

基于大数据与网络分析的长三角城市群识别研究*

Study on the Identification of Urban Agglomerations in the Yangtze River Delta Based on Big Data and Network Analysis

甄茂成 党安荣 阚长城 ZHEN Maocheng, DANG Anrong, KAN Changcheng

摘要 通过百度迁徙大数据,运用城市空间联系强度分析方法和复杂网络分析方法,构建了长三角城市群中的城际出行网络模型,测度并分析其空间结构。结果显示:(1) 城际出行网络形成“东强西弱、南北均衡”的格局态势,并形成以合肥、南京、上海、杭州、宁波为主的多中心辐射格局特征。(2) 长三角城市群的人口城际出行网络分布上呈现“一大三小”的多中心网络空间格局,上海、南京、杭州和合肥4个节点城市在整个网络中处于绝对主导地位。(3) 长三角各个城市在中心性、结构洞水平、核心—边缘结构等指标方面表现出较大的差异性。整体来看,上海、南京、杭州的网络地位较高,铜陵、宣城、池州、滁州、金华、台州、舟山等的网络地位较低。(4) 基于城际人口出行数据的长三角城市群空间结构,可以分为以上海、南京、杭州和合肥为中心的4个城市体系。

Abstract Based on the Baidu migration big data, by using the urban spatial connection analysis method and the complexity network analysis tools, this paper constructs the model of population flow network in urban agglomerations, and measures and analyzes the characteristics of the complex structure of the network. The research result shows that: (1) The population mobility network formed a migration pattern “strong in the east and weak in the west, balanced in the south and north” which is characterized by multi-center radiation pattern dominated by Hefei, Nanjing, Shanghai, Hangzhou and Ningbo. (2) The whole network presents an obvious multi-core network pattern which contains “one hubs and three subcenters”. The four node cities of Shanghai, Nanjing, Hangzhou and Hefei are absolutely dominant in the whole network. (3) Each city shows great differences in indicators such as centrality, structural hole level, and core-edge structure. On the whole, Shanghai, Nanjing and Hangzhou have higher network status, while Tongling, Xuancheng, Chizhou, Chuzhou, Jinhua, Taizhou and Zhoushan have lower network status. (4) Based on the intercity population migration data, the spatial structure of the Yangtze River Delta Urban Agglomerations can be divided into four urban systems with Shanghai, Nanjing, Hangzhou and Hefei as the centers.

关键词 长三角;大数据;城市群;百度迁徙;城际出行

Key words Yangtze River Delta; big data; urban agglomeration; Baidu migration; intercity travel

文章编号 1673-8985 (2019) 06-0008-09 中图分类号 TU981 文献标志码 A

DOI 10.11982/j.supr.20190602

作者简介

甄茂成

清华大学建筑学院 博士后

党安荣 (通信作者)

清华大学建筑学院

教授,博士生导师

清华大学人居环境信息实验室 主任

阚长城

百度时代网络技术(北京)有限公司

硕士,资深研发工程师

百度慧眼技术架构师

0 引言

城市群作为我国城市化推进过程中的主体形态和区域发展的重要空间单元^[1],对统筹整个区域的社会经济发展、基础设施布局和生态环境协调产生深远影响,是新型城镇化发展的重要空间对象。在国家层面,城市群具

有地域空间、网络空间、空间的生产与建构等多重属性,所产生的辐射带动作用已经成为拓展区域发展空间的一项重要手段^[2]。在城市和区域规划中,城市群空间结构分析是对城市群进行统筹规划的前提,通过识别与划定城市群空间结构,可以更加精确和科学地

*基金项目:本文得到政府间国际科技创新合作重点专项“数字城市规划新技术研发”(编号2017YFE0118600)子课题“数字城市规划技术集成方法研究”及清华大学中国新型城镇化研究院开放基金项目(编号TUCSU-K-17026-01)资助。

引导城市群的发展,促进城市群空间范围内人口、土地、产业、技术等要素的合理配置。然而,目前学术界和国家层面均未形成统一的城市群空间结构划分标准和规范,导致行政意义上的城市群与学术意义上的城市群存在较大偏差^[3]。因此,只有及时科学地识别城市群空间结构,才能诊断城市群发展中存在的各种问题,并提出相应的解决策略。

国内学者识别城镇群的方法主要有3种:一是选择城镇群之间关联的主要社会经济指标,根据相关标准设定临界值进行城镇群划定,通过城市群的主要特征构建一套反映社会经济发展状况的指标体系,并结合断裂点等数学模型来判断城市群发育程度^[4-5]。该方法存在指标选取随意性、对城市群空间信息表达不足等缺陷。二是通过断裂点模型、点—轴系统分析^[6]、引力模型^[7]、重力模型^[8]、水文模型^[9]等数学方法识别和计算核心城市的辐射范围,从而界定城市群边界^[10-11]。该方法过度强调城市规模所忽略的城市之间的联系强弱。三是集成指标筛选与引力模型构建的综合方法,并利用空间分析软件实现城镇群边界的识别^[12]。该方法的优点是技术路线比较完整,识别出的城市群客观性比较强,但不足之处也较明显,譬如空间可达性只能反映城市群特征的一个方面,人口、社会、经济因素对城市群的影响并未考虑其中。

上述传统的城市群识别多是依赖统计年鉴数据和路网数据^[13-14],是一种相对静态的识别,对人的活动、社会活动、经济活动等因素考虑不足,很难体现城市群网络空间的流动性和联系性。

近年来,时空大数据挖掘技术飞速发展,其所具有的数据规模巨大、数据类型多元、时空动态属性、人本视角和处理速度快等特点,能够帮助城市管理部门和学者更加精准化和量化地分析城市问题,制定规划方案。而带有地理位置属性的大数据,如社交网站数据^[15]、公交刷卡数据^[16-17]、手机信令数据^[18-19]、出租车GPS数据^[20]、微博签到数据^[21]等,广泛应用到城市地理学、行为地理学和交通地理学等

研究中。实际上,时空大数据技术为城市群空间范围的识别和空间结构分析提供了新的数据基础、方法和手段。例如,可以通过兴趣点(POI)、位置签到、导航轨迹、位置服务等数据,综合分析和识别城镇群的实际发展空间范围和虚拟活动空间范围,从而划定城镇群的实体边界和虚拟边界。鉴于此,本文试图提出一种科学识别和界定城市群空间结构的方法,利用百度迁徙人口大数据,以空间分析方法和网络分析方法为基础,进行长三角城市群的空间结构分析研究,以期为城市群界定和网络结构分析提供参考与帮助。

1 数据来源与研究方法

1.1 研究区域

研究区域为2016年5月国务院批复的长三角城市群,覆盖范围包括:上海市,江苏省的南京、无锡、常州、苏州、南通、盐城、扬州、镇江、泰州等9个城市,浙江省的杭州、宁波、嘉兴、湖州、绍兴、金华、舟山、台州等8个城市,安徽省的合肥、芜湖、马鞍山、铜陵、安庆、滁州、池州、宣城等8个城市。整个研究区域共计26个城市(见图1),是中国经济最具活力、城市层级结构最合理的城市群。

本文中人口迁徙大数据的数据收集与分析单元以地级市为主,与《长江三角洲城市群发展规划》所确定的26个城市相对应。值得注意的是,行政上的城市群与学术意义上的城市群并不一致。通过大数据对长三角城市群空间上的集聚度、城市间的联系强度、城市群网络结构等特征进行分析,审视长三角城市之间的联系程度,反映出已有城市群规划的发展现状与存在问题,对于了解该区域的城市空间体系结构具有重要意义,并能为长三角城市群的未来发展与决策提供依据。

1.2 数据来源

本文采用的数据为“百度迁徙”中的统计数据。“百度迁徙”依托百度云数据分析平台,基于定位服务(Location based service, LBS)大数据进行计算分析,通过

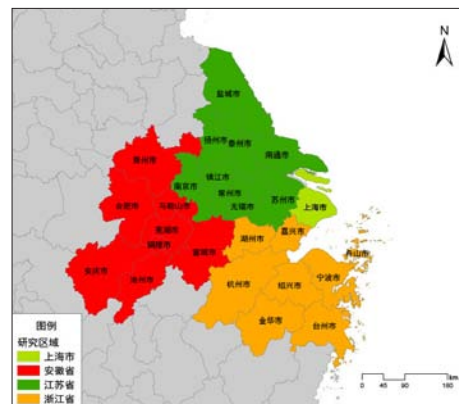


图1 研究区域范围
资料来源:笔者自绘。

手机用户的定位信息实现个人地理空间移动的可视化,整合体现不同时间尺度内人口流动的空间节点和不同节点之间的流动规模,进而勾勒出中国人口大迁徙的轨迹与特征。从“百度迁徙”中可以获取全年某一时间节点我国各城市对外人口流动的主要方向与相对大小,为研究城市或区域之间的人口联系提供了良好的一手数据资料^[22-23]。用百度迁徙数据进行城市群空间结构研究具有明显优势:首先,以城市为单元比较适合大区域的人口流动研究;其次,真实的人口流动是有效地反映城市群网络结构的重要指标;最后,依托发达的网络技术和定位技术,“百度迁徙”大数据具有更新快、采集简单的特征,相比于传统的人口流动数据更具有时效性。本文采用的人口城际出行数据由百度慧眼提供,数据采集于2019年4月10日至4月24日,26个城市之间的总迁徙量为21 815 807人次。每一次人口流入量和流出量均不同,将各城市人口的流入量和流出量求和(见表1),可以看出大多数城际人口的流入量和流出量基本持平。

1.3 技术路线

本文的技术路线如图2所示,主要分为3个步骤^[24]:(1)通过空间分析方法,结合城际出行数据计算城市联系强度,据此识别城市群空间结构;(2)基于复杂网络,对城市网络的度、平均路径长度、聚类系数、核心—边缘结构等进行分析计算,得出城市群空间结构

表1 各城市人口流入量和流出量 (单位:人次)

城市	流出量	流入量	城市	流出量	流入量
舟山市	188 874	270 076	宁波市	860 998	971 626
镇江市	682 429	750 608	南通市	592 206	785 199
扬州市	565 736	646 545	南京市	1 575 892	1 827 766
盐城市	473 425	527 431	马鞍山市	402 209	390 030
宣城市	329 214	364 557	金华市	583 230	627 718
芜湖市	463 046	475 978	嘉兴市	911 218	1 230 822
无锡市	1 369 523	1 605 140	湖州市	678 218	852 179
铜陵市	160 823	163 704	合肥市	806 399	872 017
泰州市	593 347	655 120	杭州市	2 142 646	2 587 026
台州市	429 018	415 860	滁州市	429 914	451 334
苏州市	1 711 826	2 738 118	池州市	152 779	198 718
绍兴市	953 603	1 081 056	常州市	975 578	1 065 551
上海市	3 514 393	3 314 083	安庆市	269 263	261 628

资料来源:笔者根据百度慧眼提供数据整理。

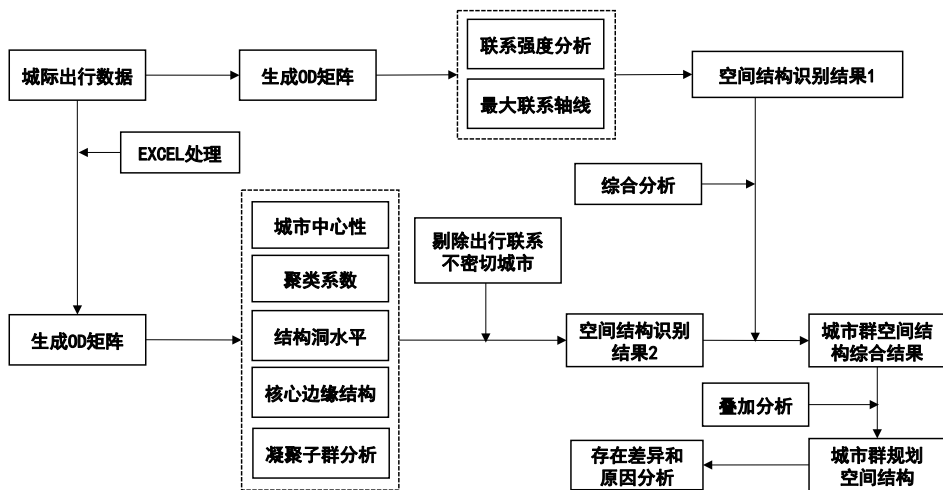


图2 基于人口城际出行数据的城市群空间结构分析技术路线
资料来源:笔者自绘。

识别结果; (3) 根据上述两者识别结果, 综合得出基于城际出行数据的城市群空间结构, 并与规划范围进行叠加分析, 找出差异。

1.4 研究方法

本文主要运用复杂网络技术, 通过建立城市间人口出行流向的矩阵关系模型, 研究城际人口出行网络复杂的流向关系和节点联系^[25]。对复杂网络结构进行刻画的主要参数包括: K-核、度、聚类系数和凝聚子群等。其中, 网络的密度主要用于刻画网络连接的总体特征, K-核和聚类系数多用于描述网络的集中性, 而凝聚子群则用于描述城市之间人

口流动而形成的组团现象。

(1) 网络的度 (Degree)

度是指与节点相连接的邻近节点的数目, 通常采用累积分布来描述网络中站点的度分布状况。其公式为:

$$P(K) = \int P(K) dx \quad (1)$$

网络的度值越大, 说明该节点的联系性越强, 中心性越突出。

(2) 接近中心度 (Closeness centrality)

接近中心度表示网络中某个行动者与其他行动者的接近程度, 即行动者能使用尽可

能短的路径与其他所有行动者产生连接。在对城市群的研究中, 接近中心度主要测量的是一个城市在具体事项上对于其他城市的主导能力。其值越大, 表明城市更大地掌握了区域合作的主动权。接近中心度模型为:

$$C_C(c_i) = (n-1) / \sum_{j=1}^n d_i(c_i, c_j) \quad (2)$$

$d_i(c_i, c_j)$ 代表城市i和城市j之间最短路径的数量, n表示城市的数量, 即网络中结点城市数量的多少。

接近中心势指整体网络图的接近集中趋势, 比如星型网络具有100%的接近中心势, 而对于完备网络、环形网络来说, 其接近中心势为0。在本文中, 接近中心势主要代表网络中的行动者在多大程度上受到少数城市控制。

(3) 聚类系数 (Clustering coefficient)

在社会网络中, 一般采用聚类系数 c_i 来反映网络局部聚类情况的参数, 它表征社会网络中因某一共同关系 (如兴趣爱好、共同利益等) 而形成的社会团体关系网络, 整个网络的聚类系数C就等于网络中所有节点聚类系数的平均值。

$$C = \frac{1}{n} \sum_i c_i \quad (3)$$

2 实证研究与结果分析

2.1 基于城市间联系强度分析的城市群识别

2.1.1 城市间联系强度分析

百度慧眼提供的城际出行数据的数量表征大小, 是个相对值。本文将人口城际出行数量作为衡量城市间联系强度的重要指标, 城市间具有双向的人口出行包括城市间的人口流入量和流出量, 用公式表达如下:

$$H = M_1 + M_2 \quad (4)$$

通过对照网络中26个城市人口的流入量和流出量, 发现两组数据排序的相对一致性。经 SPSS相关性检验, 相关系数 R^2 高达 0.93,

可以看出长三角城市间人口的流入量和流出量呈现出显著的正相关关系。

本文用流入量和流出量的和来表征城市间联系强度。利用GIS软件对人口城际出行数据进行预处理。首先,根据城市的起点和终点的经纬度坐标,生成26个城市的OD轴线。其次,用融合工具(Dissolve)将2周内所有两两有人口迁徙的市县OD轴线融合成一条,用以表征任意两个城市之间的联系强度的大小(见表2),展示了两个城市间单向的联系强度。

将城市联系强度的计算结果进行空间可视化,提取出行量大于5 000的城际人口出行数量进行分析,可以发现各个城市联系强度的格局呈现如下特征。

(1) 行政等级越高的城市在城际人口出行网络中地位越高,影响范围越广,处于迁徙总量顶端的城市是上海、杭州、南京和苏州。其中,上海是中国最大的经济中心城市,在资源、人才、创新等方面具有绝对优势,对周边地区进行辐射,形成上海—苏州、上海—南通、上海—嘉兴、上海—杭州为主的1级和2级联系强度;杭州作为中南区域的辐射中心,形成了以杭州—湖州、杭州—金华、杭州—绍兴、杭州—嘉兴等为主的2级联系强度;南京为西部区域的辐射中心,空间上也为由单点向多点发散辐射的格局,形成南京—镇江、南京—盐城、南京—常州为主的2级和3级联系强度;合肥是安徽的政治、经济中心,辐射带动作用强,与周边其他城市的联系水平较高,形成以合肥—滁州、合肥—马鞍山、合肥—芜湖、合肥—安庆为主的3级水平联系强度。

(2) 人口城际出行网络形成“东强西弱,南北均衡”的格局态势。东部区域云集上海、苏州、南通、嘉兴等城市,形成了1级和2级为主的联系强度,其中上海的联系强度处于绝对领先地位。西部区域除合肥、安庆具备3级联系强度外,池州、宣城、铜陵、滁州等城市联系强度较低,均为4级。北部区域和南部区域的城市间联系强度较为均衡,其中北部区域中的南京、盐城、镇江形成了以2级和3级联系强度

表2 部分城市联系强度计算结果(单位:人次)

城市A—城市B	强度	城市A—城市B	强度
舟山市—宁波市	85 402	无锡市—苏州市	450 357
舟山市—上海市	60 805	无锡市—常州市	331 518
舟山市—杭州市	33 522	无锡市—上海市	201 179
舟山市—苏州市	11 271	无锡市—南京市	105 367
舟山市—台州市	9 364	无锡市—泰州市	79 812
舟山市—嘉兴市	9 263	无锡市—南通市	51 705
舟山市—绍兴市	7 585	铜陵市—合肥市	38 217
镇江市—南京市	272 222	铜陵市—安庆市	29 245
镇江市—常州市	152 001	铜陵市—芜湖市	27 538
镇江市—扬州市	56 396	铜陵市—池州市	26 753
镇江市—苏州市	45 881	铜陵市—南京市	8 872
镇江市—无锡市	45 101	泰州市—扬州市	111 357
镇江市—南京市	272 222	泰州市—无锡市	81 349
镇江市—常州市	152 001	泰州市—盐城市	80 534
扬州市—舟山市	1 254	泰州市—南通市	71 745
扬州市—镇江市	60 355	泰州市—苏州市	67 533
扬州市—盐城市	26 567	泰州市—上海市	67 312
扬州市—宣城市	2 405	泰州市—南京市	66 750
扬州市—芜湖市	4 031	台州市—宁波市	146 558
扬州市—无锡市	38 083	台州市—杭州市	105 552
扬州市—铜陵市	858	台州市—金华市	54 733
盐城市—上海市	90 697	台州市—上海市	50 218
盐城市—泰州市	90 081	台州市—绍兴市	31 264
盐城市—苏州市	82 418	苏州市—上海市	1 039 164
盐城市—南通市	74 652	苏州市—无锡市	461 970
盐城市—南京市	71 116	苏州市—南通市	150 354
盐城市—无锡市	48 851	苏州市—嘉兴市	147 211
盐城市—扬州市	29 087	上海市—苏州市	1 095 745
宣城市—芜湖市	56 632	上海市—杭州市	376 605
宣城市—南京市	45 210	上海市—嘉兴市	281 101
宣城市—湖州市	44 759	上海市—南通市	242 057

资料来源:笔者根据百度慧眼提供数据整理。

为主的格局。南部区域的城市联系强度整体较低,形成了以宁波为主的3级联系网络(见图3)。

(3) 省内城市联系强度高于省区城市,上海与江苏、浙江的跨区域联系强度等级高。安徽省内合肥、滁州、安庆、芜湖、马鞍山等城市除了与南京的联系强度较为明显外,与其他城市的跨区域联系均较弱。江苏省的南京、镇江、泰州、盐城、南通等城市的联系强度为3级,其中,南通、苏州与上海的跨区域联系强度较高。浙江省的杭州、湖州、嘉兴的联系强度较高为2级,嘉兴和杭州与上海的跨区域联系强度较高。同时,上海由于其行政地位高、经济实力强、人口数量多,导致其与邻近的苏州、嘉兴、南通联系强度较高。

(4) 整体网络格局上,研究区域具有多中心辐射格局的特征。形成了以合肥、南京、上海、杭州、宁波为主的多中心格局。这几个中心城市是整个城际人口出行网络中的核心节点(见图4)。

2.1.2 首位联系分析

统计出与每个城市联系强度最大的城市,将其作为首位联系城市,并通过OD分析生成首位联系轴线,用以表征人口城际出行的空间联系特征。如图5展示的首位联系轴线,可以得出:(1)从区域层面来看,安徽省和浙江省的首位联系轴线以省内为主,并未出现跨区域联系轴线。而江苏省的首位联系轴线跨区域明显,分别与上海、滁州、马鞍山

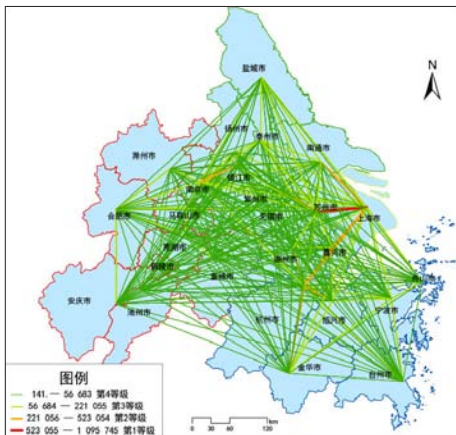


图3 城际出行联系强度
资料来源:笔者自绘。

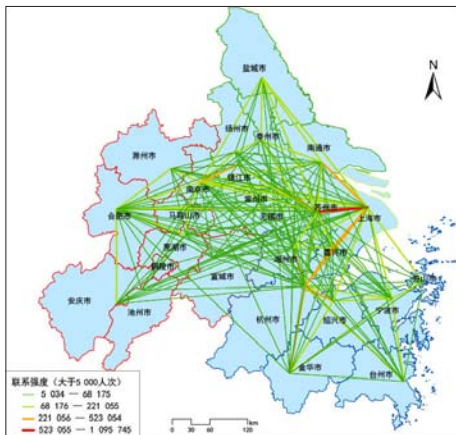


图4 城际出行联系强度 (大于5 000人次)
资料来源:笔者自绘。

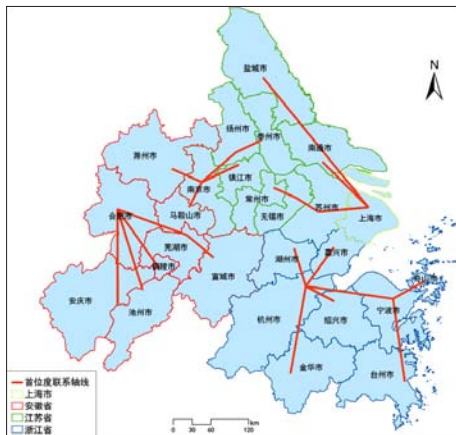


图5 首位联系轴线图
资料来源:笔者自绘。

等城市相连接。(2) 从省域尺度来看,安徽省内共有7条首位联系轴线(合肥占据4条),并以合肥为辐射中心,表现出绝对的首位城市地位。江苏省共有9条首位联系轴线,形成了以上海和南京为辐射中心的首位联系轴线格局。浙江省共有7条首位联系轴线,形成了以杭州和宁波为辐射中心的首位联系轴线格局。(3) 整个区域的首位度连接网络呈现了多中心网络的空间格局,上海、南京、杭州、合肥以及宁波作为网络中的重要节点,在人口、资源集聚和辐射方面发挥了重要作用。

为避免界定的城市群范围过大,本文采用阈值法对首位联系法所得结果进行二次筛选,以最大联系强度的10%和5%阈值做筛选,不会影响数据的质量。根据10%阈值筛选首位联系城市,如公式(4),形成新的首位联系轴线图(见图6)。结果显示:各个省的首位联系轴线数量均有所下降。其中安徽由7条降为4条,江苏由9条降为8条,浙江由7条降为6条。

$$K = 10\% \times MAX(H) \quad (5)$$

式中为K首位联系强度,MAX(H)为城际出行中的最大联系强度。根据上述城市群空间结构分析方法,将研究区域划分为4个城市体系。第一个是以上海为辐射核心的城市体系,包括上海、常州、无锡、苏州、南通5市;第二个是以南京为辐射核心的城市体系,包括

南京、滁州、马鞍山、镇江、扬州和泰州6市;第三个是以杭州为辐射核心的城市体系,包括杭州、湖州、绍兴、金华和嘉兴5市;第四个是以合肥为辐射核心的城市体系,包括合肥、安庆和巢湖3市。可以看出,基于10%门槛法的城市群空间结构分析结果对于以合肥为中心的城市组群的大小以及宁波为核心的城市体系的形成都产生了较大影响。

2.2 基于城市网络分析的城市群识别

2.2.1 城市中心性分析

构建研究区域内城市之间的人口出行矩阵,并采用Ucinet软件分析工具对矩阵进行相应计算,可以得出城市网络的结构特征参数,再借助GIS将结果进行可视化,表征各城市中心性的空间分布。

(1) 度数中心度

度数中心度(Degree)用来衡量某一个城市与其他城市交往能力的大小。值越大,中心性越强,支配力越强。以26个城市整体度数中心度平均值46.5为阈值,将其划分为前9个中心影响力较强的城市(上海、苏州、杭州、南京、无锡、合肥、常州、扬州、湖州)和后17个中心影响力较弱的城市(宣城、金华、安庆、池州、滁州等)。在中心影响力较强的9个城市中,上海、苏州、南京、杭州、合肥的度数中心度排在前列,结合上一段分析,印证了这几个中心城市是人口城际出行网络中等级较高的

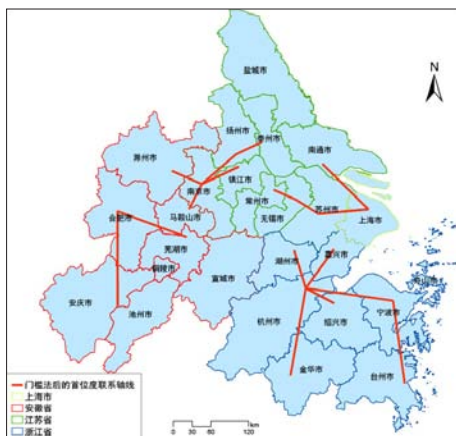


图6 10%门槛后的首位联系轴线
资料来源:笔者自绘。

城市。对各个城市的度数中心度进行可视化展示,划分为4个等级,呈现如下特征:(1) 在整个人口城际出行网络中,网络地位最高的是上海,其次是南京、杭州。这3个城市由于其行政地位高、经济实力强、人口众多,在长三角地区处于核心地位,人口的流入量和流出量大,与网络中其他城市间的联系能力强,相应的支配能力也大。(2) 网络地位最低的城市为安徽省的铜陵、宣城、池州、滁州和安庆等城市以及浙江省的金华、绍兴、舟山、宁波等南部城市。尤其是安徽省的城市加入长三角城市群时间不长,加上自身的人口、经济、规模等发展有限,在网络中的支配能力较弱。(3) 苏州、无锡、常州、扬州、湖州5个城市的平均度数中心度较高,且城市间的度数中心度差异最低,表明这些城市人口城际出行联系较强,且在网

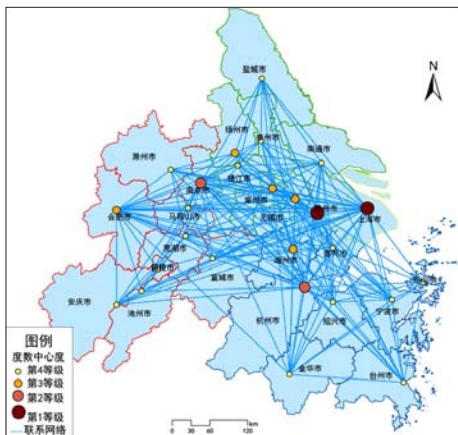


图7 各城市度数中心度格局
资料来源:笔者自绘。

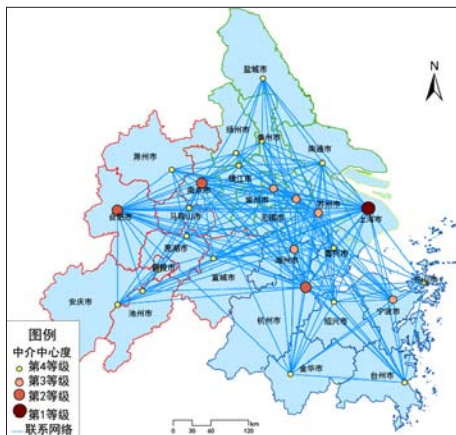


图8 各城市中介中心度格局
资料来源:笔者自绘。

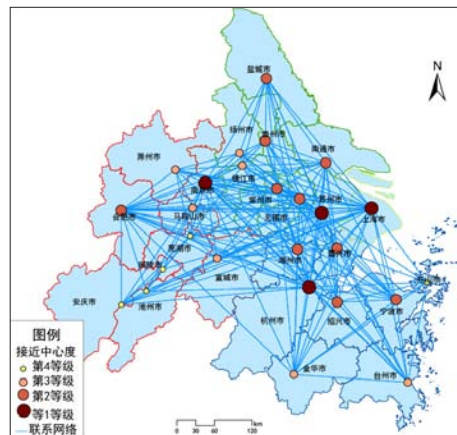


图9 各个城市接近中心度格局
资料来源:笔者自绘。

络中也发挥了较为重要的作用(见图7)。

(2) 中介中心度分析

中介中心度是指一个结点担任其他两个结点之间最短路的桥梁的次数,可以用来衡量某一个城市控制网络中其他城市之间的交往能力。一个结点充当“中介”的次数越高,其中介中心度就越大,其他城市对于该城市的依赖度便越高。通过Ucinet软件分析得到每个城市的中介中心度值,进行空间可视化后,可以发现5个方面特征。第一,上海的中介中心度以120.48位居长三角城市群之首,属于唯一的1级中介中心度,并且明显高于排名第二的合肥(83.58)。这说明上海在整个网络中处于垄断地位,在研究区域中的“桥梁”作用最强。第二,合肥、南京、杭州3个省份城市都是所在省域内最重要的中介城市,这3个城市在各省内部的城市之间的人口出行和联系中发挥重要作用,其他城市之间的联系多要依赖它们。第三,苏州、无锡、常州、嘉兴、湖州等城市具有3级中介中心度,它们在长三角城市群的城际人口出行中也发挥一定作用。第四,安徽省的池州、滁州、宣城等城市以及浙江省的金华、台州、舟山等南部城市处于长三角城市群边缘位置,这些城市中中介中心度水平较低,在整个网络联系中处于弱势地位,发挥作用有限。第五,研究区域城市的中介中心度网络格局呈现“一大三小”的格局,具体来讲,“一大”指上海,“三小”指杭州、南京、合肥,均为各区

域城市体系中的龙头城市,体现网络核心作用(见图8)。

(3) 接近中心度

通过分析26个城市的接近中心度可以发现,上海的接近中心度最高,指数达到100。苏州、南京和杭州次之,接近中心度指数都达到80。马鞍山、滁州、芜湖、铜陵、池州、舟山等城市的接近中心度最低。这表明上海、苏州、南京、杭州与其他城市之间的联系最强,人口城际出行频繁,联系紧密居于网络核心地位。而芜湖、铜陵、池州、安庆等则与长三角其他城市之间的联系较弱。图9直观反映了安徽省除合肥以外其余7个城市的接近中心度均处于最低水平,与其他城市的联系较薄弱。

2.2.2 聚类系数分析

聚类系数反映网络中节点互联水平,即与某节点相连接的两个节点是否存在连接关系。结果显示,长三角城市网络的平均聚类系数为0.668,表现出较强的集聚性。通过分析网络中每个节点的聚类系数,发现高度值节点如上海、苏州、南京等聚类系数较低,分别为0.261、0.333、0.349,表明在网络中低度节点主要倾向于与高度节点连接,从而导致产生较高的聚类系数。较高的聚类系数说明尽管大部分的节点彼此并不相连,但节点之间经过少数几步便可能发生联系,网络结构相对稳定,人口城际出行较为便捷。

2.2.3 结构洞水平分析

结构洞是指网络中的某些个体与其他个体有直接联系,但也与另一些个体不发生直接联系或关系间断的现象,从网络整体上看好像网络结构中出现了洞穴。结构洞是研究城市网络中个体在群体中关键位置和作用的方法。通过计算,得出长三角城际出行网络的结构洞水平具有几个特征(见表3)。第一,上海、苏州、杭州、南京、合肥等节点的结构洞水平较高,受限程度较低。这些城市在长三角城市群中处于枢纽位置,最能控制其他节点之间的联系。第二,马鞍山、滁州、舟山、铜陵、池州等节点的有效规模和效率得分较低,受限程度较大,表明这些城市处于网络的边缘位置。图10直观反映了各个城市的结构洞网络呈现“一大三小”(上海和南京、杭州、合肥)的特征,进一步体现“一大三小”核心节点在城际出行网络中的重要地位。此外,池州、宣城、芜湖、滁州等城市以及金华、台州、舟山等城市的结构洞水平较低,处于整个网络的边缘位置。

2.2.4 核心—边缘结构分析

核心—边缘结构是由若干元素相互联系构成的一种中心紧密相连、外围稀疏分散的特殊结构。而城市间的人口联系强度分析是识别城市网络中核心与边缘位置的一个重要依据。对各个城市人口迁徙数据进行核心—边缘结构分析,发现核心区域包括8个节点,分别是上海、南京、杭州、常州、湖州、嘉兴、苏

表3 各个城市的结构洞指标

城市	有效规模	效率	限制度	城市	有效规模	效率	限制度
马鞍山市	2.450	0.408	0.560	盐城市	2.235	0.224	0.371
滁州市	2.773	0.396	0.496	泰州市	2.389	0.239	0.367
舟山市	2.000	0.25	0.464	芜湖市	4.750	0.475	0.356
铜陵市	3.000	0.375	0.457	宣城市	5.300	0.482	0.332
池州市	2.875	0.359	0.452	湖州市	5.632	0.469	0.317
绍兴市	2.267	0.283	0.441	扬州市	4.375	0.365	0.311
金华市	2.393	0.266	0.412	常州市	5.114	0.393	0.290
镇江市	1.429	0.159	0.411	无锡市	6.771	0.451	0.266
安庆市	3.808	0.423	0.407	合肥市	7.717	0.551	0.262
台州市	2.607	0.290	0.405	南京市	10.129	0.563	0.212
南通市	1.735	0.193	0.398	杭州市	12.086	0.604	0.201
宁波市	2.853	0.317	0.391	苏州市	14.455	0.628	0.177
嘉兴市	3.529	0.353	0.371	上海市	16.270	0.651	0.161

资料来源:笔者自制。

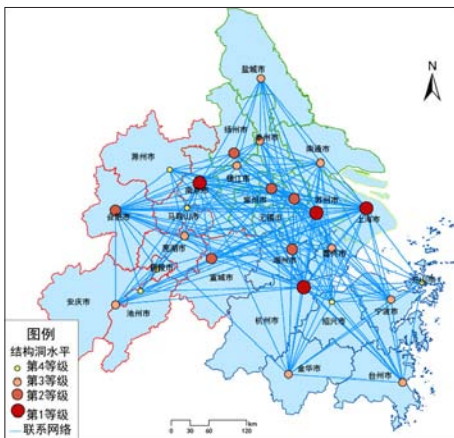


图10 各个城市的结构洞水平

资料来源:笔者自制。

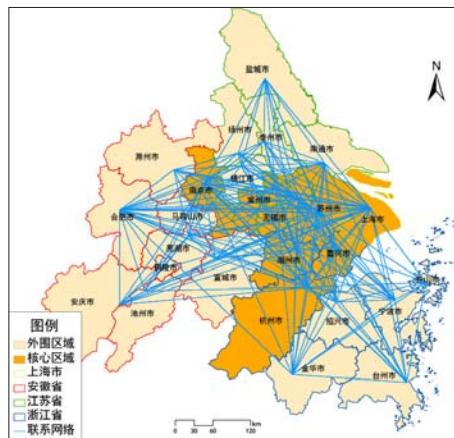


图11 各个城市核心—边缘格局

资料来源:笔者自制。

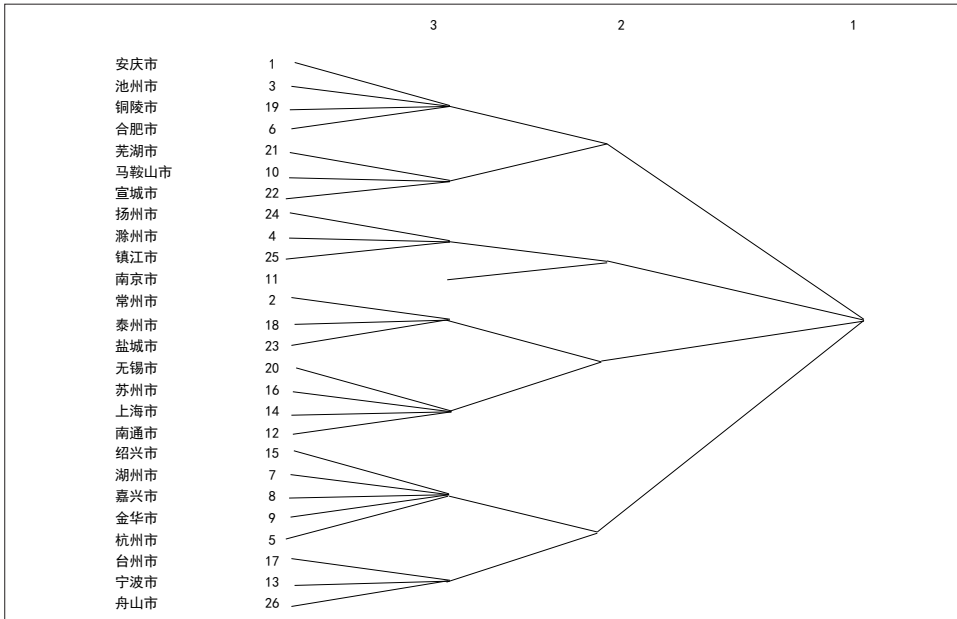


图12 凝聚子群分析结果

资料来源:笔者自制。

州和无锡,边缘节点包括18个节点,分别是安庆、池州、滁州、合肥、金华、马鞍山、南通、宁波、绍兴、台州、泰州、铜陵、芜湖、宣城、盐城、扬州、镇江和舟山。从图11可以看出,研究区域内城市所体现的核心—边缘格局具有以下特征:各个核心城市的人口出行联系强度高,在空间上表现出邻近性和连续性。同时,从整个区域的人口城际出行强度来看,安徽省的城市刚刚加入长三角城市群,目前还处于边缘位置。浙江省的金华、绍兴、舟山、宁波等南部城市也处于边缘位置。

2.2.5 凝聚子群分析

在社会网络中,凝聚子群是指网络中存在着一些行动者集合,这些集合中行动者间具有较强的、直接的、紧密的、经常的或积极的关系,其不仅影响组织成员之间的关系,还影响整个组织的发展。因而,凝聚子群内部成员之间的联系比较密切。凝聚子群分析适于揭示城市群内部各个城市之间的人口城际出行而形成的组团现象。在上述人口城际出行联系强度分析的基础上,有必要探讨整个区域各个子团结构及其构成,结果如图12,可以得出:(1) 凝聚子群具有明显的内部集聚和空间邻近特征,每个子群的成员基本由邻近城市构成。(2) 综合考虑度数中心度分析结果以及核心—边缘结构分析结果,剔除掉中心性低且处于边缘位置的城市。故将舟山、宁波、台州等中心性低且位置边缘的城市排除,将其构成的凝聚子群剔除。同样地,马鞍山、芜湖、宣城等城市构成的凝聚子群也要删除。(3) 基于复杂网络结构识别的城市体系为4个:第一个是上海为核心的城市体系,包括上海、无锡、常州、苏州、南通和泰州6市;第二个是以南京为核心的城市体系,包括南京、扬州、镇江和滁州4市。第三个是以杭州为核心的城市体系,包括杭州、绍兴、湖州、金华和嘉兴5市。第四个是以合肥为核心的城市体系,包括合肥、安庆、池州、铜陵、芜湖等5市。

2.3 城市群空间结构综合分析结果

利用百度迁移数据作为数据源,基于城

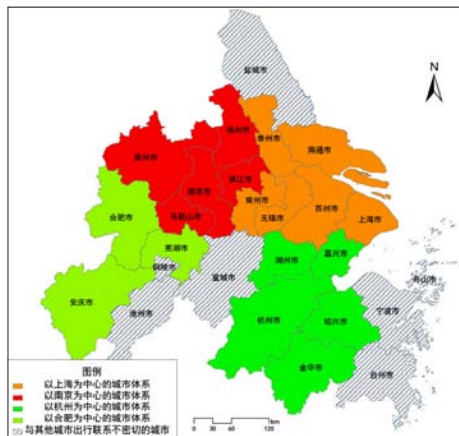


图13 基于城际人口出行数据的长三角城市群空间结构识别结果
资料来源:笔者自绘。

市联系强度的空间格局识别结果与基于复杂网络结构的识别结果基本吻合,对两者进行综合,最终得出基于城际人口出行数据的城市群识别结果。

(1) 以上海为辐射核心的城市体系:上海、常州、无锡、苏州、南通、泰州。

(2) 以南京为辐射核心的城市体系:南京、扬州、镇江、滁州、马鞍山。

(3) 以杭州为辐射核心的城市体系:杭州、绍兴、湖州、金华、嘉兴。

(4) 以合肥市为核心的城市体系:合肥、芜湖、安庆。

将识别结果与规划范围进行叠加分析(见图13),可以看到,基于城际出行数据的识别结果覆盖了大部分研究区域,规划范围内的盐城、铜陵、池州、宣城、宁波、台州、舟山等城市与几个中心城市的出行联系并不紧密。主要原因可能有二:一是基于人口城际出行的视角来判断,宁波、台州、舟山之间的人口出行联系强度较弱。二是基于人口城际出行数据的识别结果受到数据统计单元的影响比较大。由于数据的统计单元是地级市,并未达到区县级或者乡镇级,可能导致识别结果的精度不高。

3 结论与讨论

城际人口出行数据在城市群空间结构分析中有重要的作用。本文通过百度迁徙海量

个体居民出行信息数据,来构建城市之间的联系网络,采取基于GIS城市空间联系强度的方法和基于Ucinet的复杂网络方法,构建了长三角城市群城际人口出行网络模型,并分析其复杂的空间结构和网络结构。这是对城市网络研究的一个新的尝试和补充。研究发现:

(1) 基于城市空间联系强度对城市群空间结构的研究,发现城际出行网络显现出明显的层次性。行政等级越高的城市在人口出行网络中的地位越高,影响范围越广,处于出行总量顶端的城市是上海、杭州、南京和苏州。长三角城市群中的城际出行网络形成“东强西弱,南北均衡”的格局态势,并形成了以合肥、南京、上海、杭州、宁波为主的多中心辐射格局特征。

(2) 运用复杂网络计算长三角城市群26个城市人口城际出行的中心性、结构洞水平、核心—边缘结构等指标。中心性方面,各个城市的网络地位差异很大,网络地位最高的依次是上海、南京和杭州,网络地位最低的为铜陵、宣城、池州、滁州、金华、台州、舟山等城市;结构洞水平方面,上海、苏州、杭州、南京、合肥等节点的结构洞水平较高,受限程度较低;核心—边缘结构方面,上海、南京、杭州、常州、湖州、嘉兴、苏州、无锡等8市属于核心节点。

(3) 综合考虑城市节点层级分布与城市间网络联系强度,长三角城市群城际人口出行网络分布上呈现“一大三小”的多中心网络空间格局,尤其是在中心性和结构洞水平上表现得更为明显。这4个节点城市在整个网络中处于绝对主导地位,与网络中其他城市间的联系能力强,相应的支配能力也大。

(4) 基于城际人口出行数据的长三角城市群空间结构可以分为4个城市体系:以上海市为辐射核心的城市体系、以南京市为辐射核心的城市体系、以杭州市为辐射核心的城市体系和以合肥市为核心的城市体系。规划范围内的盐城、铜陵、池州、宣城、宁波、台州、舟山等城市与几个中心城市的出行联系并不紧密。

本文利用百度迁徙大数据,构建城际人口出行网络模型,为我们研究大数据驱动下的城市群空间结构提供新的分析框架、模型参考和方法借鉴。但本文仍存在几点局限:第一,数据方面。本文的数据覆盖对象具有一定局限性,百度迁徙数据并不包括人口的社会属性,人口城际出行和网络结构形成的内在机制还有待探究。同时,未来的工作需要将商业设施、公共基础设施等POI数据以及微博、QQ等热门手机软件数据应用到城市群空间结构研究中,丰富数据源,使结果更加贴近实际。第二,范围方面。本文只在规划研究区域内进行长三角城市群空间结构分析,并未考虑与城市群之外的城市间人口出行网络特征,未来工作可将研究区域周边城市纳入研究对象内。

参考文献 References

- [1] 顾朝林. 城市群研究进展与展望[J]. 地理研究, 2011, 30 (5): 771-784.
GU Chaolin. Study on urban agglomeration: progress and prospects[J]. Geographical Research, 2011, 30 (5): 771-784.
- [2] 苗长虹, 胡志强. 城市群空间性质的透视与中原城市群的构建[J]. 地理科学进展, 2015 (3): 271-279.
MIAO Changhong, HU Zhiqiang. The nature of urban agglomeration space and the construction of Zhongyuan urban agglomeration[J]. Progress in Geography, 2015 (3): 271-279.
- [3] 陈群元, 宋玉祥. 城市群空间范围的综合界定方法研究——以长株潭城市群为例[J]. 地理科学, 2010, 30 (5): 660-666.
CHEN Qunyu, SONG Yuxiang. Methods of dividing the boundary of urban agglomerations: Chang-zhu-tan urban agglomeration as a case[J]. Scientia Geographica Sinica, 2010, 30(5): 660-666.
- [4] 周一星, 史育龙. 建立中国城市的实体地域概念[J]. 地理学报, 1995, 50 (4): 289-301.
ZHOU Yixing, SHI Yulong. Toward establishing the concept of physical urban area in China[J]. Acta

- Geographica Sinica, 1995, 50 (4): 289-301.
- [5] 方创琳. 城市群空间范围识别标准的研究进展与基本判断[J]. 城市规划学刊, 2009, 182 (4): 1-6.
FANG Chuanglin. Research progress and general definition about identification standards of urban agglomeration[J]. Urban Planning Forum, 2009, 182 (4): 1-6.
- [6] 高晓路, 许泽宁, 牛方曲. 基于“点—轴系统”理论的城市群边界识别[J]. 地理科学进展, 2015, 34 (3): 280-289.
GAO Xiaolu, XU Zening, NIU Fangqu. Delineating the scope of urban agglomerations based upon the Pole-Axis theory[J]. Progress in Geography, 2015, 34 (3): 280-289.
- [7] 邓羽, 刘盛和, 蔡建明, 等. 中国中部地区城市影响范围划分方法的比较[J]. 地理研究, 2013, 32 (7): 1220-1230.
DENG Yu, LIU Shenghe, CAI Jianming, et al. Comparative study of identification methods to delineate urban spheres of influence in Central China[J]. Geographical Research, 2013, 32 (7): 1220-1230.
- [8] 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[J]. 地理研究, 2008, 27 (1): 1-12.
GU Chaolin, PANG Haifeng. Study on spatial relations of Chinese urban system: gravity model approach[J]. Geographical Research, 2008, 27 (1): 1-12.
- [9] 潘竟虎, 戴维丽. 基于空间场—水文模型的城市影响腹地界定及其空间演变分析: 以河南省地级城市为例[J]. 地球信息科学学报, 2014, 16 (5): 754-761.
PAN Jinghu, DAI Weili. Delimitation and evolvement of urban hinterland area in Henan Province based on spatial field and hydrologic model[J]. Journal of Geo-Information Science, 2014, 16 (5): 754-761.
- [10] 黄建毅, 张平宇. 辽中城市群范围界定与规模结构分形研究[J]. 地理科学, 2009, 29 (2): 181-187.
HUANG Jianyi, ZHANG Pingyu. Delimitation and fractal research on structure of entral Liaoning urban agglomeration[J]. Scientia Geographica Sinica, 2009, 29 (2): 181-187.
- [11] 王丽, 邓羽, 牛文元. 城市群的界定与识别研究[J]. 地理学报, 2013, 68 (8): 1059-1070.
WANG Li, DENG Yu, NIU Wenyuan. The definition and identification of urban agglomerations[J]. Acta Geographica Sinica, 2013, 68 (8): 1059-1070.
- [12] 许泽宁, 高晓路. 基于电子地图兴趣点的城市建成区边界识别方法[J]. 地理学报, 2016, 71 (6): 928-939.
XU Zening, GAO Xiaolu. A novel method for identifying the boundary of urban built-up areas with POI data[J]. Acta Geographica Sinica, 2016, 71(6): 928-939.
- [13] 侯贺平, 刘艳芳, 李纪伟, 等. 基于改进辐射模型的乡镇人口流动网络研究[J]. 中国人口、资源与环境, 2013, 23 (8): 107-114.
HOU Heping, LIU Yanfang, LI Jiwei, et al. Study on population mobility network among towns based on improved radiation model[J]. China Population, Resources and Environment, 2013, 23 (8): 107-114.
- [14] 王珏, 陈雯, 袁丰. 基于网络分析的长三角地区人口迁移及演化[J]. 地理研究, 2014, 33 (2): 385-400.
WANG Jue, CHEN Wen, YUAN Feng. Human mobility and evolution based on social network: an empirical analysis of Yangtze River Delta[J]. Geographical Research, 2014, 33(2): 385-400.
- [15] 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, 67 (8): 1031-1043.
ZHEN Feng, WANG Bo, CHEN Yingxue. China's city network characteristics based on social network space: an empirical analysis of sina micro-blog[J]. Acta Geographica Sinica, 2012, 67 (8): 1031-1043.
- [16] 龙瀛, 张宇, 崔承印. 利用公交刷卡数据分析北京居住关系和通勤出行[J]. 地理学报, 2012, 67 (10): 1339-1352.
LONG Ying, ZHANG Yu, CUI Chengyin. Identifying commuting pattern of Beijing using bus smart card data[J]. Acta Geographica Sinica, 2012, 67 (10): 1339-1352.
- [17] 刘耀林, 陈龙, 安子豪, 等. 基于公交刷卡数据的武汉市居住通勤特征研究[J]. 经济地理, 2019, 39 (2): 93-102.
LIU Yaolin, CHEN Long, AN Zihao, et al. Research on job-housing and commuting in Wuhan based on bus smart card data[J]. Economic Geography, 2019, 39 (2): 93-102.
- [18] 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[J]. Acta Geographica Sinica, 2016, 71 (3): 484-499.
- [19] 钟炜菁, 王德, 谢栋灿, 等. 上海市人口分布与空间活动的动态特征研究——基于手机信令数据的探索[J]. 地理研究, 2017, 36 (5): 972-984.
ZHONG Weijing, WANG De, XIE Dongcan, et al. Dynamic characteristics of Shanghai's population distribution using cell phone signaling data[J]. Geographical Research, 2017, 36(5): 972-984.
- [20] 童晓君. 基于出租车GPS数据的居民出行行为分析[D]. 长沙: 中南大学, 2012.
TONG Xiaojun. Analysis of residents, behavior based on the taxi GPS data[D]. Changsha: Central South University, 2012.
- [21] 徐敏, 黄震方, 曹芳东, 等. 基于大数据分析的城市旅游地网络结构特征及其演化模式——以新浪微博签到数据为例[J]. 地理研究, 2019, 38 (4): 937-949.
XU Min, HUANG Zhenfang, CAO Fangdong, et al. The network structure of urban tourist destination and its evolution mode based on big data analysis: taking the data of Sina weibo sign-in as an example[J]. Geographical Research, 2019, 38 (4): 937-949.
- [22] 冯章献, 张瑜, 魏冶, 等. 基于百度迁徙数据的长春市春运人口流动时空格局与动力机制[J]. 经济地理, 2019, 39 (5): 101-109.
FENG Zhangxian, ZHANG Yu, WEI Ye, et al. Spatial-temporal pattern and dynamic mechanism of population flow of Changchun City during Chunyun period based on Baidu migration data[J]. Economic Geography, 2019, 39 (5): 101-109.
- [23] 刘望保, 石恩名. 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[J]. 地理学报, 2016, 71 (10): 1667-1679.
LIU Wangbao, SHI Enming. Spatial pattern of population daily flow among cities based on ICT: a case study of "Baidu Migration"[J]. Acta Geographica Sinica, 2016, 71(10): 1667-1679.
- [24] 韦统. 基于多源数据的北部湾城市群识别与界定研究[D]. 南宁: 广西师范学院, 2015.
WEI Tong. Research on identification and definition of Beibu Gulf urban agglomeration based on multi-sourced data. [D]. Nanning: Guangxi Normal University, 2015.
- [25] 蒋小荣, 汪胜兰, 杨永春. 中国城市人口流动网络研究——基于百度LBS大数据分析[J]. 人口与发展, 2017, 23 (1): 13-23.
JIANG Xiaorong, WANG Shenglan, YANG Yongchun. Research on China's urban population mobility network based on Baidu LBS big data[J]. Population & Development, 2017, 23 (1): 13-23.