



基于手机数据的 城市居民通勤出行特征 识别及仿真研究

夏小棠 著



WUHAN UNIVERSITY PRESS
武汉大学出版社



基于手机数据的 城市居民通勤出行特征 识别及仿真研究

夏小棠 著

图书在版编目(CIP)数据

基于手机数据的城市居民通勤出行特征识别及仿真研究/夏小棠著. —武汉: 武汉大学出版社, 2018. 5

ISBN 978-7-307-20125-5

I. 基… II. 夏… III. 城市交通—交通运输管理—研究
IV. U491

中国版本图书馆 CIP 数据核字(2018)第 068308 号

责任编辑: 黄 殊 责任校对: 李孟潇 整体设计: 马 佳

出版发行: 武汉大学出版社 (430072 武昌 珞珈山)
(电子邮件: cbs22@whu.edu.cn 网址: www.wdp.com.cn)

印刷: 虎彩印艺股份有限公司

开本: 720 × 1000 1/16 印张: 11.5 字数: 164 千字 插页: 1

版次: 2018 年 5 月第 1 版 2018 年 5 月第 1 次印刷

ISBN 978-7-307-20125-5 定价: 30.00 元

版权所有, 不得翻印; 凡购我社的图书, 如有质量问题, 请与当地图书销售部门联系调换。

前 言

城市已成为人类居住的主要场所，但在城市发展中出现了越来越多的城市问题，交通拥堵即是其中较为突出的一个，而交通拥堵的主要成因之一即是由于城市居民通勤行为带来的潮汐式交通压力。从城市居民通勤行为的角度来分析，产生拥堵的原因有多种，如城市功能分区结构的不合理而导致过长的通勤距离、城市道路交通设施建设无法应对快速增长的机动车出行需求等。而要缓解或解决城市居民通勤所带来的城市交通拥堵，有两个关键点：其一是如何获取全面且相对精确的城市居民通勤出行特征，其二是如何通过定量的手段去研究分析居民的通勤行为。其中第一点随着各类城市大数据，尤其是手机数据这类位置服务数据在城市研究中的应用，已经备了较强的可行性；而就第二点而言，模型仿真已是城市交通研究中最常见的定量分析手段之一。因此，在回顾城市居民通勤行为的系列研究的基础上，本书通过对城市多源大数据及智能体模型仿真在城市研究中的应用进行解读，提出了基于手机数据来获取城市居民通勤出行特征、基于智能体模型来进行居民通勤行为仿真这一技术路线，并获得以下研究成果：

首先，基于手机话单数据、通话基站空间定位数据等，在 Hadoop 平台基础上进行了原始数据处理，并在武汉市主城区控规单元分区基础上进行了宏观层面的城市居民通勤特征分析，包括居民的主要通勤方向、平均通勤距离及其与城市交通的宏观联系。此外还从功能联系视角划分了各控规单元的主体功能，测度了不同单元分区间的联系紧密度，将其与武汉市的功能组团结构进行了对比，可验证城市功能结构规划的实施成效并为后续规划提供优化建议。

其次,选定武汉市白沙洲片区为研究对象,基于划定的居民出行小区地块及现有城市交通路网来构建智能体模型,从中微观层面去分析居民通勤出行对交通路况带来的影响。通过对居民出行时间的细粒度划分,获取工作日上午6点至9点之间四个时段的居民通勤行为数据并导入模型进行仿真。结果表明基于手机数据与智能体模型的居民通勤行为仿真能较好地模拟各主要交通路口的交通状况。此外还能基于分析数据得出较具针对性的交通优化策略,并能通过同一仿真模型来验证提出的道路优化策略能在多大程度上起到改善作用。

本书的创新主要有如下几点:

(1) 基于手机话单数据、通信基站空间定位数据以及控规单元划分,实现了城市宏观层面的居民通勤特征识别,并从通勤联系视角来重新审视了城市的空间功能结构。

以主城区控规单元为划分依据,通过手机话单数据获取了宏观层面上的居民出行OD特征,进而依托基站定位进行了空间可视化分类,不仅揭示了通勤联系视角下的城市功能区结构,还从宏观视角对通勤造成的交通拥堵进行了分析。本书的研究成果将从一定层面上反映城市规划的实施效果,并能为后期的结构调整提供依据。

(2) 构建了居民通勤行为的智能体仿真模型,基于手机数据实现了局部区域的居民通勤出行动态模拟,可以从微观层面来分析和评价城市道路布局以及城市功能区分布的合理性。

以一个典型的城市功能区域为研究对象,以小时为时间跨度,通过对一个月内工作日的通勤早高峰的4个小时的居民通勤出行数据进行模型的仿真实验,结果表明该模型有利于从微观层面来分析交通拥堵成因,其模拟结果数据也与现实观察基本保持一致,此外对于优化道路方案的仿真结果还证明了该模型可用于方案对比评价。

(3) 综合运用了大数据统计分析及智能体仿真这两个技术手段,拓展了城市研究中的技术思路。

智能体仿真模型已经逐步进入城市规划领域，因其定量分析的功能成为新的研究热点。近年来，大数据在城市研究中的应用逐渐增多，但将数据运用于真实案例的模型仿真仍较为少见。本研究对两者的综合应用不仅为城市研究提供了新的技术思路，也初步尝试了从数据统计、空间聚类到数据导入等一系列技术流程。

目 录

第一章 绪论.....	1
第一节 背景与问题.....	1
一、研究背景.....	1
二、具体研究的问题.....	2
第二节 相关领域研究综述.....	3
一、城市大数据研究.....	3
二、城市居民通勤行为研究	11
三、智能体仿真相关研究	18
第三节 研究目标及意义	28
第四节 研究方法及技术路线	30
一、研究方法	30
二、技术路线	31
第二章 多源数据及智能体建模的相关研究	32
第一节 多源数据在城市研究中的应用	32
一、多源数据的基本概念及应用	32
二、应用技术及前景分析	36
第二节 手机数据在城市中的应用	40
一、手机数据类型与特征	40
二、手机数据应用方向	41
三、相关技术及难点分析	44
第三节 智能体建模的相关理论研究	48
一、相关概念阐述	48
二、智能体建模的相关研究	52

三、智能体建模在城市研究中的作用	56
四、技术特点及应用前景	58
第四节 智能体在交通仿真研究中的应用	60
一、相关理论概述	61
二、技术重难点分析	62
第五节 本章小结	65
 第三章 手机数据处理及模型构建的技术实现	67
第一节 手机数据的获取及处理流程	67
一、手机数据概述	67
二、手机话单数据处理流程	70
第二节 数据处理平台搭建及初步校验	74
一、MapReduce 及 Hadoop 概述	75
二、Hadoop 平台搭建及数据处理	78
三、居民空间分布特征的校验与呈现	80
第三节 仿真平台选择	84
一、建模仿真平台的选择	84
二、Repast S 的功能及使用	86
三、Repast S 中的 GIS 数据导入	88
第四节 模型的技术实现	89
一、智能体的移动及路径规划	89
二、城市环境的构建实现	91
三、城市居民 Agent 的构建实现	92
第五节 本章小结	96
 第四章 基于手机数据的城市居民通勤宏观特征分析	97
第一节 通勤特征识别方法	97
一、手机数据用于通勤特征识别	97
二、通勤特征识别的流程	99
第二节 宏观视角的武汉市通勤特征识别	100
一、通勤方向识别	101

二、通勤距离特征.....	104
三、通勤特征与交通拥堵的关联分析.....	107
第三节 功能联系视角的城市空间结构.....	110
一、单元主体功能识别.....	110
二、通勤联系分区.....	112
第四节 本章小结.....	116
第五章 居民通勤出行的微观仿真及分析.....	118
第一节 城市居民通勤出行仿真框架.....	118
一、居民通勤出行模型的建模假设.....	118
二、模型的要素构成与运行流程.....	120
第二节 居民通勤出行模型的数据生成及分析.....	123
一、基础城市空间环境构建.....	123
二、居民通勤出行的数据生成及初步分析.....	123
第三节 居民通勤出行的分时仿真.....	130
一、模型仿真参数设置.....	130
二、居民通勤的分时仿真数据生成.....	131
三、仿真数据的分析及校验.....	136
第四节 基于仿真结果的交通改善策略.....	138
一、道路分段拥堵路口分析.....	139
二、交通改善策略.....	144
三、基于道路优化策略的仿真模拟.....	146
第五节 本章小结.....	155
第六章 结论.....	156
第一节 结论及贡献.....	156
第二节 研究不足及后续工作展望.....	157
参考文献.....	159

第一章 绪 论

第一节 背景与问题

一、研究背景

(一) 城市成为人类生活的主要空间区域

在人类的进化历程当中，城市化是必须要经历的过程，也标志着整个国家、社会发展的程度。从 16 世纪至今，全球城市人口大幅度增加，其占比从最初的 3% 左右，到现今的超过 50%。和发展中国家相比较，发达国家城市化程度更高，其城市人口占比更是超过七成。基于联合国的统计数据显示，到 21 世纪中叶，全球城市化水平必将会超过六成。这是一个非常庞大的数值，不仅意味着城市人口集聚增加，也表明国家、人类文化同样在迅速发展。

而在中国，随着改革开放步伐的深入，城市化进程也在持续加快。城市化进程的加速推动着整个国家的高速发展。根据相关统计数据显示，到 2016 年，我国城镇人口已超过总人口的一半，而根据相关预测，中国人口在本世纪中期将达到最高峰，估计在 16 亿人左右。届时，城镇人口将占到我国总人口的 70%。换言之，在全国范围内，城市中生活以及工作的人将超过 11 亿。

以上数据表明，随着城镇化进程在全球范围内不断推进，无论在全世界还是中国，城市都已经成为人类最为主要的集聚地。因此，城市问题也将是人类面对的最主要的问题，研究城市发展及其运行机理，不仅在当下，甚至在未来也将起到越来越重要的作用。

(二) 城市问题日益严重

在中国乃至全世界范围内,城市建设过程中出现的问题已经严重影响到整个城市化发展的质量,其主要在以下几方面体现:

第一,城市的盲目扩大,而城市原有格局、功能无法紧跟其发展,导致城市居住条件以及民众生活质量下降。

第二,城市人口持续增加,而就业岗位难以满足民众就业的需求,进而导致失业人口持续增加,带来诸多社会问题。

第三,城市中人口的急剧增加也导致人员出行和物资交流频繁。由于城市规模的持续扩大、城市经济的持续发展等原因,城市机动车发展态势迅猛,尤其是私有汽车的数量增长更为迅速。不断提升的机动车保有量为人们的日常生活带来了诸多便利,但同时也诱发不少的交通问题。大城市不得不面对交通事故频发、道路拥挤、能源短缺等一系列城市问题,这也阻碍城市生活水平的进一步提升。

唯有从城市交通内在本质出发才能分析、解决上述问题。所谓城市交通内在本质,实际上就是在民众日常生活中所衍化出来的派生物,其微观构成是城市民众个体出行行为,而这一行为在空间、时间等方面的分布,继而导致城市宏观交通现象,包括交通拥堵等。

二、具体研究的问题

我国正在持续推进和发展城市化,从宏观上说,城市规划的重要性也日渐突出,但与此同时现有的技术手段却难以满足城市规划需求,原有的城市规划方式已然落后于新的城市建设形势。无论是在设计或是实际实施层面,都缺少足够的技术作为支撑,依托于主观判断、空间美学设计所得出的成果经常面临朝令夕改的窘境^①。

^① 仇保兴.我国城镇化中后期的若干挑战与机遇——城市规划变革的新动向[J].城市规划,2010(1):9.

随着科技的进步、信息化水平的与时俱进、数字技术的广泛应用,获取数据也变得尤为简单,这进一步提升了数字定量分析的可行性。定量分析方法成为研究如何解决城市问题的有效途径。而在定量分析中,建模仿真等技术手段已经越来越多地受到众多学者的青睐。相对于单一的统一分析,建模仿真技术有利不仅有利于从微观层面来观察城市问题的产生过程,也可从更高的层面去发掘城市空间在整体上所涌现的宏观现象,更可以通过对不同方案的数据导入来对未来的城市空间变化进行预测,因此更适合对城市这样一个复杂的巨系统的研究。

要解决城市交通问题,必须对城市居民日常出行特征及其规律予以准确掌握,基于本质层面去对这一问题所在进行探索,随即通过规划和管理的手段对症结进行妥善解决。而对民众出行特征及其规律进行探究的前提条件就是能够系统全面地获取其出行行为数据。在城市居民的出行中,通勤活动是其中的重要一环。良好的通勤体验可以提升城市居民的幸福感,同时城市通勤与城市交通、城市环境等也息息相关。通勤出行造成的交通拥堵是城市交通问题中的关键问题。从微观个体出发,研究城市通勤特征及影响因素,有利于解决不合理的城市空间布局引发的城市交通问题。从城市居民通勤特征来看,城市居民通勤具备规律可重复的特性,这使得数据采集后的数据价值挖掘有了更为明确的目标。另一方面,当数据不再成为城市研究的桎梏,基于数据的城市模型仿真也能更好地用来研究及解决城市问题。

第二节 相关领域研究综述

一、城市大数据研究

随着“大数据”时代的到来,我们能够发现,不管日常生活,或思维、工作方式等都有所变化。*Nature* 杂志于 2008 年率先推出

“大数据”专刊，三年后，*Science* 杂志也推出这一专刊，对在不同领域当中大数据的研究应用进行了列举^①。大数据（big data）指的是不能在既定时间内通过常规方式来进行获取、管理以及处理的数据集合，需要全新处理方式方才可以具备更强的决策力、发现力以及优化能力的，海量、多样化、高增长率的信息资产。大数据是当下最热门的话题之一，在城市规划领域也掀起了一场革命，改变了传统城市规划的思维方式。与“大数据”经常同时出现的另一个概念是“智慧城市”，其是依托于信息以及通讯手段对城市运行系统的不同信息进行感测、分析以及整合，进而对涵盖环保、民生、城市服务等在内的不同需求进行智能响应。其于内在是依托于对信息技术的有效应用，继而确保智慧城市能够得以实现，继而改善民众的日常生活环境。“智慧城市”离不开“大数据”的有效利用，对于前者而言，后者是其基石，数据处理技术是前者的信息引擎^②。互联网时代，人们通过网络购买生活必需品代替了传统的逛街购物行为，通过网络团购和外卖软件点餐代替了传统的去饭店点菜行为，通过各类社交软件进行虚拟交流代替了传统的面对面社交行为，这些网络上的行为所产生的数据同样是一个庞大且具备价值的大数据集。来源于各类网站、手机 APP 的数据主要包括用户留言、位置共享、各类信息分享等，可深入挖掘的内容很多，也可为城市规划中的公众参与领域提供数据支持。

在城市研究领域，国外学者运用大数据分析较早，但数据量也较小，主要包括各类 Twitter、Facebook 等社交网络数据、移动手机

① 张晓祥. 大数据时代的空间分析 [J]. 武汉大学学报 (信息科学版), 2014, 39 (06): 655-659.

② 茅明睿. 大数据在城市规划中的应用: 来自北京市城市规划设计研究院的思考与实践 [J]. 国际城市规划, 2014 (06): 51-57.

通话数据、智能交通刷卡数据及城市监测设备数据的挖掘和分析^{①②③}。Klings 于 2009 年对比利时通信运营商所提供的 2500 万个通信信息进行研究分析,依托于用户手机账单地址建立起城市社会网络。2011 年,日本的 Wakamiya 等人依托于 Twitter 收集到带有地理标签的数据超过 1000 万条,这些数据包含用户发表信息的时间和所处地理位置,他们利用这些数据,运用和函数聚类法对大规模人群的行为活动进行分析,构建了日本的社会地理边界^④。Naaman 等人在 2012 年通过 Twitter 防火墙对程序进行破解,继而获得地标以及文本数据,同时结合时间序列模型对不同地区 Twitter 的关键词特征以及分布情况进行分析,最后发现能够提供不同时段关键词搜索所需信息的城市并不是很多^⑤。而 GPS 等不同类型的定位工具早已成为民众时空行为数据的采集工具,进而为对民众的出行模型进行有效研究提供契机。Mei Pok 等人对波特兰地区活动日志数据进行调查,同时依托于 GPS 获取民众出行路径数据,之后通过一些软件方式以及数学分析方式对民众活动密度、变化情况

① Y. Yuan, M. Raubal, Y. Liu. Correlating Mobile Phone Usage and Travel Behavior—A Case Study of Harbin, China [J]. Computers Environment & Urban Systems, 2012, 36 (2): 118-130.

② H. V. Roberts. Using Twitter Data in Urban Green Space Research: A Case Study and Critical Evaluation [J]. Applied Geography, 2017, 81: 13-20.

③ F. Luo, G. Cao, K. Mulligan, X. Li. Explore Spatiotemporal and Demographic Characteristics of Human Mobility via Twitter: A case study of Chicago [J]. Applied Geography, 2015, 70: 11-25.

④ W. S, L. R, S. K. Urban Area Characterization Based on Semantics of Crowd Activities in Twitter, in GeoSpatial Semantics [M]. Berlin Heidelberg: Springer, 2011: 108-123.

⑤ N. M, Z. A. X, B. S. On the Study of Diurnal Urban Routines on Twitter [C]. 6th International AAAI Conference on Weblogs and Social Media (ICWSM-12), Dublin, Ireland, 2012: 4-7.

等进行模拟^①。Edwards 等学者在基于 Google 地图的背景下,对 GPS 以及网页工具予以应用,模拟了澳大利亚的悉尼以及堪培拉两市 76 个旅游者的行动轨迹,继而对旅游者的出行路径等居民行为特征进行探究^②。

而在国内,李晨晨、姜劲松等人利用定向爬虫技术获得大众点评网的餐饮商户信息大数据,在此基础上对南京市的餐饮业空间分布格局进行基于 ArcGIS 的核密度分析、分级及影响因素研究^③。秦箫等人在研究分析大众点评网的相关商户信息之后,构建相应的评价体系,对商户口碑得分以及排名进行计算,之后通过这些信息对城市餐饮行业布局展开密度分析以及综合评价^④。此外,甄峰等人通过分析新浪微博数据信息,基于网络层面,系统地研究了我国城市网络社会空间发展特征^{⑤⑥}。

除了从城市网络社会空间这一方向以外,城市大数据的一大重要用途即是用来进行各类交通分析,包括跟城市交通有关的城市功能分区以及通勤特征分析等。针对交通的研究,传统的方式多是对人口规模、功能分区等进行分析,进而选择建设道路,基于该方式下的研究主要还是对城市及其空间协调能力进行考虑,而对于民众自身对交通条件的需求有所忽略,灵活性比较低,很容易导致交通

① K. M. L. J. Geovisualization of Human Activity Patterns Using 3D GIS: a Time-geographic Approach. Spatially Integrated Social Science. Examples in Best Practice [M]. London: Oxford University Press, 2003.

② B. Mark, M. Nick. Microscopic Simulations of Complex Metropolitan Dynamics [J]. Unpublished, 2011.

③ 李晨晨,姜劲松.基于大众点评网大数据的城市餐饮业空间分布格局研究——以南京市为例[C].沈阳:中国城市规划年会,2016.

④ 秦箫,甄峰,朱寿佳,席广亮.基于网络口碑度的南京城区餐饮业空间分布格局研究——以大众点评网为例[J].地理科学,2014(07):810-817.

⑤ 甄峰,王波,陈映雪.基于网络社会空间的中国城市网络特征——以新浪微博为例[J].地理学报,2012(08):1031-1043.

⑥ 刘朝青.基于流动空间的长三角城市社交联系研究[D].上海:上海师范大学,2013.

拥堵等突出问题。而在大数据时代,可以从城市交通管理部门、地铁站、公交公司等部门与机构来获得各类覆盖区域全面、时间跨度大且准确的交通出行数据。另外,各种通信地图类软件应用,如高德、百度、谷歌,使数据的获得方式更加多元化,同时实现了数据的可视化。利用这些交通数据可以更加方便、准确地了解城市居民的出行路线、对交通工具的使用偏好特征以及城市道路的拥堵状况^{①②}。有学者对依托于手机定位技术的时空数据采集方式以及应用研究进行了介绍,进而提出了全新的视角来探究移动信息时代下个体行为数据的收集,如通过研究民众活动日志、GPS 数据,同时结合调查、访谈等对背景两个社区的 50 位民众的相关行为信息进行统计研究,最后基于通勤弹性视角对民众通勤特征进行研究,依托于 GIS 技术刻画了多种理论通勤模型^③。随着信息时代的到来,民众的日常出行也变得尤为复杂,已然无法通过传统问卷调查来进行研究,而通过 GPS、网络日志等数据来对这一行为进行研究,被越来越多的民众所接受。龙瀛等人通过对北京一周内 IC 智能卡数据进行收集,同时与社会调查得到的出行数据相结合,采用 SCD 数据挖掘技术对北京民众通勤行为以及职住关系进行了研究^④。

随着科学技术的进步,智能手机逐步普及。当前,智能手机已然变成民众日常生活中无法缺少的东西,与此同时,手机所具有的功能能够将拥有者的日常行为进行记录,有关学者基于此,将手机视为一种探测器来进行交通研究^⑤。同传统的城市交通问题研究相

① 张维. 基于手机定位数据的城市居民出行特征提取方法研究 [D]. 南京: 东南大学, 2015.

② 宋璐. 基于手机定位数据的交通 OD 分布研究 [D]. 南京: 东南大学, 2015.

③ 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 (06): 733-744.

④ 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 [J]. 地理学报, 2012 (10): 1339-1352.

⑤ 冉斌. 手机数据在交通调查和交通规划中的应用 [J]. 城市交通, 2013. 11 (01): 72-81.

比较,通过智能手机数据开展相关研究具有众多优点,如样本量大、精度较高、连续性好等。截止到现在,海内外很多学者开始通过手机数据进行交通相关信息研究。国外一些发达国家在20世纪80年代已开始根据无线技术来对交通数据进行收集并继而予以研究。1990年之后,民众智能手机占有率的不断提高有力地促进了交通研究的发展,基于手机数据的交通优化策略研究已经逐渐成为城市探究的热点之一。

为了对依托于手机数据的交通信息以及影响其精度以及灵敏度的诸多因素进行研究分析,有研究人员通过不同的定位算法进行探究,基于速度、时间、位置的不同来研究一条单车道道路,继而针对因条件的不同而最终对信息采集所造成的影响进行分析,最后通过实验,验证了这一方式的可行性^①。Asakura等学者通过分析手机数据,继而提出民众出行轨迹跟踪标签算法,同时在日本大阪收集了大量民众的手机数据信息,依托于对这些信息的分析,最后得到民众出行轨迹,之后对比真实轨迹,发现这一方式是可行的^②。González等学者收集了超过10万个手机用户其信息,继而得出其6个月内的出行轨迹,与此同时,他们对这些轨迹展开时空数据挖掘,最终划分民众出行类型^③。同时,有些学者通过手机数据来对人口时空变化进行衡量,如通过手机话单信息来获得民众出行OD,同时和传统方式所得到的OD数据进行对比^④,进而对土地利

① R. E. Systems, P. B. Farradyne. Final Evaluation Report for the Capital—Its Operational Test and Demonstration Program [J]. Transportation Studies, 2007.

② Y. Asakura, E. Hato. Tracking Survey for Individual Travel Behaviour Using Mobile Communication Instruments [J]. Transportation Research Part C Emerging Technologies, 2004, 12 (3): 273-291.

③ M. C. González, C. A. Hidalgo. Understanding Individual Human Mobility Patterns [J]. Nature, 2008, 453 (7196): 779.

④ P. Deville, C. Linard, S. Martin, M. Gilbert, F. R. Stevens, A. E. Gaughan, V. D. Blondel, A. J. Tatem. Dynamic Population Mapping Using Mobile Phone Data [J]. Proceedings of the National Academy of Sciences of the United States of America, 2014, 111 (45): 15888.