion with taxi GPS data and travel time index [C]//Proceedings of the 15th COTA International Conference of Transportation Professionals. Beijing: [s.n.], 2015.
- [76] 童晓君. 基于出租车 GPS 数据的居民出行行为分析[D]. 长沙: 中南大学, 2012.
- TONG Xiaojun. Analysis of residents' behavior based on the taxi GPS data[D]. Changsha: Central South University, 2012.
- [77] QI Guande, LI Xiaolong, LI Shijian, et al. Measuring social functions of city regions from large-scale taxi behaviors[C]//Proceedings of 2011 IEEE International Conference on Pervasive Computing and Communications Workshops. Seattle: IEEE, 2011.
- [78] 陈晨, 赵晓华, 姚莹, 等. 基于出租车运行数据的北京城市快速路油耗分布特征及影响因素分析[J]. 北京工业大学学报, 2017, 43(12): 1900-1906.
- CHEN Chen, ZHAO Xiaohua, YAO Ying, et al. Analysis of fuel consumption distribution characteristics and influencing factors of Beijing urban expressway based on taxi operation data[J]. Journal of Beijing University of Technology, 2017, 43(12): 1900-1906.
- [79] ZHANG Mingyue, LIU Jianxun, LIU Yizhi, et al. Recommending pick-up points for taxi-drivers based on spatio-temporal clustering[C]//Proceedings of the 2012 2nd International Conference on Cloud and Green Computing. Xiangtan: IEEE, 2012: 67-72.
- [80] 李衢伶. 基于 GPS 轨迹的出租车载客路径智能推荐[D]. 湘潭: 湖南科技大学, 2014.
- LI Quling. Intelligent GPS-based taxi passenger path trajectory recommended[D]. Xiangtan: Hunan University of Science and Technology, 2014.
- [81] LIU Yizhi, LIU Jianxun, WANG Jianjun, et al. Recommending a personalized sequence of pick-up points[C]//Proceedings of the 10th Asia-Pacific Services Computing Conference. Zhangjiajie: Springer, 2016: 278-291.
- [82] 齐林. 基于 GPS 数据的出租车交通运行特性研究及应用 [D]. 哈尔滨: 哈尔滨工业大学, 2013.
- QI Lin. Research and application on taxi traffic running performance based on GPS data [D]. Harbin: Harbin Institute of Technology, 2013.
- [83] LIU Xintao, BAN Yifang. Uncovering spatio-temporal cluster patterns using massive floating car data[J]. ISPRS International Journal of Geo-Information, 2013, 2 (2): 371-384.
- [84] ZHANG Kaisheng, SUN D, SHEN Suwan, et al. Analyzing spatiotemporal congestion pattern on urban roads based on taxi GPS data [J]. The Journal of Transport and Land Use, 2017, 10(1): 675-694.
- [85] 唐炉亮, 牛乐, 杨雪, 等. 利用轨迹大数据进行城市道路交叉口识别及结构提取[J]. 测绘学报, 2017, 46(6): 770-779. DOI: 10.11947/j.AGCS.2017.20160614.
- TANG Luliang, NIU Le, YANG Xue, et al. Urban intersection recognition and construction based on big trace data[J]. Acta Geodaetica et Cartographica Sinica, 2017, 46(6): 770-779. DOI: 10.11947/j.AGCS.2017.20160614.
- [86] 江慧娟, 余洋. 出租车载客热点精细提取的改进 DBSCAN 算法[J]. 地理空间信息, 2017, 15(10): 16-20.

- JIANG Huijuan, YU Yang. Extracting fine pickup hot zone by an improved DBSCAN algorithm[J]. *Geospatial Information*, 2017, 15(10): 16-20.
- [87] SHEN Ying, ZHAO Ligang, FAN Jing. Analysis and visualization for hot spot based route recommendation using short-dated taxi GPS traces[J]. *Information*, 2015, 6(2): 134-151.
- [88] XU Xiaolong, DOU Wanchun, ZHANG Xuyun, et al. A traffic hotline discovery method over cloud of things using big taxi GPS data[J]. *Software: Practice and Experience*, 2017, 47(3): 361-377.
- [89] YANG Ci, GONZALES E J. Modeling taxi demand and supply in new york city using large-scale taxi GPS data[M]// THAKURIAH P, TILAHUN N, ZELLNER M. *Seeing Cities Through Big Data: Research, Methods and Applications in Urban Informatics*. Cham: Springer, 2017.
- [90] LI Xiaolong, PAN Gang, WU Zhaohui, et al. Prediction of urban human mobility using large-scale taxi traces and its applications[J]. *Frontiers of Computer Science*, 2012, 6(1): 111-121.
- [91] ZHENG Xudong, LIANG Xiao, XU Ke. Where to wait for a taxi? [C]// *Proceedings of ACM SIGKDD International Workshop on Urban Computing*. Beijing: ACM, 2012.
- [92] 于扬红雪. 出租车需求量预测模型的研究[D]. 大连:大连海事大学, 2017.
- YU Yanghongxue. *Research on Taxi Demand Forecasting Model*[D]. Dalian: Dalian Maritime University, 2017.
- [93] QIAN Xinwu, UKKUSURI S V. Spatial variation of the urban taxi ridership using GPS data[J]. *Applied Geography*, 2015(59): 31-42.
- [94] ASAMER J, REINTHALER M, RUTHMAIR M, et al. Optimizing charging station locations for urban taxi providers[J]. *Transportation Research Part A: Policy and Practice*, 2016(85): 233-246.
- [95] YANG Jie, DONG Jing, HU Liang. A data-driven optimization-based approach for siting and sizing of electric taxi charging stations[J]. *Transportation Research Part C: Emerging Technologies*, 2017(77): 462-477.
- [96] SHEN Jingwei, LIU Xintao, CHEN Min. Discovering spatial and temporal patterns from taxi-based floating car data: a case study from Nanjing [J]. *GIScience & Remote Sensing*, 2017, 54(5): 617-638.
- [97] HU Yujie, MILLER H J, LI Xiang. Detecting and analyzing mobility hotspots using surface networks[J]. *Transactions in GIS*, 2014, 18(6): 911-935.
- [98] 刘汇慧, 阚子涵, 孙飞, 等. 采用轨迹大数据探测短时非营运行为[J]. *武汉大学学报(信息科学版)*, 2016, 41(9): 1192-1198.
- LIU Huihui, KAN Zihan, SUN Fei, et al. Taxis' short-term out-of-service behaviors detection using big trace data[J]. *Geomatics and Information Science of Wuhan University*, 2016, 41(9): 1192-1198.
- [99] CUI Jianxun, LIU Feng, HU Jia, et al. Identifying mismatch between urban travel demand and transport network services using GPS data: a case study in the fast growing Chinese city of Harbin[J]. *Neurocomputing*, 2016(181): 4-18.
- [100] ZHANG Desheng, HE Tian, LIN Shan, et al. Taxi-passenger-demand modeling based on big data from a roving sensor network [J]. *IEEE Transactions on Big Data*, 2017, 3(3): 362-374.
- [101] 赵东保, 刘雪梅, 张弘弢. 基于大规模浮动车轨迹点数据的道路网变化检测与更新方法研究[J]. *地理与地理信息科学*, 2016, 32(2): 1-5.
- ZHAO Dongbao, LIU Xuemei, ZHANG Hongtao. Automated change detection and update of vector road network based on large scale floating car data[J]. *Geography and Geo-Information Science*, 2016, 32(2): 1-5.
- [102] PANG L X, CHAWLA S, LIU Wei, et al. On detection of emerging anomalous traffic patterns using GPS data[J]. *Data & Knowledge Engineering*, 2013(87): 357-373.
- [103] 周洋, 方志祥, 李清泉, 等. 利用经验约束规则和证据理论进行出租车异常轨迹检测[J]. *武汉大学学报(信息科学版)*, 2016, 41(6): 797-802.
- ZHOU Yang, FANG Zhixiang, LI Qingquan, et al. Anomalous taxi trajectory detection based on experiential constraint rules and evidence theory[J]. *Geomatics and Information Science of Wuhan University*, 2016, 41(6): 797-802.
- [104] XU Xiujuan, ZHOU Jianyu, LIU Yu, et al. Taxi-RS: taxi-hunting recommendation system based on taxi GPS data [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(4): 1716-1727.
- [105] LIU Jin, YU Xiao, XU Zheng, et al. A cloud-based taxi trace mining framework for smart city [J]. *Software: Practice and Experience*, 2017, 47(8): 1081-1094.
- [106] LUO Wuman, TAN Haoyu, CHEN Lei, et al. Finding time period-based most frequent path in big trajectory data [C]// *Proceedings of 2013 ACM SIGMOD International Conference on Management of Data*. New York: ACM, 2013.
- [107] WANG Jing, WANG Chaoliang, SONG Xianfeng, et al. Automatic intersection and traffic rule detection by mining motor-vehicle GPS trajectories[J]. *Computers, Environment and Urban Systems*, 2017(64): 19-29.
- [108] 杨伟, 艾廷华. 基于车辆轨迹大数据的道路网更新方法研究[J]. *计算机研究与发展*, 2016, 53(12): 2681-2693.
- YANG Wei, AI Tinghua. A method for road network updating based on vehicle trajectory big data [J]. *Journal of Computer Research and Development*, 2016, 53(12): 2681-2693.
- [109] 张健钦, 仇培元, 杜明义. 基于时空轨迹数据的出行特征挖掘方法[J]. *交通运输系统工程与信息*, 2014, 14(6): 72-78.
- ZHANG Jianqin, QIU Peiyuan, DU Mingyi. Mining



- method of travel characteristics based on spatio-temporal trajectory data [J]. *Journal of Transportation Systems Engineering and Information Technology*, 2014, 14(6): 72-78.
- [110] DENG B, DENMAN S, ZACHARIADIS V, et al. Estimating traffic delays and network speeds from low-frequency GPS taxis traces for urban transport modelling [J]. *European Journal of Transport and Infrastructure Research*, 2015, 15(4): 639-661.
- [111] ZHANG Jindong, MENG Weibin, LIU Qiangqiang, et al. Efficient vehicles path planning algorithm based on taxi GPS big data [J]. *Optik*, 2016, 127(5): 2579-2585.
- [112] 林娜, 郑亚男. 基于出租车轨迹数据的路径规划方法 [J]. *计算机应用与软件*, 2016, 33(1): 68-72.
- LIN Na, ZHENG Yanan. Path planning method based on taxi tracks data [J]. *Computer Applications and Software*, 2016, 33(1): 68-72.
- [113] 戚威, 梁伟涛, 马勇. 基于出租车轨迹数据的最优路径规划方法 [J]. *计算机应用*, 2017, 37(7): 2106-2113.
- QI Xin, LIANG Weitao, MA Yong. Optimal path planning method based on taxi trajectory data [J]. *Journal of Computer Applications*, 2017, 37(7): 2106-2113.
- [114] FENG Huifang, XU Youji. An empirical study on evolution of the connectivity for VANETs based on taxi GPS traces [J]. *International Journal of Distributed Sensor Networks*, 2016(2016): 2580465.
- [115] WU Guojun, DING Yichen, LI Yanhua, et al. Mining spatio-temporal reachable regions over massive trajectory data [C] // *Proceedings of the 2017 IEEE 33rd International Conference on Data Engineering*. San Diego: IEEE, 2017: 1283-1294.
- [116] CUI Jianxun, LIU Feng, JANSSENS D, et al. City-wide examining transport network accessibility using taxi GPS data [C] // *Proceedings of the 15th COTA International Conference of Transportation Professionals*. Beijing: ASCE, 2015.
- [117] 周洋. 基于出租车数据的城市居民活动空间与网络时空特性研究 [D]. 武汉: 武汉大学, 2016.
- ZHOU Yang. Study on human activity space patterns and network spatial temporal characteristics in urban cities using taxi trajectory data [D]. Wuhan: Wuhan University, 2016.
- [118] LIU Xi, GONG Li, GONG Yongxi, et al. Revealing travel patterns and city structure with taxi trip data [J]. *Journal of Transport Geography*, 2015(43): 78-90.
- [119] 赵鹏祥. 基于轨迹聚类的城市热点区域提取与分析方法研究 [D]. 武汉: 武汉大学, 2015.
- ZHAO Pengxiang. Research on the method of extracting and analyzing urban hotspots based on trajectory clustering [D]. Wuhan: Wuhan University, 2015.
- [120] GAO Song, WANG Yaoli, GAO Yong, et al. Understanding urban traffic-flow characteristics: a rethinking of betweenness centrality [J]. *Environment and Planning B: Urban Analytics and City Science*, 2013, 40(1): 135-153.
- [121] 周志华. 机器学习 [M]. 北京: 清华大学出版社, 2016.
- ZHOU Zhihua. *Machine learning* [M]. Beijing: Tsinghua University Press, 2016.
- [122] 林永杰, 邹难. 基于运营系统的出租车出行需求短时预测模型 [J]. *东北大学学报(自然科学版)*, 2016, 37(9): 1235-1240.
- LIN Yongjie, ZOU Nan. Short-term prediction model of taxi passenger demand based on operation systems [J]. *Journal of Northeastern University (Natural Science)*, 2016, 37(9): 1235-1240.
- [123] XU Jun, RAHMATIZADEH R, BÖLÖNI L, et al. Real-time prediction of taxi demand using recurrent neural networks [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2018, 19(8): 2572-2581.
- [124] MA Xiaolei, YU Haiyang, WANG Yunpeng, et al. Large-scale transportation network congestion evolution prediction using deep learning theory [J]. *PLoS One*, 2015, 10(3): e0119044.
- [125] PHITHAKKITNUKON S, VELOSO M, BENTO C, et al. Taxi-aware map: identifying and predicting vacant taxis in the city [C] // *Proceedings of the 1st International Joint Conference on Ambient Intelligence*. Malaga: Springer, 2010.
- [126] GONG Li, LIU Xi, WU Lun, et al. Inferring trip purposes and uncovering travel patterns from taxi trajectory data [J]. *Cartography and Geographic Information Science*, 2015, 43(2): 103-114.
- [127] ZHANG Yingjie, Li Beibei, RAMAYYA K. Learning individual behavior using sensor data: the case of GPS traces and taxi drivers [EB/OL]. [2019-05-26]. <https://ssrn.com/abstract=2779328> or <http://dx.doi.org/10.2139/ssrn.2779328>.
- [128] LI Bin, ZHANG Daqing, SUN Lin, et al. Hunting or waiting? Discovering passenger-finding strategies from a large-scale real-world taxi dataset [C] // *Proceedings of 2011 IEEE International Conference on Pervasive Computing and Communications Workshops*. Seattle: IEEE, 2011.
- [129] 梁军辉, 林坚, 杜洋. 大数据条件下城市用地类型辨识研究——基于出租车 GPS 数据的动态感知 [J]. *上海国土资源*, 2016, 37(1): 28-32.
- LIANG Junhui, LIN Jian, DU Yang. Research on the identification of urban land use under the big data condition: dynamic awareness based on taxi GPS data [J]. *Shanghai Land & Resources*, 2016, 37(1): 28-32.
- [130] YAO Baozhen, CHEN Chao, CAO Qingda, et al. Short-term traffic speed prediction for an urban corridor [J]. *Computer-Aided Civil and Infrastructure Engineering*, 2017, 32(2): 154-169.
- [131] TU Wei, LI Qingquan, FANG Zhixiang, et al. Optimizing the locations of electric taxi charging stations: a spatial-temporal demand coverage approach [J]. *Transportation*

- Research Part C: Emerging Technologies, 2016 (65): 172-189.
- [132] 温雅静. 基于热点载客区域的出租车应急调度方案研究[D]. 北京: 北京交通大学, 2014.
- WEN Yajing. Study on taxi emergency dispatch options based on taxi pick-up hotspots [D]. Beijing: Beijing Jiaotong University, 2014.
- [133] ZHANG Wangsheng, QI Guande, PAN Gang, et al. City-scale social event detection and evaluation with taxi traces [J]. ACM Transactions on Intelligent Systems and Technology, 2015, 6(3): 40.
- [134] CASTRO P S, ZHANG Daqing, LI Shijian. Urban traffic modelling and prediction using large scale taxi GPS traces [C]//Proceedings of the 10th International Conference on Pervasive Computing. Newcastle: Springer, 2012.
- [135] 韩博洋, 汪兆洋, 金蓓弘. 一种基于轨迹大数据离线挖掘与在线实时监测的出租车异常轨迹检测算法[J]. 中国科学技术大学学报, 2016, 46(3): 247-252.
- HAN Boyang, WANG Zhaoyang, JIN Beihong. An anomaly detection algorithm for taxis based on trajectory data mining and online real-time monitoring[J]. Journal of University of Science and Technology of China, 2016, 46(3): 247-252.
- [136] CHAO C, ZHANG D, CASTRO P S, et al. Real-Time Detection of Anomalous Taxi Trajectories from GPS Traces [C]// Mobile and Ubiquitous Systems: Computing, Networking, and Services. Berlin, Heidelberg: Springer-Verlag, 2011:63-74.
- [137] 吴涛, 向隆刚, 龚健雅. 路网更新的轨迹-地图匹配方法[J]. 测绘学报, 2017, 46(4): 507-515. DOI: 10.11947/j. AGCS.2017.20150479.
- WU Tao, XIANG Longgang, GONG Jianya. Renewal of road networks using map-matching technique of trajectories [J]. Acta Geodaetica et Cartographica Sinica, 2017, 46(4): 507-515. DOI: 10.11947/j. AGCS.2017.20150479.
- [138] 胡瀚, 向隆刚, 王德浩. 出租车轨迹数据的道路提取[J]. 测绘通报, 2018(7): 53-57. DOI: 10.13474/j. cnki. 11-2246.2018.0209.
- HU Han, XIANG Longgang, WANG Dehao. Road extraction based on taxi trajectory data[J]. Bulletin of Surveying and Mapping, 2018(7): 53-57. DOI: 10.13474/j. cnki. 11-2246.2018.0209.
- [139] BARBOSA H, BARTHELEMY M, GHOSHAL G, et al. Human mobility: models and applications [J]. Physics Reports, 2018(734): 1-74.
- [140] WU Huayi, YOU Lan, GUI Zhipeng, et al. GeoSquare: collaborative geoprocessing models' building, execution and sharing on Azure Cloud[J]. Annals of GIS, 2015, 21(4): 287-300.
- [141] YU Bailang, LIAN Ting, HUANG Yixiu, et al. Integration of nighttime light remote sensing images and taxi GPS tracking data for population surface enhancement [J]. International Journal of Geographical Information Science, 2019, 33(4): 687-706.

(责任编辑:张艳玲)

收稿日期: 2019-05-27

修回日期: 2019-08-30

第一作者简介: 吴华意(1966—),男,教授,研究方向为地理信息分析与挖掘。

First author: WU Huayi(1966—), male, professor, majors in geographic information analysis and mining.

E-mail: wuhuayi@whu.edu.cn

通信作者: 游兰

Corresponding author: YOU Lan

E-mail: yoyo@hubu.edu.cn