

基于等时圈形状的轨道站点客流预测研究

舒代帅

重庆交通大学 交通运输学院 重庆 400074

摘要: 准确掌握城市轨道交通的客流量变化规律是轨道交通合理规划建设和高效运营管理的前提。研究主要从等时圈形状特征的角度探索轨道站点分类方法, 提出细长度和偏心圆比两个指标对轨道站点进行分类。基于分类结果建立ARIMA模型分别对两类站点的进站客流进行预测, 最后对预测结果进行评价。结果表明, 分类后两类站点的预测效果较好, MAPE均低于10%, 其中条状等时圈类站点的预测精度高于团状等时圈类站点。

关键词: 等时圈; 形状特征; 轨道站点; 客流预测; ARIMA模型

Research on passenger flow prediction of orbit station based on isochronal circle shape

Daishuai Shu

School of Transportation, Chongqing Jiaotong University, Chongqing 400074

Abstract: It is the premise of rational planning, construction and efficient operation management of urban rail transit to accurately master the rule of passenger flow variation. This paper mainly explores the classification method of orbital stations from the Angle of isochronal circle shape characteristics, and proposes two indexes of fine length and eccentric circle ratio to classify orbital stations. Based on the classification results, the ARIMA model is established to predict the inbound passenger flow of the two types of stations respectively, and the prediction results are evaluated at last. The results show that the prediction results of the two types of sites after classification are good, and the MAPE is less than 10%. The prediction accuracy of the stripe isochronal circle sites is higher than that of the cluster isochronal circle sites.

Keywords: Isochronous circle; Shape feature; Orbital stations; Passenger flow forecast; ARIMA model

引言:

城市轨道交通作为公共交通系统的主骨架, 承载着绝大多数居民的日常出行, 它的迅猛发展为解决交通拥堵问题作出了巨大贡献。轨道站点作为城市轨道交通线网中的重要节点, 不同类型的轨道站点交通功能不同, 周边土地利用性质也存在差异, 有必要从不同角度对轨道站点进行分类, 探讨不同类型站点的差异并进行客流预测^[1]。

目前, 国内已有大量关于城市轨道交通的相关研究。在轨道站点分类方面, 徐威等以站点自身特性和建成环境为变量进行聚类得到主要因子, 采用K-means聚类对主要因子进行聚类将轨道站点分为5类; 王焕栋等基于站点服务范围内的POI数据, 以不同站点的各类POI比例、POI优势度和均匀度为指标将轨道站点分为4类; 袁发涛等采用K-means聚类算法对AFC数据进行聚类, 结

合乘客出行特征分析结果将轨道站点分为6类; 杨静等基于轨道客流数据识别出站点职住功能特征, 基于POI数据分析得到细粒度土地利用特征, 基于两个特征采用非监督学习K-means++方法将车站分为了7类^[2]。在轨道站点客流预测方面, 姚恩建等基于土地利用和客流分布的关联性, 构建了考虑站点土地利用及其站点间土地利用性质匹配度的客流分布预测模型, 具有较高的预测精度; 刘维源等基于当前客流水平构建了ARIMA-GARCH模型, 预测轨道交通未来节假日各时段客流量, 能有效识别节假日客流特征, 实现城市轨道交通节假日各时段客流预测; 朱从坤等构建了一种经验模态分解与K近邻非参数回归结合的组合预测模型, 预测结果具有较高的精度。

上述研究多侧重于轨道站点的分类或轨道站点的客流预测, 并未考虑基于站点分类的客流预测。本文基于等时圈的形状特征对轨道站点进行分类, 在分类的结果

上进行轨道客流预测, 并对预测结果进行误差评价。

一、数据获取与处理

1. 等时圈数据获取

目前, 等时圈的构建方法主要有两种: 一种是基于设施服务区的等时圈构建方法, 即通过GIS软件中的网络分析工具, 输入时间阈值, 求解设施点的服务区范围形成等时圈; 另一种是基于OD成本矩阵的等时圈构建方法, 即通过查找和测量网络中从多个起始点到多个目的地的最小成本路径, 从而形成指定时间阈值内点对点的OD对网格覆盖范围即等时圈范围。相比之下, 方式二的结果更加精确, 本文采用一种与方式二原理类似的等时圈构建方法, 即采用Python编程语言借助Mapbox地图平台的Isochrones API接口获取所需的等时圈数据。获取的详细步骤如下。

- 步骤1 指定起始点、时间阈值、出行方式;
- 步骤2 以起始点为中心生成最大可达范围;
- 步骤3 以最大可达范围为基础生成渔网点阵;
- 步骤4 计算起始点到各渔网点的出行时间成本;
- 步骤5 根据出行时间成本计算结果生成等值线;
- 步骤6 提取符合时间阈值的等值线生成等值面。

2. 等时圈数据处理

通过ArcGIS中的模型构建器进行建模将初步获取的等时圈数据分割导出; 利用Python编程语言中的GeoPandas包将矢量数据转换为方便处理的JPG图片格式, 利用OpenCV包对图片进行二值化, 以便后期进行形状特征识别。

二、等时圈的形状识别

1. 形状特征值

形状特征值主要参考景观生态学中的形状指数进行定义, 通过选取合适的形状特征值可有效的对形状特征进行描述, 以此区分等时圈的形状, 并进行形状分类识别。本文主要定义了两个形状特征指标, 分别为细长度和偏心圆比。

(1) 细长度 (α)

最小外接矩形为包括形状的最小面积的矩形, 细长度 α 是指等时圈形状的最小外接矩形长轴与短轴的比值^[3]。细长度的值可以有效描述形状区域的长短特征, 其计算公式为:

$$\alpha = \frac{L_a}{L_b} \quad (1)$$

式中: α 为所求的形状细长度, L_a 为形状最小外接矩形的长, L_b 为形状最小外接矩形的宽。部分站点等时圈的最小外接矩形如图1所示, 从中可以看出图1(b)的最小外接矩形特征与图1(a)、(c)、(d)有明显不同,

即图1(b)的最小外接矩形长、短轴长度相差较大, 因此其细长度值相较于其余等时圈偏大。

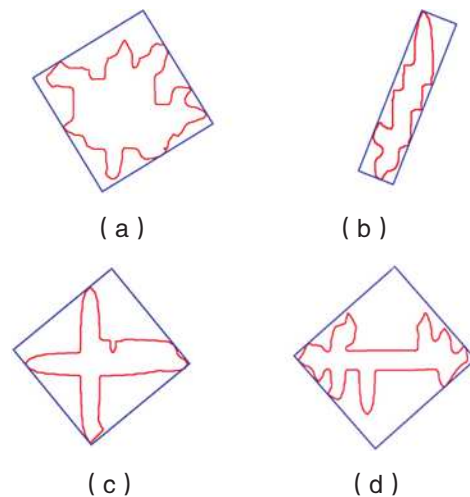


图1 部分站点等时圈的最小外接矩形

(2) 偏心圆比 (β)

形状的最大内切圆面积可以反映形状内部区域特征, 形状的最小外接圆面积可以反映形状的整体形态特征, 两者之比可以有效描述形状边缘的变化情况, 将该比值定义为偏心圆比 β , 其计算公式为:

$$\beta = \frac{S_{MC}}{S_{MCC}} \quad (2)$$

式中: β 为所求的形状偏心圆比, S_{MC} 为形状最大内切圆的面积, S_{MCC} 为形状最小外接圆的面积。部分站点等时圈的偏心圆如图2所示, 其中图2(a)的偏心圆特征与图2(b)、(c)、(d)有明显不同, 即图2(a)的最大内切圆面积与最小外接圆面积相差最小, 因此其偏心圆比相较于其余等时圈明显偏大。

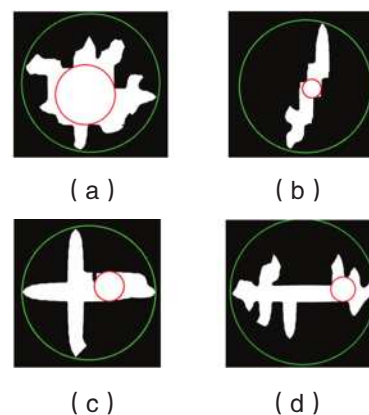


图2 部分站点等时圈的偏心圆

2. 形状特征识别

在既有的传统村落空间形态量化研究中, 普遍将村落形态分为团状、带状和指状, 且皆是先用长宽比进行初步筛分, 然后用形状指数进行细分^[4]。由2.1节可知,

上述形状特征值可分别代表等时圈的某项显著特征。将通过细长度筛分出的等时圈定义为条状等时圈, 该等时圈在形态上具有明显的方向指向性, 通常沿某条主干道展开, 且两侧多为不可跨越的建筑实体或全封闭的车行道, 形状特征临界值为 α_{\min} ; 通过偏心圆比筛分出的等时圈定义为团状等时圈, 该等时圈形态上偏向基础几何图形, 如圆形、正方形或不规则多边形, 范围内路网密度大, 10min内的空间步行可达性更高, 形状特征临界值为 β_{\min} 。

对于提取条状等时圈的细长度临界值 α_{\min} , 学界多采用浦欣成^[5]的设定即将长宽比(即本文中细长度) $\lambda=1.5$ 和 $\lambda=2$ 作为临界点, 其初步筛分原则如下:

$$\begin{cases} \lambda > 2, \text{ 带状} \\ 1.5 \leq \lambda \leq 2, \text{ 带状倾向的团状} \\ \lambda < 1.5, \text{ 团状} \end{cases} \quad (3)$$

其中, 当 $1.5 \leq \lambda \leq 2$ 时, 形状既非典型的团状, 也非典型的带状。通过对比分析部分轨道站点等时圈的细长度特点可知, 当 $\lambda < 1.5$ 时, 不能有效区分出团状等时圈, 如图1(a)、(c)(d)的细长度大小相近, 但图1(a)的等时圈形状特征与后者有较大差别。而通过偏心圆比则能有效解决这个问题, 因此本文考虑通过细长度初步筛分出条状等时圈, 通过偏心圆比进一步筛分团状等时圈, 其余等时圈暂不做分析。

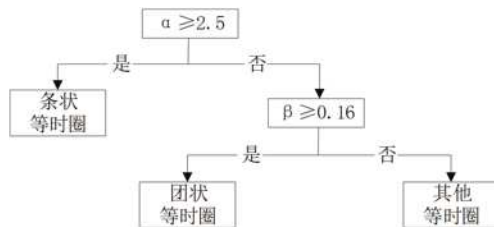


图3 等时圈形状识别过程

对于初步筛分的细长度临界值 α_{\min} , 本文首先参考式(3)中的带状筛分临界值即 $\alpha_{\min}=2$ 进行筛分, 发现部分站点等时圈的形状不满足条状等时圈的特征。为提高识别的精度, 考虑到轨道站点等时圈形状与传统村落形态的区别, 通过对计算结果的统计与分析, 将细长度的临界值确定为 $\alpha_{\min}=2.5$, 即当 $\alpha \geq 2.5$ 时, 该站点等时圈为条状等时圈。同理, 通过对偏心圆比计算结果的统计与分析, 将偏心圆比的临界值确定为 $\beta_{\min}=0.16$, 即当 $\beta \geq 0.16$ 时, 该站点等时圈形状为团状等时圈。识别过程如图3所示。

三、站点客流预测

1. ARIMA 预测模型

采用ARIMA模型进行站点客流预测, ARIMA模型是经过差分后的ARMA模型, 保证了数据的稳定性。对于

平稳序列, 若满足下面公式(4), 称其为ARMA模型。

$$y_t = \varphi_0 + \varphi_1 y_{t-1} + \dots + \varphi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (4)$$

式中: y_t 表示预测值; $\varphi_i (i=1, 2, \dots, p)$ 和 $\theta_j (j=1, 2, \dots, q)$ 表示模型参数; ε_t 表示服从独立正态分布的白噪声序列。

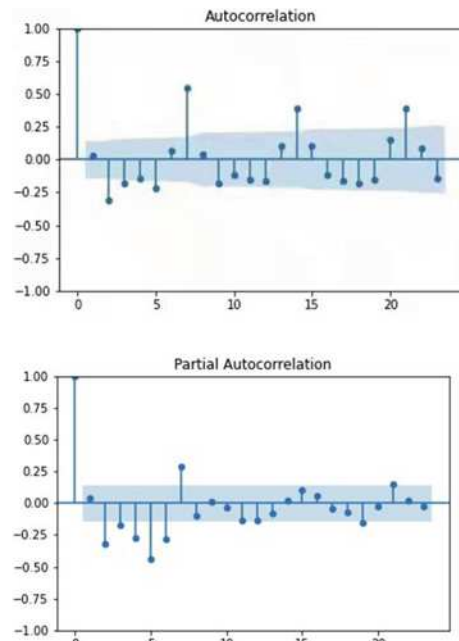
差分自回归移动平均模型作为多元线性回归模型, 是研究时间序列的重要方法之一。具体由自回归模型(AR)、移动平均模型(MA)和自回归移动平均模型(ARMA)构成。预测前需对时间序列进行平稳性检验及差分转换, 然后进行模型识别与参数标定。

模型识别一般使用自相关函数(ACF)和偏相关函数(PACF), 前者体现了样本数据的总体相关性, 后者体现了样本数据间的变量相关性^[6]。具体通过自相关函数图和偏相关函数图在滞后后期所体现的拖尾和截尾特性对模型进行识别。

参数标定一般使用贝叶斯信息准则(BIC)和赤池信息准则(AIC)。由于BIC的惩罚项大于AIC, 可以有效弥补AIC估计容量较大的数据时产生的拟合误差, 本研究以BIC准则为标准对模型参数进行标定。ARIMA模型中的p、q值通常为1-20之间的整数, 采用枚举法进行验证, 得到使BIC值最小的p、q。

2. 预测及结果分析

分别对两类站点的客流数据的拖尾和截尾情况进行判定, 两类站点的自相关函数图和偏相关函数图如图4所示。



(a) 团状等时圈类站点

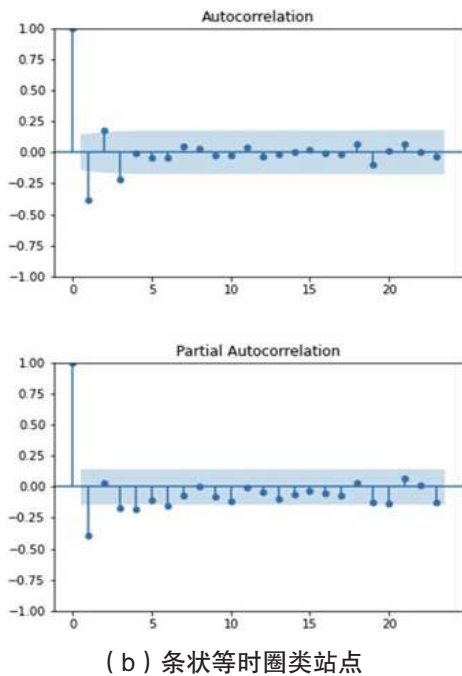


图4 客流量自相关函数和偏相关函数图

从图4可知, 两类站点的自相关函数和偏相关函数都具有明显的拖尾性, 因此使用ARMA模型进行建模。首先对实验数据进行1次差分转换为平稳性时间序列, 然后通过BIC准则确定团状等时圈类站点的模型为ARIMA(7, 1, 7), 条状等时圈类站点的模型为ARIMA(3, 1, 8), 分别运用以上ARIMA模型对两类站点的客流进行预测, 结果如图5所示。

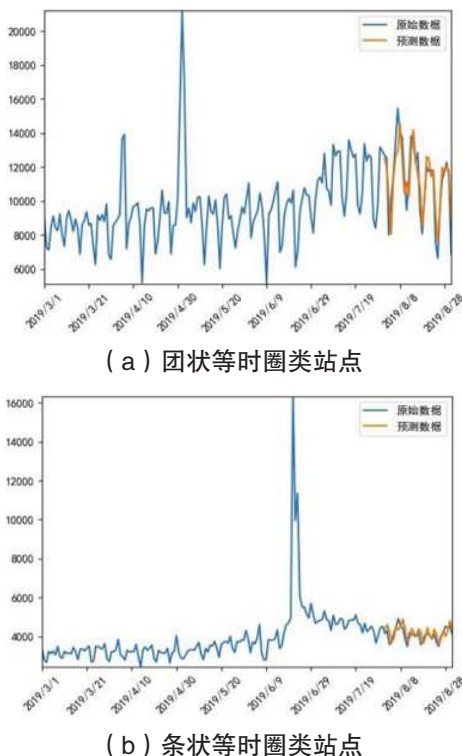


图5 ARIMA模型预测值与实际值对比

从图5看出, ARIMA模型对轨道站点客流量的预测效果较好。为进一步对两类站点的预测效果进行评价, 选取相对误差(MAPE)、绝对误差(MAE)以及均方根误差(RMSE)对预测结果的精度和拟合效果进行量化分析, 评价结果如表1所示。通过表1可以看出, 团状等时圈类站点和条状等时圈类站点的MAPE均低于10%, 由此可知对通过等时圈形状分类得到站点进行客流预测具有较好的预测精度, 说明基于等时圈形状的站点分类方法具有一定的合理性。同时, 条状等时圈类站点的误差在MAPE、MAE、RMSE三项指标上均优于团状等时圈类站点, 由此表明对条状等时圈类站点的客流量进行预测可以得到更高的预测精度。

表1 各类站点模型预测评价指标

站点类别	MAPE	MAE	RMSE
团状等时圈类站点	8.53%	866	1108
条状等时圈类站点	6.23%	253	317

四、结论

1) 通过细长度和偏心圆比两个形状指标, 有效描述轨道站点等时圈形状特征, 能较好的对轨道站点进行筛分。

2) 通过讨论形状特征值的取值范围, 确定细长度临界值点为2.5, 可以更准确的描述条状等时圈类站点。

3) 建立ARIMA模型对两类站点的进站客流进行预测, 得到了较好的预测效果, 且条状等时圈类站点的预测精度高于团状等时圈类站点。

参考文献:

- [1]但媛, 周涛, 乐伍杉, 李雪. 山地城市轨道交通步行可达性评估方法与实践[C]// 创新驱动与智慧发展——2018年中国城市交通规划年会论文集. 2018: 2096-2108.
- [2]杨静, 吴可, 张红亮, 等. 基于土地利用及客流特征的地铁车站分类[J]. 交通运输系统工程与信息, 2021.
- [3]姚春羽, 金立军, 闫书佳. 电网巡检图像中绝缘子的识别[J]. 系统仿真学报, 2012, 24(009): 1818-1822.
- [4]姚浪. 沿黄城镇带(陕北段)传统村落空间形态与优化策略研究[D]. 长安大学, 2021.
- [5]浦欣成. 传统乡村聚落二维平面整体形态的量化方法研究[D]. 浙江大学, 2012.
- [6]朱才华, 孙晓黎, 李培坤, 张景辰, 李岩. 融合车站分类和数据降噪的城市轨道交通短时客流预测[J/OL]. 铁道科学与工程学报.