

基于巢式 Logit 模型的城市轨道交通 客流分配问题研究

颜 研^{1,2}, 倪少权^{1,2}

(1. 西南交通大学 交通运输与物流学院, 四川 成都 610031;

2. 西南交通大学 全国铁路列车运行图编制研发培训中心, 四川 成都 610031)

摘要:传统的城市轨道交通网络客流分配方法使用了基于离散选择理论的多项 Logit 模型, 与其相比, 巢式 Logit 模型消除了影响因素的相关性, 具有更好的解释能力。介绍了巢式 Logit 模型的原理及公式, 以城市轨道交通网络中乘客的出行行为为研究对象, 将乘客选择过程分为换乘次数选择和换乘路径选择两个层次, 并以此为构架建立了城市轨道交通客流分配巢式 Logit 模型。通过实例, 利用 SAS 软件对模型进行求解和分析, 从概率角度用定量方法分析出行个体的出行选择行为和轨道交通网络流量分配, 对于提高城市轨道交通网络化运营组织水平、进行合理的运能配置和构建城市轨道交通票务清分系统意义重大。

关键词:无缝换乘; 随机效用; 巢式 Logit; 客流分配

中图分类号: U121 **文献标志码:** A **文章编号:** 2095-0373(2015)04-0099-05

0 引言

传统的换乘模式下, 乘客在轨道交通网络内从一条路径换乘到另一条路径, 需要重新刷卡出站刷卡进站, 这样, 进入到每条路径的客流量大小可直接通过刷卡量得到。而在无缝换乘条件下, 轨道交通内部只能记录乘客的第一次进站和最后一次出站, 无法确定乘客在轨道交通内部的换乘路径, 因此, 各路径的乘客流量大小也无法确定。研究在无缝换乘条件下城市轨道交通网络中的客流流量分配方法, 对于提高运输组织水平和客运服务水平、制订列车运行计划、进行合理的运能配置和运营成本控制意义重大。国内外针对城市轨道交通客流分配问题的研究已有较多的成果。

国内研究中, 刘剑锋等^[1]、四兵锋等^[2]分析了乘客在城市轨道交通网络中的路径选择行为, 应用多项 Logit 模型对城市轨道交通网络的客流分配问题建模; 吴祥云等^[3]借鉴道路系统的交通流量分配模型, 提出了轨道交通客流分配的均衡阻抗模型和基于 Frank-Wolf 算法的改进算法; 潘明浩^[4-5]提出拥挤条件下动态客流分配模型讨论在高峰拥堵情况下出行者的动态路径选择情况, 并用仿真的方法验证模型的有效性。目前国外对轨道交通客流分配的研究大部分借鉴了城市交通客流分配的基本理论和思想, Wardrop 提出了著名的道路网平衡的概念和定义, 为道路网客流量分配模型的建立奠定了基础; Beckmann^[6]提出了基于 Wardrop 平衡的交通分配数学规划模型; Lam 等提出了一种适用于拥挤公交网络的随机用户平衡分配模型, 并给出了相应算法。这些理论都被应用于解决城市轨道交通客流分配问题中。

本文在前人所研究多项 Logit 模型的基础上, 将巢式 Logit 模型应用于城市轨道交通网络客流分配中, 旨在克服多项 Logit 模型过大评价相似性高的选择方案的缺陷。

1 随机效用理论和巢式 Logit 模型

在城市轨道交通系统中, 为了反映决策者对各选择项满足需求的程度, 为轨道交通网络中的每一种

收稿日期: 2014-11-25 责任编辑: 刘宪福 DOI: 10.13319/j.cnki.sjztdxxb.2015.04.18

作者简介: 颜研(1990—), 女, 硕士研究生, 研究方向为交通运输规划与管理。E-mail: pyramid_z@foxmail.com

基金项目: 国家自然科学基金项目(61273242, 61403317); 中国铁路总公司科技研究计划项目(2013X006-A)

颜研, 倪少权. 基于巢式 Logit 模型的城市轨道交通客流分配问题研究[J]. 石家庄铁道大学学报: 自然科学版, 2015, 28(4): 99-103.

出行方案确定一个效用值。由于受出行者出行喜好等不可测因素的影响,面临选择项相同时个体不一定作出相同的决策,因此,各选择项的效用是一个随机变量。假定用 N 表示出行者在出行中可选择的出行方案集合,用 I 表示出行者的类型集合,则第 i 类型的出行者选择第 n 个出行方案的随机效用 U_n^i 可以表示为

$$u_n^i = V_n^i + \epsilon_n^i, n \in N, i \in I \quad (1)$$

式中, V_n^i 表示第 i 类型的出行者对于第 n 个出行方案可以确定的效用; ϵ_n^i 为随机误差项; β_n^i 为出行影响因素的待估参。

出行者选择某种出行方案的概率取决于随机误差项 ϵ_n^i 的分布形式以及出行者可以确定的出行效用 V_n^i 的取值。如果 ϵ_n^i 的随机分布已知,出行者的选择概率就可以计算出来。根据效用最大化原则,出行者选择某种出行方案的概率就是该方案的出行效用在所有可选方案中为最大的概率。这个概率的表示为

$$\left. \begin{aligned} p_n^i &= \Pr(U_n^i \geq U_p^i, p \neq n, p \in N), n \in N, i \in I \\ 0 \leq p_n^i \leq 1, \sum_{n \in N} p_n^i &= 1 \end{aligned} \right\} \quad (2)$$

MNL 模型有内在缺陷,最明显的就是 IIA 特性,即当选择枝之间存在相似性而仍然直接使用 Logit 模型的话,就会过高评价具有相似性的选择枝群,而错误地标定模型参数,导致预测偏差的问题。关于 IIA 特性,是 1959 年由研究者 Luce 发现,其中最典型的例子是红蓝巴士问题。MNL 模型存在 IIA 特性的根源在于 MNL 模型关于选择枝效用随机项完全独立结构的刚性假定^[8-9]。巢式 Logit 模型(Nested Logit Model)是 MNL 模型的一种改进形式,该模型设置了一个虚拟层,用于将一些具有类似性质的影响因素归为一类,并将其置于同一层,从而抽象为一个树状结构图,能够很好地克服 MNL 模型的 IIA 特性,如图 1。

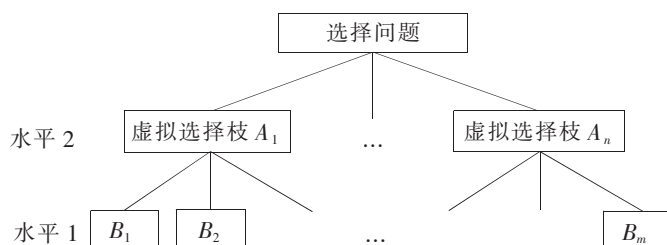


图 1 NL 树状结构图

在 NL 模型树状结构图中,上层水平 2 由虚拟选择枝 $A_1, \dots, A_n$ 组成;下层水平 1 由选择枝 $B_1, B_2, \dots, B_m$ 组成。假设出行方案是由水平 1 和水平 2 组成的选择集合,则 $U = \{(A_1 \times B_1), (A_1 \times B_2), \dots, (A_i \times B_j), \dots, (A_n \times B_m)\}$ 。设某出行方案 U 的效用为 $U(A_i, B_j)$,虚拟选择枝 A 的效用为 $U(A)$,在上层水平 2 中选择虚拟选择枝 A_i 的条件下,下层选择 B_j 时的效用为 $U(B_j | A_i)$,则有 $U(A_i, B_j) = U(A_i) + U(B_j | A_i)$ 。如果 $U(A_i)$ 和 $U(B_j | A_i)$ 均服从均值为 $V(A_i)$ 和 $V(B_j | A_i)$ 、方差为 σ_2^2 和 σ_1^2 的相互独立的概率分布,而且 $U(A_i)$ 和 $U(B_j | A_i)$ 是相互独立的,那么就有 $U(A_i, B_j) = V(A_i) + V(B_j | A_i) + \epsilon(A_i) + \epsilon(B_j | A_i)$ 。其中, $V(A_i)$ 和 $V(B_j | A_i)$ 分别表示层选择虚拟选择枝 A_i 的效用值确定项和上层选择 A_i 的条件下下层选择 B_j 时的效用值确定项; $\epsilon(A_i)$ 和 $\epsilon(B_j | A_i)$ 表示其相应的随机项,并且 $\epsilon(A_i)$ 和 $\epsilon(B_j | A_i)$ 相互独立,在其选择枝间服从 Gumbel 分布。

令 $P(B_j | A_i)$ 为上层选择虚拟选择枝 A_i 的条件下,下层选择 B_j 的概率; $P(A_i)$ 为上层选择虚拟选择枝 A_i 的概率。那么,由全概率公式得,同时选择上层虚拟选择枝 A_i 和下层选择枝 B_j 的概率 $P(A_i, B_j) = P(B_j | A_i) \cdot P(A_i)$,其公式表示为

$$P(A_i, B_j) = P(A_i) \cdot P(B_j | A_i) = \frac{\exp\{\lambda_2 [V(A_i) + V^*(A_i)]\}}{\sum_{i=1}^n \exp\{\lambda_2 [V(A_i) + V^*(A_i)]\}} \cdot \frac{\exp[\lambda_1 V(B_j | A_i)]}{\sum_{j=1}^m \exp[\lambda_1 V(B_j | A_i)]} \quad (3)$$

式中, λ_1 为只考虑下层效用的概率项方差 σ_1^2 相对应的参数, $\lambda_1^2 = \pi^2 / 6\sigma_1^2$; λ_2 为同时考虑上、下层效用的概率项方差相对应的参数, $\lambda_2^2 = \pi^2 / 6(\sigma_1^2 + \sigma_2^2)$; $U^*(A_i) = \max(V(B_j | A_i) + \epsilon(B_j | A_i))$ 表示合成效用,服从均

值为 $V^*(A_i) = \frac{1}{\lambda_1} \ln \sum_{j=1}^m \exp(\lambda_1 V(B_j | A_i))$, 方差为 σ_1^2 的 Gumbel 分布。

2 城市轨道交通客流分配巢式 Logit 模型的构建

2.1 模型假设

通过研究城市轨道交通乘客的路径选择行为,以随机效用理论为基础,结合城市轨道交通客流特点,对模型提出如下假设:出行者对路网十分熟悉,出行时都会选择出行效用最大的方案;出行者选择路径时的评判标准相同,以保证出行者有相同的出行效用函数表达形式;各选择路径效用函数的随机项相互独立且服从参数为(0,1)的 Gumbel 分布。

2.2 模型建立

通过分析乘客的出行行为过程,总结城市轨道交通路径选择的影响因素,其中,乘客自身属性因素有性别、年龄、职业和月收入水平;城市轨道交通影响因素有乘车时间、换乘时间、换乘次数和车内的拥挤度^[7-9]。如果将乘客自身属性因素和城市轨道交通影响因素所包含的所有内容放在一起考虑,形成一个综合的效用函数,就是我们通常所使用的多项 Logit 模型。如前文所述,多项 Logit 模型会过大评价类似性高的选择方案,当乘客的可选路径集中有两条甚至多条属性类似的路径,通过多项 Logit 模型计算出的路径选择概率会出现较大的偏差。因此,本文使用巢式 Logit 模型对乘客选择行为进行研究。

城市轨道交通路径选择的可选路径集中不同路径换乘次数相同的可能性最大,因此,本文将换乘次数作为一个虚拟层建立了巢式 Logit 模型。根据前人研究成果,城市轨道交通运营网络中所能接受的最大换乘次数为两次^[7],即虚拟层为:无换乘、一次换乘和两次换乘。

本文对任意假设的轨道交通网络中一个 O-D 的可选路径集进行研究,如图 2。

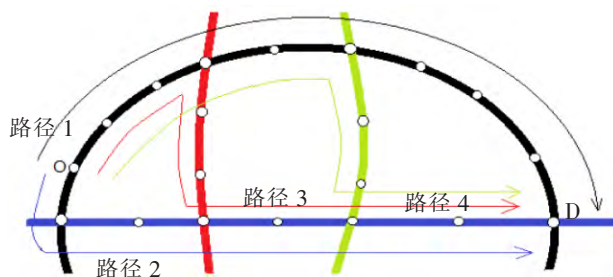


图 2 O-D 可选路径集

图 2 中的可选路径共四条,路径的属性为:路径 1,换乘次数 0,换乘时间 0 min,乘车时间为 32 min;路径 2,换乘次数 1,换乘时间 5 min,乘车时间为 22 min;路径 3,换乘次数 2,换乘时间 8 min,乘车时间为 18 min;路径 4,换乘次数 2,换乘时间 8 min,乘车时间为 22 min。

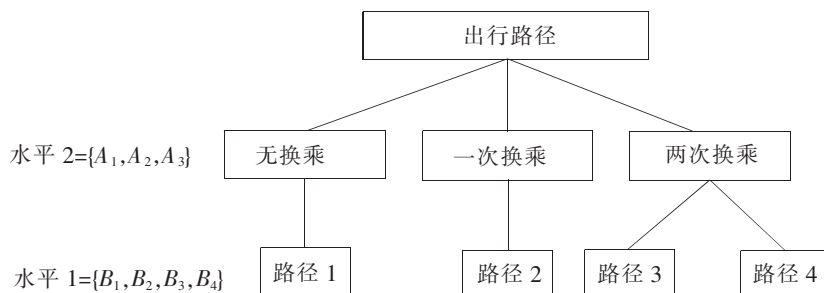


图 3 城市轨道交通乘客出行路径选择树

按照前文所述的巢式 Logit 模型分类原则,将换乘次数一致的路径分为一类,形成了城市轨道交通出行路径选择树,见图 3。影响选择树中水平 2 的因素为换乘次数,影响水平 1 的因素有性别、年龄、职业、月收入、换乘时间、乘车时间和可承受拥挤度。

2.3 模型求解

本文以成都市居民为样本对城市轨道交通出行者的路径选择行为进行问卷调查,调查内容包括出行者的性别、年龄、职业、月收入;出行者所选线路的换乘时间、乘车时间、换乘次数以及可承受拥挤度,各因素表示及取值情况见表 1。其中,出行者的性别、年龄、职业、月收入和可承受拥挤度可以通过问卷直接获得;换乘时间、乘车时间需要在通过问卷获得乘客选择的出行路径后结合路网数据和运行图数据计算得到;换乘次数可以直观地从路网数据得到。换乘时间和乘车时间的计算如下

$$E_i^{l,m} = w_i^{l,m} + 0.5 \cdot f_m, \quad \forall l, m, i \quad (4)$$

$$T_{ij}^l = \begin{cases} t_{ij}^l, & \text{站点 } i \text{ 为起始站,} \\ s_i^l + t_{ij}^l, & \text{否则} \end{cases} \quad \forall l, i, j \quad (5)$$

式中, $E_i^{l,m}$ 为乘客在换乘站 i 从线路 l 换乘到线路 m 的换乘时间; $w_i^{l,m}$ 表示乘客在换乘站 i 从线路 l 到线路 m 进行换乘的步行时间; f_m 表示线路 m 的列车平均发车间隔; T_{ij}^l 为乘客乘坐线路 l 在区间 (i, j) 的乘车时间; t_{ij}^l 为线路 l 的列车在区间 (i, j) 上的行驶时间; s_i^l 为线路 l 的列车在车站 i 的平均停站时间。

表 1 路径选择因素表示及取值

因素名称	变量	取值	待定参数
性别	x_{11}	男:1;女:2	β_1
年龄	x_{12}	1:<22;2:22~40;3:40~60;4:>60	β_2
职业	x_{13}	1:政府职员;2:企事业单位人员;3:学生;4:离退休人员;5:自由职业者;6:其他	β_3
月收入	x_{14}	1:<2 000;2:2 000~4 000;3:4 000~6 000;4:6 000~8 000;5:>8 000	β_4
换乘时间	x_{15}	统计计算得到,单位:min	β_5
乘车时间	x_{16}	统计计算得到,单位:min	β_6
可承受拥挤度	x_{17}	1:车厢内座位数还可容纳少数乘客;2:车厢内满座;3:车厢内少数人无座;4:车厢内很多人无座;5:车厢内几乎无法行走	β_7
换乘次数	x_{21}	统计得到,范围:0~2	β_8

模型参数估计过程复杂且计算量大,本文借助 SAS9.3 软件对模型参数进行估计,经过检验剔除掉对路径选择影响不大的影响因素后,得到巢式 Logit 模型各影响因素的偏回归系数表,见表 2。

表 2 剔除不相关因素后的巢式 Logit 模型参数估计和检验结果

Parameter	DF	Estimate	Standard Error	t Value	Approx Pr> t
age_L1	1	0.008 7	0.843 7	2.05	0.078 1
time1_L1	1	-0.232 5	0.046 1	-3.73	0.009 5
time2_L1	1	-0.228 3	0.117 8	-2.91	0.014 8
congest_L1	1	-0.135 4	0.241 2	-2.28	0.012 6
transfer_L2G1C1	1	-0.368 5	0.324	-3.81	0.008 3

根据参数估计和检验结果,age_L1、time1_L1、time2_L1、congest_L1 和 transfer_L2G1C1 通过 t 检验,即年龄、换乘时间、乘车时间、可承受拥挤度和换乘次数对乘客出行路径的选择产生较为显著的影响。所对应的参数估计值分别为: $\beta_2=0.008 7$, $\beta_5=-0.232 5$, $\beta_6=-0.228 3$, $\beta_7=-0.135 4$, $\beta_8=-0.368 5$ 。其中, β_2 的估计值为正值,结合到模型实例中,说明年龄偏大的乘客对路径选择的敏感度偏小;但是其绝对值很小,说明年龄对效用函数的贡献值不大。SAS9.3 同时可以计算出不同换乘次数的选择概率,在本案例中得到乘客选择路径 1-4 的概率为: $\bar{P}_1=27.7\%$, $\bar{P}_2=30.8\%$, $\bar{P}_3=29.2\%$, $\bar{P}_4=12.3\%$,这与问卷调查的结果 $P_1=28.1\%$, $P_2=32.2\%$, $P_3=29.6\%$, $P_4=10.1\%$ 十分接近。

2.4 模型对比

若使用 MNL 模型求解上述客流分配问题,即将路径选择的各影响因素:性别、年龄、职业、月收入、换乘时间、乘车时间、换乘次数和可承受拥挤度放在同一层考虑,使用 SAS9.3 软件进行参数标定和路径选择概率计算,计算后的结果为 $\bar{P}_1=23.5\%$, $\bar{P}_2=27.1\%$, $\bar{P}_3=27.8\%$, $\bar{P}_4=21.6\%$ 。通过对比发现,4 条路径被选取的概率均在 20% 以上,每条路径被选择的概率更趋于均匀化;路径 3 被选择的概率略高于路径 2 被选取的概率;路径 4 被选择的概率显著增加。这说明 MNL 模型过大评价了具有高类似性的路

径3和路径4。

3 结束语

该文简单介绍了随即效用理论和巢式 Logit 模型的公式推导,并应用于城市轨道交通客流分配问题研究中,将具有共性的出行路径选择方案放在一个巢内,消除了各影响因素之间的共性,较好地克服了多项 Logit 模型的 IIA 缺陷,具有借鉴意义。

参 考 文 献

- [1]刘剑锋,孙福亮,柏赞.城市轨道交通乘客路径选择模型及算法[J].交通运输系统工程与信息,2009(9):110-114.
- [2]四兵锋,毛保华,刘智丽.无缝换乘条件下城市轨道交通网络客流分配模型及算法[J].铁道学报,2007,29(6):52-54.
- [3]吴祥云,刘灿齐.轨道交通客流量均衡分配模型与算法[J].同济大学学报:自然科学版,2004(9):1158-1162.
- [4]Poon M H. A dynamic schedule-based model for congested transit networks[D]. China:University of Hong Kong,2001.
- [5]Poon M H,Wong S C,Tong C O. A dynamic schedule-based model for congested transit networks[J]. Transportation Research Part B. Methodological,2004(7):87-90.
- [6]Beckman M J. Studies in the Economics of Transportation[M]. American :Yale University Press ,1956.
- [7]刘剑锋.基于换乘的城市轨道交通网络流量分配建模及其实证研究[D].北京:北京交通大学,2012.
- [8]关宏志.非集计模型交通行为分析的工具[M].北京:人民交通出版社,2004.
- [9]四兵锋,高自友.交通运输网络流量分析与优化模型[M].北京:人民交通出版社,2013.
- [10]陆化普.交通规划理论与方法[M].北京:清华大学出版社,2006.

Nested Logit Model for the Traffic Assignment Problem of Urban Railway Transit

Yan Yan^{1,2}, Ni Shaoquan^{1,2}

(1. School of Traffic and Transportation, Southwest Jiaotong University, Chengdu 610031, China;

2. National Railway Train Diagram Research and Training Center, Southwest JiaoTong University, Chengdu 610031, China)

Abstract: Multinational logit model is usually used in the traffic assignment model of urban railway transit, which is based on discrete choice theory. Compared with it, nested logit model eliminates the influence of correlation factors and has a better explanation ability. This study introduces the principle and formula of the nested logit model. Through the analysis of travel choice behavior and travel utility, this paper establishes nested logit model for the traffic assignment problem of urban railway transit, which divides passengers' selection process into the selection of transfer times and the selection of transfer path. At last, this paper verifies this model by an example, which is carried out with SAS software. This study analyzes both quantitatively and qualitatively the individual travel choice behavior and the rail network traffic assignment from the angle of probability. It is of great significance to improving the level of operation organization, configuring reasonable capacity and building urban rail traffic ticket distribution system.

Key words: seamless transfer; discrete choice; nested logit; service framework; traffic assignment