

个体时空行为的规律性与可预测性研究 ——以上海市居民工作日活动为例

王 德¹, 谢栋灿², 王 灿^{1,3}, 殷振轩¹

(1. 同济大学建筑与城市规划学院, 上海 200092; 2. 上海城策行建筑规划设计咨询有限公司, 上海 200003;

3. 麻省理工学院媒体实验室, 美国 剑桥 02139)

摘 要:理解居民日常行为的规律性和随机性是空间行为研究从行为描述向行为预测推进的基础,对空间政策的制定具有重要意义。论文从上海市10个工作日的手机信令数据中提取个体的时空轨迹,从规律性的视角展开多方面分析。首先,根据活动的周期重现、前后联系等特征界定了无出行(包括仅有短距离出行)、仅有随机出行、有规律工作出行和有其他规律出行4种时空行为模式,发现后2种强规律性模式的人数、出行次数、出行距离的占比分别为49.3%、83.6%、82.0%。然后,测度了个体时空轨迹的熵值,推算出个体时空位置的平均可预测性高达0.84。这些结果均表明居民工作日的时空行为整体上具有非常强的规律性。在此基础上,论文还分析了不同行为模式、个体时空位置可预测性的空间分布,发现了强规律性模式人群在外围郊区密集分布等一系列特征,体现了城市空间环境的差异对行为特征和规律性的影响。

关 键 词:规律性;可预测性;弹性出行;手机信令数据;上海市

城市居民的日常行为一直是地理学和城市规划的研究重点^[1],是“以人为本”原则下探讨空间问题必不可少的视角。个体行为活动是自身活动需求和城市空间环境约束相互作用的结果,反过来,这些行为又时刻影响着城市空间的动态结构和环境品质。基于这样的相互关系,空间行为研究者通过活动分析法^[2]寻找空间与行为的主要矛盾,旨在通过合理的规划政策引导空间与行为相适应,改善空间环境品质。而空间政策的制定需要把握居民未来的活动状态,识别行为中具有规律性、系统性的成分,并预测政策对行为的影响。因此,本文希望从规律性的视角入手,深化对个体行为特征的认识。

随着手机的普及和移动定位技术的发展,居民日常行为的时空信息以被动的方式被记录下来,所得到的手机信令大数据具有接近总体的采样范围,为研究者呈现了更加全面的时空行为图景^[3]。特别地,手机信令数据每天24 h都在连续不断地记

录,具有无可比拟的全时动态优势。通过在较长时间段内持续追踪同一设备的记录,研究者可以在个体层面获得较完整的时空行为轨迹。因此,手机信令数据适合用于对个体时空行为规律性的分析。

目前国内基于手机信令数据的空间行为研究已经覆盖了居民的各种日常行为,包括通勤^[4-6]、购物^[7]、游憩^[8]等。但是,这些研究大多使用空间单元的集计数据,而对个体层面时空轨迹的利用有限,而后者才是研究行为规律性更适宜的数据形式。另外,一些基于传统小数据的研究利用时空轨迹数据分析了居民的频繁活动序列^[9]、活动空间^[10]、时空集聚模式^[11]等特征,该方面主要的成果包括杜洋等^[12]基于GPS轨迹(709个样本,90 d)对活动可预测性的测度,以及申悦等^[13]基于GPS轨迹和活动日志(100个样本,7 d)对通勤弹性的分析,但小数据在样本代表性上存在天然不足,且时空行为的规律性还包括更广泛的内容。

收稿日期:2020-03-30;修订日期:2020-08-13。

基金项目:国家自然科学基金项目(41771170)。[Foundation: National Natural Science Foundation of China, No. 41771170.]

第一作者简介:王德(1963—),男,教授、博士生导师,主要研究方向为城市规划方法论、空间与行为、城市大数据、城市模型。

E-mail: dewang@tongji.edu.cn

引用格式:王德,谢栋灿,王灿,等.个体时空行为的规律性与可预测性研究:以上海市居民工作日活动为例[J].地理科学进展,2021,40(3): 433-440. [Wang De, Xie Dongcan, Wang Can, et al. Regularity and predictability of individual spatiotemporal behavior: A case study of residents working day activities in Shanghai. Progress in Geography, 2021, 40(3): 433-440.] DOI: 10.18306/dlkxjz.2021.03.007

本文以手机信令多日连续记录的个体时空行为轨迹为数据,基于居民时空活动的规律性,划分居民行为模式,识别居民弹性出行,测算个体位置的可预测性,旨在深化行为的规律性研究。

1 研究数据与方法

1.1 研究数据

本文采用的数据取自上海市2014年连续2周的中国移动2G手机信令数据,共记录2400万用户的82.7亿条信令记录。首先通过对用户在夜间时段(20:00~6:00)和白天工作时段(9:00~18:00)平均位置的多日汇总,识别居住地和工作地^[4],其中,1371万用户可识别稳定的居住地和工作地。进一步,筛选10个工作日均有记录,且活动时间(6:00~24:00)记录完整性在80%以上的用户,共计237万^①。但是这些记录完整样本的空间分布与上海市居民的实际分布存在一定的偏差。为了纠正偏差,研究以街道为单元,以每个街道所有可识别居住地用户数量的4%为目标样本量,对该街道的记录完整样本进行抽样,共得到53.9万用户样本^②,其居住地分布与第六次人口普查人口分布基本吻合,具有较好的代表性。

1.2 研究方法

本文采用2种方法对个体行为的规律性进行探讨。一方面,个体的时空行为被划分为4种典型模式,各模式的命名和意义直观体现了规律性;另一方面,研究量化测度了每个个体时空位置的可预测性,作为反映规律性强弱的数值指标。2种方法都是基于活动在空间上的复现特征,但是角度和侧重点不同:行为模式的分类推断具有更强的主观性和可解释性,侧重于对“什么样的规律”这一问题的认识,为此考虑了居住地、工作地等空间细节信息;可预测性则是客观的数学计算,侧重于对“规律性有

多强”这一问题的认识,为了提高精确性,还考虑了时间上的复现特征。

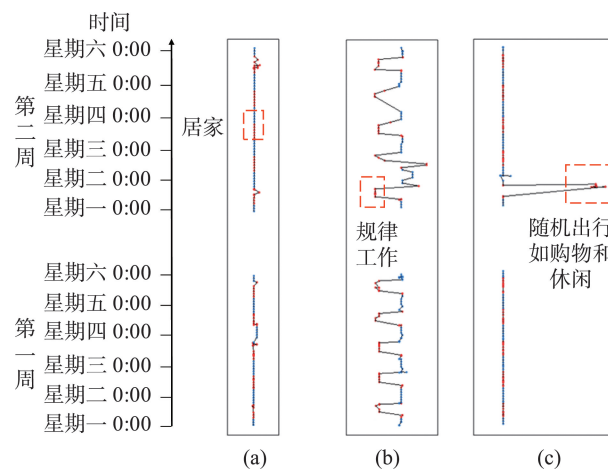
1.2.1 基于规律性视角的时空行为模式划分

虽然手机信令数据缺乏活动属性信息,但基于个体多日连续活动的时空特征,可以对活动做出合理的推断,包括居家、规律性工作和随机出行(购物、休闲等)(图1)。本文基于空间位置和规律性,将时空行为分为4类:居家、规律工作、其他规律出行和随机出行(表1),其识别流程如图2。

进一步,基于个体在连续10个工作日内,是否发生过规律工作出行、其他规律出行、随机出行,将其划分为无出行(包括仅有1 km内的短距离出行,下同)、仅有随机出行、有规律工作出行和有其他规律出行4类时空行为模式(表2)。

1.2.2 基于规律性视角的弹性出行判断和可预测性测算

出行弹性是城市交通研究中的重要概念,其描述了居民在出行决策中的自由度和选择偏好^[15]。以往的出行弹性研究主要侧重于交通方式选择的



注:图中红色点表示白天(6:00~18:00)记录,蓝色点表示夜间(18:00~6:00)记录。

图1 几种典型的时空活动类型示例

Fig.1 Examples of typical modes of actives

表1 典型活动类型的时间和空间特征

Tab.1 Spatiotemporal features of typical modes of activities

项目	居家	规律工作	其他规律出行	随机出行
时间	多天	多天;白天居多	多天	低频
空间	识别出的居住地一定范围内	识别出的工作地一定范围内	离家和工作地一定距离	离家和工作地一定距离

① 由于本文研究的对象是个体连续活动的时空轨迹,因此对数据记录的完整性要求较高。为此,基于2G手机信令数据的特性,研究以2 h为一个时间段,统计用户10个工作日记录的完整性。本文将记录较为完整的用户定义为10 d内均有记录,且活动时间(6:00~24:00)记录的完整性在80%以上,即至少72个时间段有记录的用户。

② 抽样比例为4%的确定依据是:在228个街道中,有6个街道的数据质量相对较差,居住在这6个街道记录完整性在80%以上的用户比例不足4%,其余街道均在4%以上;同时,对比国内外相关研究的数据量^[12-13],50万个以上的样本已经能较好地支撑个体层面时空行为的研究。

自由度,较少关注居民在出行时间以及目的地决策的自由度,本文从规律性视角,基于受时空约束的程度,重新界定了出行弹性概念,识别了刚性和弹性出行。具体而言,某个体的一条出行记录 (D_i, H_i, L_i) 代表了其在 D_i 天的 H_i 时段到访了位置点 L_i ,对于不同天的另一条记录 (D_j, H_j, L_j) ($j \neq i$),若 H_i 与 H_j 相距2 h以内且 L_i 与 L_j 相距1000 m以内,则表明该个体在这两天的相近时间段到访了相近的目的地。如果这样的出行记录出现了5 d以上,则认为这些出行的选择自由度较小,判

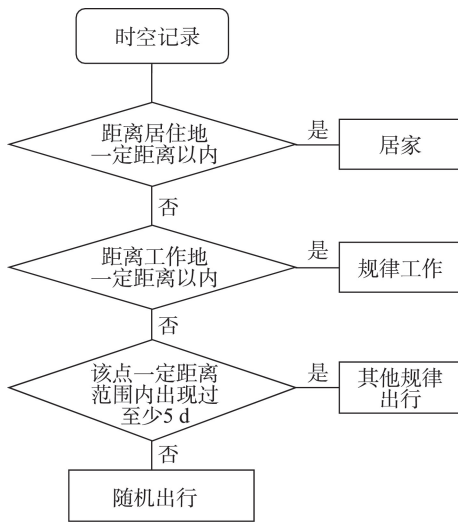


图2 活动类型识别的主要流程

Fig.2 Recognition process of single-activity mode

表2 4类时空行为模式的识别分类标准

Tab.2 Recognition criteria of four kinds of behavioral modes

时空路径模式	规律工作	其他规律出行	随机出行
无出行	无	无	无
仅有随机出行	无	无	有
有规律工作出行	有	有/无	有/无
有其他规律出行	无	有	有/无

表3 4类时空行为模式的构成

Tab.3 Composition of four kinds of behavioral modes

时空行为模式	用户数量		出行次数		出行距离		
	总量	占比/%	占比/%	人均/次	占比/%	人均/km	次均/km
无出行	102830	19.1	0.2	0.19	0.1	0.45	2.41
仅有随机出行	170096	31.6	16.1	8.80	17.9	64.15	7.29
有规律工作出行	196509	36.5	58.2	27.50	57.1	176.68	6.43
有其他规律出行	69217	12.8	25.4	34.06	24.9	218.93	6.43

③ 例如,可以简单地把个体在特定时段最常去的场所作为预测结果,也可以使用复杂的深度学习算法进行预测,后者将更好地考虑前后活动之间的复杂联系。显然,它们所得到的准确率将有差异。

定为刚性出行,反之则判定为弹性出行。因此,弹性出行的占比越高,个体行为的规律性越弱;反之,弹性出行的占比越低,个体行为的规律性越强。

可预测性是指在已知个体历史时空轨迹的条件下,预测其时空位置的准确率,个体行为的规律性越强,可预测性越高。需要注意的是,准确率很大程度上受到预测算法^③的影响,在数据科学快速发展的背景下,选择任何一种特定算法进行预测模拟都不具有充分的代表性。对此,有学者提出了基于熵值计算可预测性上界的方法^[12,16-19],其所得到的结果具有理论上最高的准确率。本文将上海市空间单元栅格化,并以2 h为间隔统计每位个体所在栅格的编号,得到位置序列。个体 i 的活动熵值 S_i 可由下式计算:

$$S_i = - \sum_{T'_i \in T_i} P(T'_i) \log_2 [P(T'_i)] \quad (1)$$

式中: T_i 是 i 的位置序列, T'_i 是 T_i 中所有可能的子序列; $P(T'_i)$ 为子序列 T'_i 在序列 T_i 中出现的概率。轨迹熵越小,则个体在固定地点之间出现占比越高,任意一段活动链的平均不确定性越小。在此基础上,根据下式所示的熵与可预测性的关系,推算个体 i 的可预测性上界 Π_i^{\max} :

$$S_i = -\Pi_i^{\max} \log_2 (\Pi_i^{\max}) - (1 - \Pi_i^{\max}) \log_2 (1 - \Pi_i^{\max}) + (1 - \Pi_i^{\max}) \log_2 (N_i - 1) \quad (2)$$

式中: N_i 是 i 活动过的位置的总数。

2 时空行为模式

2.1 时空行为模式构成

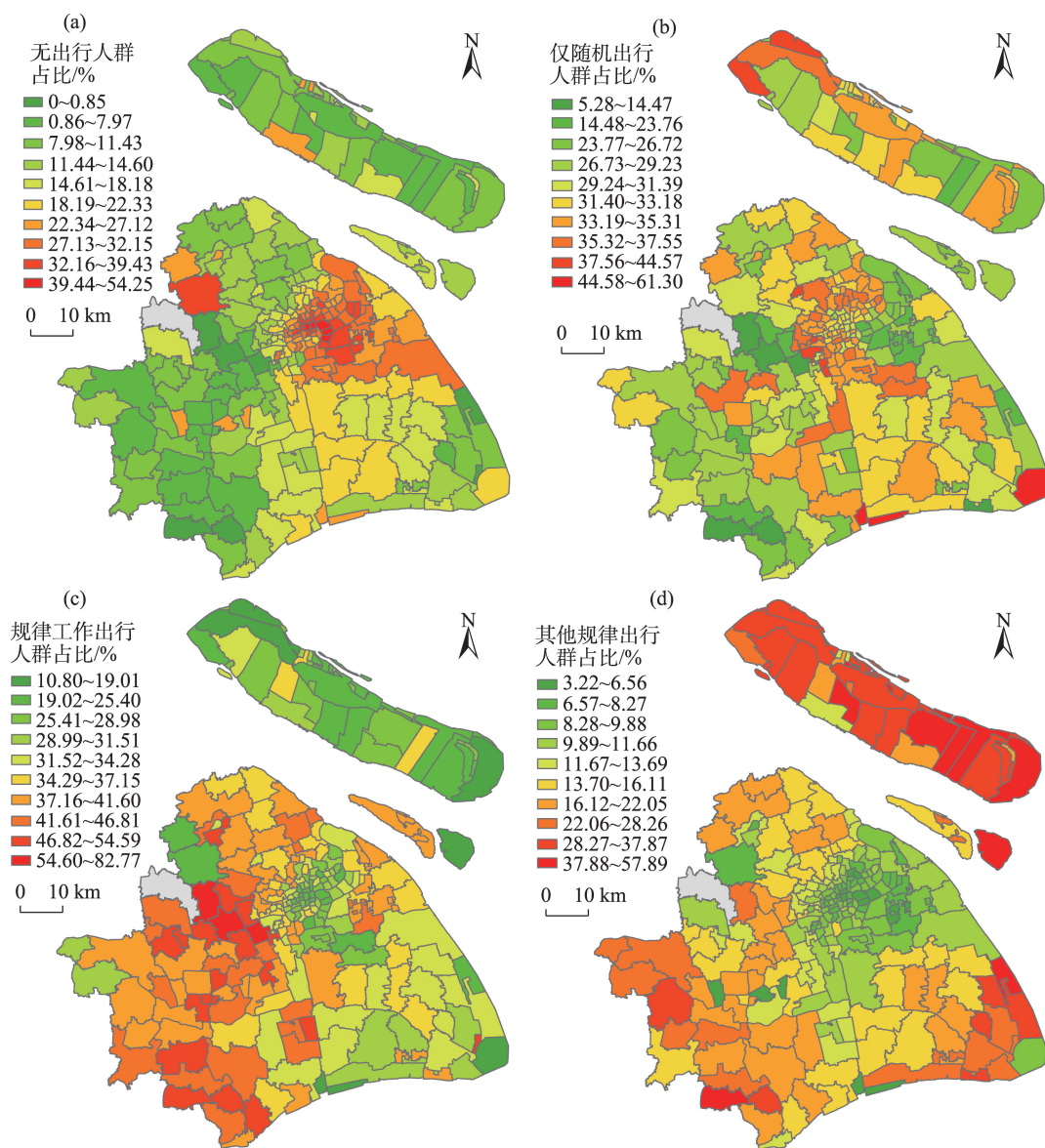
基于时空行为模式划分方法,将53.9万用户分为4类时空行为模式人群,其用户数量、出行次数和出行距离如表3所示。可以发现,有规律工作出行(36.5%)和仅有随机出行(31.6%)的用户占比较高,

而无出行(19.1%)和有其他规律出行(12.8%)的用户占比较低。进一步,可以发现各类人群的出行次数和出行距离具有较大差异,对城市公共空间和设施(如交通设施、公共服务设施等)的影响程度不同:(1)无出行人群的出行次数和出行距离均远小于其他人群,他们主要在居住地附近活动,对城市的影响相对较小;(2)有规律工作出行和有其他规律出行的人群具有最高的人均出行次数和人均出行距离,其活动具有较强的规律性,对城市的影响最大;(3)仅有随机出行的人群虽然具有最高的次均出行距离,但其人均出行次数和人均出行距离明显低于

两类有规律出行人群,对城市的影响有限。综合来看,有规律出行人群(包括有规律工作出行和有其他规律出行)的占比为49.3%,尽管未满一半,但是他们产生了高达83.6%的出行次数和82.0%的出行距离。从这个意义上可以认为,大部分的居民活动具有较强的规律性,构成了城市运行常态的主体。

2.2 各类人群的空间分布

图3显示了4种时空行为模式人群以居住地汇总在各街道中所占的比例。可以看到,不同空间的人群构成具有明显的异质性。中心城区和浦东新区主要以无出行人群为主,内外环之间的区域以仅



注:在嘉定区安亭镇出现了一块飞地式的高值区(灰色显示),但是该地区由于数据本身的问题,计算结果不作讨论分析。图6同。

图3 不同时空行为模式人群的空间分布

Fig.3 Spatial distribution of populations with different behavioral modes

有随机出行人群的占比较高,近郊区主要以有规律工作出行人群为主,而远郊区主要以有规律工作出行和有其他规律出行人群为主。这一结果反映了设施配置、职住关系、老年人比例等社会经济属性的空间差异:(1) 中心城区聚集了大量就业岗位和公共服务设施,居民有更高的机会就近通勤和获得服务,而距离小于1000 m的出行包括在无出行内;同时,中心城区的老年人比例较高,而老年人也是无出行人群的代表。(2) 内外环之间区域同时受中心城区和郊区副中心的影响,用户随机出行的概率较高。(3) 近郊区域职住分布不平衡,对中心城区有较高的依赖性;同时,部分远郊区域受松江、金山、嘉定等新城和工业区的影响,集聚了大量的有规律工作出行人群。(4) 远郊青浦、金山、浦东新区边缘和崇明等区域缺乏高等级公共服务设施,也可能与远郊居民不同的生活方式(如第二居所)有关,同样对中心城区有一定依赖;同时,在一定程度上受到基站密度下降所导致的数据本身的影响,因此有其他规律出行人群占比较高。

3 时空位置的可预测性

3.1 可预测性统计

采用1.2.2中的规则识别居民的刚性和弹性出行,发现工作日居民的活动受到显著的时空约束,67%的出行行为是刚性出行,而弹性出行行为仅占33%。图4显示了刚性出行的1 d内时间分布特征,可见刚性出行占比呈现了显著的昼夜差距,白

天(8:00~18:00)出行总量大,刚性出行占比大,约占75%左右;而夜间(22:00~6:00)出行总量小,刚性出行占比小,不足50%。可以发现,工作日居民白天对城市运行影响更大的活动主要是刚性出行,在相似的时间到达固定的地点,具有较强的规律性。

进一步,采用1.2.2中的方法测度每位居民时空轨迹的熵值和位置可预测性,定量计算个体时空行为的规律性(图5),与同类研究相似^[12,17,19],本文发现居民的工作日的时空行为具有非常强的规律性,在53.9万用户中,29.6%的个体的可预测性在0.9以上,57.7%的个体的可预测性在0.8~0.9之间,只有3.6%的个体的可预测性小于0.7,所有样本的平均可预测性达到0.84,即个体的时空位置不能由历史记录预测的成分平均只有0.16左右。究其原因,居民工作日的出行受到较强的时空约束,需要在特定时间出现在固定地点,从而大大降低了个体移动位置预测的难度。在合适的算法支持下,研究可以相

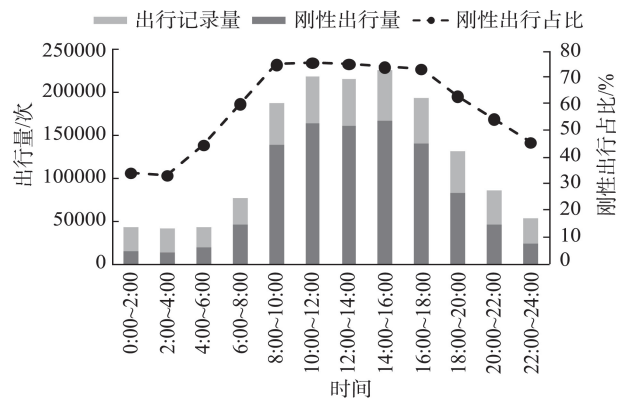


图4 刚性出行1 d内时间分布特征(以星期二为例)

Fig.4 Temporal distribution of inelastic trips

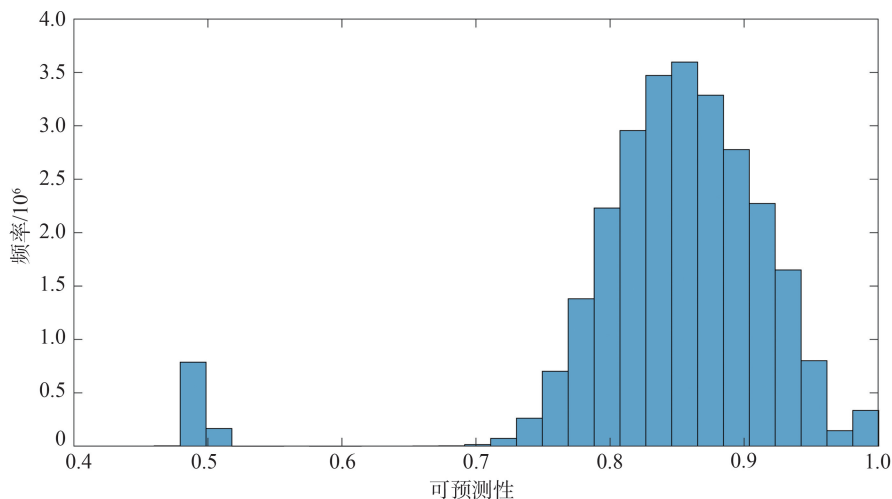


图5 可预测性的频率分布

Fig.5 Frequency distribution of predictability

当准确地把握居民未来的活动状态。

3.2 可预测性的空间分布

以居住地汇总,统计各街道个体时空位置的平均可预测性(图6),发现全部街道的平均可预测性均较高,超过0.8。其中,中心城区个体时空位置的平均可预测性相对较高;而外环附近的近郊圈层个体时空位置的平均可预测性相对较低;之后随着与中心城区距离的增加,个体时空位置的平均可预测性略微增加。值得注意的是,近郊和远郊主要以2类规律性较强的人群(有规律工作出行和有其他规律出行)分布为主,在他们的日常出行中,刚性较强的通勤占有较大的比例。所不同的是,近郊区居民受中心城区强大的商业休闲吸引力影响,产生了大量的弹性出行。图7显示了弹性出行目的地的空间分布,人民广场、南京路、五角场、徐家汇等内环内生活中心的密度最高。由此可以推测,近郊和远郊的弹性出行主要是往返中心城区的生活出行,随着与中心城区距离的增加,远郊区的此类出行受到抑制,个体时空位置的平均可预测性增加。

4 结论与讨论

行为的规律性是时间地理学,尤其是居民时空

行为研究的重要内容,通过分析居民行为的规律性和可预测性,可以为制定空间政策、构建行为预测模型、界定行为模式提供帮助^[16,20]。本文研究突破传统研究的局限,借手机信令数据的大样本、高精度的特性,以居民时空行为的规律性为视角,基于个体时空轨迹的分析,界定了4种活动类型——居家、规律工作、其他规律出行、随机出行,以及由之构成的4种行为模式——无出行、仅有随机出行、有规律工作出行、有其他规律出行,并利用熵值法定量测度了个体时空位置可预测性的上界及其时间、空间上的分布特征,验证了居民时空行为规律性的强弱,为中国大城市(尤其是上海市)居民时空行为的研究和空间政策的制定提供了一定的研究基础。

研究发现,首先,个体工作日在城市空间的活动具有很强的规律性:(1) 83.6%的出行次数和82.0%的出行距离是由49.3%的有规律工作出行和有其他规律出行的人群产生;(2) 在已知历史活动轨迹的情况下,个体时空位置具有很高的可预测性,平均可预测性高达0.84,大部分活动都是在相似场所、相似时刻持续重复出现的刚性行为。基于行为的规律性,空间政策的研究者可以对居民的将来行为状态有较好的把握,获得居民时空行为的常

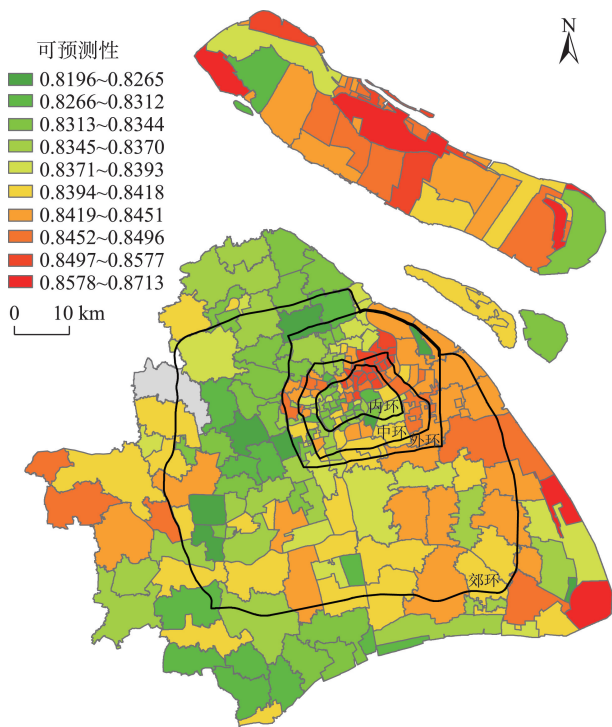


图6 可预测性的空间分布

Fig.6 Spatial distribution of predictability

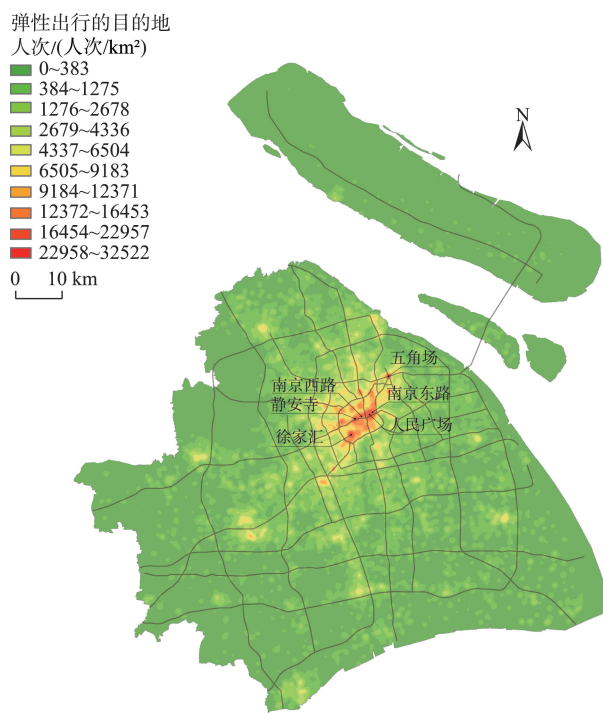


图7 弹性出行的目的地

Fig.7 Destinations of elastic trips

态;同时,强规律性有助于研究者对政策实施后的行为变化进行模拟,从而评估政策效果,使其具有更强的可预见性。

其次,不同时空行为模式人群和弹性出行占比的空间分布具有差异,中心城区以无出行和仅有随机出行人群为主,近郊以有规律工作出行人群为主,个体时空位置的平均可预测性相对较低,远郊以有规律出行和有规律工作出行为主,反映了设施配置、职住关系、老年人比例等社会经济属性的空间差异,可以用于指导城市规划和管理。

本文虽然论证了居民日常行为的强规律性,并对可预测性的上限进行了量化,但受合适算法的限制,并未尝试真正的预测模拟。在基于大数据的时空行为研究不断深入的过程中,从行为描述到行为预测的方法论跨越已经引起越来越多的关注,实现这种跨越无疑可以将大数据的价值提升到新的高度。另一方面,对于所划分的4种时空行为模式的具体行为规律,本文仅探讨了出行距离、出行强度等特征,更多特征有待于进一步深入探讨。

参考文献(References)

- [1] 章雨晴,甄峰,张永明.南京市居民网络购物行为特征:以书籍和衣服为例[J].地理科学进展,2016,35(4): 476-486. [Zhang Yuqing, Zhen Feng, Zhang Yongming. Characteristics of e-shopping behavior of Nanjing residents: A case of books and clothes. Progress in Geography, 2016, 35(4): 476-486.]
- [2] 柴彦威,沈洁.基于活动分析法的人类空间行为研究[J].地理科学,2008,28(5): 594-600. [Chai Yanwei, Shen Jie. Activity-based approach to human spatial behavior research. Scientia Geographica Sinica, 2008, 28(5): 594-600.]
- [3] 杨振山,龙瀛, Douay N. 大数据对人文—经济地理学研究的促进与局限[J].地理科学进展,2015,34(4): 410-417. [Yang Zhenshan, Long Ying, Douay N. Opportunities and limitations of big data applications to human and economic geography: The state of the art. Progress in Geography, 2015, 34(4): 410-417.]
- [4] 李祖芬,于雷,高永,等.基于手机信令定位数据的居民出行时空分布特征提取方法[J].交通运输研究,2016,2(1): 51-57. [Li Zufen, Yu Lei, Gao Yong, et al. Extraction method of temporal and spatial characteristics of residents' trips based on cellular signaling data. Transport Research, 2016, 2(1): 51-57.]
- [5] 丁亮,钮心毅,宋小冬.上海中心城就业中心体系测度:基于手机信令数据的研究[J].地理学报,2016,71(3): 484-499. [Ding Liang, Niu Xinyi, Song Xiaodong. Measuring the employment center system in Shanghai central city: A study using mobile phone signaling data. Acta Geographica Sinica, 2016, 71(3): 484-499.]
- [6] 田金玲,王德,谢栋灿,等.上海市典型就业区的通勤特征分析与模式总结:张江、金桥和陆家嘴的案例比较[J].地理研究,2017,36(1): 134-148. [Tian Jinling, Wang De, Xie Dongcan, et al. Identifying the commuting features and patterns of typical employment areas in Shanghai using cellphone signaling data: A case study in Zhangjiang, Jinqiao and Lujiazui. Geographical Research, 2017, 36(1): 134-148.]
- [7] 王德,王灿,谢栋灿,等.基于手机信令数据的上海市不同等级商业中心商圈的比较:以南京东路、五角场、鞍山路为例[J].城市规划学刊,2015(3): 50-60. [Wang De, Wang Can, Xie Dongcan, et al. Comparison of retail trade areas of retail centers with different hierarchical levels: A case study of East Nanjing Road, Wujiaochang, Anshan Road in Shanghai. Urban Planning Forum, 2015(3): 50-60.]
- [8] 方家,王德,谢栋灿,等.上海顾村公园樱花节大客流特征及预警研究:基于手机信令数据的探索[J].城市规划,2016,40(6): 43-51. [Fang Jia, Wang De, Xie Dongcan, et al. Research on dynamic change and early warning of large tourist flow based on mobile signal data analysis: A case study of Gucun Park sakura festival in Shanghai. City Planning Review, 2016, 40(6): 43-51.]
- [9] 李雄,马修军,王晨星,等.城市居民时空行为序列模式挖掘方法[J].地理与地理信息科学,2009,25(2): 10-14. [Li Xiong, Ma Xiujun, Wang Chenxing, et al. Data mining method for exploring spatio-temporal behavior sequence patterns of urban residents. Geography and Geo-Information Science, 2009, 25(2): 10-14.]
- [10] 申悦,柴彦威.基于GPS数据的北京市郊区巨型社区居民日常活动空间[J].地理学报,2013,68(4): 506-516. [Shen Yue, Chai Yanwei. Daily activity space of suburban mega-community residents in Beijing based on GPS data. Acta Geographica Sinica, 2013, 68(4): 506-516.]
- [11] 古杰,周素红,闫小培,等.居民日常出行时空集聚视角下的城市交通拥堵形成机制研究:以广州为例[J].地理科学,2012,32(8): 921-927. [Gu Jie, Zhou Suhong, Yan Xiaopei, et al. Formation mechanism of traffic congestion in view of spatio-temporal agglomeration of residents' daily activities: A case study of Guangzhou. Scientia Geographica Sinica, 2012, 32(8): 921-927.]
- [12] 杜洋,柴彦威,杨家文,等.基于GPS数据的北京市居民活动规律性与可预测性研究[J].地理与地理信息科学,2015,31(6): 47-51. [Du Yang, Chai Yanwei, Yang Ji-

- awen, et al. Predictability of resident activity in Beijing based on GPS data. *Geography and Geo-Information Science*, 2015, 31(6): 47-51.]
- [13] 申悦, 柴彦威. 基于GPS数据的城市居民通勤弹性研究: 以北京市郊区巨型社区为例 [J]. *地理学报*, 2012, 67(6): 733-744. [Shen Yue, Chai Yanwei. Study on commuting flexibility of residents based on GPS data: A case study of suburban mega-communities in Beijing. *Acta Geographica Sinica*, 2012, 67(6): 733-744.]
- [14] 谢栋灿, 王德, 钟炜菁, 等. 上海市建成环境的评价与分析: 基于手机信令数据的探索 [J]. *城市规划*, 2018, 42(10): 97-108, 120. [Xie Dongcan, Wang De, Zhong Weijing, et al. Assessment and analysis on built environment of Shanghai: An exploration based on mobile phone signaling data. *City Planning Review*, 2018, 42(10): 97-108, 120.]
- [15] 云美萍, 张元, 周源, 等. 出行弹性概念及其应用 [J]. *同济大学学报(自然科学版)*, 2013, 41(9): 1359-1365. [Yun Meiping, Zhang Yuan, Zhou Yuan, et al. Conception of travel elasticity and its application. *Journal of Tongji University (Natural Science)*, 2013, 41(9): 1359-1365.]
- [16] 康朝贵, 刘瑜, 邬伦. 城市手机用户移动轨迹时空熵特征分析 [J]. *武汉大学学报(信息科学版)*, 2017, 42(1): 63-69, 129. [Kang Chaogui, Liu Yu, Wu Lun. An analysis of entropy of human mobility from mobile phone data. *Geomatics and Information Science of Wuhan University*, 2017, 42(1): 63-69, 129.]
- [17] Song C M, Qu Z H, Blumm N, et al. Limits of predictability in human mobility [J]. *Science*, 2010, 327: 1018-1021.
- [18] Lu X, Bengtsson L, Holme P. Predictability of population displacement after the 2010 Haiti earthquake [J]. *PNAS*, 2012, 109(29): 11576-11581.
- [19] Lu X, Wetter E, Bharti N, et al. Approaching the limit of predictability in human mobility [J]. *Scientific Reports*, 2013, 3: 2923. doi: 10.1038/srep02923.
- [20] Goulet-Langlois G, Koutsopoulos H N, Zhao Z, et al. Measuring regularity of individual travel patterns [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2018, 19(5): 1583-1592.

Regularity and predictability of individual spatiotemporal behavior: A case study of residents working day activities in Shanghai

WANG De¹, XIE Dongcan², WANG Can^{1,3}, YIN Zhenxuan¹

(1. College of Architecture and Urban Planning, Tongji University, Shanghai 200092, China;

2. Shanghai Urban Architectural Planning & Design Consultancy Co., Ltd., Shanghai 200003, China;

3. MIT Media Lab, Cambridge 02139, USA)

Abstract: Understanding the regularity and randomness of residents' daily activities is fundamental for spatial behavior studies when the focus is turned from description to prediction. In this study, behavioral regularity was explored from multiple aspects based on individual spatiotemporal trajectories, which were extracted from 10 working days' mobile signaling data in Shanghai Municipality. First, spatiotemporal behavior was recognized as four modes—non-trips, random trips only, regular commutes, and other regular trips—based on periodical recurrency and time series connection. The latter two modes with strong regularity account for 49.3% in population, 83.6% in trip generations, and 82.0% in trip distance. Second, entropies of trajectories were calculated at the individual level, from which a 0.84 potential predictability of locations could be inferred. All these findings arrive at a conclusion that residents' spatiotemporal behavior in working days is strongly regular. Furthermore, the article discussed the spatial heterogeneity of behavioral modes and predictability of locations, including characteristics of spatial distributions, as well as underlying influencing factors from the environment.

Keywords: regularity; predictability; elastic trips; mobile signaling data; Shanghai