

基于大数据的城市活力区和中心城区边界识别 ——以广州市为例

宋程, 陈嘉超, 李彩霞, 艾冠韬
(广州市交通规划研究院, 广东广州 510030)

摘要: 城市活力区和中心城区是认识和研究城市的重要基础性信息, 也是城市管控政策制定的前提。传统调查数据难以准确界定边界, 城市活力区和中心城区边界的确定多以定性为主。大数据的兴起为城市功能区边界识别创造了条件。基于手机信令数据、互联网位置数据、兴趣点等多源数据, 建立人口综合密度、兴趣点密度两大静态指标, 以及与城市中心可达性、商务休闲活动驻点密度两大动态指标, 以综合识别城市活力区和中心城区边界。动态指标和静态指标的组合使用克服单一指标或仅静态指标识别的局限性。以广州市为例阐述了该方法的实用性, 识别出广州市活力区和中心城区的范围已与国际一流城市相当。

关键词: 手机信令数据; 兴趣点; 活力区; 中心城区; 边界识别; 核密度; 广州市

Urban Vitality Zone and Central District Identification Based on Big Data: A Case Study in Guangzhou City

Song Cheng, Chen Jiachao, Li Caixia, Ai Guantao

(Guangzhou Transport Planning Research Institute, Guangzhou Guangdong 510030, China)

Abstract: The urban vitality zone and the central district are important basic information to understand and research the city, and also the premise of urban management and control policies making. It is difficult to define the boundary of urban vitality zone and central district accurately using traditional survey data, which is usually the qualitative analysis. The rise of big data provides potential for boundary identification of urban functional area. Based on cellular signaling, Internet location data, POI and other multi-source data, this paper develops two static indicators of population comprehensive density and POI density, as well as two dynamic indicators of urban center accessibility and business/leisure activity point density so as to comprehensively identify urban vitality zone and central district boundaries. The combination of dynamic and static indicators overcomes the limitations of boundary identification with single index or only static index. Taking Guangzhou as an example, the paper demonstrates the practicability of the method, and concludes that the scope of Guangzhou's vitality zone and central district is equivalent to that of world-class cities.

Keywords: cellular signaling data; POI; vitality zone; central district; boundary identification; kernel density; Guangzhou

收稿日期: 2019-03-26

基金项目: 广州市“岭南英杰工程”后备人才(马小毅)培养计划科研课题项目“新时代大城市交通与土地利用(生产、生活、生态空间)协同模型重构与仿真系统研究”

作者简介: 宋程(1983—), 男, 湖北秭归人, 硕士, 主要研究方向: 交通建模与仿真。

E-mail: 510659684@qq.com

1 研究综述

城市空间结构是城市要素在空间范围内的分布和组合状态, 是城市经济结构、社会结构的空

研究分为三个阶段: 西方理论引入及中国实证研究起步期, 中国实证研究积累期, 以及目前正步入的对中国城市空间结构模式的总结及新城市空间现象研究的多元化时期。文献[2]对都市空间界定的四要素(空间要素、时间要素、流量要素和引力要素)进行了具

体量化分析,用经济势能及交通权数来确定都市圈的空间范围,通过对四个要素的综合,确定了都市圈圈域半径和圈层结构的界定方法,并以成都都市圈为例做了实证分析,得出成都都市圈的圈域半径和圈层结构。文献[3]对兰州都市圈空间界定时,将经济距离、各周边城市的生活质量引入场强和引力模型的计算中,使界定方法更具借鉴性。文献[4]以百度兴趣点(Point of Interest, POI)数据为基础数据,提出一种基于大数据技术的城市空间结构的识别方法。文献[5]以重庆市主城区40余万条POI数据为基础,利用核密度分析、自然断点法和邻近分析等方法,根据整体及不同类型POI数据的空间分布特征与聚集程度,识别城市总体及不同职能的多中心结构及其影响范围。文献[6]以辽宁省盘锦市为例,使用POI数据,结合核密度分析法分析城市中各类设施分布的空间规律,并识别城市的空间布局。文献[7]以上海市主城区为例,借助百度地图的海量时空数据,着眼于服务可及性和要素建成情况,引入可达性格局分析、等时圈分析和基于POI分类的统计分析,从服务范围、服务时效和服务能力3个维度定量考察多中心空间发展特征,探索多中心城市空间发展评估的新方法。

综上所述,关于城市空间识别研究采取的指标较为单一,部分采取多指标识别的研究也仅注重静态指标,缺乏动态指标的体现。单一指标缺陷是只能反映某一方面特征而忽视其他因素,而城市空间结构形成是多种因素的综合作用,单一指标并不能完全真实地反映城市空间形态;全部采用静态指标不能体现城市流动性对空间形态的影响。

2 城市发展边界识别考虑因素及指标选择

城市活力区是一座城市政治、经济、文化等公共活动最集中的地区,是城市功能高度聚集的区域,是城市公共活动体系的主要部分,为城市和城市所在区域集中提供经济、政治、文化等活动设施和服务空间,并在空间上有别于城市其他地区。中心城区是一个城市现状发展最为成熟的地区,以多职能综合性为主,汇聚了城市生产和服务的主要区域,是政治和行政管理中心、交通运输

中心、信息与科技中心以及人才密集之地。

城市活力区和中心城区是城市交通网络最发达、出行活动最密集的地区。边界识别不仅要考虑静态指标影响,还需考虑动态指标影响。为此,本文构建了包含2个静态指标和2个动态指标的综合识别体系:静态指标包括人口综合密度和POI密度,分别表征人口聚集和基础设施完善程度;动态指标为与城市中心可达性、商务休闲活动驻点密度,体现既有交通系统下区域交通联系的便捷程度和客流吸引强度。

1) 人口综合密度。

人口综合密度不同于居住人口密度,综合密度反映该区域单位面积常态化出现的人口规模,是居住人口和就业人员的综合密度,反映居住和就业的完善程度。一般而言,人口综合密度越高的地区是城市发展越成熟的地区。

$$pe_i = \frac{emp_i + pop_i * \alpha}{area_i}, \quad (1)$$

式中: pe_i 为区域 i 的人口综合密度/(人·km²); emp_i 为区域 i 的就业岗位数/个; pop_i 为区域 i 的居住人口数/人; $area_i$ 为区域 i 的面积/km²; α 为人口调整系数,用来反映活跃人口的影响,该值可用全市从业人员与常住人口比值标定。

2) POI密度。

POI包括医院、学校、餐饮、酒店、超市商场、主要办公楼、主要交通枢纽、公园、加油站、银行、休闲娱乐设施、政府机构等城市生活服务的基本场所,反映某一区域公共服务设施的分布密度,从一定程度上体现出区域的发展成熟度。

$$poi_i = \frac{\sum_{j=1}^n poi_{ij}}{area_i}, \quad (2)$$

式中: poi_i 为区域 i 的POI密度/(个·km²); poi_{ij} 为区域 i 的 j 类兴趣点的数量/个。

3) 与城市中心可达性。

城市中心可用市政府所在地(或标志性地点)作为代表,其可达性反映了某一地区到达城市中心的便捷程度。可达性越高代表该地区交通系统越发达,越能反映该地区的发展状况。

$$t_i = \min(t_{cari}, t_{busi}, t_{metroi}), \quad (3)$$

式中: t_i 为区域 i 到达城市中心的最小出行时间/min; t_{cari} 为区域 i 采用小汽车方式到达

城市中心的出行时间/min; t_{busi} 为区域 i 采用公共汽车方式到达城市中心的出行时间/min; t_{metro} 为区域 i 采用轨道交通方式到达城市中心的出行时间/min。

4) 商务休闲活动驻点密度。

商务休闲活动驻点密度不仅体现某一地区常态化活动人口的出行强度,也反映访问人员的出行强度,是该区域客流总体吸引强度,体现交通流在城市成熟区的聚集程度和城市成熟区最重要的经济、政治、文化活动的成熟度。某一地区人口综合密度非常大,如某工业区大量工人会产生较高的人口综合密度,但其商务休闲活动驻点密度较低,并不能代表该区域是城市最为成熟的活力区域。

$$d_i = \frac{\sum_{j=1}^n d_{ij}}{area_i}, \quad (4)$$

式中: d_i 为区域 i 的商务休闲活动驻点密度/(个·km²); d_{ij} 为区域 i 的 j 类驻点的数量/个。

3 城市发展边界识别方法研究

3.1 基于大数据的指标获取

21 世纪前 10 年,四大指标的准确获取难度较大。人口综合密度可通过统计数据获取,但大城市特别是超大、特大城市的常住人口和从业人员统计指标一方面不能最真实体现实际人口数,另一方面数据颗粒度较粗(最小颗粒度为居委会级),达不到城市发展边界识别的数据细颗粒度要求;POI 密度在大数据兴起之前更是无法获取;与城市中心可达性虽然可借助交通模型技术获取,但模型计算值精度难以保证;商务休闲活动驻点密度依靠传统方法无法获得。近年来,城市大数据的兴起为上述指标的获取提供了有效途径,数据的精细度能很好支撑城市发展边界的识别工作。各指标的获取方法如下。

1) 人口综合密度:须借助信息数据获取,如长周期的手机信令数据,通过一定算法可获取精确的城市职住空间分布,进而得到某一区域的人口综合密度,相关算法如图 1 所示。

2) POI 密度:主要通过获取百度兴趣点数据并小区域进行统计。

3) 与城市中心可达性:该指标更能体现某一区域与城市中心的联系紧密程度,数

据获取的精确性对城市发展空间识别至关重要。基于互联网位置数据是可达性获取的理想途径,如高德导航数据、腾讯导航数据等,利用互联网位置数据可获取细颗粒度、小区间的实际耗时矩阵,从而能准确识别出某一点与任意区域的出行时耗。基于互联网的耗时矩阵精度高,从时间维度上体现了出行者的实际耗时,并非交通模型计算的某一时刻道路网分配结果。一般而言,与城市中心的时耗越小,可达性越强,区域发展越成熟。

4) 商务休闲活动驻点密度:该指标必须结合手机信令数据获取,通过信令数据获得驻点分布,进而获取驻点密度。与传统驻点算法不同,研究采取了基于高频率信号触发、具有空间关联特征的手机信令数据和上网数据,通过对手机信令数据的职住对应关系(Live-Work)和时空间核聚类 Clus 分析,

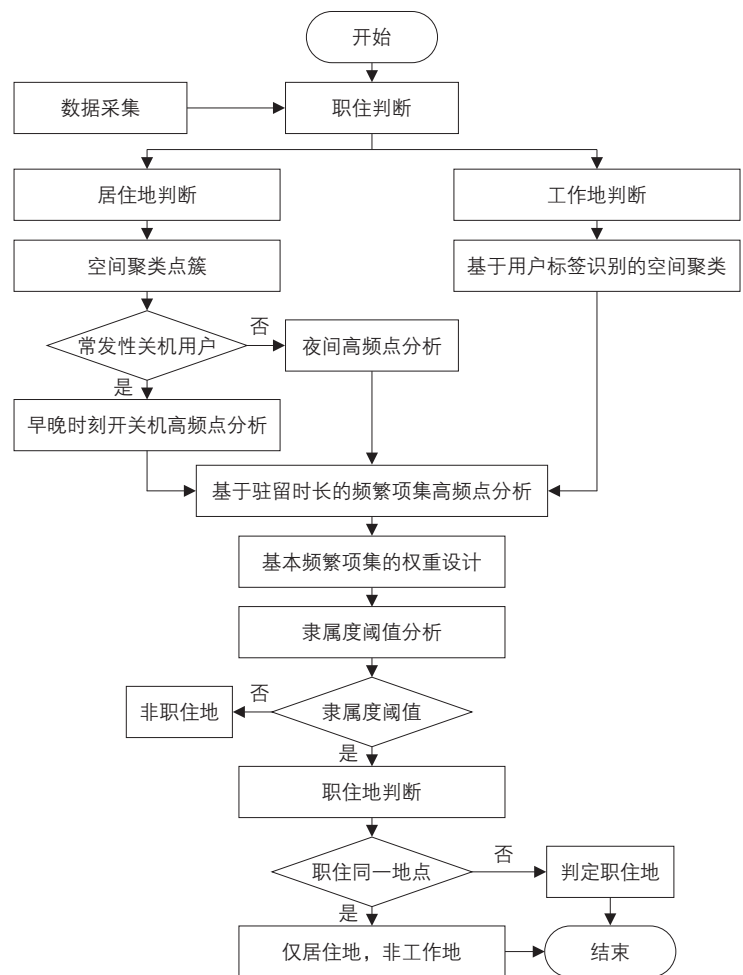


图 1 基于手机信令数据的职住地判别流程

Fig.1 Identification process of job-housing location based on cellular signaling data

提出一种满足用户出行特征的驻点判断方法，具体流程如图2所示。

3.2 综合指标的计算

4个指标从不同层面反映某一区域的开发成熟度，因此需采用一定权重对4个指标进行加权，以获取体现多种因素的综合指标，进而用于城市发展边界识别。采用专家打分法确定指标权重，通过调查20位城市规划领域相关专家，4个指标最终权重分别为0.22, 0.18, 0.33, 0.27。

确定指标权重后，为使各指标能在同一维度进行加权，还需进行指标的归一化处理。采取与阈值的比值对各指标进行归一化处理，阈值计算方法为：将某一城市划分为

若干个交通小区(颗粒度越细越好，一般超大、特大城市应达到3 000个以上，大城市应达到1 500个左右)，计算某一指标值，然后从高到低排序绘制分布曲线，采取80%分位值作为该指标的阈值，则某一区域的该指标归一化值为指标计算值与阈值的比值。据此计算各小区的综合指标

$$\bar{A}_{ij} = A_{ij} / \theta_j, \quad (5)$$

$$\lambda_i = \sum_{j=1}^n \bar{A}_{ij} * \beta_j, \quad (6)$$

式中： \bar{A}_{ij} 为交通小区*i*的指标*j*的归一化值； A_{ij} 为交通小区*i*的指标*j*的计算值； θ_j 为指标*j*的阈值； λ_i 为交通小区*i*的综合指标； β_j 为指标*j*的权重系数。

在归一化值的计算过程中，人口综合密

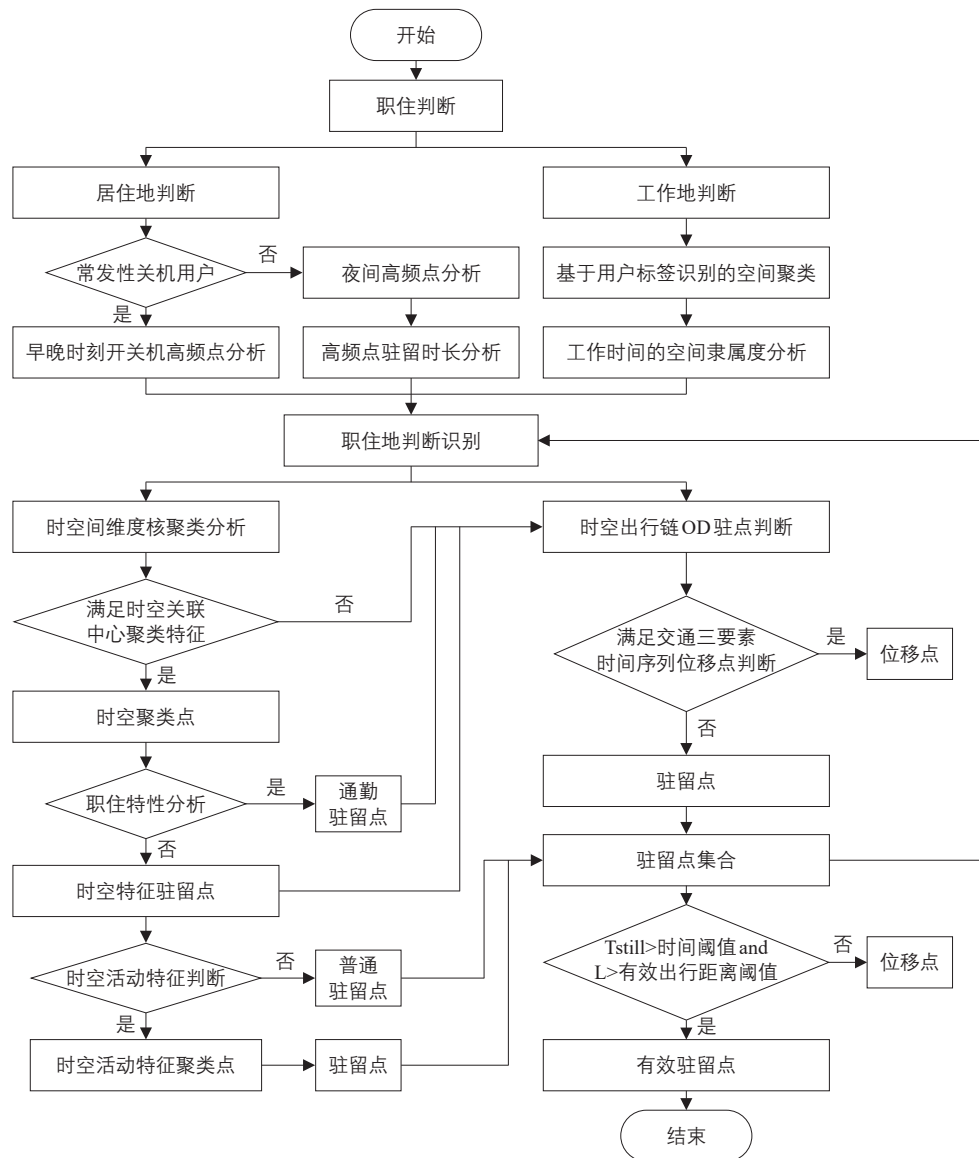


图2 基于手机信令数据的驻点判别流程

Fig.2 Identification process of stop-points based on cellular signaling data

表1 各指标具体获取方法

Tab.1 Specific acquisition methods of each index

指标	计算数据源	计算方法	参数
人口综合密度	联通手机信令数据 腾讯位置数据	基于长周期的数据识别职住点分布，聚类到3 989个交通小区，按式(1)计算密度值，按式(5)归一化处理	$\alpha = \text{广州市从业人员/常住人口} = 862/1\ 450 = 0.59$
POI密度	百度POI数据	按照获取的兴趣点数量采用式(2)计算	
与城市中心可达性	腾讯位置数据	基于腾讯位置数据获取3 989个交通小区的小汽车、公共汽车、轨道交通实际耗时矩阵，按式(3)计算	
商务休闲活动驻点密度	联通手机信令数据 腾讯位置数据	采用联通手机信令数据获取商务休闲活动驻点，由腾讯数据确定联通用户份额，扩样后按式(4)计算	市场份额为18%

度、POI密度、商务休闲活动驻点密度均为正影响值；与城市中心可达性为负影响值，应采取倒数进行归一化处理。

3.3 基于核密度法的城市发展边界识别

近年来，核密度分析法在城市热点探索方面应用广泛^[8-9]。该方法可用于计算空间点、线要素在周围领域中的密度，并对密度分布进行连续化的模拟，以图像中每个栅格的核密度值反映空间要素的分布特征^[5]。可根据每个交通小区内综合指标的核密度值估计其周围密度，并通过对不同搜索半径下的核密度分析结果进行比较，从而选取适合本研究的最优搜索半径。核密度函数计算公式为

$$f(x) = \sum_{i=1}^n \frac{1}{\pi r^2} \varphi\left(\frac{d_{ix}}{r}\right), \quad (7)$$

式中： $f(x)$ 为单位面积的综合指标核密度值； r 为搜索半径/km； n 为样本总数/个； d_{ix} 为交通小区*i*与*x*的距离/km； φ 为距离的权重。本研究基于综合指标，利用核密度法进行城市发展边界识别。

4 广州市现状发展空间识别

4.1 四大指标的获取及阈值选择

4.1.1 数据源

在广州市活力区和中心城区边界识别中，主要利用联通手机信令数据、腾讯位置数据和百度POI数据，各数据主要用途如表1所示，POI数量如表2所示。

4.1.2 各交通小区指标计算

基于各类信息数据，获取广州市3 989个交通小区的四大指标，各指标的空间分布如图3所示。

表2 广州市POI类型分布

Tab.2 Distribution of POI data in Guangzhou

分类	数量/个	比例/%
购物商场等	51 615	29.2
学校	4 742	2.7
医疗保健	4 164	2.4
休闲娱乐	3 726	2.1
酒店宾馆	3 389	1.9
银行金融	2 649	1.5
运动健身	1 317	0.7
文化场馆	543	0.3
餐饮	21 336	12.1
生活服务	17 811	10.1
公司企业	27 456	15.6
机构团体	9 193	5.2
基础设施	12 695	7.2
旅游服务	248	0.1
其他	15 581	8.8

4.1.3 指标阈值的确定及归一化处理

将四大指标的各小区值按从小到大排序后取80%分位值为该指标阈值(见图4)。其中人口综合密度阈值为2.5万人·km²，POI密度阈值为350个·km²，与城市中心可达性阈值为62 min，商务休闲活动驻点密度阈值为4.5万个·km²。对各指标进行归一化处理并按权重系数加权后得到各交通小区的综合指标如图5所示。

4.2 广州市活力区和中心城区边界识别

进行核密度算法识别，得到综合指标核密度值结果如图6所示。根据计算结果并适当考虑城市现状路网分割，划定广州市活力

区的范围为93 km²，中心城区边界约600 km²，据此指导城市交通政策制定、先行试验及交通引导，并对规划城市空间提供基础支撑。

4.3 与同类城市比较

类比纽约、伦敦、东京、巴黎4个国际一流城市，本研究所识别的广州市活力区介于东京都心6区和小巴黎范围之间；中心城区面积与北京、东京、巴黎基本相当(见表3)。从识别的城市活力区和中心城区边界来看，城市发展范围已达到国际一流城市的

水平。

5 结语

本文从多指标综合分析角度量化了城市发展边界识别方法，将现有的单一指标识别方法进行拓展，在注重人口聚集程度及设施完善程度静态指标的前提下，引入交通可达性和客流吸引强度等指标体现城市流动性对空间形态的影响，使得城市发展边界识别更符合实际，为城市管控政策实施提供更精确

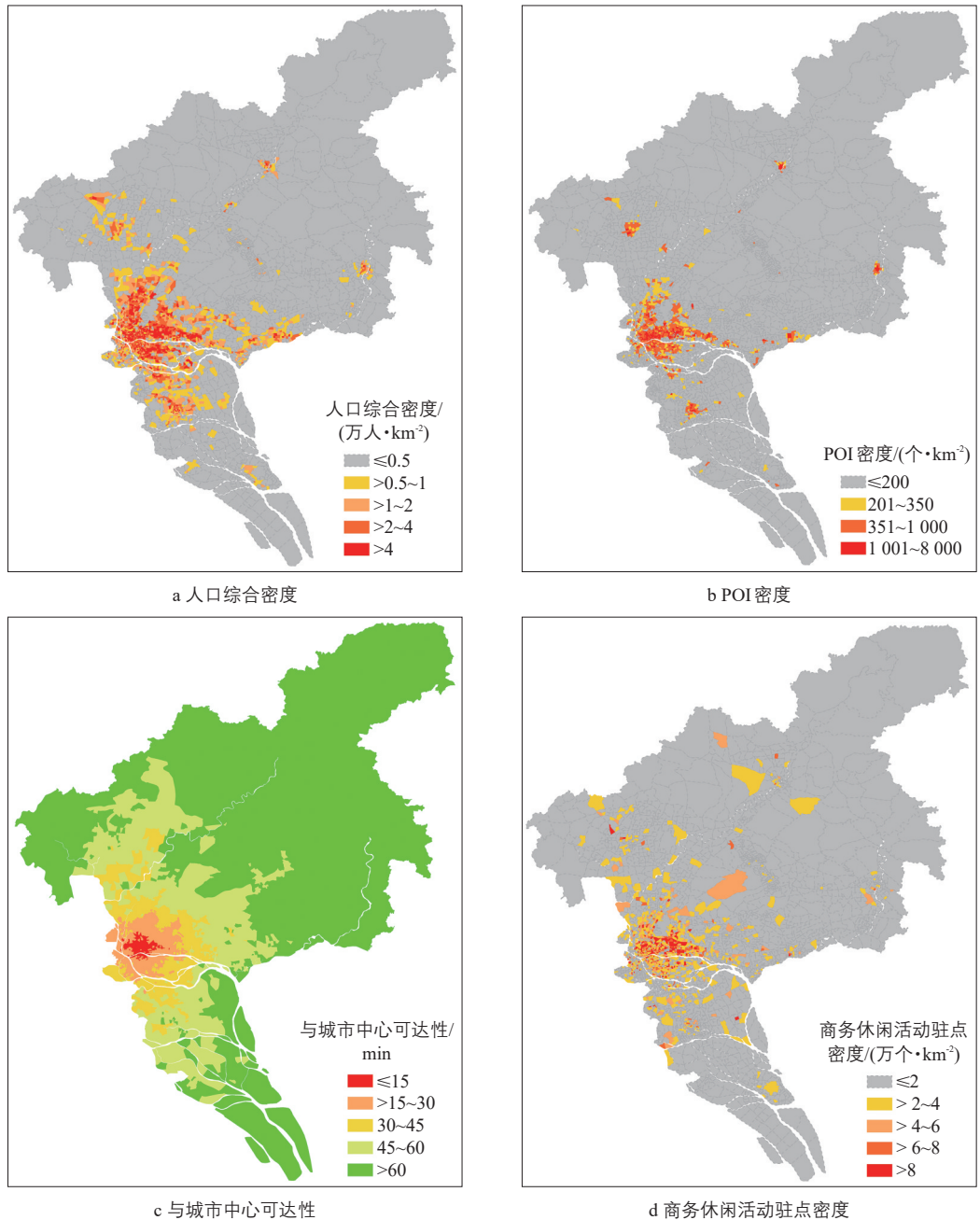


图3 四大指标的空间分布

Fig.3 Spatial distribution of four indicators

的支撑。然而，本文仅从交通角度实现了城市发展边界识别，即物理边界的识别。实际上城市发展边界除物理发展边界外，还包括经济联系、社会文化联系及认同度，该方面研究仍待进一步加强。

参考文献：

References:

[1] 周春山, 叶昌东. 中国城市空间结构研究综述[J]. 地理科学进展, 2013, 32(7): 1030-

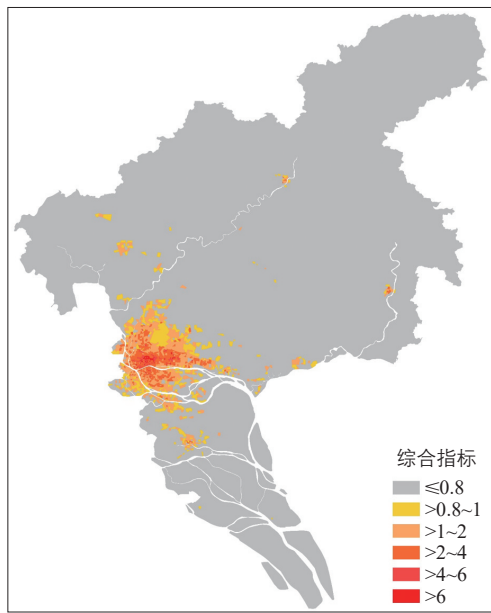


图5 综合指标空间分布

Fig.5 Spatial distribution of comprehensive indicators

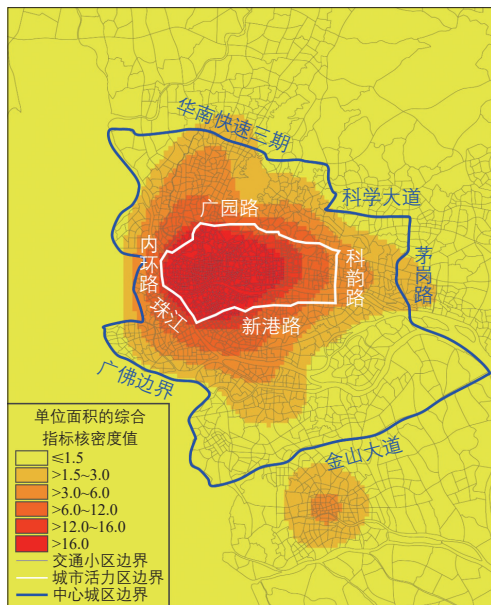


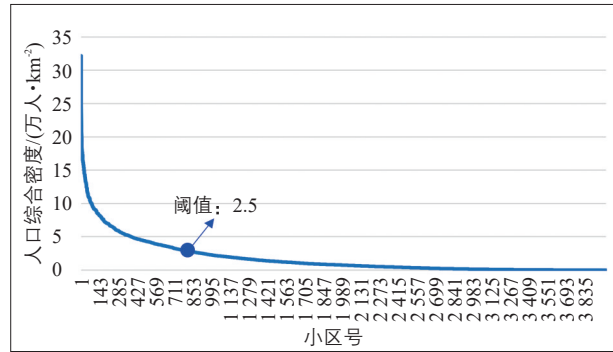
图6 基于核密度识别的广州市活力区和中心城区边界

Fig.6 Boundary of the identified vitality zone and central district in Guangzhou based on kernel density identification

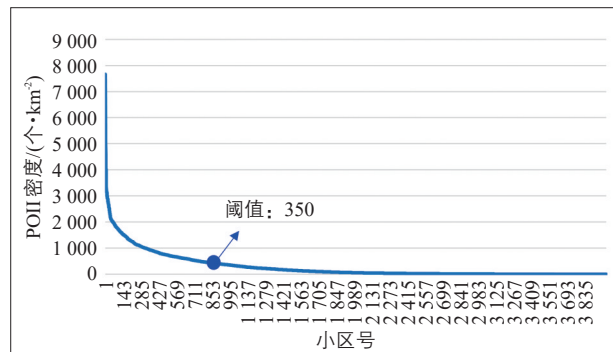
1038.

[2] 王建伟, 吴永平, 刘昕. 都市圈圈层界定方法[J]. 建筑科学与工程学报, 2007(2): 91-94.

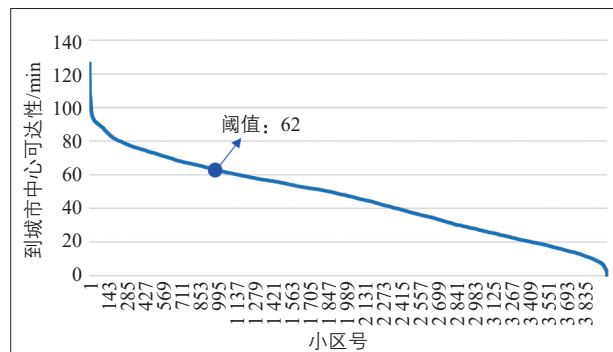
Wang Jianwei, Wu Yongping, Liu Xin. Meth-



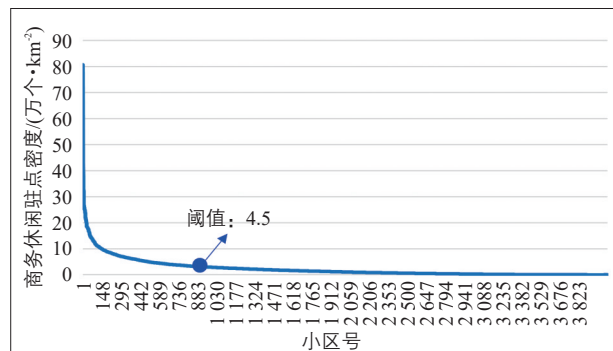
a 人口综合密度



b POI密度



c 与城市中心可达性



d 商务休闲驻点

图4 四大指标分布排序曲线

Fig.4 Distribution sorting curve of four indicators

表3 识别的广州市活力区和中心城区范围与同类城市对比
Tab.3 Comparison of the identified vitality zone and central district in Guangzhou and other compatible cities

城市	活力区面积/km ²	中心城区面积/km ²
广州	93	600
北京	93(东城区、西城区)	667(五环内)
纽约	59.5(曼哈顿)	
伦敦	27(中央伦敦)	320(内伦敦)
东京	71.8(都心6区)	617(23区部)
巴黎	105(小巴黎)	657(大巴黎)

od of Circle Region Classification of Metropolitan Region[J]. Journal of Architecture and Civil Engineering, 2007(2): 91-94.

[3] 郭爱君, 冯琦媛. 兰州都市圈空间界定方法研究[J]. 甘肃社会科学, 2009(6): 137-140.
[4] 刘鹏程, 陈超, 孔斌, 等. 基于POI数据的镇江中心城区空间结构识别与特征解析[J]. 江苏城市规划, 2017(9): 7-12.
[5] 段亚明, 刘勇, 刘秀华, 等. 基于POI大数据的重庆主城区多中心识别[J]. 自然资源学报, 2018, 33(5): 788-800.
Duan Yaming, Liu Yong, Liu Xiuhua, et al. Identification of Polycentric Urban Structure

of Central Chongqing Using Points of Interest Big Data[J]. Journal of Natural Resources, 2018, 33(5): 788-800.

[6] 明英男. 基于网络开放数据的城市空间布局识别研究: 以辽宁省盘锦市为例[C]//中国城市规划学会. 共享与品质: 2018中国城市规划年会论文集. 北京: 中国建筑工业出版社, 2018: 1073-1082.
[7] 陈映雪. 基于Webmap的多中心城市空间发展评估方法: 以上海主城区为例[J]. 上海城市规划, 2017(6): 96-102.
Chen Yingxue. Study on Polycentric Urban Spatial Development Evaluation Approaches Based on the Webmap: A Case of Shanghai Main Urban Area[J]. Shanghai Urban Planning Review, 2017(6): 96-102.
[8] Chu H J, Liao C J, Lin C H, et al. Integration of Fuzzy Cluster Analysis and Kernel Density Estimation for Tracking Typhoon Trajectories in the Taiwan Region[J]. Expert Systems with Applications, 2012, 39(10): 9451-9457.
[9] Xie Zhixiao, Yan Jun. Kernl Density Estimation of Traffic Accidents in a Network Space [J]. Computers Environment & Urban Systems, 2008, 32(5): 396-406.

(上接第37页)

参考文献:

References:

[1] 周军. 道路卡口通行数据挖掘研究[J]. 警察技术, 2014(3): 91-93.
[2] 畅玉皎, 杨东援. 基于车牌照数据的通勤特征车辆识别研究[J]. 交通运输系统工程与信息, 2016, 16(2): 77-82+112.
Chang Yujiao, Yang Dongyuan. Recognition of Vehicles with Commuting Property Using License Plate Data[J]. Journal of Transportation Systems Engineering and Information Technology, 2016, 16(2): 77-82+112.
[3] 龙小强, 苏跃江, 余畅, 等. 基于卡口车牌识别数据的车辆出行分析[J]. 交通运输系统工程与信息, 2019, 19(2): 66-72.
Long Xiaoqiang, Su Yuejiang, Yu Chang, et al. Analyzing Methods of Vehicle Travel Using Plate Recognition Data[J]. Journal of Transportation Systems Engineering and Information Technology, 2019, 19(2): 66-72.
[4] 周韬, 高杨斌, 裴洪雨. 基于卡口数据的快

速路车辆出行OD计算方法[C]//中国城市规划学会. 共享与品质: 2018年中国城市规划年会论文集. 北京: 中国建筑工业出版社, 2018: 3431-3441.

[5] 王蓓, 宁平华, 段小梅, 等. 卡口数据挖掘与城市道路交通分析[J]. 城市交通, 2019, 17(1): 89-95+6.
Wang Bei, Ning Pinghua, Duan Xiaomei, et al. Data Mining and Urban Road Traffic Analysis Based on Traffic Camera Data[J]. Urban Transport of China, 2019, 17(1): 89-95+6.
[6] 阮树斌, 王福建, 马东方, 等. 基于车牌识别数据的机动车出行轨迹提取算法[J]. 浙江大学学报(工学版), 2018, 52(5): 836-844.
Ruan Shubin, Wang Fujian, Ma Dongfang, et al. Vehicle Trajectory Extraction Algorithm Based on License Plate Recognition Data[J]. Journal of Zhejiang University (Engineering Science), 2018, 52(5): 836-844.
[7] 李敏茜. 基于轨迹数据的套牌车检测技术研究[D]. 上海: 华东师范大学, 2019.