



# 使用时序出租车轨迹识别多层次城市功能结构

姚尧<sup>1</sup> 张亚涛<sup>2</sup> 关庆锋<sup>1</sup> 麦可<sup>2</sup> 张金宝<sup>3</sup>

1 中国地质大学(武汉)地理与信息工程学院,湖北 武汉,430074

2 武汉大学测绘遥感信息工程国家重点实验室,湖北 武汉,430079

3 中山大学地理科学与规划学院,广东 广州,510275

**摘要:**地理时空大数据被广泛用于城市功能结构识别,其中功能层次性的研究对于系统理解城市功能的结构特征和分布形态具有重大意义,但相关研究仍处于空缺状态。基于时序出租车出行数据和感兴趣点数据描述居民出行模式,结合动态时间规整和 K-MEDOIDS 聚类算法识别城市的功能属性和空间结构。研究结果表明,广州市中心城区的城市功能具有明显的层次性。随着层次细致程度的提升,其功能属性由“职-住”二元结构向“职-住-休”三元结构发展;其空间结构符合环状圈层构造,功能由外围的居住游憩向中心的商业休闲过渡,并在不同的圈层上呈现出各自的功能倾向。这为城市规划人员系统理解城市功能的属性变化和结构特征提供了有效的参考价值。

**关键词:**城市功能结构;时序出租车轨迹数据;层次性;环状圈层结构;社会感知

**中图分类号:**P208

**文献标志码:**A

随着遥感和地理信息技术的发展,基于遥感影像数据的场景分类和基于地理时空大数据的社会感知被广泛应用于城市功能识别和居民行为模式分析中<sup>[1-2]</sup>。其中,场景分类方法是一种面向对象的地物识别方法,基于遥感图像的光谱、形状和纹理等特征,结合概率语义模型,能够有效识别城市地物<sup>[1]</sup>。然而,这种方法对于遥感影像的自然属性依赖性强,而城市功能不仅和地物的自然物理属性相关,在很大程度上,更是受到人群活动与社会经济因素的影响<sup>[3-4]</sup>。地理空间大数据的兴起,使得大量能体现人群活动特征的数据被应用于城市功能结构的识别中<sup>[5-8]</sup>,如百度感兴趣点(point of interests, POIs)数据、地铁公交刷卡数据和出租车轨迹数据等<sup>[6, 8-9]</sup>。其中,出租车是城市生活的重要传感器,为城市居民行为模式研究和土地利用分类研究提供了强有力的工具<sup>[10-13]</sup>;另外,也有学者基于 POIs 的频率分布对城市功能区进行识别<sup>[10-14]</sup>。

出租车轨迹数据对城市功能区识别和土地利用分类作用显著,但之前的研究对于其时序特征

的挖掘还不够充分,且缺乏对城市功能层次性的研究。如 Pan 等利用出租车轨迹数据提取的时间序列对城市功能区进行分类<sup>[15]</sup>,该模型虽然效果显著,但分类算法在度量特征之间的相似性时未考虑时间序列的特殊性。此外,城市功能具有明显的层次性,不同层次的城市功能具有较为明显的层级关系<sup>[16]</sup>,从系统功能上探讨这种层次性有助于从整体到细部逐步把握城市空间结构的功能变化和分布,但目前针对城市功能层次性的研究还较少。因此,本文引入动态时间规整描述出租车轨迹数据之间的相似性,以聚类算法挖掘其时序特征,并在广州市中心城区开展不同层次下的功能结构研究。本文模型首先对时序出租车出行数据构建出发-到达位置(orientation-destination, OD)的时序密度曲线;然后,基于动态时间规整对研究区进行聚类分析,并采用 Silhouette 值评估多层次对象;最后,结合 OD 时序曲线和 POIs 信息,对不同类别的功能区进行识别分析,在多层次的视角下讨论城市功能的功能属性和空间结构特征。

收稿日期:2018-02-08

项目资助:国家自然科学基金(41801306,41671408);武汉大学测绘遥感信息工程国家重点实验室开放基金(18S01)。

第一作者:姚尧,博士,副教授,主要从事多源时空大数据分析和精细城市模拟。yaoy@cug.edu.cn

通讯作者:张亚涛,硕士。yatau@foxmail.com

## 1 研究区和数据

### 1.1 研究区

研究区位于广州市中心城区,包括荔湾、越秀、天河和海珠4个区(见图1),是珠三角经济区的经济、政治、文化中心。据广州市2015年统计年鉴,研究区总面积为279.63 km<sup>2</sup>,约占广州市区面积的7.28%。2014年研究区内常住人口为514.38万人,约占广州市区人数的45.16%。

广州市的城市空间结构十分复杂<sup>[17]</sup>,中心城区的城市功能分布彼此交错,比如,居民区、商业区等功能区的分布虽然在总体上呈现出一定的规律性,但很多区域的功能并不单一,而是由多种功能相互作用而形成的混合功能区域。这种混合功能区域的出现吻合《马丘比丘宪章》对城市分区综合性和多功能的追求<sup>[18]</sup>,本研究也将从多层次的角度探究广州市城市土地混合利用的现状。

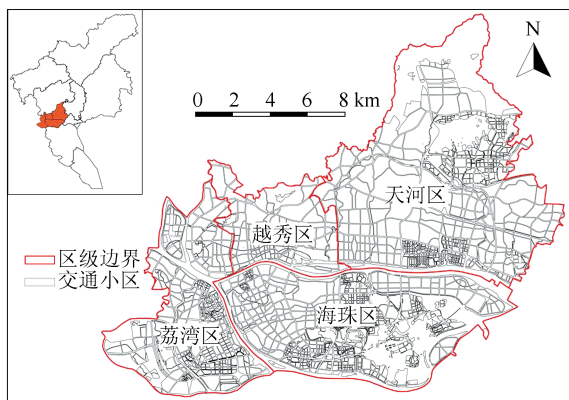


图1 广州市中心城区图

Fig. 1 Central Urban Area Map of Guangzhou

### 1.2 时序出租车轨迹数据

出租车轨迹数据源自广州市交通委员会,为了使研究结果能够反映工作日和休息日的居民出行情况,本文选取2009-05-12(星期二)和2009-05-16(星期六)的轨迹数据进行分析,共包括由14 379辆出租车产生的48 957 065条记录。其中,每条记录包含车辆的唯一ID号、入库时间、GPS信号时间、经纬度、速度、有效标示以及车辆状态,其中车辆状态包括空车、重车、点火、熄火等,GPS发送时间的间隔与车辆的状态有关,间隔4~7 s不等。相关研究表明<sup>[10]</sup>,交通分析小区(traffic analysis zone, TAZ)中具有相同或相似的功能及特征,在城市功能结构的研究中以其作为分析单元是合理的。

## 2 多层次城市功能结构识别

首先从原始的出租车轨迹数据中提取出每一条轨迹的O点和D点,并利用点密度分析以小时为间隔生成工作日和休息日共48 h的O点密度分布和D点密度分布图,之后使用基于动态时间规整的K-MEDOIDS算法对OD点时序密度向量进行聚类分析,在Silhouette值的评估下获取研究区域的多层次聚类对象,同时引入POIs的统计结果进行辅助分析,具体流程如图2所示。

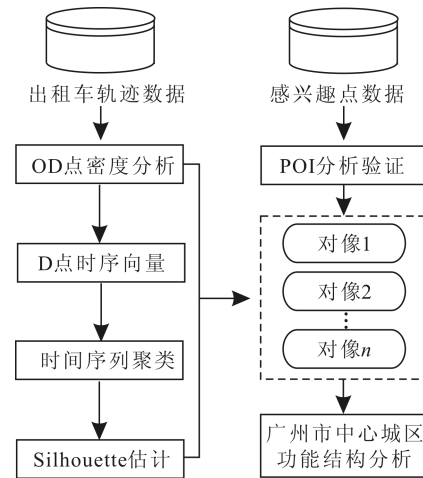


图2 城市功能结构识别流程

Fig. 2 Workflow of Urban Functional Structure Identification

### 2.1 特征数据集处理

在原始的出租车轨迹数据中,根据出租车的唯一标识符和轨迹时间顺序提取出每一条轨迹的OD点。对于任意的轨迹数据 $Tr$ ,按轨迹点生成的先后顺序有轨迹向量 $\mathbf{T} = [t_1 \cdots t_i \cdots t_j \cdots t_n]$ ,提取规则为:若 $Tr$ 中 $t_i$ 时刻为空车,而 $t_{i+1}$ 状态为重车,则 $t_{i+1}$ 时刻对应的地理位置为起点;时间后推,遇到的第一个状态为空车的时刻 $t_j$ ,则 $t_j$ 时刻对应的地理位置为轨迹终点。由此,得到每一条轨迹的O点和D点,经处理后,有效数据记录为6 687 394条。另外,由于出租车O点和D点主要分布在道路上,而居民在出行时其实际出发点和目的点多分布于道路两侧的居民区、商业区等区域,估计该距离小于1 km,可通过搜索半径为1 km的点密度分析获取TAZ小区内的O、D点密度值。以1 h作为时间间隔,最终生成每个TAZ工作日和休息日共48 h的O点密度变化曲线和D点密度变化曲线。Pan等在利用出租车轨迹上下车点进行土地利用分类时,构造了6种OD点的时序组成方式<sup>[15]</sup>:

1)  $[o_w^1 \cdots o_w^i \cdots o_w^{24}, o_r^1 \cdots o_r^i \cdots o_r^{24}]$  为 48 维 O 点向量,  $o_w^i$ 、 $o_r^i$  分别是工作日、休息日第  $i$  小时的 O 点平均数量。

2)  $[d_w^1 \cdots d_w^i \cdots d_w^{24}, d_r^1 \cdots d_r^i \cdots d_r^{24}]$  为 48 维 D 点向量,  $d_w^i$ 、 $d_r^i$  分别是工作日、休息日第  $i$  小时的 D 点平均数量。

3)  $[m_w^1 \cdots m_w^i \cdots m_w^{24}, m_r^1 \cdots m_r^i \cdots m_r^{24}]$  为 48 维 O 点和 D 点的差值向量, 其中  $m_w^i = o_w^i - d_w^i$ ,  $m_r^i = o_r^i - d_r^i$ 。

4)  $[r_w^1 \cdots r_w^i \cdots r_w^{24}, r_r^1 \cdots r_r^i \cdots r_r^{24}]$  为 48 维 O 点和 D 点的比值向量,  $r_w^i = o_w^i / d_w^i$ ,  $r_r^i = o_r^i / d_r^i$ 。

5)  $[o_w^1 \cdots o_w^i \cdots o_w^7]$  为 7 维向量,  $o_w^i$  是一周第  $i$  天 O 点数量变化。

6)  $[d_w^1 \cdots d_w^i \cdots d_w^7]$  为 7 维向量,  $d_w^i$  是一周第  $i$  天 D 点数量变化。

在对上述 6 种特征进行评估时, 48 维 O 点向量的判别能力最高; 组合多种特征评估时, 96 维 OD 点组合向量判别能力最优异, 即  $[o_w^1 \cdots o_w^{24}, o_r^1 \cdots o_r^{24}, d_w^1 \cdots d_w^{24}, d_r^1 \cdots d_r^{24}]$ 。因此, 本研究以 96 维 O、D 点时序密度向量作为城市功能区识别的输入向量进行聚类分析。

## 2.2 动态时间规整的 K-MEDOIDS 聚类

在进行时序数据分析时, 需要度量时间序列的相似性, 也即判断不同的时间序列是否表现为相似的行为<sup>[19]</sup>。目前, 时间序列聚类模型在测度相似性时根据目的不同可分为 3 类, 分别依据序列在时间、形状和变化上的相似性进行聚类分析<sup>[20]</sup>。出租车轨迹在分析城市居民的出行模式时, 着重点在于识别曲线在形状上的变化模式, 如以居住功能为主导的区域, 早高峰阶段的 O 点处于高峰、D 点处于低谷, 而以工作功能为主导的区域, 早高峰阶段的 O 点处于低谷、D 点处于高峰。动态时间规整(dynamic time warping, DTW)是基于形状相似性的时间序列度量算法, 可以实现时间序列的最佳匹配对齐, 挖掘其中灵活的相似性和差异性, 是度量时间序列相似性的最普遍使用的定量方法<sup>[19-21]</sup>。

给定时间序列  $\mathbf{P} = [p_1 \cdots p_i \cdots p_n]$  和  $\mathbf{Q} = [q_1 \cdots q_j \cdots q_m]$ , 构建  $n \times m$  的矩阵  $\mathbf{M}$ , 其元素  $(i_{th}, j_{th})$  是点  $p_i$  和点  $q_j$  之间的距离。DTW 需保证对齐  $\mathbf{P}$  和  $\mathbf{Q}$  时差异最小, 构建路径  $\mathbf{W} = [\omega_1 \cdots \omega_k \cdots \omega_K]$ , 其中  $\max(m, n) \leq K \leq m+n-1$ , 且需要满足 3 个条件: 边界条件、连续性和单调性。边界条件需保证起点和终点分别是矩阵中左下角的点  $(1, 1)$  和右上角的点  $(n, m)$ 。连续性和单调性限制  $\omega_k$  的下一个点  $\omega_{k+1}$  只能在  $\omega_k$  的上方或者右方, 不能跨越两个格网, 也不能向下方

或者左方移动。在所有满足以上 3 个条件的路径中, 选择  $d_{DTW}$  最小的一条作为输出结果, 也即度量  $\mathbf{P}$  和  $\mathbf{Q}$  中距离最小的路径, 其中  $d_{DTW}$  为:

$$d_{DTW} = \min \frac{\sum_{k=1}^K \omega_k}{K} \quad (1)$$

本研究中, DTW 在衡量交通小区 96 维 OD 点时间序列的差异性时, 可消除地块时间变化趋势中的移位等现象, 但这种移位不是毫无限制的, 故实际度量两地块的相似度时需加以局部限制。针对 96 维时间序列  $[o_w^1 \cdots o_w^{24}, o_r^1 \cdots o_r^{24}, d_w^1 \cdots d_w^{24}, d_r^1 \cdots d_r^{24}]$ , 首先施加第一重限制, 在衡量任意两个时间序列时,  $(o_w^1, o_w^1)$ 、 $(o_w^{24}, o_w^{24})$ 、 $(o_r^1, o_r^1)$ 、 $(o_r^{24}, o_r^{24})$ 、 $(d_w^1, d_w^1)$ 、 $(d_w^{24}, d_w^{24})$ 、 $(d_r^1, d_r^1)$ 、 $(d_r^{24}, d_r^{24})$  这 8 个节点在 DTW 路径中必须匹配, 保证工作日和休息日的 O、D 点匹配至对应的时间范围; 其次施加第二重限制, 在上述 8 个节点中任意相邻的两个节点, 如  $(o_w^1, o_w^1)$  和  $(o_w^{24}, o_w^{24})$ , 需保证每隔 6 h 的时间点应得以匹配, 也即在一天的时间序列中, 06:00、12:00、18:00、24:00 必须与对应的时间点匹配, 以此限制移位的范围。

以 DTW 衡量任意两对时序曲线的差异后, 本文使用 K-MEDOIDS 算法对 TAZ 进行聚类分析, 同时为了提升算法效率, 使用 AP (anomalous pattern) 方法帮助确定初始化全局聚类中心<sup>[22]</sup>。另外, 为了评估不同聚类数目结果的可靠性, 本文引入 Silhouette 评估各簇的聚类质量<sup>[23]</sup>, 在 K-MEDOIDS 算法的背景下, 假定在一个已有的聚类结果中,  $a(i)$  指代样本点  $i$  与同类中其他所有样本点的 DTW 距离的中位值,  $b(i)$  指代其他类与样本点  $i$  之间的最小 DTW 距离的中位值, 则有:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (2)$$

如果  $s(i)$  接近于 1, 意味着样本点  $i$  和现有的聚类结果十分匹配; 如果  $s(i)$  接近于 -1, 则意味着样本点  $i$  应该隶属于其邻类中。所有点的  $s(i)$  的均值越高, 则代表聚类结果越优。

## 2.3 POIs 数据辅助分析

POIs 点包含区域功能的语义信息, 在城市功能识别中发挥着重要的作用<sup>[10]</sup>, 本研究使用 POIs 作为辅助信息帮助识别广州市中心城区的 TAZ 功能属性。POIs 点的密度反映了该功能区不同种类感兴趣点的绝对量值差异, 可用于辅助判断功能区的实际功能属性; 但 POIs 的绝对量值也可能掩盖区域内具备实际主导性的属性信息, 为避免 TAZ 中不同类型 POIs 的数量差异对识别结果的影响, 同时引入 POIs 富集指数作为辅助

判别信息<sup>[24]</sup>。

POIs 点的密度表示为:

$$\text{Density}_{\text{POI}(i,j)} = \frac{\text{Num}_{\text{POI}(i,j)}}{\text{Area}_j} \quad (3)$$

式中,  $\text{Density}_{\text{POI}(i,j)}$  是第  $j$  类功能区中第  $i$  类 POI 的密度;  $\text{Num}_{\text{POI}(i,j)}$  是第  $j$  类功能区中第  $i$  类 POI 的数量;  $\text{Area}_j$  是第  $j$  类功能区的总面积。

POIs 富集指数表示为:

$$F_{i,j} = \frac{n_{i,j}/n_j}{N_i/N} \quad (4)$$

式中,  $F_{i,j}$  是第  $j$  类功能区中第  $i$  类 POI 的富集指数;  $n_{i,j}$  是第  $j$  类功能区中第  $i$  类 POI 的数量;  $n_j$  是第  $j$  类功能区中 POIs 的总数量;  $N_i$  是第  $i$  类 POI 的总数量;  $N$  是整个研究区中所有 POIs 的总数量。  $F_{i,j}$  越高, 表明第  $j$  类功能区中第  $i$  类 POI 的富集程度越高。

### 3 结果和讨论

本研究数据处理和算法实现均采用 C++ 编程实现, 使用 GDAL、Shark 等开源库, 在程序执行过程中通过启用 OpenMP 实现高速并行运算。

为了研究广州市城市功能结构的层次性, 本研究针对 96 维 OD 点向量进行多次聚类, 并从中选取 Silhouette 值较高的聚类结果进行分析。如图 3 所示, 聚类数目分别设为  $k=2\sim 10$ , Silhouette 值都大于 0.5, 具备较好的解释性, 从中选择较具代表性的 3 个层次进行详细分析。Silhouette 值可划分为 3 段:  $k=2$  时, 聚类结果最为理想, Silhouette 值高达 0.87;  $k=3\sim 4$  时, 聚类结果次之, Silhouette 值约为 0.76;  $k=5\sim 10$  时, 聚类结果再次之, Silhouette 值分布在 0.51~0.62 之间。为了使分析结果更具代表性, 本研究从 3 段中各选择 Silhouette 值最高的聚类结果进行分析, 即  $k=2, 3, 7$ 。

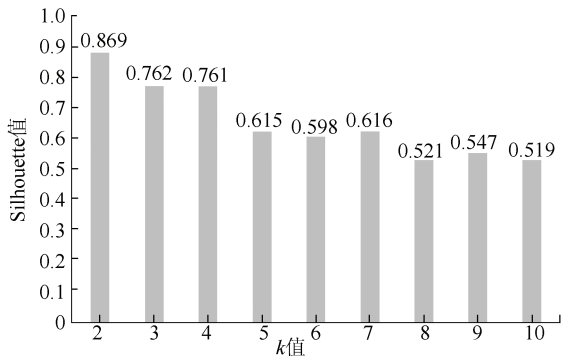


图 3 不同聚类数目  $k$  值对应的 Silhouette 估计值

Fig. 3 Silhouette Values Corresponding to Various Clustering Numbers  $k$

当聚类数目  $k=2, 3$  时, 分别如图 4、图 5 所示。图 4 中 A1 位于研究区域的中心地带, A0 呈环状围绕在 A1 四周, 构成简单的环状圈层结构。

图 5 中其聚类结果继承了  $k=2$  时的环状圈层结构, B0 区域几乎包含于 A0 的范围内, B2 区域几乎包含于 A1 的范围内, B1 区域由 A0、A1 相接部分构成。

图 6 为 A0 和 A1 的 OD 点时序变化曲线, 在工作日: A0 区域在上午时段的 O 点密度高于同期 D 点密度, 晚上时段的 D 点密度明显高于同期 O 点密度, 属于典型的居住功能区; A1 区域在 10:00、15:00、22:00 存在明显的峰值, 居民的流通高峰较为集中, 并且 17:00—24:00 时, O 点密度略高于同期的 D 点密度, 符合工作功能区的居民行为模式。在休息日: A0 区域 08:00 的 O、D 点密度是较为明显的峰值, 符合居民休息日外出游玩的特征, 且上午时段 O 点密度持续高于 D 点密度; A1 区域的 O、D 点密度曲线不存在明显的峰值, 变化幅度较工作日显著减弱, 其工作属性特征降低, 进一步验证了 A0、A1 区域的主导功能。此时, 研究区的功能体系呈现为以居民区环绕工作区的“职-住”二元环状圈层结构。

图 7 为 B0、B1、B2 类功能区的 O、D 点时序

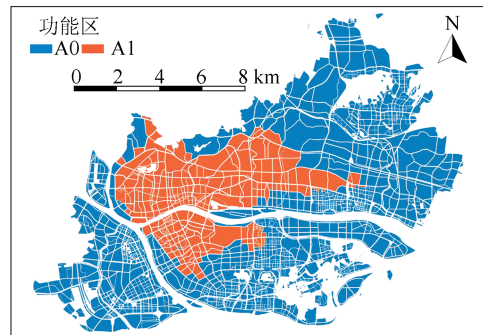


图 4  $k=2$  时广州市中心城区功能结构分布图

Fig. 4 Urban Functional Structure Map in Central Urban Area of Guangzhou in the Case of  $k=2$

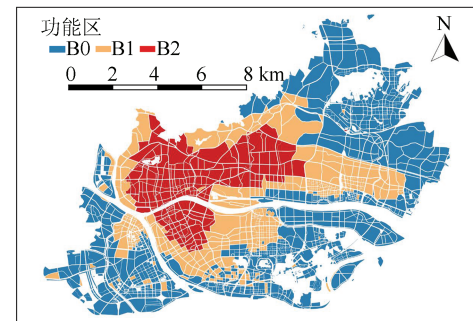


图 5  $k=3$  时广州市中心城区功能结构分布图

Fig. 5 Urban Functional Structure Map in Central Urban Area of Guangzhou in the Case of  $k=3$



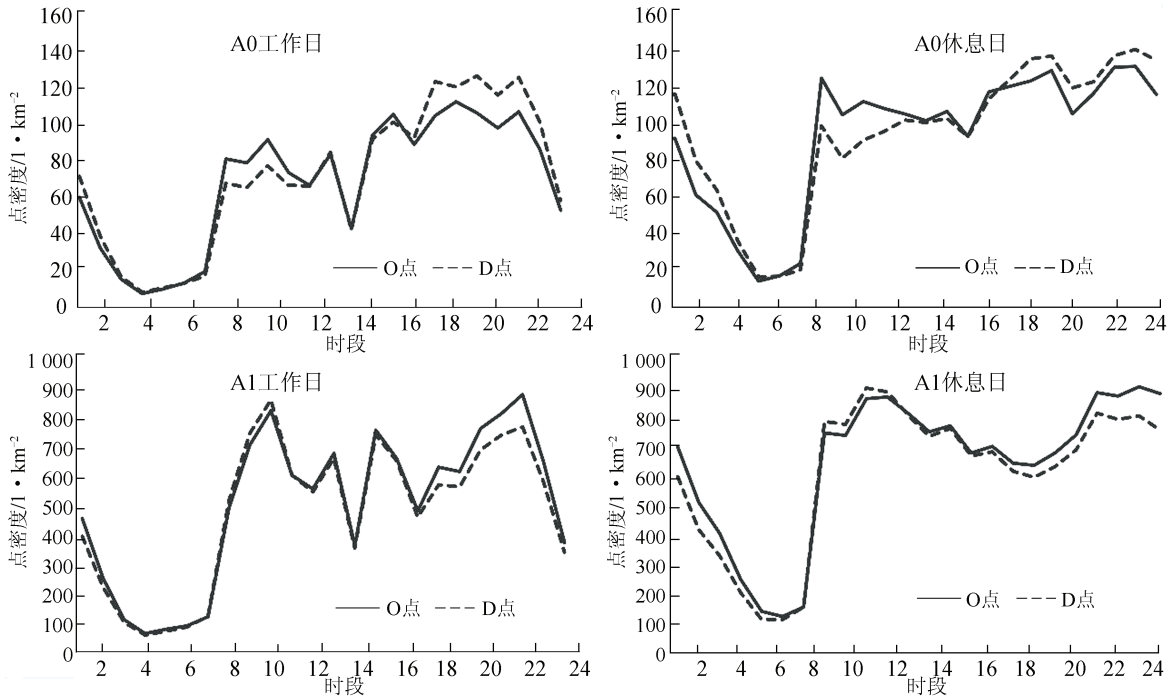


图 6  $k=2$  时广州市中心城区出租车时序曲线

Fig. 6 Taxi's Time-Series Curves in Central Urban Area of Guangzhou in the Case of  $k=2$

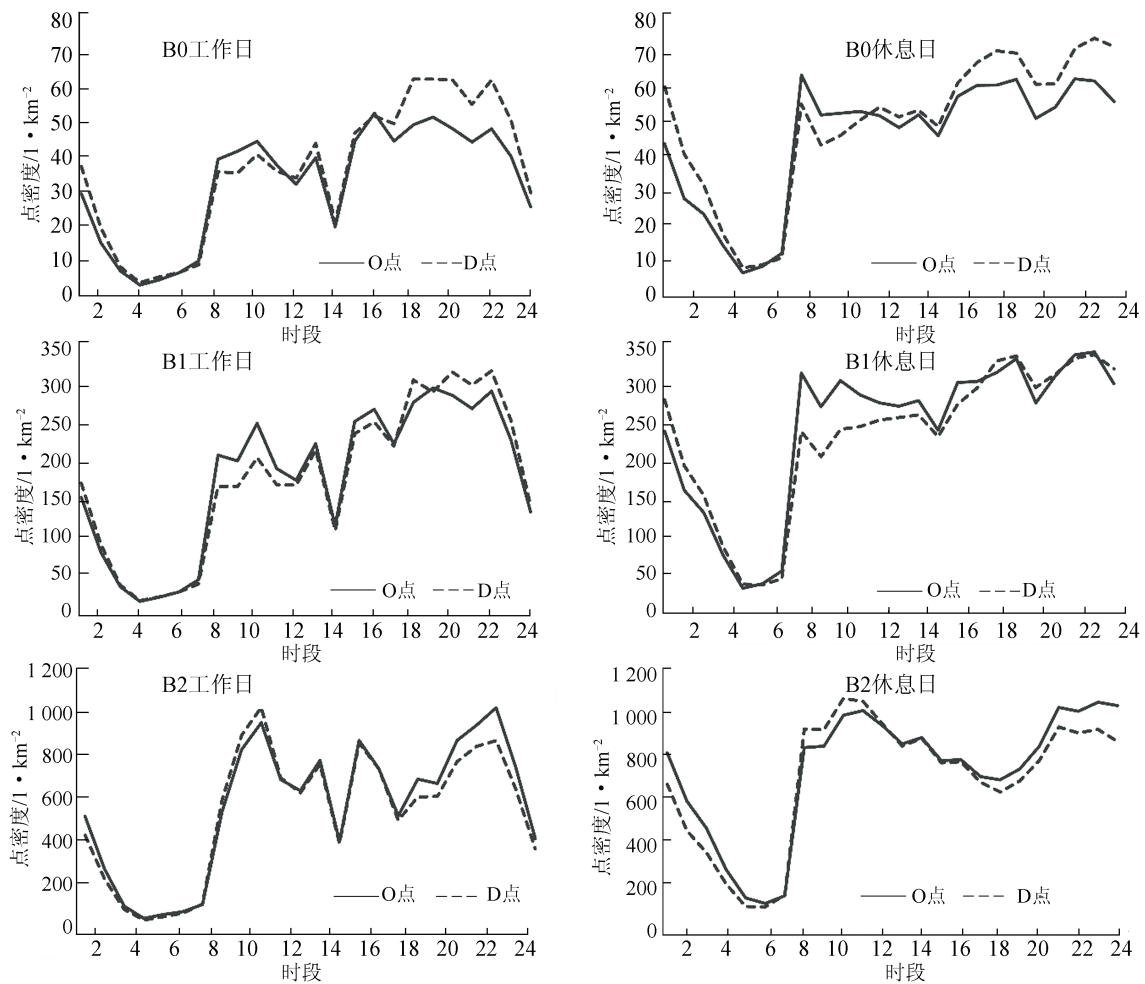


图 7  $k=3$  时广州市中心城区出租车时序曲线

Fig. 7 Taxi's Time-Series Curves in Central Urban Area of Guangzhou in the Case of  $k=3$

密度曲线, B2 区域的 O、D 点密度曲线变化模式和 A1 区域几乎一致, 但在量值上高于后者, 故 B2 是工作功能区; B0、B1 区域的 OD 点密度曲线分别在 14:00—24:00、07:00—14:00, 与 A0 区域保持一致, 推定 B0、B1 都具备居住区的功能特点。针对 B0 区域, 与 A0 相比, O、D 点密度在上午时段的变化幅度更小, 而在 17:00—24:00 时, O、D 点密度差的差异变大, 符合居住功能为主导的居民出行特点, 故 B0 是居住功能区; 针对 B1 区域, 与 A0 区域相比, 不论工作日和休息日, O、D 点密度整体较大, 且上午时段的变化幅度增大, 而下午时段的 O、D 点密度差异变小, 该曲线模式符合居民的游憩行为, 推定 B1 区域具备游憩功能, 同时 B1 区域的 O、D 点在工作日 10:00 存在明显的峰值, 符合居民的工作行为, 故最终推定 B1 区域是居住游憩工作混合功能区, 此时, 混合功能区出现, 城市功能的综合性和多元性已逐渐显现。

当聚类数目  $k=7$  时, 如图 8 所示, 由 C0~C6 组成了由外向内的环状圈层结构。从空间继承性上分析, C0 区中 96.88% 的 TAZ 属于 B0, C6 区中 98.97% 的 TAZ 属于 B2, C1 由 B0、B1 相接部分组成, C5 由 B2、B1 相接部分组成, C2、C3、C4 基本包含在 B1 内。如图 9、图 10 和表 1 所示, 考察 OD 点时序变化曲线、POIs 密度及富集指数表, C6 的曲线模式与 B2 一致, 且各类型的 POIs 密度都居于前列, 其中公司企业、金融服务及休闲娱乐的 POIs 具有高富集性, 商业配套设施十分完善, 故 C6 是中心商业区; C5 的 O、D 点曲线模式和 B1、B2 曲线的组合模式相似, 如工作日上午时段与 B1 较为一致, 晚上时段与 B2 较为一致, 同时商业性、休闲性的 POIs 数量丰富, 各类 POIs 的富集指数较为均匀, 故 C5 是商业休闲区。C0 的 O、D 曲线模式与 B0 一致, 涵盖居住区的功能属性, 同时公司企业、休闲娱乐等 POIs 类型具有很高的富集性, 故 C0 是居住游憩混合功能区; C1 区域的 O、D 曲线模式和 B0、B1 都存在差异, 而与 A0 区域的 O、D 曲线十分相似, 且该区域内地产小区的富集指数极高, 故 C1 是居住功能区。C2、C3、C4 区域主要位于 B1 内, 曲线模式和 B1 都存在相似性, 属于居住游憩工作混合区域, 在将其进一步细化为 3 种不同类型的功能区时, 三者的城市功能表现为更为复杂的特征, 每一种功能区都呈现为“职-住-休”的三元混合状态, 但其功能属性各有倾向, 如 C2 中, 生活服务的 POIs 富

集性较高, C3 中, 金融服务和休闲娱乐的 POIs 富集性较高, C4 中, 医疗教育的 POIs 富集性较高。此时, 城市功能的综合性和多元性已充分表现出来, 研究区的功能体系更为复杂, 但依然保留了  $k=3$  时以“职-住-休”为基本功能框架的特点, 同时继承了  $k=2、3$  时的环状圈层结构, 功能结构的层次性已十分明显。

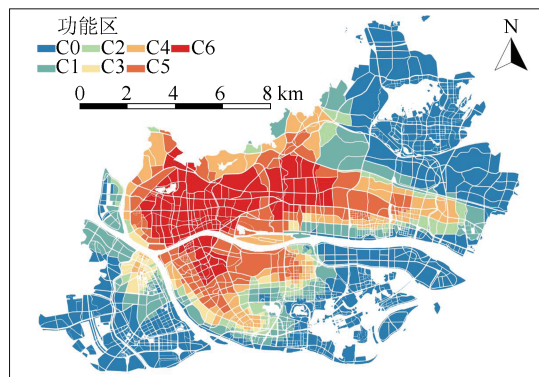
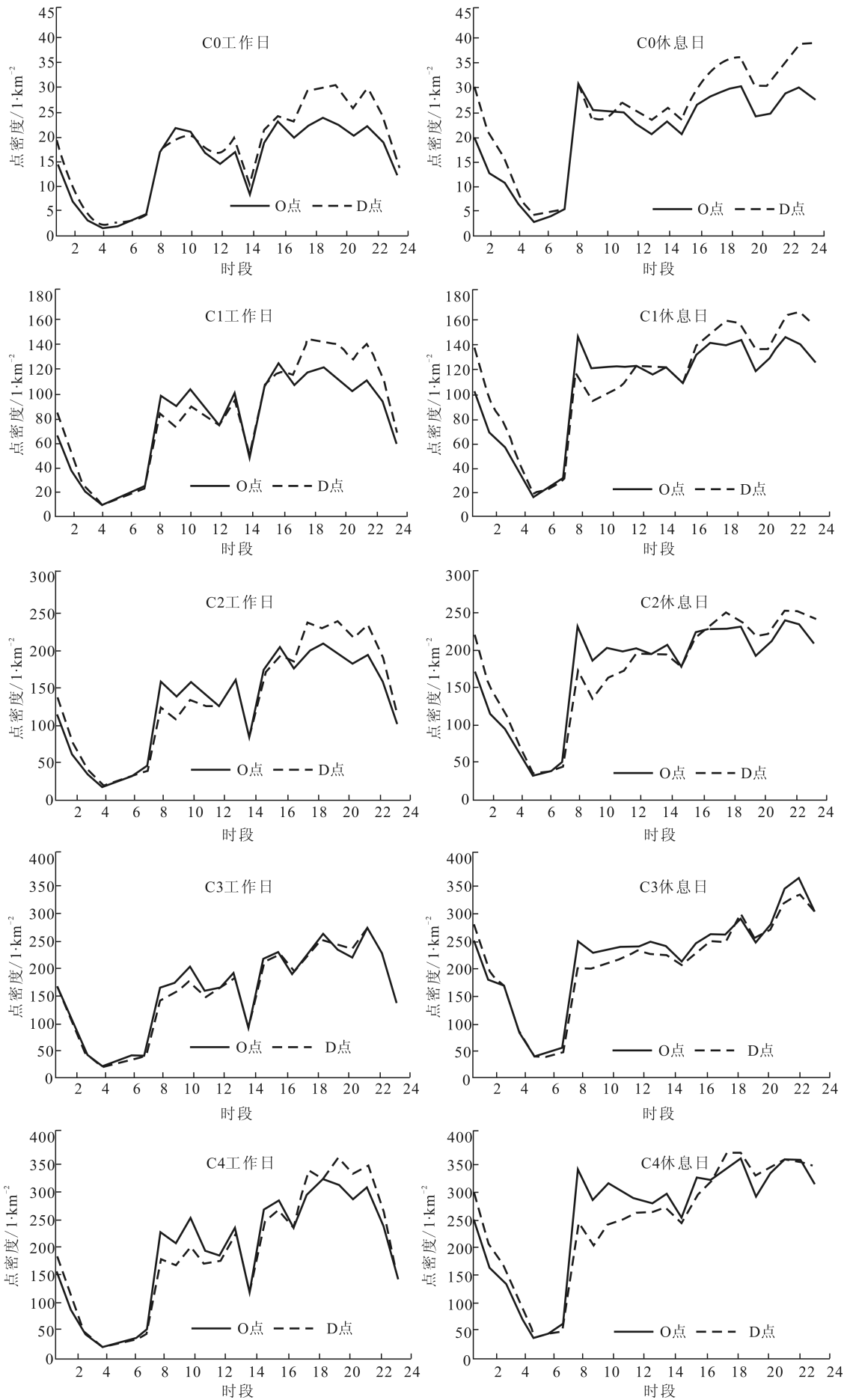


图 8  $k=7$  时广州市中心城区功能结构分布图  
Fig. 8 Urban Functional Structure Map in Central Urban Area of Guangzhou in the case of  $k=7$

总体来说, 从  $k=2$  到  $k=3$ , 再到  $k=7$ , 广州市中心城区的功能结构呈现出明显的层次性: 当  $k=2$  时, 研究区域的功能表现是“职-住”二元结构, 空间结构以环状圈层实现了中心城区的二元划分, 可为城市居民通勤提供参考; 当  $k=3$  时, 城市功能细化, 由“职-住”二元结构向“职-住-休”三元结构发展, 功能结构更为完善, 环状圈层结构的层次性提升, 由外围的居住区逐渐向中心城区的工作区过渡; 当  $k=7$  时, 城市功能在细化的同时逐渐向多元性和混合性发展, 虽然依然符合“职-住-休”的三元框架, 但土地混合利用成为城市功能的总体特征, 环状圈层结构进一步细化, 由外围的居住游憩向中心的商业休闲过渡, 混合功能区根据 POIs 的富集和主体特征表现为不同的功能属性倾向。总体来说, 本研究将城市功能和空间结构结合起来, 有效地阐述了城市功能及空间结构的多层次性特征, 这对于从系统性和层次性的特征把握城市土地利用和结构特性十分重要。

## 4 结 语

城市化进程的不断加快, 一方面造就了城市的快速扩张<sup>[25]</sup>, 另一方面导致城市内部土地利用的多样性<sup>[26]</sup>, 有效识别城市内部功能和空间结构对于城市管理和规划至关重要。本研究构建了基



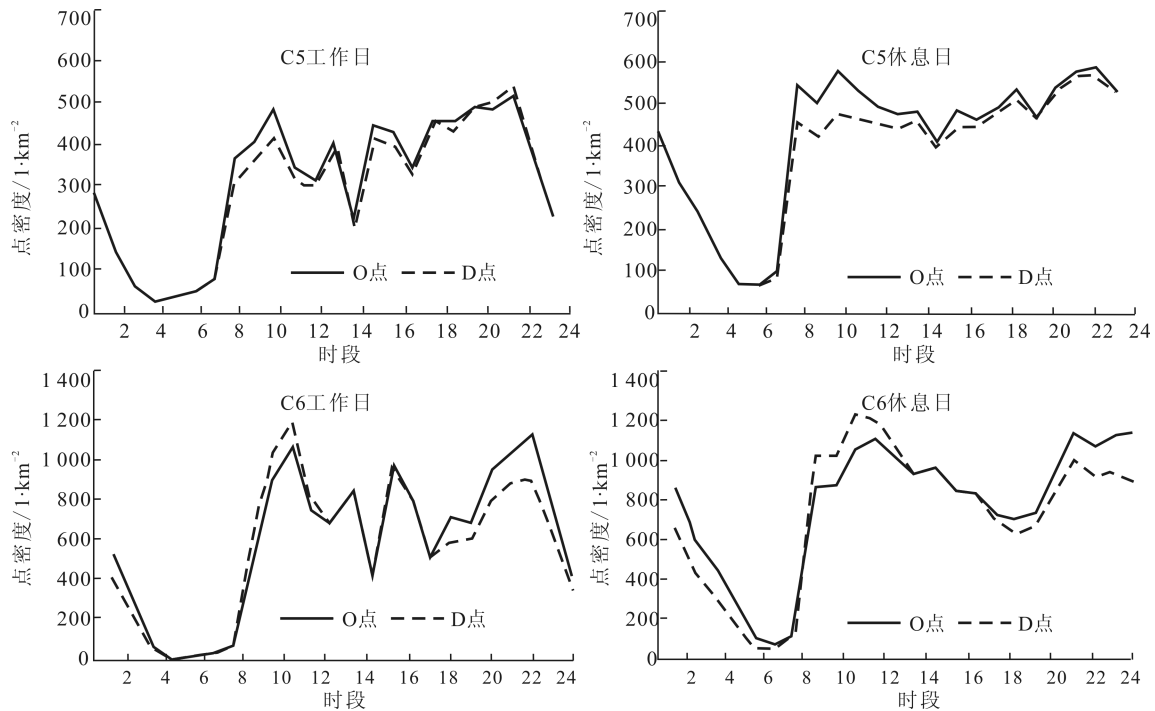
图9  $k=7$ 时广州市中心城区出租车时序曲线Fig. 9 Taxi's Time-Series Curves in Central Urban Area of Guangzhou in the Case of  $k=7$ 

表1 功能区 POIs 密度和富集指数

Tab. 1 Density and Enrichment Factor of POIs in Functional Zones

指数 类别	POIs 密度						POIs 富集指数							
	C0	C1	C2	C3	C4	C5	C6	C0	C1	C2	C3	C4	C5	C6
政府机构	6.82	27.46	36.52	40.55	36.59	61.37	72.95	0.78	1.02	0.96	0.79	1.05	1.23	0.93
公司企业	64.35	172.67	245.43	313.71	199.75	295.77	559.90	1.12	0.98	0.98	0.93	0.87	0.90	1.09
金融服务	3.85	12.24	22.48	43.10	21.99	37.19	54.53	0.69	0.71	0.92	1.31	0.98	1.16	1.09
休闲娱乐	67.36	202.93	264.09	412.41	264.38	369.40	627.63	1.01	0.98	0.91	1.05	0.99	0.96	1.04
医疗教育	13.85	42.58	55.89	63.02	65.41	87.98	109.40	1.01	1.01	0.94	0.78	1.19	1.12	0.89
生活服务	28.87	93.07	176.60	196.01	145.79	196.63	266.71	0.90	0.94	1.26	1.04	1.13	1.07	0.92
地产小区	10.85	52.82	51.07	82.95	51.32	74.73	70.30	0.95	1.51	1.03	1.24	1.13	1.15	0.69

注:颜色深度代表排序

于出租车时序轨迹数据的多层次城市功能识别模型。首先,使用点密度分析构建 TAZ 的工作日和休息日出租车时序密度曲线;其次,引入动态时间规整和 K-MEDOIDS 聚类对广州市中心城区的功能属性进行聚类分析,使用 Silhouette 值评估聚类结果的质量;最后,辅以 POIs 语义信息,探究广州市中心城区的功能特征和空间结构。从层次性上分析,其功能属性逐渐由“职-住”二元结构向“职-住-休”三元结构转变;从空间结构上分析,功能属性具备明显的环状圈层特征,随着层次的细化,该结构愈加明显;从功能变化上看,中心城区的功能属性明显存在由环状外围向中心过渡的特征,当层次低时,由居住向工作过渡,当层次稍高时,过渡之中增添了游憩功能,当层次再次提高时,功能属性过渡逐渐复杂,并在不同的圈层上呈现出不同的属性倾向。

总体来说,本研究为城市管理和规划提供了多层次分析的视角,对于城市内部土地利用模式具有重要的意义。此外,本研究尚有不足,如对于城市功能的层次性和环状圈层结构论证还不够充分,出租车作为城市交通流的重要构成对城市交通模式具有很强的代表性<sup>[10]</sup>,但公交车、地铁、自驾车等也承载了大量的交通流量,对城市交通模式的挖掘存在影响;同时,时间尺度和空间尺度的选择对于聚类结果会产生一定的影响,在今后的研究中,需要多源融合不同交通流数据在更多的城市进行城市功能结构的论证工作,并尝试着从不同的时空尺度分析中心城区的功能结构,验证研究结果的普遍性。

## 参 考 文 献

- [1] Zhong Y, Zhu Q, Zhang L. Scene Classification

- Based on the Multifeature Fusion Probabilistic Topic Model for High Spatial Resolution Remote Sensing Imagery[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53(11): 6 207-6 222
- [2] Liu Y, Liu X, Gao S, et al. Social Sensing: A New Approach to Understanding Our Socioeconomic Environments[J]. *Annals of the Association of American Geographers*, 2015, 105(3): 1-19
- [3] Yin J, Yin Z, Zhong H, et al. Monitoring Urban Expansion and Land Use/Land Cover Changes of Shanghai Metropolitan Area During the Transitional Economy (1979-2009) in China[J]. *Environmental Monitoring and Assessment*, 2011, 177 (1-4): 609-621
- [4] Liu X, He J, Yao Y, et al. Classifying Urban Land Use by Integrating Remote Sensing and Social Media Data [J]. *International Journal of Geographical Information Science*, 2017, 31 (8): 1 675-1 696
- [5] Tu W, Cao J, Yue Y, et al. Coupling Mobile Phone and Social Media Data: A New Approach to Understanding Urban Functions and Diurnal Patterns[J]. *International Journal of Geographical Information Science*, 2017, 31(12): 2 331-2 358
- [6] Yao Y, Li X, Liu X, et al. Sensing Spatial Distribution of Urban Land Use by Integrating Points-of-Interest and Google Word 2 Vec Model[J]. *International Journal of Geographical Information Science*, 2017, 31(4): 825-848
- [7] Niu Xinyi, Ding Liang, Song Xiaodong. Understanding Urban Spatial Structure of Shanghai Central City Based on Mobile Phone Data[J]. *Urban Planning Forum*, 2014(6): 61-67(钮心毅,丁亮,宋小冬. 基于手机数据识别上海中心城的城市空间结构[J]. *城市规划学刊*, 2014(6): 61-67)
- [8] Cao Rui, Tu Wei, Chao Baichong, et al. Identification and Analysis of Home and Work Regions in the Vicinity of Metro Stations Using Smart Card Data [J]. *Journal of Geomatics*, 2016, 41(3): 74-78(曹瑞,涂伟,巢佰崇,等. 基于智能卡数据的地铁周边职住用地识别与分析[J]. *测绘地理信息*, 2016, 41 (3): 74-78)
- [9] Zheng Y, Capra L, Wolfson O, et al. Urban Computing: Concepts, Methodologies, and Applications [J]. *ACM Transactions on Intelligent Systems and Technology*, 2014, 5(3): 1-55
- [10] Yuan J, Zheng Y, Xie X. Discovering Regions of Different Functions in a City Using Human Mobility and Pois[Z]. ACM, 2012
- [11] Liu Y, Kang C, Gao S, et al. Understanding Intra-Urban Trip Patterns from Taxi Trajectory Data[J]. *Journal of Geographical Systems*, 2012, 14(4): 463-483
- [12] Liu Y, Wang F, 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(1): 73-87
- [13] Liu Yu, Kang Caogui, Wang Fahui. Towards Big Data-Driven Human Mobility Patterns and Models [J]. *Geomatics and Information Science of Wuhan University*, 2014, 39(6): 660-666(刘瑜,康朝贵,王法辉. 大数据驱动的人类移动模式和模型研究[J]. *武汉大学学报·信息科学版*, 2014, 39(6): 660-666)
- [14] Jiang Yunliang, Dong Moxuan, Fan Jing, et al. Research on Identifying Urban Regions of Different Functions Based on POI Data[J]. *Journal of Zhejiang Normal University (Nat Sci)*, 2017, 40(4): 7-10(蒋云良,董墨萱,范婧,等. 基于 POI 数据的城市功能区识别方法研究[J]. *浙江师范大学学报(自然科学版)*, 2017, 40(4): 7-10)
- [15] Pan G, Qi G, Wu Z, et al. Land-Use Classification Using Taxi GPS Traces[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2013, 14(1): 113-123
- [16] Sun Zhigang. Urban Function Theory[M]. Beijing: Economy and Management Press, 1998(孙志刚. 城市功能论[M]. 北京: 经济管理出版社, 1998)
- [17] Yang Junyan, Lü Chuanting, Yang Ming, et al. Research on Urban Center System Planning of Guangzhou[J]. *City Planning Review*, 2011(10): 23-31(杨俊宴,吕传廷,杨明,等. 广州城市中心体系规划研究[J]. *城市规划*, 2011(10): 23-31)
- [18] Zhang Jingxiang. A Brief History of Western Urban Planning Thought[M]. Nanjing: Southeast University Press, 2005(张京祥. 西方城市规划思想史纲[M]. 南京: 东南大学出版社, 2005)
- [19] Gunopulos D, Das G. Time Series Similarity Measures and Time Series Indexing[J]. *ACM SIGMOD Record*, 2001, 30(2): 624
- [20] Bagnall A, Janacek G. Clustering Time Series with Clipped Data[J]. *Machine Learning*, 2005, 58(2): 151-178
- [21] Petitjean F C C O, Ketterlin A, Gan C C, et al. A Global Averaging Method for Dynamic Time Warping, with Applications to Clustering[J]. *Pattern Recognition*, 2011, 44(3): 678-693
- [22] Mirkin B. Clustering: A Data Recovery Approach [M]. 2nd ed. New York: CRC Press, 2012
- [23] Rousseeuw P J. Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis [J]. *Journal of Computational and Applied*



- Mathematics*, 1987(20): 53-65
- [24] Verburg P H, de Nijs T C, van Eck J R, et al. A Method to Analyse Neighbourhood Characteristics of Land Use Patterns[J]. *Computers, Environment and Urban Systems*, 2004, 28(6): 667-690
- [25] Wei Y D, Ye X. Urbanization, Urban Land Expansion and Environmental Change in China[J]. *Stochastic Environmental Research and Risk Assessment*, 2014, 28(4): 757-765
- [26] Liu Y, Fang F, Li Y. Key Issues of Land Use in China and Implications for Policy Making[J]. *Land Use Policy*, 2014, 40: 6-12

## Sensing Multi-level Urban Functional Structures by Using Time Series Taxi Trajectory Data

YAO Yao<sup>1</sup> ZHANG Yatao<sup>2</sup> GUAN Qingfeng<sup>1</sup> MAI Ke<sup>2</sup> ZHANG Jinbao<sup>3</sup>

<sup>1</sup> School of Geography and Information Engineering, China University of Geosciences (Wuhan), Wuhan 430074, China

<sup>2</sup> State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

<sup>3</sup> School of Geography and Planning, Sun Yat-Sen University, Guangzhou 510275, China

**Abstract:** Geospatial big data has been widely applied to distinguish urban functions. Especially, the research of functional hierarchy is of profound significance in understanding the structural characteristics and distributional forms of urban functions thoroughly, but related studies are still vacant. Therefore, this study depicts human mobility patterns based on time-series taxi trajectory data and point of interests (POIs) data, and identifies urban functional properties and spatial structures through combining dynamic time warping and *K*-MEDOIDS clustering algorithm. The results show that the urban functions in the central area of Guangzhou possess obvious hierarchical characteristics. With the refinement of hierarchies, the functional properties are developed from the dual-structure of working-living into the triple-structure of working-living-entertainment gradually. Moreover, the spatial structure is consistent with the ring-shaped structure, and its urban functions are gradually transformed from the living-entertainment function at outer rings into the commercial-entertainment function at central rings, and present different tendencies at different rings. This study provides efficient references for urban planners to understand the property changes and structural characteristics of urban functions.

**Key words:** urban functional structures; time-series taxi trajectory data; hierarchy; ring-shaped structure; social sensing

**First author:** YAO Yao, PhD, associate professor, specializes in spatiotemporal big data analytics and fine-scale urban development simulation. E-mail: yaoy@cug.edu.cn

**Corresponding author:** ZHANG Yatao, master. E-mail: yatao@foxmail.com

**Foundation support:** The National Natural Science Foundation of China, Nos. 41801306, 41671408; the Open Fund of State Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, No. 18S01.