

基于阶梯收费刷卡数据的公交下车站点 算法优化与实证评估

杨 飞¹, 姜海航¹, 郭煜东¹, 刘健国¹, 周 涛²

(1. 西南交通大学 交通运输与物流学院, 四川 成都 611756; 2. 重庆市交通规划研究院, 重庆 401147)

摘要: 以某市连续 5 日全天公交阶梯收费刷卡数据、公交出行 GPS(global positioning system)数据及公交站点数据为基础, 结合出行链算法与随机森林网络, 构建了一套公交下车站点融合分析模型。在模型中, 首先匹配 GPS 数据与公交站点数据, 确定不同时刻的公交到站信息, 再以乘客上车站点位置、出行频率、活动空间、下车点用地类型分布、下车概率为输入, 识别乘客下车站点, 最终下车站点推算率提升至 100%, 全样本有效率达 76.2%, 相比现有基于出行链的方法, 识别有效率提升 37%。

关键词: 智能公交; 公交下车站点推导; 公交刷卡数据; 随机森林网络; 出行链算法

中图分类号: U491.1

文献标志码: A

Optimization and Empirical Evaluation of Passenger Leaving Station in Bus Based on Gradient Charge Swiping Card Data

YANG Fei¹, JIANG Haihang¹, GUO Yudong¹,
LIU Jianguo¹, ZHOU Tao²

(1. School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 611756, China; 2. Chongqing Transport Planning Institute, Chongqing 401147, China)

Abstract: This paper proposed a fusion analysis model for leaving station identification, based on the 5-day continuous (24 h) bus travel data, bus travel GPS data, and bus stop data of a certain city, in combination with the travel chain method and random forest network. In the model, the GPS data and bus stop data were first matched to determine the bus arrival information at different moments, then the passenger boarding location, travel frequency, activity scope,

land use around the leaving station, and the probability of getting off the bus were applied as inputs to identify the passenger leaving station. The final calculation rate is improved to 100%, and the full sample efficiency rate reaches 76.2%. Compared with the existing methods based on the bus travel chain, the recognition efficiency is improved by 37%.

Key words: intelligent public transportation; leaving station estimation; swiping card data; random forest network; travel chain algorithm

公共交通因其成本低、能耗低、占用资源少等优点, 成为众多城市倡导的出行方式。随着相关政策^[1]的提出, 智能公交系统建设与优化受到关注。相比常规公交体系, 智能公交在指挥调度、运营管理上实现自动化、精细化操作, 大幅提升公交服务水平。准确分析公交客流下车站点, 了解客流出行特征, 摸清客流需求, 对智能公交的调度、运营、规划具有重要意义。公交 IC 卡作为各个城市智能公交的基础部分, 存储了持卡用户大量的乘车交易时空记录, 成为获取公交基础数据的新来源。基于公交 IC 卡数据的下车站点研究成为研究热点。

在下车站点推算方法中, 集计推算是一种常用、便捷的估计方法, 但该方法仅能推算出站点客流总量, 无法确定每位乘客的下车站点。个体推算方法中, 基于公交出行链的估计考虑较多, 这些方法相对容易且易于理解。2002 年, James 等^[2]根据地铁出行日志的记录设计了两个假设: 一是乘客会返回上次的目的地车站开始下一次出行, 二是乘客的最后一次出行会返回第一次出行的起点, 据此利用纽约市

收稿日期: 2021-12-16

基金项目: 国家自然科学基金(52072313); 国家重点研发计划(50908195); 重庆规划和自然资源局项目(KJ-2021007)

第一作者: 杨 飞(1980—), 男, 教授, 博士生导师, 工学博士, 主要研究方向为交通大数据与智能交通。

E-mail: yangfei_traffic@163.com

通信作者: 郭煜东(1994—), 男, 博士生, 主要研究方向为交通大数据。E-mail: guoyudong_traffic@163.com



论文
拓展
介绍

的AFC(automatic fare collection)数据实现了下车站点的推算。之后一些学者在该研究基础上进一步修正,Munizaga等^[3]研究中发现仅通过该假设易导致下车站点推算错误,在James等假设的基础上,进一步考虑了下一次公交出行的上车位置和时间并增加了额外的约束条件,优化了站点推算目标函数。Alsger等^[4]利用澳大利亚昆士兰州公交系统中含有乘客下车记录的特殊收费数据验证了出行链理论的下车站点推算模型,并研究了最大步行距离取值对推算结果的影响。He等^[5]针对出行链断裂情况(无下一次公交出行或下一次公交出行起点不能作为上次出行的终点),利用乘客历史出行数据建立了个体空间和时间的下车复合概率模型,用以推算出行链理论无法解决的记录。Kumar等^[6]放松了现有公交出行链方法的参数限制,采用多元logit模型计算连续上车站点之间的路径选择概率,以最大概率方式推算乘客的下车站点。

但基于公交出行链的经典推算方法难以估计单次、独立出行的下车站点,造成估计客流不能代表所有公交乘客真实的出行情况。一些研究通过补充多日刷卡记录推算更多的数据,其主要思路是在多日数据中寻找相似的出行记录或与他日出行记录建立连接修复断裂出行链。但该方法在海量数据中需要设定复杂的规则,同时并不能有效地提升推算率。为解决该问题 Gordon等^[7]研究中补充了公交投币箱票务数据和登车器人次数据,以此估算规则算法无法确定乘客和票务现金支付乘客的出行流量。Yan等^[8]提出应用朴素贝叶斯、SVM(support vector machine)、决策树、随机森林、KNN(k-nearest neighbor)和集成学习等机器学习算法进行下车站点的识别,杨鑫^[9]以乘客的基本上车特征作为模型的输入,采用Bi-LSTM-CRF模型对乘客的下站点实现推算。

乘客在选择某线路某站点上车后,可供选择的下车站点是有限的,因此相关研究将乘客下车站点推算视为分类问题,采用机器学习方法实现推算^[8-11]。相对于出行链的推算方法,基于机器学习的方法能够解决单次出行的下车站点识别,但也需要大量的训练数据及标签,同时预测耗时长,预测结果也存在较大偏误。对此,Yan等^[8]利用POI(point of interest)数据确定站点所在区域的土地性质,估计下车站点所在的交通小区,最终正确率达到70%以上,识别效果较优。

对于这类数据,也有学者从概率角度进行分析。

胡继华等^[12]通过计算站点上车客流,推算指向下游站点吸引权最大的站点,李佳怡等^[13]同样采用了该方法。这些方法都假设个体服从群体的下车分布概率,即乘客很有可能选择最大概率(最大吸引权)站点下车。

然而,大部分公交线路采用“一票制”收费方式,乘客下车时不会进行刷卡操作,基于公交IC卡的下车站点估计缺乏有效验证。对此,本文首先利用随机森林算法对现有出行链算法进行优化,构建了完整的公交出行下车站点估计模型。然后结合具有公交上下站点的阶梯收费刷卡数据,对两种算法的识别效果进行实证评估,保障了算法的可靠性。

1 公交下车站点估计模型

根据个体乘坐公交的行为模式,下车站点推算包括两个步骤:首先结合乘客常见行为模式,构建出行链估计模型,再针对出行链断裂等复杂情况,利用随机森林网络进行进一步优化。

1.1 出行链估计模型

个体每日出行活动具有一定规律性,其公交出行多呈现时空闭环特性。本文结合现有研究,以个体多日或连续的公交出行行为为基础,首先构建包含下述3条推断的出行链估计模型,初步推算下车站点。

(1)当乘客存在相邻的两次乘车出行,若后次的乘车站点位于前次乘车站点的下游,则乘客前次乘车的下车站点位于后次乘车站点附近。

(2)当乘客为最后一次乘车出行,若该乘车站点远离当日第一次乘车站点,则为返程出行,下车站点应位于第一次乘车站点附近。

(3)当乘客为最后一次乘车出行,若该上车站点靠近当日第一次乘车站点,则为非返程出行,下车站点位于第一次乘车的下车站点附近。

1.2 随机森林网络

由于个体出行行为相对复杂,仅依靠出行链模型不可能实现所有情况分析,尤其是单次刷卡记录和出行链断裂记录。对此本文结合随机森林网络,对两类数据的下车站点进行预测。利用公交IC卡数据推算下车站点可视为一种分类问题,即通过乘客上车时间、乘坐线路等特征筛选最可能的下车站点。随机森林网络是一种统计学习算法,通过重抽样方法从原始样本中抽取多个样本,对每个样本进行决策树建模,组合多棵决策树的预测并投票得出

最终结果。随机森林网络具有分类回归精度高、不易过拟合、高维数据适应性强等优点,适用于复杂问题的分类与拟合分析。公交下车站点估算本质可视为一种分类问题,本文以下车站点编号为目标,将上车时间、乘坐线路、上车站点、持卡类型、用地性质等个体属性作为输入,构建相应随机森林网络。公式(1)为预测函数。

$$H_{(x)} = \arg \max_Y \sum_{i=1}^k I(h_i(x)=y) \quad (1)$$

式中: $H_{(x)}$ 为随机森林对输入变量 x 的最终预测; $h_i(x)$ 为单个决策树分类模型; $I(X)$ 为示性函数。边缘函数是随机森林网络的重要属性,代表着对于输入变量 X 的分类中,正确分类得票数超过错误分类得票数的概率,边缘函数越大,随机森林分类的正确率越高,决策树与整个森林的边缘函数如公式(2)、(3)所示。

$$g(X, Y) = a_k I(h_k(X) = Y) - \max_{j \neq Y} a_k I(h_k(X) = j) \quad (2)$$

$$r(X, Y) = P_\theta(h(X, \theta) = Y) - \max_{j \neq Y} P_\theta(h(X, \theta) = j) \quad (3)$$

式中: $g(X, Y)$ 与 $r(X, Y)$ 分别为决策树与整个森林的边缘函数; $a_k I()$ 为所取平均值; $P_\theta(h(X, \theta) = Y)$ 为判断正确的分类概率; $\max_{j \neq Y} P_\theta(h(X, \theta) = j)$ 为判断错误的其他分类的概率的最大值。

公式(4)为泛化误差,表示模型对未知数据的预测能力,泛化误差越小模型预测效果越好。公式(5)为随机森林的边缘函数期望。

$$E^* = P_{X,Y}(g(x, y) < 0) \quad (4)$$

$$s = E_{X,Y} r(x, y) = E_{X,Y} (P_\theta(h(x, \theta) = y) - \max_{j \neq y} P_\theta(h(x, \theta) = j)) \quad (5)$$

公式(6)为决策树之间的相关程度,其中 $v(r)$ 、

$s(h(r))^2$ 分别为 $r(X, Y)$ 的方差与标准差。

$$\bar{\rho} = \frac{v(r)}{s(h(r))^2} \quad (6)$$

最终随机森林网络模型表达式如式(7)所示。

$$\bar{Y} = f_{RF}(\theta_1, \theta_2, \dots, \theta_T; x_1, x_2, \dots, x_n) \quad (7)$$

式中: $(\theta_1, \theta_2, \dots, \theta_T)$ 为随机森林中待标定参数; T 为待标定参数的个数; n 为输入变量维度。

2 数据分析

2.1 公交阶梯收费刷卡数据

公交阶梯收费刷卡数据是乘客下车识别的基础,相比传统公交IC卡数据,乘客在分段计费线路的上下车均要刷卡,为个体层面数据验证提供了保障。本文公交IC卡数据涵盖某市连续5个工作日城市区域全部数据,共计11条分段计费线路,215 212条刷卡数据。线路均为中长距离干线,沿城市主要客流走廊布设,占城市公交总运量的40%,具有良好的代表性。刷卡数据字段包含乘客持卡编号、刷卡时间、乘坐线路、乘坐车辆、刷卡行为等信息,其中刷卡行为用以说明乘客是上车还是下车。表1为2019年7月15日公交刷卡数据样例。

表1 公交刷卡数据样例

Tab.1 Sample of bus swiping card data

车牌	线路	卡号	时间	卡型	行为
56093	27路	6170**878	2019-07-15 9:20:32	老年免费卡	上车
56093	27路	617**126	2019-07-15 6:38:56	成人月票卡	上车
56093	27路	617**126	2019-07-15 7:25:48	成人月票卡	下车
56093	27路	617**136	2019-07-15 15:40:18	成人月票卡	上车
56093	27路	617**136	2019-07-15 16:03:54	成人月票卡	下车

2.2 公交GPS数据

公交GPS用于确定公交实时位置,判断公交到站情况。其数据字段包含车辆牌照、线路、时间、速度和经纬度等信息。表2为公交GPS数据样例。

表2 GPS数据样例

Tab.2 Sample of GPS data

车牌	线路	时间	速度/(m·s ⁻¹)	方向角/(°)	经度/(°)	纬度/(°)
川D23678	2路	2019-07-17 07:14:18	3.20	299.00	101.741 8	26.569 49
川D23678	2路	2019-07-17 07:14:33	2.55	305.00	101.741 5	26.569 71
川D23678	2路	2019-07-17 07:14:48	2.86	307.00	101.741 1	26.569 94
川D23678	2路	2019-07-17 07:15:04	3.18	288.00	101.740 7	26.570 12
川D23678	2路	2019-07-17 07:15:19	0	271.00	101.740 4	26.570 26

2.3 公交站点数据

由于公交GPS数据不包含公交进出站信息,需要联合公交站点数据,识别不同时间公交到站点。

数据字段包括站点名称,经纬度,途经线路等,表3为站点信息数据样例。

表3 公交站点数据样例

Tab.3 Sample of bus station data

站点编号	站点名称	经度/(°)	纬度/(°)	单程
10001	阿署达	101.768 644	26.543 296	阿署达专线下行
10002	阿署达	101.768 644	26.543 296	阿署达专线上行
10009	奥林匹克	101.732 886	26.567 027	109下行、113下行
10010	奥林匹克	101.732 835	26.567 112	113上行、109上行
10011	奥林匹克花园	101.733 204	26.564 308	113下行、下行
10012	奥林匹克花园	101.733 984	26.563 669	113上行、上行

2.4 模型输入特征分析

乘客公交出行中将产生多种关联信息,合理的输入参数是下车站点估计的关键,根据个体实际分析,本文选择如下特征,同时结合其他信息构造新的特征,作为随机森林的输入。

2.4.1 上车时间、上车路线、上车站

个体的一日活动通常具有规律,同一时段、同一路线及站点的活动目的相对固定,下车站点也基本相同。

为保障识别结果的准确,上车时间以小时为间隔进行离散化,上车路线、上车站点利用标签编码法(label encoding),从0开始进行重新编码。

2.4.2 持卡类型

持卡类型代表乘车人属性,能够大范围缩小目标用户,与(1)中特征联合,能够较好刻画乘客特征。持卡类型同样利用标签编码法进行编码。

2.4.3 活动半径与出行频率

活动半径表征乘客活动空间大小^[14],代表乘客站点空间特征,出行频率代表乘客对各个站点的出行偏好。图1为活动半径示意图。本文以5个工作日全部乘车数据为基础,分析乘客所有上车站点分布特征,公式(8)为活动半径计算公式。

$$\mu=\frac{1}{n}\sum_i^n r_i \tag{8}$$

式中: μ 为活动半径; r_i 为形心到上车站点*i*的距离。

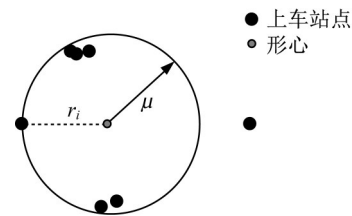


图1 活动半径计算原理

Fig.1 Calculation principle of active radius

由于活动半径 μ 与出行频率分别为连续型与接近连续型变量,直接使用难以获得准确估计结果,同时无法直接确定分类类型。对此,本文利用K-

means聚类方法对两个变量分别聚为3类,以划分不同出行特征的乘客,表4为出行频率与活动空间的最终聚类结果。

表4 出行特征划分规则

Tab.4 Classification rules of travel feature

特征	标签	类别	取值界限
出行频率 <i>f</i>	1	低频	$f\leqslant 5$
	2	中频	$6\leqslant f\leqslant 12$
	3	高频	$f\geqslant 13$
活动半径 <i>μ</i> /m	1	短距离	$\mu\leqslant 1283$
	2	中距离	$1\,283<\mu\leqslant 3\,300$
	3	长距离	$\mu>3\,300$

为了更直观地展示特征取值分布,将取值进行min-max归一化,取值界限约束到[0,1]之间。以出行频率*f*为例,转换函数如式(9)所示。

$$f_k^*=\frac{f_k-\min(f)}{\max(f)-\min(f)} \tag{9}$$

式中: f_k 为出行频率*f*中第*k*个取值; f_k^* 为*f_k*归一化后的取值; $\min(f)$ 为所有出行频率*f*中的最小值; $\max(f)$ 为所有出行频率*f*中的最大值。

图2为聚类结果示意图。可见不同活动空间与出行距离被分为9类,不同类别间的特征差异显著。

2.4.4 下游站点用地类型

不同用地类型下的站点客流吸引力存在较大差异,本文以站点周围500 m半径为活动范围,统计下游主要用地类型,图3为站点周围用地类型示意图。

由于乘客并不会在所有下游站点进行选择,本文以*R*为活动空间影响阈值,提取当前站点到*R*内特定站点*C*的用地面积,并将最大面积作为输入特征。

$$C=Rl_{\text{non}}/d_s \tag{10}$$

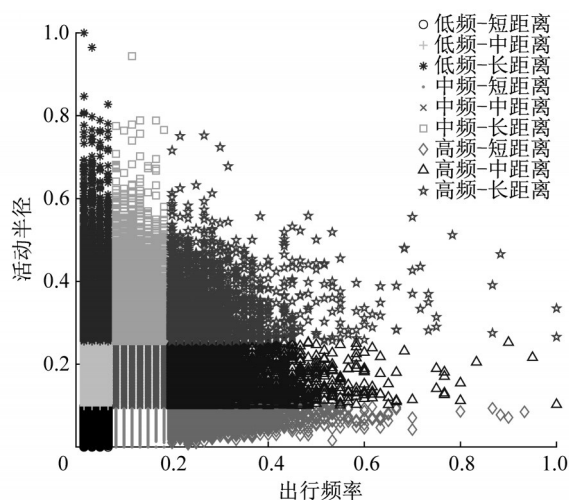


图2 公交乘客聚类

Fig.2 Clustering of bus passengers

式中: C 为阈值范围内的下游站点数; R 为活动范围直径; l_{non} 为线路非直线系数; d_s 为线路平均站间距。

2.4.5 下游站点下车概率

受站点周边区域人口、岗位和土地利用等因素的影响,不同站点的下车客流存在差异,区域内站点下车客流能够反映出区域的吸引力,吸引力较大的区域更有可能吸引乘客下车。本文将同一群体在同一线路、同一方向、同一站点的下车分布频率作为先验概率,则概率最大的站点为乘客最可能前往站点。以0.5为最大概率的下限值,若最大概率大于下限值,则最大概率对应的站点可推算为下车站点,若小

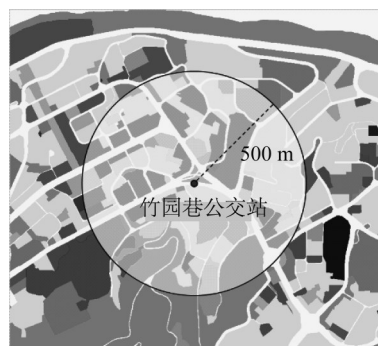


图3 站点500 m覆盖用地示意图

Fig.3 Schematic diagram of 500m site land use coverage

于下限值则需要采用其他规则重新推算。其中,本文利用乘客下车信息直接获取该先验概率,对于非阶梯收费线路,可预先调查获取目标线路各站点的下车乘客数,再统计获取目标站点下游站点的下车概率。

对于无法确定最大概率站点(概率 >0.5)的情况,可利用相邻站点概率分布进行叠加,以此反映相邻站点周边区域的下车吸引力。如图4所示,在同一站点上车的群体出行数据中提取各下车站点($s_1 \sim s_{10}$)的出行频率($f_1 \sim f_{10}$),将频率大于零的相邻下车站点划分为多个站点区域,叠加各区域内的站点频率作为该区域的下车概率 P ,则下车概率最大的区域为乘客最可能前往的区域,该区域中概率最大的站点可作为乘客的下车站点。

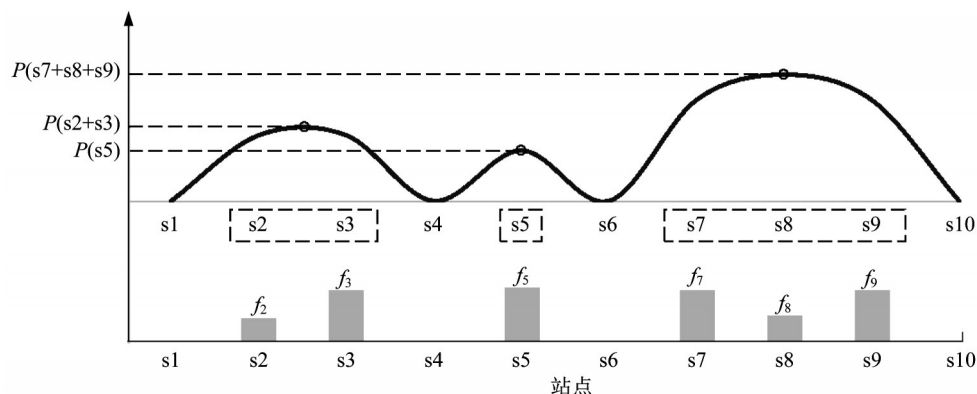


图4 站点连续区域下车概率示意图

Fig.4 Probability diagram of getting off in continuous area of station

3 实例应用

本文针对5个连续工作日内采集的215 158条公交出行数据,分别对比现有公交出行链模型、随机森林算法与本文构建融合模型的下车站点识别差

异,并结合推算率、有效率、正确率3种指标分析识别结果。其中正确率代表估计站点与实际站点一致;有效表示推算站点与实际站点误差在一个站范围内。

$$N = PC_1E \quad (11)$$

$$E_a = P_e / P \tag{12}$$

式(11)、(12)中: N 为有效估计的站点数量; P 为全部样本数; C_i 为推算率; E 为有效率; E_a 为全样本有效率; P_e 表示有效估计的站点数量。

3.1 与公交出行链模型对比

表5为公交出行链模型与本文融合模型的识别结果对比。可见公交出行链模型尽管具有较高的正

确率和有效率,但全样本有效率为55.6%,无法估计单次刷卡记录和出行链断裂记录,也难以代表全样本情况。而本文提出的融合模型能够实现100%出行数据推算,全样本有效率达到76.2%,同时样本有效率与正确率相比出行链模型差异不大,整体提升效果显著。

表5 融合方法与公交出行链模型结果对比
Tab.5 Comparison of fusion method and bus travel chain model

方法名称	全样本数	推算率/%	正确率/%	有效率/%	有效样本数	全样本有效率/%
基于公交出行链的方法	215 158	64.5	62.8	86.6	119 658	55.6
本文融合模型	215 158	100.0	52.3	76.2	163 843	76.2

3.2 与随机森林方法对比

针对公交出行链模型无法推算的35.5%剩余数据,将既有的随机森林方法与本文提出的融合模型的识别效果进行对比,如表6所示。由表6可知,两类方法均能实现100%结果估计,尽管既有的随机森林方法在公交出行起讫交通小区识别的正确率达到70%以上,然而本文以出行起讫公交站点为评价单元,相比交通小区层面识别更加困难。将既有方法应用到公交站点识别中,识别正确率约为19%,而本文提出模型通过增加乘客活动半径、出行频率等输入,将正确率提升至33.4%。两种方法的正确率都相对较低,这主要由于单次刷卡以及出行链断裂的出行记录无明显规律性,难以准确推算乘客的下车站点。但值得注意的是,本文模型在错误估计的估计下车站点中,57.2%与真实站点仅相差1个公交站,相比现有研究仍具有较大提升,说明算法仍

具有相对准确的识别效果。

表6 融合方法与随机森林方法结果对比
Tab.6 Comparison of fusion method and random forest

方法名称	样本量	推算率/%	正确率/%	有效率/%
基于随机森林的方法	76 381	100.0	18.6	37.7
本文融合模型	76 381	100.0	33.4	57.2

3.3 基于融合模型的下车估计结果

图5、6为不同出行距离下,本文提出融合模型的下车站点估计的详细结果。在2 km误差范围内,误差分布于0~100 m内占比最高,其次主要集中于200~600 m,大致为两个站点之间的距离。可能原因是乘客下车在站点周边移动后并未返回前次出行的站点上车,造成推算结果与真实站点间存在一段距离。

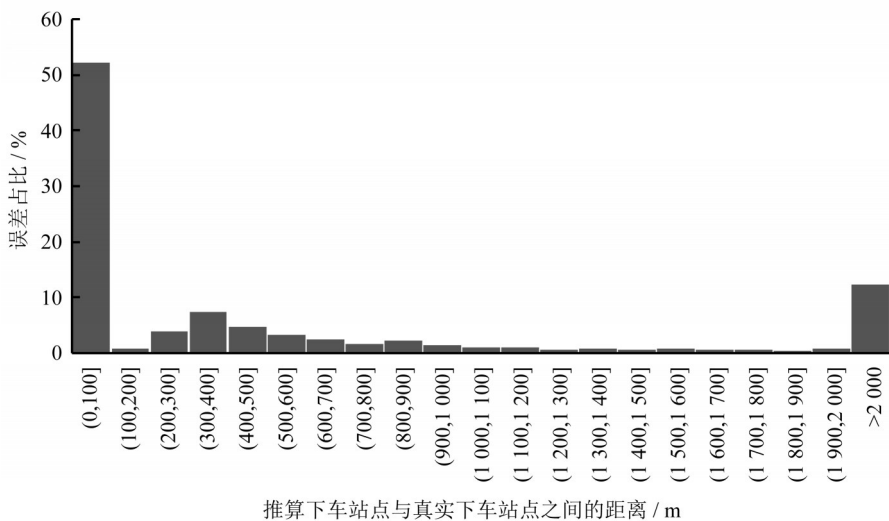


图5 下车站点短距离误差分布
Fig.5 Error distribution of short distance at leaving station

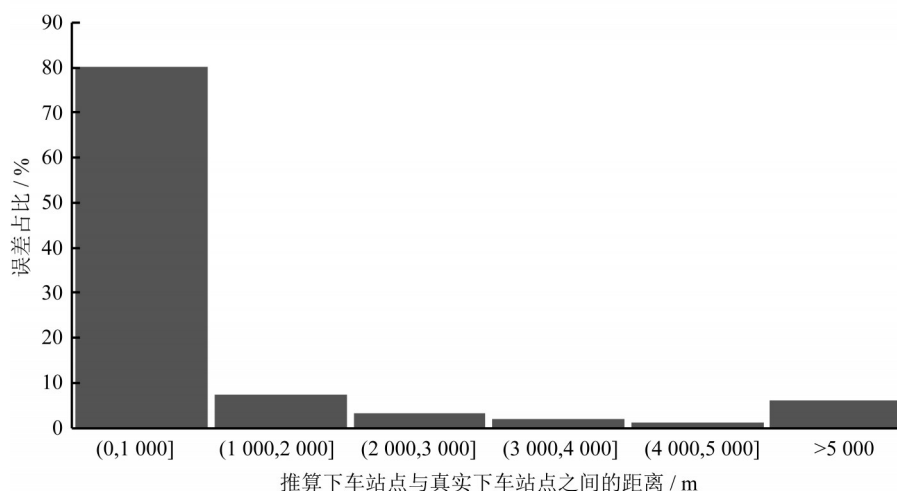


图6 下车站点长距离误差分布

Fig.6 Error distribution of long distance at leaving station

从长距离出行的误差分布来看,近80%的误差在0~1 km之内,误差在5 km以上的占6.4%。由于城市交通出行需求研究中,常以交通小区为基本单位,占地面积为1~2 km²,因此从小区尺度看,近80%的推算记录能够满足分析需求。

4 结论

准确预测公交客流下车站点是了解站点客流、线路客流以及公交起讫点的基础,有助于把握公交出行特征,对智能公交规划、调度、运营、管理具有重要的意义。

(1)本文以公交阶梯收费刷卡数据、车载GPS数据为基础,在传统公交出行链识别算法基础上,结合随机森林算法,构造了一套公交下车站点融合分析模型。首先基于公交出行链的方法,估计大部分乘客的下车站点,然后对未推算记录,利用随机森林方法进一步完成估计。

(2)构建的融合模型针对刷卡记录的下车站点成功推算率达到100%,全样本有效率达76.2%,其中近80%的记录误差在1 km范围以内。说明本文提出模型具有较高的识别效果,能够满足区域层面(交通小区)的分析需求。

作者贡献声明:

杨 飞:核心思想提炼。
姜海航:论文撰写,论文修改。
郭煜东:论文撰写,论文修改。
刘建国:数据收集处理,论文撰写。
周 涛:资料整理,思想提炼。

参考文献:

- [1] 中华人民共和国国务院. 国务院关于城市优先发展公共交通的指导意见(国发[2012]64号)[EB/OL]. [2013-01-15]. http://www.gov.cn/zhengce/content/2013-01/05/content_3346.htm.
The State Council of the People's Republic of China. Guiding opinions of the State Council on giving priority to the development of public transport in cities (No.64 Document in 2012 of State Council)[EB/OL]. [2013-01-15]. http://www.gov.cn/zhengce/content/2013-01/05/content_3346.htm.
- [2] JAMES J B, ROBERT N, ADAM R, *et al.* Origin and destination estimation in New York City with automated fare system data[J]. *Transportation Research Record Journal of the Transportation Research Board*, 2002, 1817: 183.
- [3] MUNIZAGA M A, PALMA C. Estimation of a disaggregate multimodal public transport origin - destination matrix from passive smartcard data from Santiago, Chile[J]. *Transportation Research Part C: Emerging Technologies*, 2012, 24: 9.
- [4] ALSGER A, ASSEMI B, MESBAH M, *et al.* Validating and improving public transport origin - destination estimation algorithm using smart card fare data [J]. *Transportation Research Part C: Emerging Technologies*, 2016, 68: 490.
- [5] HE L, TRÉPANIER M. Estimating the destination of unlinked trips in transit smart card fare data[J]. *Transportation Research Record*, 2015, 2535(1): 97.
- [6] KUMAR P, KHANI A, HE Q. A robust method for estimating transit passenger trajectories using automated data [J]. *Transportation Research*, 2018, 95: 731.
- [7] GORDON J B, KOUTSOPOULOS H N, WILSON N H M. Estimation of population origin - interchange - destination flows on multimodal transit networks [J]. *Transportation Research Part C: Emerging Technologies*, 2018, 90: 350.
- [8] YAN F, YANG C, UKKUSURI S V. Alighting stop determination using two-step algorithms in bus transit systems

- [J]. *Transportmetrica A: Transport Science*, 2019, 15(2): 1522.
- [9] 杨鑫. 基于IC卡数据的公交客流智能推断方法研究[D]. 北京: 北京邮电大学, 2019.
- YANG Xin. Research on intelligent bus passenger flow inferring method based on IC card data [D]. Beijing: Beijing University of Posts and Telecommunications, 2019.
- [10] JUNG J, SOHN K. Deep-learning architecture to forecast destinations of bus passengers from entry-only smart-card data [J]. *Iet Intelligent Transport Systems*, 2017, 11(6): 334.
- [11] ASSEMI B, ALSGER A, MOGHADDAM M, *et al.* Improving alighting stop inference accuracy in the trip chaining method using neural networks [J]. *Public Transport*, 2020, 12(1): 89.
- [12] 胡继华, 邓俊, 黄泽. 结合出行链的公交IC卡乘客下车站点判断概率模型[J]. *交通运输系统工程与信息*, 2014, 14(2): 62.
- HU Jihua, DENG Jun, HUANG Ze. Trip-chain based probability model for identifying alighting stations of smart card passengers [J]. *Journal of Transportation Systems Engineering and Information Technology*, 2014, 14(2): 62.
- [13] 李佳怡, 张锦, 张静文, 等. 城市公交乘客下车站点推算方法和有效性评价[J]. *武汉大学学报(信息科学版)*, 2018, 43(8): 1172.
- LI Jiayi, ZHANG Jin, ZHANG Jingwen, *et al.* An algorithm to identify passengers' alighting stations and the effectiveness evaluation [J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(8): 1172.
- [14] 闫小勇. 超越引力定律 空间交互和出行分布预测理论与方法 [M]. 北京: 科学出版社, 2019.
- YAN Xiaoyong. Theory and method of spatial interaction and trip distribution prediction based on the law of transcendental gravitation [M]. Beijing: Science Press, 2019.

(上接第308页)

- GUO Xiaofan, LI Linbo, WANG Yanli, *et al.* Satisfaction evaluation of public transit service based on entropy weight method and cloud model [J]. *Journal of Chongqing Jiaotong University (Natural Science)*, 2018, 37(9): 101.
- [19] 吴大伟. 基于层次分析法的公交发展水平评价模型探究[J]. *中国公路*, 2020(11): 178.
- WU Dawei. Research on evaluation model of public transit development level based on analytic hierarchy process [J]. *China Highway*, 2020(11): 178.
- [20] 刘丽娇. 基于五元联系数法的公交运营服务水平评价方法研究[D]. 北京: 北京交通大学, 2019.
- LIU Lijiao. Research on evaluation method of bus operation service level based on five-element connection number method [D]. Beijing: Beijing Jiaotong University, 2019.
- [21] 陈小鸿. 城市客运交通系统 [M]. 上海: 同济大学出版社, 2008.
- CHEN Xiaohong. Urban passenger transport system [M]. Shanghai: Tongji University Press, 2008.
- [22] 胡敏达. 城市BRT系统专用道与站点设计研究[D]. 西安: 长安大学, 2015.
- HU Minda. Study on the design of urban BRT lanes and stations [D]. Xi'an: Chang'an University, 2015.
- [23] 刘颖杰, 靳文舟, 康凯. 基于IC信息和概率理论的公交OD反推方法[J]. *公路与汽运*, 2010(3): 31.
- LIU Yingjie, JIN Wenzhou, KANG Kai. Bus trip OD matrix inference method based on IC information and probability theory [J]. *Highways & Automotive Applications*, 2010(3): 31.
- [24] 耿会灵. 基于轨迹与刷卡数据的公交运行可靠性评价[D]. 青岛: 山东大学, 2018.
- GENG Huiling. Bus operation reliability evaluation based on trajectory and smart card data [D]. Qingdao: Shandong University, 2018.
- [25] 尹安藤. 基于公交GPS和IC卡数据的公交OD推算[D]. 哈尔滨: 哈尔滨工业大学, 2017.
- YIN Anteng. OD matrix inference for bus trip based on GPS and IC card data [D]. Harbin: Harbin Institution of Technology, 2017.