

基于空间可达模型的名人文旅景点 k-近邻聚类算法

周啸¹ 潘娟¹ 赵太萍¹ 田江鹏²

(1. 乐山职业技术学院文化旅游学院, 四川 乐山 614000; 2. 信息工程大学地理空间信息学院, 河南 郑州 450001)

摘要: 在分析问题背景的基础上, 构建了文旅景点和交通枢纽之间的空间可达模型, 以此为基础建立了文旅景点 k-近邻空间聚类方法。以四川省内江市及张大千文旅景点为研究对象, 首先采集了市内主要交通枢纽和张大千文旅景点的地理空间数据, 建立了以空间坐标为基础的空间可达模型, 计算张大千文旅景点与交通枢纽之间的空间可达度。在搜索张大千文旅景点与交通枢纽最优空间可达度的基础上建立了空间聚类模型, 确定了各文旅景点与交通枢纽之间的空间关系, 为名人文旅景点推荐服务和旅游决策服务提供了科学量化依据, 同时也为其他旅游城市提供了参照。

关键词: 空间可达模型; 文旅景点; k-近邻聚类; 旅游规划

基金项目: 四川省社会科学重点研究基地——地方文化资源保护与开发研究中心资助课题 (四川名人文化旅游空间的模型构建, DFWH2024-012); 乐山市哲学社会科学规划项目 (基于“人工智能+文旅”创新融合模式的乐山非遗文化旅游空间决策模型构建, SKL2025B12)

DOI: doi.org/10.70693/rwsk.v2i4.380

一、研究背景

旅游交通是旅游者行前规划的重要决策因素, 是决定出行成本的必要条件。狭义的旅游交通指的是旅游者从旅游客源地出发抵达旅游目的地城市采用的交通方式及其产生的出行成本, 同时也包括从旅游城市内的出发点抵达旅游目的地景点采用的市内交通工具及其产生的出行成本。通常, 旅游者乘坐交通工具前往旅游城市时首先到达城市的重要交通枢纽, 如机场、火车站、高铁站、汽车站、公交枢纽站等, 它们是旅游者到达城市从事旅游活动的首要目的地。旅游者抵达交通枢纽后需换乘市内交通工具前往旅游景点, 从交通枢纽前往旅游景点的过程中, 乘坐的市内交通工具是产生旅游成本的主要因素。本文研究中, 对于某个旅游城市而言, 不讨论旅游客源地到旅游目的地城市之间的旅游交通和成本, 重点讨论旅游者从市内交通枢纽抵达旅游景点的交通方式和成本, 以构建市内旅游景点空间聚类模型, 分析景点基于不同交通枢纽所产生的不同交通成本。郑云扬等^[1]对南昌市内的旅游景点进行了空间可达的特征分析, 重点对景点分布状态进行了论述。黄远君等^[2]在多种交通系统模式下讨论了南宁市内的景点空间可达性, 考虑了不同交通方式下景点空间可达的不同性。李婷等^[3]对影响福州市内景点空间结构的因素进行了分析, 并对福州市景点发展提出了建议。周海涛等^[4]以省域为单位分析了内蒙古红色旅游景点的空间分布特征, 为规划红色旅游线路提供了参考依据。青果等^[5]基于空间句法构建了广西阳朔县旅游景点的空间可达模型。对景点空间可达性研究进行分析可知, 城市景点空间可达性研究主要围绕几个方面进行, 一是研究景点的总体空间可达性, 分析其空间可达特征, 对景点分布状态进行规划分析; 二是研究影响空间可达性的主要因

作者简介: 周 啸 (1986—), 男, 博士后, 教授, 高级工程师。研究方向为智慧旅游;

潘 娟 (1986—), 女, 硕士, 讲师, 研究方向为旅游管理;

赵太萍 (1990—), 女, 硕士, 讲师, 研究方向为旅游管理。

田江鹏 (1987—), 男, 博士, 讲师, 研究方向为地理信息系统

通讯作者: 周啸

素。通过对主要因素的分析做出进一步旅游规划；三是结合交通方式研究景点空间可达性，为旅游部门决策规划提供依据。

分析研究现状可知，现有的景点空间可达分析方法仅从计算空间可达度层面分析景点通达的难易程度，没有深入研究和获取景点之间、景点与交通枢纽之间的空间关系知识，也没有从数据挖掘角度为旅游管理部门和旅游者提供决策依据，复杂的空间可达度数据难以转化成决策知识。因此，景点空间可达度分析还需要融合数据挖掘方法，以深入研究旅游者与景点之间的空间关系规律，获取空间关系知识^[6]。当旅游者从不同交通枢纽所在点出发时，市内景点对于交通枢纽而言具有不同空间可达度，将产生不同空间聚类分布状态，导致出行交通成本产生较大差异，这是影响旅游决策的重要知识和依据。

二、研究方法与模型构建

以旅游城市四川省内江市及其文旅景点为例，研究目标是构建基于不同交通枢纽可达度的景点空间聚类模型，挖掘不同交通枢纽条件下文旅景点的空间分布规律，为旅游管理部门进行合理交通规划和旅游决策规划提供科学依据。研究的基本流程为：第一步，确定内江市内主要交通枢纽和文旅景点，获取其基本地理信息数据，如经纬度坐标。第二步，以每个交通枢纽为中心构建交通枢纽与文旅景点之间的空间可达模型，计算文旅景点与交通枢纽之间的空间可达度。第三步，构建以交通枢纽为中心的文旅景点空间聚类模型，分析不同交通枢纽条件下文旅景点的空间聚类特征，获取为旅游管理部门提供决策支持的空间知识。

（一）研究对象及地理信息数据采集

内江市位于四川省东南部，是重要的交通枢纽和旅游城市，以张大千文化旅游著称，市内分布着数量众多、规模各异的文旅景点。近年来，内江市交通发展取得了辉煌成就，市内交通便捷，拥有内江火车站（A1）、内江高铁站（A2）、东兴汽车客运站（A3）等重要交通枢纽，铁路运输和公路运输发达，与周围省市的联系非常密切，为旅游提供了便捷的交通条件^[7]。内江是中国国画大师张大千先生的故乡，市内以张大千为主题的文旅景点众多，代表性文旅景点有大千园（T1）、内江市文化馆（T2）、东兴区文化馆（T3）、市中区文化馆（T4）、大千广场（T5）、东兴老街（T6）、大佛寺（T7）等。以内江市主要交通枢纽和代表性文旅景点为研究对象，采集研究对象的地理信息数据如表1所示。

表1. 内江市主要交通枢纽和文旅景点地理空间坐标

	主要交通枢纽			代表性文旅景点						
	A1	A2	A3	T1	T2	T3	T4	T5	T6	T7
经度	105.037	105.084	105.070	105.059	105.051	105.074	105.065	105.071	105.074	105.062
纬度	29.587	29.609	29.603	29.593	29.576	29.590	29.579	29.592	29.589	29.583

（二）景点空间可达模型构建

空间可达的概念用于描述空间中的某一点抵达空间中另一点的难易度，通常用空间直线距离或两点间的通路距离表示。两点之间的空间可达度越高，说明两点之间越容易通达，两点之间的空间可达度越低，说明两点之间越不容易通达^[8]。当用空间经纬度表达点的坐标时，可以用经纬度计算两点之间的直线距离表达空间可达度。式

（1）为构建的交通枢纽与代表性文旅景点之间的空间可达模型。

$$S_A = \delta \times \frac{1}{\sqrt{|x_1 - x_2|^2 + |y_1 - y_2|^2}} \quad (1)$$

式(1)中 S_A 表示空间可达度, δ 表示归一化因子, 将空间可达度约束在(0,1)区间内。 (x_1, y_1) 表示交通枢纽的经纬度坐标, (x_2, y_2) 表示代表性文旅景点的经纬度坐标。

(三) 文旅景点空间聚类模型构建

空间可达度反映各文旅景点与交通枢纽间的通达难易程度, 为内江市文旅景点空间数据挖掘提供了基础数据。空间数据挖掘的过程是通过构建空间聚类算法对代表性文旅景点的潜在空间关系进行知识获取的过程^[9]。内江交通枢纽地理坐标、文旅景点地理坐标和文旅景点空间可达度均为离散无规则数据, 从中挖掘对旅游决策规划和旅游者选取景点有用的空间知识, 是建立空间数据挖掘模型的核心目标。基于内江交通枢纽地理坐标、文旅景点地理坐标和文旅景点空间可达度构建文旅景点空间聚类模型, 获得文旅景点与交通枢纽之间、文旅景点与文旅景点之间的聚类关系, 可以为旅游者抵达后确定交通枢纽最优选址、节约出行成本提供决策依据, 同时也为旅游管理部门提供决策和规划的依据。

在空间聚类算法中, k-近邻聚类具有计算简便、算法简单、易于实现的特点, 利用k-近邻聚类算法能够有效实现空间分布实体的聚类。k-近邻聚类算法中, k值的选取具有特殊性, 其数值大小决定空间聚类簇的数量^[10]。本研究中, 定义k值为内江市交通枢纽的数量, 聚类目标定义为搜索与交通枢纽具有最优空间可达度的文旅景点。根据建模基本原理, 构建式(2)所示的文旅景点空间聚类矩阵, 矩阵的每一行代表1个景点聚类簇 $C_{(i)}$, 每个元素 T_{ij} 代表文旅景点, 其中, i 为景点聚类簇的编号, j 为景点簇内景点的编号。

$$C = \begin{bmatrix} T_{11} & T_{12} & \dots & T_{1n} \\ T_{21} & T_{22} & \dots & T_{2n} \\ \dots & \dots & \dots & \dots \\ T_{m1} & T_{m2} & \dots & T_{mn} \end{bmatrix} \quad (2)$$

根据k-近邻聚类算法逻辑和内江市文旅景点空间特征, 构建以最优空间可达度模型为聚类目标函数的改进k-近邻空间聚类算法, 算法步骤如下:

Step 1: 确定k个聚类种子点, 每个种子点为1个交通枢纽, $p_{(i)}$ 为第*i*号交通枢纽, $i = 1, 2, \dots, k$ 。

Step 2: 确定文旅景点数量为 a , 标记每个文旅景点为 $T_{(x)}$, $x = 1, 2, \dots, a$ 。

Step 3: 遍历 $x = 1, 2, \dots, a$, 计算 $T_{(x)}$ 与 $p_{(i)}$ 间的空间可达度 $S_A(T_{(x)}, p_{(i)})$:

Step 3.1: 遍历 $i = 1, 2, \dots, k$, 计算 $T_{(1)}$ 与 $p_{(i)}$ 间的空间可达度 $S_A(T_{(1)}, p_{(i)})$, 取 $\max S_A(T_{(1)}, p_{(i)})$ 对应的 $p_{(i)}$, 其中 $p_{(i)}$ 的编号*i*对应第*i*个景点聚类簇。将 $T_{(1)}$ 存入矩阵第*i*行第1个元素 T_{i1} 。

Step 3.2: 遍历 $i = 1, 2, \dots, k$, 计算 $T_{(2)}$ 与 $p_{(i)}$ 间的空间可达度 $S_A(T_{(2)}, p_{(i)})$, 取 $\max S_A(T_{(2)}, p_{(i)})$ 对应的 $p_{(i)}$, 其中 $p_{(i)}$ 的编号*i*对应第*i*个景点聚类簇。

(1) 若 $T_{i1} = 0$, 将 $T_{(2)}$ 存入矩阵第*i*行第1个元素 T_{i1} ;

(2) 若 $T_{i1} \neq 0$, 将 $T_{(2)}$ 存入矩阵第*i*行第2个元素 T_{i2} ;

Step 3.3 按照相同算法, 计算 $T_{(x)}$ 与 $p_{(i)}$ 间的空间可达度 $S_A(T_{(x)}, p_{(i)})$, 判断 $T_{(x)}$ 归于第*i*个景点聚类簇, 并存储 $T_{(x)}$ 至第*i*行元素。

Step 4: 取具有最大元素的行*i*中的景点 $T_{(x)}$ 数量为 n , 定义为矩阵C的列秩, 则有其他 $\neg i$ 行的元素数量 b 一定满足 $0 \leq b \leq n$ 。

Step 5: 第*i*行的前 b 个非零元素为属于聚类簇 $C_{(i)}$ 的景点 $T_{(x)}$, 后 $n-b$ 个元素设置为0。聚类算法结束。

三、实证研究与分析

我们将模型应用于四川省区域旅游空间范围，以四川省内江市张大千名人文化旅游为例，通过对内江市交通枢纽和代表性文旅景点进行地理信息数据采集，计算各文旅景点对于每个交通枢纽的空间可达度，结果如表 2 所示。表 2 中的符号规则为：

- 内江火车站 (A1)
- 内江高铁站 (A2)
- 东兴汽车客运站 (A3)
- 大千园 (T1)
- 内江市文化馆 (T2)
- 东兴区文化馆 (T3)
- 市中区文化馆 (T4)
- 大千广场 (T5)
- 东兴老街 (T6)
- 大佛寺 (T7)

图 1 为每个交通枢纽条件下文旅景点空间可达度趋势图。分析表 2 和图 1 结果数据可知，各文旅景点对于交通枢纽的空间可达度不同，当景点与某个交通枢纽的空间可达度最大时，该文旅景点趋于该交通枢纽所在的聚类簇。由图 1 (1) - (3) 可知，旅游者选择任意交通枢纽的条件下，文旅景点的空间可达度呈现波动趋势。从空间分布结果上看，景点更趋于东兴汽车客运站 (A3) 所在的聚类簇。只有个别景点趋于内江火车站 (A1) 所在的聚类簇。而对于空间距离较远的内江高铁站 (A2)，各文旅景点与其空间亲密度最低。

表 2. 张大千文旅景点对于每个交通枢纽的空间可达度

交通枢纽	空间可达度			
	T1	T2	T3	T4
A1	0.439	0.562	0.269	0.343
	T5	T6	T7	
	0.291	0.270	0.395	
A2	T1	T2	T3	T4
	0.337	0.214	0.466	0.282
	T5	T6	T7	
A3	0.467	0.447	0.294	
	T1	T2	T3	T4
	0.673	0.303	0.735	0.408
	T5	T6	T7	
	0.905	0.687	0.464	

根据表 2 和图 1 的计算结果输出文旅景点聚类簇，如表 3 所示。由表 3 可知，归属于聚类簇 $C_{(1)}$ -A1：内江火车站的文旅景点有内江市文化馆 (T2)；归属于 $C_{(3)}$ -A3：东兴汽车客运站的文旅景点有大千园 (T1)、东兴区文化馆 (T3)、市中区文化馆 (T4)、大千广场 (T5)、东兴老街 (T6) 和大佛寺 (T7)。由内江高铁站 (A2) 构成的景点聚类簇 $C_{(2)}$ 是空集，即没有景点归属于景点聚类簇 $C_{(2)}$ 。分析聚类结果，内江市内绝大多数文旅景点集中在东兴汽车客运站所在的空间聚类簇 $C_{(3)}$ 范围内。因此，旅游者选择东兴汽车客运站 (A3) 作为抵达内江市

的交通枢纽目的地时，前往除内江市文化馆之外的文旅景点更为便捷，能够节约更多的交通成本。

内江市内所有代表性文旅景点中，仅有内江市文化馆（T2）一个文旅景点归属在景点聚类簇 $C_{(1)}$ 的空间范围内，旅游者若要前往内江市文化馆游览，且需要节约交通成本时，选择内江火车站（A1）作为抵达目的地可以使前往该景点的交通成本最低。对于内江高铁站（A2）而言，分析聚类簇计算结果和该站点在内江市内的地理空间位置可知，内江高铁站的位置距离主城区较远，因此内江高铁站距离各个文旅景点的空间距离也更远，空间可达度低，旅游者需花费的交通成本更高。从算法计算结果看，内江高铁站没有构成包含代表性文旅景点的空间聚类簇。旅游者若选择内江高铁站作为抵达内江市的交通枢纽目的地，前往文旅景点的交通成本会更高。

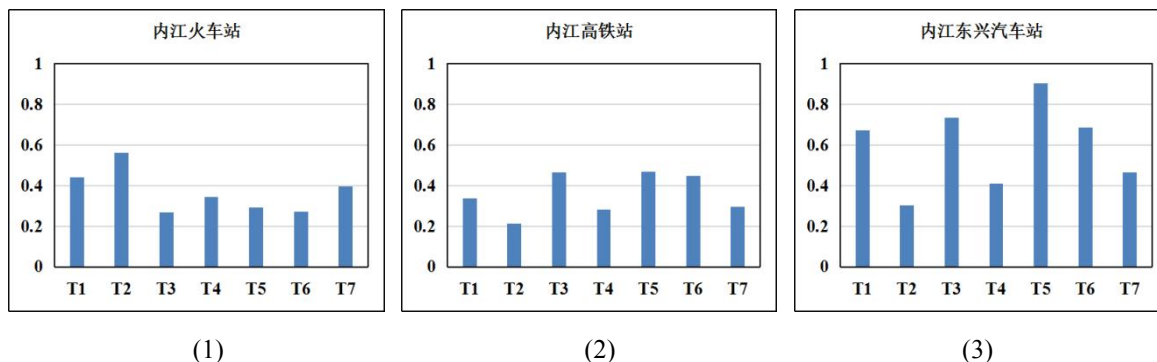


图 1.各交通枢纽条件下文旅景点空间可达度趋势图

基于表 2 和图 1 计算结果，我们进一步分析各个文旅景点的空间可达度优先级。集聚在景点聚类簇 $C_{(1)}$ 内的景点仅一个，旅游者前往内江市文化馆游览，选择内江火车站（A1）可以最大程度减少交通成本。集聚在景点聚类簇 $C_{(3)}$ 内景点较多，根据空间可达度计算结果可知，各景点的空间可达度优先级为：大千广场（T5）>东兴区文化馆（T3）>东兴老街（T6）>大千园（T1）>大佛寺（T7）>市中区文化馆（T4）>内江市文化馆（T2）。从优先级比较结果可知，旅游者选择东兴汽车客运站（A3）作为抵达内江市的交通枢纽目的地时，若聚类簇 $C_{(3)}$ 内的景点均满足旅游者兴趣，选择大千广场（T5）、东兴区文化馆（T3）、东兴老街（T6）和大千园（T1）这四个空间可达度相对更优的景点游览，可以最大程度减少交通成本。相比之下，若选择大佛寺（T7）、市中区文化馆（T4）和内江市文化馆（T2）这三个空间可达度相对较弱的景点游览，需要花费的交通成本较高。因此，从节约交通成本的角度，为旅游者推荐的文旅景点是：大千广场（T5）、东兴区文化馆（T3）、东兴老街（T6）和大千园（T1）。

表 3. 内江市文旅景点聚类簇计算结果

聚类簇	张大千文旅景点
$C_{(1)}$ -A1	内江市文化馆（T2）
$C_{(2)}$ -A2	∅（空集）
$C_{(3)}$ -A3	大千园（T1）；东兴区文化馆（T3）；市中区文化馆（T4）；大千广场（T5）；东兴老街（T6）；大佛寺（T7）

四、结论与决策建议

本文通过对内江市主要交通枢纽和代表性文旅景点进行地理信息数据采集，构建了文旅景点与主要交通枢纽之间的空间可达度模型。利用文旅景点与交通枢纽之间的空间可达度模型构建聚类目标函数，搜索与文旅景点空间可达度最优的交通枢纽，并将各文旅景点纳入交通枢纽所在的聚类簇。通过文旅景点空间聚类，得到文旅景点与交通枢纽之间的空间规律知识，为旅游者前往内江市游览文旅景点提供决策支持，为节约交通出行成本提供科学依据，也为旅游管理部门进行旅游规划和决策提供了依据。

根据研究结果提出以下旅游决策建议：

(1) 对于喜好名人文化的旅游者, 前往内江进行游览活动前, 提前掌握内江市交通枢纽和文旅景点的地理位置、分布情况以及景点特色、市内交通等信息非常必要, 有助于做好旅游规划, 确保游览过程的通畅、顺利。

(2) 内江市内绝大多数文旅景点集中在老城区东兴汽车客运站所在的空间聚类簇范围内, 因此对于省内旅游者以及内江市辖区旅游者, 乘坐客运班车前往内江后再转由市内交通前往景点, 是节约成本的最优选择。

(3) 内江高铁站处于新城, 距离文旅景点的距离较远, 因此, 高铁虽然是前往内江市的主导交通方式, 但对于节约市内交通成本并非最优选择, 旅游者应根据经济预算和时间成本综合考虑交通枢纽。

(4) 选择相对位置集中的东兴汽车客运站聚类簇包含的文旅景点, 可以在有限时间内游览更多的景点, 节约交通成本。

参考文献:

- [1] 郑云扬,李倩.南昌市旅游景点空间分布特征研究[J].中外建筑,2023,(10):63-69.
- [2] 黄远君,江培智,谢小魁等.基于多模式交通系统的旅游景区空间可达性分析——以南宁市为例[J].北部湾大学学报,2023,38(02):70-76.
- [3] 李婷,林斌,颜逸静等.福州市热门旅游景点空间结构及影响因素研究[J].闽江学院学报,2023,44(02): 103-109.
- [4] 周海涛,马钰松,樊亚宇等.内蒙古红色旅游资源空间分布及可达性分析[J].干旱区地理,2023,46(05): 814-822.
- [5] 青果,胡金龙,艾焯等.基于空间句法的旅游景点可达性研究——以广西阳朔县为例[J].河南科技学院学报(自然科学版),2022,50(03):32-40.
- [6] 肖政.基于空间数据挖掘的个性化旅游景点推荐系统研究[D].华中师范大学,2016.
- [7] 蔡袁.基于旅游目的地形象感知的内江文旅融合发展对策研究[J].内江科技,2024,45(07):19-21+76.
- [8] 简易,林岩,张书涵.基于 POI 数据的徐州市旅游景点空间分布特征及影响因素研究[J].智能建筑与智慧城市,2024,(06):36-39.
- [9] 李涛,徐江.基于 GIS 辅助的多源地理空间矢量数据挖掘方法[J].计算机仿真,2024,41(09):465-469.
- [10] 赵志忠,陈素根.基于相互 K 近邻的密度峰值聚类算法[J].安庆师范大学学报(自然科学版),2024, 30(02):41-46.

A k-nearest Neighbor Clustering Algorithm for Celebrity Cultural Tourism Attractions Based on Spatial Accessibility Model

Zhou Xiao¹, Pan Juan¹, Zhao Taiping¹, Tian Jiangpeng²

¹ School of Culture and Tourism, Leshan Vocational and Technical College, Leshan, Sichuan Province, China

² Institute of Geospatial Information, Information Engineering University, Zhengzhou, Henan Province, China

Abstract: Based on the analysis of the problem background, a spatial accessibility model between cultural tourism attractions and transportation hubs is constructed, upon which a k-nearest neighbor spatial clustering method for cultural tourism attractions is established. Taking the Neijiang city's Zhang Daqian cultural tourism attractions in Sichuan Province as the research subjects, the geographic spatial data of the city's main transportation hubs and the Zhang Daqian cultural tourism attractions are firstly collected. A spatial accessibility model based on spatial coordinates is established, and the spatial accessibility between the Zhang Daqian cultural tourism attractions and transportation hubs is then calculated. On the basis of searching for the optimal spatial accessibility between the Zhang Daqian cultural tourism attractions and transportation hubs, a spatial clustering model is established to determine the spatial relationships between each cultural tourism attraction and transportation hub. This provides a scientific quantitative basis for cultural tourism attraction recommendation services and tourism decision-making services in Neijiang, while also serving as a reference for other tourism cities in Sichuan province.

Keywords: spatial accessibility model; cultural and tourism attractions; k-nearest neighbor clustering; tourism planning