

引用格式:关庆锋,任书良,姚尧,等.耦合手机信令数据和房价数据的城市不同经济水平人群行为活动模式研究[J].地球信息科学学报,2020,22(1):100-112. [Guan Q F, Ren S L, Yao Y, et al. Revealing the behavioral patterns of different socioeconomic groups in cities with mobile phone data and house price data[J]. Journal of Geo-information Science, 2020,22(1):100-112] DOI:10.12082/dqxxkx.2020.190406

耦合手机信令数据和房价数据的城市不同经济水平人群行为活动模式研究

关庆锋¹,任书良¹,姚尧^{1,2*},梁迅¹,周剑锋¹,袁泽皓¹,戴良洋¹

1. 中国地质大学(武汉)地理与信息工程学院,武汉 430078;2. 阿里巴巴集团,杭州 311121

Revealing the Behavioral Patterns of Different Socioeconomic Groups in Cities with Mobile Phone Data and House Price Data

GUAN Qingfeng¹, REN Shuliang¹, YAO Yao^{1,2*}, LIANG Xun¹, ZHOU Jianfeng¹, YUAN Zehao¹, DAI Liangyang¹

1. School of Geography and Information Engineering, China University of Geosciences, Wuhan 430078, China;

2. Alibaba Group, Hangzhou 311121, China

Abstract: The spatial distribution characteristics and activity patterns of urban populations play essential roles in studies of spatial isolation, optimizing urban resource allocation, and so on. Because of the sensitivity of population activity data and socioeconomic data, previous studies focus mostly on the macro level. They have difficulties in dividing the socioeconomic status and quantitatively analyzing human mobility regulation. In recent years, geospatial big data, such as the mobile app data, provide us with a rare opportunity to analyze the human activity of urban internal problems. In this study, we constructed a fine-grained activity portrait of mobile phone users based on the mobile phone signaling data of Shenzhen residents, and coupled the high-resolution Shenzhen house price distribution data to achieve accurate division of people by their economic levels. Then, we extracted six activity indicators, which include the number of active locations, activity entropy, moment of inertia, travel time, travel distance, and travel speed, to quantify the spatial distribution and analyze the activity patterns of people at different economic levels. The results reveal the correlation between mobility and socioeconomic status. The distribution of people's activities at different economic levels in Shenzhen was related to the economic development of each administrative region. The results also demonstrated that three activity indicators (moment of inertia, travel distance, travel speed) were positively related to the economic level. Residents across different socioeconomic classes exhibited different travel patterns. Likely because the rich people live in the southwest of Shenzhen, but their work locations have more self-selectivity. This leads to the distribution of home and work locations in different administrative districts and the home-work distance of high-

收稿日期:2019-07-29;修回日期:2019-11-04.

基金项目:国家重点研发计划项目(2017YFB0503804);国家自然科学基金项目(41801306、41671408、41901332);武汉大学测绘遥感信息工程国家重点实验室开放基金项目(18S01);湖北省自然科学基金项目(2017CFA041). [**Foundation items:** National Key Research and Development Program of China, No.2017YFB0503804; National Natural Science Foundation of China, No.41801306, 41671408, 41901332; Open Fund of State Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, No.18S01; Natural Science Fund of Hubei Province, No.2017CFA041.]

作者简介:关庆锋(1977—),男,四川绵阳人,博士,教授,研究方向为高性能空间计算和时空大数据。E-mail: guanqf@cug.edu.cn

*通讯作者:姚尧(1987—),男,广东梅州人,副教授,研究方向为空间大数据和城市计算。E-mail: yaoy@cug.edu.cn

economic people are larger than others. For the other three activity indicators (number of active locations, activity entropy, travel time) that reflect the similar pattern of activity between different socioeconomic status, we found that people were mainly concentrated in living and working locations on weekdays. These locations share activities on weekdays for people at different socioeconomic levels. The socioeconomic status does not affect the number of daily activities nor the scheduling of activities. This study provides necessary data and policy guidance for government and urban planners.

Key words: mobile phone data; housing price data; socioeconomic status; spatial distribution characteristic; behavioral activity regulation; home-work separation; Shenzhen

***Corresponding author:** YAO Yao, E-mail: yaoy@cug.edu.cn

摘要:有效分析城市不同经济水平人群的分布特征和活动模式对优化城市资源配置和揭示空间隔离现象有着极高的价值。但人群活动和社会经济等级数据较敏感,使得以往研究仅停留在宏观层面,难以合理划分人群经济等级并对其空间分布和活动模式进行定量分析。本研究以深圳为研究区,基于空间位置关联分析方法,耦合手机信令数据和细尺度房价数据实现了人群经济水平的准确划分,通过计算活动指标定量的分析了不同经济水平人群的空间分布和行为活动特征。研究表明:深圳市不同经济水平人群活动分布与各行政区经济发展相关,呈现“南高北低,西高东低”的格局;深圳市不同经济水平人群之间活动模式存在差异,活动范围、出行距离、出行速度与经济水平存在正向相关;高经济水平人群职住地点相距较远,存在跨行政区分布的现象。本研究分析了城市不同经济水平人群的空间分布特征和活动模式,对城市规划和解决社会不平等问题具有重要的参考价值。

关键词:手机信令数据;房价数据;社会经济等级;空间分布特征;活动模式;职住分离;深圳

1 引言

中国是一个典型的流动性社会,第六次全国人口普查显示,我国城乡流动人口总量达2.21亿^[1]。大规模的流动人口为社会经济发展做出重大贡献^[2]的同时也带来许多社会问题,如疾病传播^[3]、交通拥挤^[4]、环境污染^[5]等。而随着城市化进程和社会经济水平发展的加快,城市空间与人群行为活动不断交互,城市空间变得越来越复杂,不同经济水平人群行为活动越来越多样化^[6-8]。因此,对城市不同经济水平人群行为活动的分析变得越来越重要。

在信息和通信技术普及之前,旅行问卷调查是评估和比较社会阶层和经济水平的可靠数据^[9-11]。Kanyama等^[12]研究表明,个人社会经济状况对于出行模式存在较大影响。Hanson等^[9]发现个人出行的频率与就业存在正相关,并且个人收入越高,轨迹点的离散程度越大。但由于经济水平数据的敏感性和调查问卷的高成本,使得这些研究区域较小,数据量较少,难以证明城市人群出行模式和经济水平等级是否存在普遍联系。近年来,随着经济水平、教育程度等多源数据的出现,城市人群活动模式的研究有了新的进展。Smith等^[13]通过对比爱沙尼亚不同语言群体的手机用户出行模式,发现语言对个体出行模式有着显著的影响。Xu等^[14]基于深

圳手机用户的手机信令数据,提出基于家庭位置来分析人类日常移动模式的方法,分析结果显示人们日常活动与城市内部不同地区经济差异存在相关性。

上述研究,虽然考虑到电话用户的社会人口特征(经济水平、文化程度等),但是因为难以将用户轨迹数据与社会人口特征相耦合,仍不能很好地揭示手机用户和社会人口特征之间关系。对于经济等级这一社会人口特征,目前常用的方法是将用户家庭位置与人口经济普查相关联^[15-16]。但由于人口经济普查数据尺度较大且较为敏感,并不能在精细尺度下进行不同经济水平人群行为活动模式分析。

基于对以往的研究中所存在的不足,本研究通过耦合手机信令数据和细尺度房价数据实现不同经济水平人群的准确划分,并通过计算惯性矩、活动熵、出行速度等活动指标定量的分析不同经济水平人群的空间分布和行为活动特征。这对于深入理解不同经济水平人群对城市空间利用的差异具有重要意义,可为合理配置城市资源、解决城市隔离等社会不公平问题提供宝贵参考。

2 研究区概况及数据来源

2.1 研究区概况

本研究的研究区域为中国的广东省深圳市,占

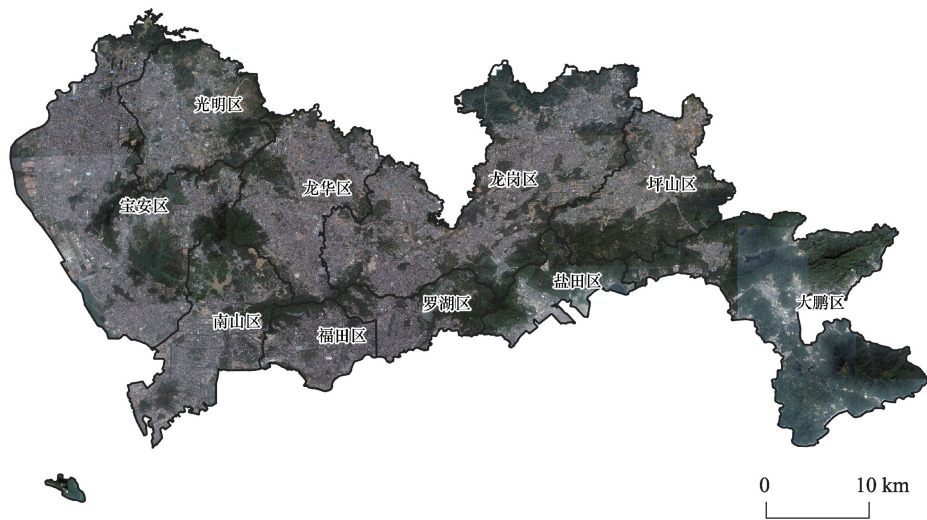
地面积达 1996.850 km², 常驻人口为 1252.83 万人。目前, 深圳共下辖 10 个(图 1)行政区^[17]。深圳的南北区域社会经济发展差异大, 人口分布不均。南部老特区(南山区、福田区、罗湖区)金融、教育较发达, 被称为“关内”, 而其他区域制造业、高新技术发达, 被称为“关外”^[18]。据统计, 深圳市是全国流动人口最多的城市^[19], 流动人口为深圳市社会经济发展贡献巨大力量的同时对社会治安^[20]、交通管理^[21]、城市环境^[22]带来了巨大的挑战。因此, 深圳独特的社会经济和人口状况使其成为研究人群流动模式的典型区域。

2.2 数据来源

本研究共收集深圳市 2013 年 3 月某一工作日的 1046.33 万人的手机信令数据, 占深圳市总人口的 80% 以上。数据的采集方式为主动采集, 采集时间如表 1 所示。为保护隐私, 本研究未获取任何个

人信息, 数据中的电话用户被随机分配用户编号。此外所有手机信令的位置信息都为用户所在服务区基站位置, 并不显示用户的真实位置信息。如图 2 所示, 深圳市共有 5943 个手机基站, 对于每个基站建立相应的泰森多边形以表示服务区域。

国内社会经济水平等敏感指标数据较难获取, 但前人研究表明, 城市尺度的住宅价格与住宅租金之间存在强相关性^[23-24]。且在短时间内租售价格比是稳定的^[25-26], 因此区域住宅价格是居民社会经济水平的有力指标。本研究使用深圳市房价分布数据(数据来自中国最大在线房产交易平台: Fang.com)与用户家庭位置相结合, 以此获得用户住宅价格并进行用户不同经济水平的分级。此外, 为解决传统房价数据, 存在空间尺度较大、偏远地区数据缺失的问题。使后续不同经济水平人群的划分更加准确, 本研究利用 Yao 等^[27-28]提出的耦合卷积神经网络和随机森林的深度学习框架。得到了分辨



注: 背景数据是天地图提供的广东省深圳市遥感图像, 空间分辨率为 5 m。

图 1 广东省深圳市示意

Fig. 1 Study area of Shenzhen city of Guangdong Province

表 1 手机信令数据样例

Tab. 1 Examples of mobile phone location records

用户 id	记录次数	记录时刻	记录位置	记录时刻	...
f5d4a*****0205	22	20120323 00:01:32	114.18** 22.64**	20120323 01:28:39	...
0bdf1*****91cb	24	20120322 23:30:13	114.21** 22.60**	20120323 00:30:15	...
1db81*****adf3	23	20120322 23:25:37	114.21** 22.60**	20120323 00:09:29	...
4cdd3*****49a3	9	20120323 12:53:30	114.09** 22.73**	20120323 02:27:50	...
556df*****439c	22	20120322 23:23:27	114.21** 22.60**	20120323 00:26:04	...
...
5790f*****c970	14	20120323 10:55:40	114.35** 22.70**	20120323 11:26:35	...

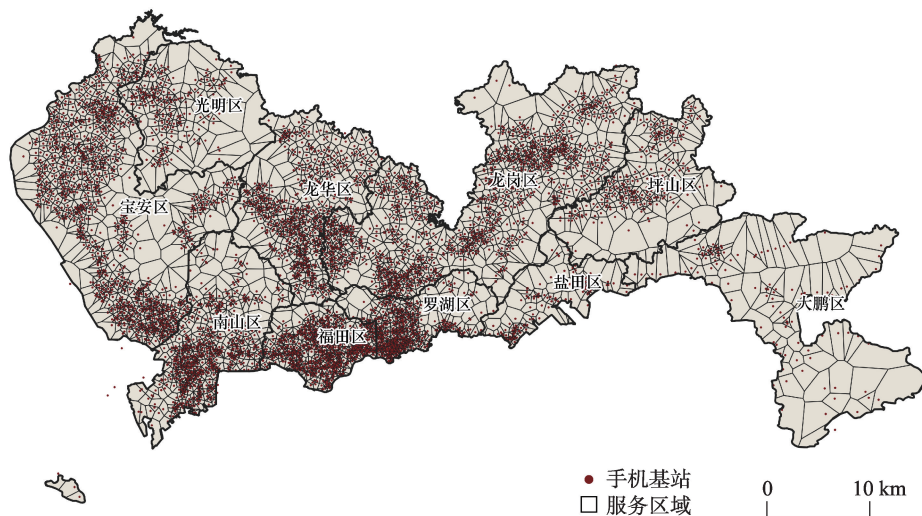


图2 2013年深圳市手机基站服务区

Fig. 2 Service area of mobile phone towers in Shenzhen in 2013

率为5 m的深圳市精细房价(图3)分布数据。

3 研究方法

如图4所示,为了更好地通过地理关系来组织和理解多源时空数据。本研究首先采用停驻点提取算法^[29-30],将用户原始位置数据提取成更有意义的停驻点位置序列。随后结合日常生活的时间规律,从位置序列中获取用户家庭位置、工作位置等活动信息。并将用户家庭位置与基于耦合卷积神经网络和随机森林所获取的精细房价数据相关联,

以获得用户的住宅价格,然后从用户住宅价格数据本身的特性出发将其经济水平分为低、中低、中、中高、高5个等级。此外,本研究计算了惯性矩^[31]、活动熵^[32]、活动位置数^[33]、出行距离、出行时间、出行速度六项活动指标,用于量化用户活动特征。已有大量研究表明,手机信令数据可以被用来开展城市地块尺度或更加精细尺度的研究^[34],对于空间定位的不确定性前人已经进行了相关研究说明^[35-36]。因此,本研究结果指示的不同经济水平人群出行差异主要为人群的固有差异。

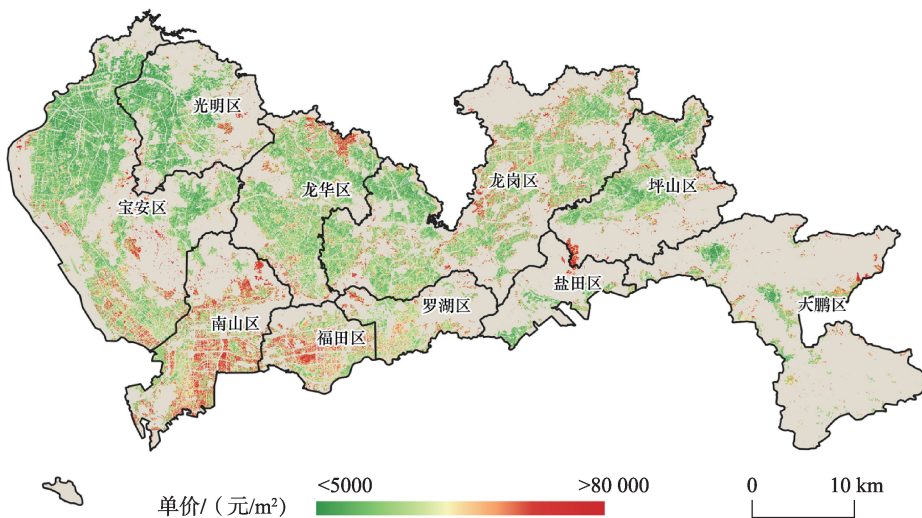


图3 2017年深圳市5 m分辨率精细房价空间分布

Fig. 3 Housing prices mapping of Shenzhen in 2017 with a 5 m resolution

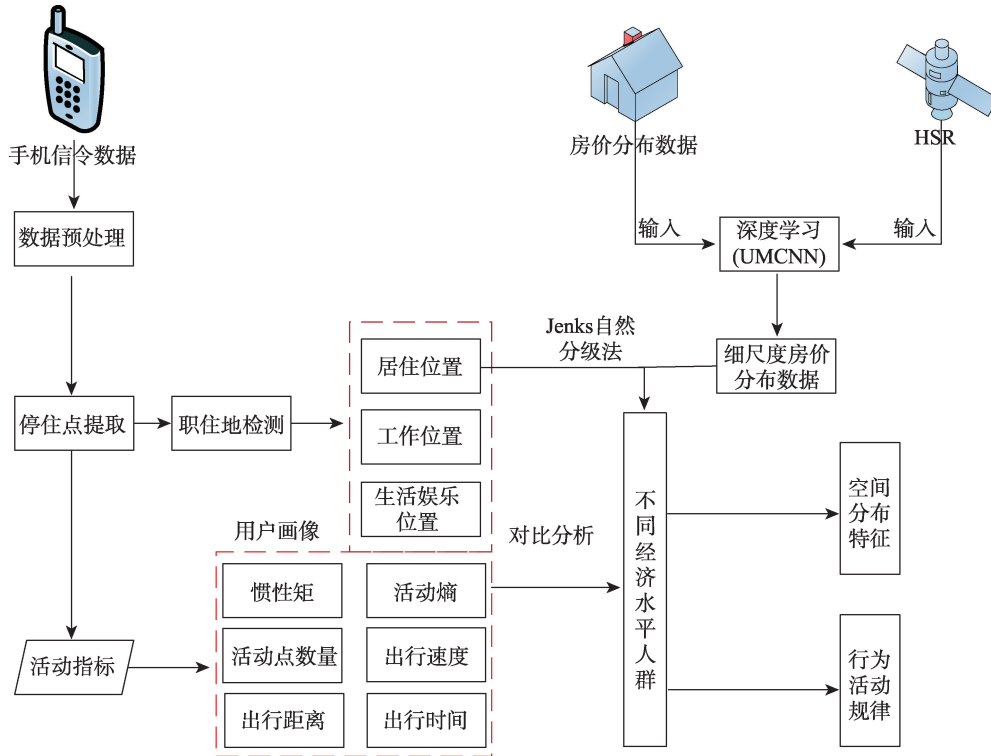


图4 耦合手机信令数据和房价分布数据的不同经济水平人群行为活动模式分析流程

Fig. 4 Framework of analyzing the behavioral patterns of people at different economic levels with mobile phone data and house price data

3.1 停驻点提取

原始用户手机信令位置数据为 $\{l_1, l_2, \dots, l_n\}$, n 为用户信令位置的采集次数。依次计算当前点和其后续所有点之间的距离,若其距离在给定距离阈值 d 内。则测量当前点与最后在距离阈值内的后续点的时间跨度,若该时间跨度大于给定时间阈值 t ,则将符合阈值条件的所有点聚类并记作 s_i 。计算公式如下:

$$l_1 l_2 \dots l_n S = s_1 \xrightarrow{t_1} s_2 \xrightarrow{t_2} \dots \xrightarrow{t_{n-1}} s_n \quad (1)$$

考虑到用户之间出行工具以及出行习惯的差异会导致距离阈值不同,本文将计算用户当日各原始位置间的平均欧氏距离作为距离阈值。在时间阈值的设置上,本研究考虑到数据采集以 1 h 为间隔,因此设置 1 h 为时间阈值。

3.2 活动指标计算

惯性矩、活动位置数、活动熵、出行时间、出行距离、出行速度是对人群活动特征量化分析的常用指标。本研究基于这些指标,通过数理统计和空间可视化的方法,实现对不同经济水平人群活动特征的分析。

3.2.1 惯性矩计算方法

原始用户手机信令轨迹数据为 $\{(l_1, t_1), (l_2, t_2), \dots, (l_n, t_n)\}$, 其中 l_i, t_i 分别表示用户当天第 i 次采集时所在的位置和时刻。惯性矩 R_g 定义如下:

$$R_g = \sqrt{\frac{\sum_{i=1}^n (\bar{l}_i - \bar{l}_c)^2}{n}} \quad (2)$$

式中: n 为用户信令位置的采集次数; i 为第几次的采集; \bar{l}_i 代表的是矢量位置; $\bar{l}_c = \sum \bar{l}_i / n$ 为用户轨迹位置的质心。 R_g 用于量化用户日常活动的空间的大小,较大的 R_g 值通常表示较大的活动空间,较小则代表用户的日常活动主要集中在较小的地理区域中。

3.2.2 活动位置数计算方法

A 为用户当日所停留的活动位置数量,公式如下:

$$A = |\text{set}(s_1, s_2, \dots, s_n)| \quad (3)$$

式中: n 为用户信令位置的采集次数。 A 较大表示用户当日活动较频繁,出行频率较高;相反, A 越小则代表用户的活动地点较少,可选择活动较为单一。

3.2.3 活动熵计算方法

基于用户停靠点,分别计算了用户在各停靠点的时间比例 $\{p_1, p_2, \dots, p_n\}$ 。并进一步计算用户当日的活动熵,公式如下:

$$p_i = \frac{\sum_{s_j=s_i} \overline{dur}_j}{\sum_{j=1}^N \overline{dur}_j} \quad (4)$$

$$H_i = -\sum_{i=1}^N p_i \log(p_i) \quad (5)$$

式中: dur 表示用户在停靠点的停靠时间; s_i 表示用户停靠点; s_j 表示相同位置的停靠点的位置点; $\sum_{s_j=s_i} \overline{dur}_j$ 表示用户在当日位置一样的 i 的停靠点时间总和; N 表示当日用户活动的位置点数量; $\sum p_i = 1$, p_i 表示用户在第 i 个位置时间占总体时间的比例。 H 越大表示用户的活动点的离散程度越高,活动多样性越大。

3.2.4 出行距离、时间和速度计算方法

停靠点检测保留了原始数据中更有意义的信息,但丢失了原始信令数据中用户的出行时间信息。因此,用户原始轨迹序列时间总和与停靠点时间总和的差即为用户出行时间。本研究首先采用几何地图匹配算法^[37]将一系列的原始经纬度,匹配为实际的道路位置 $M\{m_1, m_2, \dots, m_n\}$ 。随后,利用匹配后的实际位置以及道路网数据构建了有向图^[38]以获取实际出行距离^[39],最终计算得到出行速度。

具体公式如下:

$$T_t = T_l - T_s \quad (6)$$

$$S = s_1 s_2 \dots s_n \Rightarrow \quad (7)$$

$$M = m_1 \xrightarrow{d_1} m_2 \xrightarrow{d_2} \dots \xrightarrow{d_{n-1}} m_n \quad (8)$$

$$D = \sum_{i=1}^n d_i \quad (8)$$

$$V = \frac{D}{T_t} \quad (9)$$

式中: T_l 表示用户原始轨迹序列时间之和; T_s 为停靠点时间总和; m_i 表示道路网匹配后的用户路网位置; d_i 表示用户相邻匹配点 m_i 和 m_{i+1} 的实际路网距离; D 为用户的出行距离; V 表示用户出行速度。

3.3 职住地识别和经济水平划分

在行为活动模式的研究中,家庭和工作位置等活动点尤为重要^[40-41]。在日常活动中,人们总是从家出发,以某活动点(如工作、娱乐等)为目的。一般来说,家庭和工作是最主要的日常活动,且具

有很强的时空模式^[42]。而购物、娱乐等社会活动在日常生活中比例较少,因此本文将其统一归为生活娱乐。本研究根据用户停靠点个数及日常生活时间规律对用户停靠点进行了职住识别^[43]。若用户当日只有一个停靠点,将该停靠点记为用户居住点;若用户当日停靠点大于2个,则按照3:00-5:00的居住时间窗口,以及14:00-16:00工作时间窗口进行居住点和工作的标记;最后将没有标记的停靠点全部归为生活娱乐。

在获取用户时空活动位置后,本研究将用户居住点所在的基站区域与前文得到的深圳市细尺度的房价分布数据相耦合,基站服务区内平均住房价格即为用户住宅价格。如图5所示,本研究计算了每个手机基站服务区内所有住宅价格的标准差,并将其与整体标准差相比较,大部分手机基站服务区内标准差和整体标准差之比小于0.4。这说明手机基站服务区在很大程度上保留了深圳市房价分布异质性的特点,可以较好的代表区域用户住宅价格。随后,本文利用Jenks^[44]提出的自然间断点数据分级的方法,考虑到用户住宅价格数据自身的特点,基于数据方差与类偏差平方和等指标,将数据分成类内差异最小,类间差异最大的5类人群(图6)。

4 结果与分析

4.1 不同经济水平人群空间分布

标准差椭圆通常用于描述地理分布的空间特

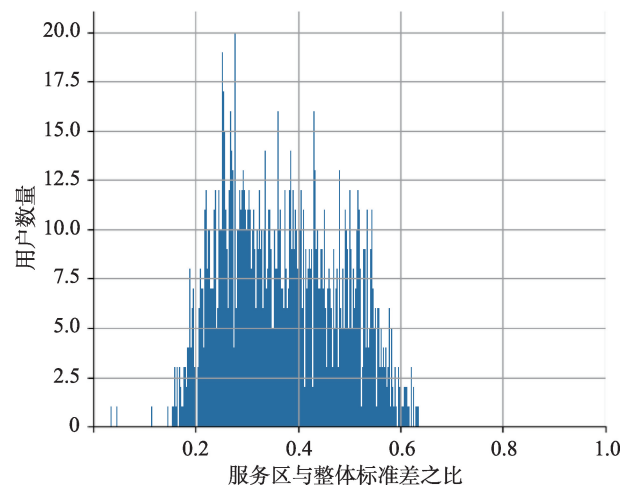


图5 深圳市手机基站服务区内房价标准差与整体房价标准差之比

Fig. 5 Histogram of the ratio between the within-cell standard deviation and overall deviation of housing price at the level of cellphone tower service areas in Shenzhen

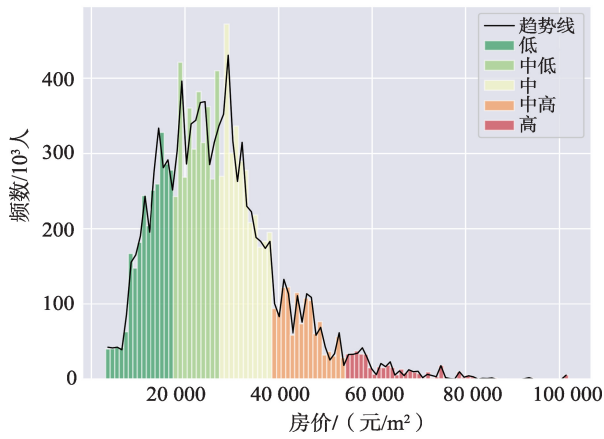


图6 深圳市基于住宅价格的不同经济水平人群划分
 Fig. 6 Dividing the population of different economic levels based on housing prices in Shenzhen

征^[45],如集中趋势、离散度和分布方向趋势。为了更好地分析不同经济水平人群日常活动的空间分布,本文引入标准差椭圆^[34]。同时为了减小在采集时手机信令数据记录位置为服务基站位置而不是真

实位置所造成的误差,本文在分析时,将同一基站所采集的用户,在服务区内随机生成相应用户点。

如图7,相同经济水平人群的居住、工作、生活娱乐标准差椭圆的中心以及长短半轴相近,这表明相同经济水平人群居住、工作、生活娱乐位置的分布模式相近。符合用户日常活动存在时间连续性的规律。而不同经济水平人群居住、工作、生活娱乐标准差椭圆的位置和形状存在较大差别。如图7(e)、图7(d),较高经济水平人群日常活动的标准差椭圆呈现东北—西南方向分布,其中心点位于深圳市南部地区,且椭圆短轴较短,长短半轴相差较大。而中等经济水平人群日常活动的标准差椭圆(图7(c))则呈现东—西分布,其中心点开始上移至深圳中部地区,且椭圆短轴变长,长短半轴长度比变小。

经济水平较低人群日常活动的标准差椭圆(图7(a)–(b))呈现西北—东南分布,其中心点位于深圳市的北部区域,且短轴较短,短半轴长度比较大。通过对比不同经济水平日常活动点的标准差椭圆,

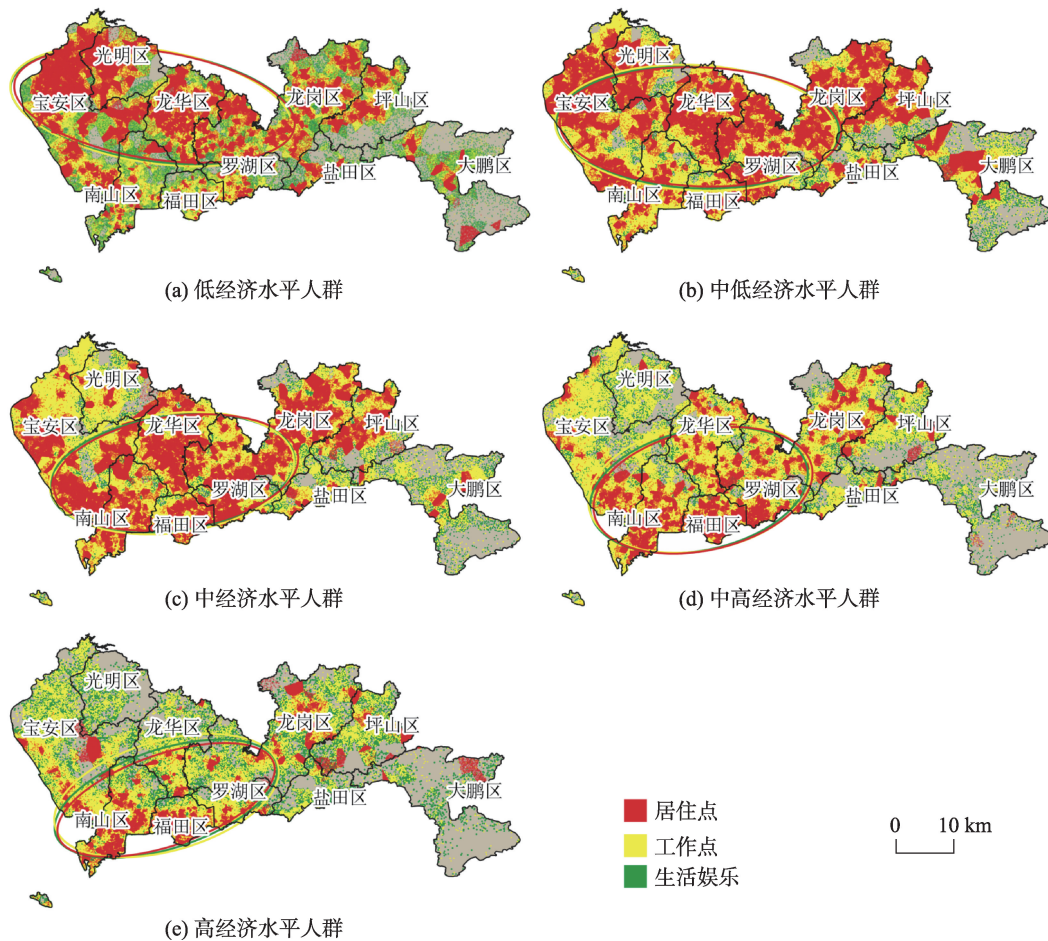


图7 深圳市不同经济水平人群活动点分布
 Fig. 7 Spatial distribution of activity points of people at different economic levels in Shenzhen

可以发现,不同经济水平人群在不同行政区的活动点分布存在较大的差异。经济水平较高的人群活动点多分布在原“关内”地区,如南山区、福田区、罗湖区,其聚集性较强,离散程度较小,具有明显的方向性。而随着经济水平降低,人群活动点的分布开始向宝安区、龙华区、龙岗区移动,且离散程度变大。而经济水平较低人群的活动点则明显在光明区和宝安区的北部以及坪山区和龙岗区聚集,和高经济水平人群相似,其分布具有明显的聚集性和方向性。

综上所述,深圳市不同经济水平人群日常活动分布的差异是由深圳市南北区域经济发展不均衡以及各区承担的功能差异造成的。福田区、南山区、罗湖区为深圳原“关内”,福田区为深圳市的中心商务区,罗湖区为深圳最早开发的城区。因此,此地多为深圳市高收入人群的活动区域。而宝安区、龙华区以及龙岗区毗邻“关内”,承担着中心城区的服务溢出,聚集了大量高新技术产业且拥有多处风景区和湿地公园,因此此地多为高经济水平人

群的居住点和其他经济水平的工作和娱乐点。而光明新区、坪山区以及大鹏区成立较晚,且存在大量工业区和科技园,因此大量外来务工人员以及工厂工人居住、工作于此。

4.2 不同经济水平人群活动特征分析

如图 8(a)、(b)不同经济水平人群活动位置数、活动熵之间差异不大。大多数工作日主要集中在居住、工作地点。这些地点是不同经济水平人群在工作日的共有活动,经济水平的高低并不会影响人们工作日的日常活动点数量以及活动时间安排。

与活动位置数和活动熵不同,不同经济水平人群之间惯性矩(图 8(c))存在较大的差异。本研究根据用户惯性矩对数分布中的拐点将其分为 5 类。第 1、2 类惯性矩(图 9(b)中绿色区域)多集中分布在深圳市宝安区、南山区、福田区、伶仃岛等西南沿岸以及龙华区、龙岗区(图 10)。这些地区经济和工业发达,有大量城市综合体、工业园、旅游产业等,生

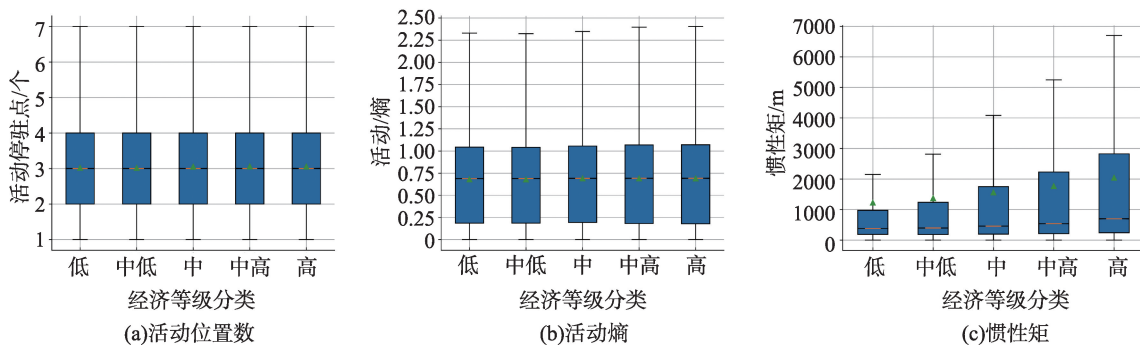
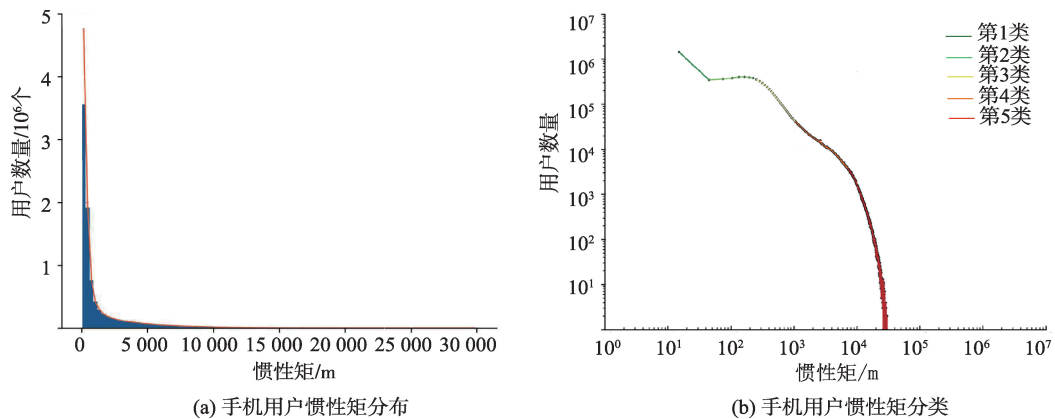


图 8 深圳市不同经济水平人群活动指标箱线图

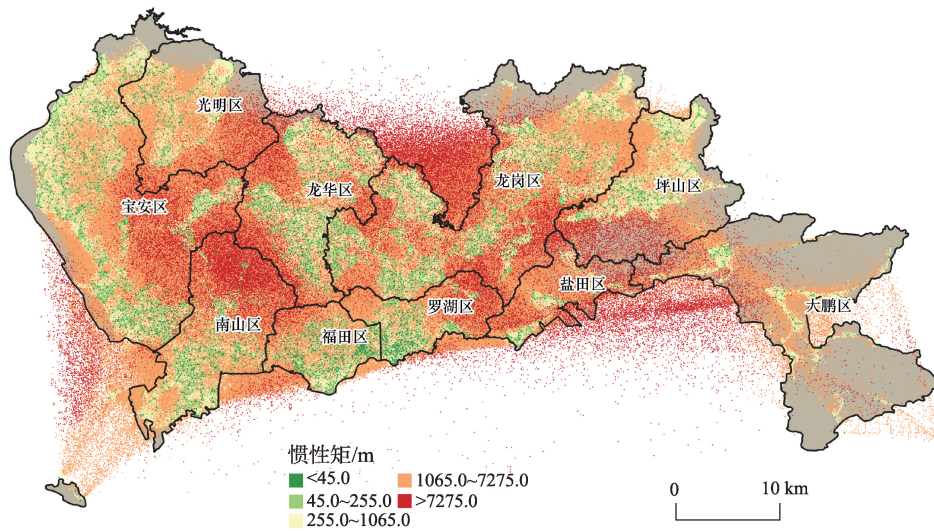
Fig. 8 Boxplots of activity indicators for different economic levels in Shenzhen



注:图(b)坐标轴为双对数轴。

图 9 深圳市全体手机用户惯性矩分布

Fig. 9 Density distribution of the moment of inertia of all Shenzhen phone users



注:图中点为惯性矩的质点,即为用户日常活动区域中心点。

图10 深圳市手机用户惯性矩空间分布

Fig. 10 Spatial distribution of the moment of inertia of all Shenzhen phone users

活配套设施较为完善。因此用户日常活动范围较小,在较小的区域便可以进行居住、工作、生活娱乐等活动。第3类惯性矩(图9(b)中黄色区域)集中于深圳各区中部。而第4、5类惯性矩的活动中心(图9(b)中红色区域)集中分布在各区交界处,如西南伶仃岛附近海域以及福田和大鹏区中间海域。其日常活动范围较大,用户在进行居住、工作娱乐等活动时,往往伴随着跨区流动。

随后,本文基于上述5类惯性矩,对不同经济水平人群活动中心及活动范围进行分析(图11)。发现随着用户经济水平的升高,第1、2、3类惯性矩在各人群所占比例不断下降,第4、5类惯性矩所占比例上升。经济水平较低用户的第2类惯性矩所占比例大于第4类,较高经济水平则相反。且与其他经济水平人群相比,高经济水平人群,第4类惯性矩所占比例最大。由此可知,随着用户经济水平的提高,用户日常活动的范围不断扩大,跨区活动占比越来越高。

如图12(a),不同经济水平人群出行时间之间差异不大,但不同经济水平用户出行距离存在较大差异(图12(b))。随着人群经济水平的提升,出行距离不断变大。且经济水平越高的人群居住、工作点往往相距较大,并存在跨区分布的现象。

随后本研究计算出行速度,揭示了不同经济水平人群的出行的空间模式。如图13,不同经济水平人群在不同行政区的出行速度存在较大的差异。

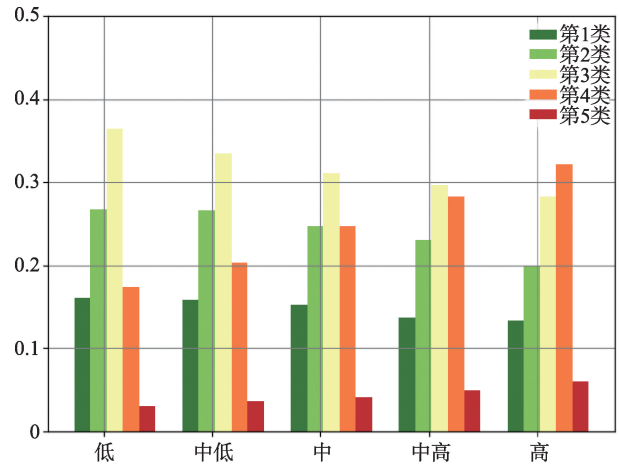


图11 深圳市不同经济水平人群惯性矩分类比例

Fig. 11 Moment of inertia of peoples in Shenzhen per economic levels

较高经济水平人群在南山区、福田区出行速度较低,而在龙华区、宝安区、龙岗区的南部则出行速度较高。而较低经济水平人群出行速度的分布则相反。

结合不同人群活动点分布可知,由于南山区、福田区多为高经济水平人群的居住点和工作点,所以在此区域活动的高经济水平人群出行速度较低,而较低经济水平人群由于居住地较远,所以出行速度较高。而龙华区、龙岗区、光明区多为低经济水平居住、工作点,所以此地区低经济水平人群日常出行速度较低。

综上所述,经济水平状况和惯性矩、出行距离、

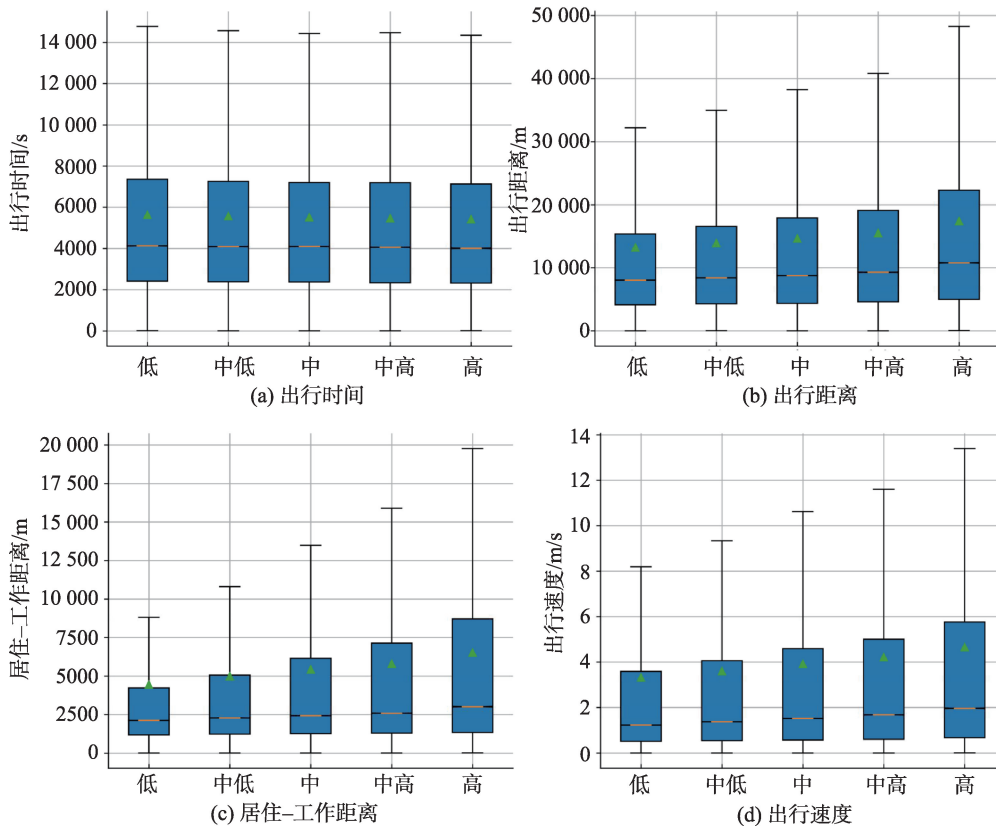


图 12 深圳市不同经济水平人群出行指标箱线图

Fig. 12 Boxplots of travel indicators for people with different economic levels in Shenzhen

职住距离、出行速度存在正向相关性(表2)。人群经济水平越高,其活动范围越大,日常出行距离越远、出行速度越快。深圳市不同经济水平人群出行模式的分布存在较大差异,经济水平较高人群出行速度呈现南高北低的分布模式。而较低经济水平人群则相反,出行速度分布为南低北高。这是由于不同经济水平人群居住和工作点分布不同导致的。经济水平较高人群多居住在经济发达的南部区域,并且可以自由选择居住点和工作点,而低经济水平却只能居住于深圳北部房价较低区域。不同经济水平人群之间,居家时间、工作时间、生活娱乐、活动点数量、活动熵等活动指标差异性较小。这与 Xu 等^[6]基于手机信令数据对新加坡和波士顿人群移动研究所得经济水平对人们活动点数量、活动熵影响不大的结果相似,但从表2中仍可以看出随着经济水平的提高,人们的居家、工作时间逐渐减少,生活娱乐时间、活动点数量开始变多。这更加说明了国内外不同经济水平人群行为活动模式的差异性和研究的重要性。

5 结论与展望

5.1 结论

如何有效分析城市不同经济水平人群的时空分布和活动模式已经成为优化城市资源,缩小贫富差距的重要内容。但是,由于经济等级数据的敏感性,已有关于人群行为活动的研究多停留在宏观层面。本研究耦合大规模手机信令数据和细尺度房价数据,实现了对经济水平人群的合理划分。通过计算惯性矩、活动位置数、活动熵、出行时间、出行距离、出行速度这6项活动指标进而对不同经济水平人群行为活动模式进行分析。主要结论包括:

(1)通过对比深圳市不同经济水平人群日常活动的标准差椭圆可知,深圳市不同经济水平人群活动分布和各区经济发展以及产业聚集相关,人群经济水平和深圳市区域经济发展均呈现“南高北低,西高东低”的分布格局。这与 Xu 等^[14]基于 13 d 的手机信令数据分析所得到的结果相似。进一步表明了,人们出行活动存在固定的分布模式,且当单日

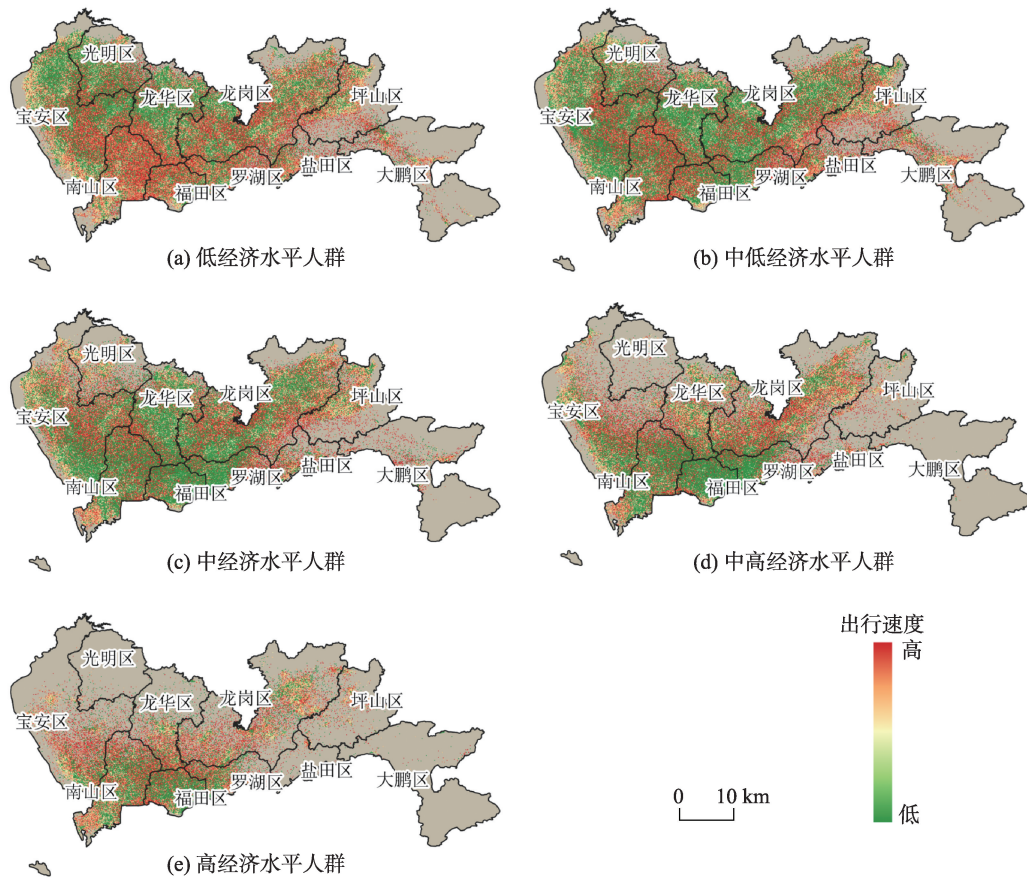


图13 深圳市不同经济水平人群出行速度分布

Fig. 13 Spatial Distribution of commuter velocity of people at different economic levels in Shenzhen

表2 深圳市不同经济水平人群活动指标对比表

Tab. 2 Comparison of activity indicators for people with different economic levels in Shenzhen

经济水平	居家时间	工作时间	生活娱乐时间	活动点数量	惯性矩	活动熵	出行时间	出行距离	职住距离	出行速度
低	1.000	1.000	0.942	0.987	0.625	0.982	0.934	0.831	0.674	0.581
中低	0.996	0.994	0.939	0.986	0.692	0.982	0.957	0.863	0.720	0.839
中	0.988	0.986	0.966	0.999	0.780	0.999	0.984	0.898	0.884	0.896
中高	0.965	0.955	0.989	1.000	0.778	1.000	1.000	0.924	0.943	0.934
高	0.943	0.945	1.000	0.999	1.000	0.998	0.997	1.000	1.000	1.000

注:为方便分析,将对比结果进行归一化处理。

手机信令数据足够多时,也可反应其分布模式。

(2)深圳市手机用户经济水平状况和惯性矩、出行距离、职住距离、出行速度存在明显的正向相关性。经济水平越高,人群活动范围越大,出行距离越长,出行速度越快。高经济水平人群的出行速度普遍较高,这使得经济水平较高的人群对居住点和工作点的选择更具有自由性。经济水平越高的人群职居地分离较远,且多出现跨区分布的现象。

(3)与惯性矩、出行距离、职住距离、出行速度不同,深圳市不同经济水平人群对于活动位置数、

活动熵、出行时间表现出非常相似的特征。这与 Xu 等^[6]基于手机信令数据对新加坡和波士顿人群移动研究所得经济水平对人们活动点数量、活动熵影响不大的结果相同。除空间活动差异外,本研究进一步从时间维度对比不同经济水平人群居家、工作、娱乐时间,发现随着经济水平变高,深圳市人群的居家和工作时间变少,娱乐时间增多。

5.2 展望

本研究未将其他社会人口统计数据纳入研究

中,如人群的年龄,性别等。这些都是影响分析的重要因素。在下一阶段,本文将继续结合其他人群社会背景信息进行深入研究。同时,对本文得到的不同类别惯性矩,后续将进行以职住地为中心的可视化等相关研究。并且本文对于惯性矩、出行距离、出行速度这3项活动指标的对比分析,可以为基于出行轨迹预测经济水平的相关研究提供依据。且本研究发现,不同经济水平人群之间活动时间安排存在差异。因此,不同经济水平人群时间活动模式应是今后研究的重点,其研究结果更有助于揭示人群对城市资源利用的差异,为解决城市隔离等社会公平等问题提供宝贵的参考。

参考文献(References):

- [1] 中华人民共和国国家统计局. 2010年第六次全国人口普查主要数据公报(第1号)[R].2011. [National Bureau of statistics of the People's Republic of China. Bulletin of main data of the sixth national population census in 2010 (no. 1)[R].2011.]
- [2] 段成荣,吕利丹,邹湘江.当前我国流动人口面临的主要问题和对策——基于2010年第六次全国人口普查数据的分析[J].人口研究,2013,37(2):17-24. [Duan C R, Lv L D, Zou X J. Major Challenges for China's Floating Population and Policy Suggestions:An Analysis of the 2010 Population Census Data.[J]. Population Research, 2013,37(2):17-24.]
- [3] Martens P, Hall L. Malaria on the move: Human population movement and malaria transmission[J]. Emerging Infectious Diseases, 2000,6(2):103-109.
- [4] 牛方曲,王芳.城市土地利用——交通集成模型的构建与应用[J].地理学报, 2018,73(2):380-392. [Niu F Q, Wang F. Modelling urban spatial impacts of land-use/transport policies[J]. Acta Geographica Sinica, 2018,73(2):380-392.]
- [5] Black R, Bennett S R G, Thomas S M, et al. Climate change: Migration as adaptation[J]. Nature, 2011,478(7370):447-449.
- [6] Batty M. The size, scale, and shape of cities[J]. Science, 2008,319(5864):769-771.
- [7] Kitchin R. The real-time city? Big data and smart urbanism[J]. GeoJournal, 2014,79(1):1-14.
- [8] Wakamiya S, Lee R, Sumiya K. Crowd-based urban characterization: Extracting crowd behavioral patterns in urban areas from twitter[J]. Pers Ubiquit Comput, 2013,17:605-620.
- [9] Hanson S, Hanson P. The travel-activity patterns of urban residents: Dimensions and relationships to sociodemographic characteristics[J]. Economic Geography, 1981,57(4):332-347.
- [10] Hanson S, Hanson P. Gender and urban activity patterns in uppsala, sweden[J]. Geographical Review, 1980,70(3):291.
- [11] Limtanakool N, Dijst M, Schwanen T. The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium and longer distance trips[J]. Journal of Transport Geography, 2006,14(5):327-341.
- [12] Carlssonkanyama A, Linden A. Travel patterns and environmental effects now and in the future: Implications of differences in energy consumption among socio-economic groups[J]. Ecological Economics, 1999,30(3):405-417.
- [13] Silm S, Ahas R. Ethnic differences in activity spaces: A study of out-of-home nonemployment activities with mobile phone data[J]. Annals of The Association of American Geographers, 2014,104(3):542-559.
- [14] Xu Y, Shaw S, Zhao Z, et al. Understanding aggregate human mobility patterns using passive mobile phone location data: A home- based approach[J]. Transportation, 2015,42(4):625-646.
- [15] Huang Q, Wong D W S. Activity patterns, socioeconomic status and urban spatial structure: what can social media data tell us?[J]. International Journal of Geographical Information Science, 2016,30(9):1873-1898.
- [16] Xu Y, Belyi A, Bojic I, et al. Human mobility and socioeconomic status: Analysis of Singapore and Boston[J]. Computers, Environment and Urban Systems, 2018,72:51-67.
- [17] 深圳市统计局.深圳统计年鉴[M].北京:中国统计出版社,2018. [Statistical yearbook of Shenzhen[M]. Beijing: China statistics press, 2018.]
- [18] 林宇川,冯健.深圳关内关外一体化过程中的边界效应及时空演变[J].热带地理,2011,31(6):580-585. [Lin Y C, Feng J. Border effect and its temporal-spatial evolution in the process of regional integration in Shenzhen[J]. Tropical Geography, 2011,31(6):580-585.]
- [19] 深圳市统计局.深圳市2015年全国1%人口抽样调查主要数据公报[R].2015. [Shenzhen statistics bureau. Shenzhen 2015 national 1% sample survey main data bulletin [R]. 2015.]
- [20] 陈刚,李树,陈屹立.人口流动对犯罪率的影响研究[J].中国人口科学,2009(4):52-61. [Chen G, Li S, Chen Y L. Population mobility and crime: An empirical analysis based on China's observation[J]. Chinese Journal of Population Science, 2009(4):52-61.]
- [21] 邵源,宋家骅.大城市交通拥堵管理策略与方法——以深圳市为例[J].城市交通, 2010,8(6):7-13. [Shao Y, Song J

- H. Traffic congestion management strategies and methods in large metropolitan area: A case study in Shenzhen [J]. *City Traffic*, 2010,8(6):7-13.]
- [22] 傅崇辉,苏杨,陆杰华,等.深圳人口与健康发展报告(2014)[R].2014. [Fu C H, Su Y, Lu J H, et al. Shenzhen population and health development report (2014)[R]. 2014.]
- [23] 郑文娟.中国城市住房价格与住房租金的影响因素及相互关系研究[D].杭州:浙江大学,2011. [Zheng W J. Study on the determinants and relationship between China's urban housing prices and rents[D]. Hangzhou: Zhejiang University, 2011.]
- [24] Cui N, Gu H, Shen T. The spatial differentiation and relationship between housing prices and rents: Evidence from Beijing in China[J]. *Geographical Research*, 2019, DOI: 10.11821/dlyj020180352.
- [25] Zhai D, Shang Y, Wen H, et al. Housing price, housing rent, and rent-price ratio: Evidence from 30 cities in China [J]. *Journal of Urban Planning and Development- asce*, 2018,144(1):4017026.
- [26] 方毅,赵石磊.房屋销售价格和租赁价格的关系研究[J]. *数理统计与管理*,2007,26(6):951-957. [Fang Y, Zhao S L. Empirical study on the relationship of sale price and rent of house in China[J]. *Application of Statistics and Management*, 2007,26(6):951-957.]
- [27] 姚尧,任书良,王君毅,等.卷积神经网络和随机森林的城市房价微观尺度制图方法[J].*地球信息科学学报*,21(2): 36-45. [Yao Y, Ren S L, Wang J Y, et al. Mapping the fine-scale housing price distribution by integrating a convolutional neural network and random forest[J]. *Journal of Geo-information Science*, 2019,21(2):168-177.]
- [28] Yao Y, Zhang J, Hong Y, et al. Mapping fine-scale urban housing prices by fusing remotely sensed imagery and social media data[J]. *Transactions in Gis*, 2018,22(2):561-581.
- [29] Ye Y, Zheng Y, Chen Y, et al. Mining individual life pattern based on location history[C]. 2009 tenth international conference on mobile data management: Systems, services and middleware, IEEE, 2009.
- [30] Zheng Y. Trajectory data mining: An overview[J]. *ACM Transactions on intelligent systems and technology*, 2015, 6(3):29.
- [31] Gonzalez M C, Hidalgo A R C, Barabasi A. Understanding individual human mobility patterns[J]. *Nature*, 2008, 453(7196):779-782.
- [32] Song C, Qu Z, Blumm N, et al. Limits of predictability in human mobility[J]. *Science*, 2010,327(5968):1018-1021.
- [33] Xu Y, Shaw S, Zhao Z, et al. Another tale of two cities: Understanding Human activity space using actively tracked cellphone location data[J]. *Annals of the American Association of Geographers*, 2016,106(2):489-502.
- [34] Xu Y, Shaw S, Yin L, et al. Understanding individual daily activity space based on large scale mobile phone locationdata[C]. *Geo-Computation 2015 Conference*, Dallas, Texas, US, 2015.
- [35] 刘瑜,肖昱,高松,等.基于位置感知设备的人类移动研究综述[J].*地理与地理信息科学*, 2011,27(4):8-13. [Liu Y, Xiao Y, Gao S, et al. A review of human mobility research based on location aware devices[J]. *Geography and Geo-Information Science*,2011,27(4):8-13.]
- [36] Miller H J. What about people in geographic information science[R]. 2005 Re-Presenting Geographic Information Systems, John Wiley, 2005:215-242.
- [37] Yin Y, Shah R R, Zimmermann R. A general feature-based map matching framework with trajectory simplification[C]. *Proceedings of the 7th ACM SIGSPATIAL International Workshop on GeoStreaming*. ACM, 2016: 7.
- [38] Hagberg A, Swart P, Chult D. Exploring network structure, dynamics, and function using network X, 2008[R]. Los Alamos National Lab, Los Alamos, NM (United States), 2008.01.
- [39] Dijkstra E W. A note on two problems in connection with graphs[J]. *Numerische Mathematics*, 1959,1(1):269-271.
- [40] Isaacman S, Becker R, Caceres R, et al. Identifying important places in people's lives from cellular network data[J]. 2011:133-151. DOI:10.1007/978-3-642-21726-5_9.
- [41] Bojic I, Massaro E, Belyi A, et al. Choosing the right home location definition method for the given dataset[J]. *arXiv: Social and Information Networks*, 2015.
- [42] Sevtsuk A, Ratti C. Does urban mobility have a daily routine? Learning from the aggregate data of mobile networks [J]. *Journal of Urban Technology*, 2010,17(1):41-60.
- [43] 曹劲舟,涂伟,李清泉,等.基于大规模手机定位数据的群体活动时空特征分析[J].*地球信息科学学报*,2017,19(4): 467-474. [Cao J Z, Tu W, Li Q Q. Spatio-temporal analysis of aggregated human activities based on massive mobile phone tracking data[J]. *Journal of Geo-information Science*, 2017,19(4):467-474.]
- [44] Jenks G F. The data model concept in statistical mapping [C]. *International Yearbook of Cartography*, 1967.
- [45] Lefever D W. Measuring geographic concentration by means of the standard deviational ellipse[J]. *American Journal of Sociology*, 1926,32(1):88-94.