

引用格式:宋辞,裴韬.北京市多尺度中心特征识别与群聚模式发现[J].地球信息科学学报,2019,21(3):384-397. [ Song C, Pei T. Exploring polycentric characteristic and residential cluster patterns of urban city from big data[J]. Journal of Geo-information Science, 2019,21(3):384-397. ] DOI:10.12082/dqxxkx.2019.180608

# 北京市多尺度中心特征识别与群聚模式发现

宋 辞<sup>1,2</sup>, 裴 韬<sup>1,2\*</sup>

1. 中国科学院地理科学与资源研究所 资源与环境信息系统国家重点实验室, 北京 100101; 2. 中国科学院大学, 北京, 100049

## Exploring Polycentric Characteristic and Residential Cluster Patterns of Urban City from Big Data

SONG Ci<sup>1,2</sup>, PEI Tao<sup>1,2</sup>

1. State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; 2. University of Chinese Academy of Sciences, Beijing 100101, China

**Abstract:** With the development of urbanization, the problem of "big city disease" is becoming more and more prominent in recent years. Since the strategy of polycentric city has been proposed, polycentric characteristics of city has become one of the most important issue in revealing spatial structure of a city. Due to the rough resolution of traditional census data, most studies have not get an insight into the fine structure of city centers and lacked empirical researches. With the widely application of LBS services, most trajectory data of human activities have been recorded to provide a profile of daily urban dynamics in real time. This information has opened up a new way to analyze the mechanism of polycentric city. Based on the reason, this study proposes a new method to identify the multi-centers in Beijing within rings, and analyzes the spatial-temporal characteristics and interactions in each center, using sina micro blog check-in data and taxi GPS data with POI data. In our study, groups of centers with different social functions can be identified from the clusters of residents' activities. From these results, 4 city centers, 16 district centers and 45 community centers have been identified within 5th rings in Beijing. The district centers can be divided into 4 groups, including cultural and recreational center, business center, education center and transportation center. And the representativeness for each types of centers are Sanlitun, Guomao, Beijing Normal University and Beijing West Railway Station. The community centers can be divided into 9 groups, including political center, residential center, administration center and business-residential center. The representativeness for each types of centers are Xidan, Wudaokou, Wanliu, Qianmen, Zhongguancun, Tianmen square, Chaoyang District Government, Beijing West Railway Station and Yonganli Community. The residential clusters can be observed in some daily hours. Number of residential clusters for district centers are commonly higher than number for community centers, while number of residential clusters in workdays area higher than that in holidays. For district centers, there are 4 classes of common cluster pattern in workdays and 5 classes of common cluster pattern in holidays. For community

收稿日期:2018-11-26;修回日期:2018-12-30.

基金项目:国家自然科学基金项目(41601430、41421001、41525004);国家重点研发计划项目(2017YFB0503801)。[ **Foundation items:** National Natural Science Foundation of China, No.41601430, 41421001, 41525004; National Key R&D Program of China, No.2017YFB0503801. ]

作者简介:宋 辞(1986-),男,湖北赤壁人,助理研究员,主要研究方向为时空数据挖掘。E-mail: songc@lreis.ac.cn

\*通讯作者:裴 韬(1972-),男,江苏扬州人,研究员,主要从事地理大数据挖掘方面的研究。E-mail: peit@lreis.ac.cn

centers, there are 5 classes of common cluster pattern in workdays and 3 classes of common cluster pattern in holidays, respectively.

**Key words:** polycentric center; residents' activities; social function; clustering process; residential cluster patterns

**\*Corresponding author:** PEI Tao, E-mail: peit@lreis.ac.cn

**摘要:**城市多中心的研究是揭开城市系统复杂空间结构的重要研究内容,然而受传统调查数据时空精度的局限,现有研究缺乏足够细致的实证分析。LBS服务的广泛应用为高效、实时、微观地刻画个体和城市之间的空间活动提供了充足的信息,使得对城市多中心的微观量化机制的研究成为可能。基于此原因,本文结合兴趣点数据以及微博签到数据和出租车GPS轨迹数据对北京市五环内的不同尺度、不同社会功能的城市中心进行了识别,并对其居民行为的群聚模式进行了分析。结果表明,北京市五环内包括4个市级中心,16个区级中心以及45个街道级中心。市级中心包括海淀科教中心,西单休闲中心,北京文化中心和国贸金融中心4大中心;区级中心主要包括文娱中心、商业中心、教育中心以及交通中心4类中心;街道级中心除了区级的四类中心外,还包括行政中心、购物中心、商住中心等9类中心。不同中心在同一时段呈现相似的群聚模式,其中工作日城市不同尺度中心的群聚模式持续时间较长,群聚密度较大,主要以通勤和饮食为主;非工作日城市不同尺度中心的群聚模式相对较分散,持续时间较短,主要以休闲娱乐为主。

**关键词:**多尺度城市中心;居民活跃度;社会功能;丛聚过程;群聚模式

## 1 引言

随着中国城市化进程的加快,“城市病”问题日益突出,严重制约了城市化的发展,许多不堪重负的大城市纷纷确定了“多中心”的空间发展战略。多中心城市不仅可以承载城市某方面的职能<sup>[1-3]</sup>,同时也对城市各区起到了连接作用<sup>[4]</sup>,是城市系统中极其重要的组成部分<sup>[5]</sup>。然而,城市各中心到底在城市空间系统中扮演何种角色,如何去量化城市中心的规模特征、职能作用,多中心的建设是否真的能缓解“大城市病”等问题,却一直缺乏深入论证,是一个难以量化的问题。

目前关于城市中心的研究主要分为2个部分:

① 研究通过定义不同指标方法试图对城市中心的界定进行量化定义。其判断指标通常包括平均人口密度,人口密度梯度<sup>[6-7]</sup>,日常人口出行量,就业密度<sup>[8-9]</sup>等;② 对城市空间结构建模,试图从理论上解释多中心城市体系形成的内在机理<sup>[10-13]</sup>。然而,这些研究还存在一定程度上的不足,一方面仅从城市发展机制角度开展研究<sup>[14-16]</sup>,缺乏实证分析,往往使得理论和模型的说服力不够,容易受到质疑;另一方面,通过普查数据或者中心模型对城市人口密度分布拟合,虽然在一定程度上提供了定量分析手段,但其采用的数据空间分辨率往往较粗,更新速度慢,难以对城市微观尺度的中心特征进行定量分析,无法表征城市各类中心的社会功能标注、日常运营模式、以及城市空间交互等特征。

随着具有LBS服务的移动计算设备广泛使用,大量具有城市居民个体标记的移动轨迹数据被记录下来,为高效、实时、微观地刻画个体和城市之间的空间活动提供了充足的数据信息<sup>[17]</sup>。个体移动轨迹数据从居民活动的角度反映了城市的多中心动态特征及其空间交互模式,使得对城市多中心的微观量化机制的研究成为可能。目前,已有许多研究通过出租车GPS轨迹数据<sup>[18-21]</sup>,交通智能卡数据<sup>[22-24]</sup>,手机数据<sup>[25-27]</sup>以及社交网签到数据<sup>[28-32]</sup>对城市空间结构进行了研究,从而体现出城市内部不同功能区之间的交互作用。

基于此原因,本研究拟采用移动轨迹数据对北京市的多尺度中心特征进行量化分析,以弥补传统方法研究城市中心的不足,从而为城市空间结构的理论模型提供实证补充。

## 2 数据源与研究方法

### 2.1 研究区概况

本研究选取了北京市五环内为研究区。研究区范围覆盖了西城区,东城区,海淀区,朝阳区,石景山区,丰台区以及大兴区(图1),总面积为668.65 km<sup>2</sup>,2017年总人口在1000万人以上。该研究区是北京市城市发展的核心区域,是经济、文化、科技、教育等发展中心,聚集了北京市全区绝大多数的金融机构、大型企业、科研单位、医疗机构等,并形成了“国贸”、“金融街”、“中关村”等相关行业密集分布的产业功能区。

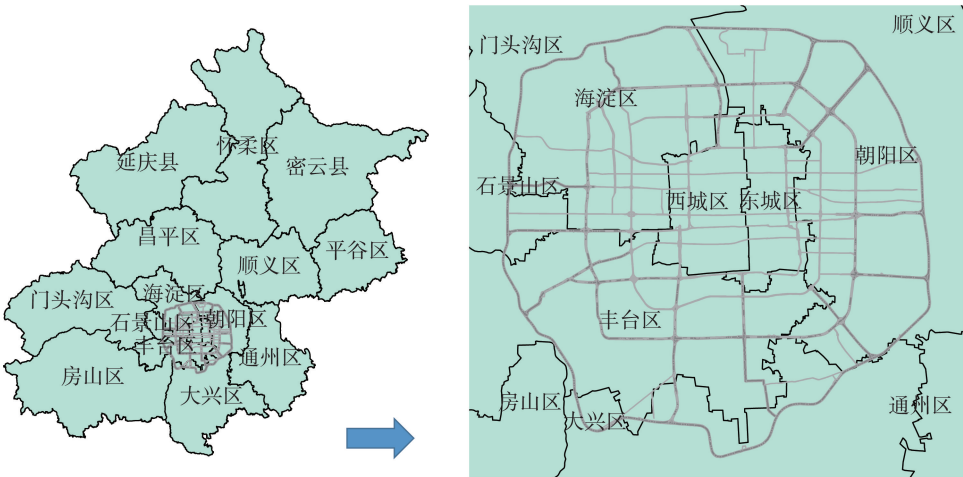


图1 研究区示意  
Fig. 1 Research area

2.2 数据源

研究数据来源包括北京市五环内的兴趣点 (POI)数据,出租车 GPS 轨迹数据以及新浪微博签到数据(2014年2月1日-2014年2月28日)。其中 POI 数据包括 14 种不同类别(交通设施,休闲娱乐,公共服务,公司,医院,小区,文教中心,旅馆酒店,汽车服务,行政单位,购物,风景旅游,餐饮以及其他),主要在一定程度上体现区域的基础设施配置以及主体社会功能。出租车轨迹数据用于提取载客点信息从而反映居民出行特征,为城市空间结构分析、土地利用规划提供支撑。新浪微博签到位置服务数据具有空间、时间、用户社会属性的数据,帮助我们分析城市居民活动特征,从而更好地认识城市的空间结构。当然,签到数据由于用户行为并不连续,使得用户轨迹并不完整,因此对居民活动特征研究也存在一定的不确定性。

2.3 研究方法

2.3.1 研究方案

本文首先通过出租车 GPS 轨迹数据以及微博签到数据提取居民出行的 OD 锚点作为人口活跃度指标,以 POI 兴趣点构建泰森多边形作为基本分析单元,采用空间层次聚类方法识别城市人口聚集中心,并根据其居民动态特征对城市中心进行了分类,得到承载不同功能的城市中心。最后识别了居民出行的从聚过程,并采用频繁模式挖掘方法对各个中心的人群集聚过程进行了分析,发现工作日/非工作日不同人群的群聚模式特征。总体研究方案如图 2 所示。

2.3.2 多尺度城市中心的空间聚类

城市中心的识别算法采用文献[33]的层次聚类的方法,以兴趣点所在的泰森多边形为最小单元,依次搜索距离较近的区域,如其活跃度满足一定的程度,就将其聚集为一类。具体算法步骤如下:

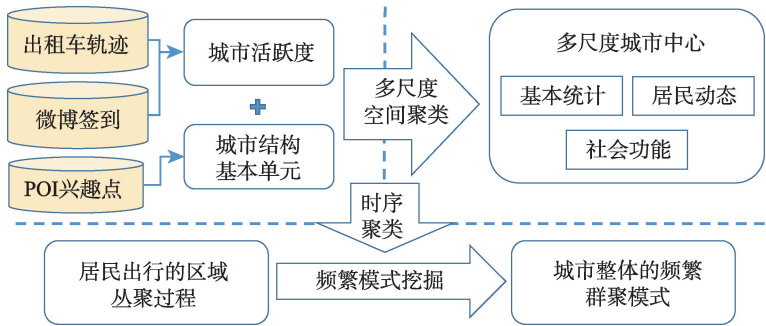


图2 总体研究方案  
Fig. 2 Research scheme

(1)对兴趣点集合  $P$  做泰森多边形,统计泰森多边形的面积  $Area_i$  以及多边形内的活跃度数量  $NumC_i$ ;

(2)对于每个兴趣点  $P_i$ ,记录其属于簇  $C_i$ 。搜索出其距离半径  $r$  内的其他兴趣点  $PSet_i$ ,并计算其距离  $DistSet_i(DistSet_i = \{Dist(P_i, P_j) | P_j \in PSet_i\})$ ;

(3)根据距离  $DistSet_i$  排序(升序);

(4)对于当前距离最近的2个兴趣点  $P_i$  和  $P_j$ ,若  $P_i \in C_i, P_j \in C_j$  且  $i \neq j$ 。统计  $C_i$  和  $C_j$  所占面积  $Area_c(Area_c = \sum_{P_m \in C} Area_m)$  以及活跃度  $NumC_c(NumC_c = \sum_{P_m \in C} NumC_m)$ 。若活跃度密度  $Dense_c \geq Dense_{AVE}$ ,则将  $C_i$  和  $C_j$  合并为新簇  $C_x, C_x = \{C_i | i, j \in x\}$ ,这里的  $Dense_{AVE}$  表示所有簇的活跃度密度平均值;

(5)所有搜索半径内的点对遍历完后,可识别城市中心的范围。

针对不同尺度的城市中心,这里采用了不同尺度的搜索半径来识别。图3(a)给出了平均活跃度密度随搜索半径变化的情况。通过统计尺度指标  $I_r = \log(Dense_r) / \log(r)$  来选择不同的半径阈值识别不同尺度的中心(图3(b))。可以发现当搜索半径  $r=50$  m时,活跃度尺度指标  $I_r$  对应着箱线图的上边缘取值(约1.716),该搜索半径可以认为是城市活跃度异常丛聚的最小尺度;当搜索半径  $r=100$  m时,活跃度尺度指标  $I_r$  对应着箱线图的上四分位数取值(约1.402),表示城市活跃度异常丛聚的最大尺度。因此,本文分别选择搜索半径为50 m和70 m(约对应着  $I_r$  的85分位数)和100 m来识别不同尺度的

城市中心范围。

### 2.3.3 群聚模式识别

这里的群聚模式表现为多个中心同时出现丛聚过程的活跃时段。因此需要先识别出各个中心的丛聚过程,然后在统计各中心共同的活跃时段发现群聚模式。这里所定义的居民出行的丛聚过程指城市中心范围内,活跃度显著较高的时段,主要强调居民活动在时间上的聚集特征。因此,这里采用DBSCAN聚类方法对各中心居民活跃时间进行聚类分析。具体步骤如下:

(1)针对特定范围的城市中心  $C_i$ ,搜索在其内的活动事件  $ESet_i$ (微博签到点或出租车到达点),记录每个活动事件的发生时间  $ESet_i.T$ 。

(2)对该中心所有活动事件的发生时间进行排序,计算该城市中心的单位时间内的平均活动密度  $\lambda_i$ 。

(3)根据泊松分布,计算该中心单位时间到达密度  $\lambda_i$  下的0.01显著水平的活跃度阈值  $n_i^0$ ,使得

$$\sum_{N=n_i^0}^{\infty} ((\lambda t)^N \cdot e^{-\lambda t}) / N! < \alpha, n_i^0 \text{ 作为活动事件丛聚的最小}$$

阈值。 $t$ 表示用于统计活动事件数量的时间长度,这里我们选择15 min(居民出行锚点的最小停留时长)作为统计基础。

(4)根据DBSCAN算法,对所有活动事件的时间进行聚类,最终得出城市各中心活动事件的丛聚过程(也就是活动事件聚集的时段)。

在识别出各个城市中心的丛聚模式后,通过时序频繁模式识别方法找出其公共的活跃时段,即可识别出群聚模式。由于各类市级中心数量较少,因此群聚模式的识别主要针对区级中心和街道级中心进行分析。

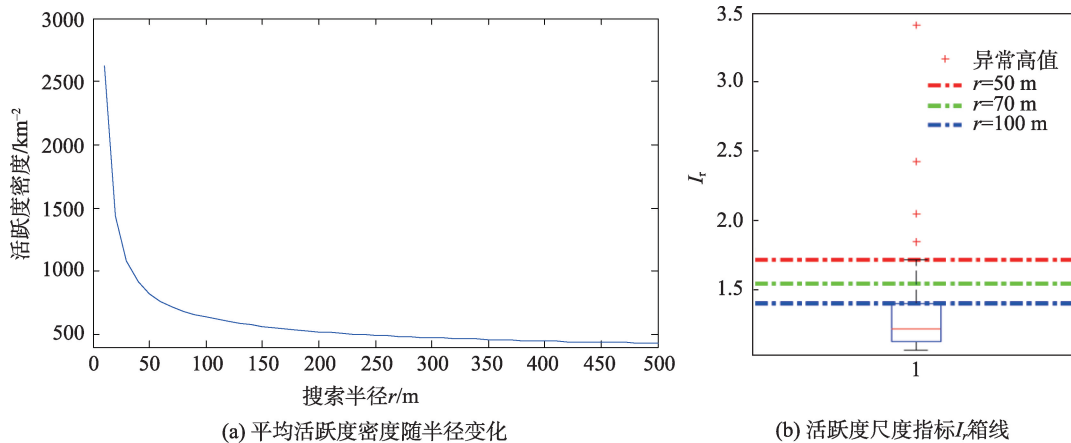


图3 平均活跃度密度随搜索半径的变化情况

Fig. 3 Variations of activity density in different radius

3 北京市多尺度城市中心特征分析

基于上述方法,对北京市多尺度城市中心进行了识别并对其居民活动特征进行了分析,包括其中心规模,活跃度时序特征等。最终得到了北京市五环范围内的4个市级中心,16个区中心和45个街道级中心。

3.1 市级中心识别结果

市级中心包括海淀科教中心,西单休闲中心,北京文化中心和国贸金融中心4大中心(图4)。四大中心中东城区文化中心面积最大达到了20.49 km<sup>2</sup>。国贸金融中心和西单休闲中心面积虽然很小,其微博签到密度却很大,尤其是西单休闲中心,密度达到5571.2次/km<sup>2</sup>。相比之下,出租车OD点密度最高的地方则在国贸金融中心,其次是东城区文化中心。而海淀科教中心的面积虽然是第二大的中心,但其签到密度和出租车OD点密度都比较小(表1)。

四大市级中心的居民活跃度时序变化特征也十分显著(图5)。就微博活跃度来看,四大中心的变化趋势都是从早上开始上升,下午或晚间开始下降。其中,西单休闲中心平均活跃度最高,国贸金融中心其次,海淀科教中心最小。国贸中心的微博活跃度在早上6-9时之间的上升最为显著,9时左

右达到第一个峰值,随后随时间波动较为明显,从9-16时,有若干个小时段的峰值,但难以看出大时段的峰值。16时以后呈波动下降趋势直到夜间。东城区文化中心的微博活跃度早上上升得相对较平稳,其上升时段也相对较长,从早上6点一直到下午14时。在14-19时变化相对较为稳定,19时以后开始缓慢下降。海淀科教中心的微博活跃度呈梯形变化,早上6时开始上升到9时后,一直处于一个



图4 北京市五环内4类市级中心  
Fig. 4 Four city centers in 5th ring in Beijing

表1 市级中心的规模特征  
Tab. 1 Basic information of city centers

序号	中心名称	中心面积/km <sup>2</sup>	微博签到密度/次/km <sup>2</sup>	出租车OD点密度/(次/km <sup>2</sup> )
1	国贸金融中心	2.88	5476.7	18 532.9
2	东城区文化中心	20.49	4050.5	11 235.4
3	海淀科教中心	12.51	2614.3	9282.2
4	西单休闲中心	1.90	5571.2	7691.5

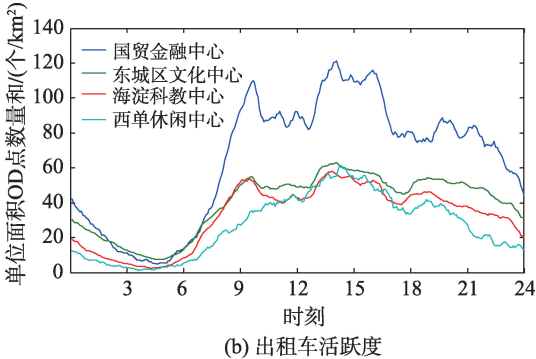
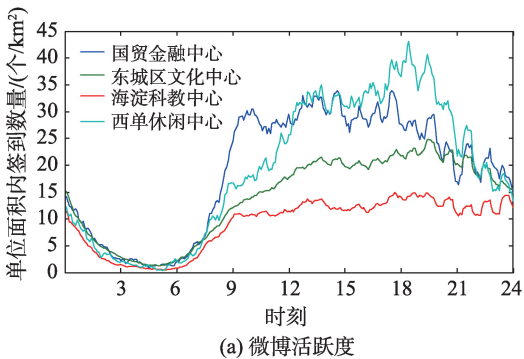


图5 四大市级中心活跃度  
Fig. 5 Activity variations of four city centers in 5th ring in Beijing

比较稳定的变化状态,一直持续到夜间,这可能是因为该区学生居多,其活动范围较为有限。西单休闲中心的微博活跃度变化存在一个双峰模式,6-13时,达到第一个峰值;随后继续上升,在下午19时达到第二个峰值,然后快速下降直至夜间。相对微博活跃度,出租车活跃度是时序变化曲线则更为平滑,且四大市级中心呈现相似的“三峰”模式——早高峰(9时),午高峰(13-16时)和晚高峰(19时/20时左右)。国贸金融中心的出租车活跃度最高,波动幅度较大。东城区文化中心和海淀科教中心的出租车活跃度变化较为相似,峰出现的时间和值都相差很小。西单休闲中心的微博活跃度最高,但出租车活跃度最低。该中心的三峰模式最不明显,整体呈现以15时为中心的对称变化,即15时以前上升,15时以后下降。其波动较其他中心相对较小。

3.2 区级中心识别结果

区级中心共16个,主要分布在北三环、四环内和东二环、三环内(图6)。面积最大的区级中心是东三环的呼家楼,约占3.52 km<sup>2</sup>,此外,工体-三里屯、王府井-东单等中心面积也都超过了3 km<sup>2</sup>;最小的是北京南站,仅有0.11 km<sup>2</sup>。微博签到密度最高的中心在西单,达到了14 168.8次/km<sup>2</sup>,其次是国贸和北京化工大学附近。除了北航外,中央民族大学、人民大学、五道口(北京语言大学和北京地质大学)、北京化工大学的微博签到密度都很高,都超过了4000次/km<sup>2</sup>。北京南站的微博签到密度虽然只有700次/km<sup>2</sup>,但其出租车OD点密度却达到了45 569.6次/km<sup>2</sup>,是出租车OD点密度最高的区中心。其次作为交通要地的北京西站出租车OD点密度也达到了17 308.6次/km<sup>2</sup>。出租车OD点密度最小的中心是西单,只有2766.2次/km<sup>2</sup>(表2)。

16个区级中心共可以分为:文娱中心,商业中心,教育中心和交通中心四类中心。这里商业中心包括了以购物为主的商业中心和金融活动为主的经济中心。其中商业中心的微博活跃度最高,接近

6000次/km<sup>2</sup>,文化娱乐中心的微博活跃度最少,仅有3000次/km<sup>2</sup>。而以北京西站和北京南站为代表的交通中心,出租车活跃度最高,达到了19 877.8次/km<sup>2</sup>,教育中心的出租车活跃度最低,约8309次/km<sup>2</sup>。文娱中心主要包括工体-三里屯,北新桥-东四2处区级中心;商业中心主要包括王府井-东单、中关村-人民大学、呼家楼、崇文门、国贸、华贸、西单;教育中心主要包括北京的各大高校——中央民族大学、五道口(北京语言大学,地质大学)、北京航空航天大学、北京师范大学和北京化工大学;交通中心则包括北京西站和北京南站两个区级中心(表2)。

各类区级中心的微博活跃度都展现了不同的变化曲线特征(图7(a))。可以看出文娱中心的工体-三里屯,北新桥-东四主要以购物、休闲、娱乐为主,工作单位较少,因此微博活跃度并没有明显早高峰,高峰时段主要在中午13-14时和晚上20时。商业中心则主要以工作地为主,居民活跃度的高峰期主要从12时到19时,之后显著下降;教育中心以高校为主,微博活跃度以学生居多。因此可以明显的看出早上9时左右的一个小高峰;此外,越到晚

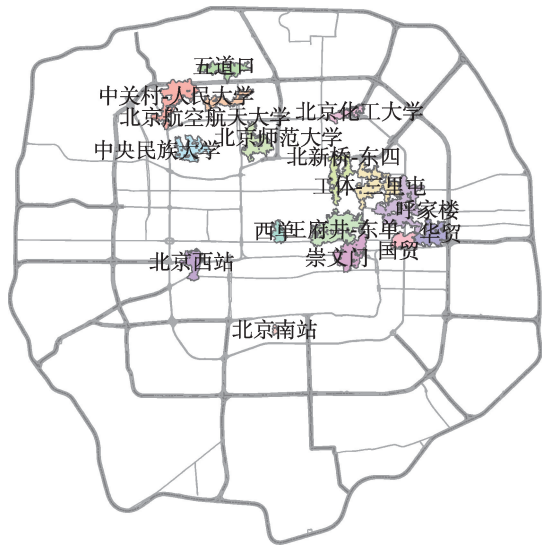


图6 北京市五环内16个区级中心  
Fig. 6 Sixteen district centers in 5th ring in Beijing

表2 各类区级中心的特征  
Tab. 2 Basic information of district centers

区级中心类别		中心名称	微博活跃度密度/(次/km <sup>2</sup> )	出租车活跃度密度/(次/km <sup>2</sup> )
文娱中心	工体-三里屯,北新桥-东四		3098.2	14 293.6
商业中心	王府井-东单,中关村-人民大学,呼家楼,崇文门,国贸,华贸,西单		5996.1	12 244.7
教育中心	中央民族大学,五道口,北京航空航天大学,北京师范大学,北京化工大学		3795.1	8309.0
交通中心	北京西站,北京南站		4145.9	19 877.8

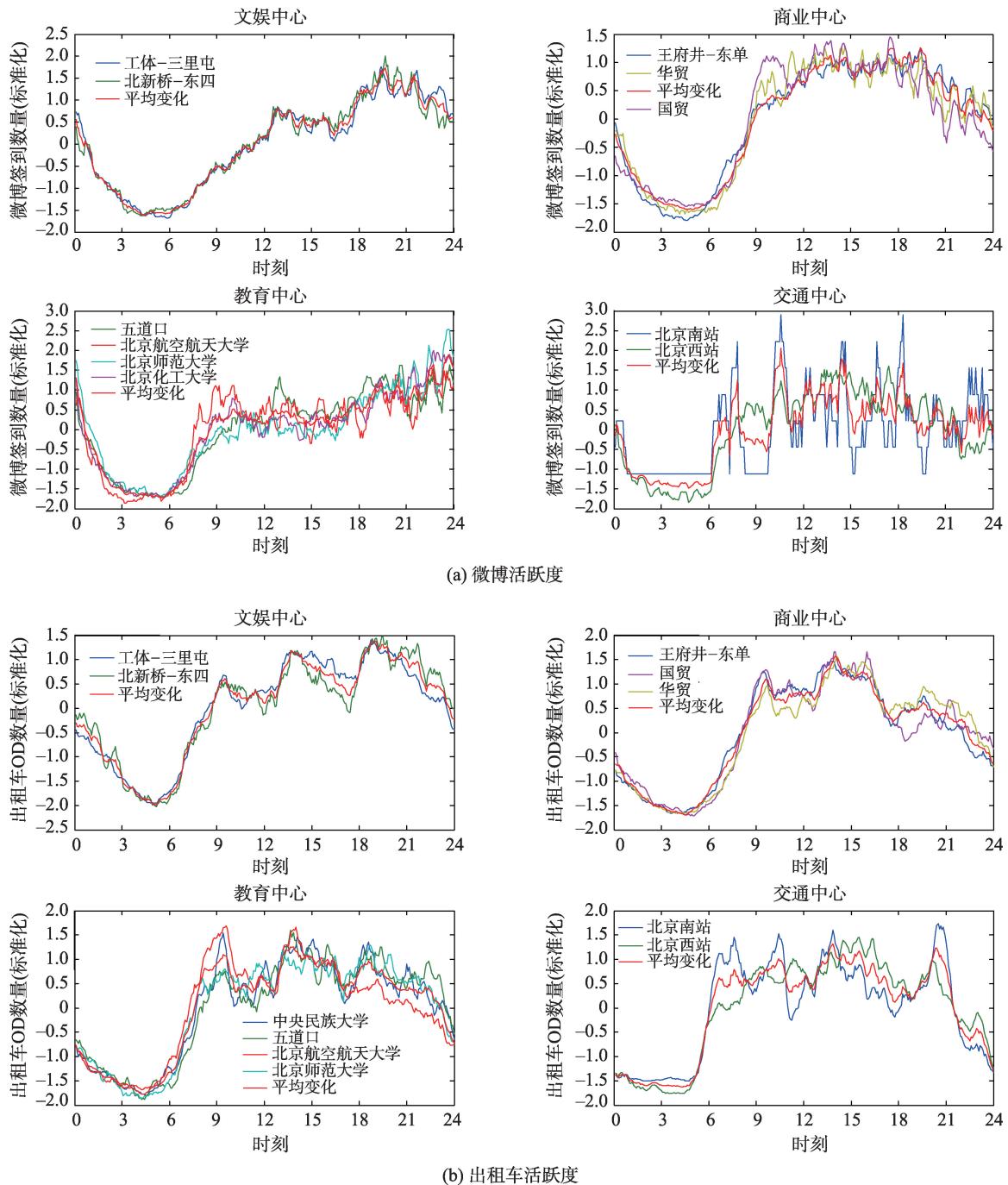


图7 4类区级中心的活跃度变化

Fig. 7 Activity variations of sixteen district centers in 5th ring in Beijing

间,其活跃度越高。交通中心的微博活跃度则展现出多峰模式,与列车车次频率密切相关。相比微博活跃度,各类区级中心的出租车活跃度都呈现了显著的多峰模式,这与出租车本身的运营特点有关(图7(b))。文娱中心的出租车活跃度具有早、午、晚3个峰值模式,分别发生在9时,14时和19时,且晚上的峰值最高(这点与微博活跃度具有一致性);商

业中心的出租车活跃度也呈现“中间高,两头低”的变化趋势,在早上10时和中午14时达到峰值,16时以后显著下降;教育中心的出租车活跃度则与微博活跃度显著不同,晚间的活跃度明显降低,活跃度峰值则主要在早上和中午时段;交通中心的出租车活跃度则同样呈现出多峰模式,在7时,10时,13时,16时和21时分别都达到了峰值。

### 3.3 街道级中心识别结果

街道级中心共识别出了45个(图8),其规模范围相对更小,最大的王府井-东单购物中心只有1.38 km<sup>2</sup>。微博签到密度较高的中心在北京南站旁边的四路通社区,国贸以及朝阳区政府,其微博签到密度都超过了20 000次/km<sup>2</sup>。其次在西单、科春社区、五道口、天安门等地,微博签到密度也超过了15 000次/km<sup>2</sup>。签到密度最小的中心则主要是分布在四环外的住宅社区,包括南湖小区,枫林绿洲,学院路,金泉广场等地区,签到密度在2000次/km<sup>2</sup>以下。出租车活跃度密度较高的街道级中心主要集中在东三环的商务中心处,其中最高的街道级中心包括东三环CBD中心,三里屯-团结湖,朝外SO-



图8 北京市五环内45个街道级中心

Fig. 8 Forty-five community centers in 5th ring in Beijing

HO,北京西站等地。此外,双榆树东里,北航,呼家楼,万柳这样微博签到较低的地方,其出租车活跃度密度也十分高,这些街道级中心往往承载了一定的居住区功能。出租车活跃度最低的街道级中心在小红门,前门等地,都没超过2000次/km<sup>2</sup>(表3)。

街道级中心由于规模相对更小,因此其社会功能更为单一,其作为城市中心的类型的划分则相对更细致一些。本文在区级中心分类的基础上将街道级中心划分为9种类型:商业中心,教育中心,住宅中心,文化中心,娱乐购物中心,政治中心,行政中心,交通中心,商住中心(表3)。住宅中心主要指居民集中居住的社区;文化中心主要指老北京建筑、字号,饮食为主的一些居民聚集中心;娱乐购物中心,则主要指以服务行业为主的休闲购物中心;商住中心指集办公,住宅为一体的居民社区。

作为政治中心的天安门,其微博活跃度是最高的,其次行政中心以及交通中心依然也占据了大量的微博活跃度;住宅中心的微博活跃度最小。交通中心同样具有最高的出租车活跃度,其次娱乐购物中心和商业中心的出租车活跃度也非常高,都在15 000次/km<sup>2</sup>以上。政治中心天安门以及文化中心由于车辆管制和交通原因,出租车活跃度最低(表3)。

图9给出了街道级中心居民活跃度的不同变化曲线。从微博活跃度变化(图9(a))可以看出商业中心和商住中心都呈现白天高,夜晚低的变化模式。其主要区别在于商业中心晚间活跃度的下降趋势更大。教育中心和住宅中心具有相似的变化模式,都呈现白天稳定变化,夜晚显著上升的趋势,但教育中心晚上活跃度的上升更为显著。文化中心和

表3 各类街道级中心的特征

Tab. 3 Basic information of community centers

街道级中心类别	中心名称	微博活跃度密度/(次/km <sup>2</sup> )	出租车活跃度密度/(次/km <sup>2</sup> )
商业中心	西单,王府井-东单,呼家楼,CBD,国贸	9424.5	15 843.7
教育中心	北师大,中日友好医院,北京联合大学,北航,学院路,中央民族大学,新外大街,五道口	4151.3	9343.3
住宅中心	朝阳园,万柳,南湖小区,金泉广场,时代国际嘉园,西什库大街,渠门内,陶然亭,阜通东大街,小红门	2720.6	9958.2
文化中心	簋街,鼓楼,前门,交道口	8704.0	8345.3
娱乐购物中心	中关村购物中心,朝阳大悦城,三里屯-团结湖,崇文门,朝外soho,工体	7105.9	16367.2
政治中心	天安门	16 126.8	7039.5
行政中心	朝阳区政府,海淀区政府	14 959.5	10 697.4
交通中心	北京西站,四路通社区	10 092.6	18 127.0
商住中心	双榆树东里,永安里,科春社区,知春里,白石桥南,枫林绿洲,东直门	6460.0	12 752.4

娱乐购物中心的变化曲线也较为相似,不同之处在于娱乐购物中心的午高峰(14时)更为明显。政治中心天安门早上7时左右有一个显著的高峰,这很可能是天安门的升国旗原因(2月)。行政中心的微

博活跃度变化较为稳定,体现出了行政机构朝九晚五的作息规律。交通中心依然体现出多峰模式,存在多个与列车时刻表对应的居民聚集时段。

从出租车活跃度看,商业中心、教育中心以及

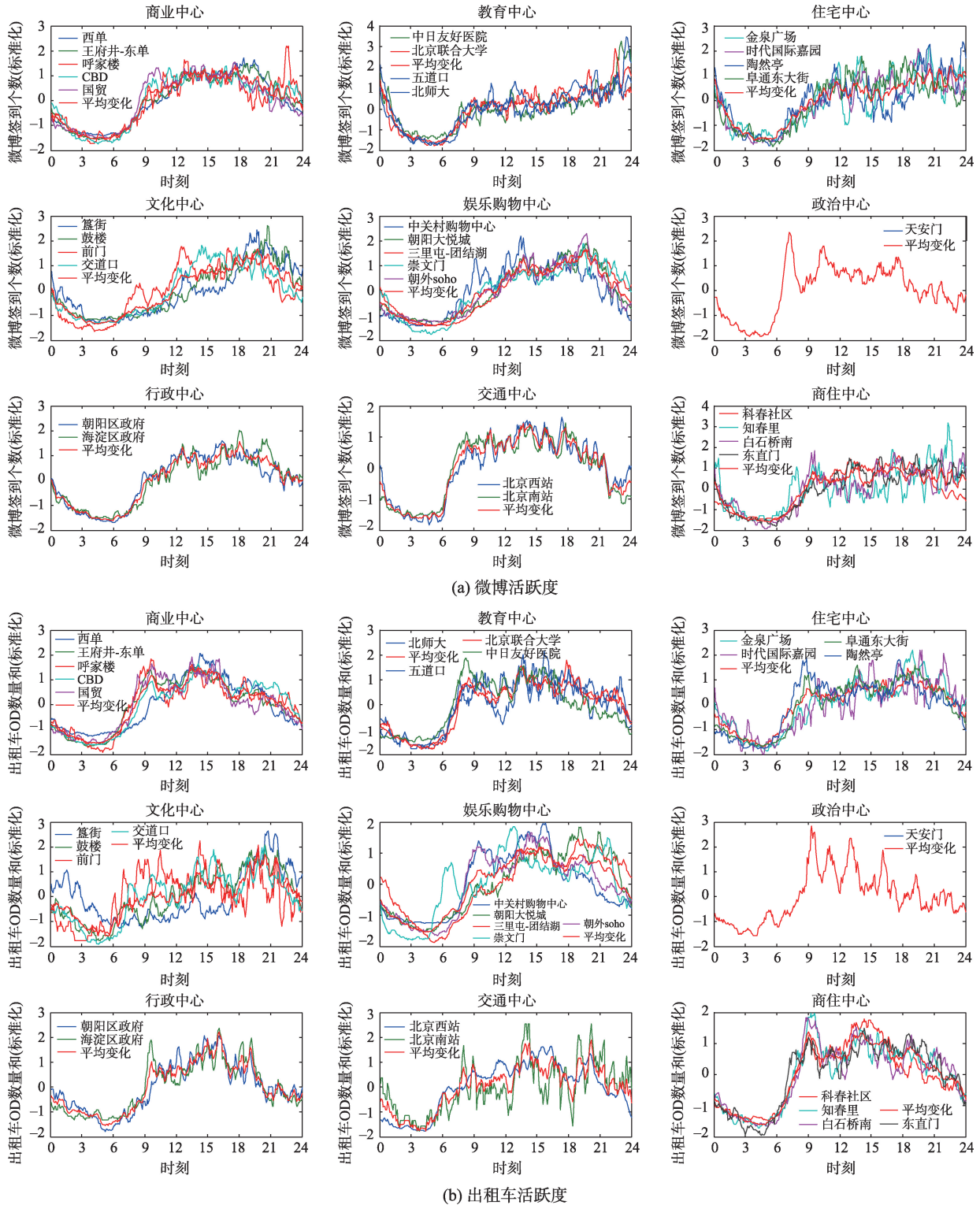


图9 各类街道级中心的活跃度变化

Fig. 9 Activity variations of Forty-five community centers in 5th ring in Beijing

商住中心的变化曲线基本一致,呈现出一个双峰模式,包括9时的一个早高峰和14时的一个午高峰(图9(b))。午高峰过后,出租车活跃度逐渐下降直至夜间。住宅中心的出租车活跃度变化曲线相对比较平稳,从早上6时开始上升,随后8时后一直保持平稳的波动变化状态,到夜间22时以后才开始下降。文化中心的出租车活跃度整体上从白天到晚上呈上升趋势,晚间21时左右达到峰值。娱乐购物中心主要在交通较为发达的道路上,其出租车活跃度峰值时段主要在13-16时之间,该时段交通相对比较便利,而早晚时段的出租车活跃度受交通影响,活跃度相对较小。政治中心天安门的出租车活跃度与微博活跃度基本保持一致,尤其是早7时晚5时的升降旗高峰,其夜间则活跃度很低。行政中心出租车活跃度在9时左右是上升最快的时段,16时左右达到了午高峰,随后开始下降,在晚上19时又达到了一个高峰。交通中心的出租车活跃度高峰期主要在8时、9时、14时和20时。

## 4 北京市多中心群聚模式分析

多个中心同时发生的丛聚过程形成了城市居民的群聚模式。群聚模式反映了不同区域的居民在同一时段的群体行为特征,这些群体行为一方面反映了不同城市中心相似的社会功能,通过资源配置对特定行为的人群产生吸引,造成群聚现象;另一方面,群聚模式也会积极有效地影响着城市空间结构的塑造。本文通过频繁模式挖掘的方法,对区级中心以及街道级中心的群聚模式进行了分析。

### 4.1 区级中心的群聚模式特征

区级中心的丛聚过程在工作日和非工作日显著不同,根据丛聚过程发生的不同时间,可以识别出工作日和非工作日不同的频繁群聚模式。表4给出了工作日内区级中心的群聚模式识别结果,可以看出工作日区级中心包含4种群聚模式;第1类模式属于

早高峰的上班群聚行为,其发生事件在早上8:40-10:20期间,发生该群聚模式的平均群聚密度是最高的,达到了311.66人/天以上。该模式主要发生在王府井-东单,中关村-人民大学,国贸,华贸,呼家楼这些区级中心。第2类模式属午间群聚行为,发生在中午13:00-16:45。该群聚模式的平均群聚密度在161.04人/天,具有相当的规模,这种群聚模式发生在北航,中关村-人民大学,呼家楼,国贸,王府井-东单这些区级中心。该群聚模式可能主要还是对应着居民的工作行为,以及部分休闲行为。第3类和第4类模式都属于晚高峰的娱乐休闲行为,该类群聚模式的时间普遍较短,分别在35 min和40 min,是群聚时间最短的模式了,且2种模式的群聚密度相近,都在90人/天以上。对比4种模式,可以看出工作日早高峰群聚模式和午间群聚模式的持续时间都较长,而2个晚高峰的持续事件普遍较短。这说明早上居民上班行为相对较为集中,各时段都容易产生丛聚过程;而晚间居民回家事件则较为分散,因此产生了多个短持续事件的群聚模式(图10(a))。

非工作日的居民出行相对更为分散,因此群聚模式相对较多,其持续时间也较短(表5)。区级中心的群聚模式共5种,可以看出由于非工作日在早高峰期间没有群聚模式(图10(b))。第1类群聚模式是中午11:35-12:20的午间餐饮聚集行为,该模式的群聚密度在176人/天,主要发生在王府井-东单,北京西站,中关村-人民大学,华贸以及崇文门。其中北京西站的群聚模式主要是由于特定的列车时刻引发的,而其他几个中心都有相当规模的餐饮兴趣点,因此该时段都发生了丛聚过程。第2类群聚模式在12:55-14:20之间,群聚密度也在182人/天,群聚地点与第一类较为类似,只是北京西站变成了北京南站,崇文门变成了北航。第3类群聚模式发生在14:20-15:25之间,其持续事件相对较长,可能对应着居民的休闲,娱乐,购物等行为。该模式发生在王府井-东单,华贸,中央民族大学等地。第4类模式发生下在下15:25-16:00之间,群聚密度在196人/天左右,发生地点则与第3类相似,只

表4 区级中心的群聚模式(工作日)

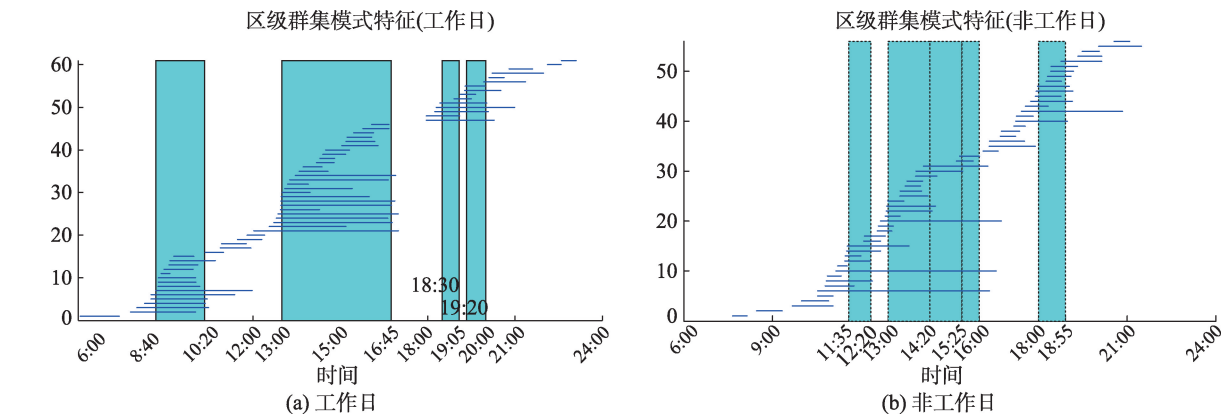
Tab. 4 Residential cluster pattern of district centers (weekday)

模式序号	起始时刻	结束时刻	平均群聚密度/(人/天)	所在中心
1	8:40	10:20	311.66	王府井-东单,中关村-人民大学,国贸,呼家楼,华贸
2	13:00	16:45	161.04	北航,中关村-人民大学,呼家楼,国贸,王府井-东单
3	18:30	19:05	90.08	北师大,中央民族大学,北新桥-东四,工体-三里屯,五道口
4	19:20	20:00	98.95	北师大,北新桥-东四,工体-三里屯,五道口,华贸

表5 区级中心的群聚模式(非工作日)

Tab. 5 Residential cluster pattern of district centers (weekend)

模式序号	起始时刻	结束时刻	平均群聚密度/(人/天)	所在中心
1	11:35	12:20	176.07	王府井-东单,北京西站,中关村-人民大学,华贸,崇文门
2	12:55	14:20	182.09	王府井-东单,中关村-人民大学,华贸,北航,北京南站
3	14:20	15:25	196.27	王府井-东单,中关村-人民大学,华贸,中央民族大学,北京西站
4	15:25	16:00	189.58	王府井-东单,中关村-人民大学,华贸,北京西站,西单
5	18:00	18:55	102.29	五道口,工体-三里屯,中关村-人民大学,中央民族大学,北新桥-东四



注:每条蓝色横线代表的是一个中心发生的丛聚过程,横线的起点终点分别表示该丛聚过程的开始时间和结束时间,纵坐标则对应着不同城市中心发生的丛聚过程编号(一个中心区可能发生多次丛聚过程)。淡蓝色矩形区域表示多个中心同时发生丛聚过程的群聚模式。

图10 区级中心的频繁群聚模式特征

Fig. 10 Residential cluster patterns in district centers

是中央民族大学变成了西单。第5类模式则对应晚高峰的群聚行为,一般是回家或者吃饭行为。该群聚模式时间较短,由18:00-18:55,群聚密度较低,仅102人/天左右,具体发生的点包括五道口,工体-三里屯,中央民族大学和北新桥-东四等。

4.2 街道级中心的群聚模式特征

街道级中心的丛聚过程在空间分布上较分散,

工作日和非工作日也呈不同特征。表6给出了工作日内街道级中心的群聚模式。街道级中心的群聚模式持续时间普遍较短,由于街道级中心面积较小,因此其群聚密度相比区级中心也要小得多(图11(a))。第1类群聚模式同样属于早高峰的上班群聚行为,但该时间主要集中于9:00-9:50之间,发生地点则以中关村购物中心,朝外SOHO,国贸等地为代表,共16个街道级中心。第2类模式发生

表6 街道级中心的群聚模式(工作日)

Tab. 6 Residential cluster pattern of community centers (weekday)

模式序号	起始时刻	结束时刻	平均群聚密度/(人/天)	所在中心
1	9:00	9:50	33.12	西什库大街,陶然亭,枫林绿洲,白石桥南,中关村购物中心,呼家楼,科春社区,朝外SOHO,王府井-东单,知春里,中央民族大学,CBD,国贸,东直门,北师大,天安门
2	13:10	13:40	60.29	崇文门,王府井-东单,西单,中关村购物中心,科春社区,永安里,国贸,中央民族大学,学院路,朝外SOHO,白石桥南,呼家楼,中日友好医院,朝阳区政府,三里屯-团结湖
3	13:40	14:45	77.49	王府井-东单,西单,中关村购物中心,科春社区,永安里,国贸,朝外SOHO,呼家楼,白石桥南,三里屯-团结湖,知春里,CBD,北京西站,双榆树东里
4	15:00	16:15	66.8	王府井-东单,西单,中关村购物中心,科春社区,永安里,朝外SOHO,呼家楼,CBD,北京西站,五道口,国贸,中日友好医院,枫林绿洲,交道口
5	18:25	19:05	31.86	朝阳大悦城,复通东大街,南湖小区,海淀区政府,北京联合大学,金泉广场,鼓楼,万柳,西单,渠门内,双榆树东里,白石桥南,三里屯-团结湖,陶然亭

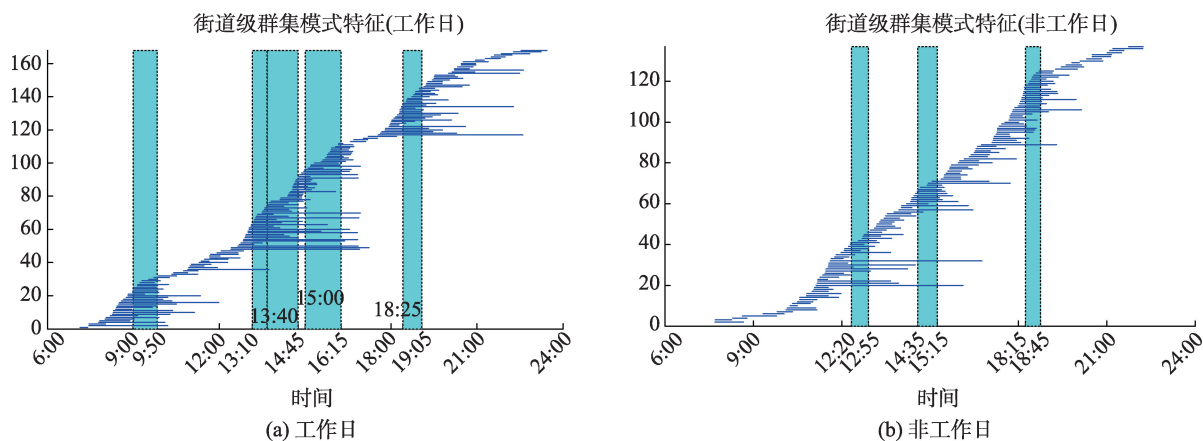
在中午 13:10–13:40,很可能对应着城市居民中午吃饭的群聚行为。其群聚密度相对较大,是早高峰群聚密度的 2 倍左右。该群聚模式主要发生在以崇文门,西单,学院路,朝阳区政府为代表的 15 个街道级中心内。第 3 类模式发生时间紧邻第 2 类模式,从 13:40–14:45,其群聚密度比第 2 类模式较高。该类中心主要发生在以王府井–东单,西单,知春里,北京西站,双榆树东里为代表的 14 个街道级中心,其群聚模式对应的行为可能是下午的上班出行集聚行为。第 4 类群聚模式从下午 15:00 持续到 16:15,发生地则在以王府井–东单,西单,中关村,枫林绿洲,交道口为代表的 14 个街道级中心。第 5 类群聚模式则是从晚上 18:25 持续到 19:05,这类群聚模式主要对应着居民下班回家或休闲娱乐的集聚行为,其群聚密度相对较低,发生地则主要以朝阳大悦城,西单,三里屯–团结湖为代表娱乐休闲中心和南湖小区,万柳,金泉广场等住宅社区为主,共包含 14 个街道级中心。

街道级中心非工作日的群聚模式则比较简单,共包含 3 类(表 7)。第 1 类模式在中午 12:20 到

12:55 之间,其平均群聚密度为 51.24 人/天。该类模式主要以城市居民的午间餐饮行为为集聚,其发生地主要集中在娱乐休闲中心以及部分住宅区内,具体包括西单,中关村购物中心,崇文门,枫林绿洲,阜通东大街等地。第 2 类群聚模式从中午 14:35 持续到 15:15,该群聚模式主要对应娱乐休闲行为,其发生地包括西单,CBD,朝阳大悦城,王府井东单,西什库大街等地。第 3 类群聚模式则在晚间 18:15 到 18:45,其平均群聚密度不到 30 人/天,主要以居民回家,或者晚间休闲行为为主,发生地主要包括朝阳大悦城,三里屯–团结湖,东直门,朝阳园,枫林绿洲,南湖小区等街道级中心(图 11(b))。

## 5 结论与讨论

本文结合北京市五环内的 POI 数据、微博签到数据和出租车 OD 点数据,采用基于点密度的层次聚类方法以及频繁模式挖掘方法对北京市五环内不同尺度、不同社会功能的城市中心进行了识别,并对其居民出行的群聚模式进行了分析,结果表明。



注:每条蓝色横线代表的是一个中心发生的丛聚过程,横线的起点终点分别表示该丛聚过程的开始时间和结束时间,纵坐标则对应着不同城市中心发生的丛聚过程编号(一个中心区可能发生多次丛聚过程)。淡蓝色矩形区域表示多个中心同时发生丛聚过程的群聚模式。

图 11 街道级中心的频繁群聚模式特征

Fig. 11 Residential cluster patterns in community centers

表 7 街道级中心的群聚模式(非工作日)

Tab. 7 Residential cluster pattern of community centers (weekend)

模式序号	起始时刻	结束时刻	平均群聚密度/人	所在中心
1	12:20	12:55	51.24	西单,中关村购物中心,崇文门,枫林绿洲,王府井–东单,永安里,CBD,学院路,北航,科春社区,阜通东大街
2	14:35	15:15	37.98	西单,CBD,中关村购物中心,科春社区,朝阳大悦城,朝阳区政府,学院路,王府井–东单,海淀区政府,西什库大街,万柳
3	18:15	18:45	27.03	朝阳大悦城,中关村购物中心,三里屯–团结湖,西什库大街,东直门,朝阳园,王府井–东单,枫林绿洲,永安里,白石桥南,CBD,南湖小区

(1)根据兴趣点上居民活跃程度分布,可以在北京市五环内识别出不同尺度的中心特征,包括4个服务全市的市级中心、服务各个区域的16个区级中心以及服务小区的45个街道级中心。

(2)根据居民活跃度时序变化特征,北京市各级中心可以分为具有不同社会功能的城市中心。市级中心主要包括海淀科教中心,西单休闲中心,北京文化中心和国贸金融中心四大中心;区级中心包括以工体-三里屯为代表的文娱中心,以王府井-东单为代表的商业中心,以中央民族大学,五道口为代表的教育中心以及以北京西站、北京南站为代表的交通中心。街道级中心大致可以分为九类,除了商业中心、文娱中心、教育中心和交通中心外,还包括天安门政治中心,以万柳、朝阳园为代表的住宅中心,以朝阳区政府、海淀区政府代表的行政中心,以及现代化集居住、工作一体的商住中心。

(3)通过居民的活跃时间分布,可以发现城市多尺度中心的居民群聚模式。其中区级中心由于空间范围较大,居民聚集时段相对较分散,街道级中心群聚模式则更为密集。区级中心发现的群聚模式在工作日包括4类,主要以早,中,晚3个时段的聚集行为为主,其持续时间较长,群聚密度较大;非工作日的群聚模式则较分散,持续时间普遍较短,群聚密度相对较小,主要以休闲娱乐行为为主要聚集目的。街道级中心的群聚模式在工作日有5类,以早午晚高峰的工作/回家出行以及吃饭行为为主要聚集目的。非工作日则群聚程度较弱,只有3类持续时间很短的群聚模式,分别在中午,下午和晚间。

城市移动大数据可以揭示城市内部的多尺度中心结构特征,解析城市系统复杂空间结构及其与居民活动的关联关系,有助于辅助城市规划,引导城市发展。然而,受数据限制,当前数据样本的代表性以及时空精度还存在一定的局限,现有数据采样密度不均。一方面,时空上的采样密度差异,可能造成统计结果在不同时间段或者空间区域内存在有偏性;另一方面,使用出租车或微博的用户来代表整个城市居民出行活动时也存在一定的有偏性,使得研究结果还存在一定程度的不确定性。未来需要进一步对大数据分析方法展开研究,结合更精细的时空大数据,对城市空间结构以及动态演化过程进行深入解析。

#### 参考文献(References):

[1] Krugman P R. The self-organizing economy[M]. Black-

well Oxford: Blackwell Publishers, 1996.

- [2] Anas A, I K. General equilibrium models of polycentric urban land use with endogenous congestion and job agglomeration[J]. *Journal of Urban Economics*, 1996,40: 232-256.
- [3] Anas A, Arnott R. Small K A. Urban spatial structure[J]. *Journal of Economic Literature*, 1996,36:1426-1464.
- [4] 韦亚平,赵民.都市区空间结构与绩效——多中心网络结构的解释与应用分析[J].*城市规划*,2006,30:9-16. [ Wei Y P, Zhao M. Spatial structure and performance of metropolis: Interpretation and application of polycentric structure[J]. *City Planning Review*, 2006,30:9-16. ]
- [5] Fujita M, Ogawa H. Multiple equilibria and structural transition of non-monocentric urban configurations[J]. *Regional Science and Urban Economics*, 1982,12:161-196.
- [6] 吴文钰,马西亚.多中心城市人口模型及模拟:以上海为例[J].*现代城市研究*,2006(12):39-44. [ Wu W Y, Ma X Y. Polycentric population density functions and modeling: The case of Shanghai[J]. *Modern Urban Research*, 2006 (12):39-44. ]
- [7] 秦贤宏,魏也华,陈雯,等.南京都市区人口空间扩张与多中心化[J].*地理研究*,2013,32:711-719. [ Qin X H, Wei Y H, Chen W, et al. Population expansion and polycentric development of Nanjing city in a period of hyper-growth [J]. *Geographical Research*, 2013,32(4):711-719. ]
- [8] 刘安国,杨开忠.克鲁格曼的多中心城市空间自组织模型评析[J].*地理科学*,2001,21(4):315-322. [ Liu A G, Yang K Z. Comments on Krugman's Edge City MO/Del[J]. *Scientia Geographica Sinica*, 2001,21(4):315-322. ]
- [9] 沈宏婷,张京祥,陈眉舞.中国大城市空间的“多中心”重组[J].*城市问题*,2005(4):25-30. [ Shen H T, Zhang J X, Chen M W. On the multicentered restructure of metropolitan space in China[J]. *Urban Problems*, 2005(4):25-30 ]
- [10] 周春山.城市空间结构与形态[M].北京:科学出版社, 2007. [ Zhou C S. Urban spatial structure and morphology [M]. Beijing: Science Publishing, 2007. ]
- [11] Champion A G. A changing demographic regime and evolving polycentric urban regions: Consequences for the size, composition and distribution of city populations[J]. *Urban Studies*, 2001,38:657-677.
- [12] McMillen D P. Polycentric urban structure: The case of Milwaukee[J]. *Economic Perspectives- Federal Reserve Bank of Chicago*, 2001,25:15-27.
- [13] Meijers E J, Burger M J. Spatial structure and productivity in US metropolitan areas[J]. *Environment and Planning A*, 2010,42:1383-1402.
- [14] 冯健,周一星.中国城市内部空间结构研究进展与展望[J].*地理科学进展*,2003,22(3):304-315. [ Feng J, Zhou Y

- X. A review and prospect on urban internal spatial structure research in China[J]. *Progress in Geography*, 2003,22(3):304-315. ]
- [15] 孙斌栋,石巍,宁越敏.上海市多中心城市结构的实证检验与战略思考[J].*城市规划学刊*,2010(1):58-63. [ Sun B D, Shi W, Ning Y M. An empirical study on the polycentric urban structure of Shanghai and strategies in future [J]. *Urban Planning Forum*, 2010(1):58-63. ]
- [16] 张纯,易成栋,宋彦.北京市职住空间关系特征及变化研究:基于第五和六次人口普查和2001和2008年经济普查数据的实证分析[J].*城市规划*,2016,40(10):59-64. [ Zhang C, Yi C D, Song Y. Characteristics of job-housing spatial relationship and changes in Beijing: An empirical study based on data from the 5<sup>th</sup>, 6<sup>th</sup> population census and economy census in 2001 and 2008[J]. *City Planning Review*, 2016,40(10):59-64 ]
- [17] Batty M. "Smart cities, big data" [J]. *Environment and Planning B: Planning and Design*, 2012,39:191-193.
- [18] Chen C, Zhang D, Castro P S, et al. Real-time detection of anomalous taxi trajectories from GPS traces[C]. *Mobile and Ubiquitous Systems: Computing, Networking, and Services*, Berlin: Springer, 2012:63-74.
- [19] Fang Z X, Shaw S L, Tu W, et al. Spatiotemporal analysis of critical transportation links based on time geographic concepts: A case study of critical bridges in Wuhan, China [J]. *Journal of Transport Geography*, 2012,23:44-59.
- [20] Sun L, Zhang D Q, Chen C, et al. Real Time anomalous trajectory detection and analysis[J]. *Mobile Networks & Applications*, 2013,18:341-356.
- [21] Zheng K, Zheng Y, Yuan N J, et. al. On Discovery of Gathering Patterns from Trajectories[C]. *IEEE 29<sup>th</sup> International Conference on Data Engineering (ICDE)*, 2013:242-253.
- [22] Roth C, Kang S M, Batty M, et al. Structure of urban movements: Polycentric activity and entangled hierarchical flows[J]. *Plos One*, 2011,6(1):1-9.
- [23] Zhong C, Huang X F, Arisona S M, et al. Inferring building functions from a probabilistic model using public transportation data[J]. *Computers Environment and Urban Systems*, 2014,48:124-137.
- [24] 龙瀛,张宇,崔承印.利用公交刷卡数据分析北京职住关系和通勤出行. *地理学报*,2012,67(10):1339-1352. [ Long Y, Zhang Y, Cui C Y. Identifying commuting pattern of Beijing using bus smart card data[J]. *Acta Geographica Sinica*, 2012,67(10):1339-1352. ]
- [25] Ratti C, Frenchman D, Pulselli R M, et al. Mobile landscapes: Using location data from cell phones for urban analysis[J]. *Environment and Planning B-Planning & Design*, 2006,33:727-748.
- [26] Calabrese F, Colonna M, Lovisolo P, et al. Real-time urban monitoring using cell phones: A case study in rome [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2011,12:141-151.
- [27] Pei T, Sobolevsky S, Ratti C, et al. A new insight into land use classification based on aggregated mobile phone data[J]. *International Journal of Geographical Information Science*, 2014,28:1988-2007.
- [28] Birkin M, Malleon N. Microscopic simulations of complex metropolitan dynamics[C]. *Amsterdam: Complex City workshop*, 2011.
- [29] Wakamiya S, Lee R, Sumiya K. Urban area characterization based on semantics of crowd activities in twitter[J]. *Geospatial Semantics*, 2011,6631:108-123.
- [30] Hollenstein L, Purves R S. Exploring place through user-generated content: Using Flickr tags to describe city cores [J]. *Journal of Spatial Information Science*, 2010,1:21-48.
- [31] Cranshaw J, Schwartz R, Hong J I, et al. The Livehoods Project: Utilizing social media to understand the dynamics of a city[C]. *Dublin: International AAAI Conference on Weblogs and Social Media*, 2012:58-65.
- [32] Liu Y, Wang F H, Xiao Y, et al. Urban land uses and traffic 'source-sink areas': Evidence from GPS-enabled taxi data in Shanghai[J]. *Landscape and Urban Planning*, 2012,106:73-87.
- [33] Guo D S, Zhu X, Jin H, et al. Discovering spatial patterns in origin-destination mobility data[J]. *Transactions in GIS*, 2012,16:411-429.