

基于改进两步移动搜索法的公交站点可达性分析

刘正彪

【摘要】研究公交站点可达性对公交站点选址规划的合理性具有重要意义。本文以高斯距离衰减函数为基础,以 Python、ArcGIS、高德路径规划 API 为支撑,对传统两步移动搜索法进行改进。以成都双流区公交站点为例,从供需角度评价公交站点的空间可达性,发现双流区公交站点可达性存在南北差异明显、资源分配不均等问题。公交站点密集地区受到高出行需求的影响,实际可达性较低;站点分散区域的居民出行需求相对较低,局部反而出现高可达性。本文丰富了城市公交站点可达性评价案例,提出的基于高德路径规划 API 的高斯两步移动搜索法不仅可应用于城市公交站点选址规划,也可为提高公共交通服务水平提供思路参考。

【关键词】公共交通;可达性评价;两步移动搜索法;路径规划;距离衰减

1 引言

随着我国社会经济不断发展和城镇化持续推进,城市交通出行需求越来越高,小汽车保有量高速增长,城市交通拥堵越来越严重。公共交通以其大运量、低能耗、土地利用率高等特点逐渐成为解决城市交通拥堵的首选方案。公共汽车作为公共交通的重要组成部分,其服务水平高低决定着公共交通的吸引力水平,公交站点可达性作为衡量公共交通服务水平的指标之一,对构建公共交通体系、优化公交站点布局、提高居民出行效率、带动站点周边经济发展具有重要意义。

公共交通可达性一般分为单个站点可达性和交通网络可达性^[1],单个站点可达性指出行者使用不同交通方式到站点的便捷程度;交通网络可达性指从某地出发,利用公共交通网络到达目的地的便捷程度,一般在一定时间、距离阈值内所获得的概率作为度量指标。针对公交站点可达性的研究主要分为两个方面:一是从公交站点供给能力出发,通过网络分析法、重力模型法等方法研究公交站点的服务范围和服务能力;二是从居民对公交站点服务的获得性出发,基于空间句法、最小距离法等方法研究从交通需求到公交站点的最短距离。LI 等以空间句法为基础,利用广义出行费用函数对地铁站点进行了综合可达性分析^[2];姚志刚等运用考虑了路网距离缓冲区方法中距离衰减的因素,对传统直线缓冲区法在计算站点可达性方面进行了优化^[3];陈锦渠等基于乘客广义出行费用和空间句法,对站点可达性从易达性和可动性两个方面进行定量研究^[4],这些研究在一定程度上反映了公交站点的可达性情况,但较少从居民出行需求视角分析供给点与需求点之间的可达性。

采用两步移动搜索法(2SFCA)对可达性进行分析时,全面考虑了供给和需求两方面,

在诸多公共服务设施可达性评价中广泛应用，但受限于传统时空数据精度的不足，传统两步移动搜索法中的供给点到需求点之间的距离为两点之间的直线距离^[5]，本文通过 World Pop 网站获取开源人口数据，基于第七次全国人口普查数据进行人口校正，通过 ArcGIS 将校正后的人口数据关联到渔网栅格，使用高德路径规划 API 计算供给点到需求点的距离成本；传统两步移动搜索法中超过搜索阈值的供给点均不可达，在搜索阈值内的供给点均可达，未考虑到出行意愿会随着出行时间、距离的增加而降低本文，本文引入高斯距离衰减函数对此进行修正，提高其计算结果的精确性、有效性；最后通过高斯混合聚类分析造成公交站点可达性空间差异性的原因。

2 模型与方法

2.1 研究框架

从供需角度出发，影响公交站点可达性的因素主要有以下三方面：一是公交站点的供给能力，公交站点的供给能力由空间位置以及站点线路数量、站点发车间隔等因素决定；二是居民需求水平，人口分布密集会导致站点供给不足，人口分散则会降低公共交通使用效率；三是交通距离成本，出行距离越远，出行意愿越低。本文构建了一个城市公交站点可达性评价及聚类分析的研究框架，主要研究内容包括供需分析、可达性评价与聚类分析等。首先供给点为研究范围内所有公交站点，其搜索阈值为公交站点服务范围，需求点为研究范围内人口栅格数据，并通过第七次全国人口普查数据进行校正；其次，利用高德路径规划 API 计算供需点间的距离成本，运用改进高斯两步移动搜索法计算公交站点可达性；最后，从可达性程度、人口密集程度以及距离成本 3 个方面进行高斯混合聚类分析，剖析公交站点可达性差异化的空间分布及其成因，为优化公交站点布局提供参考。

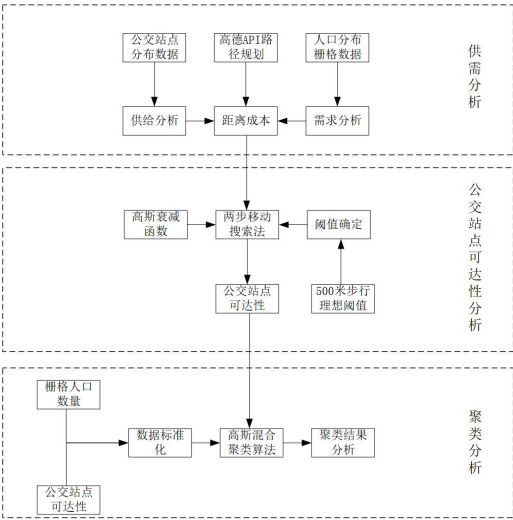


图 1 公交站点可达性研究框架

2.2 基于改进高斯两步移动搜索法的公交站点可达性评价分析

两步移动搜索法综合考虑了供给和需求两个方面，在公共设施（如公园绿地、学校、医疗机构以及商务办公等）的可达性研究方面得到了广泛应用，可以全面对公交站点可达性进行计算^[6]。通过距离衰减函数优化传统两步移动搜索法距离成本中，利用高斯函数建立空间距离衰减函数效果最好，该方法基于公交站点服务水平确定公交站点的搜索阈值范围，得到阈值内的供需比，从需求点角度计算出搜索阈值范围内的可达性^[7]，具体计算步骤如下：

Step 1 针对研究区域每一个公交站点 j ，以 j 作为中心点，以 d_0 作为搜索半径，搜索该区域全部人口数量，采用高斯函数根据距离衰减规律进行权值赋值，将这些权值进行人口数量汇总相加得到公共设施的供需比 R_j ：

$$R_j = \frac{S_j}{\sum_{k \in \{d_{ij} \leq d_0\}} G(d_{ij}) D_k}$$

式中： S_j 为公交站点服务水平； D_k 指各需求单元人口数量； d_{kj} 为地点 k 、 j 之间的路径规划距离， $G(d_{ij})$ 作为高斯距离的衰减函数以数学形式表示：

$$G(d_{ij}) = \frac{e^{-\frac{1}{2} \times \left(\frac{d_{ij}}{d_0}\right)^2}}{1 - e^{-\frac{1}{2}}} \quad (d_{ij} < d_0)$$

Step 2 将任意一个公交站点 i 作为需求点，以出行者前往该公交站点的最大路径距离 d_0 作为半径，构建搜索区域 I ，查找所有搜索区域内的公交站点 j ，基于高斯衰减函数，将这些公交站点的供需比 R_j 进行汇总并求和，从而得到了供给点基于距离成本的公交站点可达性 A_i^D ，该值越大表示可达性程度越高：

$$A_i^D = \sum_{j \in \{d_{ij} \leq d_0\}} G(d_{ij}) R_j$$

Step 3 传统的两步移动搜索法是基于距离成本来构建的可达性计算模型。距离成本可细分为两部分：一是时间成本，二是距离成本。其中，时间成本可以进一步分为两点之间的直线距离和两点之间的路径距离，时间成本可以用供给和需求两点间出行时间表达，基于路径规划的时间成本兼顾路径选择，道路畅通性等因素，可以更好地反映出可达性研究的现实价值。本文采用基于路径规划的时间成本进行可达性计算模型，搜索范围是到达公交站点的极限通行时间 t_0 （替换路网极限距离 d_0 ），将供需点间的路径规划通行时间 t_{ij} 作为衡量 OD 通行成本的指标，进一步改进可达性模型：

$$A_i^T = \sum_{j \in \{t_{ij} \leq t_0\}} G(t_{ij}) \frac{S_j}{\sum_{k \in \{t_{kj} \leq t_0\}} G(t_{ij}) D_k}$$

式中： A_i^T 为基于时间成本的公交站点可达性。

2.3 高斯混合聚类分析

聚类分析可以将需要分析的数据按照某些相似特征进行分类,有助于研究数据集内在的相似性与差异性。聚类模型中的高斯混合聚类模型收敛速度快,适用于大规模数据集,故采用高斯混合聚类算法对公交站点进行聚类分析。采用高斯混合聚类分析方法,获得研究区域内公交站点可达性及其供需组合模式,并从供给、需求及出行成本等方面,剖析影响可达性差异的原因,识别不同供需组合模式下影响可达性的关键制约性因素,并针对不同组合方式提出有针对性的政策建议。

将公交站点分为 k 类,每个类别的站点都是由单个单高斯分布生成,但具体站点归属哪一种高斯分布未知,故假设各站点分别用 k 个单高斯分布混合模型来表示,即站点由高斯混合模型产生的概率密度函数如下:

$$p_M(x) = \sum_{i=1}^k \alpha_i \cdot p(x|\mu_i, \Sigma_i)$$

式中, x 是维度为 n 的随机向量; i 为单高斯模型的数量; μ_i 为第 i 个单高斯分布的均值; Σ_i 为第 i 个单高斯分布的方差; α_i 为混合系数; $p(x|\mu_i, \Sigma_i)$ 为第 i 个单高斯分布的概率密度函数,可表示为:

$$p(x|\mu_i, \Sigma_i) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_i|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_i)^T \Sigma_i^{-1}(x-\mu_i)}$$

使用高斯混合模型进行聚类分析的具体步骤如下:

1) 估计数据由每个单高斯分布生成的概率,对于每个数据 x_j 来说,它由第 i 个单高斯分布生成的概率为:

$$p_M(z_j = i|x_j) = \frac{a_i P(x_j|\mu_i, \Sigma_i)}{\sum_{l=1}^k a_l P(x_j|\mu_l, \Sigma_l)}$$

2) 对每个高斯分布参量进行估计,并用极大似然法对其参数进行估计。

①样本均值 μ_i

可令 $\gamma_{ji} = p_M(z_j = i|x_j)$, 得到:

$$\mu_i = \frac{\sum_{j=1}^m \gamma_{ji} x_j}{\sum_{j=1}^m \gamma_{ji}}$$

②协方差矩阵

$$\Sigma_i = \frac{\sum_{j=1}^m \gamma_{ji} (x_j - \mu_i)(x_j - \mu_i)^T}{\sum_{j=1}^m \gamma_{ji}}$$

③混合系数

$$\alpha_i = \frac{\sum_{j=1}^m \gamma_{ji}}{m}$$

3) 重复迭代直至似然函数收敛。具体算法流程如图 2 所示。

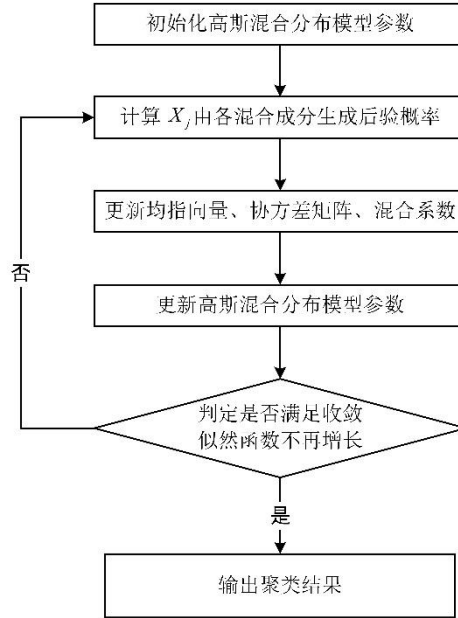


图 2 高斯混合聚类算法流程图

3 公交站点可达性评价——以成都市双流区为例

3.1 研究区域与数据预处理

成都市是典型的平原城市，其城市发展模式为中心向四周摊大饼结构，中心区域发展迅速，在三环内轨道网络与公交线网组成密集的公共交通网络，路网密度大，体系完善，相对于三环内，三环外区域轨道线网覆盖范围降低，公共汽车能够很好地衔接轨道交通，两者高效互补有助于公共交通发展，方便居民出行。

本文选取三环外的双流区公交站点作为研究范围，双流区为天府新区的核心，良好的公交站点可达性能够促进城市经济发展，缓解城市交通拥堵，通过分析双流区公交站点居民的空间可达性可为提高居民出行效率、城市公共交通站点布局规划提供决策参考。

主要数据及来源如下：

(1) 公交站点数据。基于高德开发平台中的地图 API (Application Programming Interface), 利用 Python 数据爬取工具, 定向爬取双流区范围内的 1124 条公共线路, 13959 个公交站点数据。

(2) 人口数据。通过 World Pop 官方网站获取 2020 年的中国 100 米分辨率的人口数据, 根据成都市双流区第七次人口普查数据进行校正, 利用 ArcGIS10.8 对研究区域进行 200 米

×200 米的渔网切割，共得到 14462 个格网。将校正后的人口数据关联到 200*200m 的网格上，用以表示居民的出行需求。

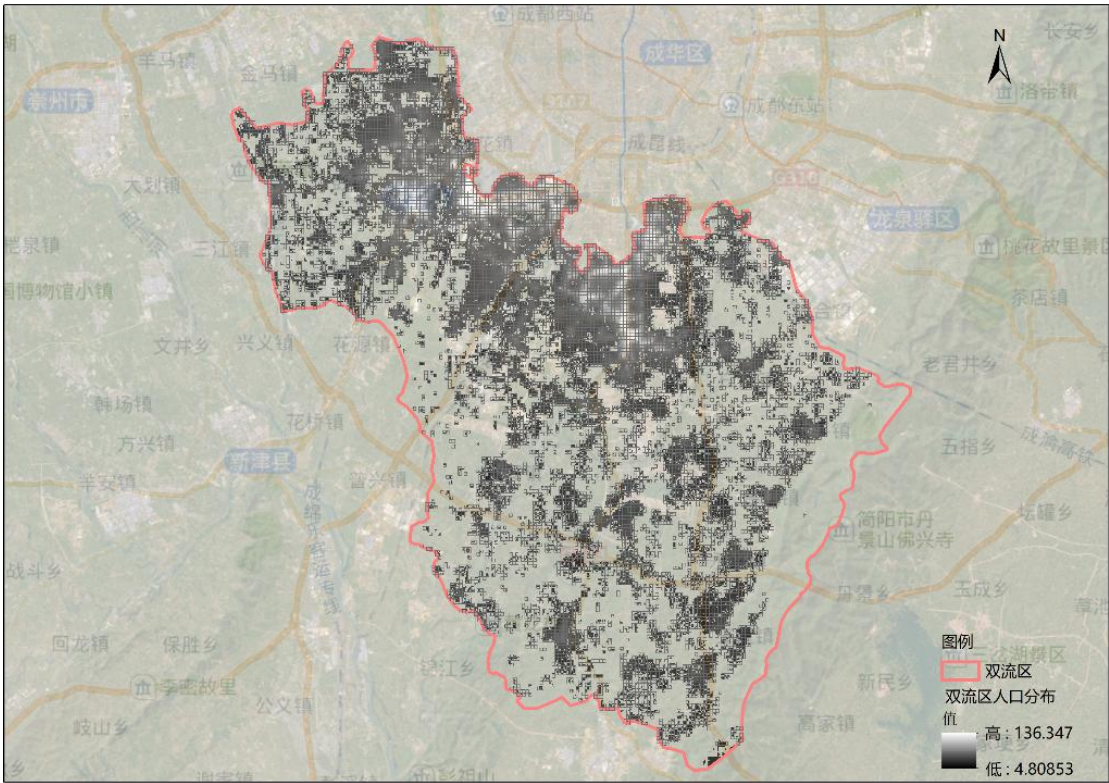


图 3 双流区人口分布

3.2 评价结果与分析

在《城市综合交通体系规划标准》中对城市公共汽车（电车）的覆盖区域作出以下规定：300 米城市公交站点服务半径应覆盖超过 50%的城市建设用地；500 米城市公交站点服务半径应覆盖超过 90%的城市建设用地。本文以 500 米作为服务半径阈值，对公交站点的可达性进行计算。在改进高斯两步移动搜索法计算公式中，分别考虑了供给量、需求量以及规划时间成本，计算得到成都市双流区公交站点可达性计算结果，分析其可达性差异性。

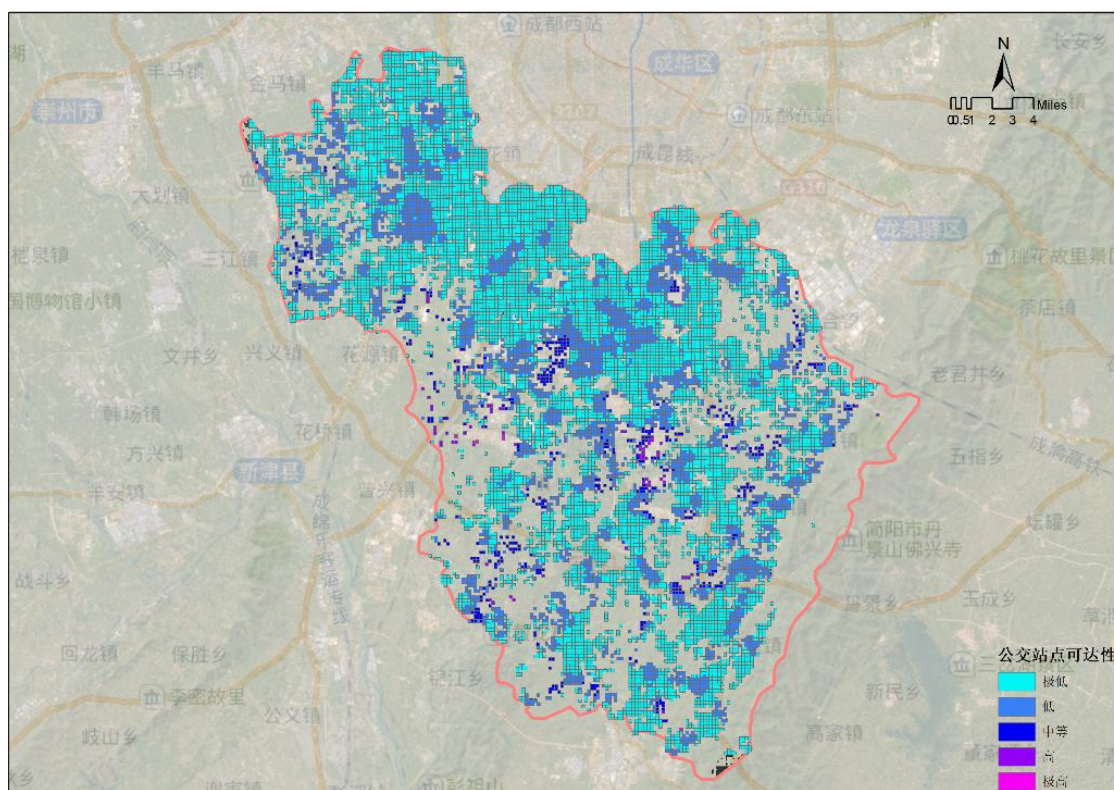


图 4 公交站点可达性分布（500 米搜索阈值数据统计）

表 1 公交站点可达性统计表

可达性水平	栅格数量	栅格占比	人口数量	人口占比
极低	10119	69.970%	1072398	73.288%
低	3791	26.214%	376111	25.704%
中等	475	3.284%	13371	0.914%
高	71	0.491%	1314	0.090%
极高	6	0.041%	73	0.005%

从可达性空间分级结果来看，双流区公交站点可达性由北向南呈现出由密集到疏散，将可达性计算结果关联到空间栅格数据上，使用自然间断点法将其分为 5 级，可以发现双流区北部最强，这是由于成都市城市发展是由北向南，北部公共交通发展相对完善；南部由于人口相对分散，其可达性水平较低；东部由于地形因素，人口基本沿龙泉山脉分布；西部人口分散且人口数量较少，故可达性水平较低。

为进一步识别并剖析双流区公交站点可达性差异及成因，在具体的研究单元中研究公交站点可达性差异性，以便更全面、更具针对性地制定可达性优化的政策建议，对 14462 个关联人口数据的栅格进行高斯混合聚类法分析，聚类指标包括公交站点可达性与各渔网所关联的人口数量，如图 5、表 2 所示为聚类高斯混合聚类结果，本文根据以栅格平均可达性水平（3.97）为界限，高于该界限的区域称为高可达性区，反之，为低密度区；以栅格平均人口

数量（101）为阈值，高于该阈值的区域为高密度区，反之，为低密度区。根据各项指标均值对各分区进行命名：低可达性低人口密度区、高可达性低人口密度区、低可达性高人口密度区、高可达性高人口密度区。

由聚类结果可知，大部分高可达性高人口密度区（栅格占比 30.34%）分布在双流区北部，人口密度高，公交站点可达性高；部分低可达性高人口密度区（栅格占比 9.51%）人口密度较高，导致公交站点可达性水平较低；大部分低可达性低人口密度区（栅格占比 56.09%）处于相对较远的地理位置，人口较为分散，集中分布在双流区南部，高可达性低人口密度区域主要集中在公交站点集中的区域。

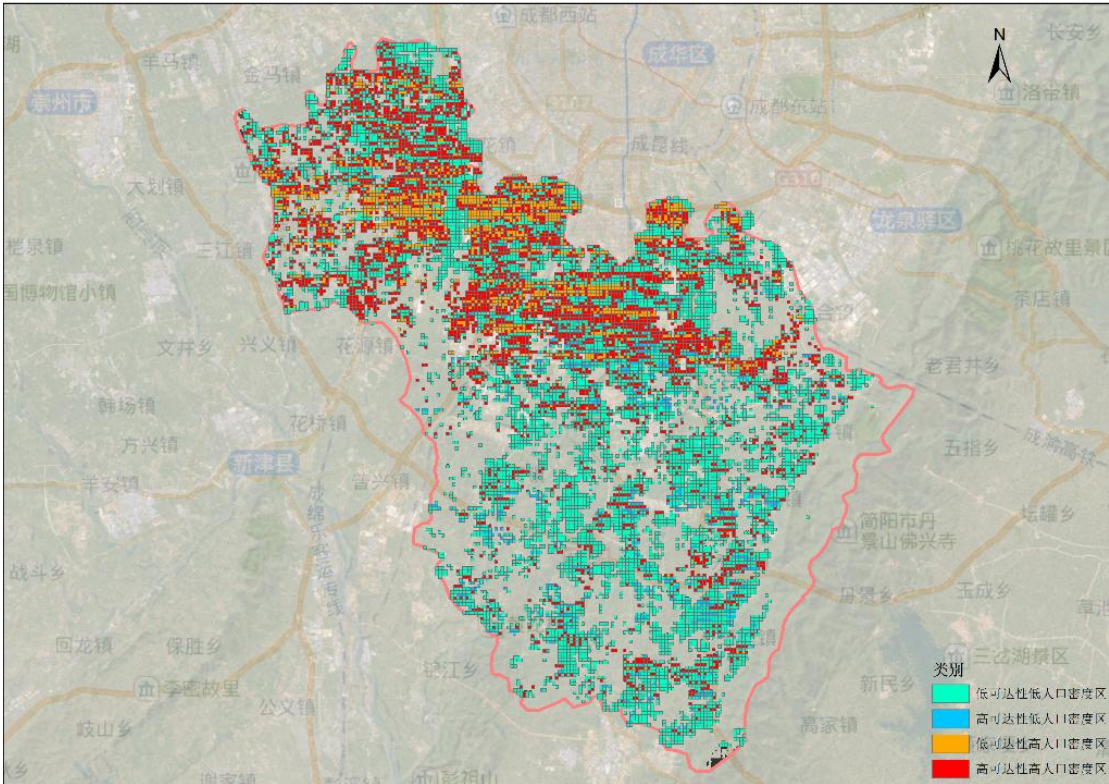


图 5 公交站点可达性聚类结果

表 2 聚类结果统计						
类别	可达性均值	人口数量	人口占比	栅格数量	栅格占比	人口密度
第一类（低可达性低人口密度区）	3.67	623464	42.61%	8112	56.09%	76.85
第二类（高可达性低人口密度区）	6.10	40720	2.78%	586	4.05%	69.48
第三类（低可达性高人口密度区）	3.32	259066	17.70%	1376	9.51%	188.27
第四类（高可达性高人口密度区）	4.43	540021	36.91%	4388	30.34%	123.07

4 结论与讨论

本文以成都市双流区为例,针对公交站点设置 500 米搜索半径,以校正后 world pop 人口数据关联的渔网数据评价居民对公交站点的需求量,采用改进高斯两步移动搜索法对公交站点进行可达性评估,通过高斯混合聚类算法分析公交站点可达性与人口聚集程度组合模式的差异及其成因。得出如下结论:

(1) 针对传统两步移动搜索法中忽略随着距离的增加出行意愿随之降低的弊端,本文使用高斯衰减函数进行改进,同时通过高德路径规划 API 替代了直线距离成本;使用第七次全国人口普查数据对 World Pop 官方网站的人口数据进行校正,从供给到需求两个方面提高公交站点可达性计算准确性,为医疗、学校、公园绿地等公共服务设施的可达性评价以及大数据在城市公共交通专项规划领域的应用提供了思路。

(2) 双流区公交站点可达性和人口密度均呈现出以北部为高值聚集区向南部逐渐分散的分布特征。当考虑供需关系时,双流区北部则出现因人口密度大而呈现出的可达性较低;双流区南部由于人口分散呈现出可达性水平较低,外围边缘地区人口点状聚集导致部分区域出现高可达性高人口密度。综上,在双流区北部由于人口高度聚集地区应补充性规划公交站点,在南部外围边缘区域应结合人口聚集中心适当建设公交站点。

本文对公交站点可达性评价方法及模型成因分析等方面具有一定创新之处,但建模过程中仍然存在不足之处,如仅用居民人口数表示公交需求量,以公交服务半径表示供给量,没有考虑不同公交站点供给能力差异性,同时未考虑居民出行时交通方式的选择对公交可达性产生的影响。未来研究可利用大数据获取不同人群出行方式的选择,从而提高可达性计算精度和有效性,为公交站点选址规划提供更有价值的决策依据。

参考文献

- [1] MALEKZADEH A, CHUNG E. A review of transit accessibility models: Challenges in developing transit accessibility models[J]. International Journal of Sustainable Transportation, 2020, 14(10): 733-748.
- [2] LI L, REN H, ZHAO S, et al. Two dimensional accessibility analysis of metro stations in Xi'an, China[J]. Transportation Research Part A: Policy and Practice, 2017, 106: 414-426.
- [3] 姚志刚,傅宇豪,张俊青. 站点可达性方法对公交不平等测度的影响研究[J]. 交通运输系统工程与信息, 2021, 21(03): 206-213.
- [4] 陈锦渠,吴雨遥,殷勇等. 城市轨道交通站点可达性研究[J]. 交通运输工程与信息学报, 2020, 18(04): 38-45.

[5] 张政, 蔡军. 基于改进两步移动搜索法的轨道站点可达性研究——以天津市津城区域为例[C]//中国城市规划学会城市规划新技术应用学术委员会. 夯实数据底座·做强创新引擎·赋能多维场景——2022年中国城市规划信息化年会论文集. 广西科学技术出版社, 2022:9.

[6] 仝德, 孙裔煜, 谢苗苗. 基于改进高斯两步移动搜索法的深圳市公园绿地可达性评价[J]. 地理科学进展, 2021, 40 (07) : 1113-1126.

[7] 袁红霞, 徐菱, 严余松等. 基于改进两步移动搜索法的轨道交通接驳可达性分析[J]. 交通运输系统工程与信息, 2022, 22 (04) : 113-119.

作者简介

刘正彪, 男, 硕士, 江苏省城市规划设计研究院, 助理工程师。电子邮箱:
769205681@qq.com