

# 基于多源数据融合的高铁枢纽多模式 换乘客流分担率估计

马晓磊, 刘 兵, 姚李亮

(北京航空航天大学 交通科学与工程学院, 北京 102206)

**摘要:** 为实现高铁客运枢纽换乘客流分担率的精准辨识, 研究结合多模式公共交通大数据, 提出了一类基于广义出行链的高铁客运枢纽换乘方式选择模型。通过对不同公共交通方式换乘阶段的关联和融合, 提取以高铁客运枢纽为端点的个体广义出行链, 分析高铁客运枢纽换乘客流的时空分布特征; 综合考虑高铁客运枢纽乘客的个体经济社会属性、主观心理因素及其个体出行特征对换乘方式选择行为的影响, 并结合个体广义出行链的特征, 基于多指标多原因(MIMIC)模型和多元logit(MNL)模型提出高铁客运枢纽乘客换乘方式选择模型。以北京南站实际数据作为模型的输入, 得到了换乘客流方式分担率的估计值。比对分析估计值和真实值可知, 误差在可接受范围内。

**关键词:** 交通规划; 高铁枢纽换乘客流分担率; 多模式公共交通; 广义出行链; 换乘方式选择行为

中图分类号: U491

文献标志码: A

## Estimation Passenger Transfer Demand Multimodal Split in a High-Speed Railway Hub Based on Multi-Source Data Fusion

MA Xiaolei, LIU Bing, YAO Liliang

(School of Transportation Science and Engineering, Beihang University, Beijing 102206, China)

**Abstract:** In order to accurately identify travel mode split rate of transfer passenger flow in high-speed rail hubs, a travel mode choice model is proposed based on a generalized trip chain model based on multimodal public transportation big data. Through the correlation and fusion of different public transportation modes in the transfer phase, the individual generalized trip chain with the high-speed rail hub as the origin is extracted. Then, the temporal-spatial distribution characteristics of the

transfer flow in high-speed rail hubs are analyzed. Comprehensively considering the influence of individual economic and social attributes, and subjective psychological factors in combination with individual travel characteristics on transfer mode choice behavior, a transfer mode choice model for passengers in high-speed rail hubs is proposed based on multiple indicators and multiple causes (MIMIC) and multi-nominal logit (MNL) model. Taking the actual data of Beijing South Railway Station as the input, the estimated travel mode split rate of the commuting passenger flow is obtained. A comparison and analysis of ground true data indicates that the estimation error is within an acceptable range.

**Key words:** transportation planning; passenger flow split rate in high speed railway hubs; multimodal public transit; generalized trip chain; transfer travel mode choice behavior

随着我国高铁客运枢纽建设日渐完善, 各配套设施逐步开放, 枢纽区域的出行需求急剧增长, 高铁客运枢纽逐步成为衔接不同区域交通网络的关键节点。但当前我国高铁客运枢纽对复杂的客流时空分布特征适应能力较弱, 多模式公交之间无法根据客流实际需求执行高效协同的服务机制。同时由于缺乏系统科学的多模式换乘客流方式分担率估计的方法, 加剧了联合调度多模式交通接驳运力资源的难度。随着各种新兴技术的发展完善, 城市居民日常出行活动促使交通系统每天产生的多源海量出行数据, 为估计高铁枢纽换乘客流方式分担率提供了新的视角。因此, 结合多模式交通大数据, 提出高铁枢纽换乘客流方式分担率精准估计的方法对于提高高铁客运枢纽的服务水平、优化运力资源配置具有重

收稿日期: 2021-12-16

基金项目: 国家重点研发计划(2021YFB1600100); 北京市自然科学基金(8212010)

第一作者: 马晓磊(1985—), 男, 教授, 博士生导师, 工学博士, 主要研究方向为交通大数据分析与公共交通。

E-mail: xiaolei@buaa.edu.cn



论文  
拓展  
介绍

要意义。

针对乘客换乘出行方式分担率的研究,主要集中在衔接换乘行为影响因素分析和出行方式选择建模上面。如贾洪飞等<sup>[1]</sup>考虑出行时间、费用、乘客的性别、年龄等因素,基于MNL(multi-nominal logit)模型对乘客选择公交和小汽车的行为进行了研究。近年来,很多研究者逐渐意识到,乘客出行方式的选择受诸多无法直接观测的因素影响,主要为个体对出行方式的感知及其生活方式等内在的主观感受因素<sup>[2]</sup>。这些因素对出行者的出行方式选择行为有着显著的影响<sup>[3]</sup>。程龙等<sup>[4]</sup>、Paulssen等<sup>[5]</sup>、姚恩建等<sup>[6]</sup>、马书红等<sup>[7]</sup>在出行方式选择的研究中引入了心理潜变量,研究结果表明引入主观感受影响因素的出行方式选择模型预测结果更加精准、更具有解释能力。然而,上述研究主要借助问卷调查等人工获取的数据,基于离散选择模型对乘客出行方式选择行为进行研究,数据获取的质量较低、难度较大、样本量较小,研究的成果具有一定的局限性<sup>[8]</sup>。近年来随着大数据融合与挖掘算法的普及以及计算技术的不断发展,公交与地铁刷卡数据被广泛用于交通需求及交通方式辨识的研究中。马晓磊等<sup>[9]</sup>利用北京公交刷卡及GPS数据,基于贝叶斯决策树及马尔科夫链对公交乘客上车站点的推算进行了研究。周雨阳等<sup>[10]</sup>利用公交IC卡数据对长距离公共交通乘客出行的时空分布进行了研究,其成果对改善公交部分线路服务水平提供了有价值的参考建议。张郑等<sup>[11]</sup>基于RP(revealed preference)调查和广州地铁刷卡数据,对地铁乘客出行方式的衔接行为和路径选择行为进行了研究。Ma等<sup>[12]</sup>为识别公交乘客的出行模式,使用公交IC卡数据构建出行链,并利用聚类算法与粗糙集理论对出行模式进行聚类与分类,研究结果对认识交通需求在时空维度的分布具有重要意义。

综上所述,基于交通出行数据的交通需求辨识与公共交通协同换乘的研究,取得了一定成果。但同时使用多方式出行数据进行换乘方式选择行为辨识的研究,缺乏在时空域范围同时使用公交、地铁刷卡数据、出租车数据以及移动互联等多源异构数据,对个体出行链进行挖掘,尤其是以客运枢纽为起点的出行链的挖掘。

因此,本文通过纠正错误刷卡记录、修补异常站点信息等数据处理手段,完成不同公共交通方式换乘阶段的关联和融合,提取以高铁客运枢纽为端点的个体广义出行链,分析高铁客运枢纽城内换乘客

流的时空分布特征;并在此基础上,综合考虑高铁客运枢纽乘客的个体经济社会属性、出行特性及其主观心理因素对换乘出行方式选择行为的影响,建立基于个体广义出行链的换乘行为选择模型,最终估计换乘客流方式分担率,探索高铁客运枢纽乘客的换乘行为特性。

## 1 高铁枢纽乘客换乘行为建模

模型构建的整体思路为:结合公交、地铁、出租车等多模式交通出行数据,在数据清洗、推断和融合的基础上,提出个体广义出行链提取的方法;综合考虑高铁枢纽乘客的个体经济社会属性及其个体主观心理因素对换乘出行方式选择行为的影响,构建多指标多原因模型(MIMIC),研究影响因素的相互关系;并将换乘乘客的主观感受变量设为解释变量,结合个体经济社会属性信息和广义出行链的特征,建立高铁客运枢纽换乘方式混合选择模型。模型构建的框架如图1所示。

### 1.1 个体广义出行链提取

利用公共交通大数据和出租车轨迹数据,提取高铁枢纽换乘乘客的广义出行链,其目的在于挖掘乘客的个体出行特征,如出行时间、出行距离、出行费用等,用于高铁枢纽乘客换乘方式选择行为模型的构建。由于公共交通运营数据的割裂,不同公共交通出行方式产生的出行记录关联不高,现有研究通常只对公交-公交、地铁-地铁等相同交通方式间的换乘行为进行研究<sup>[13]</sup>。因此,本文通过深入剖析多源公共交通数据的采集规则,提出了公共交通大数据关联融合及个体广义出行链提取的方法。提取公共交通个体出行链的具体步骤如下。

#### 1.1.1 确定数据关联融合的方法

结合换乘乘客采取地铁、公交两类公共交通出行方式产生的出行记录数据,确定数据关联融合的方法。具体步骤为:首先分别对公交IC卡刷卡数据和地铁AFC数据中的编号编码规则进行分析,确定两类数据之间的关联列;其次,将关联数据中代表地铁、公交进出站时间的字段进行字符串提取,统一转换为标准时间格式。最后,根据两类数据的关联规则,选取相同字段,将两类刷卡数据进行合并,并增加出行方式类型标识,完善出行链出行方式属性信息。

#### 1.1.2 确定换乘时空阈值

本文考虑高铁枢纽换乘乘客采用公共交通的三

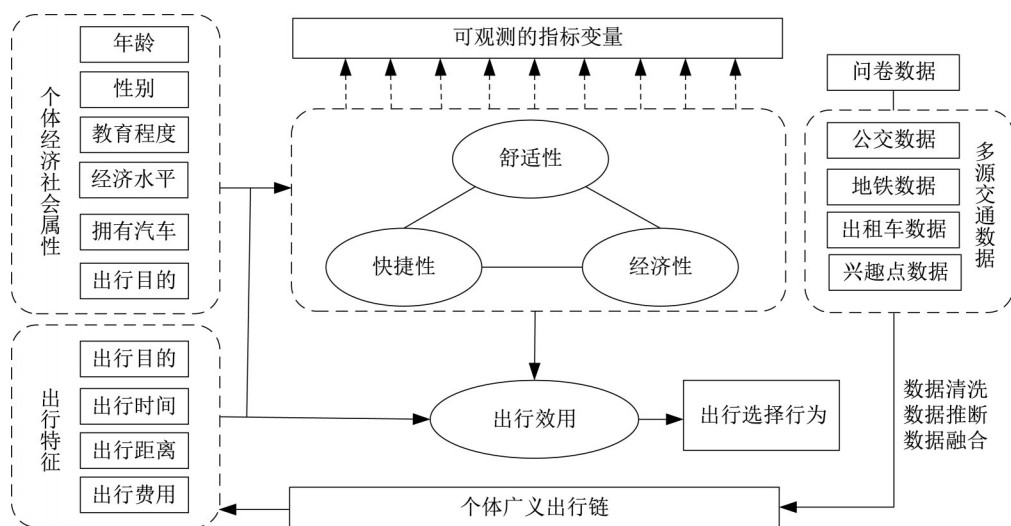


图1 换乘选择模型构建的框架

Fig.1 Framework of travel mode choice model

种换乘模式:公交-公交(B-B)、公交-地铁(B-M)、地铁-公交(M-B),并假设乘客以步行方式完成换乘,以此对乘客换乘出行的过程进行简化。考虑换乘时间和换乘距离对乘客换乘方式选择行为的影响,首先对换乘空间阈值和换乘时间阈值进行定义。

换乘空间阈值是指相邻两阶段换乘的站间距,以500 m作为空间限制;换乘时间阈值是指相邻两阶段换乘乘客所能接受的最大换乘时间。若两阶段之间的站间距、换乘时间超过了换乘空间或时间阈值,则判定为两次出行;若两阶段的时空距离处于阈值范围内,则认为两阶段之间乘客并未进行其他活动,属于一次出行,即属于同一条公共交通出行链。

研究针对3种换乘模式分别给出换乘时间阈值的计算方法。B-M模式的换乘时间阈值根据公交下车刷卡时间与地铁进站刷卡时间的时间差来确定,M-B模式的换乘时间阈值根据地铁出站刷卡时间与公交上车刷卡时间的时间差来确定,B-B模式的换乘时间阈值根据公交下车刷卡时间与下一阶段公交上车刷卡时间的时间差来确定<sup>[14]</sup>。

### 1.1.3 提取全域范围内的公共交通出行链

结合上述数据融合的规则及定义的两类换乘阈值,利用大数据处理软件,即可提取得到单一乘客的出行链。通过对换乘乘客出行过程的复现,可验证提取规则和融合思路的合理性。

### 1.2 换乘出行方式选择影响因素解析

本文旨在研究高铁枢纽换乘乘客的出行方式选择行为,不考虑乘客的到站行为,即不考虑乘客到站时间对其出行时间的影响。本文将影响高铁枢纽换乘乘客出行方式选择行为的因素划分为可直接观测

度量的显变量和不可直接观测度量的潜变量。

显变量包括了个体经济、社会属性和出行链出行特征属性。其中,个体经济、社会属性主要包括年龄、性别、学历程度、是否拥有私家车、出行目的、出行方式等。出行链出行特征属性主要包括基于个体广义出行链提取得到的出行时间、出行费用、出行距离、出行方式、出行目的地等。

潜变量通常用于反映乘客出行方式选择行为的心理偏好,衡量出行者对出行过程的满意度及效用值。研究参考相关文献的研究成果<sup>[12]</sup>,以舒适性、快捷性和经济性作为换乘出行方式选择的潜变量因素。

潜变量模型和离散选择模型共同构成混合选择模型,如图2所示。

### 1.3 出行效用函数与出行分担率计算

单一的离散选择模型,如MNL, NL, R-MNL模型通常是将影响乘客换乘方式选择行为的心理因素转化为可直接量化的指标或采取问卷打分的形式对其进行量化,无法客观体现心理变量之间的内在联系。而多指标多原因(multiple indicators and multiple causes, MIMIC)模型通过构建潜变量与指标变量之间的关联关系,对乘客换乘方式选择行为的心理因素进行更细粒度的刻画<sup>[15]</sup>。因此,研究引入MIMIC和MNL模型进行高铁枢纽乘客换乘方式选择行为模型的构建。

#### 1.3.1 MIMIC模型的构建

MIMIC模型的优势在于不需要进行严格的前提假设,也不需要被各种条件严格约束<sup>[16]</sup>,即能够对个体经济、社会属性、个体出行特征和代表个体主观



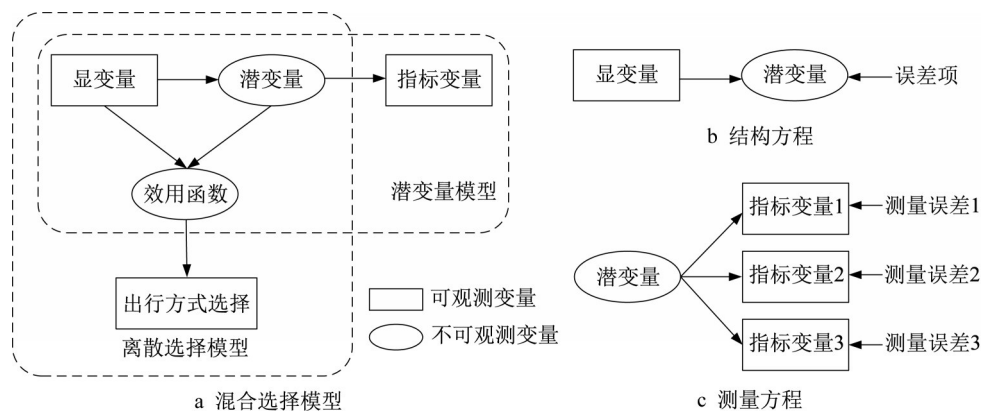


图2 混合选择模型结构示意图

Fig.2 Framework of hybrid choice model

心理感受的潜变量进行统一度量。MIMIC 模型中潜变量与指标变量之间的对应关系如表1所示。

表1 潜变量与指标变量的对应关系

Tab.1 Correspondence between latent variables and index variables

潜变量	选择偏好	指标变量
舒适性	选择换乘方式时, 很在意出行方式是否舒适	$f_1$
	愿意为更舒适的出行方式花费更多的时间或金钱	$f_2$
	很讨厌拥挤、劳累等不舒适的出行方式	$f_3$
快捷性	选择换乘方式时, 很在意出行方式是否快捷	$q_1$
	愿意为更快地到达目的地舍弃其他出行体验	$q_2$
	舍弃某种不够快捷(长时间换乘或堵车)出行方式	$q_3$
经济性	选择换乘方式时, 很在意出行方式是否经济	$e_1$
	即使要花更多时间也会选择更便宜的出行方式	$e_2$
	即使更不舒适也会选择更便宜的出行方式	$e_3$

MIMIC 模型之所以被称之为多因果多指标模型, 原因在于其测量方程部分实质为多指标模型。潜变量“经济性”的测量方程式如下:

$$e_1 = \lambda_{e1}\eta_e + \epsilon_{e1} \quad (1)$$

$$e_2 = \lambda_{e2}\eta_e + \epsilon_{e2} \quad (2)$$

$$e_3 = \lambda_{e3}\eta_e + \epsilon_{e3} \quad (3)$$

式(1)~(3)中: $\eta_e$ 为“经济性”潜变量; $\epsilon_{e1}$ 、 $\epsilon_{e2}$ 、 $\epsilon_{e3}$ 为测量方程的误差项; $\lambda$ 为因子载荷。

结构方程部分实质为多因果模型, 是对出行者个体经济、社会属性、出行特征等可观测的外生变量与潜变量之间的关系分析, 如下:

$$\eta_{fi} = \gamma_1 x_i + \gamma_2 x_i + \cdots + \gamma_n x_i + \zeta_i \quad (4)$$

式中: $\gamma$ 、 $\zeta_i$ 为待估计参数; $x$ 为年龄、性别等可观测的个体属性与出行特征; $i$ 为被调查个体。

由此, “舒适性”和“快捷性”的MIMIC模型如下:

$$f_i = \lambda_{fi}\eta_f + \epsilon_{fi} \quad (5)$$

$$f_2 = \lambda_{f2}\eta_f + \epsilon_{f2} \quad (6)$$

$$f_3 = \lambda_{f3}\eta_f + \epsilon_{f3} \quad (7)$$

$$\eta_{fi} = \gamma_1 x_i + \gamma_2 x_i + \cdots + \gamma_n x_i + \zeta_i \quad (8)$$

$$q_1 = \lambda_{q1}\eta_q + \epsilon_{q1} \quad (9)$$

$$q_2 = \lambda_{q2}\eta_q + \epsilon_{q2} \quad (10)$$

$$q_3 = \lambda_{q3}\eta_q + \epsilon_{q3} \quad (11)$$

$$\eta_{qi} = \gamma_1 x_i + \gamma_2 x_i + \cdots + \gamma_n x_i + \zeta_i \quad (12)$$

### 1.3.2 出行效用函数构建

本文将换乘乘客的出行链划分为简单出行链、复杂出行链两类。简单出行链是指乘客以高铁枢纽站为起点, 以一种出行方式(公交、地铁或出租车)完成一次出行的过程。复杂出行链是指乘客至少采用两种出行方式完成一次出行的过程。同时, 乘客的复杂出行链只考虑地铁和公交之间的换乘, 即乘客的整个过程仅借助公共交通完成。

每条出行链都可看作是一种出行方式选择方案。当乘客选择第 $j$ 种出行方式选择方案且为简单出行链时, 其效用函数 $U_j$ 可表示为

$$U_j = \sum_i a_{ji} X_{ji} + \sum_m b_{jm} S_{jm} + \sum_n c_{jn} \eta_{jn} + \epsilon_j \quad (13)$$

式中: $I$ 为出行者个人社会经济属性集合; $i$ 为个体经济社会属性; $M$ 为出行特征属性集合; $m$ 为出行特征属性; $N$ 为潜变量集合; $n$ 为潜变量特征属性; $a_{ji}$ 、 $b_{jm}$ 、 $c_{jn}$ 均为待估参数; $X_{ji}$ 为个体经济社会属性 $i$ 的量值; $S_{jm}$ 为第 $m$ 个出行链出行特征的量值; $\eta_{jn}$ 为潜变量; $\epsilon_j$ 为误差项。

当乘客选择第 $j$ 种出行方式选择方案且为复杂出行链时, 其效用函数 $U_{ik}$ 可表示为

$$U_{ik} = \sum_i a_{iki} X_{ki} + \sum_m b_{ikm} S_{km} + \sum_n c_{ikn} \eta_{kn} + \epsilon_{ik} \quad (14)$$

式中: $k$ 为不同的出行方式; $t$ 为复杂出行链的多个阶段(简单出行链 $t$ 为1,复合出行链 $t$ 为各阶段之和)。

### 1.3.3 出行分担率计算

根据效用最大化理论,出行者选择换乘出行方案 $j$ 的概率 $P_j$ 为

$$P_j = \frac{\exp(\sum_i^I a_{ji} X_{ji} + \sum_m^M b_{jm} S_{jm} + \sum_n^N c_{jn} \eta_{jn} + \epsilon_j)}{\sum_j \exp(\sum_i^I a_{ji} X_{ji} + \sum_m^M b_{jm} S_{jm} + \sum_n^N c_{jn} \eta_{jn} + \epsilon_j)} \quad (15)$$

表2 北京南站出行链数据提取结果

Tab.2 Trip chain extraction result at Beijing South Railway Station

日期	简单出行链			复合出行链	
	地铁出行/人次	公交出行/人次	出租车出行/车次	地铁-公交出行/人次	公交-地铁出行/人次
2018-10-10	41 440	7 223	2 233	5 509	183
2018-10-11	39 598	7 458	2 239	5 301	217
2018-10-12	47 240	7 311	2 089	6 174	261
2018-10-13	38 347	6 750	2 140	5 259	150
2018-10-14	54 241	7 325	2 276	8 111	97
2018-10-15	45 436	8 360	2 014	5 716	285

基于提取得到的个体公共交通出行链,进行乘客换乘方式选择行为及其出行特征分析。

#### 2.1.1 公交接驳客流特征

图3为北京南站出站乘客选择公交为换乘方式的群体出行目的地分布。宏观来看,选择公交出行的乘客出行目的地集中在北京中部(东城区、西城

## 2 数据处理与分析

### 2.1 多模式接驳客流特征分析

基于本文提出的个体广义出行链的提取方法,对已有数据进行处理并对以北京南站为出行链起点的数据进行筛选。在研究时间范围内,提取得到的有效出行链见表2。由表2可知,2018年10月10日至15日期间,共提取以北京南站及其附近站点为起点的公共交通个体广义出行链34.8万条(单位为人次),出租车出行链1.3万条(单位为车次)。

区)、南部区域(丰台区、大兴区、房山区),少量乘客通过乘坐公交的方式前往北城。主要原因在于,北京南站周边的公交站点所涉及的14条线路中,70%分布在北京南站南侧。选择公交出行的乘客很少会再通过换乘前往北部区域。

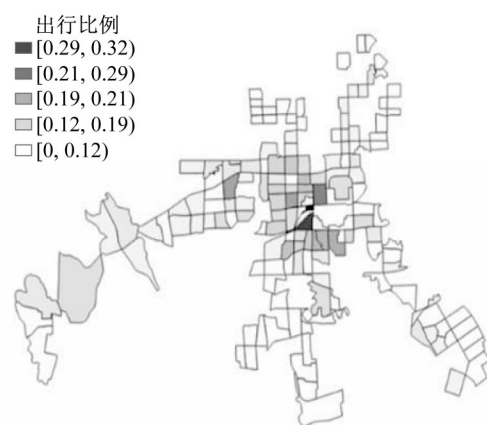


图3 公交接驳乘客出行热力图

Fig.3 Heatmap of bus transferring passengers

图4给出了公交接驳乘客的时间和距离分布。从图4a可以看出,出行距离在5 km以下的短途行程占比约为60.0%;距离在15 km以下的中短途行程占比约为86.0%,极少数南站乘客(2.3%)会在进行超远途市内出行时选择公交车作为换乘方式。从图4b可以看出,公交车刷卡人数,具有明显的早晚高

峰,分别为早上的7:00—9:00点和傍晚的16:00—18:00点。

#### 2.1.2 地铁客流特征

图5给出了北京南站出站乘客选择地铁为换乘方式的群体出行目的地分布。从图5中可以看出,与换乘公交的出行群体的目的地分布不同,地铁乘



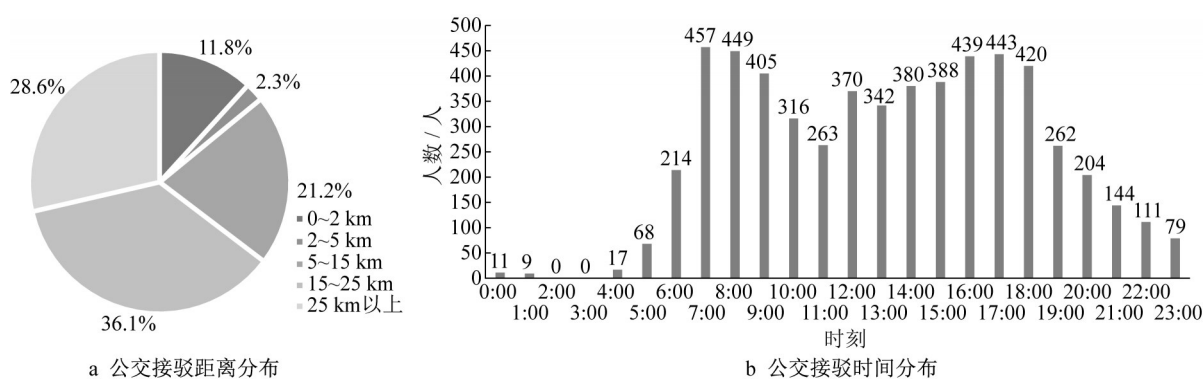


图 4 公交接驳乘客时间和距离分布

Fig.4 Trip time and distance distributions of bus transferring passengers



图 5 地铁接驳乘客出行热力图

Fig.5 Heatmap of subway transferring passengers

客的目的地分布更加广泛,目的地基本均匀分布在北京全域的地铁沿线上,前往北京北部区域的乘客数量与前往南部区域的乘客数量并无明显差别。

图 6 给出了地铁接驳乘客时间和距离分布。图 6a 所示,出行距离在 5 km 以下的短途行程占比仅为 7.1%;而距离在 5~15 km 以及 15~25 km 中长途行程占比分别约为 37.9% 和 43.3%,合计占比超过

80.0%;距离在 25 km 以上的超远途市内出行占比达到 11.7%,显著高于公交出行的 2.3%。从出行距离上看,出行距离远的乘客更偏好选择高速、换乘少、换乘快的地铁。图 6b 所示,采用地铁出行的客流同样具有明显的早晚高峰,高峰时段和公交出行的高峰时段基本一致,早高峰的集聚现象尤为显著。

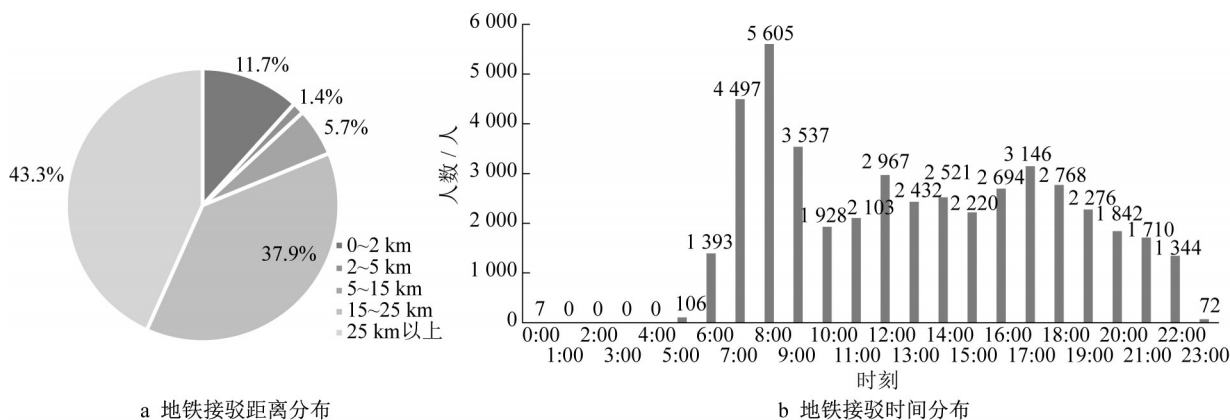


图 6 地铁接驳乘客时间和距离分布

Fig.6 Trip time and distance distributions of subway transferring passengers



### 2.1.3 出租车接驳客流时空特征

图7给出了北京南站出站乘客选择出租车为换乘方式的群体出行目的地分布。宏观上,与换乘公

共交通相比,出租车乘客的目的地分布更加广泛,遍布北京全域,且出行热力基本集中在中心城区。这和出租车灵活、便捷的特点有关。



图7 出租车接驳乘客出行热力图

Fig.7 Heatmap of taxi transferring passengers

图8给出了出租车接驳乘客时间和距离分布。从图8a可以看出,出行距离在5 km以下的短途行程占比约为7.1%;出行距离在5~15 km以及15~25 km中长途行程占比分别约为43.2%和33.1%,合计占比约76.0%,与地铁在此出行距离的行程占比相似;然而,出租车行程在超长途市内出行的占比超过16.0%,大于公交和地铁在此出行距离范围内出

行占比的总和。这说明长距离出行的乘客更倾向选择出租车出行。从整体上看,随着出行距离的增大,对应区域的出行热力变小。出租车在北京地区不同时段的载客车次数据分布如图8b所示。从图8b可以看出,采用出租车出行的客流同样具有明显的高峰,但此高峰主要集中在中午12:00—15:00点和夜晚20:00—24:00点。

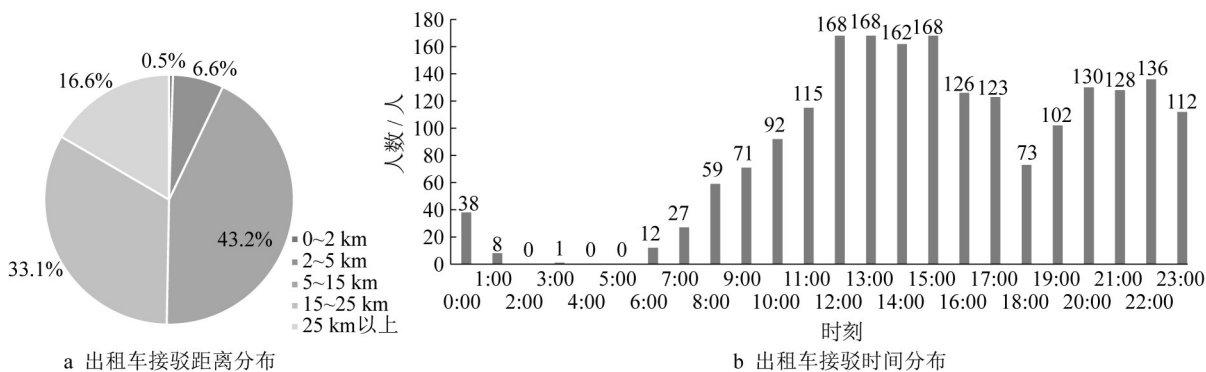


图8 出租车接驳乘客时间和距离分布

Fig.8 Trip time and distance distributions of taxi transferring passengers

图9给出了同一天出租车出行链和公共交通复杂出行链出行时间分布的对比。从图9可以看出,出租车接驳乘客数量的高峰期恰好是公共交通接驳乘客人次的低谷。其原因可能是因为不同出行时段公共交通的运力配置不同。

### 2.2 个人属性与出行意愿调查数据

考虑到列车到站时间和接驳客流的换乘出行时间对出行方式选择的影响,本文在不同日期的3个时段分别对前往北京南站的乘客进行问卷调查。调查

时段为8:00—12:00、13:00—17:00、19:00—24:00。调查问卷实际下发的数量为238份,其中,有效问卷201份,有效问卷回收率达85.0%。从年龄上看,被调查者中男性占比54.2%,女性占比45.8%;18以下群体占比8.9%,18~30岁群体占比45.8%,30~40岁群体占比22.4%,40~60岁群体占比18.9%,60岁以上群体占比4.0%。从受教育程度上看,高中及以下群体占比10.0%,大专及本科学历群体占56.2%,研究生群体占33.8%。从个体属性的角度上看,

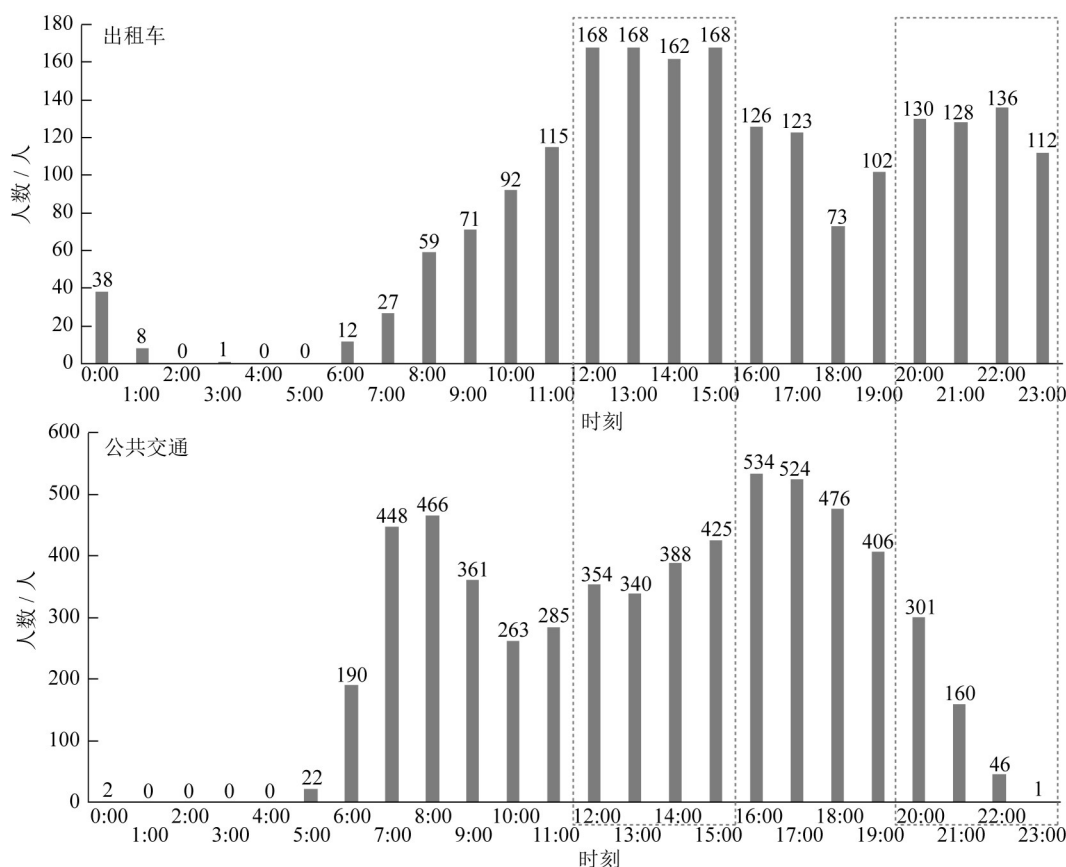


图9 出租车与公共交通接驳乘客时间分布对比

Fig.9 Comparison of trip time distributions between taxi and bus passengers

21.9%的被调查者拥有私家车,45.7%的群体来京目的为工作或公务出行等。调查样本的整体分布比较均匀,部分属性的分布略显集中,如被调查者的年轻群体占比较高、高学历群体占比较高。

### 3 接驳客流分担率估计结果分析

本文借助 AMOS 软件将问卷调查数据和出行链数据导入,运用最大似然估计进行模型估计。模型结构如图 10 所示,图中椭圆表示本文设置的潜变量“舒适性”“快捷性”和“经济性”;“性别”“年龄”“教育程度”“汽车出行”“出行目的”为代表个体经济社会属性的显变量。通过对观测变量和潜变量之间测量关系的分析,可得到潜变量的适配值表达式,进而可将潜变量和显变量同时作为解释变量构建 MNL 模型。图 10 中箭头上的数字为路径系数,用于表示变量之间的关系是否显著。在进行路径系数检验时,一般要求  $P$  值小于 0.05,否则代表两变量之间的影响不显著,需对路径进行删改。

#### 3.1 MIMIC 模型的参数估计

MIMIC 模型的拟合度评价指标如表 3 所示。从

表 3 可以看出,近似均方根误差(RMSEA)、拟合优度指数(GFI)、卡方值基本符合模型拟合度评价的要求,表明此模型和样本数据的适配情况良好。

表 4 为 MIMIC 模型潜变量和个体经济、社会属性的参数估计结果。由表 4 可得,教育程度越高的乘客对经济性的要求越低,对舒适性、快捷性的要求越高。本文以是否拥有私家车来衡量乘客的经济水平,拥有私家车的乘客(经济水平较高)对经济性的要求较低,符合假设预期。同时可以看出,出行目的对出行方式选择的影响显著,出行目的为公务出行的乘客,对经济性的要求显著更低,对舒适性和快捷性的要求更高。

表 5 为潜变量和指标变量的因子载荷系数。从表 5 中可以看出其系数均为较大的正值,由此说明选取的潜变量观测指标对潜变量有着显著的正向影响。

运用主成分分析法和凯撒正态化最大方差法,对成分矩阵(也称因子载荷矩阵)进行标准化求解,得到潜变量的适配值,见表 6。结果表明,本文引入的主观心理感受潜变量与个体经济、社会属性有显著关联,选取的可观测变量可较好地解释潜变量。



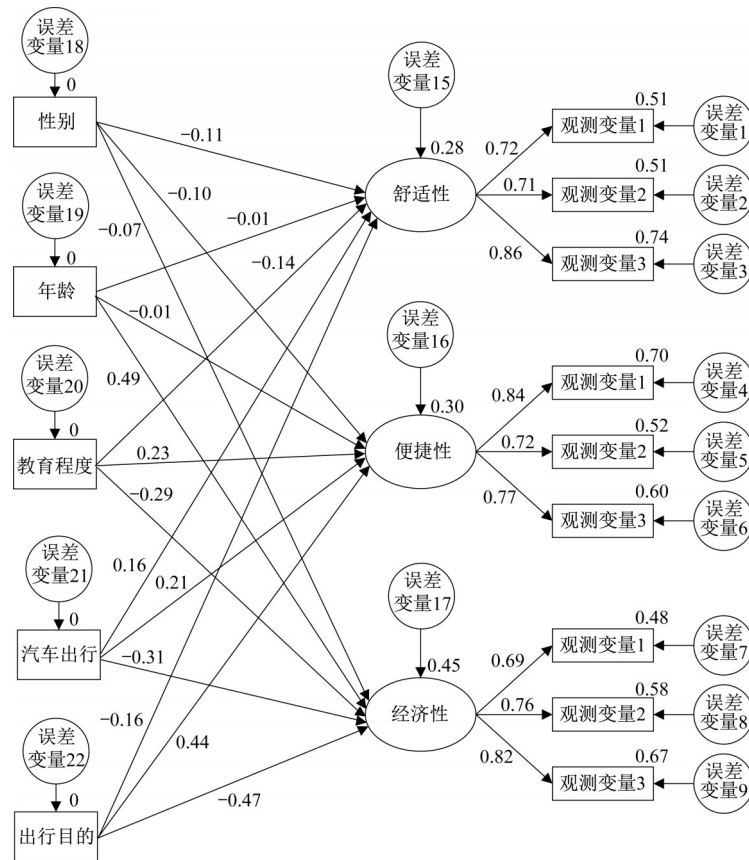


图10 换乘方式选择潜变量模型结构示意图

Fig.10 Structure of transfer mode choice model with latent variables

表3 结构方程模型拟合指数

Tab.3 Goodness of fit index for structural equation model

拟合度评价指标	$\chi^2/df$	RMSEA	GFI
参考值	<3	<0.08	>0.9
模型拟合值	2.072	0.068	0.914

表4 潜变量参数估计结果

Tab.4 Parameter estimation results of latent variables

名称	变量	$\eta_f$	$\eta_q$	$\eta_e$
性别	gender			-0.154*
年龄	age			0.162*
教育程度	education	0.145*	0.226***	-0.287***
是否有私家车	car		0.214***	-0.313***
出行目的	purpose	0.486**	0.441***	-0.471***

注:“\*”表示 $P$ 值小于0.05;“\*\*”表示 $P$ 值小于0.01;“\*\*\*”表示 $P$ 值小于0.001。

### 3.2 换乘出行方式标定结果

将个体经济、社会属性数据和广义出行链数据进行关联后,使用SPSS软件对混合选择模型进行求解,得到模型参数的标定结果,见表7。由表7可知,出行目的对乘客出行方式选择行为的影响显著,表明公务出行的群体更倾向选择出租车出行;出行时

表5 模型指标变量参数估计结果

Tab.5 Parameter estimation results of index variables

观测变量	$\eta_f$	$\eta_q$	$\eta_e$
$f_1$	0.716		
$f_2$	0.712		
$f_3$	0.86		
$q_1$		0.835	
$q_2$		0.724	
$q_3$		0.774	
$e_1$			0.691
$e_2$			0.761
$e_3$			0.818

表6 潜变量适配值表达式的成分矩阵及结果

Tab.6 Component matrix and results of latent variable adaptation expression

观测变量	成分			适配值
	1	2	3	
$f_1$	-0.047	0.385	0.029	0.385
$f_2$	-0.019	0.388	0.057	0.388
$f_3$	-0.087	0.446	0.052	0.446
$q_1$	0.457	-0.126	0.075	0.457
$q_2$	0.381	0.014	0.104	0.381
$q_3$	0.387	-0.013	0.082	0.387
$e_1$	0.205	0.011	0.500	0.500
$e_2$	0.029	0.049	0.398	0.398
$e_3$	-0.014	0.081	0.398	0.398

间对选择出租车换乘选择行为也有显著正向影响,出行时刻越晚的出行者更倾向于选择出租车出行;在出行费用方面,出行费用高的倾向选择出租车出行,出行费用低的更倾向于选择公共交通出行,与实际情况相符。

### 3.3 出行分担率估计结果

将上述模型标定的结果代入离散选择模型,各

种换乘出行方式的分担率计算结果见表8。为了便于对比分析,表中对一次换乘的复杂出行链和简单出行链进行了合并,并省略了选择私家车、网约车以及其他出行方式的分担率。

从表8中可以看出,预测结果和真实值的绝对误差基本在3%以内,由此可知模型估计的精度较高。

表7 换乘混合选择模型参数标定结果

Tab.7 Parameter calibration results of transfer hybrid choice model

类别	特征变量	换乘出行方案选择模型参数				
		地铁	公交	出租车	地铁—公交	公交—地铁
个人经济社会属性	经济水平	-0.723	-0.598	-1.166	-0.684	-0.623
	性别		-0.061	0.214		-0.112
	年龄		0.347			0.102
	受教育程度	0.113	-0.184	0.216	-0.122	-0.345
	是否拥有私家车	0.063	-0.173	0.331	-0.185	-0.214
出行特征	目的地	-0.212		0.352	-0.059	
	出行时刻		-0.284	0.314		-0.166
	距离	0.298		0.231	0.684	0.103
	成本	-0.371	-0.554	0.659	-0.243	-0.302
潜变量	$\eta_f$	0.231	0.124	0.444	0.263	-0.064
	$\eta_q$	0.329	0.041	0.628	-0.241	-0.469
	$\eta_e$	-0.231	-0.164	-0.455	-0.284	-0.063

表8 北京南站换乘选择方式出行分担率估计结果

Tab.8 Estimated results of mode splits at beijing south railway station

日期	公交分担率%			地铁分担率%			出租车分担率%		
	估计值	真实值	误差值	估计值	真实值	误差值	估计值	真实值	误差值
10月10日	20.1	22.7	-2.6	47.2	50.0	-2.8	15.2	13.1	2.1
10月11日	23.1	24.4	-1.3	46.4	49.7	-3.3	13.9	13.6	0.3
10月12日	19.8	22.3	-2.5	53.4	54.6	-1.2	15.6	11.7	3.9
10月13日	20.1	22.8	-2.7	49.3	50.1	-0.8	15.4	13.5	1.9
10月14日	15.2	19.4	-3.8	59.9	56.9	3.0	12.6	11.4	1.2
10月15日	22.3	25.0	-2.7	53.7	51.0	2.4	13.3	11.1	2.2

## 4 结论

传统对综合交通枢纽出行者换乘选择行为的推断,大多基于个体问卷或单一行程数据,得到的结果精度较低,应用范围受限。因此,本文基于多模式交通数据,提出了一种综合考虑出行者心理潜变量、个体经济、社会属性以及个体出行链出行特性的换乘方式选择行为模型,对高铁客运枢纽接驳客流的出行分担率进行了估计。主要结论如下:

(1)该模型的本质是MIMIC模型和MNL模型结合的混合选择模型,经样本数据拟合的结果可以看出,考虑的潜变量(舒适性、快捷性、经济性)对出行者换乘出行方式的选择有显著影响。

(2)通过对上下层模型参数估计结果进行分析,

得到了个体特征、心理潜变量和出行方式之间的相互影响关系。其中,从潜变量模型标定的结果可以看出,经济性要求较低的出行者更倾向选择出租车出行,更不倾向选择复杂出行链出行,且对舒适度要求较高;从离散选择模型的标定结果可以看出,出发时间越晚或出行费用越高的出行者越倾向于选择出租车出行。

(3)将参数估计结果和样本数据带入下层模型,估计得到高铁枢纽换乘乘客的出行方式分担率,通过与真实值对比,可以看出估计结果和真实值的绝对误差在3%以内,在可接受范围内。

尽管本文对丰富的多源交通数据进行了较为充分的应用,但由于不同交通方式之间的使用记录存在数据壁垒,如对乘坐公交后打车或骑乘共享单车

的乘客无法识别,使得部分出行链的构建还不够完整,同时对高铁客运枢纽换乘方式分担率的研究也仅限于公共交通和出租车。因此后续研究将进一步考虑私家车、共享单车等出行数据,以丰富乘客广义出行链的构成,提高换乘方式分担率估计的精度。

#### 作者贡献声明:

马晓磊:研究方案构思,整体思路设计,算法、数据处理指导,文稿修改。

刘 兵:研究成果整理,论文书写,方法梳理,算法设计。

姚李亮:算法设计,模型实现,数据处理与分析。

#### 参考文献:

- [1] 贾洪飞,孙宝凤,王苏娟,等.基于计划行为理论的换乘决策机理解析[J].吉林大学学报(工学版),2012,42:1481.  
JIA Hongfei, SUN Baofeng, WANG Sujuan, *et al.* Analysis of transfer decision mechanism based on planned behavior theory [J]. Journal of Jilin University (Engineering and Echnology), 2012, 42: 1481.
- [2] SRINIVASAN K K, MAHMASSANI H S. Analyzing heterogeneity and unobserved structural effects in route-switching behavior under ATS: a dynamic kernel logit formulation[J]. Transportation Research Part B: Methodological, 2003, 37: 793.
- [3] DE VOS J, SINGLETON P A, GARLING T. From attitude to satisfaction: introducing the travel mode choice cycle [J]. Transport Reviews, 2021, 42: 1.
- [4] 程龙,陈学武,杨硕.基于态度-行为模型的低收入通勤者出行方式选择[J].交通运输系统工程与信息,2016,16:176.  
CHENG Long, CHEN Xuewu, YANG Shuo. Travel mode choice of low-income commuters based on attitudinal behavior model [J]. Transportation System Engineering and Information Technology, 2016, 16: 176.
- [5] PAULSSEN M, TEMME D, VIJ A, *et al.* Values, attitudes and travel behavior: a hierarchical latent variable mixed logit model of travel mode choice[J]. Transportation, 2014, 41: 873.
- [6] 姚恩建,陈伟迪,卢天伟,等.考虑出行者选择偏好的出行方式选择模型[J].北京交通大学学报,2020,44:42.  
YAO Enjian, CHEN Weidi, LU Tianwei, *et al.* Travel mode choice model considering traveler's choice preference [J]. Beijing Jiaotong University, 2020, 44: 42.
- [7] 马书红,李阳,岳敏.考虑出行链的城际乘客换乘选择行为研究[J].北京交通大学学报,2020,44:74.  
MA Shuhong, LI Yang, YUE Min. Research on intercity passenger transfer choice behavior considering trip chain [J]. Beijing Jiaotong University, 2020, 44: 74.
- [8] LEE D, DERRIBLE S, PEREIRA F C. Comparison of four types of artificial neural network and a multinomial logit model for travel mode choice modeling [J]. Transportation Research Record, 2018, 2672: 101.
- [9] 马晓磊,刘从从,刘剑锋,等.基于公交IC卡数据的上车站点推算研究[J].交通运输系统工程与信息,2015,15:78.  
MA Xiaolei, LIU Congcong, LIU Jianfeng, *et al.* Research on bus boarding station based on IC card cata [J]. Traffic And Transportation Systems Engineering and Information Technology, 2015, 15: 78.
- [10] 周雨阳,赵倩阳,陈艳艳,等.基于IC卡数据的长距离公交客流特征分析[J].重庆交通大学学报(自然科学版),2018,37:80.  
ZHOU Yuyang, ZHAO Qianyang, CHEN Yanyan, *et al.* Analysis of long-distance bus passenger flow characteristics based on IC card data [J]. Journal of Chongqing Jiaotong University (Natural Science), 2018, 37: 80.
- [11] 张政,陈艳艳,梁天闻.基于网约车数据的城市区域出行时空特征识别与预测研究[J].交通运输系统工程与信息,2020,20:89.  
ZHANG Zheng, CHEN Yanyan, LIANG Tianwen. Research on spatial and temporal characteristics of urban travel based on ride-hailing data [J]. Transportation Systems Engineering and Information, 2020, 20: 89.
- [12] MA X, LIU C, WEN H, *et al.* Understanding commuting patterns using transit smart card data [J]. Journal of Transport Geography, 2017, 58: 135.
- [13] CHANG X, WU J, HE Z, *et al.* Understanding user's travel behavior and city region functions from station-free shared bike usage data [J]. Transportation Research Part F: Traffic Psychology and Behaviour, 2020, 72: 81.
- [14] XING Y, WANG K, LU J J. Exploring travel patterns and trip purposes of dockless bike-sharing by analyzing massive bike-sharing data in Shanghai, China [J]. Journal of Transport Geography, 2020, 87: 102787.
- [15] 靳海涛,金凤君,陈卓,等.基于换乘链断裂点时空信息的公交换乘行为识别[J].交通运输工程学报,2018,18:176.  
JIN Haitao, JIN Fengjun, CHEN Zhuo, *et al.* Identification of bus transfer behavior based on spatio-temporal information of transfer chain breaking point [J]. Journal of Traffic and Transportation Engineering, 2018, 18: 176.
- [16] 王俊兵.基于出行链的公交乘客出行特征分析[D].北京:北京交通大学,2017.  
WANG Junbing. Travel characteristics analysis of bus passengers based on travel chain [D]. Beijing: Beijing Jiaotong University, 2017.