

DOI:10.16799/j.cnki.esdqyh.2022.08.051

出租车轨迹数据在综合交通调查中的应用方法

代位

[上海市政工程设计研究总院(集团)有限公司城乡规划局,上海市 200092]

摘要:出租车轨迹数据是综合交通调查中重要的信息化数据来源,能应用在包括出租车调查、车速调查等多项交通调查工作中。系统梳理了出租车轨迹数据在综合交通调查中的应用场景、应用要求,设计了出租车轨迹数据的分析框架和处理流程,重点介绍了“相邻定位点最短路径匹配”的地图匹配计算方法。以某市为例,介绍了出租车轨迹数据在该市综合交通调查中的应用和现状特征结论,相关分析起到了完成交通调查、支持交通规划任务的效果。

关键词:综合交通调查;出租车;轨迹数据;地图匹配
中图分类号: U491.1+1 **文献标志码:** B

文章编号: 1009-7716(2022)08-0190-04

0 引言

城市综合交通调查是为明确城市交通需求和运行特征而进行的调查,为城市综合交通体系规划编制、交通模型标定与校验等工作提供基础数据,包含居民出行调查等9项内容^[1-2]。具备条件的城市可采用信息化数据对调查项目和内容进行适当调整,信息化数据包括视频或检测数据、车辆定位数据、移动通信数据等。

出租车轨迹数据是信息化数据的一种,属于车辆定位数据,具有以下几个特点:一是分布广泛,出租车在城市各处流动;二是采集连续,出租车在运营期间不间断返回定位数据;三是应用面广,其数据挖掘结构既能反映出租车运营特征,也可反映道路交通运行情况。因此,出租车轨迹数据在交通调查中应用价值价高。

当前对出租车轨迹数据的研究主要集中在地图匹配方面,大量文献研究了车辆轨迹点与路段匹配的精确算法^[3],包括几何方法^[4]、隐马尔可夫模型^[5]以及其他方法^[6-7],为出租车轨迹数据的处理方法积累了大量的经验。近年一些研究倾向于挖掘出租车轨迹数据所包含的居民出行^[8-10]、通勤活动^[11]、路网状态^[12]等信息。但缺乏对出租车轨迹数据在交通调查中应用的梳理。

本文从综合交通调查的应用角度出发,梳理出租车轨迹数据在交通调查中的应用场景及要求,分

析数据处理框架,并涉及适宜交通规划项目的数据处理方法,最后以实际项目为例介绍应用情况。

1 基于交通调查场景的出租车轨迹数据处理方法

1.1 出租车轨迹数据的交通调查应用场景分析

根据《城市综合交通调查技术标准》(GB/T 51334—2018)^[1](下称《标准》)和《城市综合交通体系规划交通调查导则》^[2](下称《导则》),可应用到出租车轨迹数据的交通调查主要有:出租车调查(《导则》中为“商用车调查”出租部分)、居民出行调查、道路交通调查、出行生成源调查。各项调查内容见表1。

表1 出租车轨迹数据在交通调查中的应用场景及要求

调查名称	应用场景	应用要求	数据性质
出租车调查	计算出租车运营里程、载客里程、载客次数、每次载客上下车地点和时间	抽样率根据出租车总数量取5%~10%,且不低于300辆	主要数据
道路交通调查	将载客情况下的出租车视为浮动车,计算道路各时段的平均行程车速	应包含全日高峰时段,必要时包含连续的12h或24h	主要数据
出行生成源调查	校核问卷调查中以出租车方式到离生成源的时间和空间分布	宜包含出行高峰时段的连续12h	补充数据
居民出行调查	校核问卷调查中以出租车方式出行的次数、时耗、距离、时空分布,及出租车分担率、OD分布	样本应分布均匀	补充数据

1.2 出租车轨迹数据处理分析框架

1.2.1 数据结构及处理框架

出租车轨迹数据至少应包含定位时间、经纬度、

收稿日期:2021-11-22

作者简介:代位(1995—),男,硕士,助理工程师,从事交通规划工作。

载客状态等字段,典型的轨迹数据字段及对应的数据类型如表2所示。其中定位时间、经纬度确定出租车辆的时空位置,载客状态的变化判断出租车辆订单转换,行驶速度或距离确定车辆速度。

表2 出租车轨迹数据在交通调查中的应用场景及要求

字段	含义	数据类型	字段	含义	数据类型
id	编号	整型	lon	定位经度	浮点型
time	时间	时间戳	lat	定位纬度	浮点型
status	载客状态	布尔型	speed	瞬时速度, km/h	浮点型
location	定位状态	布尔型	distance	行驶里程, m	浮点型

出租车轨迹数据在交通调查应用中的处理流程包括数据预处理、数据处理、数据分析三个步骤,分析框架见图1。

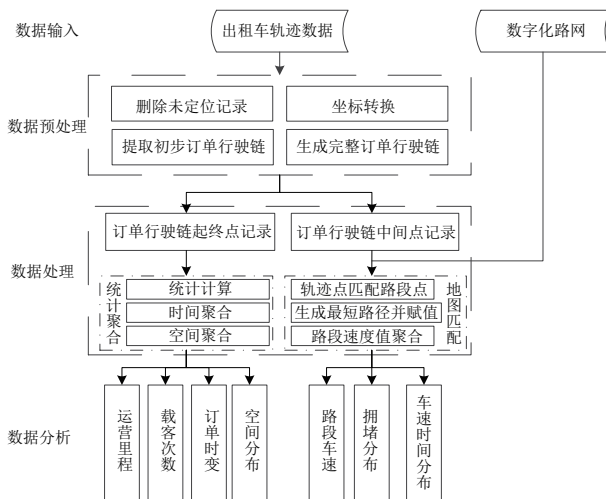


图1 出租车轨迹数据处理框架

1.2.2 数据预处理方法

出租车轨迹的原始数据会存在几个问题:

一是未定位的数据:往往产生于一辆出租车启动或关闭定位功能的短时间内,这段时间出租车信号很弱,无法获得准确坐标,定位状态字段[LOCATION]显示FALSE。需删除该部分数据。

二是坐标系问题:不同出租车服务商采用的定位数据坐标系可能不同,需要与交通调查所选用的坐标系统一。

三是有无客状态标记错误数据:在实践中,发现有部分出租车辆的载客状态字段[STATUS]的“摇摆”现象。例如*t*时刻显示“载客”,在*t+Δt*时刻(Δt 为轨迹数据的记录时间步长,本文中为15s)显示“空驶”,在*t+2Δt*时刻又显示“载客”。该部分数据需要通过完整的载客行驶链判断修正。

因而数据预处理包括:

一是删除未定位的数据;

二是坐标转换,将全部轨迹点经纬度坐标,转换为交通调查的统一坐标;

三是根据载客状态的变化提取初步订单行驶链,当载客状态在*t*时刻从“空驶”变为“载客”,在*t+NΔt*时刻(*N*为时间步长个数)从“载客”变为“空驶”,则*t*至*t+NΔt*时间内的记录构成一条初步订单行驶链 L_i (*i*为初步载客出行链编号);

四是完整订单行驶链处理,由于可能存在载客状态字段的“摇摆”现象,需要判断前后初步订单行驶链的时间间隔,记相邻两条初步订单行驶链 L_i 和 L_{i+1} 的终点和起点时间分别为 $t_{L_i}^s, t_{L_{i+1}}^s$,设定订单间时间阈值 θ ,若 $t_{L_{i+1}}^s - t_{L_i}^s \leq \theta$,则将 L_i 和 L_{i+1} 合并,继续与下一条链进行判断,若 $t_{L_{i+1}}^s - t_{L_i}^s > \theta$,则 L_i 标记为完整订单行驶链, L_{i+1} 进行下一步判断。

1.2.3 数据处理方法

根据数据分析需求不同,出租车数据处理分为订单行驶链的起终点处理和中间点处理。

(1) 订单行驶链的起终点处理

起终点处理目的是得到各车辆在每次订单的起点位置、起点时间、终点位置和终点时间,主要方法是对预处理后的完整订单行驶链起终点进行统计聚合处理,方法比较简单,本文不展开介绍。

(2) 订单行驶链的中间点处理

中间点处理的目的是根据出租车的实时速度判断路段车速。本文采用一种较为简单的“相邻定位点最短路径匹配法”实现地图匹配计算过程。过程介绍如下:

a. 路网数字化处理。为和交通调查中其他交通系统匹配,采用统一的路网,一般为CAD路网文件。在ArcGIS软件中导入CAD路网,转换坐标后,采用“要素转线”功能将路网打断,生成较为细致的路网图层。在该图层中,利用字段计算器中“shape.wkt”函数,生成线要素文本[wkt]字段。将[wkt]字段中的数据提取出来,获得路段点(交点或端点)的坐标、路段与路段点的关联关系。此处记路网*G*:

$$G = \{P, E\} \tag{1}$$

式中:*P*为路段点集合 $P = \{p_1, p_2, \dots, p_i, \dots, p_n\}$; *E*为路段集合 $E = \{e_1, e_2, \dots, e_i, \dots, e_m\}$; 对于路段集合中任意路段 e_i ,又由点集中的两元素 p_j (起点)、 p_k (终点)对应,即 $e_i = (p_j, p_k)$ 。

b. 相邻轨迹点匹配路段点。按距离最短原则,对于出租车载客行驶链 L_i 中的任意相邻两轨迹点 l_q 和 l_{q+1} ,匹配路段点集*P*中最近的两个路段点 p_q 和

p_{q+1} 。为减少匹配计算量,可划定每个轨迹点的匹配范围:记轨迹点 l_q 坐标分别为 x_q, y_q , 设定匹配坐标计算阈值 Δx 和 Δy , 则生成路段点集 P 的子集 P_{pq} 满足 $P_{pq}=\{p_i(x_i, y_i) | x_q - \Delta x \leq x_i \leq x_q + \Delta x, y_q - \Delta y \leq y_i \leq y_q + \Delta y\}$, 进而在 P_{pq} 中进行路段点的匹配。

c. 生成匹配路段点的最短路径并赋速度值。对于路网 G , 寻找路段点 p_q 和 p_{q+1} 间的最短路径 $E_q = \{e_q^1, e_q^2, \dots, e_q^n\}$, 对最短路径集 E_q 中的每一条路段记录定位点此时的时刻和速度。若车辆在某时段速度很低(例如发生了拥堵情况), 导致匹配的路段点相同, 即 $p_q = p_{q+1}$, 则将速度值赋给上一时段。同样, 为减少计算量, 可划定最短路径计算范围: 合并计算路段点 p_q 和 p_{q+1} 时的点集 P_{pq} 和 P_{pq+1} , 得到点集 $P_{\Sigma q}$, 并得到其关联的路段集 $E_{\Sigma q}$, 构成路网子集 $G_{\Sigma q} = \{P_{\Sigma q}, E_{\Sigma q}\}$, 进而在 $G_{\Sigma q}$ 中计算最短路径 E_q 。

d. 聚合出结果。所有车辆的轨迹点匹配完成后, 计算各时段每条路段的平均运行速度值。由于不同等级路段直接比较运行速度意义不大, 还可以计算路段运行指数。本文定于路段运行指数 = 该段道路上最高运行速度 / 统计时段内平均运行速度。

2 出租车轨迹数据在某市综合交通调查中的应用

以某市连续一周工作日的 1 816 辆巡游出租车数据进行相关交通调查工作。该市中心城区出租车总规模为 5 150 台, 其中巡游车 2 750 台, 网约车 2 400 余台, 出租车万人拥有率为 18.7 台 / 万人。样本比例为 35%, 满足抽样率要求。

2.1 出租车调查

经过计算, 该市出租车平均日载客 19 单, 行驶里程 163.2 km。平均每单载客时间 14.3 分钟, 出行距离为 6.7 km(见图 2)。平均里程利用率为 78%, 空驶率较低, 出租经营状态良好。从早 8 时至晚 21 时, 出租车订单比例均在 5% 以上, 波动较小; 出租车出行高峰发生在 17~18 时, 比例达 7.2%, 案例城市出租车订单时变特征见图 3。

从出租车订单的起点和终点分布(图 4、图 5)来看, 出发、到达地空间分布大体趋同, 主要分布在高铁北站、机场、火车站等重要交通枢纽区域, 以及大学城、解放路商业中心、北京路商务中心等商贸活动片区, 反映出租车出行主要服务进出城旅客和商贸活动乘客的特点。出租车订单起终点空间分布也可运营公司科学调配车辆资源提供决策依据。

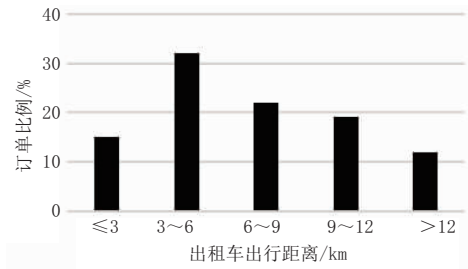


图2 案例城市出租车出行距离分布

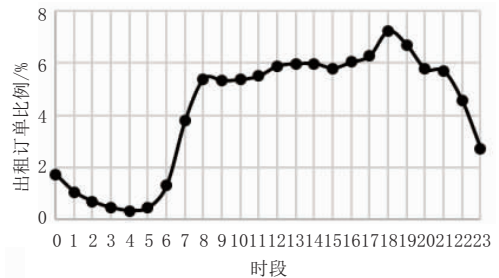


图3 案例城市出租车订单时变特征

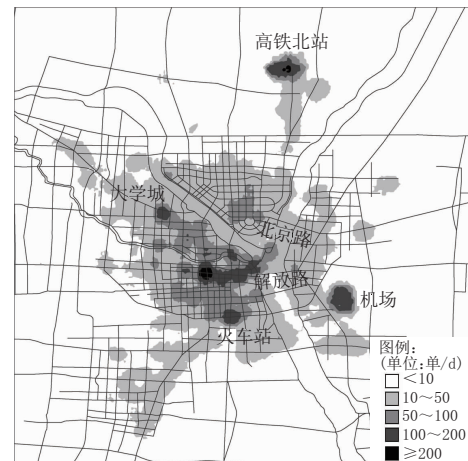


图4 案例城市出租车订单起点空间分布

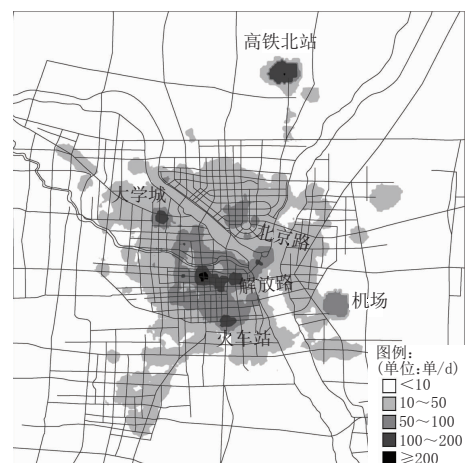


图5 案例城市出租车订单终点空间分布

2.2 车速调查

通过地图匹配和各条道路聚合计算, 结合不同道路等级对应的设计速度, 得到各时段各条路段平均行驶速度和拥堵指数分布。计算结果表明:

一是速度空间分布, 高峰小时路段拥堵指数分

布(如图6)表明,通达路桥、沂蒙路桥拥堵严重,北京路桥、解放路桥行驶缓慢,城区过河交通压力较大;横向兰山路、金雀山路、陶然路、凤凰大街,纵向蒙山大道、通达路、沂州路等路段较为拥堵。



图6 工作日晚高峰小时路段拥堵指数分布

二是速度时间分布,从全市出租车车速变化特征(图7)可看出,存在两个明显“低谷”时段,分别为7~8时和17~18时,且晚高峰运行车速更低,路网更拥堵。

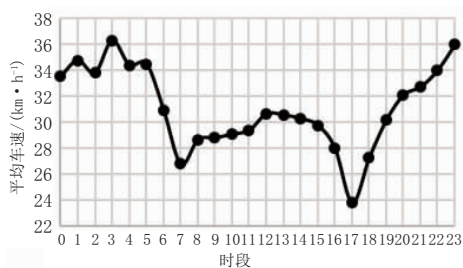


图7 案例城市出租车车速时变特征

2.3 出行生成源调查

以高铁北站为出行生成源,可研究其区域内出租车到达时空分布。以到达地分布为例(见图8),可看出,从高铁北站出发的出租车,终点集中分布在大学城区域,此外北京路商务区、解放路商业区分布较多,机场、河东区域、罗庄区域分布较少。

2.4 居民出行调查

将每笔出租车订单的起终点分别对应到交通小区(或中区、大区)中,按区域聚合,便可得到区域间的出租车出行OD矩阵和期望线分布。图9是该城市交通大区的出租车出行期望线图,可看出,兰山老城区与北城新区联系最为紧密,出租车出行量最大;与河东区北部、罗庄区联系较紧密;其他区域间出租车出行量不大。该数据结果,可为居民出行调查

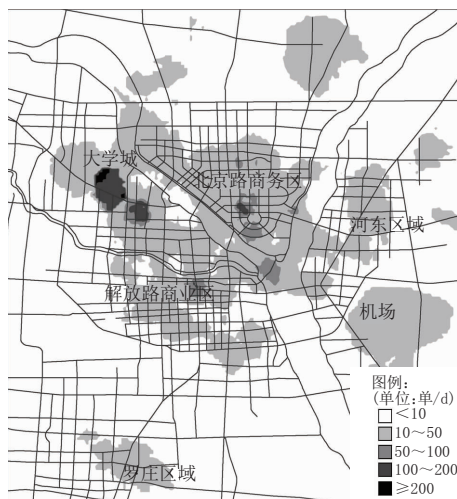


图8 以高铁北站为终点的出租车起点分布中的出租车出行分布提供校正依据,亦可为完善交通模型提供支持。

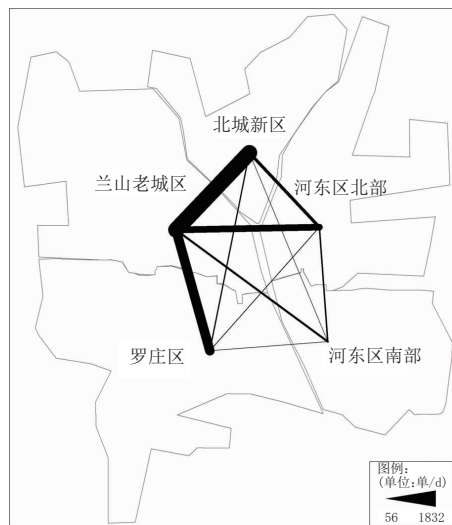


图9 交通大区出租车出行期望线图

3 结语

本文系统梳理了出租车轨迹数据在交通调查中的应用场景及应用方法,重点介绍了数据处理流程、关键算法,及其在某城市出租车调查、车速调查、出行生成源调查和居民出行调查中的应用结论。本文中的应用方法适用于综合交通规划、相关交通专项规划中的交通调查,为信息化手段进行交通调查提供了有效手段。今后将持续深入挖掘出租车轨迹数据的开发应用。

参考文献:

[1] GB/T 51334—2018,城市综合交通调查技术标准[S].
 [2] 中华人民共和国住房和城乡建设部.住房和城乡建设部关于印发城市综合交通体系规划交通调查导则的通知[EB/OL].(2014-9-25)[2021-11-21].http://www.mohurd.gov.cn/wjfb/201409/t20140929_219194.html.

试验阶段,该位置锁扣型钢地下连续墙围护结构深层水平位移与一般的钢筋混凝土制地下连续墙相比较小。监测结果表明锁扣型钢地下连续墙在本项目深基坑围护施工中是可行的。

4 锁扣型钢地连墙工法与常规地连墙工法的造价指标对比

选取深度为60 m,厚度为1 200 mm,幅宽为3.2 m的单幅地连墙进行比较。表2为单幅常规地墙和锁扣型钢地下墙的材料用量和单价表。表3为单幅常规地连墙和锁扣型钢地连墙的综合单价。经计算,单幅锁扣型钢地连墙造价高1.2倍。由于其刚度大和薄壁化的特点,通过对壁厚的进一步设计优化,后期造价可进一步减少。

表2 单幅地连墙地连墙的材料用量和单价表

材料	材料用量	单价	备注
钢筋	40 t	6 000 元/t	常规地连墙
混凝土	230.4 m ³	736 元/m ³	常规地连墙
锁扣型钢构件	88 t	7 370 元/t	锁扣型钢地连墙
大流动性混凝土	230.4 m ³	1 200 元/m ³	锁扣型钢地连墙

表3 单幅地连墙的综合单价 单位:元/m³

	材料费	机具费	人工费	综合单价
常规地连墙	1 777.7	554.1	78.9	2 410.7
锁扣型钢地连墙	4 014.9	1 091.1	197.5	5 303.5

5 结语

为实现在狭窄的城市空间内修建超深、高刚度的地下连续墙,本项目进行了锁扣型钢地下连续墙施工法的试验研究,得出以下结论:

(1)该工法钢构件工厂化制作,提高了施工质量和施工效率,显著提升现场安全文明施工水平,降低现场制作带来的安全隐患。有利于施工集中管理,平衡施工场地、机械及劳动力等各种资源,降低施工成

本。设备、生产和人员的管理难度降低,降低管理成本。

(2)锁扣型钢地下墙与一般的钢筋混凝土制地下墙相比无需钢筋笼胎膜场地,能够节约施工场地。围挡时能节约1~2根社会车道,明显改善施工时对社会影响。特别适合于城市建筑密集、施工场地狭小、基坑挖深大的施工项目。

(3)锁扣型钢地下连续墙采用CT锁扣作为型钢构件连接件,止水效果好,开挖面墙体无渗漏。

(4)试验阶段,与一般的钢筋混凝土制地连墙相比,单幅锁扣型钢地连墙造价高1.2倍。由于其刚度大和薄壁化的特点,通过对壁厚的进一步设计优化后,材料用量可进一步减少,造价能够控制在传统工法的115%左右。如遇中心城区施工场地狭小的情况时,其经济指标优于传统工法。

综上所述,锁扣型钢地连墙施工法具有钢构件工厂化制作、止水效果好、节约施工场地的优势。在中心城区场地面积狭小的条件下,该工法能最大限度地有效利用施工场地面积,实现快速施工且符合安全文明施工的要求,适用于传统技术难以施工的狭窄地区的地下工程建设。通过本文的工程实例,我们认为该施工法具有可行性,可为类似工程提供经验和参考。

参考文献:

[1] 方银钢.软土地区深大基坑变形控制措施研究[J].研究成果,2019(7):13-15.

[2] 贾坚,谢小林,翟杰群,等.软土基坑变形控制的微扰动技术[J].上海交通大学学报,2016,50(10):1651-1657.

[3] 赵启嘉.预制地下连续墙在地铁工程中的应用研究探讨[J].工程技术,2015(21):96-97.

[4] 田崎和之.钢制地下连续墙施工法的设计与施工[J].探矿工程译丛,1996(4):6-11,42.

[5] 谈海斌,陈超平.钢制地下连续墙的设计、施工要点[J].公路工程,2012(12):218-221.

(上接第193页)

[3] 高文超,李国良,塔娜.路网匹配算法综述[J].软件学报,2018,29(2):225-250.

[4] 曾嘉邨,孙立双,王晓明.北京出租车GPS轨迹数据地图匹配算法研究[J].北京测绘,2019,33(3):255-260.

[5] 谢彬彬.基于隐马尔科夫模型的出租车OD时序轨迹的道路匹配算法研究[D].长安大学,2020.

[6] 盛彩英,席唱白,钱天陆,等.浮动车轨迹点地图匹配及插值算法[J].测绘科学,2019,44(8):106-112.

[7] 张校慧,孙凯,职保平,等.针对复杂道路网络的车辆轨迹地图匹配算法[J].测绘科学,2018,43(8):110-115.

[8] 周英迪,张兴国,潘晓芳,等.基于出租车轨迹的居民出行时空分析[J].城市勘测,2020(3):17-23.

[9] 邸少宁,朱杰,郑加柱,等.出租车轨迹数据的南京人群出行模式挖掘[J].测绘科学,2021,46(1):203-212.

[10] 罗钧韶,潘嘉杰.基于GPS数据挖掘的出租车出行特征分析[J].交通与运输,2020,33(S2):49-54.

[11] 罗琼,舒红,徐亚瑾,等.移动轨迹数据支持下的城市居民通勤活动分析[J].武汉大学学报(信息科学版),2021,46(5):718-725.

[12] 杨丁山.基于浮动车轨迹数据的城市路网交通状态判别及关键节点识别[D].重庆:重庆交通大学,2020.