

姚建盛,刘艳玲,吴忠军.基于数字足迹的桂林游客时空行为特征[J].湖南科技大学学报(自然科学版),2022,37(4):84-93. doi:10.13582/j.cnki.1672-9102.2022.04.011

YAO J S, LIU Y L, WU Z J. Spatial-Temporal Behavior Characteristics of Tourists in Guilin Based on Digital Footprint [J]. Journal of Hunan University of Science and Technology (Natural Science Edition), 2022, 37(4): 84-93. doi:10.13582/j.cnki.1672-9102.2022.04.011

基于数字足迹的桂林游客时空行为特征

姚建盛,刘艳玲*,吴忠军

(桂林理工大学 旅游与风景园林学院,广西 桂林 541004;广西旅游产业研究院,广西 桂林 541004)

摘要:基于网络游记构建数字足迹数据库,利用季节集中指数、月际变动指数、旅游偏好指数、最邻近指数和社会网络分析等方法 and 可视化技术,从时间、空间和网络结构方面分析桂林游客的时空行为特征.分析结果表明:游客来桂林旅游的时间存在一定的季节性,停留天数以 3~4 d 为主;桂林景区的空间分布不均衡,访问偏好也不均匀,热门景区包括漓江、遇龙河、西街、两江四湖、龙脊梯田;从社会网络分析角度看,桂林核心景区为漓江、遇龙河、西街和两江四湖,基于空间聚类方法可划分 3 个强连通社区.最后,针对桂林游客时空分布的不均衡性提出政策建议.

关键词:时空行为;数字足迹;桂林;社会网络分析

中图分类号:TP391.1; F590

文献标志码:A

文章编号:1672-9102(2022)04-0084-10

Spatial-Temporal Behavior Characteristics of Tourists in Guilin Based on Digital Footprint

YAO Jiansheng, LIU Yanling, WU Zhongjun

(College of Tourism and Landscape Architecture, Guilin University of Technology, Guilin 541004, China;

Institute of Guangxi Tourism Industry, Guilin 541004, China)

Abstract: This paper constructs a digital footprint database based on online travel notes, and analyzes the spatial-temporal behavior characteristics of tourists in Guilin from the perspectives of time, space and network structure by using the methods of season concentration index, tourism preference index, nearest neighbor index, social network analysis, and visualization technology. Results show that there is a certain seasonal concentration of tourists to Guilin, and the number of days of stay is about 3 or 4 days. The spatial distribution of scenic spots in Guilin is disequilibrium, and the visit preference is uneven. The popular scenic spots include Lijiang River, Yulong River, West Street, Two Rivers and Four Lakes, and Longji Terrace. Besides, from the perspective of social network analysis, the core scenic spots in Guilin are Lijiang River, Yulong River, West Street, Two Rivers and Four Lakes, and there are three strongly connected community based on spatial clustering algorithm. Finally, this paper proposes some policy suggestions on the disproportion space-time distribution of tourists in Guilin.

Keywords: spatial-temporal behavior; digital footprint; Guilin; social network analysis

收稿日期:2021-01-15

基金项目:国家自然科学基金资助项目(72162011);广西科技重点研发计划资助项目(桂科 AB17195028);桂林市科学研究与技术开发计划资助项目(20180102-2;20170220);广西中青年教师基础能力提升项目资助(2019KY0292)

*通信作者,E-mail:2018014@glut.edu.cn

游客是旅游活动的主体,也是旅游产业的服务对象,只有掌握了游客的行为特征并理解其内在机理,才能加强旅游目的地的建设、提升旅游服务水平^[1]。游客时空行为(移动性)是游客在旅游活动中最主要的体现,是游客的自身情况、旅游动机、行为偏好和选择决策的外在表现,研究游客时空行为能更好地理解游客的旅游活动,对分析及预测旅游市场具有重要意义,是旅游研究的核心命题之一^[2]。

随着网络技术和网络应用的普及,商业化旅游平台逐渐成为旅游者出行前获取信息、选择线路的重要工具,同时也成为旅游者出行中和出行后分享旅游经历、表达旅游感受的重要媒体。Girardin 等^[3]首次提出“数字足迹”概念,并指出旅游数字足迹是游客在旅游中发送的信息和通话记录,以及旅游后在网络发表的文字图片等信息。李君轶^[4]在国内较早研究旅游数字足迹的内涵和特点,指出旅游数字足迹是游客在旅游活动中遗留在互联网、通讯网和物联网等信息设施上的具有定位信息或者地理标签的电子痕迹,这些电子痕迹可以反映游客的时空运动轨迹。

旅游数字足迹为研究者研究游客行为提供了一个全新的视角。近年来,研究者使用不同的数据源,如GPS^[5]、移动电话数据^[6]、地理标记照片^[7-8]、微博^[9]、游记^[10-13]、游记+照片^[14]和网络评论^[15]等研究游客移动模型^[6]、网络形象感知^[7]、旅游流网络结构^[8-9,11]、时空行为特征^[12,14]、游客满意度^[15]和网络舆情^[16]等。桂林作为国际旅游胜地,相关的研究却很少,与其旅游名城地位不符。为此,本文以桂林及其下辖市县为案例地,以网络游记为数据源,研究来桂林游客的时空行为特性,为桂林的旅游规划、形象提升和市场营销提供科学依据。

1 数字足迹库案例

1.1 案例地

桂林自古就享有“山水甲天下”的美誉,是国家批复建设的国际旅游胜地、对外开放国际旅游城市、全国旅游创新发展先行区、中国建成世界级旅游目的地的示范区等,是联合国世界旅游组织/亚太旅游协会旅游趋势与展望国际论坛永久举办地、世界旅游组织推荐中国最佳旅游城市之一。2019年,桂林接待国内外游客13 833.66万人次、实现旅游总消费1 874.25亿元。旅游业是桂林的支柱性产业,因此,研究桂林的旅游需求具有重要意义。

截至2019年,桂林下辖6区10县、代管1个县级市,旅游资源丰富,包括AAAAA级景区4家,AAAA级景区37家,AAA级景区29家和多家非A级知名景区,如印象刘三姐、桂林千古情、天湖生态旅游度假区、三千漓山水人文度假区等。

1.2 数字足迹库

携程网是中国最大的在线旅行服务提供商,拥有大规模的注册用户,在线旅游信息保存相对完整,信息更新速度快,适合作为样本来源库。通过对比不同类型的社交网络平台及各大旅游网站的游记内容,发现携程网上关于桂林的网络游记的数量较大,信息较全,适合本研究,因此,本文选择携程网平台作为研究的数据源。基于Python语言的网络爬虫技术采集2014.01.01—2018.12.31的游记样本共5 840篇,利用Python语言的Pandas库对游记进行清洗和预处理。

首先,清洗数据源。按照以下标准筛选游记:有游览路径;有游览天数,并且游览天数小于20 d;以桂林旅游为主,即删除游览路径中包含大量非桂林景区的记录。最后获得有效样本为2 057篇。

然后,进行数据预处理并构建数字足迹库。以桂林及其辖下市县的景区为研究对象,对数据进行如下处理:(1)提取2 057条游览路径,总长度包括19 254个景区节点,平均游览路径长度为9.4,去重后包括985个不同景区的名称;(2)合并相同景区的不同名称和位于同一个景区的不同兴趣点(POI),如表1所示;(3)删除非桂林景区(很多游览路径记录很详细,从出发地开始一直到返程地,因此,路径中包含非桂林景区);(4)统计景区在游览路径中出现的次数,作为景区的游览频度;(5)利用Python语言调用高德地图API获取景区经纬度。桂林和阳朔在整个游览路径中大量出现,但是二者为旅游目的地,包含众多景区,在统计分析时去掉桂林和阳朔,最后获得桂林景区79个。将以上信息整理后存入SQLite数据库,构建数字足迹时空数据库。

表1 部分景区名称合并

统一名称	合并名称
漓江	阳朔漓江, X-漓江桂林段, 桂林漓江, 漓江竹筏, 漓江游船, 漓江渔火, 漓江精华游, 九马画山, 黄布倒影等
遇龙河	遇龙河漂流, 阳朔遇龙河, 遇龙桥
西街	阳朔西街, 老西街, 桂林市阳朔县西街, 桂林阳朔西街
龙脊梯田	古壮寨, 龙脊古壮寨, 龙脊古壮寨梯田, 古壮寨, 大寨, 龙脊大寨, 大瑶寨, 平安寨, 平安壮寨, 平安壮族梯田, 龙脊平安, 黄洛瑶寨, 天下第一长发村, 七星伴月, 金坑梯田, 平安寨, 黄洛瑶寨, 金竹壮寨
象山公园	象山景区, 象鼻山, 象山, 桂林象鼻山, 橡皮山
靖江王府	独秀峰王城景区, 靖江王城, 独秀峰, 独秀峰·王城, 靖江王府
七星景区	七星岩, 七星公园, X-桂林七星岩, 七星岩牌坊广场, 桂林七星公园

2 研究方法

2.1 季节集中指数

季节集中指数又称季节性时间强度指数(季节强度指数), 可对网络游记发文时间的季节集中程度加以定量分析, 从而间接分析桂林旅游的季节集中程度, 季节集中指数的计算公式为

$$I = \sqrt{\frac{\sum_{i=1}^{12} (x_i - \bar{x})^2}{12}} \quad (1)$$

式中: x_i 为各月网络游记发文量占全年关注度总数的比重; \bar{x} 为各月网络游记发文量占全年网络游记发文量总数的比重的平均值, 假设服从平均分布, 则 $\bar{x} = 1/12 \approx 0.083$.

指数值 I 越大, 表明网络游记发文量的季节集中程度越高, 季节性差异越大; 相反, 指数值 I 越趋于 0, 表明网络游记发文量的季节集中程度越低, 季节分布越均匀.

2.2 月际变动指数

变动指数相当于将网络游记发文量标准化, 同样反映网络游记发文量在统计周期内的变化程度, 但能更好地观测数据的变化情况, 月际变动指数的计算公式为

$$S_i = \frac{y_i}{\frac{1}{12} \sum_{k=1}^{12} y_k} \times 100\% \quad (2)$$

式中: S_i 为第 i 月的网络游记发文量变动指数, $0 \leq S_i \leq 1$; y_i 为年内第 i 月的网络游记发文量; y_k 为第 k 月的网络游记发文量.

S_i 值越接近于基准值 100%, 表明网络关注度的月分布越均匀, 季节性越弱; 反之, 网络关注度的季节性越强.

2.3 旅游偏好指数

定量分析旅游偏好的方法有很多种, 本文采用孙根年^[17]提出的旅游偏好指数来分析游客对桂林市旅游景区的游览偏好, 旅游偏好指数的计算公式为

$$P = \frac{n_i}{X} \times 100\% \quad (3)$$

式中: n_i 为第 i 个景区在游览路径中出现的次数; X 为全部景区在游览路径中出现次数的总和.

P 的取值范围在 0~1, P 值越趋近于 1, 则游客对该景区的旅游偏好程度越强.

2.4 最邻近指数

点状要素有 3 种空间分布类型: 聚集、均匀和随机. 均匀分布的点状要素的最邻近距离最大, 随机分布次之, 凝聚分布最小. 最邻近距离是表示点状事物在地理空间中相互邻近程度的地理指标, 能很好地反映点状要素的空间分布特征, 计算方法是实际最邻近距离与理论最邻近距离(即随机分布时的理论值)的比值, 计算公式为

$$R = \frac{\bar{r}_1}{\bar{r}_E} = 2\sqrt{D}\bar{r}_1; \quad (4)$$

$$\bar{r}_E = \frac{1}{2\sqrt{\frac{n}{A}}} = \frac{1}{2\sqrt{D}}. \quad (5)$$

式中: R 为最近邻指数; \bar{r}_1 为最近点之间距离的平均值; \bar{r}_E 为理论最近邻距离; D 为点密度; A 为区域面积; n 为研究对象的数目.

$R=1$ 时,点要素随机分布; $R>1$ 时,点要素均匀分布; $R<1$ 时,点要素凝聚分布.

2.5 社会网络分析法

社会网络分析法是由社会学家根据数学方法、图论等发展起来的定量分析方法,有助于把个体间的关系、“微观”网络与大规模的社会系统的“宏观”结构结合起来,是旅游流研究的常用方法.本文以数字足迹库中的景区为节点,以路径为边,构建桂林旅游流网络.通过社会网络分析节点度、中心性和社区划分等,度量节点的重要性和节点之间的联系,从而可有效地揭示出各个景区之间的空间联系规律.

3 试验与结果分析

3.1 时间分析

统计数字足迹库中2014年—2018年每个月发文量的平均值,计算季节集中指数为0.012,月际变动指数的范围为0.065~0.095.从数值上看,桂林旅游没有表现出季节集中性.可视化游记数量的月际变动指数如图1所示.由图1可知:桂林旅游具有季节性,每年的4月—10月是旅游旺季,1月—3月和11月—12月相对为淡季,这和桂林的实际旅游情况比较吻合.这是因为月际变动指数的最高值(0.095)和最低值(0.063)的差值仅为0.03,所以从数值上看不到季节性,而可视化技术局部放大了 y 轴的这一区间,才观测到其中的微小差异.因此,桂林游记的发文数量和桂林游客的数量具有微弱的正向相关性.

图2是对数字足迹库中,游客游览天数和分布密度的统计和可视化.由图2可知:来桂林游览游客的出游天数呈幂律分布,其中以停留3~4 d为主,其次为停留5, 1, 2, 6和7 d,其他停留天数占比较小,平均停留天数为4.1 d.来桂林旅客住宿的比例较大,停留7 d以上的游客占比较小,说明来桂林深度旅游处在初级阶段,短期的观光旅游、休闲旅游在桂林总体旅游中的占比较大.

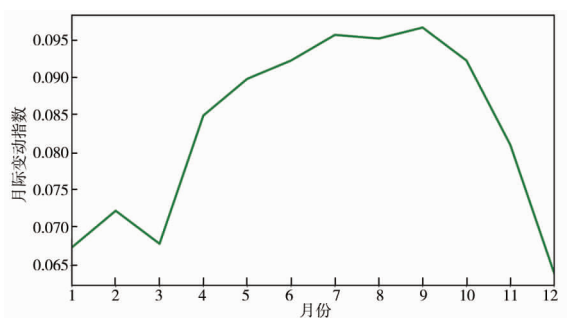


图1 游记数量月际变动指数

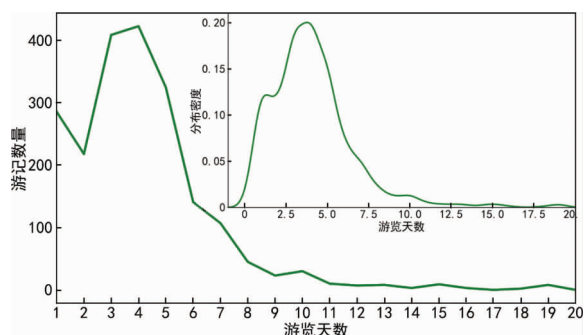


图2 游客游览天数分布

3.2 空间分析

3.2.1 空间分布

提取数字足迹库中整理的79个桂林景区、经纬度信息和桂林市辖区的几何形状数据,计算桂林景区的最近邻指数为0.377 $8<1$,表明桂林景区分布的聚集特征.图3为以高德地图为底图可视化桂林市辖区游客到访景区的分布图.由图3可知:桂林景区的分布南北狭长,主要沿江分布,北起资源八角寨景区,南到荔浦荔江湾景区,东到龙脊梯田景区,西到兴安灵渠和乐满地主题乐园.

景区分布不均衡,总体呈现出“大分散、小聚集”的空间分布格局,主要聚集区在桂林市区和阳朔,以及从市区到阳朔的漓江两岸和桂阳公路两侧,灵川、龙胜、资源、兴安和荔浦等下辖市县有零散分布,永福、

灌阳、全州、恭城和平乐等县的景区由于在游记的轨迹路线中出现的频率较低,没有具体统计.

3.2.2 景区访问偏好

以数字足迹库中景区在游览轨迹路径中出现的次数作为景区被访问次数的模拟值,按照式(3)计算景区旅游偏好指数,绘制景区访问偏好等值线图(图4)和景区访问偏好指数及其累计值占比(图5).由图4可知:桂林景区的访问偏好并不均衡,呈幂律长尾分布.由图5可知:从累计占比看,21.52%的热门景区占据了80.58%的访问量,可见桂林景区的访问偏好基本上符合二八定律(帕累托法定律).从等值线图可直观观测到访问热度从北到南分别集中在龙胜、桂林市、漓江沿岸和阳朔.

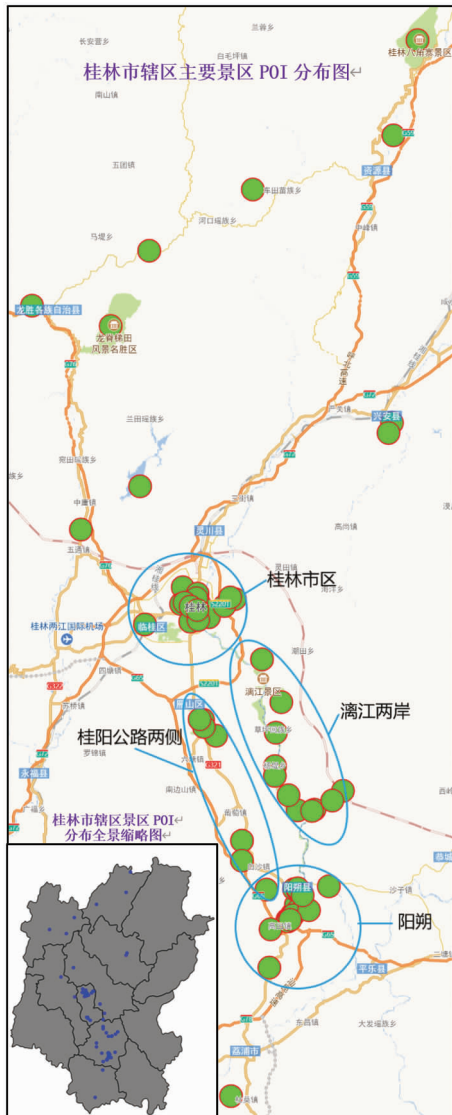


图3 桂林景区 POI 分布

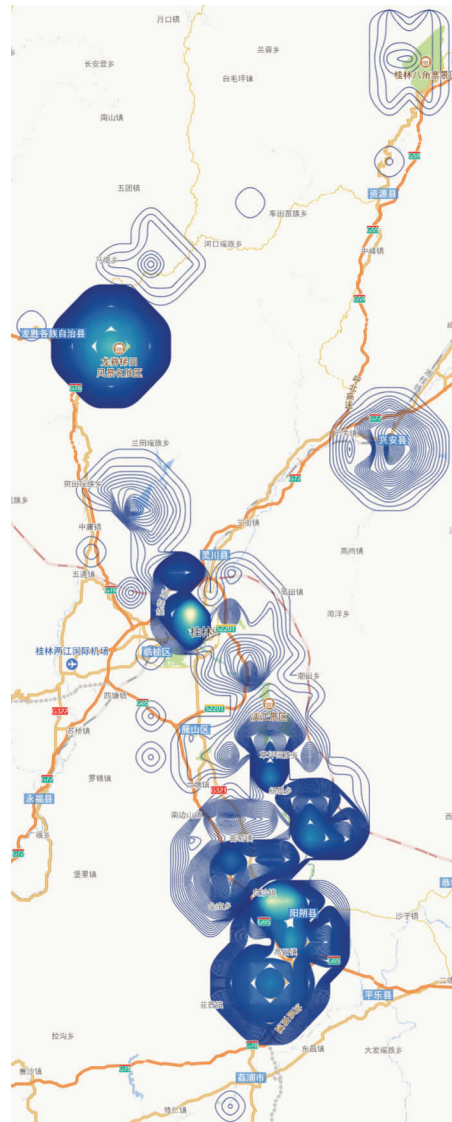


图4 景区访问偏好等值线

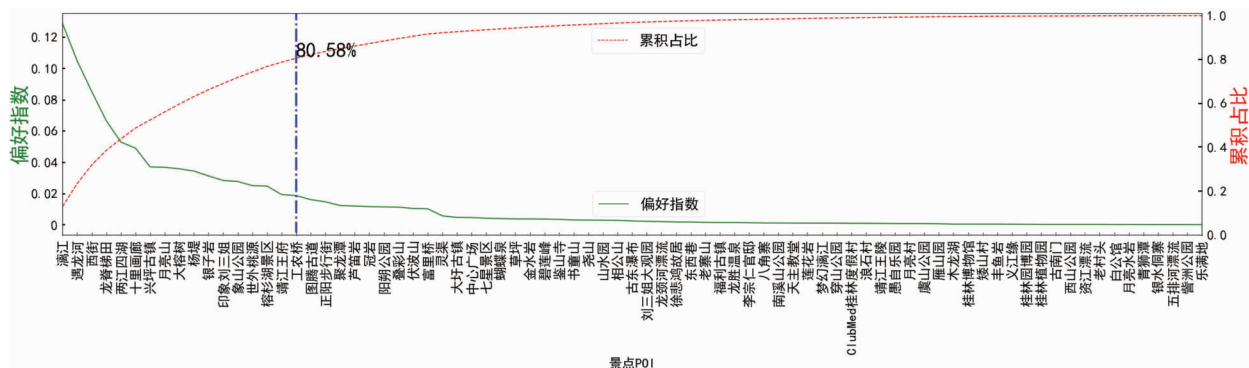


图5 景区访问偏好

排名前5的热门景区分别为漓江、遇龙河、西街、龙脊梯田和两江四湖,是桂林旅游的精华。“桂林山水甲天下,阳朔山水甲桂林”,漓江联通桂林和阳朔,是桂林山水风光的典型代表;遇龙河被称为小漓江,以田园风光著称,和城市的快节奏生活形成巨大反差,近年来游览热度逐年增加;与遇龙河的自然风光相比,西街是拥有1400多年历史的历史文化街区,融合中西特色的商业和文化,是阳朔最重要的人文旅游景点之一;龙胜梯田距今至少有2300多年的历史,是全球重要的农业文化遗产,堪称世界梯田原乡,是民族风情和自然风景的完美融合;两江四湖环城水系再现桂林“水城”的繁荣景象,号称“东方威尼斯”。

3.3 社会网络分析

从数字足迹库中提取游览轨迹路径,以景区为节点(79个),以景区间的轨迹路径为边(1215条),利用Python语言的复杂网络分析包NetworkX构建景区客流社会网络,并进行相关计算,通过Gephi软件进行可视化构建景区客流社会网络。其中,节点大小和边的权重(厚度)分别用节点度(和其他节点直接连接的次数)和边出现的次数度量,计算网络密度为0.2023,平均路径长度为2,网络直径为5,网络传递性为0.5468,网络平均度为15.577。可见,桂林作为我国较早建设的国际旅游城市,市辖区内景区的网络通达性较好,景区间的联系较紧密。

3.3.1 k-核分解

k-核分解是以节点度衡量节点的重要性并排序。景区社会网络分层如图6所示。网络按节点度的大小布局,节点度越大,离网络核心越近;反之,节点度越小,离网络核心越远。由于网络节点度的分布域较大,最小值为1,最大值为93(入度和出度和),因此,严格按照k-核分解以1为粒度进行层次划分,层次较多而且意义不大。论文借鉴k-核分解思想,按照度值将网络划分4个层次,网络分层如表2所示。核心层是桂林旅游的精华景区;次核心层是从核心层扩散的重要景区,如兴坪古镇是漓江的延伸,十里画廊、世外桃源、银子岩是遇龙河的延伸,印象刘三姐是西街的延伸,靖江王府、象山公园、正阳步行街是两江四湖的延伸;第三层是桂林旅游热度居中的景区;边缘层是网络末端,由于地理等原因导致访问热度较低。

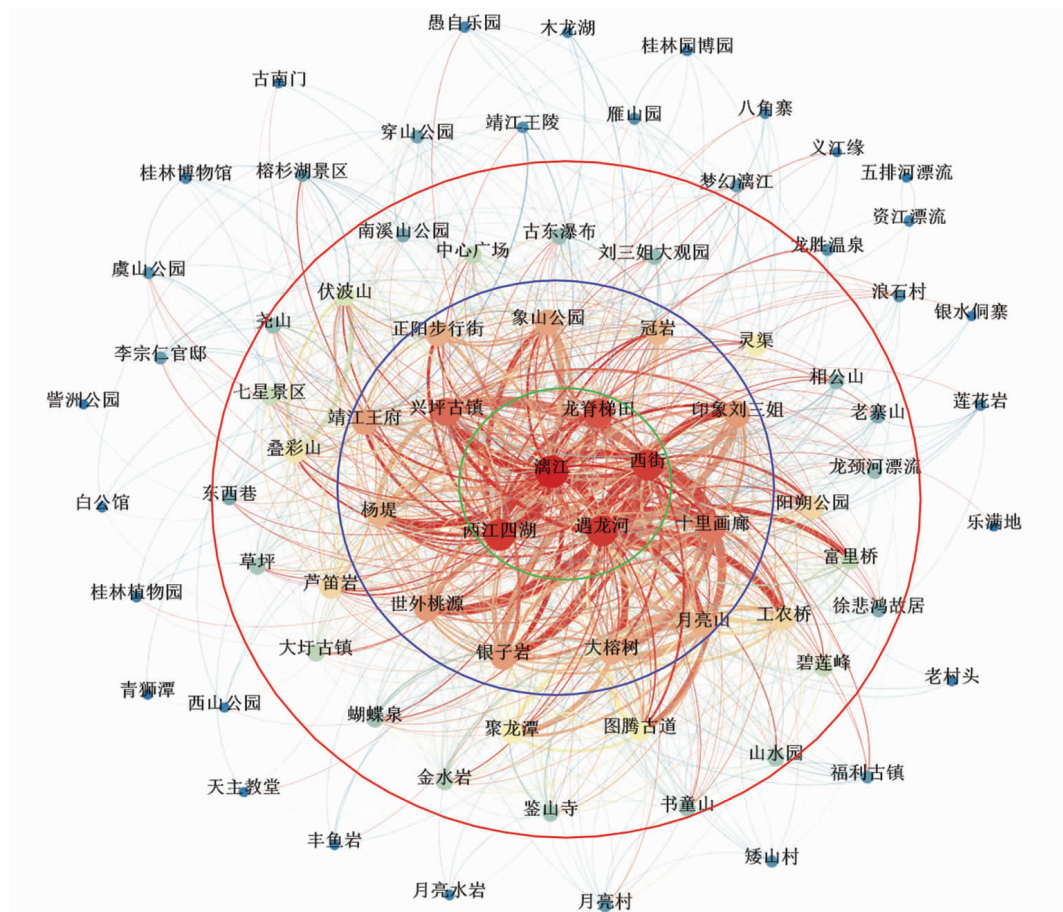


图6 景区社会网络分层

表2 网络分层

网络分层	节点集合
核心层节点	漓江、遇龙河、西街、两江四湖、龙脊梯田
次核心层节点	兴坪古镇、十里画廊、世外桃源、银子岩、印象刘三姐、靖江王府、象山公园、月亮山、大榕树、正阳步行街、杨堤、冠岩、工农桥、阳朔公园、芦笛岩、叠彩山、灵渠、图腾古道、聚龙潭、伏波山、七星景区、碧莲峰、中心广场、富里桥、大圩古镇
第三层节点	镇、金水岩、山水园、草坪、鉴山寺、蝴蝶泉、书童山、古东瀑布、尧山、刘三姐大观园、龙颈河漂流、相公山、徐悲鸿故居、南溪山公园、老寨山、东西巷
边缘层节点	其他节点

3.3.2 中心性分析

度中心性通过景点之间直接连接的数量,表达了景点和其他景点的联系程度,在有向网络中分为入度中心性和出度中心性.依据旅游流网络中节点的出度中心性和入度中心性可分析景点在网络中的重要性.当出度中心性和入度中心性都大时,表明景区位于网络的核心区;当出度中心性大而入度中心性小时,表明景点处于网络的辐射区;当出度中心性小而入度中心性大时,表明景点处于网络的聚集区;当出度中心性和入度中心性都小时,表明景区处于网络的边缘,对促进景区间游客流动的作用不大.表3列出中心性大于0.2的节点,图7可视化了入度中心性和出度中心性都大于0.3的景点.由图7可知:漓江是连通桂林市区和阳朔的水道,并且是桂林山水旅游资源最典型的代表,位于桂林旅游流网络的最核心;遇龙河、西街是阳朔的核心景区,两江四湖是桂林市区的核心景区;龙脊梯田的入度中心性明显高于出度中心性,趋向聚集区;其他景区分布在对角线附近,出入度比较均衡,说明在实际交通网络和社会网络中都处于中间过渡区域.

表3 网络部分节点中心性

排名	入度中心性(取值>0.2)		出度中心性(取值>0.2)		中介中心性(取值>0.2)		接近中心性(取值>0.5)	
	景区	中心性值	景区	中心性值	景区	中心性值	景区	中心性值
1	漓江	0.61	漓江	0.60	漓江	0.60	漓江	0.70
2	两江四湖	0.61	遇龙河	0.53	遇龙河	0.53	遇龙河	0.70
3	遇龙河	0.60	西街	0.52	西街	0.52	龙脊梯田	0.70
4	龙脊梯田	0.60	两江四湖	0.52	两江四湖	0.52	两江四湖	0.70
5	西街	0.60	兴坪古镇	0.47	兴坪古镇	0.47	西街	0.69
6	兴坪古镇	0.52	十里画廊	0.47	十里画廊	0.47	兴坪古镇	0.66
7	十里画廊	0.48	银子岩	0.45	银子岩	0.45	十里画廊	0.64
8	杨堤	0.44	龙脊梯田	0.45	龙脊梯田	0.45	杨堤	0.62
9	印象刘三姐	0.44	世外桃源	0.44	世外桃源	0.44	正阳步行街	0.62
10	正阳步行街	0.44	大榕树	0.44	大榕树	0.44	印象刘三姐	0.62
11	象山公园	0.43	靖江王府	0.43	靖江王府	0.43	象山公园	0.61
12	世外桃源	0.42	印象刘三姐	0.42	印象刘三姐	0.42	靖江王府	0.61
13	靖江王府	0.42	工农桥	0.42	工农桥	0.42	世外桃源	0.61
14	芦笛岩	0.42	象山公园	0.4	象山公园	0.40	芦笛岩	0.61
15	银子岩	0.40	冠岩	0.4	冠岩	0.40	银子岩	0.60
16	月亮山	0.40	月亮山	0.4	月亮山	0.40	月亮山	0.60
17	冠岩	0.38	叠彩山	0.38	叠彩山	0.38	冠岩	0.60
18	大榕树	0.36	阳朔公园	0.38	阳朔公园	0.38	大榕树	0.59
19	阳朔公园	0.32	聚龙潭	0.38	聚龙潭	0.38	图腾古道	0.57
20	图腾古道	0.31	正阳步行街	0.36	正阳步行街	0.36	灵渠	0.57
21	灵渠	0.31	杨堤	0.35	杨堤	0.35	叠彩山	0.56
22	叠彩山	0.29	灵渠	0.34	灵渠	0.34	工农桥	0.56
23	工农桥	0.29	图腾古道	0.32	图腾古道	0.32	阳朔公园	0.56
24	伏波山	0.26	芦笛岩	0.29	芦笛岩	0.29	伏波山	0.55
25	聚龙潭	0.26	伏波山	0.26	伏波山	0.26	碧莲峰	0.55
26	中心广场	0.23	金水岩	0.25	金水岩	0.25	中心广场	0.54
27	七星景区	0.23	七星景区	0.25	七星景区	0.25	七星景区	0.54
28	碧莲峰	0.23	富里桥	0.23	富里桥	0.23	草坪	0.53
29	富里桥	0.22	碧莲峰	0.23	碧莲峰	0.23	大圩古镇	0.53
30	大圩古镇	0.22	中心广场	0.22	中心广场	0.22	富里桥	0.53

续表 3

排名	入度中心性(取值>0.2)		出度中心性(取值>0.2)		中介中心性(取值>0.2)		接近中心性(取值>0.5)	
	景区	中心性值	景区	中心性值	景区	中心性值	景区	中心性值
31	草坪	0.21	山水园	0.22	山水园	0.22	聚龙潭	0.53
32	—	—	大圩古镇	0.22	大圩古镇	0.22	刘三姐大观园	0.53
33	—	—	—	—	—	—	古东瀑布	0.52
34	—	—	—	—	—	—	尧山	0.51
35	—	—	—	—	—	—	蝴蝶泉	0.51
36	—	—	—	—	—	—	东西巷	0.51

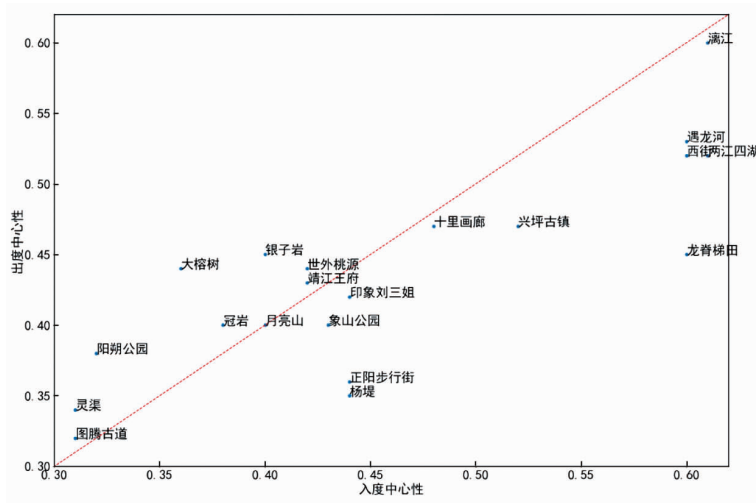


图 7 部分景区度中心性

度中心性基于节点间的直接联系度量节点的重要性,而中介中心性和接近中心性是从路径上考虑节点的重要性.中介中心性指的是 1 个景点担任其他 2 个景点之间最短路径的桥梁次数,反映景点在旅游流网络互动中对其他旅游节点的控制和依赖程度,中介中心性越高的景点,对其他旅游节点的控制性越强,则越是网络中的关键节点.接近中心性考量每个景区到其他景区的最短路径的平均长度,接近中心性的值越大,表明节点对节点之间客流流动的观察视野越好,可以设置旅游集散中心等旅游公共基础设施.表 3 给出网络中节点中心性大于平均值的中心性和排名,在所有排名中都位于前 5 的核心景区有漓江、遇龙河、西街和两江四湖,在所有排名中都位于前 10 的次核心景区(不包括核心景区)有十里画廊、兴坪古镇和龙脊梯田,在所有排名中都位于前 20 的重要景区(不包括核心和次核心景区)有印象刘三姐、靖江王府、象山公园、银子岩、世外桃源、正阳步行街、杨堤、冠岩、大榕树和月亮山.

3.3.3 空间聚类分析

通过空间聚类算法分析桂林景区有 3 个强连通社区,模块度为 0.347,平均聚类系数为 0.64.为理清网络社区划分结构,可视化边的权重大于等于 50 的子图,结果如图 8 所示.图 8 中以景区和路径在游览轨迹路径中出现的次数作为节点大小和边的粗细,以地理空间划分,可观察到 3 个社区.由图 8 可知:(1)以两江四湖为核心的城区风光带,主要包括象山公园,其他没有显示的景区如榕杉湖景区、古南门、伏波山、叠彩山等景区;(2)以漓江为核心的沿江风光带,主要包括杨堤和兴坪,其他没有显示的景区如古东瀑布、草坪、大圩古镇、相公山等;(3)以遇龙河为核心的阳朔风景区,也是最大的网络社区,主要包括西街、十里画廊、银子岩、印象刘三姐、月亮山、大榕树、世外桃源、工农桥、富里桥和图腾古道等,还有其他没有显示的景区.桂林经典游览路线是以漓江为纽带连通市区和阳朔的路线,如两江四湖—漓江—西街—遇龙河—十里画廊.

龙脊梯田景区从网络度上看位居第 4,但是由于不在桂林旅游的黄金带上,没有划入骨干网络.图 9 可视化了边的权重大于 20 的子图,并且只保留龙脊梯田和其他景区的相互连接.由图 9 可知:从桂林核心区(漓江、两江四湖、遇龙河和西街)都有去龙脊梯田的路径,其中,从两江四湖去龙脊梯田是主流路径,这与市区离龙脊较近、交通便利有关;从龙脊梯田到漓江的游客较多,到两江四湖的较少;从龙脊梯田到西街

和遇龙河的游客更少,在这个子图中没有体现。

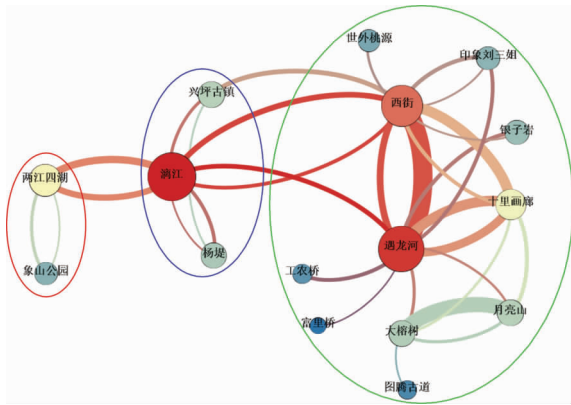


图8 社区检测

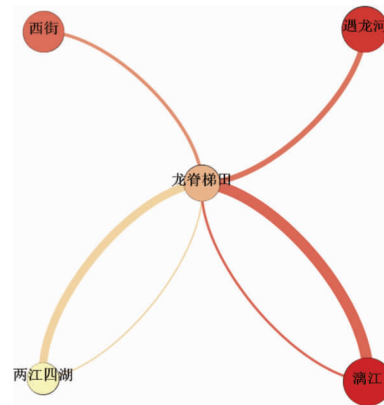


图9 龙脊梯田和其他景区的连接

3.3.4 社会网络分析应用

通过k-核分解、中心性和空间聚类社会网络分析方法,从不同角度分析桂林旅游景区的重要性和作用,为桂林旅游线路规划、旅游资源开发和旅游基础设施建设(如旅游集散中心、购物点等)提供科学依据。

漓江是桂林喀斯特地貌旅游资源的代表,并且在地理上是连接桂林市区和阳朔的纽带,从网络度、k-核分解和聚类分析看,漓江都是桂林旅游的核心,同时具有聚集和扩散功能,也是社会网络路径的枢纽。因此,从陆路(桂阳公路)和水陆(漓江)充分开发漓江的旅游资源、旅游集散中心等基础设施,针对桂林其他旅游景区和旅游产品进行营销,对桂林游客分流、增强旅游业竞争力和宣传力具有重要作用。

“阳朔山水甲桂林”,由此可以看出,阳朔是桂林重要的旅游目的地,其中,遇龙河和西街分别作为阳朔山水自然风景和人文景观的典范,已成为阳朔乃至桂林的核心景区。阳朔周边的旅游资源丰富,因此,增强阳朔及其周边乡村旅游的交通等基础设施建设,能进一步增加阳朔自驾游游客的数量,不仅能增强阳朔的分流作用,而且能进一步提升桂林的旅游形象。

两江四湖·象山景区是国家AAAAA级景区,是桂林城区的主要自然风景,也是桂林旅游的核心。靖江王城是桂林最重要的人文景观,是桂林作为历史文化名城的重要象征,而同是国家AAAAA级景区的独秀峰·王城景区却不在整个网络的核心位置。可见,“桂林山水甲天下”宣传了桂林的自然风景而忽略了桂林厚重的人文历史。因此,桂林旅游应该加强桂林历史文化名城的宣传力度,增强人文景观资源的开发、活化、利用和营销,促进文旅融合,从而提升桂林旅游的核心竞争力。

相比而言,龙脊梯田景区的入度中心性(0.6)明显高于出度中心性(0.45),具有较高的接近中心性(0.7)和较低的中介中心性(0.45),这表明龙脊梯田位于网络的聚集区,并且从桂林多个景区直接到达龙脊梯田景区的概率较高,但是从龙脊梯田到桂林其他景区的概率较低,该景区并不在桂林旅游网络路径的核心。这是由龙脊梯田的地理位置决定的,可以通过建设道路等旅游基础设施,增强龙脊梯田在网络路径中的核心位置,从而进一步优化桂林游客流的时空分布,减少节假日游客流的拥堵,降低旅游流管理成本,提升游客游览体验。

基于以上分析,一方面,进一步提升桂林核心景区的知名度应充分发挥核心景区的带动作用;另一方面,通过建设旅游基础设施和旅游资源开发等,增强次核心层的扩散和延伸作用,将游客从核心景区向边缘景区输送,整体提升桂林旅游的竞争力;此外,还应加强桂林人文景观的宣传,提升桂林历史文化名城的知名度,增强桂林旅游的核心竞争力。

4 结论

1)基于旅游数字足迹分析桂林游客时空行为,发现桂林旅游在时间和空间上的分布都不均衡,这主要是受桂林核心旅游资源(即漓江)的影响。时间上,漓江在冬季进入枯水期;空间上,很难辐射到远离漓江的景区。这也是一些其他受自然景观影响的旅游目的地所面临的共性问题,建议地方政府可采取政策进行

人工干预,以减小目的地内旅游时空分布的不均衡性。

2) 建议桂林市政府通过兴修水库等水利工程破解漓江水困难题;利用南北气候和民族文化差异,针对北方游客,打造冬游桂林旅游项目;通过建立交通、旅游服务中心等旅游基础设施,改善桂林景区现有的网络结构,以增强漓江对其他景区的辐射作用;利用新媒体营销策略包装非核心景区,打造新的“网红”打卡地,使桂林旅游向多核心发展,提升整体的市场竞争力。

3) 论文研究局限和未来研究计划。

(1) 数据源单一。论文只选择携程网作为数据源,事实上,其他 OTA 网站和社交网络,如马蜂窝、去哪儿网、猫途鹰和微博等都包含大量的游客时空数据。不同数据源拥有不同的用户群体,因此,多源数据分析结果可以更全面地展示用户时空行为特征。

(2) 数据类型单一。论文仅从文本中提取游客时空数据,而 OTA 网站(如马蜂窝等)的照片和微博的签到数据等都具有丰富的时空信息。多类型数据分析可以增加分析结果的全面性。

(3) 研究案例单一。论文仅以桂林为案例,不同旅游目的地受旅游资源和特征的影响,游客具有不同的时空行为特征。多个案例分析能更好地总结归纳游客旅游目的地的时空行为特征,为旅游目的地的管理、服务和营销建立更一般的理论基础。

4) 未来将考虑基于多源异构数据、多个案例分析游客移动特征,融合移动和 PC 端多个 OTA 网站及社交网站的文本、图片、签到等时空数据,总结出更普适的游客时空行为特征,从而更好地指导区域旅游发展。

参考文献:

- [1] Hannam K, Butler G, Paris C M. Developments and key issues in tourism mobilities[J]. *Annals of Tourism Research*, 2014, 44:171-185.
- [2] 黄潇婷,李玟璇,张海平,等.基于 GPS 数据的旅游时空行为评价研究[J].*旅游学刊*,2016,31(9):40-49.
- [3] Girardin F, Blat J, Calabrese F, et al. Digital Footprinting: Uncovering Tourists with User-Generated Content[J]. *IEEE Pervasive Computing*, 2008, 7(4):36-43.
- [4] 李君轶.旅游数字足迹:在线揭示游客的时空轨迹[J].*思想战线*,2013,39(3):103-107.
- [5] Zheng W, Zhou R, Zhang Z, et al. Understanding the tourist mobility using GPS: How similar are the tourists? [J]. *Tourism Management*, 2019, 71:54-66.
- [6] Zhao X, Lu X N, Liu Y Y, et al. Tourist movement patterns understanding from the perspective of travel party size using mobile tracking data: A case study of Xi'an, China[J]. *Tourism Management*, 2018, 69:368-383.
- [7] Zhang K, Chen Y, Li C. Discovering the tourists' behaviors and perceptions in a tourism destination by analyzing photos' visual content with a computer deep learning model: The case of Beijing[J]. *Tourism Management*, 2019, 75:595-608.
- [8] 范梦余,陈怡宁,张辉.呼伦贝尔游客 POI 空间格局与移动轨迹研究——基于地理标记照片的实证分析[J].*干旱区资源与环境*,2019,33(6):203-208.
- [9] 贺小荣,李宗幸,李启明,等.基于数字足迹的风景区旅游者时空结构特征研究——以赴张家界景区的旅游者为案例[J].*湖南师范大学自然科学学报*,2018,41(1):11-17.
- [10] 张鲜鲜,李婧晗,左颖,等.基于数字足迹的游客时空行为特征分析——以南京市为例[J].*经济地理*,2018,38(12):226-233.
- [11] 韩冬,黄丽华.基于旅游数字足迹的旅游流网络结构研究——以内蒙古自治区为例[J].*干旱区资源与环境*,2018,32(3):192-197.
- [12] 刘震,戴泽钊,楼嘉军,等.基于数字足迹的城市游憩行为时空特征研究——以上海为例[J].*世界地理研究*,2019,28(5):95-105.
- [13] 马晓路,向玉成,曾雪晴.旅游数字足迹下川西藏区景区空间结构研究[J].*西藏研究*,2019(1):64-72.
- [14] 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[J].*旅游学刊*,2016,31(12):41-50.
- [15] Zhao Y, Xu X, Wang M. Predicting Overall Customer Satisfaction: Big Data Evidence From Hotel Online Textual Reviews[J]. *International Journal of Hospitality Management*, 2019, 76(1): 111-121.
- [16] 周倩,姜磊,程宇航,等.情感倾向驱动的旅游酒店大数据网络舆情[J].*湖南科技大学学报(自然科学版)*,2020,35(4):67-73.
- [17] 孙根年.国际旅游支付方程、支付等级与旅游偏好[J].*地理学与国土研究*,2001,17(1):50-54.