

基于ICT的中国城市间人口日常流动空间格局 ——以百度迁徙为例

刘望保, 石恩名

(华南师范大学地理科学学院, 广州 510631)

摘要: 随着互联网时代的来临, 网络数据已越来越成为表征居民地理行为的重要载体, 用户迁移、社交网络、移动通信等地理行为大数据成为城市联系研究的重要数据来源。“百度迁徙”大数据通过LBS技术, 全程、动态、即时和直观地记录了城市之间的人口日常流动轨迹。通过采集“百度迁徙”数据库中2015年一季度(2月7日至5月16日)国内369个城市之间的逐日的人口流动数据, 分“季度平均、春运期间(春节前)、春运期间(春节后)、劳动节、周末和工作日”6个时间段, 从人流集散层级、人流集散网络体系的分层集聚、人口日常流动空间格局及其与“胡焕庸线”之间的关系等角度分析各时间段的的城市之间的人口日常流动相关特征与空间格局。研究发现, “百度迁徙”大数据清晰地显示了春运期间中部和沿海地区之间的人口流动格局。人流集散中心主要分布在京津冀、长三角、珠三角和成渝4大城市群中, 并与其城市等级有较强的一致性。人口日常流动集散体系呈明显的分层集聚, 京津冀、长三角、珠三角、成渝和乌鲁木齐5大集散体系在各时间段基本得到体现, 而华中、东北、西南和福建沿海等地区并未出现高层级集散城市和高等级集散体系, 与这些区域在国家区域发展战略中的地位在一定程度上不相匹配。“胡焕庸线”能较好地反映国家层面的城市之间人口日常流动格局, 反映了地理环境对城市间人口日常流动的深刻影响。城市之间的人口流动强度是体现区域经济联系强度、城市等级和网络结构等的重要指标, 此项研究可为形成国家区域经济发展新格局和促进区域平衡发展提供参考。

关键词: 百度迁徙; 日常流动; 集散体系; 春运; 中国

DOI: 10.11821/dlxb201610001

1 引言

2014年中央电视台《据说春运》节目中的“百度看迁徙”播出后备受关注。“百度迁徙”大数据利用基于地理位置的服务(Location Based Service, LBS)技术, 首次全程、动态、即时、直观地展示了2014年春运期间全国人口大迁徙的轨迹和特征。目前, “百度迁徙”平台已经更新, 大众可实时查看国内城市之间的人口流动状况。近年来, 随着3S技术、移动互联网、物联网等技术的快速发展, 使包括地理位置、社会属性、移动轨迹、迁移过程和交互模式等信息的人类时空行为数据综合、连续观测成为可能^[1], 既可

收稿日期: 2015-08-17; 修订日期: 2016-01-30

基金项目: 国家自然科学基金项目(41001088); 教育部人文社科青年项目(09YJC840016); 国家重点基础研究发展计划(973)前期研究专项(2014CB460614) [Foundation: National Natural Science Foundation of China, No.41001088; Fund of Social Sciences Research, Ministry of Education, No.09YJC840016; The National Basic Research Program of China (973 Program), No.2014CB460614]

作者简介: 刘望保(1975-), 男, 博士, 教授, 中国地理学会会员(S110007720M), 研究方向为城市社会地理。

E-mail: wbliu@scnu.edu.cn

通讯作者: 石恩名(1991-), 男, 硕士生, 研究方向为城市社会地理。

通过网络数据挖掘等技术获取带有位置信息的社交网络信息,如新浪微博签到^[2]和点评网^[3]等,也可利用GPS或LBS设备,结合GIS或网络日志采集和分析居民地理行为数据^[4],抑或者获取政府开放大数据,如公交刷卡数据^[5]、浮动机动车^[6]等交通出行行为大数据,这些海量时空数据使基于此的人类移动性(human mobility)研究成为可能^[7-9]。

目前,研究者们主要用包括带有位置服务的社交网络和移动通信等大数据来研究城市之间的互动关系。在社交网络数据方面,国外学者主要用Twitter和FourSquares等大型社交网站的签到(check-in)数据,而国内学者主要用微博签到或好友关系数据来进行相关研究。Naaman等^[10]利用美国不同城市用户在Twitter上发帖的关键词的特征和分布状况来分析城市之间的网络活动等级及其联系,其他类似研究如利用多源LSS(Location sharing services)(如FourSquares、Gowalla等数据)来分析用户城市之间的时空足迹^[11]。国内学者利用微博数据分析城市之间的互动联系^[12-14],利用新浪微博好友数据来分析城市网络结构和体系^[12]和利用移动定位社交(LBSNS)签到数据分析城市联系^[14]。移动通信数据也成为分析城市之间的社会网络联系的重要数据来源^[15-18]。Krings等^[15]利用比利时6个月内2500万个用户通信信息和用户手机账单地址对应的邮编来分析城市之间的通讯联系强度和等级体系;Kang等^[17]利用黑龙江1个月内19亿个通话记录数据,构建起省和国家两个层面的区域之间的通信联系网络格局。以往城市联系研究主要采取从城市之间的功能联系来解读城市联系或采用交通流来分析城市联系格局。如采用全球生产者服务企业的办公网络来分析城市之间的社会经济联系^[19],或采用航班和货运量^[20-22]、铁路流量^[23]等交通流来研究城市之间的联系格局。跨城市间的地理行为大数据的数据量大、准确性强、整合分析能力强等方面优势是传统城市联系数据无法比拟的,因而受到城市联系研究越来越多的青睐。

随着互联网时代的来临,网络数据越来越成为表征居民社会活动的重要载体^[24],利用用户迁移、社交网络、移动通信等地理行为大数据来研究城市间互动关系已成为重要趋势。与传统行为调查数据相比,这些地理行为大数据及其分析方法以其强大的信息搜集、整合、分析和处理优势,在样本量大、数据多源和分析预测性强等方面具有传统数据无法比拟的优势,已经越来越成为城市空间与人类行为模式研究的重要数据来源。当然,地理行为大数据往往具有一定缺陷,如存在被采集群体偏向性较强和个体社会经济属性缺失或不真实等问题,本文采取的“百度迁徙”大数据也存在这些缺陷,大部分未接入ICT的群体行为未被记录,也缺乏迁徙者的个体社会经济属性,无法建立起迁徙行为与个体社会经济属性之间的相关性。

2 数据来源与分析思路

2.1 数据来源与评估

百度地图LBS开放平台为数十万款APP提供免费、优质的定位服务,是国内LBS数据源最广的数据与技术服务平台。“百度迁徙”数据来源于百度地图和第三方用户的定位数据统计,实时、动态、直观地展示了区域间人口日常流动,真实地记录了数以亿计的人口迁徙轨迹。本文利用“百度迁徙”平台,获取了2015年2月7日至5月16日(共99天)的全国369个城市之间的逐日人口迁移数据,经统计,此期间城市之间的人流量合计达到1.21亿人次。

由于预处理原因,此项数据在分析城市间的人口日常流动方面存在如下缺陷:①大量未接入百度地图和第三方用户的群体的日常流动行为未被记录;②只记录了用户8小

时的空间位移,尽管现在高铁、飞机和私家车出行方式所占比重较高,大部分的出行耗时在8小时内,但可预见,仍有相当规模的空间位移必然未被记录;③大部分的出行被拆解,未被完整识别,如从A地出发到最终目的地B地途径了C地,同时用户在C地也有定位信息,此次出行会被拆解为A-C和C-B两次出空间位移,由于此项数据只是统计城市间的人口流动量,没有用户的ID信息,因此无法识别用户的完整出行。

基于以上分析,与以往用于城市联系研究的城市之间的经济功能联系解读和交通流量数据相比,“百度迁徙”大数据由于无法识别实际的城市之间的人口互动(联系),在分析重力模型检验和城市体系及其边界识别上存在较大缺陷。此项数据首次通过海量数据全程、动态、即时地展示了城市之间的人口流动格局,而更适合用于分析基于ICT群体的城市之间的日常人口流动分析,这也是其他数据无法比拟的。

2.2 分析思路

数据的采集时间跨度包含春运期、“五一”劳动节等重要节假日及周末、工作日等时间段,不同时间段的用户出行主要目的差异较大,相应地,其人口日常流动的特征和空间格局必然存在较大差异,为此,本文选取2015年“季度平均(98天)、春运期间(春节前,2月7日至19日)、春运期间(春节后,2月20日至3月15日)、劳动节(选取5月1日至3日)、周末(除春运期和劳动节整周外的所有周末)和工作日(除春运期和劳动节整周外的所有工作日)”等时间段(以下分别简称为“季度平均、春运期(春节前)、春运期(春节后)、劳动节、周末和工作日”),为便于比较,采取各时间段的日均人流量。经统计,“季度平均、春运期间(春节前)、春运期间(春节后)、劳动节、周末和工作日”各时间段的日均人流量分别为124万、265万、185万、120万人,97万和92万人,可见,各时间段的差异已经凸显,春运期间,尤其是春运期(春节前)的日均人流量最大,其次是劳动节,而周末的日均人流量比工作日稍大。本文试图从人流集散层级、人流集散网络体系的分层集聚、人口流动空间格局及其与“胡焕庸线”之间的关系等方面比较不同时间段的城市间人口日常流动的相关特征和空间格局差异。鉴于春运期间人流量巨大(占整个季度人流总量的64.89%)和人口流动格局的典型性,本文还简要分析了春运期间的人口日常流动的特征和空间格局。

3 分析结果

3.1 春运期间区域之间人口日常流动空间格局

图1以省为单元统计出2015年一季度逐日的人口净迁入量^①的时间序列分布图,并突出显示了春运期间的分布格局,归纳其主要特征为:①清晰地展示了春节前沿海省份向中部省份的大规模“回乡潮”和春节后中部省份向沿海省份的大规模“返工潮”。②各省份流入流出高峰期的时间节点存在较大差异,人口流出高峰时间点发生最早的省份是广东,在春节前6天就迎来了人口流出高峰,其他人口流出大省一般在春节前2~3天达到高峰。广东又是人口流入高峰时间最迟的省份,人口流入高峰发生在元宵节前后,而其他沿海省份一般在正月初六、七左右,究其原因在于广东省春运期间人口流动以农民工为主,农民工的迁徙时间自由度较高,大部分会选择在家过完元宵节后回广东工作。

图2以城市为单元展示了春运期间城市之间的日均净迁入量分布图(箭头表示迁入方向),图形直观地反映了京津冀、长三角、珠三角、成渝4大城市群在春运期间城市间

① 区域间人口净迁入量=区域间迁入量-区域间迁出量,本文仅保留净迁入量“正值”部分,且“正值”表示迁入方向。

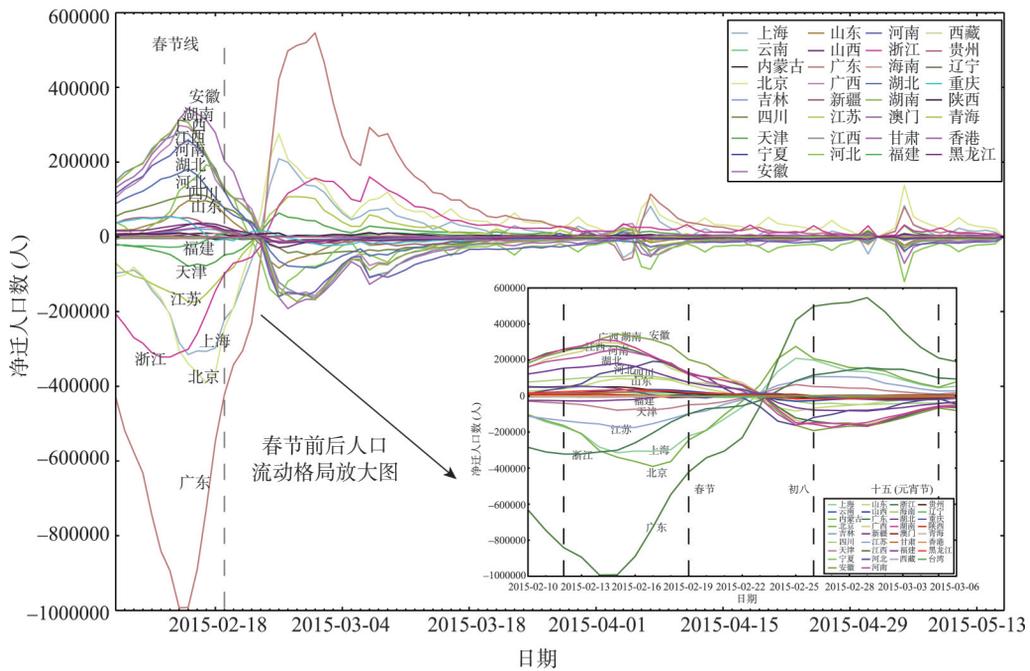


图1 2015年春运期间中国各省份净迁入人口分布

Fig. 1 The net immigration population distribution of each province during "Spring Festival Travel Rush" in 2015

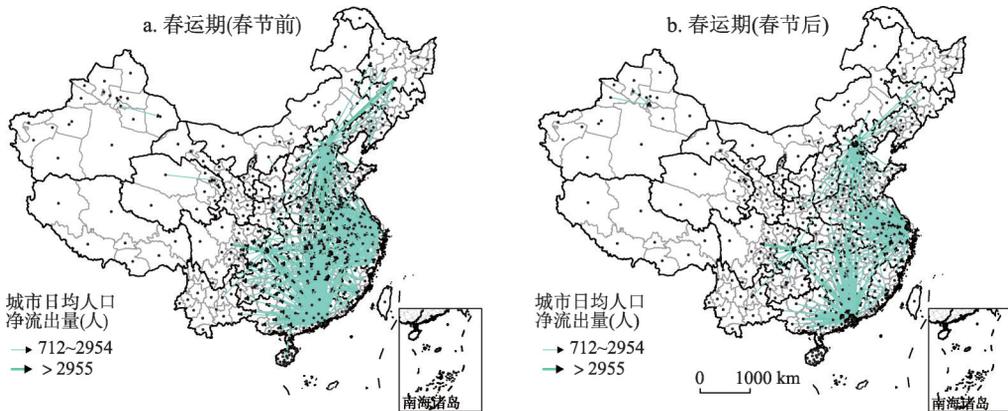


图2 2015年春运期间中国城市之间人口日常流动净流量空间格局

Fig. 2 The spatial pattern of urban net daily immigration flow in "Spring Festival Travel Rush" in 2015

人口日常流动格局中的重要地位。首先，从春运期人口日常净流入、净流出的空间格局来看，净流量在第2层级（即 ≥ 712 人）以上的流出、流入城市（春节前流出、春节后流入）均属于4大城市群的核心城市。其次，4大城市群有明显的辐射范围界限，京津冀主要覆盖华北地区、华东和东北部分区域；长三角主要覆盖华东地区、中部安徽和河南的部分地区；而珠三角主要覆盖湖南、江西、安徽、广西等地区；成渝城市群主要覆盖四川、重庆及湖北、贵州的部分地区；辐射范围体现空间临近性特征。

3.2 城市间人口日常流动集散层级与网络体系

根据各城市的人口集散总量（即流入流出总量）和城市间人口日常流量，利用 ArcGIS 9.3 软件绘制各时间段的人口集散层级和网络体系分布专题图。由于工作日的日

均人流量规模最小,因此,集散层级和城市间人流量的区间划分统一按照“工作日”的规模分级(采取自然分级法(Natural Breaks Jenks))来确定,其中,集散层级分为4级,包括:0~6051人、6052~26667人、26668~61107人和 ≥ 61108 人;城市之间的日均人流量区间划分为5级,包括:0~222人、223~1042人、1043~2957人、2958~7108人和 ≥ 7109 人。

3.2.1 人口集散量排前10位城市和流动路线 表1显示了各时间段人流集散量排前10位的城市和流动路线,归纳其特征主要为:①除春运期2个时间段外,其他时间段人流集散两排前10位城市基本类似,主要包括北京、上海、天津、苏州、深圳、广州、重庆、廊坊等,这些城市为全国主要人流集散中心。顾朝林等利用重力模型,用摩擦系数 $b = 1$ 和 $b = 2$ 分别揭示国家和省区尺度上的网络联系,得到2003年I级节点城市为北京、上海、广州、重庆、武汉、哈尔滨、沈阳等7个^[25],而甄峰等利用新浪微博好友关系来确定

表1 2015年各时间段中国人口日常流动集散层级排前10位的城市和流动路线

Tab. 1 The urban daily flow levels and flow paths in each period (top 10) in 2015

类型	人口日常流动集散层级		城市间人口日常流动路线	
	集散层级排前10位城市(人)	占人口总流量的比重	流量排前10位的流动路线	占人口总流量的比重
季度平均 (98天日均)	北京(216146)、上海(160084)、苏州(74938)、天津(74245)、深圳(71704)、广州(61828)、廊坊(50944)、东莞(45764)、杭州(41474)、重庆(37679)	合计834805人,占集散总量的33.66%	北京至廊坊(18806)、廊坊至北京(18727)、上海至苏州(14321)、苏州至上海(13871)、保定至北京(12625)、北京至保定(9445)、天津至北京(7824)、北京至天津(7156)、乌鲁木齐至昌吉(5947)、昌吉至乌鲁木齐(5807)	合计114528人,占总量的9.20%。
春运期 (春节前,2月7日-19日日均)	北京(367560)、上海(305944)、深圳(226306)、广州(163108)、东莞(1581560)、重庆(153976)、苏州(147339)、天津(104282)、杭州(90823)、佛山(76196)	合计1793689人,占集散总量的33.93%	北京至廊坊(23054)、北京至保定(21479)、廊坊至北京(19890)、上海至苏州(19736)、苏州至上海(16613)、广州至赣州(11252)、成都至重庆(10976)、广州至衡阳(10674)、深圳至赣州(10512)、北京至张家口(10442)	合计154627人,占总量的5.83%。
春运期 (春节后,2月19日-3月15日日均)	北京(273091)、上海(204596)、重庆(125319)、深圳(117612)、广州(96387)、苏州(94860)、天津(89253)、东莞(78445)、杭州(57403)、廊坊(50721)	合计1187686人,占集散总量的32.15%	廊坊至北京(18082)、北京至廊坊(17640)、保定至北京(16865)、苏州至上海(13549)、上海至苏州(12723)、北京至保定(8635)、天津至北京(8029)、重庆至成都(7763)、张家口至北京(7011)、昌吉至乌鲁木齐(6959)	合计117256人,占总量的6.32%。
劳动节 (5月1日-3日日均)	北京(208402)、上海(147981)、天津(95964)、重庆(74761)、苏州(62092)、南京(48519)、廊坊(47493)、杭州(40266)、广州(38510)、武汉(36880)	合计800868人,占集散总量的33.43%	上海至苏州(19009)、北京至廊坊(18388)、廊坊至北京(15909)、北京至天津(13790)、北京至保定(12792)、天津至北京(12665)、苏州至上海(11652)、保定至北京(11623)、乌鲁木齐至昌吉(7191)、上海至杭州(7001)	合计130017人,占总量的10.80%。
周末 (除春运和五一整周外所有周末日均)	北京(170975)、上海(119200)、天津(68588)、苏州(55916)、重庆(52944)、廊坊(49437)、广州(34393)、南京(33829)、深圳(33786)、杭州(30765)	合计649833人,占集散总量的33.40%	北京至廊坊(19359)、廊坊至北京(17271)、上海至苏州(16040)、苏州至上海(12175)、保定至北京(11548)、天津至北京(6689)、北京至保定(8303)、北京至天津(8128)、乌鲁木齐至昌吉(6208)、昌吉至乌鲁木齐(5729)	合计113429人,占总量的11.68%。
工作日 (除春运和五一整周外所有工作日日均)	北京(170201)、上海(114997)、天津(61107)、苏州(52259)、重庆(48549)、廊坊(48517)、深圳(35606)、广州(35606)、南京(28668)、杭州(26668)	合计622516人,占集散总量的33.91%	廊坊至北京(18941)、北京至廊坊(17331)、苏州至上海(13445)、上海至苏州(12276)、保定至北京(12095)、天津至北京(7109)、北京至保定(5887)、北京至天津(5551)、昌吉至乌鲁木齐(5213)、廊坊至天津(5073)	合计102921人,占总量的11.16%。

注:数据经“百度迁徙”数据库整理得。

城市网络社会空间及其等级关系,确定北京、上海、广州、深圳、成都、天津、杭州、武汉和福州等城市为区域级网络联系中心以上城市^[12]。总体来看,本文研究结论与相关研究有较大相似性,但廊坊、苏州等城市等级明显要高,而武汉、福州、成都等城市确没有出现在前10位城市中。廊坊与苏州城市等级尽管不如武汉、福州和成都等城市,但高等级城市之间的重要中转城市的人口集散功能层级明显偏高。西北地区、华中地区、福建沿海地区和东北地区均未出现排前10位城市。珠三角的东莞和佛山2个城市在春运期间(2个时间段)也排前10位。② 这些城市主要为京津冀(北京、天津和廊坊)、长三角(上海、苏州、南京和杭州)、珠三角(深圳、广州)和成渝(重庆)4大城市群的中心城市。③ 各时间段,排前10位城市的人流集散量占所有城市的总流量的比重均比较稳定,季度平均、春运期(春节前)、春运期(春节后)、劳动节、周末和工作日的相应比重分别为33.66%、33.93%、32.15%、33.43%、33.40%和33.91%,均约占1/3左右。④ 流量排前10位的流动线路以位置临近的中心城市为主,其中,季度平均、劳动节、周末和工作日4个时间段均以北京和廊坊、上海和苏州、北京和保定、北京和天津、乌鲁木齐和昌吉之间的日常流动为主,而春运期间,尤其是春运期(春节前)珠三角中心城市(广州与深圳)与周边城市(赣州与衡阳)流动量进入前10位,反映了春运对城市间人口日常流动格局的影响,也可以看出,在春运期间,赣州和衡阳与珠三角的人员联系最为密集,是珠三角辐射中部省份的最主要的中心城市。

3.2.2 人口流动集散层级的空间分布 图3显示了各时间段的人流集散层级的空间分布,归纳其特征主要为:① 层级较高的城市主要分布在沿海地区,尤其是京津冀、长三角、珠三角和成渝等城市群,在京沪高铁沿线呈带状分布;中西部地区高层级城市主要为省会城市。② 在西北地区,乌鲁木齐是最大的集散中心,除工作日时间段外,其他时间段乌鲁木齐的集散层级均为第三级(即集散规模在26668~61107人之间);其他集散中心还主要包括昌吉、西宁和海东等城市,而兰州仅在春运期(即春节前、后2个时间段)和劳动节期间达到第2层级(即集散规模在6052~26667人之间)。③ 中西部地区规模较大的集散中心主要包括重庆、西安、成都,尤其是重庆在季度平均、工作日和劳动节3个时间段达到第3层级,而在春运期2个时间段和周末达到第4层级,成为中西部地区最大的集散中心。④ 海西地区和东北地区作为国家经济发展重心地区之一,但并未出现较高集散层级城市,6个时间段中,只有温州在春运期(春节前)集散量为61494人,达到了最高层级区间。

进一步分析6个时间段的差异,主要表现在:① 以相同的区间划分标准,人流量越大,各城市的集散层级相应越高,春运期2个时间段高层级城市数量明显增加,且分布更为密集;② 春运期间,尤其是春运期(春节前),湖南、江西、广西、河南等中部省份主要城市集散层级大幅度提高,体现了春运期中部省份和东部沿海省份之间的人口大规模流动的趋势。

3.2.3 人口日常流动集散体系分层集聚 借助ArcGIS软件中的网络分析(network analysis)工具,根据城市人口集散层级和城市之间的人口流动强度,制作城市人口日常流动集散体系专题地图,并识别其分层集聚现象(表2,图4)。集散体系分层集聚的识别标准为:① 集聚中心的识别:有明显的集聚中心,且集聚中心层级在第3层级及其以上(即集聚规模 ≥ 26668 人);② 边界识别:城市之间的人流量在第3层级及其以上(即 ≥ 1043 人),且连接数 > 2 。根据以上标准,识别出各时间段的集散体系分层集聚格局。

各时间段的人口日常流动集散体系数量和覆盖范围差异较大。首先,京津冀、长三角、珠三角、成渝和乌鲁木齐5大体系在各时间段基本能被识别(除珠三角体系在“周

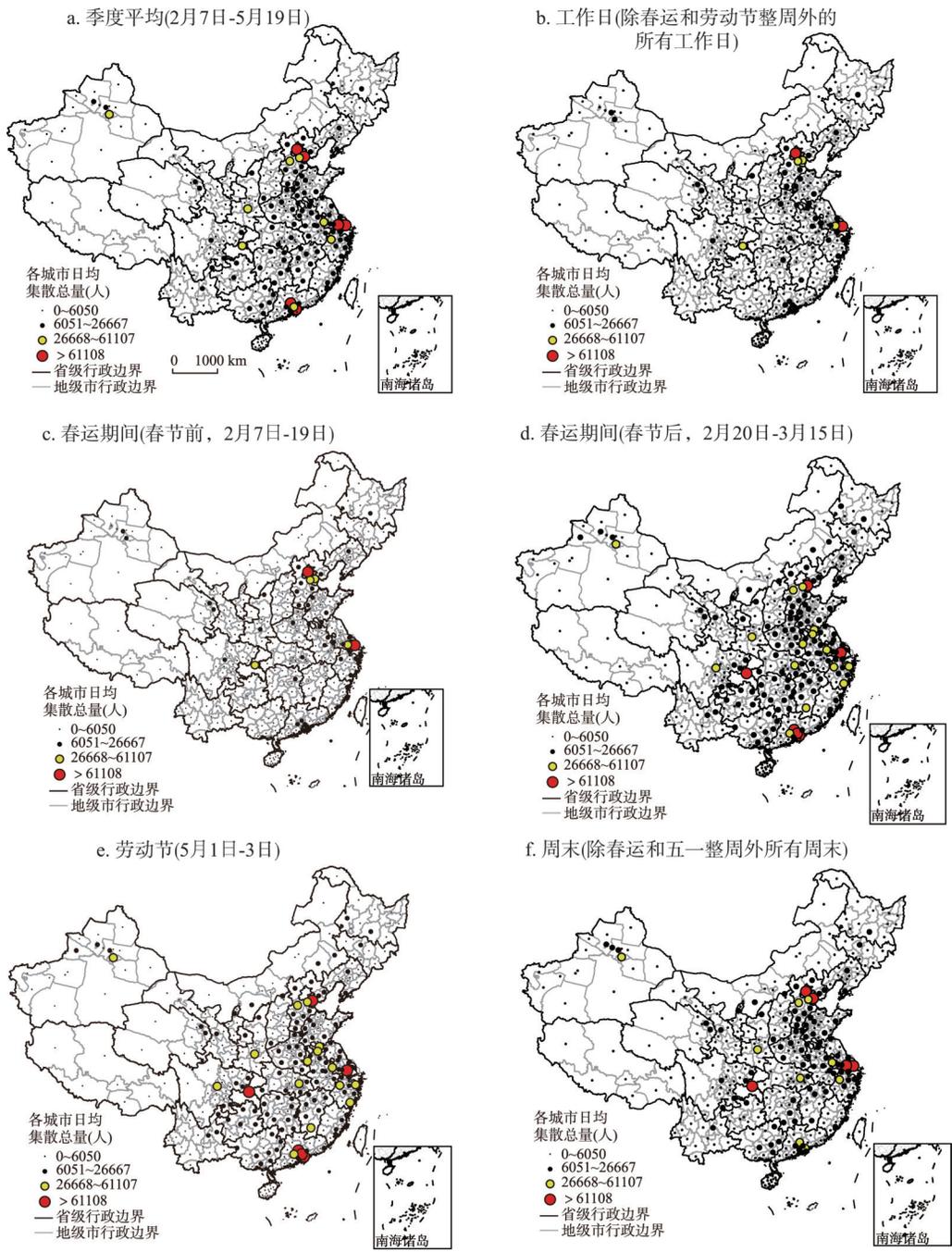


图3 2015年各时间段中国城市人口日常流动集散层级分布
 Fig. 3 The spatial patterns of urban population daily flow levels in 2015

末”时间段和乌鲁木齐体系在“工作日”时间段未被识别外), 而西安体系、武汉体系、厦门体系在部分时间段被识别。与以往研究^[12]相似, 京津冀、长三角、珠三角、成渝4大体系在本研究中基本得以体现, 但从“百度迁徙”数据分析中, 中部地区(武汉)、福建东南沿海(海西地区)的集散体系只在春运期间和五一假期得以体现, 且集聚规模不

表2 2015年各时间段中国人口日常流动集散体系的识别
Tab. 2 Main network systems of population daily flow in each period in 2015

时间段	集散体系	核心城市(第3层级及其以上)	时间段	集散体系	核心城市(第3层级及其以上)	
季度平均	京津冀体系	北京、天津、廊坊、保定	工作日	京津冀体系	北京、天津、廊坊	
	长三角体系	上海、苏州、杭州、南京		长三角体系	上海、苏州、南京	
	珠三角体系	深圳、广州、东莞		珠三角体系	深圳、广州	
	成渝体系	重庆		成渝体系	重庆	
	乌鲁木齐体系	乌鲁木齐				
春运期间(春节前)	京津冀体系	北京、天津、廊坊、保定	春运期间(春节后)	京津冀体系	北京、天津、廊坊、保定	
	长三角体系	上海、苏州、杭州、宁波、温州、南京、阜阳、金华、无锡、周口、嘉兴、宿州、信阳、商丘、徐州、六安、亳州、台州		长三角体系	上海、苏州、杭州、南京、阜阳、宁波、徐州、温州、宿州	
		珠三角体系			深圳、广州、东莞、佛山、赣州、武汉、中山、玉林、衡阳、贵港、永州、南宁、黄冈、长沙、邵阳、惠州	珠三角体系
	成渝体系	重庆、成都		成渝体系	重庆、成都	
	西安体系	西安		西安体系	西安	
	乌鲁木齐体系	乌鲁木齐		乌鲁木齐体系	乌鲁木齐	
	厦门体系	厦门、赣州、鹰潭				
	劳动节	京津冀体系		北京、天津、廊坊、保定	周末	京津冀体系
长三角体系		上海、苏州、南京、杭州	长三角体系	上海、苏州、南京、杭州		
珠三角体系		深圳、广州	成渝体系	重庆		
武汉体系		武汉	乌鲁木齐体系	乌鲁木齐		
成渝体系		重庆				
西安体系		西安				
乌鲁木齐体系		乌鲁木齐				

注：数据经“百度迁徙”数据库整理得。

大，而东北地区基本不存在较明显的集散体系。新疆的城市首位度较高，人口分布分散，乌鲁木齐成为区域最大的人口集散中心，形成了乌鲁木齐体系，这在除“工作日”外的5个时间段均被识别。春运时期（春节前、春节后）2个时间段，由于京津冀、长三角和珠三角3大体系覆盖范围广，诸多次级体系被覆盖而未被识别。其次，从覆盖范围来看，人流量越大，集散体系覆盖范围越大，相应地核心城市数量越多。春运期2个时间段，京津冀、长三角、珠三角和成渝4大集散体系覆盖范围基本囊括了整个中、东部地区。

3.2.4 人口日常流动格局与“胡焕庸线”的关系 以图4叠加了“胡焕庸线”，可以看出，各时间段的高层级的人口集散城市和识别的集散体系均主要分布在“胡焕庸线”的东南壁。以“胡焕庸线”为界统计出东南壁内部和西北壁内部各城市之间的人口流量，统计得到2015年“季度平均、工作日、春运期（春节前）、春运期（春节后）、劳动节和周末”6个时间段的东南壁内部和西北壁内部人流量分别为86.23万和8.26万、81.42万和6.36万、201.27万和10.93万、129.20万和10.57万、107.28万和8.02万、85.83万和6.98万，东南壁内部人流量分别是西北壁相应流量的10.43倍、12.79倍、18.42倍、12.22倍、13.38倍和12.29倍。可见，各时间段均体现了“胡焕庸线”东南壁内部人流量远高

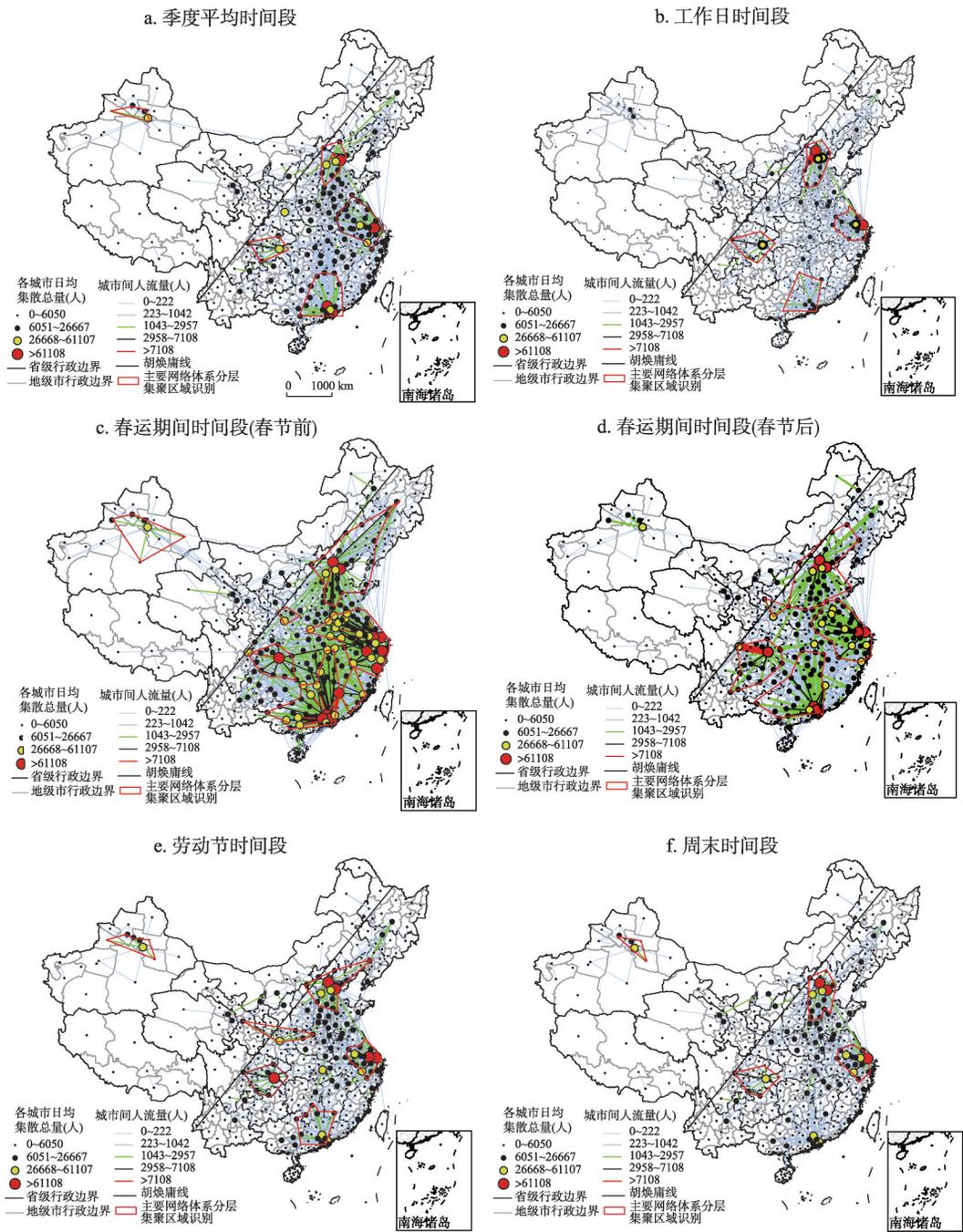


图4 2015年各时间段中国城市人口日常流动集散体系识别

Fig. 4 The identification of urban daily flow network system in each period in 2015

于西北壁的相应流量的规律, 从两端流量的比值来看, 除春运期外, 其他时间段基本在10~13左右, 比值相对稳定, 戚伟等利用改革开放以来国家四次分县人口普查数据研究证实“胡焕庸线”两端人口分布格局的稳定性^[26]。由此可见, “胡焕庸线”不仅反映了中国相对稳定的人口空间分布格局, 也反映了相对稳定的国家层面整体城市间的人口日常

流动空间格局。

基于ICT的人口流动格局与区域自然地理环境、人口与产业活动、城镇化发展及互联网发展等要素密切相关,“胡焕庸线”两端形成了相对稳定的人口流动格局,反映了自然地理环境对基于ICT的城市间人口日常流动空间格局的深刻影响。两端的自然地理环境差异影响两端的产业与城镇化发展,加上历史因素的影响,造成区域发展不平衡,进一步影响城市间的人口流动,另外,区域互联网发展水平和居民移动互联网使用习惯的差异也是造成两端城市间人口日常流动规模和格局差异的重要原因。

4 结论与讨论

4.1 结论

鉴于“百度迁徙”数据自身固有的缺陷,本文仅对接入ICT群体城市之间的人口日常流动的相关特征和空间格局进行尝试性研究,诸多相关研究内容,如未接入ICT群体的人口流动、人口日常流动行为与个体社会经济属性之间的相关性等值得后续的补充研究。本文主要研究思路为:利用“百度迁徙”大数据中2015年一季度(采集时间为2月7日-5月16日)城市之间的日均人流量,分“季度平均、春运期(春节前)、春运期(春节后)、劳动节、工作日和周末”6个时间段,从人流集散层级、人流集散网络体系的分层集聚、人口流动空间格局及其与“胡焕庸线”之间的关系等角度比较各时间段的差异。得出结论主要如下:

(1)“百度迁徙”大数据清晰地展示了中国春运期间人口流动空间格局,春运期间人流量大,春节前后中部省份与沿海省份之间的大规模“回乡潮”和“返工潮”得以体现,京津冀、长三角、珠三角和成渝4大城市群在春运期间人口流动空间格局中占据重要地位,净迁出、净迁入规模较大的城市(春节前为净迁出、春节后为净迁入)绝大部分分布在4大城市群中。

(2)各时间段主要的人流集散中心有较高的相似性,北京、上海、天津、苏州、深圳、广州、重庆、廊坊等为主要人流集散中心。春运期,特别是春运期(春节前)人流量大,各城市集散层级明显提升,尤其是中部省份重要城市的集散层级得到大幅度提升。各时间段主要的人流集散中心与城市等级有较强的一致性,主要分布在4大城市群中,而介于高等级城市之间的中转城市,如廊坊、苏州等人流集散层级偏高。华中、福建沿海、东北等地区中心城市,如武汉、福州、厦门、哈尔滨、沈阳、大连等城市的集散层级并不高,与所承担区域职能不相匹配。西北地区最大的集散中心是乌鲁木齐,而非兰州,相反,西北地区的西宁和海西2个城市由于旅游业的发展,其集散层级甚至比兰州要高。

(3)城市之间的人口日常流动集散体系呈现明显的分层集聚。京津冀、长三角、珠三角、成渝和乌鲁木齐5大集散体系在各时间段基本能体现,而西安体系、厦门体系和武汉体系只在部分时间段得以体现。与国内城市体系相关研究不同的是,“百度迁徙”反映出华中、东北、西南和福建沿海等并未出现高等级的集散体系,仅有武汉体系、西安体系、厦门体系在部分时间段得以体现。

(4)“胡焕庸线”也能较好地反映国家层面的城市之间人口日常流动格局。各时间段高层级人流集散城市和集散体系均主要分布在“胡焕庸线”的东南壁,除春运期(春节前)外的其他时间段,东南壁内部人口流量是西北壁相应流量的10~13倍,比值相对稳

定,反映了自然地理环境对城市之间的人口日常流动格局的深刻影响。

4.2 讨论

2015年国家政府工作报告中指出要统筹实施“四大板块”和“三个支撑带”战略组合,“四大板块”为东部地区、西部地区、中部地区和东北地区,而“三个支撑带”为“一带一路”、京津冀协同发展、长江经济带,形成全国区域经济发展新格局。京津冀、长三角和珠三角三大城市群在全国城市网络体系中已占据主导地位,而“一带一路”规划将新疆和福建分别作为丝绸之路经济带和21世纪海上丝绸之路的核心区,然而,从城市间人口日常流动空间格局来看,新疆和福建的中心城市的人口集散功能层级仍较低,也未形成完善的人口集散体系,未来需要增强新疆和福建各城市,尤其是核心城市的人口集散功能,并通过构建便捷的交通促进区域人口集散体系的发展完善。长江经济带形成了长三角、长江中游和成渝城市群,从城市间人口联系网络体系格局来看,长三角和成渝两大城市群网络体系发育较为完善,而长江中游的武汉、长沙、南昌等中心城市人口集散功能层级较低,城市网络体系发育不完善,不利于长江经济带的整体发展。“胡焕庸线”也描绘了城市间人口流动格局,如何突破这种格局,促进中西部地区城市人口集散功能的快速发展是重大课题。

参考文献(References)

- [1] Ke Wenqian, Yu Zhaoyuan, Chen Wei, et al. Architecture and key issues for human space-time behavior data observation. *Geographical Research*, 2015, 34(2): 373-383. [柯文前, 俞肇元, 陈伟, 等. 人类时空行为数据观测体系架构及其关键问题. *地理研究*. 2015, 34(2): 373-383.]
- [2] Wang Bo, Zhen Feng, Zhang Hao. The dynamic changes of urban space-time activity and activity zoning based on check-in data in Sina Web. *Scientia Geographica Sinica*, 2015, 35(2): 151-160. [王波, 甄峰, 张浩. 基于签到数据的城市活动时空间动态变化及区划研究. *地理科学*, 2015, 35(2): 151-160.]
- [3] Qin Xiao, Zhen Feng, Zhu Shoujia, et al. Spatial pattern of catering industry in Nanjing urban area based on the degree of public praise from Internet: A case study of Dianping.com. *Scientia Geographica Sinica*, 2014, 34(7): 810-817. [秦箫, 甄峰, 朱守佳, 等. 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. *地理科学*, 2014, 34(7): 810-817.]
- [4] Qin Xiao, Zhen Feng, Xiong Lifang, et al. Zhu Shoujia. Methods in urban temporal and spatial behavior research in the Big Data era. *Progress in Geography*, 2013, 32(9): 1352-1361. [秦箫, 甄峰, 熊丽芳, 等. 大数据时代城市时空行为研究方法. *地理科学进展*, 2013, 32(9): 1352-1361. doi: 10.11820/dlkxjz.2013.09.005.]
- [5] Long Ying, Zhang Yu, Cui Chengyin. Identifying commuting pattern of beijing using bus smart card data. *Acta Geographica Sinica*, 2012, 67(10): 1339-1352. [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. *地理学报*, 2012, 67(10): 1339-1352.]
- [6] Zhou Suhong, Hao Xinhua, Liu Lin. Validation of spatial decay law caused by urban commercial center's mutual attraction in polycentric city: Spatio-temporal data mining of floating cars' GPS data in Shenzhen. *Acta Geographica Sinica*, 2014, 69(12): 1810-1820. [周素红, 赫新华, 柳林. 多中心化下的城市商业中心空间吸引衰减率验证-深圳市浮动车GPS时空数据挖掘. *地理学报*, 2014, 69(12): 1810-1820. doi: 10.11821/dlxb201412007.]
- [7] Wu Jiansheng, Huang Li, Liu Yu, et al. Traffic flow simulation based on call detail records. *Acta Geographica Sinica*, 2012, 67(2): 1657-1665. [吴健生, 黄力, 刘瑜, 等. 基于手机基站数据的城市交通流量模拟. *地理学报*, 2012, 67(2): 1657-1665.]
- [8] Lu Zi, Zhao Yahong, Wu Shifeng et al. The time distribution and guide analysis of visiting behavior of tourism website users. *Acta Geographica Sinica*, 2007, 62(6): 621-630. [路紫, 赵亚红, 吴士锋, 等. 旅游网站访问者行为的时间分布及导引分析. *地理学报*, 2007, 62(6): 621-630.]
- [9] Mao Xia, Xu Rongrong, Li Xinhua, et al. Fine grid dynamic features of population distribution in Shenzhen. *Acta Geographica Sinica*, 2010, 65(4): 443-453. [毛夏, 徐蓉蓉, 李新硕, 等. 深圳市人口分布的细网格动态特征. *地理学报*, 2010, 65(4): 443-453.]

- [10] Naaman M, Zhang A X, Brody S, et al. On the study of diurnal urban routines on Twitter. 6th International AAAI Conference on Weblogs and Social Media (ICWSM-12). Dublin, Ireland, 2012: 4-7.
- [11] Cheng Z, Caverlee J, Lee K, et al. Exploring millions of footprints in location sharing services//ICWSM' 11: Proceedings of The International AAAI Conference, 2011: 81-88.
- [12] Zhen Feng, Wang Bo, Chen Yingxue. China's city network characteristics based on social network space: An empirical analysis of Sina micro-blog. *Acta Geographica Sinica*, 2012, 67(8): 1031-1043. [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. *地理学报*, 2012, 67(8): 1031-1043.]
- [13] Liu Zhaoqing. City social relation in Yangtze River Delta based on the space of flows [D]. Shanghai: Shanghai Normal University, 2013. [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.]
- [14] Liu Yu, Sui Zhengwei, Kang Chaogui, Gao Yong. Uncovering patterns of inter-urban trip and spatial interaction from social media check-in data. *PLOS One*, 2014, 9(1): 1-11. e86026.
- [15] Krings G, Carabrese F, Ratti C, et al. Urban gravity: A model for inter-city telecommunication flows. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(7): 1-8. L07003.
- [16] Silm S, Ahas R. The seasonal variability of population in Estonian municipalities. *Environment and Planning A*, 2010, 42(10): 2527-2546. a43139.
- [17] Kang C, Zhang Y, Ma X, et al. Inferring properties and revealing geographical impacts of intercity mobile communication network of China using a subnet data set. *International Journal of Geographical Information Science*, 2013, 7(3): 431-448. <http://dx.doi.org/10.1080/13658816.2012.689838>.
- [18] Yuan Y, Raubal M, Liu Y. Correlating mobile phone usage and travel behavior: A case study of Harbin, China. *Computers, Environment and Urban Systems*, 2012, 36(2): 118-130. doi: 10.1016/j.compenvurbsys.2011.07.003.
- [19] Castells Maunul. Local and global: Cities in the network society. *Tijdschrift Voor Economische en Sociale Geografie*, 2002, 93(5): 548-558.
- [20] Goetz A R. Air passenger transportation and growth in the US urban system 1950-1987. *Growth and Change*, 1992, 23(2): 218-2421.
- [21] Matsumoto H. International urban systems and air passenger and cargo flows: Some calculations. *Journal of Air Transport Management*, 2004, 10(4): 239-247.
- [22] Zhou Yixing, Hu Zhiyong. Looking into the network structure of Chinese urban system from the perspective of air transportation. *Geographical Research*, 2002, 21(3): 276-286. [周一星, 胡智勇. 从航空运输看中国城市体系的空间网络结构. *地理研究*, 2002, 21(3): 276-286.]
- [23] Dai Teqi, Jin Fengjun, Wang Jiao'e. Spatial interaction and network structure evolvement of cities in terms of China's railway passenger flow in 1990s. *Progress in Geography*, 2005, 24(2): 80-89. [戴特奇, 金凤君, 王姣娥. 空间相互作用与城市关联网络演进: 以我国20世纪90年代城际铁路客流为例. *地理科学进展*, 2005, 24(2): 80-89.]
- [24] Shi Bole, Wang Wei, Xiao Yanghua. Information network mining and analysis: Opportunities and challenges. *Software Industry and Engineering*, 2014(4): 16-21. [施伯乐, 汪卫, 肖仰华. 信息网络挖掘与分析: 机遇与挑战. *软件产业与工程*, 2014(4): 16-21.]
- [25] Gu Chaolin, Pang Haifeng. Study on spatial relations of Chinese urban system: Gravity Model Approach. *Geographical Research*, 2008, 27(1): 1-12. [顾朝林, 庞海峰. 基于重力模型的中国城市体系空间联系与层域划分. *地理研究*, 2008, 27(1): 1-12.]
- [26] Qi Wei, Liu Shenghe, Zhao Meifeng. Study on the stability of Hu Line and different spatial patterns of population growth on its both sides. *Acta Geographica Sinica*, 2015, 70(4): 551-566. [戚伟, 刘盛和, 赵美凤. “胡焕庸线”的稳定性及其两侧人口集疏模式差异. *地理学报*, 2015, 70(4): 551-566. doi: 10.11821/dlxb201504004.]

Spatial pattern of population daily flow among cities based on ICT: A case study of "Baidu Migration"

LIU Wangbao, SHI Enming

(College of Geography Science, South China Normal University, Guangzhou 510631, China)

Abstract: With the advent of the Internet era, network data has become an important carrier characterizing residents' geography behavior. The residents' migration, social network, mobile communications and other geographic behavior big data have become an important data source for urban interactive relationship research. "Baidu Migration" big data can fully, dynamically, immediately and visually record population migration trajectory with LBS technology. Through collecting population daily flow among 369 cities in China during the period from February 7 to May 16 in 2015 in "Baidu Migration" and extracting six periods with "Quarter average, Spring Festival Transportation (before Festival), Spring Festival Transportation (after Festival), Labor Day, weekends, workdays", this paper tries to analyze and compare the characteristics and spatial patterns of daily flow among cities from the aspects of "population daily flow distribution levels, flow distribution layers network aggregation, spatial patterns and its relationship with the 'Hu Huanyong Line' for population flow". This paper analyzes the characteristics and spatial pattern of population daily flow among cities in each period. The result shows that "Baidu migration" big data clearly shows the pattern of population flow between the central and coastal areas in China during the period of Spring Festival Transportation. Main flow assembling centers are distributed in the urban agglomerations of Beijing-Tianjin-Hebei, Yangtze River Delta, Pearl River Delta and Chengdu-Chongqing, and those centers have strong coherence with those urban hierarchies in each period. Clear hierarchical structure and level distinction can be identified in the Yangtze River Delta, Pearl River Delta, Beijing-Tianjin-Hebei, Chengdu-Chongqing and Urumqi assembling systems in each period. But Central China, Northeast China, Southwest China and coastal Fujian do not present a pattern of higher level of urban assembling centers and high hierarchical assembling systems and those conditions do not match the status in national regional development strategy of those areas. The "Hu Huanyong Line" can reflect the spatial patterns of population daily flow at national level, and the profound influence of geographical environment on the population daily flow among cities. The intensity of population flow among cities is an important indicator of intensity of regional economic relationship, urban hierarchy and network structure. Spatial patterns of population flow showed in this paper can provide reference for the formation of the new pattern of regional economic development and the promotion of regional balanced development.

Keywords: "Baidu Migration"; population daily flow; assembling system; Spring Festival Transportation; China