

基于主成分分析法的城市轨道交通系统聚类分析

彭涌涛^{1,2}, 陈中治³, 赵尘¹, 王欣南³

(1. 南京林业大学, 江苏 南京 210037; 2. 南京交通职业技术学院; 3. 中交第二公路勘察设计研究院有限公司)

摘要:城市轨道交通系统科学分类是合理选择系统制式的基础, 该文针对目前定性分类居多的现状, 基于北京、上海、广州、深圳、南京、重庆、长沙等城市 77 条轨道交通线路的预测日均客流量、最高设计速度、线路规划长度、车站设置数量、平均车站间距、单节车厢载客量、车辆宽度 7 个方面的规划数据, 将主成分提取的 3 个公因子解释为运输能力、最高速度、车辆型号 3 个主要因子, 运用 k -means 聚类方法将城市轨道交通系统分为地铁、轻轨、市域快轨、市郊铁路、现代有轨电车和城市专用轨道系统 6 类制式, 分类结果与线路实际的功能定位基本吻合。并从服务区域、运量、速度等方面例举了各类制式的技术特征, 从定量分析角度验证了运量是城市轨道交通系统区别于其他公共交通方式的最显著的特征指标。

关键词:城市轨道交通; 系统制式; 主成分; 聚类

1 前言

从 1863 年世界上第一条地铁在英国伦敦诞生以来, 城市轨道交通随着城市化进程不断发展, 截止到 2018 年中国共有 35 个城市开通运营轨道交通线路 5 767 km。运营的系统制式也从单一的地铁逐渐丰富到现今的地铁、轻轨、单轨、市域快轨、市郊铁路、磁悬浮轨道交通、现代有轨电车、APM 等制式并存。系统制式的多样化为城市轨道交通选型决策提供了更多的可能, 根据城市现状选择合理可持续发展的轨道交通制式, 成为轨道交通规划建设中的关键一环。鉴于目前系统制式多样, 同一制式称谓不同、同一制式标准不同的现状, 国内外学者对系统制式的分类展开了系统研究, 以期城市轨道交通科学选型提供理论依据。

日本曾根悟将城市轨道交通分为有轨电车、市郊有轨电车、地下铁道、铁路电车和新交通系统等几种制式; 曾任德国杜塞尔多夫地区铁路公司总经理的 Georg Puettner 将欧洲大都市轨道交通系统分为轻轨 (LRT)、地铁、市域线 (S-Bahn)、区域线; 美国 Vuchic 对各种公共交通方式按线路专用程度、系统技术和运营方式等特点来进行分类和定义; 陈炎按照运量和速度两个维度, 将城市轨道交通系统划分为地铁、轻轨、有轨电车、市域快轨和特殊功能专线; 周

立新按照空间地理位置和服务区域特征将城市轨道交通线路分为市区线、市域线、市郊线、区域线, 提出“地域”与“运量技术等级”相结合的“二元”城市轨道交通系统分类思路; 魏庆朝等归纳分析了 CJJ/T 114—2007《城市公共交通分类标准》和《城市轨道交通工程项目建设标准》的分类方法, 建议从驱动、支承、导向方式将现有的城市轨道交通系统进行分类, 形成特征鲜明、直观的命名方法; 魏运等分析了城市都市区、市域、主城区、组团内的交通特征和需求, 认为《城市轨道交通工程项目建设标准》的分类解决了主城区轨道交通类型选择的问题, 城市其他区域轨道交通需求超出了该分类方法所覆盖的范围, 需要进一步细化; 孔令洋从服务区域、运输能力、旅行速度 3 个方面将城市轨道交通分为地铁系统、轻轨系统、市域快线系统、有轨电车系统和特殊专用系统; 廖亚莎从运输能力和服务区域两个方面将地铁、轻轨、单轨、有轨电车划分为市区线, 磁悬浮系统属于市域线, 市郊铁路属于市郊线, AGT、有轨电车属于局域线; 李世元根据相关城市轨道交通标准、行业标准及目前中国城市轨道交通运用情况, 对目前中国城市轨道交通制式进行梳理, 提出轨道交通系统分为地铁系统、轻轨系统、有轨电车系统、城市快轨系统。

综上所述, 目前关于城市轨道交通系统制式的分类研究主要从运量、速度、区域 3 个方面开展了定性研

收稿日期: 2020-11-05 (修改稿)

基金项目: 江苏省高校哲学社会科学基金项目 (编号: 2016SJB630058)

作者简介: 彭涌涛, 男, 硕士, 副教授, E-mail: pyt760815@163.com

究,关于定量研究的文献还不多见。为此,该文试图从预测日均客流量、最高设计速度、线路规划长度、车站设置数量、平均车站间距、单节车厢载客量、车辆宽度等方面归纳整理城市轨道交通线路规划阶段的基础数据,采用 SPSS 分析软件提取主成分进行定量分析,验证目前定性分类的合理性,并且针对城市及城市群快速发展,以及 CJJ/T 114—2007《城市公共交通分类标准》和《城市轨道交通工程项目建设标准》中有关城市轨道交通系统的分类已经不能完全覆盖现有轨道交通系统类型的具体情况,提出系统制式的分类标准及具体类型。

2 原始变量选择

通过国内外关于城市轨道交通系统制式定性分类研究可知,运输能力、运行速度、服务区域是表征系统制式特征的主要指标。该文针对北京、上海、广州、深圳、南京、重庆、长沙等城市 77 条轨道交通线路规划阶段的基础数据,选择预测日均客流量、最高设计速度、线路规划长度、车站设置数量、平均车站间距、单节车厢载客量、车辆宽度共 7 个指标作为聚类分析的原始变量。部分代表性的线路数据如表 1 所示。

表 1 城市轨道交通线路规划阶段代表性数据

序号	轨道交通线路名称	车型	编组	车辆宽度/ m	单节车 厢载客 量/人	预测日均 客流/ (万人次)	最高设 计速度/ (km·h ⁻¹)	线路规 划长 度/km	车站设 置数 量/座	平均车 站间 距/km
1	南京地铁 1 号线	A 型车	6A	3.00	310	98.00	80	21.72	16	1.36
2	南京地铁 S1 号线	B 型车	6B	2.80	240	9.60	100	35.80	8	4.48
3	南京河西有轨电车	100%低地板 现代有轨电车	5 节	2.65	380	0.25	70	7.76	13	0.64
4	苏州高新有轨电车 1 号线	100%低地板 现代有轨电车	5 节	2.65	380	0.55	70	18.19	22	0.87
5	上海地铁 1 号线	A 型车	8A	3.00	310	107.00	80	32.59	27	1.25
6	上海地铁 5 号线	C 型车	4C	2.60	290	11.00	80	17.20	11	1.72
7	上海地铁浦江线	APM	4 节	2.80	240	3.10	80	6.64	6	1.33
8	北京地铁 6 号线	B 型车	8B	2.80	240	88.00	100	52.90	34	1.61
9	北京地铁机场线	L 型车	4L	2.80	180	3.30	110	28.10	4	9.37
10	北京地铁 S1 号线	中低速磁浮	6M	2.80	150	0.26	120	10.20	8	1.46
11	广州地铁 4 号线	Lb 型车	4L	2.80	180	36.00	90	60.03	24	2.61
12	广州地铁 APM	APM	3 节	2.80	240	4.60	60	3.94	9	0.49
13	重庆轨道交通 2 号线	单轨	6 节	2.90	150	36.00	75	31.36	25	1.31
14	长沙轨道交通磁浮快线	中低速磁浮列车	3 节	2.80	120	1.40	100	18.55	5	4.64
15	温州轨道交通 S1 号线	D 型	6D	3.30	330	4.30	140	53.50	18	3.15
...

处理,计算公式如下:

3 原始变量的标准化处理

由于预测日均客流量、最高设计速度、线路规划长度、车站设置数量、平均车站间距、单节车厢载客量和车辆宽度 7 个变量的量纲均不同,并且数值上相差较大,采用 Z-score 标准化方法对以上变量进行标准化

$$Z=\frac{x_{ij}-u_j}{\sigma_j}$$
(1)

$$u_j=\frac{1}{n}\sum_{i=1}^nx_{ij}$$
(2)

$$\sigma_j=\left[\frac{1}{n-1}\sum_{i=1}^n(x_{ij}-u_j)^2\right]^{\frac{1}{2}}$$
(3)

式中: x_{ij} 为原始数据; u_j 为第 j 列样本数据的均值; σ_j 为第 j 列样本数据的标准差。

理后其结果为一个 77 行、7 列的矩阵。部分标准化结果如表 2 所示。

原始数据包含 77 条线路,7 个变量,经标准化处

表 2 原始变量标准化结果

序号	Z 预测日均客流	Z 最高设计速度	Z 线路规划长度	Z 车站设置数量	Z 平均车站间距	Z 车辆宽度	Z 单节车厢载客量
1	1.844 04	-0.482 42	-0.271 42	0.072 03	-0.459 77	0.975 90	0.732 26
2	-0.448 39	0.451 30	0.517 29	-0.792 31	0.894 38	-0.195 18	-0.112 66
3	-0.690 86	-0.949 28	-1.053 40	-0.252 10	-0.772 27	-1.073 49	1.577 18
4	-0.683 08	-0.949 28	-0.469 15	0.720 28	-0.672 45	-1.073 49	1.577 18
5	2.077 43	-0.482 42	0.337 48	1.260 49	-0.507 52	0.975 90	0.732 26
6	-0.412 08	-0.482 42	-0.524 61	-0.468 18	-0.303 53	-1.366 26	0.490 86
7	-0.616 95	-0.482 42	-1.116 13	-1.008 40	-0.472 79	-0.195 18	-0.112 66
8	1.584 72	0.451 30	1.475 16	2.016 79	-0.351 27	-0.195 18	-0.112 66
9	-0.611 76	0.918 15	0.085 97	-1.224 48	3.016 74	-0.195 18	-0.836 87
10	-0.690 60	1.385 01	-0.916 72	-0.792 31	-0.416 37	-0.195 18	-1.198 98
11	0.236 23	-0.015 56	1.874 55	0.936 37	0.082 75	-0.195 18	-0.836 87
12	-0.578 05	-1.416 13	-1.267 38	-0.684 27	-0.837 37	-0.195 18	-0.112 66
13	0.236 23	-0.715 85	0.268 58	1.044 41	-0.481 48	0.390 36	-1.198 98
14	-0.661 04	0.451 30	-0.448 99	-1.116 44	0.963 82	-0.195 18	-1.561 09
15	-0.585 83	2.318 72	1.508 77	0.288 11	0.317 13	2.732 52	0.973 67
...					

4 聚类因子提取

为了减少分析的变量数量,同时又要保留原有变量的大部分信息,通过主成分分析提取相互独立有用的变量信息,避免因变量之间相互关联造成分类评价的误差。

(1) KMO 和 Bartlett 检验

77 条线路,7 个变量采用 SPSS 软件计算,其结果见表 3,由表 3 可得:KMO 的检验值为 0.536,其值大于 0.5,说明变量之间存在明显的相关性;Bartlett 检验的 Sig.=0,说明数据结构效度较高,相关矩阵之间有共同因素存在,因子分析有效。

表 3 KMO 和 Bartlett 的检验结果

取样足够度的 Kaiser—Meyer—Olkin 度量	近似卡方	Bartlett 的球形度检验	
		df	Sig.
0.536	255.3	21	0.000

(2) 公因子提取

采用主成分分析法提取公因子,提取结果如表 4 所示。由表 4 可知:各原始变量因子提取数值均为 0.5 以上,说明公因子能有效提取原始变量信息。

表 4 公因子方差

变量编号	项目	初始	提取
1	Zscore(预测日均客流)	1.000	0.732
2	Zscore(最高设计速度)	1.000	0.799
3	Zscore(线路规划长度)	1.000	0.899
4	Zscore(车站设置数量)	1.000	0.908
5	Zscore(平均车站间距)	1.000	0.777
6	Zscore(车辆宽度)	1.000	0.767
7	Zscore(单节车厢载客量)	1.000	0.809

表 5 为 7 个变量的解释总方差。由表 5 可知:从 7 个变量中提取特征值大于 1 的 3 个公因子,3 个公因子的累积方差贡献率达到 81.308%,即提取的公因子

可以解释原始变量 81.308%的信息,符合公因子提取要求。

表 5 解释的总方差

变量 编号	初始特征值			提取平方和载入			旋转平方和载入		
	合计	方差/%	累积方差/%	合计	方差/%	累积方差/%	合计	方差/%	累积方差/%
1	2.609	37.273	37.273	2.609	37.273	37.273	2.238	31.968	31.968
2	1.898	27.112	64.385	1.898	27.112	64.385	1.952	27.879	59.848
3	1.185	16.923	81.308	1.185	16.923	81.308	1.502	21.461	81.308
4	0.491	7.020	88.328						
5	0.414	5.917	94.244						
6	0.326	4.652	98.896						
7	0.077	1.104	100.000						

(3) 公因子解释

公因子 Factor1 在车站设置数量、线路规划长度、预测日均客流 3 个变量的载荷系数较大,可以将公因子 Factor1 解释为运输能力;公因子 Factor2 在最高设计速度和平均车站间距 2 个变量载荷系数较大,可以将公因子 Factor2 解释为最高速度;公因子 Factor3 在单节车厢载客量、车辆宽度 2 个变量载荷系数较大,可以将公因子 Factor3 解释为车辆型号(表 6)。

进行聚类分析,经过 spss 软件计算得出不同 K 值的聚类结果,考虑了线路的实际情况的 ANOVA 表,结果如表 7 所示。

表 6 成份得分系数矩阵

项目	成份		
	1	2	3
Zscore(预测日均客流)	0.376	-0.097	-0.077
Zscore(最高设计速度)	0.029	0.461	0.015
Zscore(线路规划长度)	0.384	0.286	-0.029
Zscore(车站设置数量)	0.424	-0.078	-0.078
Zscore(平均车站间距)	-0.068	0.430	0.002
Zscore(车辆宽度)	-0.034	0.110	0.568
Zscore(单节车厢载客量)	-0.154	-0.097	0.635

5 轨道交通聚类分析

(1) 聚类分析及结果

采用 K-means 聚类方法,以提取的 3 个公共因子作为聚类变量,分别取聚类 K=4、5、6 对 77 条线路

表 7 ANOVA(K=6)表

项目	聚类		误差		F	Sig.
	均方	df	均方	df		
REGR factor score 1 for analysis 1	9.049	5	0.433	71	20.888	0.000
REGR factor score 2 for analysis 1	9.458	5	0.404	71	23.391	0.000
REGR factor score 3 for analysis 1	12.652	5	0.179	71	70.508	0.000

最终确定 K=6 时的聚类结果,如表 8 所示。将 77 条线路分为城市专用轨道系统(对应聚类 5)、城市现代有轨电车(对应聚类 4)、市郊铁路(对应聚类 2)、市域快轨(对应聚类 6)、轻轨系统(对应聚类 1)、地铁

系统(对应聚类 3)。

(2) 各类系统制式的特征分析

① 地铁系统:服务于特大、大城市中心区域,最高速度一般为 80 km/h,运量为 3 万~7 万人次/h,线路长度小于 40 km,车站平均间距为 0.8~1.2 km,一般采用普通钢轮钢轨系统的 A/B 型车,属于高、大运量城市轨道交通系统。

② 轻轨系统:服务于中等城市核心区域或特大、

表 8 每个聚类中的案例数

聚类						有效	缺失
1	2	3	4	5	6		
20.000	11.000	24.000	5.000	2.000	15.000	77.000	0.000

大城市客流不太大的区域,最高速度一般为 80 km/h,运量为 1 万~3 万人次/h,线路长度一般小于 30 km,车站平均间距为 0.6~1.0 km,可采用普通钢轮钢轨系统的 B/C 型车、直线电机车辆、跨座式单轨车辆、中低速磁悬浮车辆等,属于中运量城市轨道交通系统。

③ 市郊铁路:服务于中心城区边缘到卫星城之间的区域,最高速度一般可达 160 km/h,运量为 20 万~45 万人次/d,线路长度为 50 km 以上,车站平均间距为 5~8 km,可以采用普通钢轮钢轨系统的 CRH6 型车、D 型车,也可以采用 CRH1、CRH2、CRH3、CRH5 等普通动车组。

④ 市域快轨:服务于卫星城与主城之间、卫星城与卫星城之间的区域,线路在中心城区设有车站,采用大站快车和站站停运营模式,最高速度一般为 140 km/h,运量为 2 万~4 万人次/h,线路长度为 50 km 以上,车站平均间距为 2~5 km,可以采用普通钢轮钢轨系统的市域 A、B、D 型车,CRH6 动车组。

⑤ 城市现代有轨电车:服务于中小城市核心城区或作为特大、大城市轨道交通骨干线路的补充,一般采用地面敷设方式、与地面其他交通方式混行,服从路面交通管理,最高速度一般为 70 km/h,运量为 0.6 万~1.0 万人次/h,线路长度为 20 km 左右,车站平均间距为 0.5~0.8 km,目前普遍采用 100% 低地板车辆。

⑥ 城市专用轨道系统:服务于组团内部、特定区域内部、机场内、旅游景区内、以及其他具有点对点特征的专用轨道交通线路。可以采用 APM 自动导轨系统,最高速度一般为 80 km/h,运量为 0.2 万~0.4 万人次/h,线路长度、车站平均间距均较短。

6 结 论

将主成分分析方法和 K-means 聚类方法相结合,选取中国有代表性的 77 条城市轨道交通线路规划阶段的基础数据进行了聚类分析,得出以下主要结论:

(1) 确定了运输能力、最高速度、车辆型号为城市轨道交通系统分类的主要特征指标,验证了运量是城市轨道交通系统区别于其他公共交通方式最显著的特征指标。

(2) 结合城市轨道交通系统定性分类和聚类分析结果,将城市轨道交通系统分为城市专用轨道系统、城市现代有轨电车、市郊铁路、市域快轨、轻轨系统、地铁系统 6 类。其分类结果与运营线路实际的功能定位基本吻合。

(3) 从服务区域、运量、速度、车站间距等方面分析了地铁、轻轨、市域快轨、市郊铁路、城市现代有轨电车、城市专用轨道系统的技术特征。

参考文献:

- [1] 中国城市轨道交通协会. 中国城市轨道交通年度统计分析报告[R], 2018.
- [2] Haruya Hirooka. The Development of Tokyo's Rail Network[J]. Japan Railway & Transport Review, 2000(3): 22-31.
- [3] Georg Puettnner. 德国城市轨道交通的各种制式系统[J]. 城市轨道交通研究, 2007(10).
- [4] 魏庆朝,等. 城市轨道交通制式分类及适用性[J]. 都市快轨交通, 2017(1).
- [5] 陈炎. 城市轨道交通系统分类方法研究[J]. 都市快轨交通, 2015(6).
- [6] 周立新. 城市轨道交通系统分类法的探讨[J]. 城市轨道交通研究, 2005(2).
- [7] 王静,李波,李鹏,等. 基于主成分分析法的网级路面养护管理[J]. 中外公路, 2018(2).
- [8] 魏运,等. 基于功能分类的城市轨道交通系统需求[J]. 都市快轨交通, 2013(1).
- [9] 孔令洋. 城市轨道交通系统型式选择研究[D]. 北京交通大学博士学位论文, 2009.
- [10] 廖亚莎. 山地中心城市轨道交通制式的技术经济评价与选择研究[D]. 重庆交通大学硕士学位论文, 2015.
- [11] 李世元. 轨道交通制式选型在工程中的应用[C]. 中国铁道学会工程分会第七届线路专委会第二次会议论文集, 2017.
- [12] 胡兵,等. GM(1,1)灰色模型优化方法研究[J]. 中外公路, 2018(4).
- [13] 陈彬科,陈林杰,吴建辉,等. 新建地铁下穿既有轨道车站施工方案研究[J]. 中外公路, 2018(6).
- [14] 彭智勇,杨秀仁. 基坑分块开挖参数对邻近地铁盾构隧道的变形影响分析[J]. 中外公路, 2019(2).