

doi: 10.3969/j.issn.1672-6073.2025.01.013

# 基于多源数据的城市轨道交通 乘客路径选择估计方法

朱 国<sup>1</sup>, 章澜岚<sup>1</sup>, 刘家俊<sup>2</sup>, 杨皓帆<sup>2</sup>, 殷立超<sup>2</sup>, 张 宁<sup>2</sup>, 张恒文<sup>3</sup>

(1. 南京熊猫信息产业有限公司, 南京 210008; 2. 东南大学 智能运输系统研究中心轨道交通研究所, 南京 210018; 3. 中车城市交通规划设计研究院有限公司, 南京 211800)

**摘要:** 城市轨道交通的成网运行, 带来了客流路径选择的多样性。依据现有自动售检票系统(AFC)交易数据和概率推断方法, 难以准确获知乘客的网络路径选择, 完成轨道交通网络客流分配和票务清分等工作。利用网络站点信息构建城市轨道交通拓扑网络, 搜寻 OD 对可行路径集合, 并利用 AFC 交易数据、手机信令数据、列车时刻表数据等多源数据构建非线性优化模型, 推断乘客出行的路径选择。基于南京地铁网络进行验证, 结果表明: 模型具有有效性和鲁棒性。本研究可以为城市轨道交通运营及票务清分提供指导。

**关键词:** 城市轨道交通网络; 自动售检票系统; 多源数据; 出行路径选择; 非线性优化  
中图分类号: U231 文献标志码: A 文章编号: 1672-6073(2025)01-0099-07

## Estimation Method of Urban Rail Transit Passenger Route Choices Based on Multi-source Data

ZHU Guo<sup>1</sup>, ZHANG Lanlan<sup>1</sup>, LIU Jiajun<sup>2</sup>, YANG Haofan<sup>2</sup>, YIN Lichao<sup>2</sup>, ZHANG Ning<sup>2</sup>, ZHANG Hengwen<sup>3</sup>

(1. Nanjing Panda Information Industry Co., Ltd., Nanjing 210008; 2. ITS Rail Transit Research Institute of Southeast University, Nanjing 210018; 3. CRRC P&amp;C Institute Co., Ltd., Nanjing 211800)

**Abstract:** The network operation of urban rail transit has introduced diversity in passenger route choices. It is difficult to accurately determine passengers' network route choices based on existing Automatic Fare Collection (AFC) transaction data and probabilistic inference methods. This difficulty affects tasks such as rail transit network passenger flow allocation and ticket clearing. This study utilizes network station information to construct an urban rail topology network. The proposed method searches for feasible path sets for Origin-Destination (OD) pairs and uses multi-source data, including AFC transaction data, mobile signaling data, and train schedule data, to build a nonlinear optimization model to infer passengers' travel route choices. Experiments based on the Nanjing Metro network show that the model is effective and robust. This study can provide guidance for urban rail transit operations and ticket clearing.

**Keywords:** urban rail transit network; automatic fare collection system; multi-source data; travel route choice; nonlinear optimization

## 0 引言

随着城市化进程的加速, 城市交通拥堵问题日益突出, 城市轨道交通作为缓解拥堵的重要手段, 规模

与网络化程度不断提升。乘客路径选择估计作为票务清分、客流分析的关键, 传统方法如交通调查, 耗时耗力且数据质量难以保证; 而新兴技术如 5G、移动支付等使得更加多元的大数据能够被城市轨道交通系统

收稿日期: 2024-03-02 修回日期: 2024-10-16

第一作者: 朱国, 男, 高级工程师, 从事大数据平台研究, guo.zhu@panda.cn

通信作者: 张宁, 男, 副教授, 从事城市轨道交通运营管理工作, 331061643@qq.com

基金项目: 中国城市轨道交通协会重大专项(CAMET-KY-2022007); 国家重点研发计划资助项目(2020YFB1600701)

引用格式: 朱国, 章澜岚, 刘家俊, 等. 基于多源数据的城市轨道交通乘客路径选择估计方法[J]. 都市轨道交通, 2025, 38(1): 99-105.  
ZHU Guo, ZHANG Lanlan, LIU Jiajun, et al. Estimation method of urban rail transit passenger route choices based on multi-source data[J]. Urban rapid rail transit, 2025, 38(1): 99-105.

采集, 这些数据可以作为分析乘客出行选择行为的有力支撑, 是进一步研究更加准确、高效的乘客路径选择估计方法的基础。

目前, 有关城市轨道交通乘客路径选择估计方法的研究主要包括基于概率确定技术的方法<sup>[1]</sup>和基于用户轨迹重构的方法<sup>[2]</sup>两类, 其中, 基于出行轨迹重构的方法是指通过对乘客的不完整出行信息进行分析并与交通网络的已有信息进行匹配, 从而推断出其在交通网络中的路径选择, 其数据基础主要是手机信令数据<sup>[3]</sup>。目前手机信令数据在应用于地铁系统进行用户出行轨迹重构时的主要难点在于信令数据存在定位失准, 以及如何利用 OD 对间有效路径上部分手机用户的选择推断所有乘客的路径选择的问题。

本文首先建立城市轨道交通网络物理模型并构建有效路径搜索算法, 然后基于手机信令数据, 通过路径相似性计算完成乘客种子路径选择概率估计, 最后

综合利用 AFC 数据、列车时刻表数据构建基于 OD 对乘客旅行时间分布和站点出站客流时间分布的乘客路径选择概率扩算系数计算方法, 得到轨道交通系统中所有乘客的路径选择概率。

## 1 城市轨道交通拓网络模型

### 1.1 城市轨道交通网络拓扑分析

“拓扑”指将实体抽象成与其大小、形状无关的“点”, 把连接实体的线路抽象成“线”, 表示这些“点”“线”之间关系的图即为拓扑结构图。城市轨道交通网络是由地铁站点和轨道线路构成的网状拓扑结构<sup>[4]</sup>。

本文将轨道交通网络中的地铁站点视为拓扑结构中的“点”, 站点间的通行线路视为“线”, 在换乘站设置虚拟站台作为两条线路的中间“点”, 将换乘线路作为连接换乘站内部“点”的“线”, 如图 1 所示。

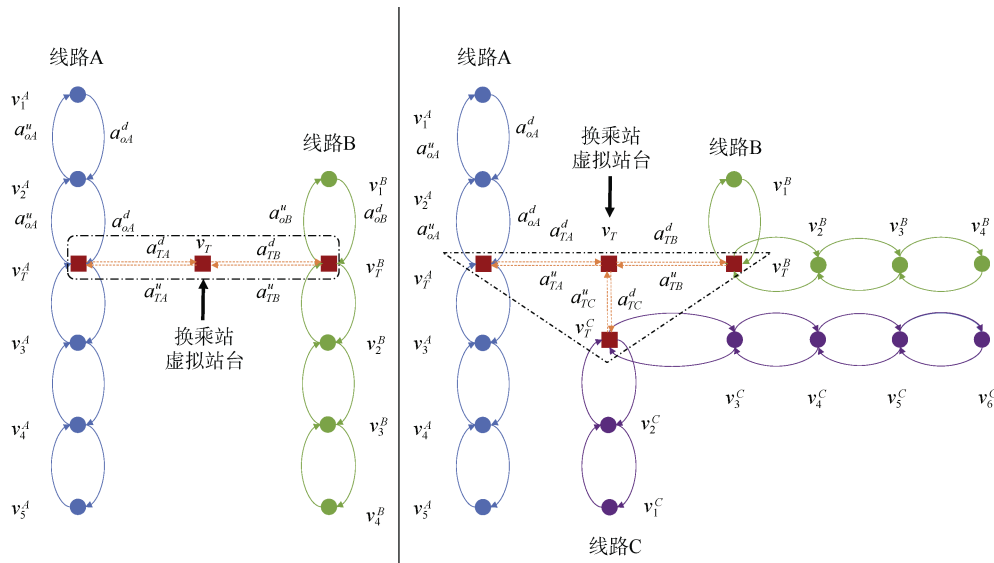


图 1 城市轨道交通网络拓扑图

Figure 1 Topology diagram of urban rail transit network

网络中节点的连通性和距离可表示为

$$d_{ij} = \begin{cases} a_{ij} & c_{ij} = 1 \\ +\infty & c_{ij} = 0 \end{cases} \quad (1)$$

式中,  $d_{ij}$  为站点  $i, j$  的有向连边的长度;  $a_{ij}$  为站点间的距离;  $c_{ij}$  为 0-1 变量; 若  $i, j$  直接连通,  $c_{ij}=1$ , 反之,  $c_{ij}=0$ 。由此得到的拓扑网络结构式为

$$G(V, A) = \phi(V, A, M_c, M_d) \quad (2)$$

式中,  $V$  为网络节点集合;  $A$  为网络连边集合;  $M_c$  为网络中各个节点的邻接矩阵;  $M_d$  为节点间的距离矩阵;  $\phi$  为拓扑网络中的“点”与“线”构成的有向图的

映射关系。

### 1.2 城市轨道交通网络模型构建

城市轨道交通网络由“网”“线”“站”3 个层次构成, 其要素之间的关系为

$$\begin{cases} V_{l_j} = \{v_{l_j}^1, v_{l_j}^2, \dots, v_{l_j}^n\} & v_{l_j}^i \in V \\ L = \{l_1, l_2, \dots, l_k\} \\ V = \bigcup_{j=1}^k V_{l_j}, l_j \in L \end{cases} \quad (3)$$

式中,  $L$  为网络线路集合;  $k$  为网络线路数量;  $i, j$  关

别为站点序号和线路序号。

由式(3)各层次要素间的关系以及网络拓扑有向图  $G(V, A)$  可以得到城市轨道交通网络模型为

$$\mathcal{N} = \{\delta(L, V), G(V, A)\} \quad (4)$$

式中,  $\mathcal{N}$  表示城市轨道交通网络;  $\delta(L, V)$  为式(3)中网络要素间的空间逻辑映射关系。

### 1.3 网络 OD 对间的可行路径搜寻

#### 1.3.1 广度优先搜索算法

广度优先搜索算法<sup>[5]</sup>是连通图的一种遍历策略,其基本思想是从图中的一个顶点开始,辐射状地优先遍历其周围较广的区域,它并不考虑结果的可能位置,彻底地搜索整张图,直至找到结果为止。算法的基本流程为,从根节点开始,沿着树(图)的宽度遍历树(图)的节点。如果所有节点均被访问,则算法终止。

#### 1.3.2 路径筛选及可行路径集构建

在城市轨道交通网络中,乘客出行的路径选择不仅考虑路径的出行时间或距离,也会考虑出行的便利程度。因此,在构建网络 OD 对的可行路径集时,需要同时考虑路径的出行时间或距离和换乘次数。在本研究中,通过设置路径的最大换乘次数剔除乘客不可能选择的路径。因此,已知需要研究的 OD 信息,即可通过此方法得到相应的可行路径集,进而可以针对这些路径的选择概率进行估计。

## 2 基于手机信令数据的种子路径选择概率估计

### 2.1 手机信令数据特征分析

手机信令数据主要是由手机在蜂窝小区的基站间切换所产生的,城市轨道交通网络中产生的信令数据主要包括移动台识别码(mobile station ID, MSID)、时间戳(timestamp)、移动基站编码(base station ID, BSID)以及信令事件编码(event ID)等<sup>[6]</sup>。

城市轨道交通移动通信网络是城市移动通信网络的一部分,该网络通过专有基站小区进行覆盖,并采集城市轨道交通网络中乘客出行过程中产生的手机信令数据<sup>[7]</sup>。在基站的布设和分布上具有以下特征<sup>[8]</sup>: ①轨道交通站点间的基站覆盖区域交叉范围小; ②地面基站和地下基站之间不会产生交叉干扰; ③手机用户在乘坐列车经过轨道交通线路轨行区时,可以通过位置更新和信令数据进行识别。

### 2.2 移动用户出行轨迹重构

通过分析手机用户的信令数据,可以了解乘客在

城轨交通网络中的出行情况<sup>[9]</sup>。首先,将信令基站与地铁站的经纬度信息进行匹配,通过比较两者之间的距离和相似度等指标,可以找到与该乘客信号强度最匹配的地铁站<sup>[10]</sup>,从而确定乘客所在的位置。之后通过查询网络线路站点关系字典了解乘客在网络中的出行轨迹。将乘客在不同时间段内的位置信息进行整合和分析,可以重构出乘客在网络中的出行轨迹。

### 2.3 路径相似性计算及乘客路径匹配

#### 2.3.1 路径时空相似因子定义

为了实现乘客的正确路径分配,引入路径相似因子  $CF_p$  评估基于手机信令数据重构后的出行轨迹与已知的 OD 对间可行路径之间的相似性,将乘客分配至与其重构轨迹相似度最高的路径上。

$$CF_p = \lambda_1 CF_s + \lambda_2 CF_t$$

$$\text{s.t.} \quad \lambda_1 + \lambda_2 = 1, \lambda_1, \lambda_2 \in [0, 1] \quad (5)$$

式中,  $CF_s$ 、 $CF_t$  分别为重构出行轨迹与 OD 对  $\omega$  间可行路径的空间和时间相似度,具体表达式如式(6)和式(7)所示;  $\lambda_1$ 、 $\lambda_2$  分别为空间和时间相似度的对应的权重系数。

$$CF_s = \ln \sum_{\pi_x, q_\omega \in R(i, j)} \left( \frac{L_{\pi_x, q_\omega}}{L_{\pi_x} L_{q_\omega}} \right)^\gamma \quad (6)$$

$$CF_t = \ln \sum_{\pi_x, q_\omega \in R(i, j)} \left( \frac{T_{\pi_x, q_\omega}}{T_{\pi_x} T_{q_\omega}} \right)^\eta \quad (7)$$

式中,  $\pi_x$  为在路径  $\pi$  上乘客  $x$  重构出行轨迹;  $R(i, j)$  为站点  $i$  和站点  $j$  之间的可行路径集合;  $q_\omega$  为 OD 对  $\omega$  间的可行路径;  $L_{\pi_x, q_\omega}$  为重构乘客出行轨迹与可行路径间的空间区间交集的距离;  $L_{\pi_x}$ 、 $L_{q_\omega}$  分别为重构乘客出行轨迹和可行路径的距离;  $T_{\pi_x, q_\omega}$  为重构乘客路径的出行时间区间与 OD 对  $\omega$  间可行路径上列车运行时间区间的交集的时间长度;  $T_{\pi_x}$  为乘客出行时间区间长度;  $T_{q_\omega}$  为可行路径上列车运行时间区间长度;  $\gamma$  和  $\eta$  为正常数。

#### 2.3.2 重构出行轨迹与 OD 对间可行路径匹配

通过计算移动用户的出行轨迹与其对应 OD 之间可行路径的相似度因子,将乘客匹配到最合适的路径上,从而确定移动用户在网络中选择的出行路径。通过进一步统计不同路径上的移动用户乘客人数,计算出在不同 OD 对之间选择路径的概率<sup>[11]</sup>,即

$$P_{r_j} = \frac{n_j}{n_\omega} \quad (8)$$

式中,  $n_j$  为移动用户乘客在 OD 对  $\omega$  选择第  $j$  条可行路径的人数;  $n_\omega$  为 OD 对  $\omega$  上的移动用户乘客人数。

### 3 基于 AFC 和列车时刻表数据的种子路径选择概率扩算系数计算

基于信令数据的乘客出行的种子选择概率只能代表部分用户的选择概率。因此, 基于自动售检票系统 (AFC)<sup>[12]</sup> 数据及列车时刻表数据对已有的种子概率进行扩算, 推测网络中所有乘客的路径选择概率。

#### 3.1 OD 对旅行时间分布估计

对于某个 OD 对, 可能有多条可行路径, 首先需要对路径的旅行时间分布进行估计。某个路径的旅行时间  $T_r$  主要包括进站时间  $T_o$ 、车内时间  $T_r$ 、出站时间  $T_d$ , 若有换乘行为, 则包含换乘时间  $T_x$ 。假设每个组成部分都服从独立的高斯分布, 由独立高斯分布的可加性可知, 乘客的路径旅行时间分布也服从高斯分布。

为估计上述高斯分布中各组成部分的参数, 首先依据路径的空间特征, 将乘客出行路径分类为无换乘单路径、单换乘单路径、多换乘单路径、多路径 4 类, 以便对路径时间的组成进行分解分析。 $T_o$  和  $T_d$  可以通过将无换乘单路径出行的乘客 AFC 数据与列车时刻表匹配得到, 随后再利用单换乘单路径出行的乘客 AFC 数据, 根据推测出的站点  $T_o$  和  $T_d$  信息, 进行乘客与列车的匹配, 根据乘客上下车时间得到  $T_x$ , 可以结合列车时刻数据得到  $T_r$ 。将 4 个分布相加即可推断路径的旅行时间分布, 进行 OD 对旅行时间分布和站点出站客流的时间分布估计<sup>[13]</sup>。

以第 2.3.2 节中计算的 OD 对间网络移动用户路径选择概率  $p_{r_j}$  作为种子选择概率矩阵, 可得 OD 对的旅行时间分布的高斯混合模型表达式为

$$p_\omega(t) = \sum_{j=1}^J p_{r_j} N(\mu_{r_j}, \sigma_{r_j}^2) \quad (9)$$

式中,  $j$  为 OD 对  $\omega$  间可行路径的编号;  $J$  为 OD 对间的可行路径数;  $N(\mu_{r_j}, \sigma_{r_j}^2)$  为路径  $r_j$  的旅行时间分布。

#### 3.2 站点出站客流时间分布估计

同样利用 OD 对间网络移动用户路径选择概率  $p_{r_j}$  计算某时间段  $\Delta t$  内的 OD 对  $\omega$  中路径  $r_j$  的乘客人数, 即

$$F(r_j, \omega, I_k) = p_{r_j} \times D_\omega(I_k) \quad (10)$$

式中,  $D_\omega(I_k)$  为时间间隔  $I_k$  内在站点  $v_i$  进站的人数,  $I_k$  长度即为  $\Delta t$ 。

式(11)用来表示在时间段  $h$  内从站点  $v_i$  进入地铁网络的乘客选择路径  $r_j$ , 并在时间段  $t$  内从站点  $v_j$  出站的概率。其中参数的选取如图 2 所示,  $t_+$ 、 $t_-$  为起讫点站时间轴上  $t'_+$  或  $t'_-$  所在区间的前后端点。

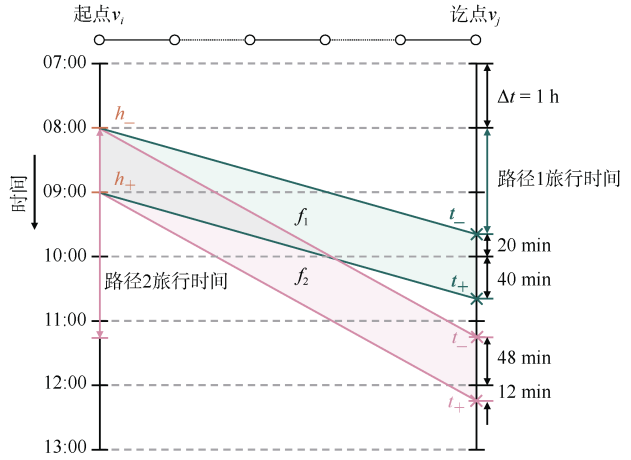


图 2 某多路径 OD 乘客旅行路径时空示意  
Figure 2 Spatiotemporal schematic of multi-path OD passenger travel routes

$$m_{jt,ih}^r = \begin{cases} (t'_+ - t_-) / \Delta t, t_- < t'_+ < t_+ \\ (t_+ - t'_-) / \Delta t, t_- < t'_- < t_+ \\ 0, \text{其他} \end{cases} \quad (11)$$

随后, 可以计算出时间间隔  $I_k$  内站点  $v_i$  到  $v_j$  的估计出站客流量  $E_\omega^*(F, t)$ , 即

$$E_\omega^*(F, t) = \sum_{r \in r_\omega} \sum_{h \leq t} m_{jt,ih}^r \alpha_{r_j} \times F(r_j, \omega, h) \quad (12)$$

式中,  $\alpha_{r_j}$  为 OD 对  $\omega$  之间路径  $r_j$  的种子选择概率的修正比例系数。

#### 3.3 融合旅行时间和出站客流的双目标优化模型

使用 3.1 小节和 3.2 小节所得的 OD 对旅行时间分布以及出站客流的时间分布, 可以计算 OD 对  $\omega$  之间路径  $r_j$  的种子选择概率的修正比例系数  $\alpha_{r_j}$ 。OD 对的估计旅行时间分布  $\hat{\beta}(t)$ 、出站客流的时间分布  $E_\omega^*(F, I_k)$  可分别表示为

$$\hat{\beta}(t) = \sum_{j=1}^J \alpha_{r_j} p_{r_j} N(\mu_{r_j}, \sigma_{r_j}^2), j \in J \quad (13)$$

$$E_\omega^*(F, I_k) = \sum_{r \in r_\omega} \sum_{h \leq t} m_{jt,ih}^r \alpha_{r_j} p_{r_j} \times D_\omega(I_k) \quad (14)$$

为了得到最优的修正比例系数以使估计路径选择

概率最接近真实值, 本文构建估计和观测的 OD 对旅行时间分布、出站客流的时间分布之间的综合差异最小化的符合概率约束的优化模型。其中, 旅行时间分布的差异使用 KL 散度量, 而出站客流的时间分布使用均方根误差(RMSE)计算。KL 散度及 RMSE 的计算公式分别为

$$D_{KL}[\beta(t) \parallel \hat{\beta}(t)] = \sum_{k=1}^K \beta(I_k) \lg \frac{\beta(I_k)}{\hat{\beta}(I_k)} \quad (15)$$

$$RMSE(E_\omega, E_\omega^*) = \sqrt{\frac{1}{n} \sum_{k=1}^K (E_\omega(I_k) - E_\omega^*(I_k))^2} \quad (16)$$

式中,  $\beta(t)$  为 OD 对观测旅行时间分布, 将其按照固定时间间隔离散为  $K$  个区间;  $E_\omega(I_k)$  为时间区间  $I_k$  内出站客流量观测值。随后, 计算估计和观测的旅行时间分布和出站客流的时间分布之间的综合差异最小化。使用 min-max 归一化方法, 分别计算两者的最小值, 再使用二者与各自的最小值的差与本指标最大最小值之差的比之和作为目标函数, 即

$$D_{KL-\min} = \min D_{KL}[\beta(t) \parallel \hat{\beta}(t)] \quad (17)$$

$$RMSE_{\min} = \min RMSE(E_\omega, E_\omega^*) \quad (18)$$

$$\min \omega_1 \frac{D_{KL}[\beta(t) \parallel \hat{\beta}(t)] - D_{KL-\min}}{D_{KL-\max} - D_{KL-\min}} + \omega_2 \frac{RMSE(E_\omega, E_\omega^*) - RMSE_{\min}}{RMSE_{\max} - RMSE_{\min}} \quad (19)$$

$$\text{s.t. } \omega_1 + \omega_2 = 1, 0 \leq \alpha_{r_j} p_{r_j} \leq 1, \sum_{j=1}^J \alpha_{r_j} p_{r_j} = 1, j \in J$$

式中,  $\omega_1$  和  $\omega_2$  分别为旅行时间分布和出站客流的时间分布所占综合差异度的权重, 由旅行时间和出站客流对模型的影响程度确定。使用 scipy 中的序列最小二乘法(SLSQP), 求解使得优化目标函数最小时的参数, 即  $\alpha_{r_j}$ 。

## 4 案例分析

截至 2022 年底, 南京地铁已开通地铁线路 12 条, 总长 449 km, 换车站共 16 座, 构成覆盖南京全市 11 个市辖区的地铁网络<sup>[14]</sup>。本文选择小市站到新街口站与新街口站到小市站作为研究的 OD, 由于两 OD 的可行路径在行程时间上较为接近, 仅根据行程时间难以对乘客路径选择进行判断, 在实际的路径选择判断过程中具有典型性。

图 3 中表示了这两个 OD 的 4 条可行路径。

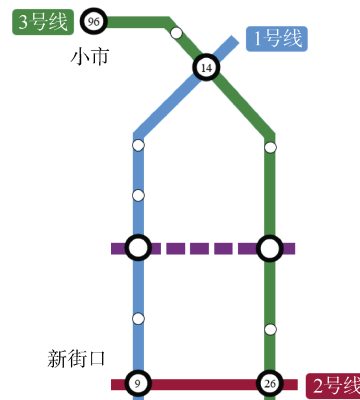


图 3 实验网络示意

Figure 3 Experimental network

### 4.1 基于合成 AFC 数据的模型有效性验证

由于乘客的真实路径选择情况无法通过确切数据得到, 因此本文基于乘客出行链逻辑, 合成 AFC 数据, 使用这些数据对模型有效性进行验证。将小市站到新街口站作为 OD 对 1, 将新街口站到小市站作为 OD 对 2。OD 对 1 中两条路径为 Path\_1\_1: 96→26→9、Path\_1\_2: 96→14→9。OD 对 2 中两条路径为 Path\_2\_1: 9→26→96、Path\_2\_2: 9→14→96。按照如表 1 所示的参数设置生成路径时间的估计分布, 如 3.1 节所述,  $T_o$  与  $T_d$  可利用真实 AFC 数据得到, 而  $T_x$  和  $T_r$  分别通过列车时刻表和时间约束计算得到。路径种子选择概率根据手机用户轨迹重构与路径匹配得到, 所采用的手机信令数据来自中国移动运营商, 渗透率为 67%。然后, 基于合成的 AFC 数据估计 OD 对旅行时间分布, 进而估计路径选择概率。由于路径真实选择概率无法被观测, 验证实验中通过预先定义的真实值作为基准, 将估计的概率与路径真实选择概率进行比较以验证模型的性能。

表 1 合成数据参数设置

Table 1 Synthetic data parameter settings

路径	参数	$T_o/s$	$T_r/s$	$T_x/s$	$T_d/s$
Path_1_1	$\mu$	144.85	960	245.55	102.46
	$\sigma$	96.29	30	264.76	112.20
Path_1_2	$\mu$	144.85	840	408.51	102.46
	$\sigma$	96.29	30	237.04	112.20
Path_2_1	$\mu$	136.69	948	294.63	108.50
	$\sigma$	88.82	30	215.40	95.15
Path_2_2	$\mu$	136.69	825	467.29	108.50
	$\sigma$	88.82	30	250.80	95.15

本实验共生成 2 000 条 AFC 数据, 得到两个 OD 4 条路径的扩算系数、估计选择概率和真实选择概率, 如表 2 所示。结果表明估计误差均在 5% 以内, 模型具有有效性。

表 2 实验结果

Table 2 Experimental results

路径	种子选择 概率 $p_{r_j}$	扩算系数 $\alpha_{r_j}$	估计选择 概率 $p'_{r_j}$	真实选择 概率 $r_j$	相对 误差/%
Path_1_1	0.600	1.169	0.701	0.700	0.143
Path_1_2	0.400	0.746	0.299	0.300	0.333
Path_2_1	0.800	0.737	0.590	0.600	1.667
Path_2_2	0.200	2.050	0.410	0.400	2.500

### 4.2 模型敏感性分析

通过模拟估计偏差对模型性能的影响评估模型敏感性, 包括旅行时间和样本数量两个方面。并在此过程中将传统的基于 KL 散度的单目标优化方法<sup>[4]</sup>与本文提出的结合 KL 散度和出站客流分布的双目标优化模型进行对比。

行程时间方面, 将偏差添加到 OD 对乘客路径旅行时间的标准差中以模拟有偏差的旅行时间, 取偏差百分数  $\rho = \pm 10\%, \pm 20\%, \pm 30\%$ , 修改后的标准差  $\sigma^* = \sigma \times (1 + \rho)$ 。生成带偏差的 AFC 数据, 模型估计得到的选择概率与预设的真实选择概率的相对误差如图 4 所示 (展示 Path\_1\_1, Path\_2\_1)。结果表明: 当偏置标准偏差达到真实值的 30% 时, 本模型对偏置输入具有鲁棒性, 相对误差均在 10% 以内。在实验中, 双目标优化的

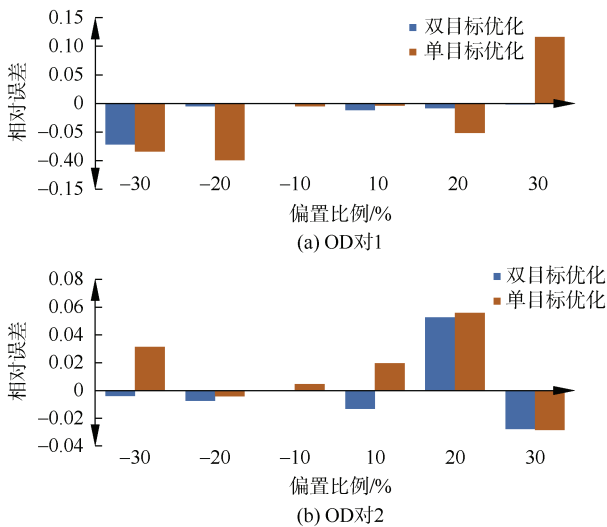


图 4 旅行时间偏差下模型估计误差

Figure 4 Model estimation error under travel time bias

实验结果均优于或等于单目标优化, 且从图 4 中可以看出, 双目标优化对于数据偏置的敏感程度低于单目标优化, 稳定性更佳。

输入不同数据量, 模拟样本数量偏差。分别输入 500、1 000、3 000、5 000 和 10 000 条数据, 合成数据的模型估计选择概率如图 5 所示。结果表明: 当样本量不同时, 本模型具有鲁棒性, 相对误差均在 10% 以内。

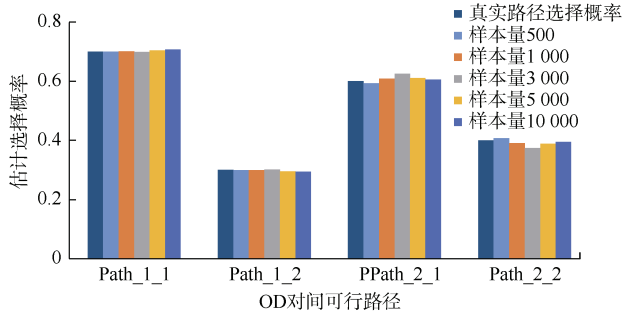


图 5 样本数量偏差下模型估计结果

Figure 5 Model estimation error under sample quantity bias

### 4.3 基于真实 AFC 数据的模型验证

在验证了模型有效性和鲁棒性的基础上, 本文利用南京地铁工作日和周末的 AFC 交易数据, 测试所建模型, 并与使用信令数据估计的路径选择估计先验分布进行对比。实验结果如图 6 所示, 估计和观测旅行时间分布的对比如图 7 所示。结果表明: 根据本文所提出的方法计算所得的乘客旅行时间参数、路径选择概率能较好地拟合真实情况下的乘客旅行时间分布, 表明本文提出的方法能较好地还原乘客在地铁线网中的出行情况。同时, 与先验分布相比, 根据本模型估计的路径选择概率计算的 OD 旅行时间分布更加吻合真实的 OD 旅行时间分布。

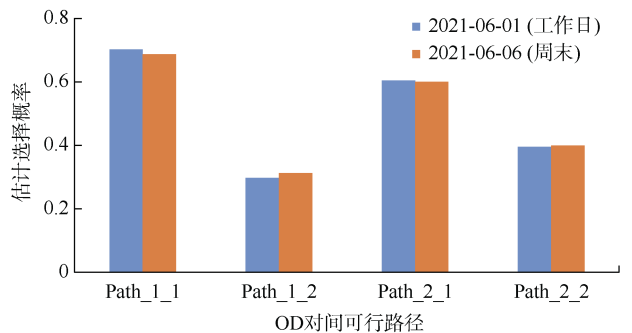


图 6 不同日期 AFC 数据估计路径选择概率

Figure 6 Path choice probability estimated by AFC data on different days

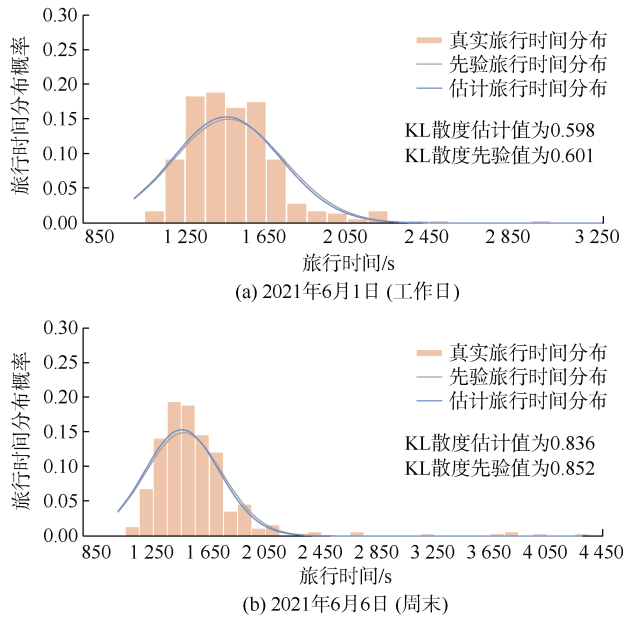


图7 工作日和周末估计和观测旅行时间分布

Figure 7 Estimated and observed travel time distributions for a weekday and a weekend.

### 5 结束语

本文通过城市轨道交通网络拓扑结构、地铁手机信令数据对乘客轨迹进行分析，并针对信令数据无法覆盖所有乘客的问题，通过构建融合旅行时间和出站客流的双目标优化模型求解扩算系数，实现轨道交通乘客路径选择估计，实验结果表明：本文所提的纯数据驱动的方法能够有效并准确地估计乘客在网络中的路径选择行为，并且在数据输入存在偏差时，模型具有稳定性。

#### 参考文献

[1] KUSAKABE T, ASAKURA Y. Behavioural data mining of transit smart card data: a data fusion approach[J]. Transportation research part C: emerging technologies, 2014, 46: 179-191.

[2] JO H H, KARSAI M, KARIKOSKI J, et al. Spatiotemporal correlations of handset-based service usages[J]. EPJ data science, 2012, 1(1): 10.

[3] LUO Qin, XU Tao, LIN Huazhen, et al. Modeling of passenger's time-space path estimation in urban rail transit network based on mobile signaling[C]//CICTP 2018. Beijing, China. American Society of Civil Engineers, 2018: 1076-1085.

[4] LIU J, MA Z, WEN L, et al. Urban rail transit fare reconciliation method using multi-source data[C]. Transportation Research Board 102th Annual Meeting, 2023.

[5] BUNDY A, WALLEN L. Breadth-first search[M]//Catalogue of Artificial Intelligence Tools. Berlin, Heidelberg: Springer Berlin Heidelberg, 1984: 13.

[6] 胡必松. 基于手机信令的铁路车站客流特征提取技术研究[J]. 铁道标准设计, 2022, 66(5): 145-151.  
HU Bisong. Research on passenger flow characteristics technology of railway station based on mobile phone signaling data[J]. Railway standard design, 2022, 66(5): 145-151.

[7] 方恒堃. 多元数据在北京轨道交通线网规划中的应用[J]. 都市快轨交通, 2021, 34(1): 37-44.  
FANG Hengkun. Application of big data in Beijing rail transit network planning[J]. Urban rapid rail transit, 2021, 34(1): 37-44.

[8] OLIVEIRA E M R, VIANA A C. From routine to network deployment for data offloading in metropolitan areas[C]//2014 Eleventh Annual IEEE International Conference on Sensing, Communication, and Networking (SECON). June 30-July 3, 2014, Singapore. IEEE, 2014: 126-134.

[9] ASAKURA Y, HATO E. Tracking survey for individual travel behaviour using mobile communication instruments[J]. Transportation research part C: emerging technologies, 2004, 12(3/4): 273-291.

[10] PU Yichao. Verification and optimization of metro fare clearing models based on travel route reconstruction[J]. New metro, 2020, 1(1): 34-47.

[11] 孙舒森, 赵嘉伟, 张宁. 轨道交通路网仿真中乘客出行路径的确定[J]. 铁路通信信号工程技术, 2020, 17(12): 45-50.  
SUN Shumiao, ZHAO Jiawei, ZHANG Ning. Determination of passenger travel path in rail transit network simulation[J]. Railway signaling & communication engineering, 2020, 17(12): 45-50.

[12] 张义鑫, 林磊, 张宁, 等. 基于云平台和微服务架构的城市轨道交通 AFC 系统[J]. 都市快轨交通, 2021, 34(6): 131-139.  
ZHANG Yixin, LIN Lei, ZHANG Ning, et al. AFC system of the urban rail transit cloud platform based on micro-service deployment[J]. Urban rapid rail transit, 2021, 34(6): 131-139.

[13] ZHAO Juanjuan, ZHANG Fan, TU Lai, et al. Estimation of passenger route choice pattern using smart card data for complex metro systems[J]. IEEE transactions on intelligent transportation systems, 2017, 18(4): 790-801.

[14] 蔡玉萍, 黄娟, 常炳阳, 等. 南京轨道交通与城市空间发展互动思考[J]. 都市快轨交通, 2023, 36(4): 6-12.  
CAI Yuping, HUANG Juan, CHANG Bingyang, et al. Reflections on the interaction between Nanjing rail transit and urban space development[J]. Urban rapid rail transit, 2023, 36(4): 6-12.

(编辑: 傅依萱)