

多源数据融合的山地轨道交通站点分类研究

邹庆茹, 郭雪莉

(重庆交通大学 交通运输学院, 重庆 400074)

摘要:为探究山地轨道交通车站客流波动变化特征,利用自动售检票(AFC)数据、兴趣点(POI)数据、路网拓扑结构及车站周边信息,提出考虑山地特征的轨道交通车站类型辨识方法。首先,基于AFC数据和POI数据,采用肘部分析法确定合理的聚类数量,利用K-means聚类法将车站分为3大类;其次,利用路网拓扑数据和车站周边的公-轨数据进行二次聚类,最终将车站分为6个子类:综合类山地型站点、综合类非山地型站点、就业类山地型站点、就业类非山地型站点、居住类山地型站点、居住类非山地型站点。分类结果显示,综合类非山地型相对综合类山地型地势平缓,居民活动聚集性特征明显;居住类山地型与居住类非山地型相比,路网密度较小、受地势影响较大、土地开发程度低;就业类非山地型与就业类山地型相比,周边公交便捷程度高、乘客出行活跃,但停车场数量偏少。

关键词:车站分类;数据挖掘;城市轨道交通;多源数据;聚类

中图分类号: U291 **文献标志码:** A **文章编号:** 2095-0373(2025)01-0097-07

0 引言

山地城市由于特殊的空间地理环境,城市道路布局受限,导致地面交通拥堵问题十分严峻。各大山地城市积极构建以城市轨道交通为骨干的城市公共交通系统。相比平原城市,山地城市轨道交通占公共交通出行的比例较高,如东京都市圈轨道交通占公共交通总量的比例高达77.7%^[1],香港达41%。重庆作为国内典型的山地城市,其城市规划发展中提出2025年轨道交通出行率有望达到50%^[2]。为更好地把握客流需求特征,研究山地城市轨道交通车站类型对提升轨道交通市场占有率具有重要意义。

目前城市轨道交通车站类型识别研究主要分为站点客流特征挖掘和站点周围建成环境2大类。一部分学者通过对站点客流特征挖掘,将站点分为不同类型。余丽洁等^[3]结合高峰小时客流、站点规模和服务设施评价等分类指标,把西安地铁2号线17座车站采用聚类算法分为5类;谭章智等^[4]基于PCA法与地铁客流时序峰值波动形态,将广州市地铁站分成了8个类型;GAN et al^[5]以南京轨道交通站点为例,以小时客流量作为特征值,将其划分为就业导向、混合导向、居住导向等7大类别;袁发涛等^[6]为探究城市轨道交通各站点乘客出行特征,提出一种基于AFC数据挖掘的站点分类研究方法;另一部分研究学者通过站点周围建成环境的研究对轨道站点进行划分。MA et al^[7]基于遥感土地利用数据,利用层次分析法,识别站区的TOD开发类型;路羨乔等^[8]将建成环境因素引入传统的节点-场景模型中并结合GIS对北京市站点进行分类划分。现有研究中,同时考虑站点客流特征和站点周围建成环境的研究较少,且大部分学者对站点分类研究集中在平坦地区,未考虑山地城市的适用性。因此,结合山地城市的空间地理特色研究轨道站点分类尤其重要,准确识别站点类型可以对站内设施及人力资源配备、列车运行班次进行差异性调整和优化,不仅可以防范进出站瞬时大客流风险,还可以促进站点的个性化运营,能对不同站点周边土地周边站域提出差异化的组织管理方案和土地利用开发策略,对引导城市建设与轨道交通建设协同发展具有积极的促进作用。

重庆是典型的山地城市,以重庆轨道交通已运营的9条线路所辖站点为对象,具体包括1、2、3、4、5、6、10号线、环线以及国博线的194个车站(换乘站不重复计算),从开发强度、车站客流数据、换乘便捷性

收稿日期:2024-10-16 责任编辑:车轩玉 DOI:10.13319/j.cnki.sjztdxzb.20240258

基金项目:国家自然科学基金(52302386);中国博士后科学基金(2023M730430);重庆市教委项目(KJQN202300741)

作者简介:邹庆茹(1987—),女,副教授,研究方向为轨道交通运营。E-mail:308378567@qq.com

邹庆茹,郭雪莉.多源数据融合的山地轨道交通站点分类研究[J].石家庄铁道大学学报(自然科学版),2025,38(1):97-103.

和步行可达性 4 个方面对山地轨道交通车站进行类型辨识,为客流特征的精细化、差异化分析奠定基础。

1 车站分类指标选取

结合重庆的道路网络结构及步行体系特点,构建“以轨道交通站点为中心,半径 500 m 为站域关键影响区域”^[9],确定了以下 4 大类数据作为车站类型辨识的指标:

(1)开发强度指标。POI 数据能够有效揭示车站周边区域的开发密度和职位分布状况,根据文献[10],对核心站点区域提取的兴趣点进行分类,涵盖住宅用地、商业用地、特殊用地、公共管理与公共服务用地、交通运输用地等类别。其中商业用地 POI 138 887 个,住宅用地 POI 72 878 个,公共管理与公共服务用地 POI 40 215 个,交通运输用地 POI 23 766 个。特殊用地 POI 数量相对较少,在此不进行统计。

(2)换乘便捷性指标。根据高德 POI 数据计算出轨道交通车站 500 m 范围内公交线路数量,以此衡量站点周边城市公交与轨道交通换乘的便捷性,通过停车场的数量反映 P-R 换乘模式下的便捷性。提取车站周围公交车线路共计 3 585 条,停车场数量共 3 316 个。

(3)步行可达性指标。利用 Open Street Map 城市交通网络数据,步行可达性通过站点所覆盖的路网长度来衡量。一般站点周边道路网长度越长,站点附近的路网结构越健全,站点可达性也就越高^[11]。山地城市周边土地路线的曲折程度和高低差异,可通过道路增长系数与平均纵坡来反映,如式(1)、式(2)所示。共计提取站点覆盖范围内道路信息 3 842 条,进而计算出道路长度、道路增长系数及平均坡度。

$$I = (H - H_0) / L_r \times 100\% \quad (1)$$

式中, I 为道路平均纵坡; H 为站点高程; H_0 为 500 m 范围内乘客出发点高程, L_r 为 500 m 范围内乘客出行长度。

$$C_r = L_r / L_s \quad (2)$$

式中, C_r 为道路增长系数; L_r 为 500 m 范围内乘客步行长度; L_s 为乘客步行起终点的直线距离。

(4)车站客流指标。来源于 2023 年的城市轨道交通自动售检票系统的车站客流数据。选取各站点早晚高峰进出站客流,反映不同时段的客流变化^[12]。

结合上述 4 大类指标,具体选取 11 个指标作为车站分类的指标,此外为了减少不同指标间量纲不一致带来的影响,对数据进行标准化处理,确保各项数据指标的相对可比性。采用 Z-score 方式进行数值标准化处理,即将每个变量值与其平均值的差值除以该变量的标准差。经过标准化处理后变量的平均值变为 0,标准差为 1,以此消除量纲和数量级的影响,处理后结果见表 1(数据类型均为连续型数据)。

表 1 站点聚类因素

指标类型	具体分类指标	单位	数据范围	标准化后数据范围
开发强度	商业 POI 数量	个	[22,6 395]	[-0.739 1,5.419 8]
	住宅 POI 数量	个	[8,444 8]	[-0.605 2,5.911 1]
	公共管理与公共服务 POI 数量	个	[0,188 4]	[-0.733 7,5.800 2]
	交通运输 POI 数量	个	[0,157 2]	[-0.686 7,5.712 9]
换乘便捷性	公交线路数量	条	[2,58]	[-1.520 2,3.881 5]
	停车场数量	个	[1,88]	[-0.818 2,4.912 7]
步行可达性	路网长度	m	[53.794,543.755]	[-0.882 6,5.635 9]
	道路平均纵坡	%	[-101.88,35.54]	[-1.790 7,2.695 9]
	道路增长系数		[1,41.97]	[-5.330 8,1.931 7]
车站客流	早高峰进站客流	人/h	[20,14 258]	[-0.852 8,6.307 9]
	晚高峰进站客流	人/h	[18,20 943]	[-0.663 6,7.732 7]
	早高峰出站客流	人/h	[15,17 149]	[-0.679 1,6.708 3]
	晚高峰出站客流	人/h	[12,15 943]	[-0.807 3,6.806 5]

2 基于聚类的车站类型辨识

运用 K-means 聚类算法对各车站类型进行无监督分类。为了更加准确描述山地城市环境对城市轨道交通

通车站特征的影响,将以上 4 类因素分 2 次进行聚类。首先,从站点功能性划分及客流特征角度,根据站点客流指标和开发强度指标对站点进行聚类;其次,为进一步研究山地道路环境对站点分类的影响,利用换乘便捷性指标和步行可达性指标进行二次聚类。

2.1 基于站点客流和开发强度的一次聚类

结合客流刷卡数据和 POI 的各项指标对轨道城市交通站点进行分析聚类。利用肘部法对距离数量确定,结合 Python 代码对本文数据样本进行分析,确定聚类数目 K ,如图 1 所示。

由图 1 可以发现,当 K 值由 1~3 逐渐增大时,其畸变程度呈现出明显的波动;当 K 值超过 3 之后,畸变程度的变动幅度明显减小。因此,选定最佳聚类数为 3,利用 K-means 算法进行聚类,聚类结果如图 2 所示。

针对聚类结果中的 3 类站点,通过图 3 和图 4 的聚类结果展示了各类工作日站点的进出站客流^[13]和 POI 开发类型,对其具体类型进行辨识及分析。

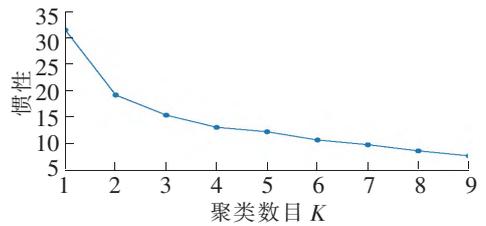


图 1 不同聚类数量下的误差平方和

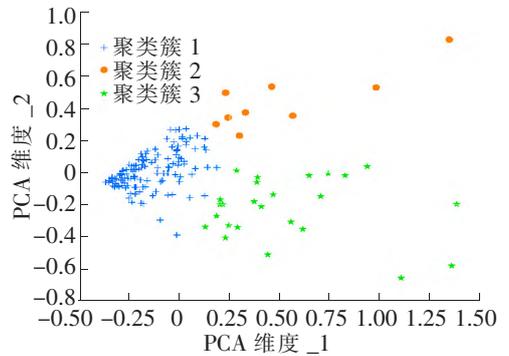


图 2 K-means 聚类结果图

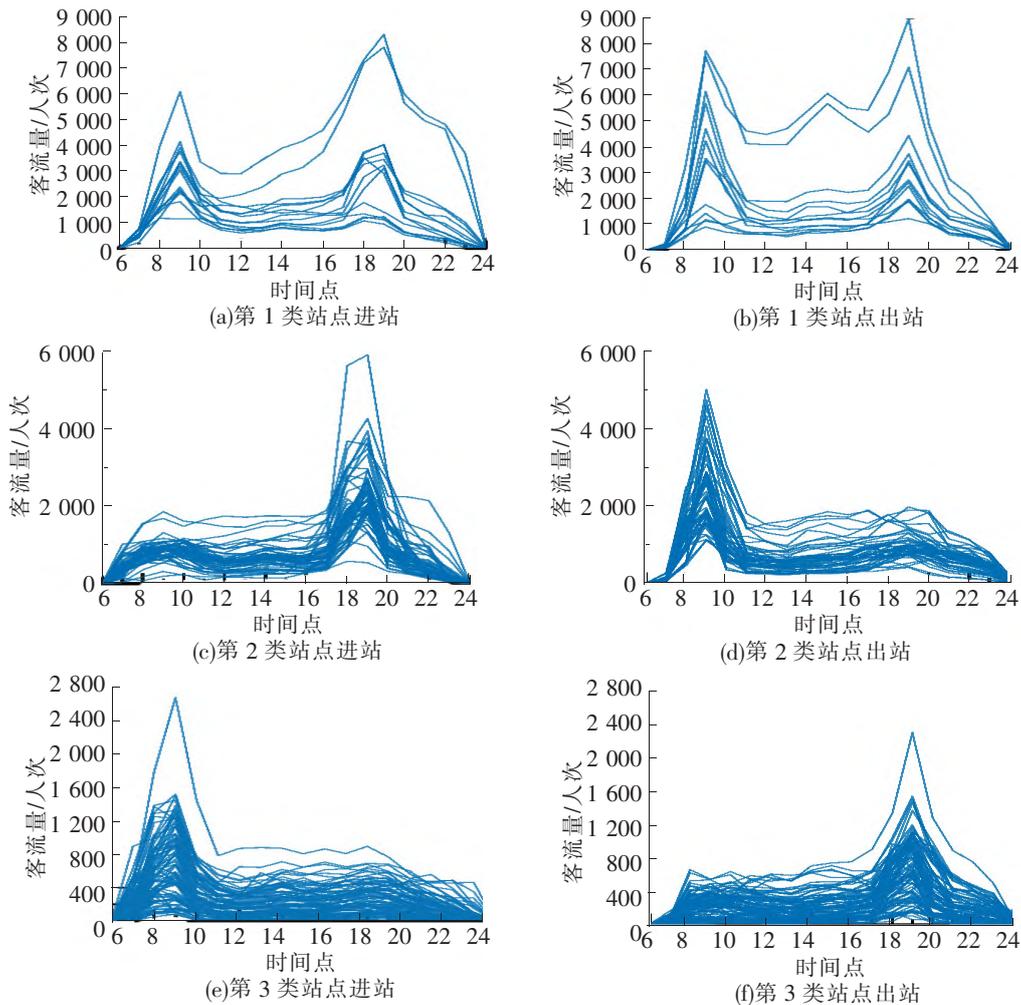


图 3 工作日站点的进出站客流

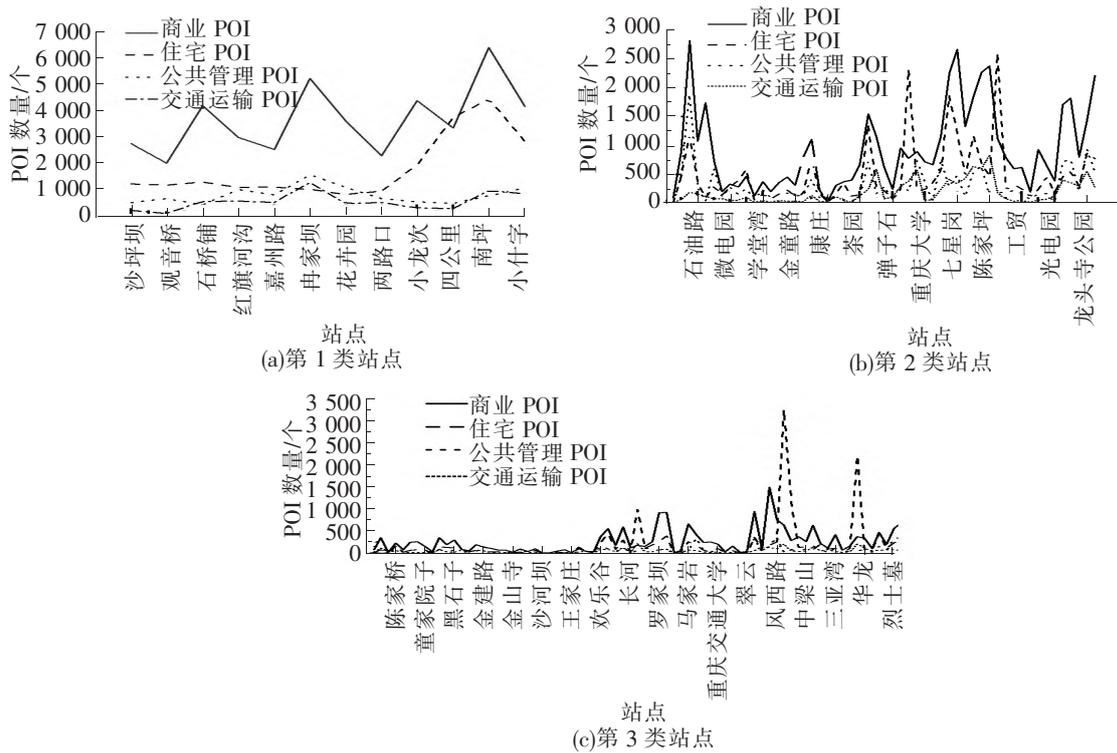


图 4 POI 类型分布图

(1) 第 1 类站点在工作日的客流量展示出明显的双峰特征。早晚高峰十分显著,这主要是因为用地性质以商业服务为主^[14],且客流量要高于其他 2 类站点。将此类站点定义为综合类站点。

(2) 第 2 类站点的客流量曲线呈现单峰形态,早高峰时段出站客流和晚高峰进站客流显著较高。该区域中商业设施和服务的 POI 所占比例较高,周围公交数量和停车场数量较大,这可能是因为公共交通为了满足居民通勤要求,还有在就业集聚地带,居民也可能会选择自行开车通勤。因此,这类站点被确定为就业类站点。

(3) 第 3 类站点客流曲线呈现为单峰形态,但其高峰时段与第 2 类站点相反,早晨的进站客流量和晚上的出站客流量显著较高。周边地域构成以居住用地为主导,周围用地开发程度不高。将此类站点定义为居住类站点。

2.2 基于换乘便捷性因素和步行可达性因素的二次聚类

根据现有站点分类对每个大类进行基于山地轨道特征、换乘便捷性和步行可达性的聚类分析,用肘部法确定不同类型站点进一步聚类的聚类数目,如图 5 所示。

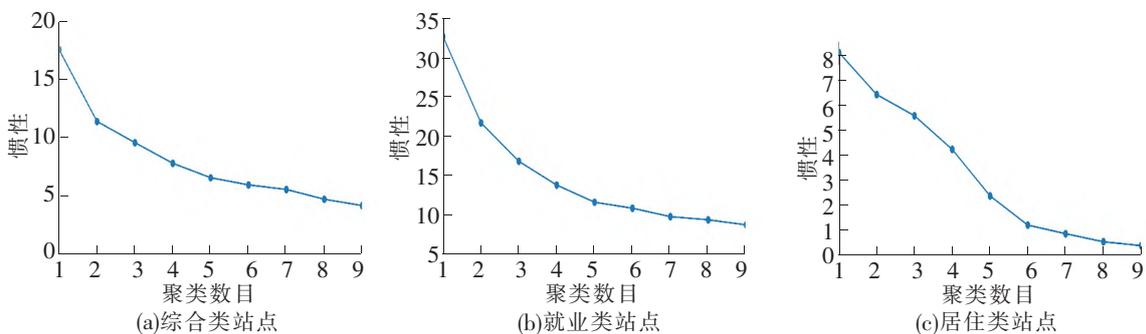


图 5 不同类型站点聚类肘部图

由图 5 可知,不同类型站点畸变程度在 2 时达到最大,当超过 2 之后畸变程度明显减小,且通过对分为 2 类和 3 类之后的聚类效果显示,在聚类数量为 2 时,效果达到最优。故此,选定不同类型站点的聚类类别数为 2 类。

采用 K-means 聚类对站点类型进一步划分,最终将轨道站点聚为 6 类,以步行可达性的 3 个因素为示例,展示聚类结果如图 6 所示。综合类站点客流量较大且周围土地利用率高;就业类站点客流高峰在早晚,早高峰时段出站客流和晚高峰进站客流显著较高。该区域中商业设施和服务的 POI 所占比例较高;居住类站点早晨的进站客流量和晚上的出站客流量显著较高且周边地域构成以居住用地为主导,站点覆盖范围内停车场数量和公交线路密度比较低。在各大类站点内,根据二次聚类对山地特征的分析,具有明显山地特征的称之为山地型,不明显的称为非山地型,因此将重庆市 194 个轨道站点的分类结果通过 Arcgis 进行了可视化展示,如图 7 所示。

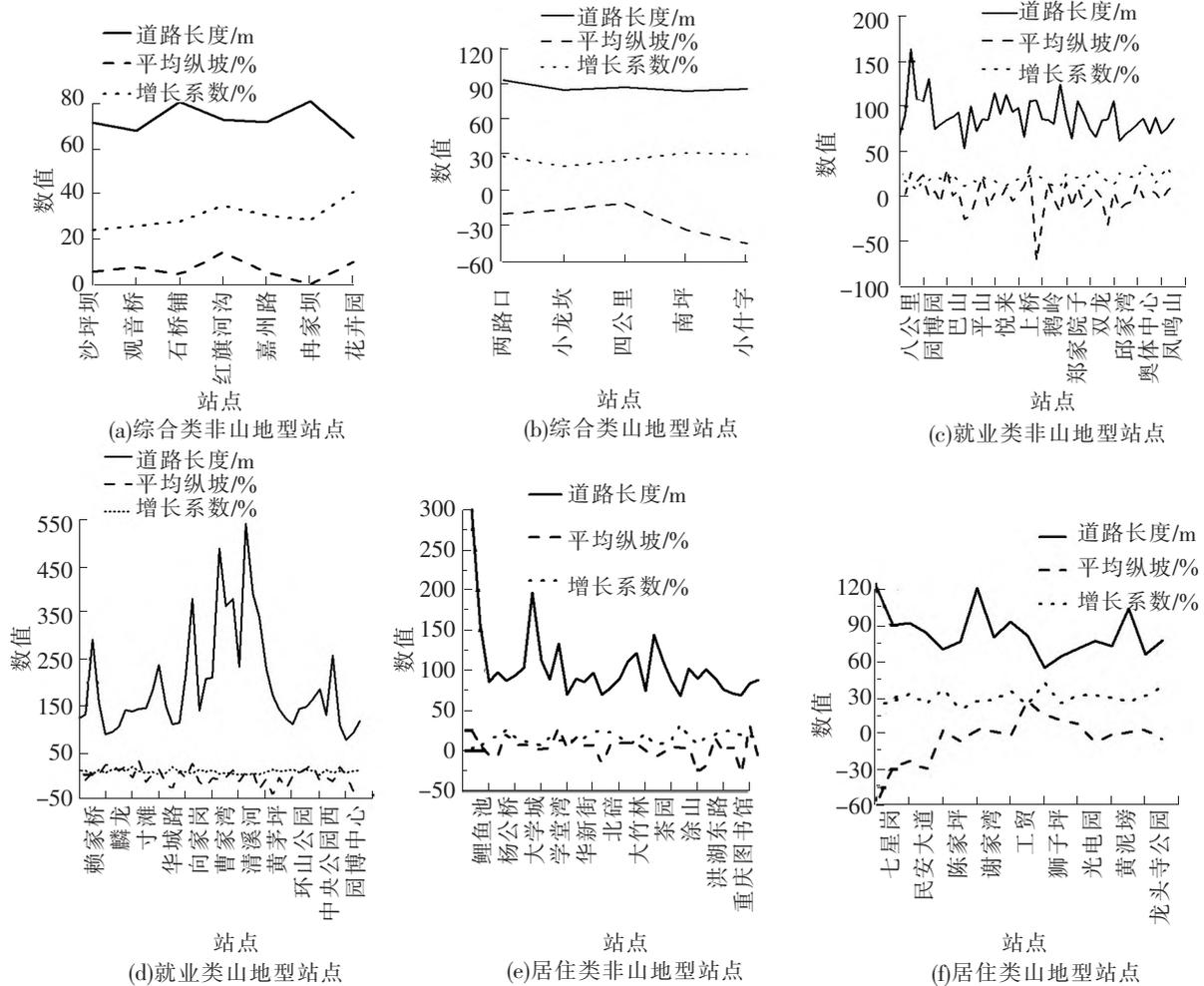


图 6 不同类别的步行可达性指标结果图

(1)综合类站点中,平均纵坡较大,站点周围坡度较高,道路增长系数高,有明显山地特征的定义为综合类山地型站点,反之为综合类非山地型站点。综合类非山地型站点早晚高峰进出站客流量要高于综合类山地型站点,地势平坦的地方居民活动更加聚集,土地混合利用程度更高,乘客出行更加便捷。

(2)就业类站点中,道路网密度低,站点覆盖范围内公交线路密度较低,停车场数量多,早高峰出站和晚高峰进站客流量比较显著,客流呈现单峰形态的为就业类山地型站点,而在就业类非山地型站点中,站点周围公交线路网密度和停车场数量都比较高。就业类非山地型站点有较平缓的

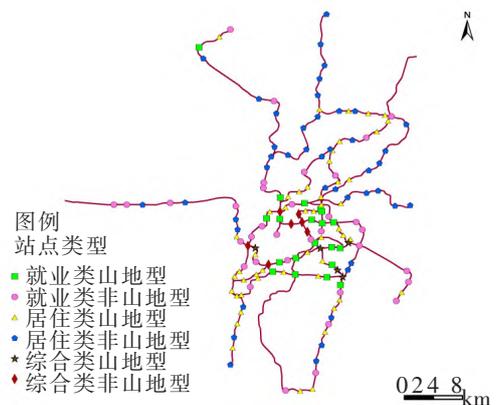


图 7 重庆市轨道交通站点分类

地势,更利于站点周围城市规划和土地开发利用,尤其是在商业 POI 聚集的站点,客流的单峰状态更加明显。就业类山地型站点停车场数量比较高,是由于在山地地形区乘客为了方便出行,选择 P-R 的换乘方式。

(3)居住类站点中,道路网密度低,平均纵坡度高,居住类 POI 高的定义为居住类山地型站点,反之为居住类非山地型站点。站点周围乘客出行受地势影响比较大,居住类非山地型站点地势平缓,进出站客流量要高于居住类山地型站点。居住类 POI 山地型站点要高于非山地型站点,是由于山地型站点周围土地开发程度比较低,房价、物价等方面要低于开发密度高的地区。同样在地势较为平坦的地区,公交线路网密度要比地势不平的地区高。

3 结论

利用重庆市 AFC 数据和 POI 数据,运用肘部法则结合 K-means 聚类,通过分析站点早晚高峰进出客流与站点开发强度,将重庆市轨道交通站点分为综合、居住和就业 3 大类。引入山地地形因素指标,考虑居民出行便捷程度,在原有聚类类型下进行山地指标聚类,将每大类分为山地型和非山地型,得出结论:重庆市 194 个轨道交通站点可分为综合类非山地型、综合类山地型、居住类非山地型、居住类山地型、就业类山地型及就业类非山地型 6 类。该分类结果对山地城市的城市交通规划具有一定参考意义,对城市轨道交通发展和站点周围土地利用有积极的意义,更好地提高居民出行效率。后续研究可在本文研究基础上考虑更多元的数据,如手机信令、共享单车等,或者可以结合共轨换乘进一步分析城市轨道交通站点时空分布特征。

参 考 文 献

- [1]拓道论道. 教科书级典范! 日本东京都市圈市域(郊)铁路特点及其启示[DB/OL]. (2019-06-20) [2024-04-09]. https://www.sohu.com/a/321902473_281835.
- [2]胡虹,古福. 人民网-重庆频道:到 2025 年重庆轨道交通将占公共交通出行超 50% [DB/OL]. (2022-01-16) [2024-04-09]. <http://cq.people.com.cn/n2/2022/0116/c365401-35097389.html>.
- [3]余丽洁,李岩,陈宽民. 基于谱聚类的城市轨道交通站点分类方法[J]. 交通信息与安全,2014,32(1): 122-125.
- [4]谭章智,李少英,黎夏,等. 基于地铁客流的广州地铁站点类型识别[J]. 热带地理,2017,37(1): 102-111.
- [5]GAN Z, YANG M, FENG T, et al. Understanding urban mobility patterns from a spatiotemporal perspective: Daily ridership profiles of metro stations[J]. Transportation,2020, 47: 315-336.
- [6]袁发涛,陈通箭,魏剑波. 基于 AFC 数据的轨道站点分类研究[J]. 道路交通与安全,2021,21(1):48-52.
- [7]MA X, WU Y J, WANG Y, et al. Mining smart card data for transit riders' travel patterns[J]. Transportation Research Part C: Emerging Technologies, 2013,36:1-12.
- [8]路羨乔,王晶,张政,等. 基于扩展节点-场所模型的北京轨道微中心地铁站点评价与分类[J]. 科技导报,2023,41(24): 41-51.
- [9]杜光远,吴瑞,谭桂菲,等. 基于客流特征的城市轨道交通站点分类:以青岛市为例[J]. 交通运输研究,2023,9(5): 63-72.
- [10]王凯,赵民,林坚,等. 城市用地分类与规划建设用地标准:GB 50137—2011[S]. 北京:中国建筑工业出版社,2010.
- [11]马壮林,杨兴,胡大伟,等. 城市轨道交通车站客流特征影响程度分析[J]. 清华大学学报(自然科学版),2023,63(9):1428-1439.
- [12]李岩辉. 城市轨道交通客流时空分布特性研究[J]. 中国铁路,2021(2):65-75.
- [13]解坤,张俊芳. 基于 KMO-Bartlett 典型风速选取的 PCA-WNN 短期风速预测[J]. 发电设备,2017,31(2): 86-91.
- [14]邵治锦,刘雯. 基于 POI 数据的商业空间集聚与轨道交通站点耦合研究:以重庆市中心城区为例[C]//人民城市,规划赋能:2022 中国城市规划年会论文集. 北京:中国建筑工业出版社,2022.

Classification of Mountain Rail Transit Stations with Multi-source Data Fusion

ZOU Qingru, GUO Xueli

(College of Traffic and Transportation, Chongqing Jiaotong University, Chongqing 400074, China)

Abstract: In order to explore the changing characteristics of passenger flow fluctuation in mountain rail transit stations, a method for identifying rail transit station types considering mountainous features was proposed using the automatic fare collection (AFC) data, point of interest (POI) data, road network topology, and the data of the mountain station. Firstly, based on AFC data and POI data, the reasonable cluster number was determined according to elbow analysis, and the K-means clustering method was used to divide stations into three categories; Secondly, using the road network topology data and public-rail data for secondary clustering, the stations were ultimately divided into six sub-categories: comprehensive mountain stations, comprehensive non-mountain stations, employment mountain stations, employment non-mountain stations, residential mountain stations and residential non-mountain stations. The classification results show that the comprehensive non-mountain type has a relatively flat terrain and obvious characteristics of residential activity aggregation compared to the comprehensive mountainous type; Compared with non-mountainous residential type, mountainous residential type has lower road network density, is more affected by terrain, and has lower degree of land development; Compared with employment non-mountain type, the employment mountain type has higher convenience in surrounding public transportation and active passenger travel, but the number of parking lots is relatively small.

Key words: station classification; data mining; urban rail transit; multi-source data; clustering

~~~~~  
(上接第 21 页)

**Abstract:** Taking the shield tunnel project of Suzhou Rail Transit Line 11 as the background, a numerical calculation model was established by the finite element software ABAQUS to analyze the displacement and variation of the electric tower in short distance under four working conditions: the foundation of the tower and the tower body without reinforcement (no reinforcement), the foundation of the tower reinforced by grouting (grouting reinforcement), the tower body reinforced by anchors only (anchors reinforcement), and the foundation of the tower combined with the anchors of the tower body reinforced by grouting (grouting combined with the anchors, the actual reinforcement), and the results were verified by on-site measurement. The results show that, after grouting and reinforcing the electric tower, the settlement distribution pattern of the reinforced section will change when the shield tunneling passes through it, and the maximum settlement value will shift to the side far away from the reinforced area; Under the actual reinforcement condition, the variation law of the settlement of the tower pile foundation caused by the shield tunneling through the tower is basically the same, and the difference between the maximum settlement value (average value) and the measured value is 16.1%; When the shield tunneling passes through the electric tower under each reinforcement condition, the horizontal displacement of the electric tower pile foundation in the vertical and parallel direction of the tunnel is 1.81, 1.73, 1.63 and 1.15 mm and 1.83, 1.81, 1.86 and 1.44 mm respectively; Grouting combined with anchor cable reinforcement is the most effective in controlling the settlement and horizontal displacement of the pile foundation.

**Key words:** shield tunnel; pile deformation; reinforcement measures; numerical analysis