引文格式：张红,李玥,邓雯,等. 利用网络游记分析成都市旅游景区空间结构[J]. 地理空间信息, 2024, 22(7): 31-35. doi:10. 3969/j.issn. 1672-4623.2024.07.007

利用网络游记分析成都市旅游景区空间结构

张　红[1]，李　玥[2*]，邓　雯[2]，王　艺[2]
（1. 华东师范大学 地理科学学院，上海 200241；2. 西南交通大学 地球科学与环境工程学院，四川 成都 611756）

摘　要：从携程网爬取了2010—2019年成都市游记数据，构建旅游景区关键词文本共现网络，采用数量统计、空间分析、复杂网络分析等方法，挖掘成都市旅游景区游客到访的时空分异特征。结果表明，成都市旅游景区到访频率呈长尾分布，空间上形成以青城山—都江堰、金牛—武侯主城区为高热度中心的“双核摄动”格局；成都市旅游景区文本共现网络具有较高的集聚系数和较短的平均路径，等级圈层结构和马太效应较显著，热门景区对邻近热门景区的空间溢出效应显著，但对邻近低等级景区的带动效果有限；成都市旅游景区空间结构由条带状和团簇式向轴辐式和网络化方向转变，在区域旅游一体化发展中的作用日益显著。

关键词：旅游景区；网络游记；空间结构；文本共现；复杂网络分析
中图分类号：P208　　**文献标志码：**B　　**文章编号：**1672-4623（2024）07-0031-05

Spatial Structure Analysis of Tourist Attractions in Chengdu Using Online Travel Notes

ZHANG Hong[1], LI Yue[2], DENG Wen[2], WANG Yi[2]
(1. School of Geographic Sciences, East China Normal University, Shanghai 200241, China;
2. Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu 611756, China)

Abstract: In this article, we extracted the Chengdu travel notes from 2010 to 2019 from Ctrip, and constructed a co-occurrence network of tourist attractions by text analysis. Then, we employed the approaches of statistical analysis, spatial analysis and complex network analysis to explore the spatio-temporal characteristics of tourists visiting behaviors to Chengdu. The results show that the visiting frequencies of tourist attractions in Chengdu follow a long-tailed distribution. Spatially, a dual-core pattern is formed with centers of Qingcheng Mountain-Dujiangyan and Jinniu-Wuhou. The text co-occurrence network of tourist attractions in Chengdu shows high values of agglomeration coefficient and low average shortest path length. Besides, the hierarchical pattern and the Matthew effect of tourist attractions in Chengdu are obvious, the spatial spillover effect of popular tourist attractions to the other neighboring tourist attractions is obvious while for those not well-known tourist attractions is very limited. In general, the spatial structure of tourist attractions in Chengdu is shifted from strip-like and clustering to hub-and-spoke and network, indicating that Chengdu plays increasing role in the regional tourism integration development.

Key words: tourist attraction, online travel note, spatial structure, text co-occurrence, complex network analysis

旅游景区是旅游活动的核心和空间载体，其空间结构能反映区域旅游一体化发展水平。传统旅游景区空间结构研究多基于统计年鉴或问卷调查数据，其时空分辨率较低且更新较慢，难以准确、完整地反映游客的时空分布和流动性特征[1]。游客的GPS或移动基站定位数据虽能精确刻画游客活动的时空格局[2]，但数据获取成本较高。近年来兴起的以游记、点评、问答与攻略等为代表的旅游数字足迹数据，因其获取相对便利、现势性强且信息丰富而成为新的旅游数据源，在旅游行为分析、旅游流分析和旅游情感认知研究等方面发挥重要作用，为理解旅游景区空间结构提供了新视角。

旅游数字足迹是旅游者在旅游期间和旅游结束后的记录性信息[3]，如评论、旅游照片、网络游记等[4-7]，反映了旅游者的活动与感知，也为旅游景区的空间格局[8-10]、旅游流[11-16]、旅游目的地形象感知[17-20]等提供了新视角。研究发现，基于语义或地点信息构建的旅游数字足迹网具有集聚性、非均衡性、小世界与等级特征；但多集中于国家、省份或城市群尺度，对单个城市的研究不多。本文以“中国最具幸福感的

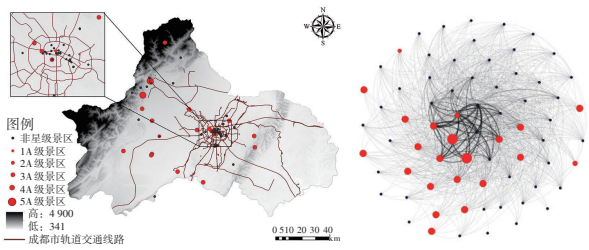
收稿日期：2022-12-28。
项目来源：国家自然科学基金面上项目（42171420）；四川省科技支撑计划资助项目（2020YJ0325）；中央高校基本科研业务费资助项目（2021ECNU-YYJ015）；华东师范大学教学改革与研究资助项目（2022HSJG006）。
第一作者简介：张红（1981—），博士，副教授，研究方向为空间复杂性与城市演化，E-mail：hzhang@re.ecnu.edu.cn。
通信作者：李玥（1998—），硕士研究生，研究方向为空间分析、旅游地理学，E-mail：985161530@qq.com。

城市”——四川省成都市为例，从携程网收集了2010—2019年成都市游记数据，并基于景点文本共现构建了旅游数字足迹网并分析旅游景区的空间结构特征，揭示其空间效应，以期能服务于成都市旅游规划与开发。

1 数据来源与研究方法

1.1 研究区概况与数据来源

成都市位于四川盆地西部，地处龙门山脉、邛崃山脉、成都平原和川中丘陵的交界地带；旅游资源丰富，被誉为“中国文化名城”“美食之都”“中国最具幸福感城市”[21]。携程网是我国最大的在线旅游服务网站。本文从携程网收集了成都市2010—2019年游记数据，共20 190条，每条游记包含用户名、发布时间、标题、游玩天数、出行时间、花费、玩法、游记文本等信息；删除目的地信息缺失、广告和重复游记等，最终得9 900条游记和79个旅游景区（图1a）；再利用Python jieba工具分词，构建旅游景区文本共现网络，包括71个节点和1 260条连线（图1b）。



a 成都市旅游景区空间分布　　b 成都市旅游景区共现网络

图1 成都市旅游景区空间分布与共现网络

1.2 研究方法

本文从统计非均衡性、空间异质性、网络关联复杂性3个方面刻画旅游景区的空间结构特征。

1）统计非均衡性的测度。赫芬达尔指数可反映规模指标的集聚程度，计算公式为：

$$H=\sum_{i=1}^{n}P_i^2 \tag{1}$$

式中，P_i为某地区或某时间段数值占比；H的取值在0~1之间，越接近于1，表示数值分布越集中，反之越分散[22]。

2）空间异质性的刻画。反距离权重插值是将离散点数据转换为连续曲面，从而可视化要素空间分布的一种方法[23]。以插值点与样本点间距离（或距离平方）的倒数为权重估算插值点数值，离插值点越近的样本点权重越高。本文基于距离的平方进行插值。

$$Z_{S_0}=\sum_{i=1}^{n}\lambda_i Z_{S_i} \tag{2}$$

式中，Z_{S_0}为插值点S_0的预测值；n为样本点数量；λ_i为样本点权重，随样本点与插值点距离的增大而减小；Z_{S_i}为样本点数值。

3）网络关联复杂性的描述。常见的节点中心性指标包括度中心性、中介中心性和邻近中心性。度中心性是指与它直接相连的节点总数[24]；中介中心性用于测度节点在其他网络节点连通中中介作用的大小；邻近中心性刻画了某节点与网络中其他所有节点联系的方便程度。

$$K_i=\sum_{j=1,i\neq j}^{n}L_{ij} \tag{3}$$

$$B_i=\sum_{s,t\in V,s,t\neq i}\frac{o(s,t|i)}{o(s,t)} \tag{4}$$

$$C_i=n-1\Big/\sum_{j=1,i\neq i}^{n}d_{ij} \tag{5}$$

式中，K_i为度中心性；n为网络节点总数；L_{ij}为节点i与j的连边数；B_i为中介中心性；$o(s,t)$为节点s与t之间的最短路径数量；$o(s,t|i)$为节点s与t的最短路径中经过节点i的数量；C_i为邻近中心性；d_{ij}为节点i到j的最短路径长度。

中心性仅考虑了单个节点在网络中的角色，而集聚系数则可描述节点的邻域完整性[25]，计算公式为：

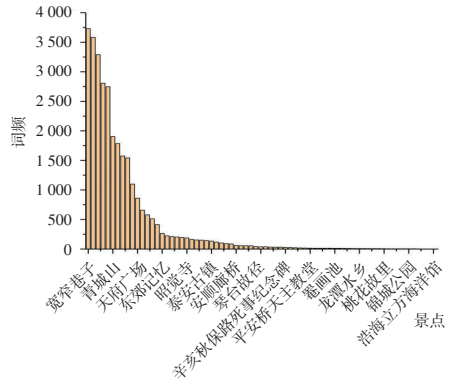
$$CC_i=2E_i/k_i(k_i-1) \tag{6}$$

式中，E_i为节点i邻域的实际连边数量；k_i为节点i的度值；CC的取值范围为0~1，当所有节点完全连通时，CC为1。网络的集聚系数为网络中所有节点集聚系数的平均值。

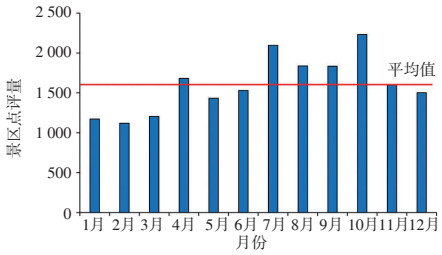
2 旅游景区点评量的统计与空间特征

2.1 景区点评量的统计分布

对旅游景区进行词频统计发现，其点评量呈长尾分布（图2a），这与参考文献[26]、[27]的结论一致。到访量排名前五的景区分别为宽窄巷子、锦里、春熙路、武侯祠和熊猫基地，都是成都市文化的代表，印证了城市文化是游客目的地认知的基础与核心。此外，历史人文积淀深厚和休闲娱乐设施齐全的景区也倍受青睐，如宽窄巷子、青城山、天府广场等，且这些景区的空间可达性也较高。成都市旅游景区各月点评量呈波动增长（图2b），春季点评数量较少，4月后点评热度明显回升，7—10月增至峰值，11—12月略有回落；赫芬达尔指数为0.087，表明各月份间点评量整体差异较小，这与成都市气候四季皆宜、人文景观丰富有关。



a 旅游景区点评热度柱状图

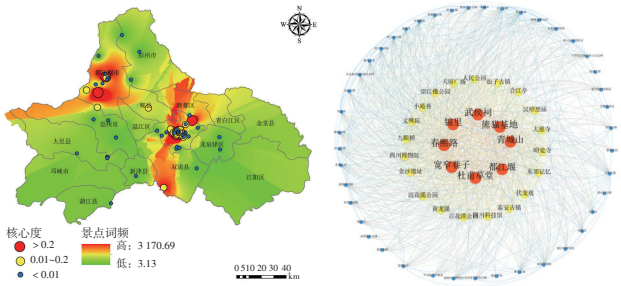


b 点评热度季节分布图

图2　成都市旅游景区点评热度与月份分布统计图

2.2　景区点评量的空间分布

成都市旅游景区点评热度呈现紧密团块型和条带型两种分布模式（图3a），高点评热度区集聚在西北部的青城山—都江堰名山群以及市中心的宽窄巷子—天府广场景区群，其中青城山—都江堰景区以道教文化和水利工程名扬天下，邻近的赵公山、灵岩山和熊猫谷的点评热度也较高，显示出较强的扩散效应。



a 成都市旅游景区结构等级空间分布

b 成都市旅游景区共现网络核心边缘结构

图3　成都市旅游景区点评量核密度与核心边缘结构图

审图号：GS(2016)1760

3　成都市旅游景区文本共现网络的结构特征

3.1　文本共现频次

成都市点评量排名前七的旅游景区文本共现频次见表1，共现频次较高的景点对为宽窄巷子—锦里、锦里—武侯祠、宽窄巷子—春熙路，这3对景区彼此邻近，除武侯祠外，其余3个景区均免费，再次表明距离或可达性是旅游线路设计的主要因素之一。

表1　成都市点评量排名前七的旅游景区文本共现频次统计

	宽窄巷子	锦里	都江堰	武侯祠	春熙路	青城山	人民公园
宽窄巷子	0	2 702	1 361	1 962	1 986	962	933
锦里	2 702	0	1 380	2 298	1 906	983	864
都江堰	1 361	1 380	0	1 151	971	1 617	443
武侯祠	1 962	2 298	1 151	0	1 479	836	704
春熙路	1 986	1 906	971	1 479	0	726	747
青城山	962	983	1 617	836	726	0	341
人民公园	933	864	443	704	747	341	0

3.2　网络特征

成都市旅游景区共现网络包含71个节点和1 260条连线，网络直径和特征路径长度分别为3和1.498，密度和集聚系数分别为0.507和0.842，表明成都市旅游景区共现网络密度较高且结构较紧凑。与北京[15]、上海[27]、杭州[28]等城市类似，成都市旅游景区网络呈现典型的核心—外围圈层结构（图3b），核心圈层包含锦里、宽窄巷子、武侯祠、春熙路、都江堰、杜甫草堂、熊猫基地、青城山8个景区，集中分布于都江堰—青城山片区和主城区；中间圈层包含酒吧文化“九眼桥”、人文揽胜金沙遗址、文殊院、大慈寺等20个景区；边缘圈层多为农家乐、公园或古镇型景区，散见于城市各处，受可达性与资源季节性影响，各景区彼此间以及与其他景区的共线路关系较稀疏。因此，景区共访特征受资源特色、知名度、可达性和季节性的影响，高等级景区对边缘景区的辐射带动作用较弱。

3.3　景区中心性

度中心性、中介中心性和邻近中心性排名前十的旅游景区见表2，都江堰、宽窄巷子和春熙路的各种中心性数值均较高，证明了旅游景区知名度与可达性的叠加效应。各旅游景区的平均度值为35，说明游客旅游线路多样。景区中介中心性差异显著，最高为都江堰和杜甫草堂，即多数旅游者倾向于以这两个景区作为转换景区，主要是由于其便捷的可达性和旅游套票促销措施。建川博物馆、蔚然花海等10个旅游景区的中介中心性为0，这些景区离市区较远或旅游特色不够鲜明，可替代性强。景区邻近中心性数值普遍较高，表明各景区线路组合灵活性强。邻近中心性数值较高的景区为都江堰、春熙路、宽窄巷子、青城山等，最低的为清水河公园，再次证实了知名景区的空间溢出效应。

3.4　结构洞分析

节点的结构洞水平能刻画其在网络中的竞争力。节点的有效规模越大、效率越高且限制度越小，其结构洞优势越显著。成都市点评热度排名前二十的景区

表2 中心性排名前十的景区

排名	度中心性	中介中心性	邻近中心性
1	都江堰	都江堰	都江堰
2	宽窄巷子	杜甫草堂	春熙路
3	春熙路	黄龙溪	宽窄巷子
4	武侯祠	青城山	青城山
5	青城山	春熙路	熊猫基地
6	熊猫基地	宽窄巷子	武侯祠
7	锦里	熊猫基地	杜甫草堂
8	杜甫草堂	武侯祠	锦里
9	人民公园	锦里	人民公园
10	黄龙溪	人民公园	黄龙溪

的结构洞指标见表3，宽窄巷子、锦里、都江堰、武侯祠、春熙路等点评热度较高的景区具备很强的旅游竞争力，游客更倾向于选择打卡著名旅游景区，同时去往邻近高等级或特色鲜明的景区；换而言之，高等级景区更容易与邻近高等级景区共享旅游流，而其对邻近低等级景区的空间溢出效应有限。

表3 成都市点评热度排名前二十的旅游景区结构洞指标统计

排名	景区名称	有效规模	效率	限制度
1	宽窄巷子	34.622	0.533	0.256
2	锦里	31.941	0.507	0.262
3	都江堰	35.926	0.544	0.266
4	武侯祠	32.435	0.507	0.271
5	春熙路	33.750	0.519	0.267
6	青城山	33.351	0.521	0.275
7	人民公园	31.789	0.513	0.266
8	杜甫草堂	32.368	0.514	0.275
9	熊猫基地	32.569	0.509	0.277
10	文殊院	27.388	0.489	0.267
11	天府广场	29.276	0.514	0.267
12	黄龙溪	32.792	0.547	0.253
13	九眼桥	29.755	0.513	0.262
14	大慈寺	27.110	0.512	0.260
15	金沙遗址	29.685	0.521	0.258
16	东郊记忆	23.323	0.496	0.264
17	四川博物院	23.642	0.525	0.253
18	汉昭烈庙	20.940	0.465	0.280
19	浣花溪公园	25.280	0.527	0.252
20	伏龙观	22.019	0.512	0.275

4 结 语

游记文本数据记载了游客出行的时间、到访地点、特色与评价等信息，为理解旅游行为、目的地营销、分析旅游资源的空间结构等提供了新型数据源。传统旅游空间结构研究局限于国家、省份或城市群尺度，对单个城市尤其是平原城市的研究不多，也缺少对景区点评热度、互访关系和空间效应的分析。本文从携程网收集了2010—2019年成都市游记文本数据，综合采用统计、空间与网络分析方法，揭示了成都市旅游景区到访率的时空分异性、旅游线路组合的多样性和邻近性特征。结果表明：①成都市各旅游景区点评量呈长尾分布特征，极少数知名度高的景区成为成都市旅游的王牌产品，包括宽窄巷子、锦里、春熙路、武侯祠和熊猫基地等，得益于丰富的人文景观，成都市旅游景区季节性淡旺季不明显；②成都市旅游景区点评热度空间分异显著，呈现紧密团块型和条带型相结合的分布模式，集中于青城山—都江堰名山群和宽窄巷子—天府广场景区群，表明景区等级与邻近性的叠加效应明显；③知名度高的景区对邻近高等级或特色鲜明景区的空间溢出效应明显，对邻近低等级景区的带动效果有限。

未来研究将考虑新浪微博、抖音、快手等多种网络数据源，并结合问卷调查、实地访谈数据，从全生命周期角度挖掘旅游者旅游行为特征、评价旅游景区的开发绩效等。此外，旅游点评数据语义十分丰富，对于游记主题的挖掘、情感认知、线路设计等问题的研究仍待深入。

参考文献

[1] Mou N, Zhang Y, Makkonen T, et al. Tourists' Digital Footprint: the Spatial Patterns of Tourist Flows in Qingdao, China[J]. Tourism Management, 2020, 81(105):104-414

[2] 郎月华，李仁杰，傅学庆. 基于GPS轨迹栅格化的旅游行为空间模式分析[J]. 旅游学刊，2019，34(6):48-57

[3] Girardin F, Calabrese F, Fiore F, et al. Digital Footprinting: Uncovering Tourists with User-generated Content[J]. IEEE Pervasive Computing, 2008, 7(4):36-43

[4] 蔚海燕，戴泽钒，许鑫，等. 上海迪士尼对上海旅游流网络的影响研究:基于驴妈妈游客数字足迹的视角[J]. 旅游学刊，2018，33(4):33-45

[5] 秦静，李郎平，唐鸣镝，等. 基于地理标记照片的北京市入境旅游流空间特征[J]. 地理学报，2018，73(8):1 556-1 570

[6] Peng H, Zhang J, Liu Z. et al. Network Analysis of Tourist Flows: A Cross-provincial Boundary Perspective[J]. Tourism Geographies, 2016, 18(5):561-586

[7] Park S B, Kim J, Lee Y K, et al. Visualizing Theme Park Visitors' Emotions Using Social Media Analytics and Geospatial Analytics[J]. Tourism Management, 2020, 20(5):20-23

[8] 陈志云，刘敏纯，钟广锐，等. 基于多源数据的梅州市旅游资源综合评价分析[J]. 地理空间信息，2021，19(4):51-55

[9] 高彩霞，刘家明，高岩，等. 京津冀非物质文化遗产资源的空

间格局及旅游开发研究[J]. 地理与地理信息科学,2021,37(3):103-108
[10] 刘畅,孙彩歌,樊风雷. 基于多源数据的拉萨旅游空间结构分析[J]. 地理空间信息,2020,18(8):25-30
[11] Gao Y, Ye C, Zhong X, et al. Extracting Spatial Patterns of Intercity Tourist Movements from Online Travel Blogs[J]. Sustainability,2019,11(13):1-18
[12] Jin C, Cheng J, Xu J. Using User-generated Content to Explore the Temporal Heterogeneity in Tourist Mobility[J]. Journal of Travel Research,2018,57(6):779-791
[13] 李磊,陶卓民,陆林,等. 贵州省避暑旅游流网络结构特征及其影响因素[J]. 地理研究,2021,40(11):3 208-3 224
[14] 雷家琳. 基于网络游记的胶东半岛旅游流空间网络结构研究[D]. 青岛:青岛大学,2021
[15] 周李,吴殿廷,虞虎,等. 基于网络游记的城市旅游流网络结构演化研究:以北京市为例[J]. 地理科学,2020,40(2):298-307
[16] 姚云霞,管卫华,李在军. 江苏省入境旅游流的时空演变及影响因素分析[J]. 旅游科学,2016,30(5):52-62
[17] 王素洁,黄楷伊,董玉洁. 中国视觉旅游形象研究:基于社会网络视角[J]. 旅游科学,2018,32(2):66-79
[18] 董引引,曲颖. 滨海目的地“投射形象网络机制”对比研究差异化竞争定位的契机:以三亚与厦门为例[J]. 旅游科学,2021,35(5):43-61
[19] 黄雪. 基于网络游记的兰州市旅游空间与形象认知分析[D]. 兰州:兰州大学,2019
[20] Song Y T, Li C Y, Chen C C. Online Social Construction of Taiwan's Rural Image: Comparison Between Taiwanese Self-representation and Chinese Perception[J]. Tourism Management,2020,76(20):103 968
[21] 四川省旅游发展委员会. 四川旅游年鉴2016[M]. 成都:四川科学技术出版社,2017
[22] 朱付彪,陆林,於冉,等. 都市圈旅游空间结构演变研究:以长三角都市圈为例[J]. 地理科学,2012,32(5):570-576
[23] 李正泉,吴尧祥. 顾及方向遮蔽性的反距离权重插值法[J]. 测绘学报,2015,44(1):91-98
[24] 钟柯. 中国列车运行网络节点中心性度量及演化分析[D]. 北京:北京大学,2012
[25] 周宏浩,谷国锋. 外部性视角下中国城市网络演化及其环境效应研究[J]. 地理研究,2022,41(1):268-285
[26] 杨敏,李君轶,杨利. 基于旅游数字足迹的城市入境游客时空行为研究:以成都市为例[J]. 旅游科学,2015,29(3):59-68
[27] 李倩,曲凌雁. 城市旅游流网络结构特征及其影响因素:以上海市为例[J]. 世界地理研究,2021,30(1):114-124
[28] 李晶,葛玉辉. 基于数字足迹的城市旅游流网络结构演化研究:以杭州市为例[J]. 资源开发与市场,2021,37(5):620-628

（上接第30页）

入填补了POI缺少规模信息的漏洞，能够更加精准地识别城市各功能区的分布，提高规划编制的科学性。然而，POI面积研究存在局限性：目前仅能获得POI面积信息，仅使用网格POI数量与POI簇中POI数量的对比判别网格是否应赋予POI簇的功能具有一定偶然性，无法精确识别POI所在地块的边界。POI面积信息的进一步利用、基于面积信息的POI簇边界识别、除面积之外其他信息的获取等可能成为今后研究关注的重点。

参考文献

[1] 张婷,王红原. 基于POI数据的南宁市城市功能区识别及可视化研究[J]. 中外建筑,2021(3):102-107
[2] 窦旺胜,王成新,薛明月,等. 基于POI数据的城市用地功能识别与评价研究:以济南市内五区为例[J]. 世界地理研究,2020,29(4):804-813
[3] 段松江,丁忆. 基于影像与GIS数据的现状建设用地精细化提取[J]. 地理空间信息,2017,15(11):57-59
[4] 柴宝惠,李培军,张瑞洁,等. 基于Landsat数据和DMSP/OLS夜间灯光数据的城市扩展提取:以天津市为例[J]. 北京大学学报(自然科学版),2016,52(3):475-485
[5] 冉钊,周国华,吴佳敏,等. 基于POI数据的长沙市生活性服务业空间格局研究[J]. 世界地理研究,2019,28(3):163-172
[6] 李莹,涂志德,刘艳芳,等. 基于时间序列相似度的城市功能区识别研究[J]. 地理空间信息,2021,19(1):22-29
[7] 曹根榕,顾朝林,张乔扬. 基于POI数据的中心城区“三生空间”识别及格局分析:以上海市中心城区为例[J]. 城市规划学刊,2019(2):44-53
[8] 康雨豪,王玥瑶,夏竹君,等. 利用POI数据的武汉城市功能区划分与识别[J]. 测绘地理信息,2018,43(1):81-85
[9] 胡晓鸣,黎小龙,蔚芳. 基于POI的城市功能区及其混合度识别研究:以重庆市核心城区为例[J]. 西南大学学报(自然科学版),2021,43(1):164-173
[10] 张玲. POI的分类标准研究[J]. 测绘通报,2012(10):82-84
[11] 浩飞龙,王士君,冯章献,等. 基于POI数据的长春市商业空间格局及行业分布[J]. 地理研究,2018,37(2):366-378