

引用格式: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, et al. 2017. Spatio-temporal analysis of aggregated human activities based on massive mobile phone tracking data, 19(4): 467-474.] DOI:10.3724/SPJ.1047.2017.00467

基于大规模手机定位数据的群体活动时空特征分析

曹劲舟¹,涂伟^{2,3,4*},李清泉^{1,2,3,4},曹瑞⁵

1. 武汉大学 测绘遥感信息工程国家重点实验室,武汉 430079;2. 深圳大学 土木工程学院空间信息智能感知与服务深圳市重点实验室,深圳 518060;3. 海岸带地理环境监测国家测绘地理信息局重点实验室,深圳 518060;4. 深圳大学智慧城市研究院,深圳 518060;5. 宁波诺丁汉大学国际博士创新研究中心,宁波 315100

Spatio-temporal Analysis of Aggregated Human Activities Based on Massive Mobile Phone Tracking Data

CAO Jinzhou¹, TU Wei^{2,3,4*}, LI Qingquan^{1,2,3,4} and CAO Rui⁵

1. State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China; 2. Shenzhen Key Laboratory of Spatial Smart Sensing and Services, College of Civil Engineering, Shenzhen University, Shenzhen 518060, China; 3. Key Laboratory for Geo-Environmental Monitoring of Coastal Zone of the National Administration of Surveying, Mapping and GeoInformation, Shenzhen 518060, China; 4. Smart City Institute, Shenzhen University, Shenzhen 518060, China; 5. International Doctoral Innovation Centre, University of Nottingham, Ningbo 315100, China

Abstract: Urban space and the behavior of human activities constantly interact with each other. Investigation on distribution of aggregated human activities and spatio-temporal change benefits data-driven policy-making in urban planning and urban governing. In the era of big data, with the development of information and communication technologies, it is possible to collect city-scale data with high resolution in space and time by various location-aware devices and sensors. Exploration of spatial-temporal activities attracts a lot of attention. By taking about 10 million one-day tracking data of mobile phone users in Shenzhen, China as an example, this paper firstly identified their stay locations according to spatial and temporal rules to generate stay trajectory for each individual and recovered activity semantic information by labelling activity types for each stay locations. Then, the significant differences in patterns of distributions of stay locations and their activities were analyzed. Spatial and temporal distributions of different human activities were explored, respectively. The study shows that the distribution of stay locations and activities is obviously heterogeneous. The average number of stay locations of an individual per day is 2.1, while the average number of activities an individual engaged in per day is 3.4. This study furthermore suggests that different types of activities have temporal variance and spatial heterogeneity. The temporal distribution fluctuates significantly over 24 hours, which is in accordance with daily routine. The spatial distribution overall obeys “space power law”, and the spatial distribution of social activity, which has a faster-down tail, shows a more obvious pattern of spatial segregation than the other two activities. The study revealed the diversity and heterogeneity of spatial and temporal distribution of human aggregated activities in urban space, which is meaningful in analyzing human activities research and facilitating urban traffic optimization and urban planning.

收稿日期 2016-11-04;修回日期:2016-12-24.

基金项目:国家自然科学基金项目(41401444、41371377、41671387);深圳大学青年教师科研启动项目(2016065);国土资源部城市土地资源监测与仿真重点实验室开放基金资助课题(KF-2016-02-009)。

作者简介:曹劲舟(1991-),男,湖南益阳人,博士生,主要从事时空大数据分析 with 挖掘研究。E-mail: caojinzhou@whu.edu.cn

*通讯作者:涂伟(1984-),男,湖北黄冈人,博士,助理教授,主要从事时空大数据分析研究。E-mail: tuwei@szu.edu.cn

Key words: mobile phone tracking data; trajectory analysis; spatial-temporal big data; aggregated human activities; spatial-temporal pattern

*Corresponding author: TU Wei, E-mail: tuwei@szu.edu.cn

摘要 城市空间与居民行为不断交互,相互影响。探究城市空间中的群体活动分布及其时空变化能够帮助数据驱动的城市规划与城市治理。基于大数据的时空间群体活动研究是当前时空大数据研究的一个热点。本文以深圳市为例,基于约1000万手机用户在某一工作日的基站尺度的手机定位数据,识别用户停留位置和停留活动,重建活动语义信息,分析用户的停留点和停留活动的分布差异,研究群体活动的时空分布模式,探讨人群活动模式的多样分布特征。研究表明:停留位置和活动分布存在差异,每人每天平均的停留个数约为2.1个,而每人每天平均从事的活动约为3.4个;不同类型的活动在时间上存在波动;群体活动存在空间分异特征,整体上服从“空间幂律”。本研究揭示了城市空间中群体活动的多样性及其时空分布特征,对于城市居民活动研究、城市交通优化和城市规划具有重要的意义。

关键词 手机定位数据;轨迹分析;时空大数据;群体活动;时空特征

1 引言

城市是动态且复杂的^[1]。城市空间不仅由土地、建筑和其他基础设施组成,也与城市内居民的活动和移动行为息息相关^[2-3]。城市空间与居民行为不断交互,相互影响。随着城市化和社会变革的加速,城市空间变得复杂,人群流动和活动也变得快速化和多样化。因此,深入探究人们在城市空间的活动规律及其时空变化,可以更好地理解群体活动与城市空间的相互作用,为人类行为学研究^[4]、城市规划布局^[5]、公共交通优化^[6]等领域提供辅助决策信息。传统的群体活动分析数据主要依靠收集出行调查问卷获得,然而此种依赖人工的方式样本量少,搜集时间长,耗时耗力^[7],因此亟需可靠、易取的时空数据来快速捕捉群体活动的特征。

随着信息通信技术(Information Communication Technology, ICT)和泛在感知技术(Ubiquitous Sensing Technology)的飞速发展,使实时、高效记录人类时空位置信息成为可能,产生了海量、多源的人类时空活动数据^[8-9]。这些海量数据弥补了传统方式的不足,具有高精度、高频次、高效率、低成本等特点^[10]。由于智能手机的广泛使用,移动运营商通过通讯信号将基站尺度的时空位置记录下来,基于智能手机的移动定位数据为群体时空活动特征研究提供了新的解决途径^[11]。

目前,基于手机定位数据的时空活动识别主要从个体和群体2个层面出发^[12]。个体层面的识别主要根据一定的时空规律和先验信息对时空轨迹链中不同的位置进行语义活动标记^[13-15];群体层面则从时空活动位置与土地利用的空间关系出发,在空间集聚的尺度上来推断活动的类型^[16-19]。这些研

究都从不同角度探究了群体活动特征与城市空间的相互作用的关系,但是由于手机数据缺乏语义信息(类别、位置持续时间等),活动轨迹空间尺度不一,活动特征难以提取,因而基于大数据环境下的群体活动精准识别与分析仍存在巨大的挑战。

本文利用深圳市海量手机定位数据,通过识别用户停留位置和停留活动、标记活动语义信息,分析用户的停留点和停留活动的分布差异,研究群体活动的时空分布模式,探讨人群活动模式的多样分布特征。这对于深入理解城市空间中不同类型活动的时空分布的多样性和差异性具有重要意义,对于辅助城市交通优化和城市规划具有参考价值。

2 研究区概况及数据源

本文选择深圳作为研究区。深圳位于广东省南部,南临香港,北接东莞、惠州,总面积1996 km²。如图1所示,深圳市下辖8个行政区,其中罗湖、福田、南山原属经济特区,2010年之后特区范围扩大到整个辖区。截止2015年末常住人口约137.891万人,人口密度每平方公里达到10 000人,位列全国第一,流动人口总数在全国最多^[20-21]。因为历史发



图1 研究区域

Fig. 1 The study area

展的原因,深圳的南北区域经济社会发展差异大,人口分布不均匀,原特区部分(罗湖、福田、南山)在科技、金融、教育等领域高度发达,人口分布集中,旧称为“关内”,而其他行政区被称为“关外”,其制造业较为发达,人口分布较少。随着深圳城市化加速和人口的持续扩张,关内关外差异在逐渐缩小,跨区域群体出行逐渐增多,人群活动变得更为复杂。

本文采用2012年3月某一工作日约1000万用户的移动手机定位数据,移动运营商以1h左右的采样间隔通过移动基站记录手机用户通信的基站编号,获取用户所在位置。数据由移动运营商进行了匿名化处理,以保护用户隐私。通过对每一条数据记录的时间进行小时整数化处理,得到包含用户id、小时数、所在基站id、基站位置经度、基站位置纬度等属性的数据序列,如表1所示。

表1 移动手机定位数据示例

Tab. 1 Examples of mobile phone location records

用户id	小时数	基站id	基站经度	基站纬度
536****	0	19**	114.14**	22.60**
536****	1	19**	114.14**	22.60**
536****	2	54**	114.12**	22.58**
...
536****	23	14**	114.14**	22.60**

注:为了隐私保护,具体数值以*号标识

3 研究方法

为了研究群体活动的时空分布模式,探讨群体活动模式的多样分布特征,以及在时空上的分异,本文首先通过对原始数据的预处理和停留轨迹提取,通过构建一定的时空规则实现对人类时空活动精确识别,标记活动语义信息;随后对不同类型的活动的时空特征进行分析。具体处理流程(图2)主要包括:①基于手机基站尺度的时空轨迹生成;②时空停留轨迹提取;③家庭-工作-社会活动识别;④群体活动时空特征分析。

3.1 基于手机基站尺度的时空轨迹生成

为了从海量的手机定位数据中生成完整的时空轨迹序列,首先需要对原始数据进行预处理,过滤掉不符合需求的数据,主要步骤包括:①去除重复数据;②去除缺少属性的数据;③去除时间和空

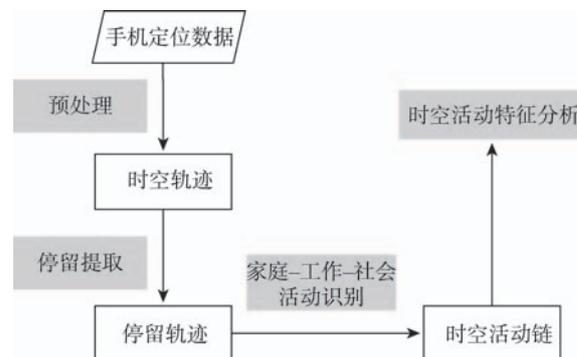


图2 基于大规模手机定位数据的群体活动时空特征分析流程图

Fig. 2 Framework of spatio-temporal analysis of aggregated human activities using cellphone location data

间尺度不在研究范围的数据;④根据每个用户的原始数据分布,剔除用户出现时间少于18h(即数据点序列小于18个)的用户记录。

对于符合标准的数据,按照用户和时间排序,可以得到完整的个体轨迹数据集。一条个体轨迹通常表示为有带地理坐标和时间标记的点序列,如式(1)所示。

$$Tr = \{P_1, P_2, \dots, P_n\} \quad (1)$$

式中: P_i 表示个体的第 i 个位置记录点,如式(2)所示。

$$P_i = (x, y, t) \quad (2)$$

式中: t 表示小时数; x, y 表示手机基站位置坐标。

3.2 时空停留轨迹提取

人们通常在一天中大部分时间处于停留状态,由于活动性质的不同,人们停留的时空属性(停留地点数量、停留时间段和时长)会存在差异^[22]。

停留点是由满足一定时空约束条件的一系列连续的手机位置点所确定。对于时序点轨迹,需要将空间接近和时间邻近的时序点聚类成更为抽象的停留点轨迹。为了简便,每一个抽象停留点的位置由停留点内位置点个数最多的位置确定。考虑到用户 i 的序列点轨迹为 $Tr_i = \{P_1, P_2, \dots, P_n\}$, (n 为位置记录点个数),将第一个时序点 P_1 加入候选停留点轨迹中的第1个停留点;随后计算时序点轨迹中每一个点与已有候选停留点轨迹中的停留点的空间距离。若空间距离小于设定的阈值(本研究设为500m),则将该点加入到候选停留点,否则该点设为新的候选停留点,直到时序轨迹中 n 个位置点全部计算完毕,得到候选停留点轨迹,如式(3)所示。

$$Tr'_i = \{S'_1, S'_2, \dots, S'_m\} \quad (3)$$

式中： S'_i 表示个体的第*i*个停留点，如式(4)所示。

$$S'_i = (x, y, t_{start}, t_{end}) \quad (4)$$

式中： x, y 表示第*i*个停留点的基站位置的坐标； t_{start} 表示该停留开始时间； t_{end} 表示停留结束时间。

对于候选停留点轨迹中的所有停留点，若该点的开始时间与结束时间的差值小于设定阈值(本研究设为1 h)，则认为该点不是真正停留点，将其从候选停留点轨迹中移除，最后得到完整的停留点轨迹。

需要注意的是，考虑到基站空间分布的不均匀性和服务范围的差异以及邻近基站之间信号跳跃的因素，本文利用ArcGIS软件生成手机基站的Voronoi多边形，并计算任意相邻基站距离的分布。如图3所示，约93%的基站距离小于500 m(如图3中红点位置)，因此选择500 m作为空间距离阈值，能最大限度满足不同基站间的差异性。

3.3 家庭-工作-社会活动识别

对于人们一天的日常规律来说，假设人们总是

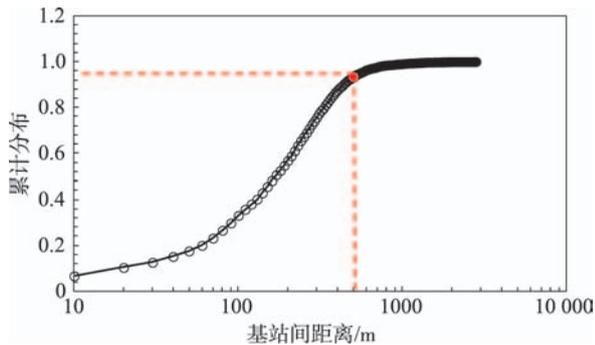


图3 任意相邻基站距离累计分布

Fig. 3 Cumulative distribution of the distance between any adjacent base tower stations

从家出发，通过出行移动到其他位置进行某种活动(如工作)，最后总是会回到家的位置。因此，家庭和工作活动是人类最主要的2类日常活动，且具有很强的时空规律^[23]。而对于其他如购物、休闲、健身等活动，在居民日常生活中发生的比例较少，且规律性不强，在本研究中统一归为社会活动。从用户一天内的停留点轨迹中识别出居家和工作的位置，本研究提出根据用户在一天内停留点的个数对识别过程进行分类处理，如图4所示。基于这样的假设，根据深圳市居民生活习惯，本研究对用户活动的识别窗口时间划分如下：对于居家位置，取夜晚时段(0-7时)为识别窗口；对于工作位置，取白天时段(9-17时)为识别窗口。具体流程如下：

(1)对于一天内只有1个停留点个数的用户，认为该停留点为居家位置。

(2)对于一天内拥有2个停留点的用户，分别将2个停留点的持续时间与2个识别窗口进行匹配，将占比更多的停留点设为居家位置或者工作位置。

(3)对于一天内停留点停留个数大于2的用户，将所有停留点的持续时间分别与2个识别窗口进行匹配，若有停留点的持续时间落在识别窗口内，并占识别窗口时间长度占50%以上，则认为匹配成功，作为候选居家或工作位置；找到匹配时间最长的居家或工作活动位置作为该用户的居家和工作活动位置；若没有匹配成功，则认为该用户没有找到居家或工作活动位置。

(4)对于停留点轨迹中的所有停留点，将居家位置的停留点，标记为家庭活动；将工作位置的停留点，标记为工作活动；没有被标记为居家或工作位置的停留点，则全部标记为社会活动，最后生成完整的时空活动链。

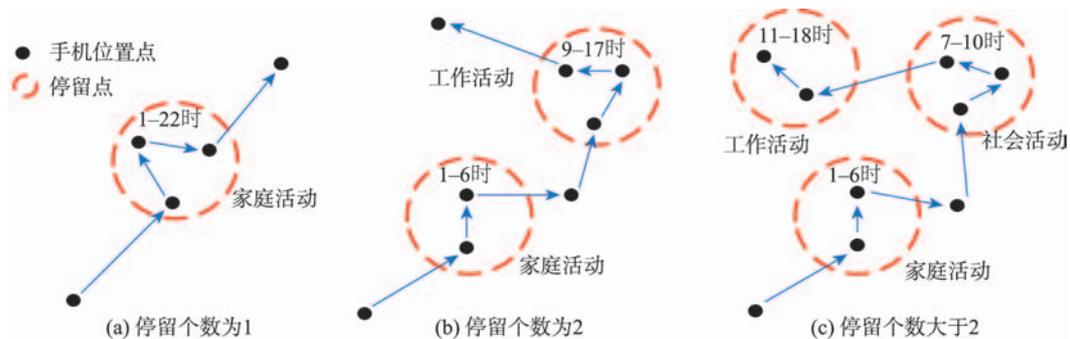


图4 从停留点轨迹中识别家庭-工作-社会活动方法

Fig. 4 Identification of home, work and social activities from stay points trajectory

4 结果与分析

4.1 停留位置与停留活动分析

利用本文所述的手机用户停留轨迹提取和活
动识别方法处理上述约 1000 万用户的手机位置数
据,获得所有用户的停留位置和活
动,进行统计分
析(图 5)。结果表明:用户一天的停留位置数量非

常有限,97.7%的用户不超过 4 个;停留个数为 2 的
用户群体最多,所占比例为 36.8%,个数为 1 和 3 的
次之,所占比例分别为 28.3%和 24%。不同时刻,手
机用户停留的次数存在差异(图 6),夜间时段停留
明显多于白天时段,其中有 3 个明显的停留低谷时
间分别为 7-9 时、11-14 时和 17-19 时,即交通出行
早高峰、午平峰和晚高峰时段。

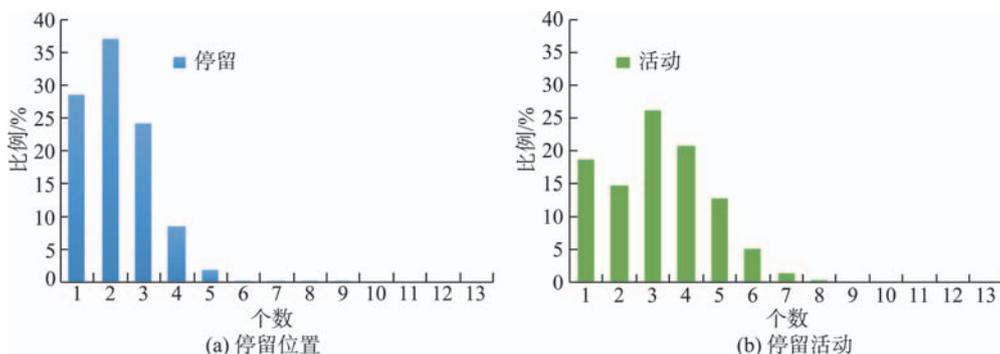


图 5 一天内停留位置和停留活动统计分布

Fig. 5 The statistics of daily stay points and activities

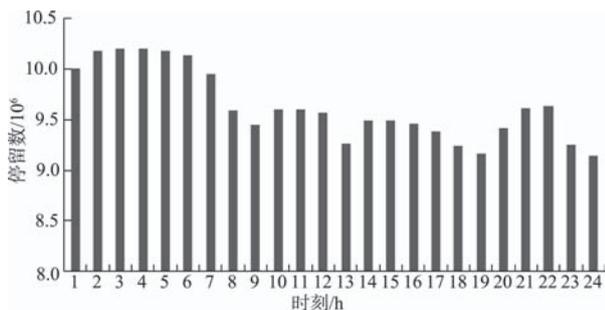


图 6 停留个数不同时间段分布

Fig. 6 Daily temporal distribution of number of stay points

统计所有用户一天内从事的活动个数,实验结
果表明:用户一天进行的活动个数同样有限,约
93%的用户每天进行的活动不超过 5 个;活动个
数为 3 的用户占比最多,达 26%;其次是 4 个和 1 个,分
别达 20.7%和 18.6%。

总体上看,用户停留位置和停留活动在分布上
存在明显差异。每人每天平均停留的位置个数约
为 2.1 个。每人每天平均从事的活动约为 3.4 个。
每人每天平均在一个停留位置不止从事一个活动。

4.2 家庭-工作活动分布特征分析

通过活动识别,提取了全市域一天内的个体活
动约 3436 万个,其中家庭活动约有 1579 万,工作活
动约有 920 万。本研究成功识别出 98%的用户(约

1059 万)的居家位置 and 94%的用户(1015 万)的工作
位置。其中,71%的用户具有不同的居家和工作位
置,21%的用户具有相同的居家和工作位置。有 6%
的用户只识别出了居家位置,2%的用户只有工作
位置。只有 0.3%的用户没有被识别出居家和工作
位置(图 7)。由此可知,平均每人每天在居家位置
从事家庭活动的个数为 1.49 个,说明有一部分人
在一天之中往返于居家位置和其他位置,在居家位置
形成了多次停留。

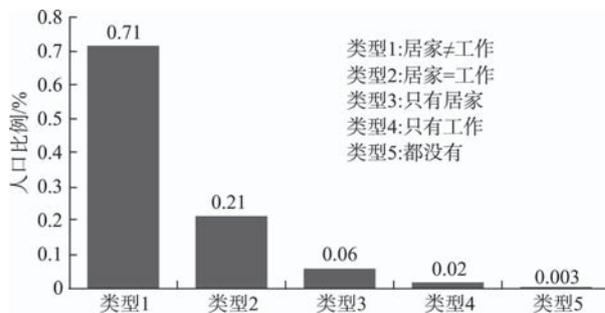


图 7 5 种类型的居家和工作位置识别结果的人口比例

Fig. 7 Percentage of population by five types of detection of home and work locations

为了评价上述方法识别出的居家位置分布的
准确性,将用户的居家位置人口分布与 2010 年街
道级深圳市人口普查数据进行比较(图 8),发现居

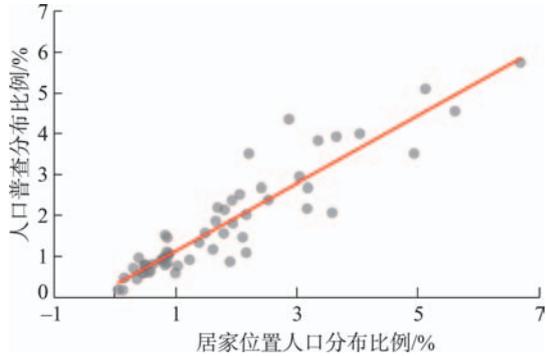


图8 街道级别居家位置人口分布与2010人口普查分布比例相关性分析

Fig. 8 Correlation between the spatial distributions of population based on home locations at street level and the population distribution from 2010 census data

家位置人口分布与真实人口分布,相关系数达到0.92,在置信度水平95%下存在显著相关性。因此,本文方法识别的居家位置分布与人口普查结果在整体上有较好的一致性。

4.3 不同活动时空特征分析

分别进行家庭、工作、社会活动的时空统计,分析不同活动类型在数量强度、空间分布、一天内时间波动上的差异。整体上看,在一天24 h内,群体活动强度存在明显的随时间变化的规律(图9),与人的作息规律基本一致,其中家庭和工作两类活动在24 h变化上呈现“凹凸”现象。其中,0-7时是睡眠时间,此时家庭活动占总活动的比例很高,平均约有87%;从7时开始,人们逐渐离开居家位置,进行其他日常活动,大部分人开始出行,因为出行不构成停留,总活动强度开始减少,并在出行高峰(7-9时、11-14时和17-19时)形成3个明显的低谷;在7时开始,工作活动开始逐步增强,家庭活动开始逐

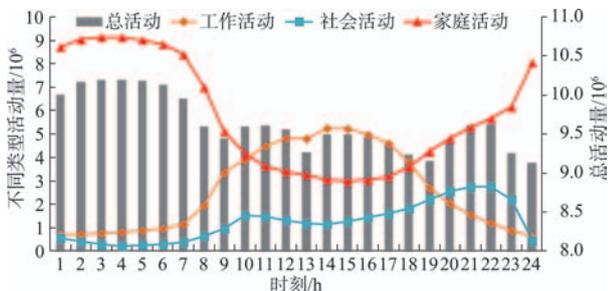


图9 总活动量和工作、社会、家庭活动量不同时间段变化分布

Fig. 9 Daily temporal distribution of the volume of total activity, work activity, social activity and home activity

步减弱,并在10时形成第一个交叉,之后工作活动数量开始高于家庭活动。在18时,大部分人下班返回家中,又形成第二个交叉,之后家庭活动数量又高于工作活动,并且差距逐渐变大。在10-18时,平均约有35%的人进行家庭活动,49%的人进行工作活动。相较于家庭和工作活动,社会活动在一天中处于缓慢波动状态,变化相对不明显;从7时开始缓慢增加,并在22时左右达到高峰,说明活动类型(购物、餐饮、休闲娱乐等)变得丰富,活动越加频繁。

图10展示了总体活动的空间分布。由图10可知,不同行政区的活动密度存在较大差异,原“关内”地区活动密度明显高于原“关外”地区,呈现“南强北弱”的特征。其中,密度最大的是福田区,其次是罗湖区;福田区是深圳的中心商务区,罗湖是深圳最早开发的城区,是深圳人口的主要集聚地,大规模人口的集聚必然会导致活动的集聚和活动类型的丰富。龙岗和宝安由于聚集了大量规模较大的工厂,就业人口较多,因而活动强度也相对较强。

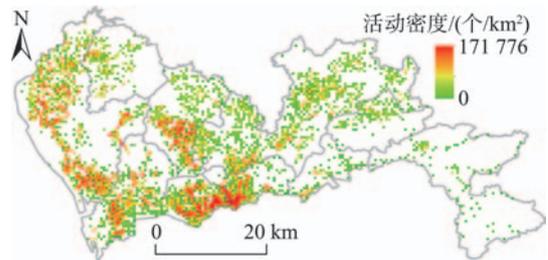


图10 总体活动密度空间分布

Fig. 10 Spatial distribution of activity density

图11为3种不同类型的活动密度的空间分布,不论是家庭活动、工作活动还是社会活动,在整体上都呈现出“空间幂律”的分布特征,即大量活动只分布在少数区域,而大部分区域的活动都很弱,空间分异特征明显。这是由于城市发展历史特点所决定的,也与城市不同区域常住人口分布、产业布局、交通设施规划等相关。

活动密度空间强度的互补累积分布如图12所示,考虑到不同类型的活动密度数值上差异较大,通过除以每类活动的密度均值进行归一化。从图12可看出,不同类型的活动密度空间强度分布非常相似,在双对数坐标系下,呈现为截断幂律分布,说明在空间上,活动密度存在“空间幂律”分布特征。而且,社会活动的分布在尾部下降速度更快,其次是家庭活动和工作活动,说明社会活动的“空间幂律”特征更明显,空间分异更强,而工作活动相对于其他两类活动空间分布上更集中。

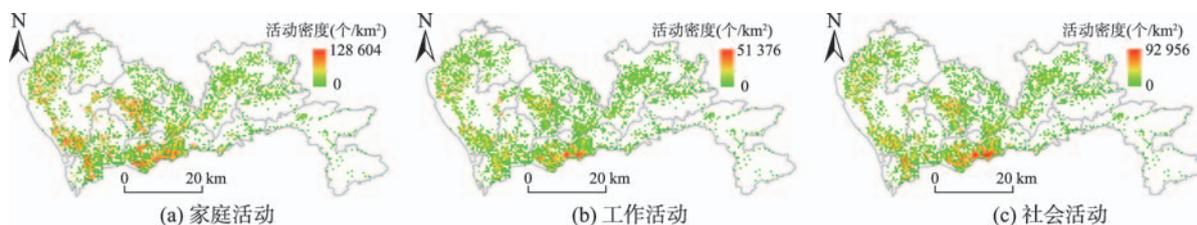


图 11 家庭、工作、社会活动密度空间分布

Fig. 11 Spatial distributions of densities of home, work and social activities

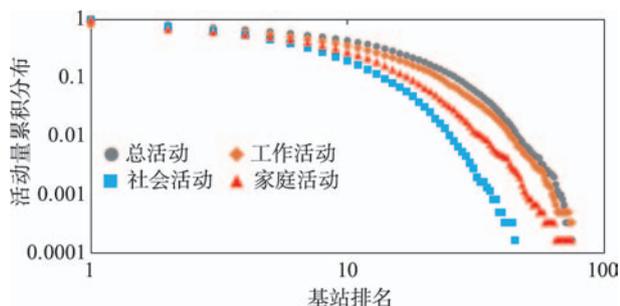


图 12 不同类别活动密度排名的互补累积分布

Fig. 12 Complementary CDF of ranks of activity densities normalized by the mean in different categories

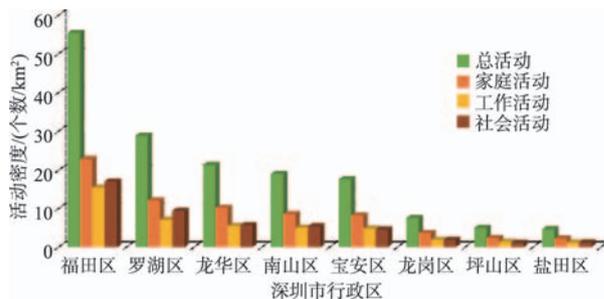


图 14 不同行政区活动密度分布统计

Fig. 14 The statistics of distribution of activity densities in different administrative districts

图 13、14 分别展示了不同行政区的活动总量和活动密度。由图可知,宝安区因其行政面积最大,人口总量也较多,使得活动量较其他行政区最多;而福田区因其面积较小,人口密度大,因而在家庭、工作、社会活动的密度也是最高。盐田区是深圳滨海旅游区和东部港口区域,常住人口少,在该区主要是短暂停留,流动性强,在活动量和活动密度上均为最少。

5 结论

本文基于约 1000 万手机用户在某一工作日的基站尺度的手机定位数据,通过识别用户停留位置

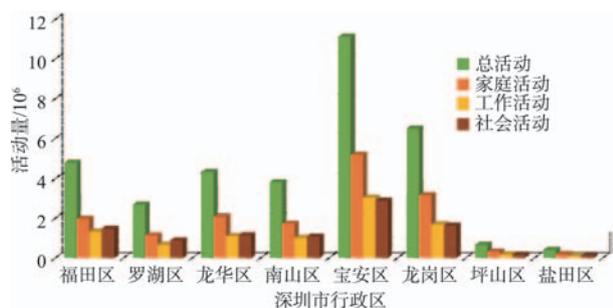


图 13 不同行政区活动量分布统计

Fig. 13 The statistics of distribution of activity volumes in different administrative districts

和停留活动,分析群体活动时空分布模式特征。本文首先通过设定一定时空约束条件得到用户的停留轨迹,再识别出用户在一天内的家庭、工作和社会活动类型,通过该方法识别出 98% 用户的居家位置和 94% 用户的工作位置,最后通过分析不同用户的停留位置和停留活动的统计特征,探讨了用户的停留点和停留活动分布的差异。

基于深圳市手机定位数据的实验结果表明,大量用户在一天内只会进行少量停留和从事少量活动,停留位置和活动分布存在差异,每人每天平均 2.1 个停留,但是每人每天平均从事的活动数却为 3.4 个,表明每人每天平均在一个停留位置不止从事一个活动。通过对不同类型的活动进行时空统计,研究了群体活动的时空分布模式,发现不同的活动在一天内不同时间段存在符合日常作息规律的波动,在空间上由于历史发展、功能产业布局等原因,在不同行政区存在空间分异特征,在整体空间分布上存在“空间幂律”特征。本文结论很好地揭示了城市空间中不同类型的群体活动及其时空分布的多样性和差异性,相较于传统的居民调查等方法,具有样本量大、覆盖范围广、成本低等优势,对于分析深圳市居民出行活动行为、辅助城市交通优化和城市规划具有重要的意义。

手机定位数据由于手机基站位置空间分辨率

低,会导致停留轨迹识别错误的问题;同时,短期采样的数据由于其可重复周期短,对基于时空约束的活动识别结果会存在一定的误差。今后将结合其他来源时空轨迹数据、居民出行调查数据和土地利用数据,对活动识别的准确度进一步进行验证,同时研究不同群体的活动时空分布及其与城市功能结构的相互作用和影响。

参考文献(References):

- [1] Batty M. The size, scale, and shape of cities[J]. *Science*, 2008,319(5864):769-771.
- [2] Kitchin R. The real-time city? Big data and smart urbanism[J]. *GeoJournal*, 2013,79(1):1-14.
- [3] Jiang S. Deciphering human activities in complex urban systems: Mining big data for sustainable urban future[D]. Cambridge MA: Massachusetts Institute of Technology, 2015.
- [4] Noulas A, Scellato S, Lambiotte R, et al. A tale of many cities: Universal patterns in human urban mobility[J]. *PLoS ONE*, 2012,7(5):e37027.
- [5] Zhong C, Arisona S M, Huang X, et al. Detecting the dynamics of urban structure through spatial network analysis [J]. *International Journal of Geographical Information Science*, 2014,28(11):2178-2199.
- [6] Fang Z, 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.
- [7] Chen J, Shaw S L, Yu H, et al. Exploratory data analysis of activity diary data: A space-time GIS approach[J]. *Journal of Transport Geography*, 2011,19(3):394-404.
- [8] Yue Y, Lan T, Yeh A G O, et al. Zooming into individuals to understand the collective: A review of trajectory-based travel behaviour studies[J]. *Travel Behaviour and Society*, 2014,1(2):69-78.
- [9] 李清泉,李德仁.大数据 GIS[J].*武汉大学学报·信息科学版*,2014,39(6):641-644. [Li Q Q, Li D R. Big data GIS [J]. *Geomatics and Information Science of Wuhan University*, 2014,39(6):641-644.]
- [10] Calabrese F, Diao M, Di Lorenzo G, et al. Understanding individual mobility patterns from urban sensing data: A mobile phone trace example[J]. *Transportation Research Part C: Emerging Technologies*, 2013,26:301-313.
- [11] Cao R, Tu W, Cao J, et al. Comparison of urban human movements inferring from multi-source spatial-temporal data[J]. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2016,XLI-B2:471-476.
- [12] 周涛,韩筱璞,闫小勇,等.人类行为时空特性的统计力学[J]. *电子科技大学学报*,2013,42(4):481-540. [Zhou T, Han X P, Yan X Y, et al. Statistical mechanics on temporal and spatial activities of human[J]. *Journal of University of Electronic Science and Technology of China*, 2013, 42(4):481-540.]
- [13] Liu F, Janssens D, Cui J, et al. Characterizing activity sequences using profile Hidden Markov Models[J]. *Expert Systems with Applications*, 2015,42(13):5705-5722.
- [14] Jiang S, Ferreira J J, Gonzalez M C. Discovering urban spatial-temporal structure from human activity patterns [A]. *Proceedings of the ACM SIGKDD International Workshop on Urban Computing*[C]. New York, NY, USA: ACM, 2012:95-102.
- [15] Isaacman S, Becker R, Cáceres R, et al. Identifying important places in people's lives from Cellular Network data [A]. *Pervasive Computing*[M]. Springer Berlin Heidelberg, 2011:133-151.
- [16] Demissie M G, Correia G, Bento C. Analysis of the pattern and intensity of urban activities through aggregate cellphone usage[J]. *Transportmetrica A: Transport Science*, 2015,11(6):502-524.
- [17] Noulas A, Mascolo C, Frías-Martínez E. Exploiting four-square and cellular data to infer user activity in urban Environments[A]. *MDM (1)'13*[C]. 2013:167-176.
- [18] Soto V, Frías-Martínez E. Automated land use identification using cell-phone records[A]. *Proceedings of the 3rd ACM International Workshop on MobiArch*[C]. New York, NY, USA: ACM, 2011:17-22.
- [19] Hasan S, Ukkusuri S V. Urban activity pattern classification using topic models from online geo-location data[J]. *Transportation Research Part C: Emerging Technologies*, 2014,44:363-381.
- [20] 深圳市人民政府公报 2016 年第 39 期(总第 983 期)[R]. <http://www.sz.gov.cn/zfgb/2016/gb983/>.
- [21] 深圳市 2015 年全国 1% 人口抽样调查主要数据公报 [EB/OL].http://www.szjtj.gov.cn/xxgk/tjsj/tjgb/201606/t20160614_3697000.htm.
- [22] 徐金垒,方志祥,萧世伦,等.城市海量手机用户停留时空分异分析——以深圳市为例[J].*地球信息科学学报*, 2015,17(2):197-205.[Xu J L, Fang Z X, Shaw S L, et al. The spatio-temporal heterogeneity analysis of massive urban mobile phone users' stay behavior: A case study of Shenzhen city[J]. *Journal of Geo-Information Science*, 2015,17(2):197-205.]
- [23] 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.