

中国市域旅游流网络结构空间分异及其效应研究

——基于携程旅行网的大数据挖掘

方叶林¹, 黄震方², 李经龙¹, 程雪兰¹, 苏雪晴¹

(1. 安徽大学商学院, 合肥 230601; 2. 南京师范大学地理科学学院, 南京 210023)

摘要: 大数据驱动的旅游流网络结构研究是旅游流深化研究的主要方向之一。利用网络爬虫技术抓取携程旅行网上中国市域尺度的旅游线路及相关数据, 在构建 O-D 矩阵的基础上, 揭示网络结构指标的空间分异及其效应。中国市域旅游流网络具有以下特征: 网络同质性较小; 不同主题类旅游流流量总体上遵从东多西少、东南多西北少的总体格局; 中国市域旅游流网络形成了五大高值集聚区, 分别为长三角城市群、泛珠三角地区、云南内部、青海—甘肃交界地区、北疆地区; 中国市域旅游流网络还具有显著的资源—经济指向性, “结构”效应总体不显著。未来需要进一步重视区域旅游网络化的发展特征, 在尊重空间规律的基础上, 发挥核心组团的辐射作用, 发挥“资源—经济”的双核驱动, 促进跨区域旅游合作。

关键词: 旅游流; 网络结构; 空间分异; 效应; 大数据

近年来, 大数据已经渗透到全球范围内社会经济的各个领域^[1], 为人文地理学研究带来新的机遇, 被称为地理学“第四范式”^[2]。带有地理空间信息的数据成为大数据的重要组成部分, 契合人文地理学“计量转向”及人本主义发展趋势^[3], 逐渐掀起了一股研究热潮。旅游作为现代社会重要的生活方式和社会经济活动, 被深深打上了互联网大数据的“烙印”^[4]。旅游流是旅游地理学研究的核心问题之一^[5], 数据获取难易程度是制约旅游流研究的基本问题^[6], 大数据自身具有的规模性、多样性、高速性、价值性等特点, 能够揭示社会经济现象的复杂性^[7], 逐渐成为旅游流研究中重要的数据来源。

国外学者对旅游流的研究较早, 相关研究主要集中在旅游流概念内涵与分析框架^[8]、旅游流流量模拟及预测^[9]、旅游流空间结构模式^[10]、旅游流的影响因素研究^[11]、旅游流流动效应^[12]等方面。旅游地不断发展以及旅游与信息技术的不断融合, 使得旅游流时空演化具有显著的网络结构特征。国外文献对旅游流网络结构的研究主要以社会网络理论及方法为基础, 进而揭示网络结构特征。如 Hwang 等^[13]用网络分析法, 研究了不同客源地区游客在美国进行多目的地游程的特征; Shih^[14]对中国台湾南投地区的自驾车旅游网络结构进行了研究, 并对旅游基础设施的建设提出了建议; Miguéns 等^[15]对全球旅游流网络结构进行了深入分析。国内学者杨兴柱等^[16]的研究较早涉及到城市旅游流网络结构, 认为城市旅游流网络结构主要包括旅游节点结构与旅游网络结构, 并构建了城市旅游流网络结构评价指标体系; 刘法建等^[17]构建了中国入境旅游流网络结构评价指标, 并利用 QAP

收稿日期: 2020-08-16; 修订日期: 2020-12-01

基金项目: 国家自然科学基金项目 (42171238, 41601142); 安徽省自然科学基金项目 (2108085MD125); 安徽省哲学社会科学规划项目 (AHSKQ2020D64)

作者简介: 方叶林 (1986-), 男, 安徽巢湖人, 博士, 副教授, 硕士生导师, 研究方向为旅游地理与区域经济。

E-mail: fangyelin2006@126.com

进行网络机理分析；马耀峰等^[18]选取不同时间节点，对中国主要城市入境旅游流网络结构时空演化进行分析；穆成林等^[19]以京福高铁为例，探讨高铁开通对黄山市旅游目的地网络空间结构的影响。近年来，国内学者一方面努力拓展不同时空尺度的旅游流网络结构对比研究，另一方面尝试进行多学科交叉研究，总结旅游流网络结构时空演化规律，均取得显著进步。从空间尺度看，涉及到城市群尺度^[20]、省域尺度^[21]、城市内部及旅游景点尺度^[22]。从研究方法上看，现有国内文献一般利用社会网络方法分析网络结构基本特征，并结合位序—规模分布^[23]、核心—边缘理论^[24]、拓扑分析^[25]、空间场强效应^[26]等方法进行研究，拓展了旅游流网络结构研究的理论与方法。

数据是科学研究工作开展的重要基础。利用大数据挖掘技术获取旅游流数据，进而分析旅游流网络结构时空分异规律，是该领域深化研究的主要方向之一，近年来取得丰富的研究成果^[27]。随着“网络范式”的转向和“流空间”研究的不断升温^[28]，与此相关的数据挖掘技术，已经逐步深入到旅游研究的各个领域。相关学者分别基于网络关注度^[29]、互联网游记数据^[30]、微博签到数据^[31]、手机信令数据^[32]、地理图片数据^[33]构建旅游流网络并进行结构特征分析，均取得一定的进展。总结以往旅游流研究文献可以发现：一是先前较多关注旅游流流量的研究，数据样本量有限，采用的方法也较为传统，相对忽视了旅游流网络结构及其空间分异规律；二是研究数据尺度相对较小，涉及到“大范围、小尺度”的研究相对不足，市域尺度的全国旅游流流量及网络结构特征研究相对匮乏；三是相对忽视了旅游流网络结构的现实效应，网络结构与现实游客之间具有怎样的关系，文献涉及较少。

“网络化”是“点轴”系统发展的结果^[34]，是区域空间结构演化的高级形式^[35]，是地理要素在空间上一种效率较高的布局模式。网络化发展是区域旅游高质量发展的基本特征之一，而旅游流是研究区域旅游网络化发展态势的重要手段。大数据驱动的区域旅游流网络结构研究，既是旅游地理学理论深化研究的重要科学问题，也是区域旅游高质量发展面临的基本现实问题。本文基于大数据爬虫技术，抓取中国市域尺度的旅游线路数据，构建旅游流网络并分析网络结构特征，进而揭示中国市域尺度旅游流网络空间分异规律，进一步提出相关对策建议。

1 研究方法与数据来源

1.1 研究方法

1.1.1 网络密度

网络密度是用来描述网络中所有节点联系的稠密程度，网络密度 D 一般用网络中的实际联系数量占理论联系数量的比例来衡量（%）； L 表示网络中实际联系数量（个）； n 为网络中的节点数量（个）；则网络密度 D 可以表示为^[36]：

$$D = 2L/[n(n-1)] \quad (1)$$

1.1.2 节点度

度是复杂网络最基本的统计指标，节点度表示与该节点所有连接边 T_{ij} 的总数^[37]。对于有向网而言，度分为出度与入度，出度表示节点对网络资源的控制能力，入度表示节点对网络资源的吸引能力，计算公式如下^[37]：

$$O_i = \sum T_{ij} \quad (2)$$

$$I_i = \sum_j T_{ji} \quad (3)$$

$$D_i = O_i + I_i \quad (4)$$

式中： O_i 表示节点*i*的出度； I_i 表示节点*i*的入度； D_i 表示节点*i*的度数。

1.2 数据来源

数据主要来源于两方面：一是携程旅行网 (<https://www.ctrip.com>) 上的旅游线路挖掘数据。携程旅行网是主流旅游网站，网站上包含了大量的旅游线路及相关信息。聘请专业的大数据挖掘团队，对城市与“周边”的旅游线路数据进行挖掘、转换、整理、建网，建立出发地—目的地 (Origin-destination, O-D) 网络矩阵；抓取时间截至2020年6月30日，一共抓取了110823条原始旅游线路数据，进一步建立相关描述性统计数据库，作为旅游流网络结构分析的原始数据。二是矢量数据，主要来源于中国科学院资源环境科学与数据中心 (<http://www.resdc.cn>)。其中市域 (包括地级市、自治州、地区及省会城市) 个数根据数据收集情况进行了相关微调。考虑到部分行政区划调整、部分地区与其他地区基本没有旅游线路数据，选择其中299个地级市构建网络矩阵。本文的原始数据为299×299的矩阵数据，以及相关人数、主题等数据。

2 结果分析

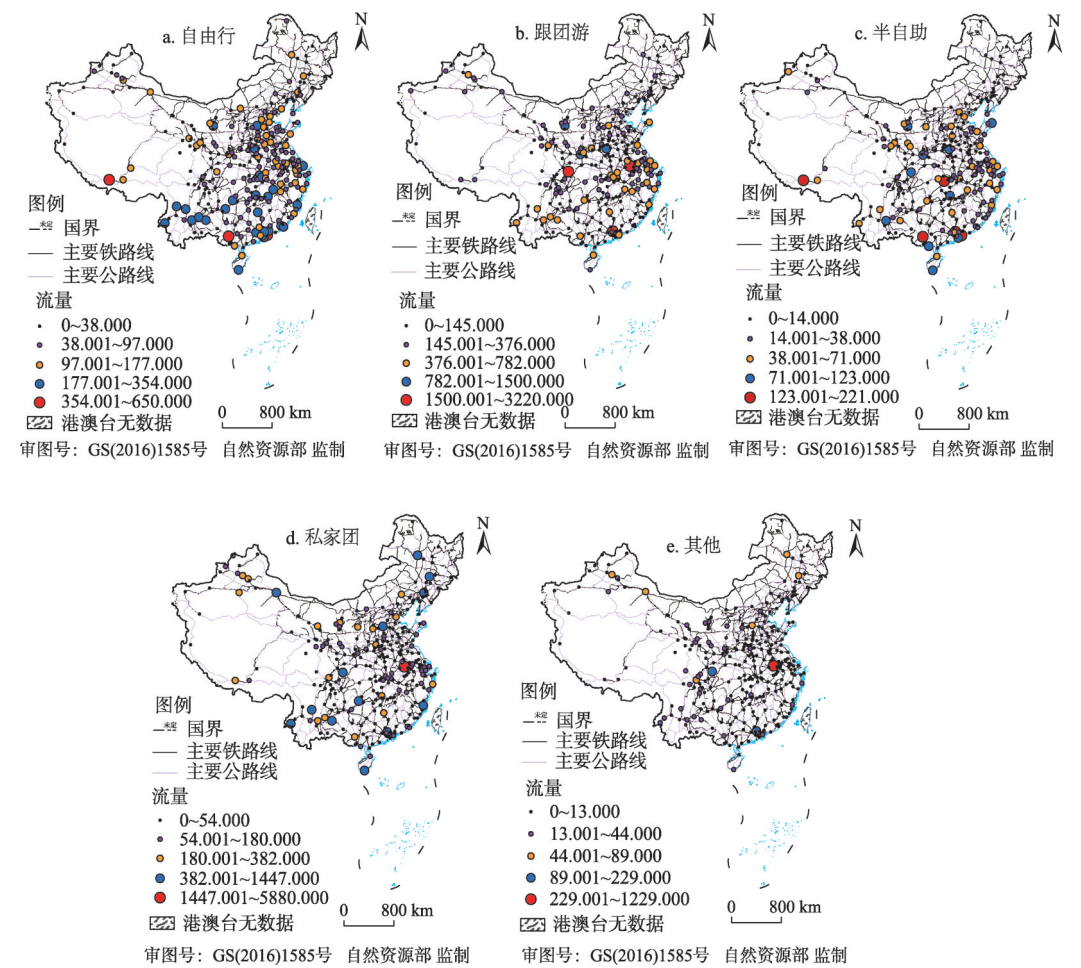
2.1 网络结构空间分异

2.1.1 样本特征

根据携程旅行网上抓取到的线路主题，从中分解出旅游线路类型，进而提炼不同类型旅游流流量的空间分布规律。携程旅行网将旅游线路主题分为自由行、跟团游、半自助、私家团以及其他主题等，不同主题类旅游空间分布如图1所示。从图1可以看出：(1) 不同主题类旅游流流量总体上遵从东多西少、东南多西北少的总体格局，尤其集中在华北及华南片区，东部的京津冀、长三角与珠三角基本上连成一个整体。(2) 不同主题类旅游流流量高值分布具有“沿线、围城”的特征，旅游流流量的高值区主要集中在铁路、公路沿线，以及省会城市周边，交通与配套服务仍是当前主题类旅游发展十分重要的保障因素。(3) 从省域分布看，不同主题类旅游流流量高值区一方面分布在经济发达的沿海省份，如山东、江苏、浙江、福建、广东；另一方面分布在旅游资源特色显著的中西部省 (自治区)，如云南、广西、新疆等地。(4) 从不同主题旅游流流量的密度分布看，自由行逐渐成为国内旅游的主流方式，跟团游仍占据相对较多的份额，其他类型的旅游主题，如蜜月婚拍、野外探险等主题旅游也逐渐兴起。此外，携程旅行网根据现实旅游市场发展特点，专门推出了私家团，私家团具有“私人订制”的性质，在主流旅游线路中占据了越来越重要的比例。

2.1.2 网络结构计算结果

利用节点出度、入度与度公式计算中国城市旅游流网络结构特征，其中，排名前60位的城市分别如表1所示。从出度的计算结果来看，排名前10位的地区分别为：成都 (3806)、资阳 (3464)、德阳 (2787)、南京 (2732)、楚雄 (2502)、玉溪 (2482)、曲靖 (2467)、杭州 (2410)、昆明 (1826)、上海 (1591)，基本上位于长三角城市群与云南境内。这些地区在网络中处于关键节点位置，对网络资源的控制能力较强。从入度的计算结果来看，排名前10位的分别为：上海 (7484)、杭州 (7430)、苏州 (7343)、乐山



注：本图基于自然资源部标准地图服务系统下载的标准地图制作，底图无修改，下同。

图1 不同主题旅游流流量的空间分布
Fig. 1 Spatial distribution of tourist flow of different themes

(5451)、成都 (5131)、北京 (4792)、大理 (4092)、丽江 (4000)、南京 (2919)、乌鲁木齐 (2758)。这些城市大部分具有深厚的历史文化底蕴，旅游资源丰富，在网络中具有较强的吸引能力。节点度的结果为节点入度与出度的和，从计算结果看，排名前10位的城市分别为：杭州 (9840)、上海 (9075)、成都 (8937)、苏州 (8398)、乐山 (5721)、南京 (5651)、北京 (5369)、丽江 (4624)、昆明 (4227)、大理 (4192)。网络节点入度总体上大于出度，表明节点的吸引能力大于控制能力。

表1的计算结果还反映了“大范围、小尺度”旅游流网络结构存在的部分规律——网络节点的重要性受区域经济与资源禀赋交互驱动。节点在网络中能否具备优势“位置”，一方面受宏观经济因素驱动，宏观经济因素为区域旅游发展提供设施保障，促进节点在网络中的地位逐渐凸显，对出游动机具有一定的“推力”；另一方面受资源禀赋驱动，优越的旅游资源吸引了大量的外来旅游者，使得节点在网络中具有较为重要的作用，对游客出游动机具有一定的“拉力”。

旅游流流量具有规模—位序特征，已被大部分学者的研究所证实^[23]。为了进一步验

表1 各地区网络结构指标计算结果（排名前60位）

Table 1 Calculation results of network structure indicators in various regions (Top 60)

地区	出度	地区	出度	地区	入度	地区	入度	地区	度	地区	度
成都	3806	台州	849	上海	7484	吐鲁番	1047	杭州	9840	武汉	1763
资阳	3464	上饶	847	杭州	7430	武汉	831	上海	9075	长沙	1734
德阳	2787	镇江	834	苏州	7343	扬州	823	成都	8937	常州	1604
南京	2732	合作	812	乐山	5451	景德镇	757	苏州	8398	青岛	1516
楚雄	2502	嘉兴	784	成都	5131	珠海	735	乐山	5721	厦门	1466
玉溪	2482	黄山	784	北京	4792	威海	720	南京	5651	广州	1431
曲靖	2467	贵阳	770	大理	4092	北海	521	北京	5369	天津	1358
杭州	2410	益阳	763	丽江	4000	烟台	490	丽江	4624	扬州	1330
昆明	1826	运城	759	南京	2919	九江	484	昆明	4227	嘉峪关	1309
上海	1591	济南	756	乌鲁木齐	2758	太原	479	大理	4192	合肥	1260
无锡	1554	湖州	739	无锡	2556	宜昌	463	无锡	4110	宁波	1245
常州	1489	丽水	702	昆明	2401	广州	459	资阳	3464	开封	1227
合肥	1255	南通	680	桂林	2318	济南	415	乌鲁木齐	3258	深圳	1218
宁波	1181	安庆	672	黄山	2279	南宁	397	黄山	3063	昌吉	1213
昌吉	1175	呼和浩特	671	银川	2050	恩施	396	德阳	2787	徐州	1177
徐州	1174	芜湖	662	重庆	2034	承德	395	桂林	2603	济南	1171
温州	1162	滁州	652	张掖	2028	延安	373	玉溪	2549	温州	1164
白银	1127	铜陵	651	西安	1936	南昌	359	楚雄	2516	白银	1142
盐城	1063	马鞍山	646	洛阳	1755	深圳	347	曲靖	2479	吴忠	1101
苏州	1055	兰州	641	天门	1727	德令哈	331	重庆	2369	景德镇	1093
蚌埠	1047	绍兴	637	兰州	1636	大连	304	西安	2353	盐城	1066
吴忠	1026	都匀	634	郑州	1509	伊宁	246	兰州	2277	蚌埠	1047
广州	972	丽江	624	西宁	1482	喀什	243	银川	2120	奎屯	1013
定西	951	包头	619	嘉峪关	1307	三亚	230	张掖	2051	临夏	977
奎屯	946	咸阳	588	贵阳	1226	枣庄	217	西宁	2004	镇江	927
武汉	932	北京	577	长沙	1165	惠州	195	贵阳	1996	烟台	911
吐鲁番	932	宜兴	573	厦门	1157	中山	193	吐鲁番	1979	南宁	892
临夏	924	石家庄	573	青岛	1109	固原县	159	郑州	1966	合作	878
深圳	871	长沙	569	开封	1079	安阳	155	洛阳	1923	上饶	860
金华	849	铜川	566	天津	1069	绍兴	153	天门	1806	运城	857

证旅游流网络结构是否也具有类似的规模一位序特征，分别以中国城市旅游流网络的出度、入度、度为纵轴，以各节点以上指标的排名为横轴，绘制散点图，结果如图2所示。利用对数函数对散点图进行拟合，发现曲线 R^2 分别为0.703、0.914、0.821，旅游流网络结构具有较强的位序—规模特征。网络结构数值较大的节点，所占比例较少；而网络结构数值较小的节点，所占比例较多，网络结构指标具有显著的“长尾特征”，网络同质性较差。

2.1.3 网络结构的空间分布

为了进一步分析旅游流网络结构的空间特征，将计算结果用GIS空间插值（克里格插值法）进行可视化表达，结果如图3所示。从空间插值的结果可以发现：出度与入度

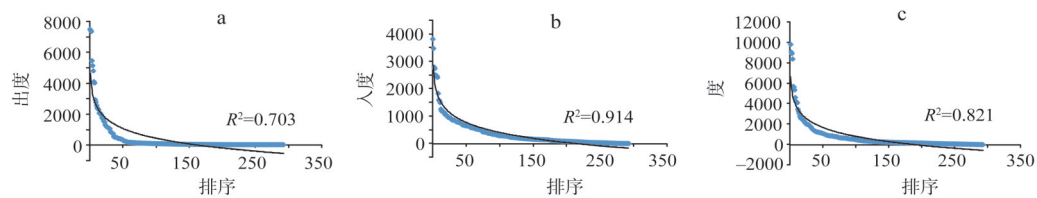


图2 网络结构指标的幂指特征

Fig. 2 The power index characteristics of network structure index

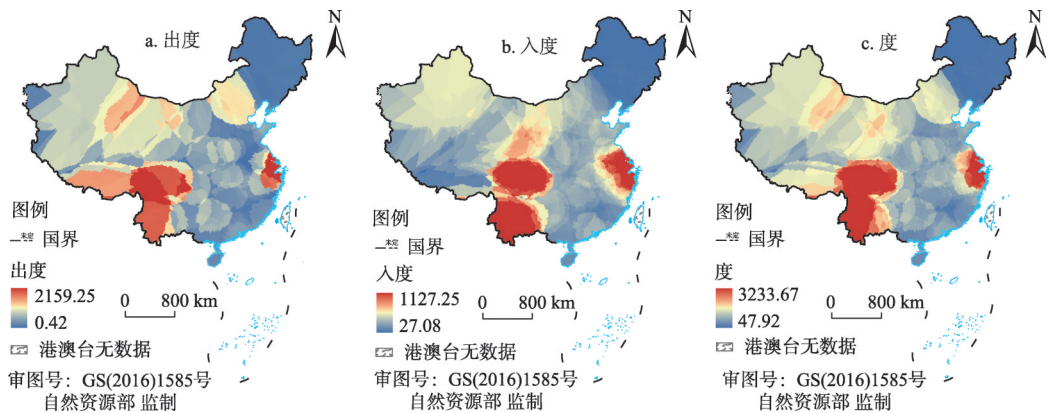


图3 网络结构指标的空间分布

Fig. 3 Spatial distribution pattern of network structure index

的空间分布具有一定的相似性，高值区主要集聚在长三角城市群、成渝经济圈、云南省及其周边，表明游客目的地与客源地具有一定的重叠性。长三角城市群与成渝经济圈是我国重要的经济增长中心，也是中国的主要客源地与目的地，位于中国旅游流网络的核心位置。云南省是我国旅游资源十分丰富的地区之一，也是重要的旅游目的地，同时，大部分西南地区的旅游线路都涉及到云南境内的城市，因而其点出度数值相对较大。此外，中国市域旅游流网络在东北地区形成一个低值的“洼地”，东三省及内蒙古东北部城市节点在网络中的地位相对较低。

需要指出的是，作为中国旅游经济增长核心之一的珠三角地区在网络节点中的位置并不凸显，其主要原因包括：一是和携程旅行网推出的旅游线路有关，携程旅行网主要推出了自由行、跟团游、半自助、私家团等主题旅游线路，从挖掘的线路主题看，这些旅游线路主打生态环境，与城市旅游、商务旅游等旅游主题关系不强。二是旅游线路仅仅以国内旅游为主，没有考虑出境与入境旅游，而这些旅游形式在珠三角旅游市场中占据了重要份额。所以，与长三角及中西部旅游大省相比，珠三角地区在网络中虽然具有优势，但有时优势地位不显著。

2.1.4 网络结构的组团特征

利用密度计算公式对旅游流网络数据进行密度分析，网络总体密度为1.189，标准差为21.666。网络密度相对较小，标准差相对较大，表明网络同质性较小，网络中存在较多的“小团体”，旅游流网络具有一定的层级特征。为了进一步分析中国城市旅游流网络结构的层级特征，以流量为断点值绘制旅游流网络，按照自然断裂点分类方法（Jenks）

将旅游流分为四类, 流量断点值分别为 $GE=8$ 、 $GE=39$ 、 $GE=91$ 、 $GE=171$, 结果如图4所示。随着断点值的增加, 网络密度逐渐下降, 当断点值取值为171时, 大致可以发现中国城市旅游流网络结构高值区分布的基本规律。旅游流网络结构高值区具有显著的组团特征, 总体上分为五大组团。

组团一: 一强多超型网络。主要以长三角城市群为核心向北及向西辐射, 向北延伸至以大连为中心的辽东半岛次核心、以青岛—烟台—威海一线为核心的山东半岛次核心, 以及以北京—天津为核心的京津冀次核心; 向西以郑州为核心的河南次核心, 以及以武汉为中心的次核心。该组团网络结构特征显著, 在网络中的核心地位优势明显, 是中国市域旅游流网络最为密集、旅游活动最为活跃的地区。组团二: 多核心串联型网络。大部分以华南省份的省会为核心串联, 主要包括贵阳—南宁—贺州—广州—厦门一线。省会城市及关键节点在网络中的地位十分重要, 受地形与行政区划影响, 该组团内分布着较多小团体网络。组团三: 近四边形封闭型网络。主要分布在云南省内, 核心节点包括丽江、大理、玉溪、昆明等地, 旅游流网络结构形成相对封闭的四边形, 向周边地区的辐射力度有限, 一方面受云南高原山地边境省份的影响, 另一方面, 这些地区也是外界去云南旅游的主要地区, 核心节点内部联系较多, 外部联系较弱。组团四: 多节点扇形网络。主要位于青海、甘肃交界处, 以西宁—兰州为中心向周边辐射。扇形开口

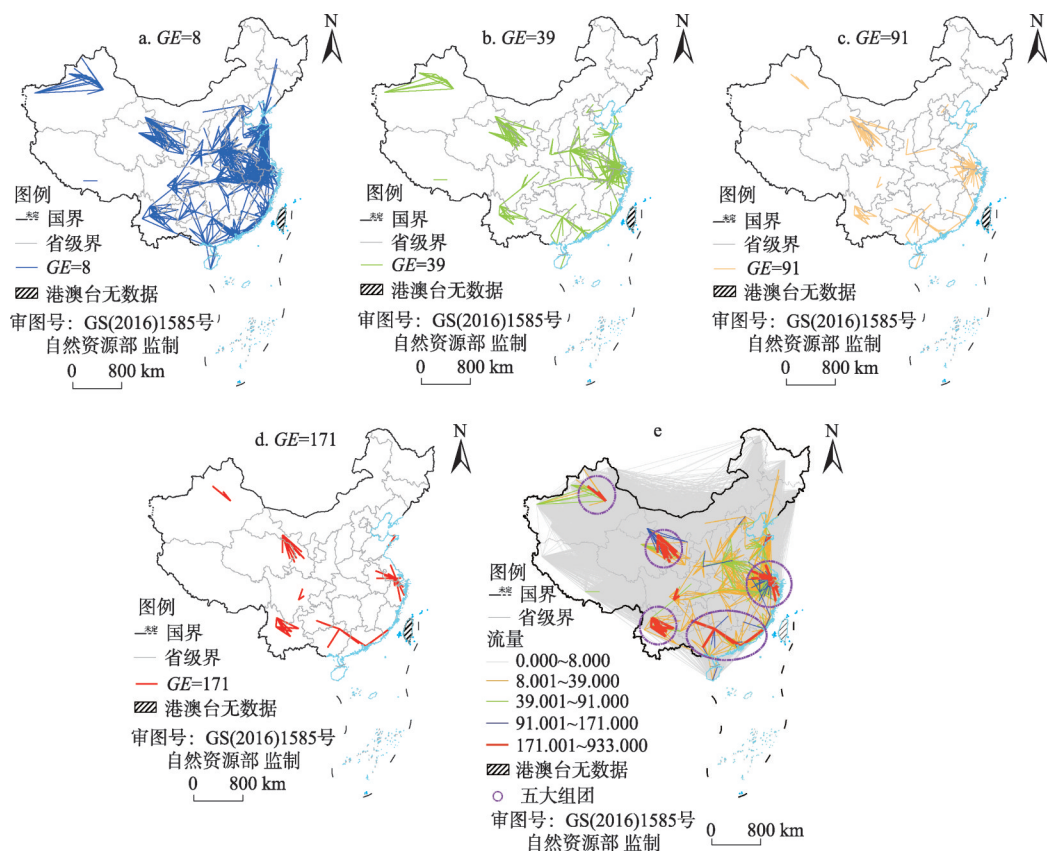


图4 不同断点值约束下城市旅游流网络空间分布

Fig. 4 Spatial distribution of urban tourism flow network under different break-point values

方向主要面向东中部,也表明了东中部地区的客源市场对该组团旅游发展具有十分重要的作用,该组团的网络密度相对较低。组团五:节点串联型网络。主要位于新疆北部,包括库尔勒、乌鲁木齐、吉昌、石河子等地区,同时在边境也有零星节点。该组团旅游流节点串联组合明显,组团密度较低。这些地区主要位于北疆,集中了新疆大部分优势旅游资源。

2.2 网络结构效应

网络结构与现实游客有何关系,旅游流网络结构有何现实效应?根据抓取的旅游线路数据,从中提取节点的游客人数进行克里格空间插值,结果见图5a。通过对比图5a与图4可以发现,网络结构指标与游客数量存在一定的错位现象。利用网络各节点入度、出度、度与游客数量的散点图,发现网络结构指标与游客数量总体上具有一定的正相关性,不过显著性不强(R^2 分别为0.099、0.053、0.094),这种关系在局部地区是否如此,需进一步验证。利用网络节点度与游客数的平均值,将游客人数与节点度之间的关系划分为四种类型,分别为:高高型、低高型、低低型、高低型,划分结果如图5b所示。

(1) 高高型:城市游客人数与节点度排名均超过均值。位于该区间的地区主要有:珠海、深圳、丽江、厦门、德阳、宜昌、益阳、常州、景德镇、黄山、上饶、无锡、苏州、湖州、宁波、温州、运城、定西、广州、南宁、昆明、成都、长沙、合肥、武汉、上海、杭州、西安、郑州、西宁、济南、天津、乌鲁木齐、天门等地区。位于该区间城市个数占到总数的11.41%,表明节点在网络中处于较为有利的位置,与其他节点的交流互动较为频繁,结构优势得到发挥,同时也吸引了较多的游客,网络结构具有一定的正向效应。

(2) 低高型:城市游客人数较少,但节点度却超过了均值。位于该区间的地区主要有:大理、楚雄、玉溪、桂林、乐山、金华、丽水、嘉兴、绍兴、临夏、洛阳、开封、徐州、蚌埠、扬州、镇江、盐城、嘉峪关、张掖、白银、青岛、威海、烟台、吐鲁番、

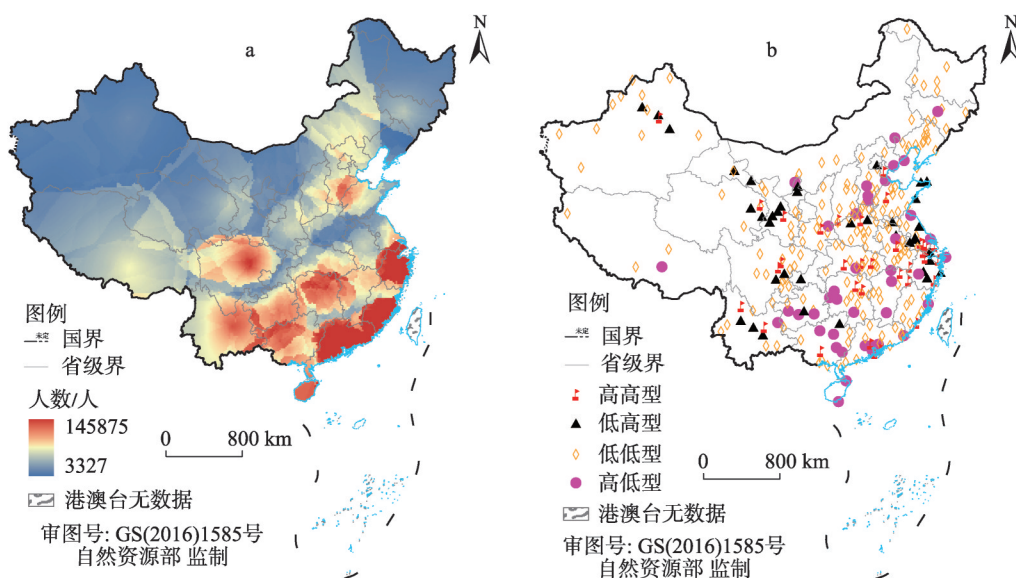


图5 游客人数的空间插值及其与网络结构的关系

Fig. 5 Spatial interpolation of the number of tourists and its relationship with network structure

奎屯、昌吉、合作、同仁、共和、吴忠、台州、资阳、贵阳、重庆、南京、银川、兰州、北京等地区。位于该区间城市个数占到总数的12.75%，表明这些地区虽然在网络中具有较好的位置，但“结构优势”并没有得到充分发挥。

(3) 低低型：城市游客人数与节点度排名均低于均值。位于该区间的城市在网络中位置较差，与其他节点的互动交流较少，同时游客人数不多。位于该区间的地区主要有：个旧、百色、梧州、清远、钦州、北海、茂名、湛江、阳江、河源、惠州、汕头、汕尾、攀枝花、保山、潞西、昭通、遵义、凯里、兴义、娄底、邵阳、湘潭市、萍乡、衡阳、郴州、贺州、韶关、新余、宜春、吉安、三明、南平、宁德、龙岩、莆田、梅州、漳州、泉州、日喀则、绵阳、巴中、遂宁、南充、雅安、自贡、内江、宜宾、泸州、荆门、恩施、孝感、常德、岳阳、麻城、六安、塔河、济源、宜兴、仙桃、潜江、防城港等地区。位于该区间的城市最多，占到总数的61.74%，反映了当前中国城市旅游流网络中大部分节点的位置没有达到最优。

(4) 高低型：城市游客人数较多，但节点度排名却低于均值，城市游客人数与城市在网络中的位置存在一定“脱钩”关系。位于该区间的地区主要有：三亚、贵港、肇庆、东莞、佛山、玉林、江门、中山、揭阳、六盘水、安顺、都匀、曲靖、铜仁、怀化、永州、河池、柳州、赣州、荆州、吉首、九江、衢州、舟山、晋城、渭南、连云港、淮南、南通、唐山、秦皇岛、沧州、邢台、邯郸、赤峰、阿拉善、海口、福州、拉萨、南昌、石家庄、长春等地区。位于该区间的城市占到总数的14.09%。

通过以上分析可以发现，以上四种类型交错分布在全国各地。低低型与低高型的地区占到总数的75.83%，表明大部分节点在网络中位置处于不利地位，尤其低低型城市节点占到总数的61.74%，反映出中国市域旅游流网络结构不平衡现象较为严重，大部分节点在网络中的结构与游客数处于不利地位。不同于中小尺度的旅游流网络，中国市域旅游流网络结构效应不显著。

3 结论与讨论

3.1 结论

旅游流网络结构及其空间分异规律，一直是旅游流研究的热点问题。利用大数据挖掘技术抓取携程旅行网上旅游线路相关数据，构建中国市域旅游流网络，分析网络结构的特征与网络结构效应。研究结论主要有：

(1) 中国市域旅游流网络同质性较小。从流量角度看，不同主题类旅游流流量总体上遵从东多西少、东南多西北少的格局，尤其集中在华北及华南片区；从网络结构视角看，节点度的高值区主要集中分布在长三角城市群、成渝经济圈、云南省及其周边。

(2) 中国市域旅游流网络组团特征明显。选择不同的流量断点值进行细分，中国市域旅游流网络形成了五大高值集聚区，分别为长三角城市群（一强多超型网络）、泛珠三角地区（多核心串联型网络）、云南内部（近四边形封闭型网络）、青海—甘肃交界处（多节点扇形网络）及北疆地区（节点串联型网络）。

(3) 中国市域旅游流网络具有显著的资源—经济指向型。一方面，不同主题类旅游流流量高值的分布具有“沿线、围城”的特征；另一方面，节点出度、入度的计算结果均表明：节点在网络中能否具备优势“位置”，受旅游资源禀赋与区域经济“推—拉力”

的共同作用。

(4) 中国市域旅游流“结构”效应不显著。一方面,网络结构指标具有显著的幂律特征,使得网络结构优势集中在少数节点;另一方面,利用网络节点度与游客数的平均值,将游客人数与节点度之间的关系划分为四种类型,分别为:高高型、低高型、低低型、高低型,低低型与低高型的地区占到总数的四分之三以上,大部分节点在网络中的位置处于不利地位。

虽然研究的样本量较大,但仅仅抓取携程旅行网上旅游线路及相关数据,其他主流旅游网站是否如此,仍需进一步加强对比研究。利用大数据抓取旅游流信息,进行“大尺度、小范围”的旅游流网络结构及其效应研究,是未来旅游流研究的趋势之一。

3.2 相关启示

旅游流网络是促进内循环的重要动力,网络式空间结构是高质量发展方式对生产要素在空间布局上的基本要求之一。基于上述结论,本文主要启示有:(1) 尊重旅游流网络的空间分布规律,发挥核心组团的辐射带动作用。继续发挥长三角城市群、泛珠三角地区、云南地区、青海—甘肃交界地区、北疆地区等网络组团对整个中国地级市旅游流网络的带动作用,进而促进整个网络的均质化发展。(2) 进一步加强跨区域旅游合作,针对网络同质性较差、网络中存在的诸多“小团体”现象,大尺度的跨区域旅游合作尤为重要,努力促进东中西各城市节点“互为目的地,互为客源地”。(3) 进一步强化“经济—资源”的双核驱动作用,提炼资源特色,大力发展全域旅游,夯实地区旅游发展物质基础。(4) 重视区域旅游的“网络化”发展态势。“十四五”时期旅游业高质量发展必须重视区域旅游“网络化”发展趋势,按照网络结构特征合理布局旅游要素,促进“十四五”时期区域旅游转型升级与提质增效。

参考文献(References):

- [1] LYNCH C. Big data: How do your data grow. *Nature*, 2008, 455(7209): 28-29.
- [2] 程昌秀,史培军,宋长青,等. 地理大数据为地理复杂性研究提供新机遇. *地理学报*, 2018, 73(8): 1397-1406. [CHENG X C, SHI P J, SONG C Q, et al. Geographic big-data: A new opportunity for geography complexity study. *Acta Geographic Sinica*, 2018, 73(8): 1397-1406.]
- [3] 甄峰,王波. “大数据”热潮下人文地理学研究的再思考. *地理研究*, 2015, 34(5): 803-811. [ZHEN F, WANG B. Re-thinking human geography in the age of big data. *Geographical Research*, 2015, 34(5): 803-811.]
- [4] PAVLOVICH P. The evolution and transformation of a tourism destination network: The Waitomo Caves, New Zealand. *Tourism Management*, 2003, 24(2): 203-216.
- [5] 钟士恩,张捷,周强,等. 农村居民国内旅游流空间集中性. *地理研究*, 2009, 28(6): 1562-1571. [ZHONG S E, ZHANG J, ZHOU Q, et al. Spatial concentration of domestic tourist flow of Chinese rural residents. *Geographical Research*, 2009, 28(6): 1562-1571.]
- [6] AHAS R, AASA A, ROOSE A, et al. Evaluating passive mobile positioning data for tourism surveys: An Estonian case study. *Tourism Management*, 2008, 29(3): 469-486.
- [7] 杨振山,龙瀛, NICOLAS D. 大数据对人文—经济地理学研究的促进与局限. *地理科学进展*, 2015, 34(4): 410-417. [YANG Z S, LONG Y, NICOLAS D. Opportunities and limitations of big data applications to human and economic geography. *Progress in Geography*, 2015, 34(4): 410-417.]
- [8] WILLIAMS A V, ZELINSKY W. On some patterns in international tourist flows. *Economic Geography*, 1970, 46(4): 549-567.
- [9] KIM J H, MOOSA I A. Forecasting international tourist flows to Australia: A comparison between direct and indirect methods. *Tourism Management*, 2005, 26(1): 69-78.

- [10] COSHALL J. Spectral analysis of international tourism flows. *Annals of Tourism Research*, 2000, 27(3): 577-589.
- [11] KIM S S, AGRUSA J, LEE H, et al. Effects of Korean television dramas on the flow of Japanese tourists. *Tourism Management*, 2007, 28(5): 1340-1353.
- [12] PARTE-ESTEBAN L, ALBERCA-OLIVER P. Tourist flow and earnings benchmarks: Spanish hotel industry. *Journal of Hospitality & Tourism Research*, 2016, 40(1): 58-84.
- [13] HWANG Y H, GRETZEL U, FESENMAIER D R. Multicity trip patterns: Tourists to the United States. *Annals of Tourism Research*, 2006, 33(4): 1057-1078.
- [14] SHIH H Y. Network characteristics of drive tourism destinations: An application of network analysis in tourism. *Tourism Management*, 2006, 27(5): 1029-1039.
- [15] MIGUÉNS J I L, MENDES J F F. Travel and tourism: Into a complex network. *Physica a Statistical Mechanics & its Applications*, 2008, 387(12): 2963-2971.
- [16] 杨兴柱, 顾朝林, 王群. 南京市旅游流网络结构构建. *地理学报*, 2007, 62(6): 609-620. [YANG X Z, GU C L, WANG Q. Urban tourism flow network structure construction in Nanjing. *Acta Geographica Sinica*, 2007, 62(6): 609-620.]
- [17] 刘法建, 张捷, 陈冬冬. 中国入境旅游流网络结构特征及动因研究. *地理学报*, 2010, 65(8): 1013-1024. [LIU F J, ZHANG J, CHEN D D. The characteristics and dynamical factors of Chinese inbound tourist flow network. *Acta Geographica Sinica*, 2010, 65(8): 1013-1024.]
- [18] 马耀峰, 林志慧, 刘宪锋, 等. 中国主要城市入境旅游网络结构演变分析. *地理科学*, 2014, 34(1): 25-31. [MA Y F, LIN Z H, LIU X F, et al. The evolution of network structure of inbound tourist in major cities of China. *Scientia Geographica Sinica*, 2014, 34(1): 25-31.]
- [19] 穆成林, 陆林. 京福高铁对旅游目的地区域空间结构的影响: 以黄山市为例. *自然资源学报*, 2016, 31(12): 2122-2136. [MU C L, LU L. The impact of Beijing - Fuzhou High-speed Rail on the spatial structure of tourist destination districts: A case study of Huangshan city. *Journal of Natural Resources*, 2016, 31(12): 2122-2136.]
- [20] 刘大均. 长江中游城市群旅游流空间格局及发展模式. *经济地理*, 2018, 38(5): 217-223. [LIU D J. Spatial pattern and development model of tourist flow in urban agglomeration in the Middle Reaches of the Yangtze River. *Economic Geography*, 2018, 38(5): 217-223.]
- [21] 戢晓峰, 戈艺澄, 陈方. 基于公路交通流大数据的节假日旅游流时空分异特征: 以云南省2017年7个节假日为例. *旅游学刊*, 2019, 34(6): 37-47. [JI X F, GE Y C, CHEN F. Spatial and temporal differentiation characteristics of holiday tourism flow based on highway traffic flow big data: A case study of 7 holidays in 2017 in Yunnan province. *Tourism Tribune*, 2019, 34(6): 37-47.]
- [22] 李磊, 陆林, 穆成林, 等. 高铁网络化时代典型旅游城市旅游流空间结构演化: 以黄山市为例. *经济地理*, 2019, 39(5): 207-216. [LI L, LU L, MU C L, et al. Spatial structure evolution of tourist flow in typical tourist cities in the period of high-speed rail networking: A case study of Huangshan city. *Economic Geography*, 2019, 39(5): 207-216.]
- [23] 杨国良, 张捷, 艾南山, 等. 旅游流齐夫结构及空间差异化特征: 以四川省为例. *地理学报*, 2006, 61(12): 1281-1289. [YANG G L, ZHANG J, AI N S, et al. Zipf structure and difference degree of tourist flow size system: A case study of Sichuan province. *Acta Geographica Sinica*, 2006, 61(12): 1281-1289.]
- [24] 王金莹, 吴晋峰, 唐澜, 等. 基于SNA的中国入境欧洲旅游流网络性质和结构特征研究. *人文地理*, 2013, 28(6): 147-153. [WANG J Y, WU J F, TANG L, et al. Research on the properties and structural characteristics of inbound European tourist flows network in China based on SNA. *Human Geography*, 2013, 28(6): 147-153.]
- [25] 崔大树, 孙杨. 基于分形维数的湖州旅游景区系统空间结构优化研究. *地理科学*, 2011, 31(3): 337-343. [CUI D S, SUN Y. A fractal-dimension-based study on optimization of spatial structure of Huzhou tourism scenic system. *Scientia Geographica Sinica*, 2011, 31(3): 337-343.]
- [26] 李创新, 马耀峰, 张颖, 等. 时空二元视角的入境旅游流集散空间场效应与地域结构: 以丝路东段典型区为例. *地理科学*, 2012, 32(2): 176-185. [LI C X, MA Y F, ZHANG Y, et al. The spatial field effect and regional structure of concentration and diffusion of inbound tourism flows on spatial and temporal scale: Case of typical district of the eastern part of the Silk Road. *Scientia Geographica Sinica*, 2012, 32(2): 176-185.]
- [27] DAMING L, LIANBING D, ZHIMING C. Statistical analysis of tourist flow in tourist spots based on big data platform and DA-HKRV M Algorithms. *Personal and Ubiquitous Computing*, 2020, 24(1): 87-101.

- [28] ALBERT N. Centralized decentralization of tourism development: A network perspective. *Annals of Tourism Research*, 2012, 40(5): 235-259.
- [29] 高楠, 张新成, 王琳艳. 中国红色旅游网络关注度时空特征及影响因素. *自然资源学报*, 2020, 35(5): 1068-1089. [GAO N, ZHANG X C, WANG L Y. Spatio-temporal characteristics and influencing factors of Chinese red tourism network attention. *Journal of Natural Resources*, 2020, 35(5): 1068-1089.]
- [30] 罗秋菊, 梁思贤. 基于数字足迹的自驾车旅游客流时空特征研究: 以云南省为例. *旅游学刊*, 2016, 31(12): 41-50. [LUO Q J, LIANG S X. Temporal and spatial characteristics of self-driving tourist flows based on tourism digital footprints: A case study in Yunnan province. *Tourism Tribune*, 2016, 31(12): 41-50.]
- [31] 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. *地理研究*, 2019, 38(4): 937-949. [XU M, HUANG Z F, CAO F D, 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. *Geographical Research*, 2019, 38(4): 937-949.]
- [32] 顾秋实, 张海平, 陈旻, 等. 基于手机信令数据的南京市旅游客源地网络层级结构及区域分异研究. *地理科学*, 2019, 39(11): 1739-1748. [GU Q S, ZHANG H P, CHEN W, et al. Hierarchical network structures and regional differentiations of tourist source destinations of Nanjing based on cellular signaling data. *Scientia Geographica Sinica*, 2019, 39(11): 1739-1748.]
- [33] 秦静, 李郎平, 唐鸣镝, 等. 基于地理标记照片的北京市入境旅游流空间特征. *地理学报*, 2018, 73(8): 1556-1570. [QIN J, LI L P, TANG M D, et al. Exploring the spatial characteristics of Beijing inbound tourist flow based on geo-tagged photos. *Acta Geographica Sinica*, 2018, 73(8): 1556-1570.]
- [34] 李小建, 李国平, 曾刚, 等. *经济地理学*(第二版). 北京: 高等教育出版社, 2006: 186-187. [LI X J, LI G P, ZENG G, et al. *Economic Geography* (Edition 2nd). Beijing: Higher Education Press, 2006: 186-187.]
- [35] 李国平, 王志宝. 中国区域空间结构演化态势研究. *北京大学学报: 哲学社会科学版*, 2013, 50(3): 148-157. [LI G P, WANG Z B. A study on the evolutionary trend of the regional spatial structure in China. *Journal of Peking University: Philosophy and Social Sciences*, 2013, 50(3): 148-157.]
- [36] 马丽, 张博. 中国省际电力流动空间格局及其演变特征. *自然资源学报*, 2019, 34(2): 348-358. [MA L, ZHANG B. The spatial distribution and evolution of interprovincial electricity flow in China. *Journal of Natural Resources*, 2019, 34(2): 348-358.]
- [37] 盛科荣, 张红霞, 侣丹丹. 中国城市网络中心性的空间格局及影响因素. *地理科学*, 2018, 38(8): 1256-1265. [SHENG K R, ZHANG H X, LI D D. The spatial pattern and influential factors of urban network centrality in China. *Scientia Geographica Sinica*, 2018, 38(8): 1256-1265.]

Research on the spatial differentiation and effects of network structure in tourism flow in Chinese cities: Big data mining based on Ctrip

FANG Ye-lin¹, HUANG Zhen-fang², LI Jing-long¹, CHENG Xue-lan¹, SU Xue-qing¹

(1. School of Business, Anhui University, Hefei 230601, China; 2. College of Geographic Science, Nanjing Normal University, Nanjing 210023, China)

Abstract: The research on the structure of tourism flow network driven by big data is one of the main directions of deepening research. By using web crawler technology to capture travel routes and related data on Ctrip, and then constructing a 299×299 tourism flow of O-D matrix in Chinese cities, this article reveals the spatial differentiation and effects of network structure.

The results show that: (1) The tourism flow network homogeneity is small in Chinese cities. Tourism flow of different themes follows the overall pattern of "more in the east and less in the west, and more in the southeast and less in the northwest". The high-value areas of the network degree are mainly concentrated in the Yangtze River Delta Urban Agglomeration, the Chengdu-Chongqing Economic Circle, as well as Yunnan province and its surrounding areas. A low-value "depression" is formed in Northeast China. (2) The tourism flow network has obvious group characteristics in Chinese cities. It has formed five high-value agglomeration areas, namely: the Yangtze River Delta Urban Agglomeration (one strong and multi-super network), the Pan-Pearl River Delta Region (multi-core series network), and Yunnan (closed quadrilateral network), Qinghai-Gansu Junction Area (multi-node and fan-shaped network), and northern Xinjiang (node-series network). (3) The tourism flow network has significant resource-economic orientation in Chinese cities. The distribution of high value of tourism flow under different themes has the characteristics of "along the line and besieged city". Whether a node can have an advantageous "position" in the network is affected by the "push-pull" force of regional economy and tourism resources. (4) Generally, the "structure" effect is not obvious in the tourism flow network of Chinese cities. On the one hand, the network structure index has significant power-law characteristics, making the network structure advantage embodied in a few nodes; on the other hand, according to the average value of the network node degree and the number of tourists, the relationship between them can be identified into 4 types: high-high type, low-high type, low-low type, and high-low type. Most of the nodes are at a disadvantage position in the network. The high-quality development of tourism in the period of "14th Five-year Plan (2021-2025)" must attach importance to the development trend of networking characteristics of urban tourism, and reasonably arrange tourism elements according to the networking characteristics. For one thing, we should pay more attention to the development law of tourism flow network, promote cross-regional tourism cooperation; for the other, we should give full play to the radiation role of the core group and the dual-core drive of "resource-economy" of Chinese urban tourism.

Keywords: tourism flow; network structure; spatial differentiation; effect; big data