

# 基于社交媒体数据的旅游移动模式提取

孙奇 张毅<sup>†</sup> 赵鹏飞 吴梦彤

北京大学地球与空间科学学院遥感与地理信息系统研究所, 北京 100871; <sup>†</sup> 通信作者, E-mail: zy@pku.edu.cn

**摘要** 提出一种从社交媒体大数据中提取个体旅游时空行为,再基于海量旅游时空行为挖掘群体城市间移动模式的方法。采集4000多万条到访过苏州市用户的全球地理微博,从中提取88270条旅游时空行为轨迹,识别出5类36种城市间旅游移动模式。结果表明,提取的移动模式符合LCF理论模型;除简单移动模式外,还存在更加复杂的复合移动模式。基于大数据能够得到更全面更精准的旅游移动模式,有助于旅游管理者了解游客动向及偏好,调整目的地营销策略,优化旅游资源配置,为游客提供更好的服务。

**关键词** 旅游移动模式; 社交媒体; 地理大数据; LCF 模型

## Travel Movement Pattern Extraction Based on Social Media Data

SUN Qi, ZHANG Yi<sup>†</sup>, ZHAO Pengfei, WU Mengtong

Institute of Remote Sensing and Geographic Information Systems, School of Earth and Space Sciences, Peking University, Beijing 100871; <sup>†</sup> Corresponding author, E-mail: zy@pku.edu.cn

**Abstract** The authors propose a method to extract individual travel spatiotemporal behaviors from social media data, and then mine the group based on massive spatiotemporal behaviors. This study collects more than 40 million global geographic microblogs from users who have visited Suzhou, extracts 88270 tourism trajectories, and identifies 36 classes of inter-city tourism movement patterns in five categories. It is found that the extracted patterns conform to the LCF theoretical model; besides the simple movement patterns, there are more complex composite movement patterns. Based on big data, more accurate tourism movement patterns can be obtained, which helps tourism managers understand tourists' trends and preferences, adjust destination marketing strategies, optimize tourism resource allocation, and provide better services for tourists.

**Key words** tourism movement pattern; social media; geographic big data; LCF model

人类移动模式指人群移动的时空规律,研究移动模式有助于理解人类移动行为和城市空间结构,在城市规划和交通管理等方面具有重要意义<sup>[1]</sup>。旅游移动模式指游客离开常住地后,在各目的地城市旅游的时空轨迹呈现的特定移动规律<sup>[2]</sup>,了解游客的移动模式对目的地营销和规划有重要意义<sup>[3-4]</sup>。从旅游行程中提取的移动模式能够帮助旅游管理者更好地了解游客的决策行为<sup>[5]</sup>、消费习惯<sup>[6-7]</sup>以及旅游偏好<sup>[8-9]</sup>,甚至可以基于用户偏好的移动模式来调整营销策略<sup>[10-11]</sup>,推出旅行产品<sup>[12]</sup>,推荐热门线路<sup>[13-14]</sup>。

有关旅游移动模式的理论研究已经比较成熟。Mercer<sup>[15]</sup>于1970年首次定义旅游流的概念。Lue等<sup>[16]</sup>对旅游路线的空间模式进行系统的分析,并划分为单目的地模式、往返模式、基营模式、区域环游模式和完全环游模式5种。2002年,Stewart等<sup>[17]</sup>将这一模式体系称为LCF模型。Oppermann<sup>[18]</sup>将旅游移动模式划分为7种,包括两种单目的地模式和5种多目的地模式。Flognfeldt Jr<sup>[19]</sup>通过研究7年的游客出行数据,将旅游移动模式划分为一日游、度假游、基营游和观光游4种类型。Lew等<sup>[20]</sup>定义3种目的地城市内部的移动模式:点点模式、环状模式

和复杂模式。Tussyadiah等<sup>[21]</sup>开发兰卡斯特(Lancaster)模型来解释多目的地旅行套餐的最佳组合。这些理论模型为后续研究(如时空约束<sup>[22]</sup>、游客体验<sup>[23]</sup>、目的地熟悉度<sup>[24]</sup>和节点结构<sup>[25]</sup>等)奠定了坚实的基础。

研究者采用多种来源的数据进行实证研究,通常用调查问卷和访谈的方式获取实验数据<sup>[24,26-27]</sup>。一些研究者将GPS数据与访谈数据相结合,得到粒度更细的轨迹信息和游客信息<sup>[28-31]</sup>。地理信息系统(geographic information systems, GIS)也广泛地应用于旅游流和空间模式中<sup>[32-33]</sup>。Yang等<sup>[34]</sup>通过探索性的数据分析,发现入境和国内旅游流的全局莫兰(Moran's *I*)指数具有显著的空间正相关关系,并利用Moran显著性水平图提取4个重要的入境旅游热点。

通常用一些回归模型来分析各种因素变量在旅游移动模式中的作用。Zhang等<sup>[35]</sup>从供给侧的视角,提出一个模型来探讨并解释国际旅游流动的影响因素(如自然条件和基础设施等)。Koo等<sup>[36]</sup>发现,休闲旅游的移动模式选择受旅游环境影响,而较低的机票价格会减弱旅游移动模式选择偏好的影响。Yang等<sup>[37]</sup>使用嵌套logit模型来评估旅游决策的影响因素,发现空间配置也是影响游客目的地选择的因素之一。Le-Klähn等<sup>[38]</sup>通过一个双变量probit模型,发现影响游客交通方式的多种因素,如访问地区、教育背景和游客动机。

传统的问卷调查方法费时费力以及GPS设备需要回收等原因,使得近年来研究者将更多的注意力放在目的地城市内的移动模式方面。随着GPS定位设备和移动互联网的普及,收集人们的出行轨迹愈加便捷,使大数据驱动的人类移动模式成为研究热点<sup>[1]</sup>。研究人员通过移动定位数据<sup>[39-41]</sup>、社交媒体数据<sup>[42-45]</sup>、旅游博客数据<sup>[46-47]</sup>以及其他来源的数据<sup>[48-49]</sup>,对游客的空间行为有了更加深入的了解。与传统的数据源相比,大数据研究节省时间,降低人工成本,具有更高的时空分辨率<sup>[41,50]</sup>。

庞大的数据量也导致研究方法的转变。研究者通过社会网络分析和复杂网络方法,计算得到目的地网络的中心性指数、密度、中介中心性和出度中心性<sup>[51-53]</sup>等指标来衡量旅游移动模式的特征。网络motif指出现频率非常高的子图,通常用于挖掘常见的旅游移动模式<sup>[54]</sup>。Liu等<sup>[53]</sup>基于twitter数据,发现居民和游客的主要移动模式是短距离移动,而

用户均匀度呈现的“T”形核心-外围结构可能与州际公路有关。

在基于大数据的旅游行为研究中,根据游客流量构建目的地网络是一种常用的方法<sup>[46,51,55]</sup>。现有的研究中,较多关注目的地城市网络体现的空间结构或目的地城市在移动模式中所处的位置,较少关注游客群体呈现的移动模式。本文通过海量个体旅游时空轨迹,挖掘游客在目的地城市间的不同移动模式,着重关注旅游行为中游客呈现的移动模式。一方面,用大数据方法验证和细化传统的旅游移动模式理论模型;另一方面,研究结果可以帮助当地旅游管理者了解游客的旅游行为。本文以苏州市为研究案例,以新浪微博为数据源,通过确定常居城市、提取游客出行行为以及识别游客出行移动模式,提取到访苏州游客的城市间移动模式。

## 1 研究区域与数据

本文选择苏州市进行案例研究。苏州市以其独特的园林景观,被誉为“中国园林之城”。苏州市又处在经济高度发展、交通高度连通的长江三角洲(简称长三角)地区,周边上海、南京和杭州等城市都具有丰富的旅游资源,到访苏州的游客通常会同时选择这些城市。例如,“马蜂窝旅游”中就有这样的推荐线路:上海→苏州→乌镇→杭州。

本文采用的数据源是新浪微博数据。新浪微博是提供微型博客服务的社交网络网站,其中主要有两种地理标注数据,分别是地理微博(geo-tagged weibo)和签到微博(check-in weibo),内容结构如表1所示。签到微博是地理微博的一个子集,其中增加了一个POI字段,记录微博的签到POI。为了尽可能地还原游客的时空轨迹,本研究使用全部的地理微博数据。

根据新浪地理微博的抓取策略,首先选择一个城市(苏州),然后抓取苏州市的POI数据。在此基础上,抓取苏州市的签到微博,从而得到签到用户。最后,抓取苏州市签到用户在全球的地理微博。我们共采集苏州市171045条POI数据,并据此获取2710510条签到微博,涉及748847名用户。根据这些用户信息,提取到分布在全球的48409276条地理微博。

## 2 研究方法

本文提出的从社交媒体数据中提取旅游移动模

表 1 地理微博和签到微博的数据结构及示例  
Table 1 Data structure and example of geo-tagged weibo and check-in weibo

类型	用户 ID	微博 ID	文本内容	时间	坐标	POI 信息
地理微博	2663431322	3557229653678318	踏上旅途, 让心飞扬, 第二站。我 在这里: <a href="http://t.cn/zjAymMc">http://t.cn/zjAymMc</a>	2013-05-08 15:51:42	31.323 120.611	NULL
签到微博	2803567942	3605499255854876	周庄纸箱王有人在歌唱 我在这里: <a href="http://t.cn/zHDwJV3">http://t.cn/zHDwJV3</a>	2013-07-29 20:37:52	31.116 120.846	周庄古镇

式的处理流程包含以下 4 个步骤。

1) 反向地理编码: 根据地理微博的经纬度坐标, 得到用户发布微博时所在的地市级行政区。

2) 确定用户常居地: 根据用户的地理微博发布时间序列, 计算得到用户最可能的常居城市。

3) 提取旅游时空行为: 基于空间移动、时间间隔和狭义旅游等规则, 对用户的地理微博时间序列进行划分, 并提取游客的时空行为数据。

4) 识别城市间的移动模式: 对旅游行为进行形式化表示, 并提取其中出现的旅游移动模式。

## 2.1 用户常居地的确定

世界旅游组织将旅游定义为基于观光休闲目的, 离开常居城市, 去往他处短暂逗留(不超过一年)的活动<sup>[56]</sup>。因此, 识别常居城市是研究旅游移动模式的重要前提。中国大陆共计 333 个地市级行政区, 本文采用信息熵来衡量各个城市成为游客常居城市的可能性。对于一个特定的用户  $U$ , 通过反向地理编码得到该用户到访的城市集合  $C=\{c_1, c_2, \dots, c_n\}$ 。对于每个城市  $c_k$ , 得到该用户在该城市发布的微博总数  $N_k$ 。用  $n_{k,m}$  表示该用户  $m$  月份在城市  $c_k$  发布微博的总数, 并用以下公式计算熵值  $E_k$ :

$$E_k = \sum_p -p \ln(p) = \sum_{m=1}^{12} -\frac{n_{k,m}}{N_k} \ln\left(\frac{n_{k,m}}{N_k}\right) \quad (1)$$

根据信息熵的原理, 数据分布越均衡, 熵值越大; 用户在常居城市停留时间最长, 在全年范围内的微博数量分布会更加均匀, 所对应的信息熵值最大。因此, 最大熵值对应的城市就确定为用户的常居城市。

## 2.2 旅游行为的提取

确定常居城市后, 需要对用户的连续微博序列进行划分, 得到该用户的若干次旅游出行时空行为。本文采用以下 4 个规则划分用户的旅游行为。

1) 常居地规则: 根据世界旅游组织的定义, 旅游行为发生在游客离开惯常环境情况下。因此, 我们规定游客的旅游活动发生在常居城市之外。如果

常居城市出现在游客的城市序列中, 那么这个序列就会从常居城市处被分隔开, 即如果常居地出现  $k$  次, 则用户的微博序列将被划分为  $k+1$  个子序列。

2) 时间间隔规则: 用时间间隔阈值来衡量相邻两条微博是否属于同一次出行行为。如果相邻两条微博的时间间隔超过给定的阈值, 那么就假定游客已结束当前出行行为, 并开始下一次出行行为。采用时间间隔阈值这一参数, 可以防止用户返回常居城市但未发布微博而导致序列过长的情况。本文将时间间隔阈值设定为 3 天。

3) 简化规则: 合并出行行为中连续的同一点, 即相同城市。

4) 旅游规则: 国内将旅游分为广义旅游和狭义旅游<sup>[57]</sup>。广义旅游指离开常居城市, 在他处短暂逗留的所有活动; 狭义旅游则规定短暂逗留活动中必须包含游憩活动。本文采用狭义旅游的概念, 规定用户在出行过程中至少到访过一个景区, 该次出行才是旅游行为。如果其行程未包含景区, 则用户在该次出行中不是游客身份, 该次出行也不是旅游行为。

常居地规则给出空间约束条件, 时间间隔规则给出时间约束条件, 简化规则减少出行行为表示的复杂度, 旅游规则给出出行动机的限制。通过上述 4 个规则, 可以从社交媒体数据中提取游客的若干次出行, 即旅游时空行为轨迹, 将其表示为目的城市的时间序列。

下面以一个用户的行为为例, 说明如何从游客的微博序列中划分出多个出行行为(Trip)。如图 1 所示, 从 20 天的时间线上可以提取 3 次出行行为。第 1 次出行中, 虽然用户在第 3 天和第 7 天没有发布微博, 但处于时间间隔规则允许的范围内, 因此这段出行是连续的。第 2 次出行中, 用户在第 9 天和第 12 天返回常居城市, 因此该次出行只包含第 10 天和第 11 天。第三次出行中, 虽然用户在第 16 天后没有返回常居城市, 但超过时间间隔的阈值, 因此该次出行在第 16 天后结束。最终, 每次出行

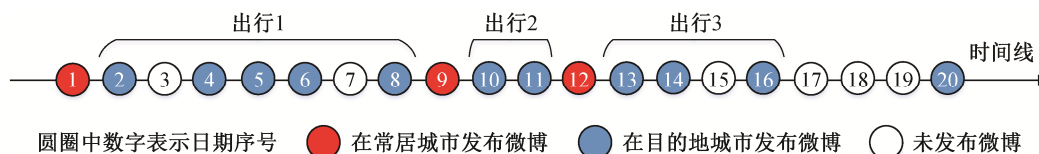


图 1 以某个用户的微博序列为例, 进行行为划分

Fig. 1 A user's weibo timeline, used as an example

(Trip)都可以表示为如下形式:

$$\text{Trip}_k = \{R, c_1, c_2, \dots, c_N, R\}, \quad (2)$$

其中,  $R$  是用户的常住城市;  $c_i$  是行程中第  $i$  个目的地城市, 与  $c_{i+1}$  表示的城市不同。

### 2.3 城市间旅游移动模式的识别

对提取的旅游行为, 需要用统一的形式表示。本文用有向图表示游客旅行过程中的城市序列: 用有向图中的节点表示游客的常居城市和目的地城市, 用节点之间的有向边表示游客在相邻两个城市间的移动行为。根据当前行为中游客到访目的地城市的先后顺序, 确定有向图中节点的序号。例如, 序号为 1 的节点表示游客到达的第一个目的地城市。

然后, 将有向图转换成一个邻接矩阵。为了便于存储, 将方阵按行展开, 转换为二进制字符串, 最后转换为十进制值, 作为唯一图 ID (GID)。旅游移动模式的表示方法如图 2 所示。

在利用有向图抽象地表示游客在常居城市与目的地城市间移动的拓扑关系基础上, 识别旅游行为的移动模式。在有向图中, 只保留游客访问节点的顺序关系, 不保留节点的其他信息(如城市名称等)。将游客的旅游行为表示为对应的 GID, 相同的 GID 即为同一个模式。例如, 两个游客分别有旅游行为“北京市-上海市-苏州市-北京市”和“上海市-无锡市-苏州市-上海市”, 这两种行为都可以表示成图 2(a)中的节点关系, 即从常居地出发, 访问两个不同的城市后, 返回常居城市。虽然两个行为中

包含的城市不完全相同, 但其拓扑结构相同, 属于同一个移动模式, 即 GID=140 的模式。

## 3 研究结果

### 3.1 旅游行为数据

根据旅游的定义以及社交媒体的稀疏性特征, 本文选择符合以下 3 个条件的用户作为研究对象: 1) 非苏州市居民; 2) 在常居城市至少发布 50 条微博; 3) 在出行过程中至少到达过一个景区。我们共得到 58720 个用户。从这些用户中提取 88270 个旅游时空行为, 如表 2 所示。随着目的地城市数量的增加, 出行次数迅速减少, 相应的平均旅行天数不断地增加。大多数旅游时空行为只有一个目的地城市, 其平均逗留时间为 1.1 天。

### 3.2 城市间移动模式

从 88270 个旅游时空行为中识别出 853 种城市间移动模式。选择出现频次超过 25 次的模式作为频繁模式, 最终识别出 36 种旅游移动模式。目的地城市数量是影响游客旅游行为及移动模式的重要因素, 目的地的数量会影响游客的出行规划(如全程旅行天数)。因此, 尽管一些模式之间仅在目的地城市数量方面存在差异, 但本文都单独进行表示(图 3)。

根据 LCF 理论模型, 按照拓扑结构特征, 可以将图 3 中 36 种旅游移动模式分为 5 类。

表 2 按目的地城市数量划分的出行数量和平均旅行时间  
Table 2 Quantity and duration of trips with different number of destination cities

目的地城市数量	出行次数	平均旅行天数
1	67634	1.1
2	8746	2.8
3	4627	4.3
4	2592	5.2
5	1625	6.3
6	1063	7.3
>7	1983	10.1
总计	88270	-

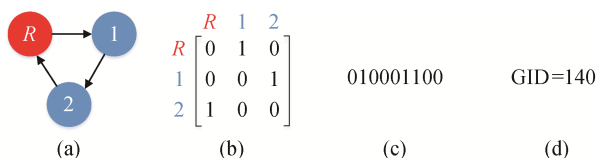
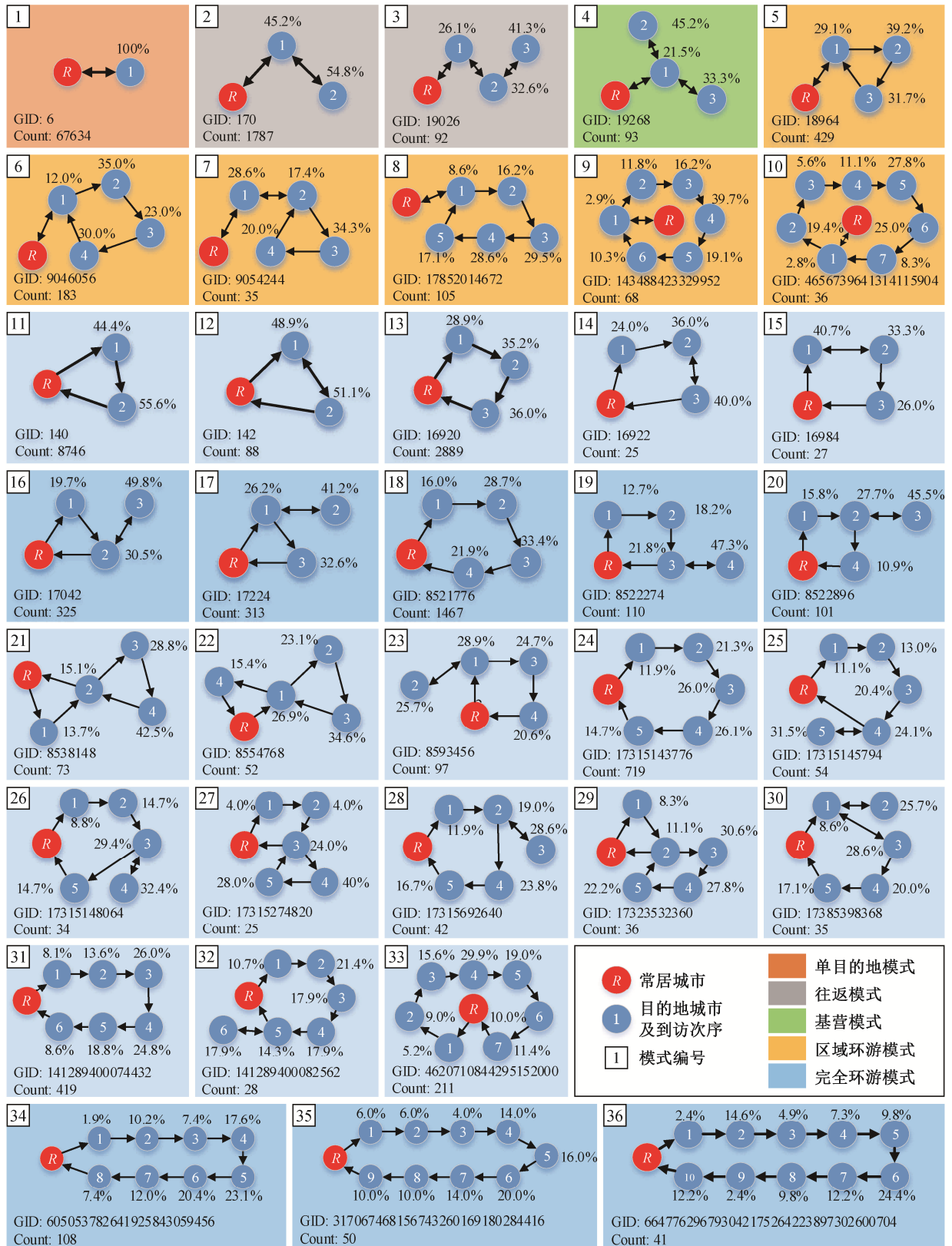


图 2 不同的旅游移动模式表示方法  
Fig. 2 Different representation methods of travel movement patterns



百分数表示苏州市出现在当前节点的概率, GID 表示子图的唯一标识序号, Count 表示子图在数据集中出现的次数

图 3 36 种旅游移动模式的识别结果  
Fig. 3 Recognition results of 36 travel movement pattern

第 1 类(模式 1): 单一目的地模式。该模式是最简单的一种旅游移动模式, 仅由常居城市和单个目的地城市组成, 是苏州旅游行为的主导模式, 占总数的 76.62%。

第 2 类(模式 2 和 3): 往返模式。该模式在路径上具有往返特性。第一个目的地城市通常具有交通枢纽的功能。从常居城市到第一个目的地的距离一般较远, 而目的地城市之间的距离较短。往返模式中目的地城市数量较少, 只有两个子模式与之匹配, 分别包含两个和 3 个目的地城市。

第 3 类(模式 4): 基营模式。该模式的特点是以一个目的地城市为基地, 将游客的常居城市与其他几个目的地城市连接起来, 整体上呈放射状。基营模式通常出现在经济不发达地区, 游客需要选择区域中接待要素聚集的相对发达的城市作为基地。苏州市所处的长三角地区经济较发达, 该模式数量很少。

第 4 类(模式 5~10): 区域旅游模式。该模式的特点是门户目的地(游客离开常居城市后到达的第一个目的地<sup>[58]</sup>)和离开目的地(游客返回常居城市之前到达的最后一个目的地)是同一个城市, 称为枢纽目的地。其他目的地和枢纽目的地形成一个单向环路。通常, 枢纽目的地是本地区交通最便利、经济最发达的城市(如长三角地区的上海市)。区域旅游模式更加多样化, 包括 6 个子模式。其中, 模式 7 (GID=9054244)比较特殊: 游客首先按照往返模式到达第 2 个目的地城市, 进行一次区域环游, 最后返回第 2 个目的地城市。可以将该模式视为往返模式与区域旅游模式这两个简单模式的组合。

第 5 类(模式 11~36): 完全环游模式。该模式的特点是门户目的地和离开目的地是不同的城市, 且所有目的地大致形成一个单向环路。实验结果中多目的地城市的旅游行为大多属于该模式, 并表现出更多的形态, 大致分为单环型、单环+支路型和双环型 3 个亚型。

单环型(模式 11~15, 18, 24, 31 和 33~36)对应 LCF 模型中的简单单环类型, 有多个子模式, 目的地城市数量为 2~10 个, 目的地城市之间可能出现往返现象(模式 12, 14 和 15)。

单环+支路型(模式 16, 17, 19, 20, 23, 25, 26, 28, 30 和 32)是简单单环型与往返模式和基营模式相结合, 在完全环游的过程中, 游客以其中一个目的地为城市基地, 到达另一个独立目的地城市后返回,

继续完成环游。单环+支路型子模式之间的差异除与目的地城市数量有关外, 还与基地城市在环路中的绝对位置有关。

双环型(模式 21, 22, 27 和 29)是区域环游模式与完全环游模式相结合, 游客在一条大型环路上进行局部环游。双环型中子模式的拓扑结构比其他类型复杂, 随着目的地城市数量和拓扑复杂度增加, 对应模式的出现频率逐渐降低。

根据数据采集方式, 游客肯定访问过苏州市, 即本文识别的模式是到访苏州游客的城市间移动模式。表 3 显示到访苏州的旅游行为中单目的地、往返、基营、区域旅游和完全环游 5 类模式的数量和比例, 可以看出, 单目的地旅游移动模式超过 70%, 是苏州游客的主导模式, 客源主要来自周边城市, 特别是上海; 多目的地移动模式中, 完全环游模式占绝对优势(18.26%), 其次是区域旅游模式(2.13%), 选择往返模式和基营模式的行为总计只有 1%左右。另外, 出现更复杂的旅游移动模式, 可以用多个简单模式的组合来解释。虽然旅游移动模式多种多样, 但游客倾向于选择简单的路线旅行。在目的地城市数量确定的情况下, 选择不同拓扑形式的游客数量排序为单环型>单环+支路型>往返型, 符合以较低的成本游览更多目的地的游客消费心理。

## 4 结束语

本文提出大数据驱动的旅游移动模式提取方法, 其中两个核心部分为旅游行为提取和城市间移动模式识别。旅游行为提取方法结合社交媒体数据的特点, 基于时空约束及旅游规则, 精准地获取游客的出行行为。城市间移动模式识别方法则通过有向图表示城市间的拓扑结构, 发现游客的移动模式。从新浪微博数据中提取 88270 个到访过苏州的

表 3 不同移动模式的行为数量和比例  
Table 3 Number and proportion of different movement patterns

旅游移动模式	数量	占比/%
单目的地模式	67634	76.62
往返模式	1879	2.13
基营模式	93	0.11
区域环游模式	856	0.97
完全环游模式	16115	18.26
无明显规律的其他行为	1693	1.92
合计	88270	100.00



旅游时空行为,并从中识别出5大类36种城市间的频繁移动模式。与基于小数据的研究<sup>[24,59]</sup>相比,基于社交媒体数据的研究具有数据量大、时间跨度长等优势,得到的旅游移动模式类型更加丰富,模式表示更加细粒度。

通过社交媒体提取的到访苏州游客的城市间旅游移动模式覆盖LCF模型中全部5类简单移动模式,验证了LCF理论模型。同时,实验结果中出现由简单模式组合形成的复合旅游移动模式,表明游客不是单纯地遵循一个简单模式在目的地城市之间流动,可能会出于不同的动机,甚至综合多个动机,对旅游线路有不同的选择,形成复杂的城市间移动轨迹<sup>[16]</sup>。本研究验证、细化并丰富了LCF模型中的旅游移动模式类型。

受多种因素的影响,本研究中旅游移动模式的选择与以往研究结果<sup>[5,11,22,60]</sup>存在差异。到访苏州游客的城市间移动模式中,单目的地模式占据主导地位;多目的地模式中,完全环游模式占据主导地位。这种模式选择差异是目的地城市的区位、社会经济水平、旅游资源配置、文化差异、旅游动机和旅游时间等因素影响的结果<sup>[16]</sup>。选择单目的地模式的客源以周边城市为主,旅游动机主要是周末休闲度假;选择多目的地模式的客源以远距离的城市为主,旅游动机主要是观光<sup>[61]</sup>。

社交媒体为旅游行为研究提供了新数据源,通过社交媒体大数据得到的频繁旅游移动模式能够更全面更精准地描述游客在目的地城市间的移动行为,有助于旅游决策者把握游客动向及偏好,调整目的地营销策略,优化旅游资源配置,为游客提供更好的服务,也有助于旅游管理者快速地获取目标人群,进行精准营销。例如,可以针对热门的单目的地旅游移动模式,推出苏州深度游、休闲多日游等旅游产品;可与周边城市联合营销,根据多目的地旅游移动模式中枢纽城市的位置,精准地投放联合营销广告,设计环型旅行线路来推动区域旅游发展;同时,可以根据群体的频繁移动模式,将热门线路推荐给其他旅游者,减少游客出行前和出行中的决策时间,提升旅游体验。

社交媒体驱动的旅游移动模式研究存在一些局限性:1)只使用社交媒体数据,不能完整地刻画整个游客群体,存在数据偏差问题;2)社交媒体不能精准地刻画用户的完整移动过程,存在数据稀疏性问题。未来研究中,需要将社交媒体大数据与其他

数据(如信令数据等)相融合,更快速、更准确地获取不同类型游客的精准移动轨迹,以期得到更加精准的旅游移动模式。

## 参考文献

- [1] 杨喜平,方志祥.移动定位大数据视角下的人群移动模式及城市空间结构研究进展.地理科学进展,2018,37(7):880-889
- [2] Zhao X, Lu X, Liu 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. *Tourism Management*, 2018, 69: 368-383
- [3] Tideswell C, Faulkner B. Multidestination travel patterns of international visitors to Queensland. *Journal of Travel Research*, 1999, 37(4): 364-374
- [4] Hwang Y H, Fesenmaier D R. Multidestination pleasure travel patterns: empirical evidence from the American Travel Survey. *Journal of Travel Research*, 2003, 42(2): 166-171
- [5] Popp L, McCole D. Understanding tourists' itineraries in emerging rural tourism regions: the application of paper-based itinerary mapping methodology to a wine tourism region in Michigan. *Current Issues in Tourism*, 2016, 19(10): 988-1004
- [6] Mok C, Iverson T J. Expenditure-based segmentation: Taiwanese tourists to Guam. *Tourism Management*, 2000, 21(3): 299-305
- [7] Carr N. Poverty, debt, and conspicuous consumption: university students tourism experiences. *Tourism Management*, 2005, 26(5): 797-806
- [8] Goličnik B, Ward Thompson C. Emerging relationships between design and use of urban park spaces. *Landscape and Urban Planning*, 2010, 94(1): 38-53
- [9] Batra A. Senior pleasure tourists: examination of their demography, travel experience, and travel behavior upon visiting the bangkok metropolis. *International Journal of Hospitality and Tourism Administration*, 2009, 10(3): 197-212
- [10] Miah S J, Vu H Q, Gammack J, et al. A big data analytics method for tourist behaviour analysis. *Information and Management*, 54(6): 771-785
- [11] Charles Chancellor H. Applying travel pattern data to destination development and marketing decisions. *Tourism Planning and Development*, 2012, 9(3): 321-332
- [12] Li X, Meng F, Uysal M. Spatial pattern of tourist

- flows among the Asia-Pacific countries: an examination over a decade. *Asia Pacific Journal of Tourism Research*, 2008, 13(3): 229–243
- [13] Sun X, Huang Z, Peng X, et al. Building a model-based personalised recommendation approach for tourist attractions from geotagged social media data. *International Journal of Digital Earth*, 2019, 12(6): 661–678
- [14] Gavalas D, Konstantopoulos C, Mastakas K, et al. Mobile recommender systems in tourism. *Journal of Network and Computer Applications*, 2014, 39(1): 319–333
- [15] Mercer D C. The Geography of Leisure — a contemporary Growth-Point. *Geography*, 1970, 55(3): 261–273
- [16] Lue C C, Crompton J L, Fesenmaier D R. Conceptualization of multi-destination pleasure trips. *Annals of Tourism Research*, 1993, 20(2): 289–301
- [17] Stewart S I, Vogt C A. Multi-destination trip patterns. *Annals of Tourism Research*, 2002, 24(2): 458–461
- [18] Oppermann M. A model of travel itineraries. *Journal of Travel Research*, 1995, 33(4): 57–61
- [19] Flognfeldt Jr T. Traveler geographic origin and market segmentation: the multi trips destination case. *Journal of Travel & Tourism Marketing*, 1999, 8(1): 111–124
- [20] Lew A, McKercher B. Modeling tourist movements: a local destination analysis. *Annals of Tourism Research*, 2006, 33(2): 403–423
- [21] Tussyadiah I P, Kono T, Morisugi H. A model of multideestination travel: implications for marketing strategies. *Journal of Travel Research*, 2006, 44(4): 407–417
- [22] Kang S. Associations between space–time constraints and spatial patterns of travels. *Annals of Tourism Research*, 2016, 61: 127–141
- [23] Shoval N, Schvimer Y, Tamir M. Real-time measurement of tourists’ objective and subjective emotions in time and space. *Journal of Travel Research*, 2018, 57(1): 3–16
- [24] Yu Y, Li X, Qiu C. Understanding Chinese domestic travelers’ trip patterns to Mount Fanjing. *Journal of China Tourism Research*, 2017, 13(2): 141–160
- [25] Zeng B. Pattern of Chinese tourist flows in Japan: a social network analysis perspective. *Tourism Geographies*, 2018, 20(5): 810–832
- [26] Birney B A. Criteria for successful museum and zoo visits: children offer guidance. *The Museum Journal*, 2010, 31(4): 292–316
- [27] Smallwood C B, Beckley L E, Moore S A. An analysis of visitor movement patterns using travel networks in a large marine park, north-western Australia. *Tourism Management*, 2012, 33(3): 517–528
- [28] Orellana D, Bregt A K, Ligtenberg A, et al. Exploring visitor movement patterns in natural recreational areas. *Tourism Management*, 2012, 33(3): 672–682
- [29] Edwards D, Griffin T. Understanding tourists’ spatial behaviour: GPS tracking as an aid to sustainable destination management. *Journal of Sustainable Tourism*, 2013, 21(4): 580–595
- [30] De Cantis S, Ferrante M, Kahani A, et al. Cruise passengers’ behavior at the destination: investigation using GPS technology. *Tourism Management*, 2016, 52: 133–150
- [31] East D, Osborne P, Kemp S, et al. Combining GPS & survey data improves understanding of visitor behaviour. *Tourism Management*, 2017, 61: 307–320
- [32] Connell J, Page S J. Exploring the spatial patterns of car-based tourist travel in Loch Lomond and Trossachs National Park, Scotland. *Tourism Management*, 2008, 29(3): 561–580
- [33] Wu C L, Carson D. Spatial and temporal tourist dispersal analysis in multiple destination travel. *Journal of Travel Research*, 2008, 46(3): 311–317
- [34] Yang Y, Wong K K F. Spatial distribution of tourist flows to China’s cities. *Tourism Geographies*, 2013, 15(2): 338–363
- [35] Zhang J, Jensen C. Comparative advantage: explaining tourism flows. *Annals of Tourism Research*, 2007, 34(1): 223–243
- [36] Koo T T R, Wu C L, Dwyer L. Transport and regional dispersal of tourists: is travel modal substitution a source of conflict between low-fare air services and regional dispersal?. *Journal of Travel Research*, 2010, 49(1): 106–120
- [37] Yang Y, Fik T, Zhang J. Modeling sequential tourist flows: where is the next destination?. *Annals of Tourism Research*, 2013, 43: 297–320
- [38] Le-Klähn D T, Roosen J, Gerike R, et al. Factors affecting tourists’ public transport use and areas visited at destinations. *Tourism Geographies*, 2015, 17(5): 738–757



- [39] 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
- [40] Nilbe K, Ahas R, Silm S. Evaluating the travel distances of events visitors and regular visitors using mobile positioning data: the case of Estonia. *Journal of Urban Technology*, 2014, 21(2): 91–107
- [41] Raun J, Ahas R, Tiru M. Measuring tourism destinations using mobile tracking data. *Tourism Management*, 2016, 57: 202–212
- [42] Long J A, Nelson T A. A review of quantitative methods for movement data. *International Journal of Geographical Information Science*, 2013, 27(2): 292–318
- [43] Lucas J P, Luz N, Moreno M N, et al. On-Tour Interactive Travel Recommendations. *Tourism Management*, 2015, 2(4): 89–94
- [44] Narangajavana Y, Callarisa Fiol L J, Ángel Moliner Tena M, et al. The influence of social media in creating expectations. An empirical study for a tourist destination. *Annals of Tourism Research*, 2017, 65: 60–70
- [45] Wu X, Huang Z, Peng X, et al. Building a spatially-embedded network of tourism hotspots from geo-tagged social media data. *IEEE Access*, 2018, 6: 21945–21955
- [46] Leung X Y, Wang F, Wu B, et al. A social network analysis of overseas tourist movement patterns in Beijing: the impact of the Olympic games. *International Journal of Tourism Research*, 2012, 14(5): 469–484
- [47] Chung H C, Chung N, Nam Y. A social network analysis of tourist movement patterns in blogs: Korean backpackers in Europe. *Sustainability*, 2017, 9(12): 1–19
- [48] Versichele M, de Groote L, Claeys Bouuaert M, et al. Pattern mining in tourist attraction visits through association rule learning on Bluetooth tracking data: a case study of Ghent, Belgium. *Tourism Management*, 2014, 44: 67–81
- [49] Zoltan J, McKercher B. Analysing intra-destination movements and activity participation of tourists through destination card consumption. *Tourism Geographies*, 2015, 17(1): 19–35
- [50] Zheng W, Huang X, Li Y. Understanding the tourist mobility using GPS: where is the next place?. *Tourism Management*, 2017, 59: 267–280
- [51] Asero V, Gozzo S, Tomaselli V. Building tourism networks through tourist mobility. *Journal of Travel Research*, 2016, 55(6): 751–763
- [52] Shih H Y. Network characteristics of drive tourism destinations: an application of network analysis in tourism. *Tourism Management*, 2006, 27(5): 1029–1039
- [53] Liu Q, Wang Z, Ye X. Comparing mobility patterns between residents and visitors using geo-tagged social media data. *Transactions in GIS*, 2018, 22(6): 1372–1389
- [54] Lozano S, Gutiérrez E. A complex network analysis of global tourism flows. *International Journal of Tourism Research*, 2018, 20(5): 588–604
- [55] Hu F, Li Z, Yang C, et al. A graph-based approach to detecting tourist movement patterns using social media data. *Cartography and Geographic Information Science*, 2019, 46(4): 368–382
- [56] 吴必虎, 俞曦. 旅游规划原理. 2版. 北京: 中国旅游出版社, 2010
- [57] 吴必虎, 宋子千. 旅游学概论. 北京: 中国人民大学出版社, 2009: 26–33
- [58] Lew A A, McKercher B. Trip destinations, gateways and itineraries: the example of Hongkong. *Tourism Management*, 2002, 23(6): 609–621
- [59] Scuderi R, Dalle Nogare C. Mapping tourist consumption behaviour from destination card data: what do sequences of activities reveal?. *International Journal of Tourism Research*, 2018, 20(5): 554–565
- [60] Lue C C, Crompton J L, Stewart W P. Evidence of cumulative attraction in multideestination recreational trip decisions. *Journal of Travel Research*, 1996, 35(1): 41–49
- [61] Xue L, Zhang Y. The effect of distance on tourist behavior: a study based on social media data. *Annals of Tourism Research*, 2020, 82(2): 102916