

体时代,记忆工作法在技术上获得了更为有效的平台工具。传播学和营销学领域中公众意识与个体意识的关系体现在很多方面,例如,个体记忆和经历如何塑造目的地形象<sup>⑥</sup>,旅游者对景点或酒店的在线点评如何整合形成大众对此的印象<sup>⑦</sup>,基于个体游客的旅游攻略如何形成该旅游产品的普遍感知,等等。通过大数据分析的方法,个体旅游者对某个旅游目的地、景点、酒店、路线的个体记忆能够整合为大众对此的集体记忆。

在旅游行为研究领域,记忆工作法同样能够起到较大的作用,例如,公众动机如何通过个体动机的整合来达成,如何通过旅游过程中每一天的个体体验最终形成整体的公众体验等。

(第一作者系复旦大学旅游学系副教授;第二作者系复旦大学旅游系硕士研究生)

## 微博数据在旅游研究中的应用及展望\*

李 渊 付 航(厦门大学建筑与土木工程学院城市规划系 福建  
厦门 361000)

### 引言

“大数据”时代的到来以及各类社交软件的广泛应用使得微博数据的应用探索受到旅游学科学者的关注。

对于旅游目的地官方微博而言,对微博运营效果的评估,有助于检验旅游部门官方微博的运营情况。对微博传播过程中的3个重要步骤——微博的发布、微博的传播、微博的接收的相关研究,有利于指导旅游目的地官方微博更好地传播旅游信息,进一步提高旅游目的地官方微博的运营效果。

旅游者自愿发布的微博签到数据完整地记录了包括旅游者活动的地理信息、

---

<sup>⑥</sup> 王纯阳,黄福才.基于SEM的旅游目的地形象影响因素研究——以张家界为例[J].经济管理,2010(3):92~100.

<sup>⑦</sup> 高宝俊,孙含琳,王寒凝.在线评论对酒店订满率的影响研究[J].旅游学刊,2016(4):109~117.

\* 基金项目:国家自然科学基金面上项目(项目编号:41671141);福建省自然科学基金面上项目(项目编号:2015J01226);中央高校基金项目(项目编号:2012121033)。

时间信息、文本信息等相关内容，其时空活动数据揭示了旅游者时空间分布状况，按照时间序列进一步构建旅游者移动轨迹能够得到大样本量的旅游者活动数据库。

对于旅游者发布微博时附带的文本信息，通过自然语言处理技术能够分析旅游者感知以及态度，帮助旅游目的地获得旅游者反馈，据此改善旅游者的旅游体验。

目前，关于微博数据应用的综述文章主要是 Dong Li 等人提出的基于微博等众源数据对于城市规划的研究框架<sup>①</sup>，而有关微博数据在旅游研究中的应用的相关综述文章较为缺乏。

本文首先介绍微博数据的获取方式，然后从微博的传播、基于微博数据的旅游者时空间行为、微博文本语义分析 3 个方面进行相关文献的梳理，最后，提出微博签到数据在旅游研究应用中的发展方向。

## 一、微博数据的采集

随着新浪微博的快速发展以及广泛使用，对其研究也逐步深入。显然，微博基础数据的获取对研究具有非常重要的意义。根据新浪授权标准以及最新的微博加密方式，目前主要有 3 种微博数据采集方式：（1）基于微博 API 接口获取微博数据，此方法适用于需要特定条件下的数据且数据量不大时，在需要数据量较大和长时间进行数据获取的情况下，稳定性和数据完整性会受到影响；（2）基于网络爬虫进行微博数据抓取，克服了 API 接口可能出现的波动以及数据完整性降低的缺点，但在只需要部分特定数据时，后期处理工作较为繁琐；（3）通过网页采集器针对微博编写适当的采集规则进而实现对数据的获取，采集量相对平稳，但由于是根据规则获取数据，数据完整性会受到影响<sup>②③</sup>。3 种数据采集方案都能有效地对数据进行采集，在实际研究中，需要针对所需数据的类型，采用不同的采集方案。

---

① Li D, Long Y. A crowd-sourced data based analytical framework for urban planning [J]. *China City Planning Review*, 2015, 24 (1) : 49~57.

② 卢体广, 刘新, 刘任任. 微博数据通用抓取算法 [J]. *计算机工程*, 2014 (5) : 12~16.

③ 徐雁飞, 刘渊, 吴文鹏. 社交网络数据采集技术研究与应用 [J]. *计算机科学*, 2017 (1) : 277~282.

## 二、微博的传播

微博拓宽了旅游信息的传播渠道,为旅游目的地宣传旅游信息提供了全新的平台,相关学者对微博传播的过程进行了研究。

就微博的发布而言,刘大均等人研究了旅游目的地官方微博的空间分布特征,运用位序—规模法则、多元回归等方法,发现了国内旅游目的地官方微博的空间分布格局和以黑河—腾冲线为界的明显的梯度分布特征,并详细分析了其影响因素<sup>④</sup>。李欣欣等人研究了对旅游目的地官方微博运营效果的评估方法,通过对旅游局官方微博特征数据进行主成分分析,得出了影响旅游局官方微博营销效果的3种主成分——覆盖度、影响度、活跃度及其与微博相关指标的量化关系<sup>⑤</sup>。静恩明等人对旅游目的地的官方微博的受关注度进行了研究,选取微博上有关该旅游目的地的网络游记数量代表该目的地在网络上的受关注度,采用核密度分析法和均值比率方法分析旅游目的地的网络关注度空间格局特征,发现了明显的空间分异与极化特征<sup>⑥</sup>。

就微博的传播而言,微博转发是微博传播的主要方式。唐佳等人研究了影响微博传播的因素,以指导旅游目的地官方微博发布旅游信息的策略,分析了影响微博转发的特征变量,利用多分 Logistic 回归方法,发现了微博活动信息、粉丝数和图片是3个对微博转发水平影响最大的因素<sup>⑦</sup>。周芳如等人研究了用户转发微博形成的网络信息流空间,采用社会网络分析法和地图法,对旅游信息在网络信息流空间中传播的时空特征进行了分析<sup>⑧</sup>。

就微博的接收而言,旅游者是旅游信息的接收者,是微博传播过程的最后一个环节。段淼然对旅游目的地官方微博粉丝的分布情况进行了分析,以确定旅游目的地的客源市场状况;他通过分析微博粉丝来源地分布,发现来源省份

---

④ 刘大均,胡静,程绍文,等. 中国旅游微博空间分布格局及影响因素——以新浪旅游微博为例[J]. 地理科学, 2015(6): 717~724.

⑤ 李欣欣,翟晓鹏. 基于特征数据的旅游局官方微博营销效果研究[J]. 武汉理工大学学报(信息与管理工程版), 2014(2): 251~254.

⑥ 静恩明,郭风华,李仁杰,等. 基于新浪旅游博客的河北省A级景区网络关注度研究[J]. 地理与地理信息科学, 2015(3): 118~122.

⑦ 唐佳,李君轶. 基于多分 Logistic 回归的旅游局官博转发影响因素研究[J]. 旅游学刊, 2015(1): 32~41.

⑧ 周芳如,吴姗姗,吴晋峰,等. 旅游信息微博传播网络结构及时空特征[J]. 经济地理, 2016(6): 195~203.

和景点微博关系构成的矩阵所代表的旅游者外出旅游能力的出游驱动有向流，并分析旅游流空间分布特征，探索各区域之间的旅游联系<sup>⑨</sup>。

总体而言，微博作为旅游目的地传播旅游信息的重要媒介，对其传播过程的研究为旅游目的地评估自身官方微博运营情况、结合自身旅游资源、明确市场营销对象与方向、制定旅游规划提供了理论支持。对微博传播的研究一方面可以通过探索促进微博信息传播的因素来改善旅游信息的传达；另一方面可以从信息传播形成的流空间的角度研究分析微博传播的特征，帮助旅游目的地确定客源市场，进一步确定市场营销的方向。同样，对微博粉丝分布情况的研究也有助于分析客源分布特征，从而针对性地开展宣传活动。

### 三、基于微博签到数据的时空间行为

微博签到数据包含大量的人群分布和人群移动的时空间数据。基于微博数据的时空间行为研究是传统时空间行为研究的重要发展，以传统调查难以采集到的大样本量、长时间、全天候的数据丰富了对个体时空间行为和旅游目的地内部空间之间关系的理解。

关于人群分布特征，张子昂等人研究了景区内部的旅游者时空分布，运用空间核密度分析法，并通过性别特征和来源地归属进行分类讨论<sup>⑩</sup>。陈宏飞等人研究了旅游者到达公园的时空分布特征，以城市公园为研究对象，探讨了不同空间分布的旅游者到达公园的时空分布规律和不同空间分布的旅游者在不同时段到达公园游览的比例<sup>⑪</sup>。

关于人群移动，唐佳等人研究了不同人群移动的时空规律，研究旅游者的时间分布模式，并与当地居民比较，发现其与居民的活动规律的差异以及客流的时间分布模式与空间特征的较强对应关系<sup>⑫</sup>。李君轶等人研究了人群移动形成的旅游流空间，探讨了旅游者的时空共现和旅游流空间特征，并提出基于社

---

<sup>⑨</sup> 段森然，陈刚，于靖，等. 基于新浪微博的省域出游驱动力空间分布特征[J]. 地域研究与开发，2015（2）：96~102.

<sup>⑩</sup> 张子昂，黄震方，靳诚，等. 基于微博签到数据的景区旅游活动时空行为特征研究——以南京钟山风景名胜区的为例[J]. 地理与地理信息科学，2015（4）：121~126.

<sup>⑪</sup> 陈宏飞，张玥，刘广，等. 基于微博的城市公园游客来源时空分布研究[J]. 西北大学学报（自然科学版），2016（2）：292~297.

<sup>⑫</sup> 唐佳，李君轶. 基于微博大数据的西安国内游客日内时间分布模式研究[J]. 人文地理，2016（3）：151~160.

会感知计算、运用大数据进行的旅游者行为研究框架<sup>⑬</sup>。

总体上,基于微博的时空间行为研究主要对人群分布和人群移动两个方面进行了研究,是传统的GPS调查、活动日志调查等时空间行为调查方式的重要补充。目前,其研究方法与传统研究类似,但是由于微博签到具有获取方式简单、成本低的优势,能够在城市尺度,甚至更大的省域、全国尺度进行研究,有效弥补了传统时空行为调研的研究范围以及尺度的局限。已有研究在宏观的省域、全国尺度以及微观的景区内部尺度相对缺乏,有待进一步加强。

#### 四、微博文本语义分析

旅游者发布微博时经常会附上一定的文本信息,通过自然语言处理技术能够对这些文本信息进行挖掘。张思豆等人通过判断微博文本的情感倾向,研究了旅游者文本信息的情感倾向值与空气质量的相关关系,发现了PM<sub>2.5</sub>、SO<sub>2</sub>等污染物与旅游者情感的负相关关系<sup>⑭</sup>。

类似的对于网络文本的语义分析研究主要为基于博客、游记、点评等旅游者发布的信息文本对旅游行为的研究。其中,有基于旅游博客,结合内容分析与共现网络方法,研究旅游目的地形象认知特征的研究<sup>⑮</sup>;有基于旅游网站游记分析旅游目的地感知形象和意象的研究<sup>⑯</sup>;有根据网站在线评论和酒店位置评分等因素分析不同档次酒店订满率的研究<sup>⑰</sup>。

将网络文本的研究扩展到微博文本时要注意微博文本简短、格式不严谨、无效信息较多的特点,分析其适用的研究方法和研究范围。随着自然语言处理技术的发展,通过语义分析对微博文本进行研究,对于舆情控制、市场营销、旅游者体验反馈等都有重要意义。

---

⑬ 李君轶,唐佳,冯娜. 基于社会感知计算的游客时空行为研究[J]. 地理科学, 2015(7): 814~821.

⑭ 张思豆,李君轶. 基于微博大数据的游客情感与空气质量关系研究——以西安市为例[J]. 陕西师范大学学报(自然科学版), 2016(4): 102~107.

⑮ 郭风华,王琨,张建立,等. 成都“五朵金花”乡村旅游地形象认知——基于博客游记文本的分析[J]. 旅游学刊, 2015(4): 84~94.

⑯ 徐小波,赵磊,刘滨谊,等. 中国旅游城市形象感知特征与分异[J]. 地理研究, 2015(7): 1367~1379.

⑰ 高宝俊,孙含琳,王寒凝. 在线评论对酒店订满率的影响研究[J]. 旅游学刊, 2016(4): 109~117.

## 五、国外微博类数据的研究

国外有 Facebook、Twitter 等众多社交应用平台，其数据特征与微博有一定相似性。

其中，Facebook 作为社交网络服务平台，重点体现出用户之间的社交关系，主要被用来作为旅游目的地分析市场的工具。Ji Hoon Park 等人研究了韩国政府如何通过 Facebook 进行旅游目的地的提升与市场营销<sup>⑮</sup>；Mariani 研究了意大利的目的地管理组织（Destination Management Organisations）如何将 Facebook 运用到旅游目的地的提升和市场营销中<sup>⑯</sup>；Alvarez 等人基于 Facebook 研究了社交网络作为参与社区建设的一种潜在工具其背后的运作机制<sup>⑰</sup>。

Twitter 作为重要的社交网络服务平台，用户会分享大量的短文本和带有地理信息的图片内容。Padawangi 等人基于 Twitter 含有地理信息的图片数据，研究了 Twitter 地理信息在应对印度尼西亚雅加达城市洪水灾害中的作用<sup>⑱</sup>；Alvin Chua 基于 Twitter 带有地理信息的图片分析了意大利南部的旅游流<sup>⑲</sup>。

此外，还有其他的一些社交应用。Flickr 也是重要的图片分享平台。Donaire 通过研究旅游者上传 Flickr 的照片，分析旅游者关注的元素，并对旅游者进行聚类分析<sup>⑳</sup>；Trip Advisor 是全球旅游评论网站，Ganzaroli 基于 Trip Advisor 的在线评价数据，研究了评论内容对于酒店品质的影响<sup>㉑</sup>。

---

⑮ Park J H, Lee C, Yoo C, et al. An analysis of the utilization of Facebook by local Korean governments for tourism development and the network of smart tourism ecosystem [J]. *International Journal of Information Management*, 2016, 36 (6) : 1320~1327.

⑯ Mariani M M, Di Felice M, Mura M. Facebook as a destination marketing tool: Evidence from Italian regional destination management organizations [J]. *Tourism Management*, 2016 (54) : 321~343.

⑰ Alvarez L, Borsi K, Rodrigues L. The role of social network analysis on participation and placemaking [J]. *Sustainable Cities and Society*, 2017 (28) : 118~126.

⑱ Padawangi R, Turpin E, Herlily, et al. Mapping an alternative community river: The case of the Ciliwung [J]. *Sustainable Cities and Society*, 2016 (20) : 147~157.

⑲ Chua A, Servillo L, Marcheggiani E, et al. Mapping Cilento: Using geotagged social media data to characterize tourist flows in southern Italy [J]. *Tourism Management*, 2016 (57) : 295~310.

⑳ Donaire J A, Camprubí R, Galí N. Tourist clusters from Flickr travel photography [J]. *Tourism Management Perspectives*, 2014 (11) : 26~33.

㉑ Ganzaroli A, De Noni I, van Baalen P. Vicious advice: Analyzing the impact of trip advisor on the quality of restaurants as part of the cultural heritage of Venice [J]. *Tourism Management*, 2017 (61) : 501~510.

微博数据与 Facebook、Twitter、Flickr 等数据有一定的相似性, 由于微博在中国的广泛应用, 在国内旅游研究大数据中发挥着重要作用, 可以将国外学者在微博类数据研究方面的探索有效地借鉴到微博数据的研究上, 并针对中国不同特点的旅游目的地开展不同类型的研究。

## 六、结论和讨论

对于微博传播的研究主要是针对旅游信息发布主体, 即旅游目的地官方微博的运营情况评估的研究, 以及微博传播的特征和作为微博接收者的旅游者来源分布的研究。主要目的在于指导旅游目的地更好地进行旅游信息的宣传, 并分析客源市场, 进行下一步的市场营销活动。

对于基于微博签到数据的时空行为的研究, 其研究尺度主要为中观的城市尺度, 缺乏宏观的全国、省域尺度以及微观的景区、街区内部尺度, 没有充分体现微博数据时空信息丰富、样本量大、采集较简单、成本低的优势, 有待进一步在传统时空行为研究领域中较难开展的宏观及微观尺度进行更加深入的研究。研究问题也应从目前的揭示旅游者活动规律进一步深入到与具体物质空间的互动关系, 并结合景区规划管理面临的实际问题进行研究。

对于微博文本的语义分析还处于探索阶段, 不同于现在已有的针对博客长文本、网络游记、在线评论等网络文本的大量研究, 微博文本具有文字简短、表达直接的优点, 同时也有文本格式较为自由、无效信息较多的缺点, 需要预先进行数据提炼及筛选工作, 可能导致提炼后的数据有效信息较少。所以, 需要区别对待微博文本与其他大数据文本, 结合其他网络文本的研究方法, 针对其特性对文本价值进行充分挖掘。

目前, 尽管微博签到数据在旅游研究中已经取得了一定进展, 但对于微博数据的适用范围有待探索, 对于微博数据的认识理解有待深化。在微博传播上, 尽管通过对微博的研究已经能够对旅游目的地官方微博的运营提供指导, 但如何理解作为旅游信息发布者的旅游目的地官方微博与作为旅游信息接收者的旅游者之间的社交关系仍然有待深入研究。在时空行为研究上, 尽管大样本、高精度的微博签到数据能对时空行为进行更精确的研究, 但如何通过理解时空行为和旅游目的地内部空间之间的关系来引导旅游目的地进行旅游者参观路线规划、旅游者疏散引导等旅游规划管理工作, 仍然需要进一步的研究将其落实到旅游实际运作中。对于微博文本, 如何针对其特性充分挖掘文本的价值有待进一步探索。同时, 国外对于微博类数据进行了相当数量的研究, 如何吸收借

鉴其有益的探索尝试，并针对微博数据的特点和国内旅游发展的特点展开相应研究，也是需要重点考虑的问题。

最后，尽管微博数据丰富了旅游研究的数据源，但是如何根据微博数据的新特点找到其最适宜的旅游研究的方向，如何全面、系统地理解这类新型数据来弥补传统的旅游研究方法的不足，以及如何结合两者来解决旅游学科的实际问题仍然是微博数据应用所面临的重要挑战。

（第一作者系厦门大学建筑与土木工程学院城市规划系副教授；第二作者系厦门大学建筑与土木工程学院城市规划系硕士研究生）



### 本刊声明

1. 本刊所发作品仅为作者观点，不代表编委会或编辑部的立场。
2. 欢迎对本刊所登文章开展学术讨论。
3. 本刊对所发作品享有中文专有出版权，请勿一稿多投。
4. 欢迎转载本刊文章，并按规定付酬及注明出处。
5. 本刊对来稿保留修改权，有特殊要求者，请事先声明。
6. 本刊对所发论文享有汇编权、翻译权、电子出版权及信息网络传播权、转让权，如有异议，请事先声明。