

# 基于游客体验的旅游活动碳足迹研究

顾畑甜, 马彪

东华大学管理学院, 上海  
Email: gttabb@126.com

收稿日期: 2020年11月16日; 录用日期: 2020年12月13日; 发布日期: 2021年2月5日

## 摘要

旅游业带动经济增长的同时, 也会产生碳排放引发环境问题。为了响应国家碳中和的目标, 需要减少旅游过程中的碳排放。相较于之前的研究以统计数据为基础, 本文以游记为基础, 提出了词性规则和依存关系相结合的抽取方法, 抽取游记中旅游活动和对应情感词, 再用旅游活动的碳足迹模型分析个人和地区的旅游活动碳排放, 为旅游业的减排工作提供决策依据。

## 关键词

旅游碳足迹, 活动 - 情感词对, 词性规则, 依存关系

# Research on Carbon Footprint of Tourism Activities Based on Tourist Experience

Tiantian Gu, Biao Ma

Department of Management, Donghua University, Shanghai  
Email: gttabb@126.com

Received: Nov. 16<sup>th</sup>, 2020; accepted: Dec. 13<sup>th</sup>, 2020; published: Feb. 5<sup>th</sup>, 2021

## Abstract

While tourism promotes economic growth, it will also produce carbon emissions and cause environmental problems. In response to the national goal of carbon neutrality, carbon emissions from tourism need to be reduced. Compared with previous studies based on statistical data, this paper proposes a method based on travel notes, which combines part of speech rules and dependency relationship, extracts tourism activities and corresponding emotional words in travel notes, and then analyzes individual and regional carbon emissions of tourism activities with the carbon

**footprint model of tourism activities, which provides decision-making basis for the emission reduction of tourism industry.**

## Keywords

**Travel Carbon Footprint, Target-Opinion Pair, Speech Rules, Dependency Parsing**

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

温室气体排放和全球变暖问题早已成为全球关注的重点。2016年《巴黎协定》提出在2051至2100年间,全球实现碳中和。2020年习近平主席在第七十五届联合国大会一般性辩论上,强调中国将采取更加有力的政策和措施,争取2060年前实现碳中和。减少碳排放,实现碳中和,已成为刻不容缓的问题。

旅游行业快速发展,拉动经济增长的同时,也会带来相应的环境压力,增加碳排放引起气候变化。旅游的过程伴随着碳排放的问题,旅途中的交通、各项旅游活动、旅游住宿等都会产生碳排放。为了响应国家早日实现碳中和的号召,研究如何减少旅游过程碳排放成为低碳旅游乃至碳中和旅游的发展目标。

Weidema [1]提出的旅游碳足迹模型包括三部分:旅游交通、旅游住宿和旅游活动碳。旅游碳足迹即旅游者在交通、住宿、餐饮、参与娱乐活动等过程中直接和间接产生二氧化碳排放量的总和。目前对碳足迹的研究主要是通过人文社会经济数据的调查进行研究[2],尤其是研究旅游活动碳足迹,往往通过问卷调查、访谈等获得一手数据资料,需要花费大量人力且覆盖面较小,抽样调查可能导致研究结果的偏差。互联网海量的游记文本中,包含了游客大量的旅游活动。因此,本文考虑以游记数据为基础,通过抽取方法获得旅游活动、研究旅游活动的碳足迹,并结合游客对旅游活动的情感倾向,为当地低碳旅游的发展提供科学依据。

## 2. 相关工作

### 2.1. 旅游活动特征的研究现状

旅游活动碳足迹模型表示,游客参与不同的旅游活动产生的碳足迹值不同。运动体验类的旅游活动比休闲度假类的活动产生的碳排放量更高。为了更准确地估算旅游活动的碳足迹,需要先挖掘游客在旅游地的旅游活动。姚治国[3]采用市场调研的方法,调研了海南省游客的出行距离、停留时间、旅游活动等信息。一些学者[2][3][4]通过查询各地区的旅游统计年鉴中的数据来估算分析旅游活动的碳足迹。在旅游领域的个性化推荐中,学者为了提高推荐的精确性,会将游客的历史旅游记录作为推荐依据。De等[5]研究游客分享在社交网络上的旅游照片,分析游客旅游活动;ARASE等[6]分析游客发布的照片以及所带的标签和标题,YIN等[7]分析游客在景点的签到数据,数据包括签到地名称、地理位置、签到时间等,构建游客活动画像。Yu等[8]基于LBSNs (Location-Based Social Network Services)的签到数据获得游客所在地点的信息,并对地点的类型细分,更加细粒度分析游客的浏览行为。吴清霞等[9]认为游客在每个POI (Point of Information)停留的时间可以反映他对此类POI的偏好程度,将旅游地点和停留时间共同作为活动特征。文献[10][11]分析了游记中的图片和文字描述,挖掘出游客的历史行为轨迹,作为路线推

荐的依据。孙文平等[12]首先基于结构化知识库构建景点信息,获得景点的平均浏览时长,再从游记中抽取旅游天数、旅游路线、游客类型等属性,利用景点位置属性计算景点间交通时间,详细刻画游客的浏览行为,提出一种有效的旅游路线推荐方案。

上述方法主要是分析游客浏览旅游景点的行为轨迹,缺少细粒度的旅游过程的研究,比如旅游活动、活动体验等。而游客在各景点也会参与不同的旅游活动,各项旅游活动的碳排放系数具有差异,分析游客的具体活动能比较准确地估算碳排放。游客在游记中的情感表达反映了游客的活动体验,侧面体现了游客对此活动的兴趣度和满意度。游客的满意度高,未来再次参加此活动以及其他游客参加此活动的概率更大,因此分析游客群体对某旅游活动的情感可以用于预测此活动未来的走势。综上所述,本文将从旅游活动和活动情感的角度,细粒度分析游客的旅游行为,估算旅游活动中的碳排放情况。

## 2.2. 活动特征 - 情感词对抽取方法的研究现状

用户的情感表达往往是针对某个对象的。对长文本仅做篇章级的情感分析,只能粗粒度地了解用户整体的情感,不能得知用户情感抒发的对象,深入挖掘用户的兴趣点。细粒度的情感挖掘,可以将文本中的情感词和它对应的评价对象(特征)作为词对抽取出来,更加高效,准确地挖掘用户感兴趣的点。文献[13]采用关联规则将名词或名词短语作为特征,将距离特征词最近的形容词作为情感词抽取,此类方法没有考虑特征词和情感词的长距离搭配。廖祥文等[14]采用词对齐模型抽取特征 - 情感词对,再考虑特征和情感词的依存关系,构建多层关系图模型,利用随机游走算法筛选出置信度高的词对。王素格等[15]利用词对间的依存句法关系,构建特征和情感词的组块规则,设计候选特征的识别方法,抽取出特征 - 情感词对。张凌等[16]先以情感词为基准搜索带标注的情感对象,通过计算先验概率进行词性剪枝和模板筛选,提取原子情感对象。运用基于斯坦福句法树的依存规则,将原子情感对象扩展为完整的情感要素。张建华等[17]构建名词短语的词性规则抽取模板,再结合八种依存关系抽取,结合两种方法的优点,在数据集上表现较好。邱云飞等[18]根据中文词性特征构建规则提取评价对象,结合句法结构树提取并筛选评价对象,基于核心句法路径筛选评价搭配,得到最终的评价对象和评价词对。江腾蛟等[19]针对金融评论评价对象构成复杂,情感词多为动词的特点,提出结合了语义角色标注和依存句法分析的抽取规则,并对虚指评价对象和隐式评价对象进行识别和替换。

对当前的特征 - 情感词对抽取方法的梳理中发现,现有的方法仅考虑了特征是名词或名词短语的情况。这种抽取方式适用于较短的评论文本,因为评论文本的评价对象以名词为主。而针对较长的游记,此类抽取方法会缺失特征是动词或动词短语的情况,导致语义不完备。游记中游客的情感表达并不仅仅是对名词事物,也会针对旅途中的行为活动。因此本文针对游记等长文本,提出特征词以动词、动词短语为主的词对抽取方法,获取游客旅游活动和对应的活动情感词对。

## 3. 基于游客体验的旅游活动碳足迹模型

本节主要介绍了旅游活动 - 情感词的抽取方法和旅游活动碳足迹模型。首先采用本文词对抽取方法获得旅游活动 - 情感词的词对,然后改进旅游活动碳足迹模型,加入情感倾向系数,得到本文的基于游客体验的旅游活动碳足迹模型。

### 3.1. 旅游活动和情感词的抽取方法

抽取旅游活动及其对应情感词,以词对的形式展示。首先基于词性规则进行提取,再结合依存关系的规则,弥补单纯运用词性规则存在的抽取遗漏和不完整的问题,最后对两种方法获得的词对进行筛选过滤,最终实现旅游活动情感特征的抽取。

### 3.1.1. 活动特征和情感词的词性分析

在前人的许多研究中,情感的承载者被称为评价对象。评价对象的词性被限定为名词、名词短语或从句,适用于大部分的评论文本。本文的研究对象是旅游文本,其中包含了大量游客活动行为的描述,并且这些活动是有价值的,包含了游客的内在偏好。与传统的评价对象的词性不同,本文的评价对象以游客活动特征为主,当词性为动词、动词性短语时,是游客的显性行为。如“吃甜品让人幸福”,“吃甜品”这个动词性短语是游客的一个活动特征,承载了他的情感“幸福”。

对大量的游记分析发现,游客的活动特征并不是全都以动词和动词短语的形式来展现的。一些游客对行为的描述较少,对事物、景物的描写较多,描述对象是名词或名词短语。本文认为对于事物的描写是一种隐性信息,能够据此推断出游客的行为活动。“烤鸭的味道很地道”中虽未出现动词,但可以从“味道”一词推断该游客吃饭的行为。因此将抽取名词和名词短语作为隐性的活动特征。综上所述,本文定义抽取的活动特征的词性为动词、动词短语和名词、名词短语。

王治敏等[20]对汉语情感词研究发现,情感词的词性一般是形容词、动词、名词、副词等。在游记文本中,情感词的词性也覆盖了上述词性。不同词性的情感词在句子中充当不同的语法成分。

1) 形容词是情感词最常见的词性,当前许多商品评论的情感词对的抽取方法也是以抽取形容词情感词为主。“吹着凉风也挺惬意”,“惬意”作为形容词情感词修饰了“吹着凉风”这个行为。

2) 动词情感词,一般在句子中充当谓语。“我喜欢上海滩的夜景”,“喜欢”作为心理动词情感词,修饰的对象是“上海滩的夜景”。

3) 名词情感词,在游记文本中一般充当句子的谓语、定语。“半山温泉给我们的感觉就像人间仙境”,这里“人间仙境”用作修饰“半山温泉”的情感词。

4) 副词情感词,一般在句中修饰形容词和动词,在句子中充当状语和补语,对情感强度有影响。这里不单独将副词情感词作为修饰词。

### 3.1.2. 基于词性规则抽取活动情感特征

对活动特征和情感词的词性分析可知,活动特征分为显性活动特征和隐性活动特征。显性特征的词性,以动词和动词短语为主。隐性特征的词性以名词和名词短语为主。情感词的词性以形容词、动词、名词为主。基于词性构建规则模板,在旅游领域的适用性较强,能够较为准确的提取词对。抽取范围,以逗号为分解符,作为一个分句。在每个分句中抽取游记中的行为特征和相应情感词,构成词对(活动特征,情感词)。基于哈尔滨工业大学的语言技术平台(Language Technology Platform, LTP) [21]的词性标注,对大量的游记文本进行分析总结,分别制定了显性和隐性活动特征的词性规则。词性的缩写含义, n (名词)、v (动词)、a (形容词)、d (副词)、u (助词)

#### 1) 活动特征为动词短语

传统的词性模板未考虑动词短语,本文分析游记后发现,动词短语时通常包括两种形式, v + n 和 v + v 的短语结构。如果仅抽取动词 v 会丢失宾语,造成本语意缺失。例如“欣赏风景”是一个完整的动词短语,仅考虑动词“欣赏”时并不可知其所指的对象。因此需要考虑 v + n 和 v + v 两种结构的特征词。情感词应用的场景,包括:

#### ① 形容词情感词

a) 充当谓语时,通常是“特征词 + 副词 + 情感词”的形式;

b) 充当状语时,通常是“情感词 + 地 + 特征词”的形式。

#### ② 动词情感词

活动特征是动词短语时,一般充当谓语,是“副词 + 情感词 + 特征词”的形式。

## ③ 名词情感词

当特征词是动词短语时, 名词情感词较少, 这里不予考虑。

基于上述分析, 当特征词为动词短语时构建模板如表 1。

**Table 1.** Extracting part of speech templates of verb phrases

**表 1.** 提取动词短语的词性模板

序号	规则	例句	抽取词对
1	(v + n) + d + a	欣赏/v 风景/n 也/d 不错/a	(欣赏风景, 不错)
2	(v + v) + d + a	逛街/v 购物/v 非常/d 累/a	(逛街购物, 累)
3	a + u + (v + n)	悠闲/a 地/u 欣赏/v 风景/n	(欣赏风景, 悠闲)
4	a + u + (v + n)	欢乐/a 地/u 去/v 吃饭/n	(去吃饭, 欢乐)
5	d + v + (v + n)	非常/d 喜欢/v 喝/v 奶茶/n	(喝奶茶, 喜欢)
6	d + v + (v + v)	很/d 想/v 体验/v 骑马/v	(体验骑马, 想)

## 2) 活动特征为动词

当特征词是一个单独的动词时, 抽取时仅需考虑动词 v, 其情感词应用场景与特征词是动词短语时相似, 可以构建模板如表 2。

**Table 2.** Extracting part of speech templates of verb

**表 2.** 提取动词的词性模板

序号	规则	例句	抽取词对
1	v + d + a	爬山/v 也/d 悠闲/a	(爬山, 悠闲)
2	v + a	服务/v 贴心/a	(服务, 贴心)
3	a + u + v	刺激/a 地/u 漂流/v	(漂流, 刺激)
4	d + v + v	最/d 期待/v 射箭/v	(射箭, 期待)

## 3) 特征词为名词

当特征词是一个单独的名词时, 抽取对象为名词 n。情感词的应用场景包括:

## ① 形容词情感词

a) 充当谓语时, 通常是“特征词 + 副词 + 情感词”或“特征词 + 情感词”的形式;

b) 充当定语时, 通常是“情感词 + 的 + 特征词”的形式。

## ② 动词情感词

a) 充当谓语时, 通常是“副词 + 情感词 + 特征词”的形式;

b) 充当定语时, 通常是“情感词 + 的 + 特征词”的形式。

## ③ 名词情感词

当特征词是名词时, 一般充当定语, 是“情感词 + 的 + 特征词”的形式。

基于上述分析, 当特征词为名词时构建模板如表 3。

### 3.1.3. 基于依存关系抽取活动情感特征

基于词性规则抽取特征和情感词对, 对于短文本的效果较好。在以标点为分隔符的分句中, 情感词和评价对象距离较近, 运用词性模板抽取的准确率较高, 但仍存在一些问题。第一, 当句子结构较复杂时, 词性模板适用性较差。第二, 词性模板的抽取结果中, 可能存在特征和情感词抽取不完整的问题。

因此, 将进一步根据依存关系, 制定句法规则抽取词对。

**Table 3.** Extracting part of speech templates of nouns  
**表 3.** 提取名词的词性模板

序号	规则	例句	抽取词对
1	n + d + a	食材/n 颇为/d 新鲜/a	(食材, 新鲜)
2	n + a	服务员/n 态度/n 恶劣/a	(态度, 恶劣)
3	a + u + n	九寨沟/n 最大/a 的/u 村寨/n	(村寨, 最大)
4	d + v + n	非常/d 喜欢/v 奶茶/n	(奶茶, 喜欢)
5	v + u + n	喜爱/v 的/u 小吃/n	(小吃, 喜爱)
6	n + u + n	风情/n 的/u 小镇/n	(小镇, 风情)

句法分析是用来确定句子的结构以及句子中词与词之间的依存关系的。句子中的词语如果存在依存关系, 就构成一个依存关系对。每个依存关系对中, 一个词语处于支配地位, 称为核心词(支配词), 另一个被支配的词语称为修饰词(从属词)。一般从核心词到修饰词用一个有向弧表示依存关系。在依存弧上方标注关系类型。根据哈尔滨工业大学的语言技术平台(Language Technology Platform, LTP)定义的依存关系类型, 可以分为 24 类。常见的依存句法关系标注类型, 如表 4 所示。

**Table 4.** Common meaning of dependency parsing  
**表 4.** 常见依存关系及含义

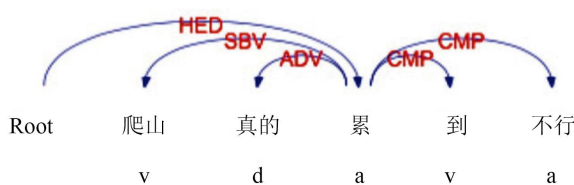
标注	含义	标注	含义
SBV	主谓关系	VOB	动宾关系
ATT	定中关系	ADV	状中结构
CMP	动补结构	COO	并列结构
FOB	前置宾语	POB	介宾关系
WP	标点	HED	核心

**(一) 补充抽取词对**

在游记中句子的结构比较复杂, 特征词和情感词虽在同一分句中, 但是距离较远, 词性规则不能很好的覆盖到所有情况。而特征词和情感词之间存在直接或间接的依存关系, 通过分析依存关系, 可以进一步补充抽取特征 - 情感词对。抽取范围, 以逗号为分解符, 作为一个分句。在每个分句中抽取游记中的行为特征和相应情感词, 构成词对(活动特征, 情感词)。

**规则 1:** 分句中存在 SBV 结构时, 且核心词是形容词, 特征 - 情感词对可表示为(SBV 结构的修饰词, SBV 结构的核心词)。

例: 如图 1 所示, “爬山真的累到不行”, “爬山 - 累” 构成 SBV 结构, 可抽取特征 - 情感词对(爬山, 累)。



**Figure 1.** Dependency parsing of SBV structure  
**图 1.** SBV 结构的依存句法分析

**规则 2:** 分句中存在 VOB 结构时, 当情感词是动词, 修饰的宾语是特征词, 特征 - 情感词对表示为 (VOB 结构的修饰词, VOB 结构的核心词)。

例: 如图 2 所示, “我喜欢上海滩的夜景”这句话中, 有 VOB 结构“喜欢 - 夜景”, “喜欢”是一个动词情感词, 其修饰的对象是“夜景”, 特征 - 情感词对表示为(夜景, 喜欢)。

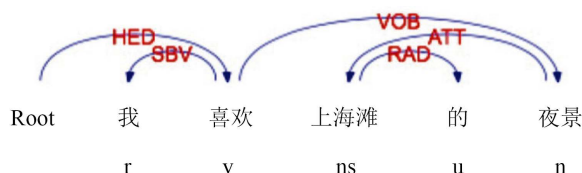


Figure 2. Dependency parsing of VOB structure  
图 2. VOB 结构的依存句法分析

## (二) 识别完整的特征词和情感词

在抽取特征词的时候, 抽取的是情感词直接修饰的特征词, 而句子中经常在特征词前面会有多个修饰词来描述特征词, 使特征词的描述更加完整。前文中基于词性规则或者依存关系的抽取中, 只抽取了部分特征词, 会导致语义不完整, 因此需要识别完整的特征词。

**规则 3:** 当特征词有定语修饰时, 即有定中关系 ATT 时, 将 ATT 的修饰词作为特征的一部分, 得到完整的特征词。

例: 如图 3 所示, 上文中的例句“我喜欢上海滩的夜景”, 根据规则 1 获得词对(夜景, 喜欢), 特征词“夜景”存在定语修饰“上海滩”, 可以扩展得到完整的特征词, 最终特征 - 情感词对抽取为(上海滩夜景, 喜欢)。通过规则 3 可以扩展名词, 得到的特征词是名词短语。

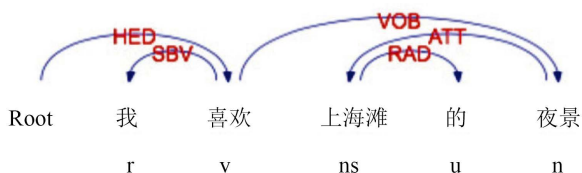


Figure 3. Dependency parsing of ATT structure  
图 3. ATT 结构的依存句法分析

部分游记中, 作者会对多个不同特征词赋予相同的感情, 此时多个特征词之间存在并列关系, 组成多个特征 - 情感词对。

**规则 4:** 当特征词存在并列结构 COO 时, 抽取为多个特征 - 情感词对。

例: 如图 4 所示, “坐船、放灯都很有趣”, 根据规则 1 可以获得初步的词对(坐船, 有趣), 特征词“坐船”存在并列结构“放灯”, 可以得到 2 个特征情感词对(坐船, 有趣)和(放灯, 有趣)。

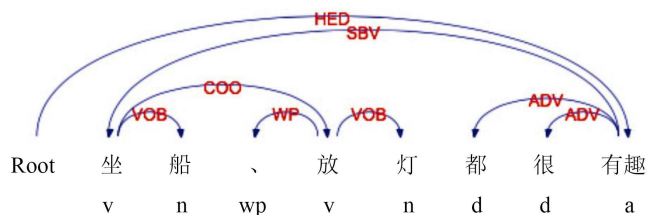


Figure 4. Dependency parsing of COO structure with feature words  
图 4. 特征词是 COO 结构的依存句法分析

不仅特征词可能会有多个并列, 作者抒发情感时, 情感也是错综复杂的, 会出现对同一特征词多种情感交错, 使用多个情感词。此时情感词之间存在并列关系, 可以组成多个特征 - 情感词对。

**规则 5:** 当情感词存在并列结构 COO 时, 抽取为多个特征 - 情感词对。

例: 如图 5 所示, “清晨的西湖格外宁静、朦胧和凉爽”, 根据规则 1 可以获得初步的词对(西湖, 宁静), 情感词“宁静”存在并列结构“朦胧”和“凉爽”, 可以得到 3 个特征情感词对(西湖, 宁静), (西湖, 朦胧)和(西湖, 凉爽)。

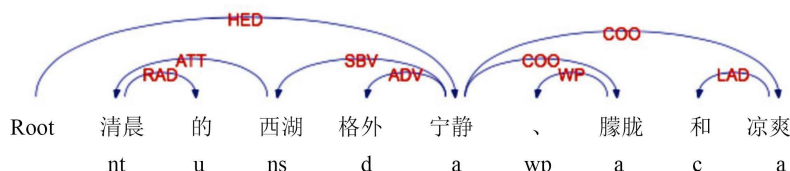


Figure 5. Dependency parsing of COO structure with affective words  
图 5. 情感词是 COO 结构的依存句法分析

### 3.1.4. 过滤筛选词对

鉴于词性模板中, 针对情感词是名词和动词的规则会引用一些噪音数据, 出现虽然符合规则但名词和动词并非情感词的情况, 因此需要对情感词是名词和动词的词对进行筛选。本文基于大连理工情感词汇本体构建了旅游领域的情感词典。考虑到该词典中情感词的覆盖面较广, 应用于旅游领域时仍会有噪音数据需要剔除。因此, 选取其中部分情感词作为种子词, 采用 PMI 算法[22]进一步拓展了适用于旅游领域的情感词, 包括 346 个名词情感词和 389 个动词情感词。通过词性规则和依存关系初步抽取的词对结果中, 如果情感词是名词或动词, 需要判断该情感词是否存在于情感词典中, 若情感词典中包含该词则加入最终词对集, 否则过滤该词对。在得到最终词对集后, 如果词对集中存在特征词和情感词都相同的词对时, 删除重复词对。

## 3.2. 旅游活动碳足迹模型

旅游碳足迹是指旅游者在交通、住宿、餐饮、各类娱乐活动等过程中直接和间接产生 CO<sub>2</sub> 排放量的总和。按照“自上而下”的分析方法, 旅游过程产生的碳足迹可以分解为旅游交通碳足迹、旅游住宿碳足迹和旅游活动碳足迹。游客参与不同的旅游活动会产生不同大小的碳排放。旅游活动的碳足迹模型是用于估算不同旅游活动中产生的碳排放量。

### 1) 传统旅游活动碳足迹模型

传统的旅游活动碳足迹模型一般用于估算某一地区的旅游活动碳排放量[4], 计算公式为:

$$CA = \sum_j (N_t \times W_j \times P_j) \tag{1}$$

其中 CA 为旅游活动产生的碳排放量,  $N_t$  为  $t$  时期的游客总数,  $W_j$  代表第  $j$  类出行目的的游客比例,  $P_j$  代表第  $j$  类旅游活动的单位碳排放系数(千克/人)。

### 2) 基于游客体验的个人旅游活动碳足迹模型

个人旅游活动碳足迹模型表示游客个人在旅途中产生的碳排放。可以从个体出发, 引导每位游客减少旅游碳排放, 总体达到节能减排的目的。本文在传统的旅游活动碳足迹模型[4]中加入情感倾向系数  $S$ , 计算公式为:

$$CP = \sum_j (M_t \times A_j \times P_j \times S_j) \tag{2}$$



其中  $CP$  为某位游客的旅游活动碳排放量,  $M_t$  代表  $t$  时期  $g$  该游客的旅游活动次数,  $A_j$  代表游客参加第  $j$  类旅游活动的占比,  $P_j$  代表第  $j$  类旅游活动的单位碳排放系数,  $S_j$  代表该游客对第  $j$  类旅游活动的情感倾向系数。

### 3) 基于游客体验的地区旅游活动碳足迹模型

地区旅游活动碳足迹模型可以分析某地区不同旅游活动的碳排放量, 指导旅游地的节能减排工作。地区旅游活动碳足迹模型可以看作是大量旅游个体在同一旅游地的碳排放量, 同样加入了情感倾向系数  $S$ , 计算公式为:

$$CR = \sum_j (N_t \times A_j \times P_j \times S_j) \quad (3)$$

其中  $CR$  为该地区的旅游活动碳排放量,  $N_t$  为  $t$  时期的游客总数,  $A_j$  代表游客中参加第  $j$  类旅游活动的占比,  $P_j$  代表第  $j$  类旅游活动的单位碳排放系数,  $S_j$  代表第  $j$  类旅游活动的情感倾向系数。

### 4) 情感倾向系数 $S$

情感倾向系数  $S$  主要是计算游客情感的积极或消极程度, 反映了游客对某类活动的兴趣度和满意度, 一定程度上可以用于预测旅游活动的未来发展趋势。系数  $S$  的区间设定为  $[0,1]$ , 分值越接近 1 表明情感积极程度越高, 分值越接近 0 表明情感消极程度越高。在地区旅游活动碳足迹模型中,  $S$  越大表示此活动受到游客的普遍好评, 未来有更多游客参与此类活动的趋势; 在个人旅游活动碳足迹模型中,  $S$  越大表示该游客更加偏好此类活动, 未来会更频繁参与此类活动。

因此, 对旅游活动的情感进行情感分析, 判断其正负极性, 给出情感值。基于百度 AI 平台中的情感分析功能, 输入词对, 得到每项旅游活动的情感值  $t_i$ , 情感值的区间为  $[0,1]$ , 分值越接近 1 表明情感积极程度越高, 分值越接近 0 表明情感消极程度越高。

对于相同的活动,  $S$  则为该活动的平均情感倾向系数, 每次活动的情感值记为  $t_i$ ,

$$S = \frac{1}{n} \sum_{i=1}^{i=n} t_i \quad (4)$$

## 4. 旅游活动碳足迹实验分析

本节分别对个人旅游活动碳足迹模型和地区旅游活动碳足迹模型进行了实验分析, 验证碳足迹模型的效果, 并对个人和地区旅游政策提供理论指导。

### 4.1. 数据来源

本文从携程网上爬取了旅游游记作为实验数据, 分别选取 2019 年关于海南的 10 篇旅游游记和 2019 年某位游客的 10 篇游记作为分析语料。分别分析游客个人的旅游活动碳排放情况和海南省旅游活动的碳排放情况。

### 4.2. 个人旅游活动碳足迹实验

根据参考文献[3]的数据, 不同旅游活动的碳排放系数存在差异。基于内燃发动机动力系统的旅游活动类型具有较高的碳排放系数, 基于自然条件开展的旅游活动的碳排放系数较小。各项旅游活动对应的碳排放系数如表 5 所示。

选取 2019 年一位游客的 10 篇游记作为语料, 按照本文方法抽取词对, 获得 523 个词对。按照表 5 的旅游活动分类, 将词对汇总统计后, 该游客总共有 186 次旅游活动, 其中自然观光类占 24.1%, 文艺观赏类占 14.1%, 普通运动类占 53.8%, 高端活动类占 8.1%。

**Table 5.** Carbon emission coefficient of tourism activities  
**表 5.** 旅游活动的碳排放系数

大类	子类	碳排放系数(千克/人·次)
自然观光	植物园、公园、自然观光等	0.417
文艺观赏	文艺演出、节庆会展等	0.172
普通运动	潜水、游泳、冲浪、温泉、漂流等	2.24
高端活动	高尔夫、游轮、游艇等	15.3

情感倾向系数 S 的计算, 主要基于百度 AI 平台中的情感分析功能, 输入情感词计算情感值, 最后再得到该旅游活动总体情感倾向系数。

根据 3.2 中的公式(2), 可以估算这位游客的旅游活动碳排放量如表 6, 其中“碳排放量 1”未考虑情感倾向系数, “碳排放量 2”增加了情感倾向系数的影响, “碳排放量排名 1”是按照“碳排放量 1”中的数据降序排列, “碳排放量排名 2”是按照“碳排放量 3”中的数据降序排列。

**Table 6.** Carbon emissions of individual tourist activities  
**表 6.** 游客个人旅游活动的碳排放量

大类	旅游活动占比(%)	情感倾向系数	碳排放量 1 (千克)	碳排放量 2 (千克)	碳排放量排名 1	碳排放量排名 2
自然观光	37.7	0.72	29.24	21.05	3	3
文艺观赏	40.3	0.87	12.89	11.21	4	4
普通运动	16.1	0.62	67.08	41.59	2	2
高端活动	5.9	0.58	167.90	97.38	1	1

#### 4.3. 地区旅游活动碳足迹实验

对游记文本数据预处理后, 按照本文方法抽取旅游活动 - 情感词对, 总共获得 466 个词对。

先筛选过滤非四大类活动的词(如: 吃饭), 同样按照表 5 的旅游活动分类, 将词对汇总统计后, 自然观光类占 24.1%, 文艺观赏类占 14.1%, 普通运动类占 53.8%, 高端活动类占 8.1%。

估算海南省整体的旅游活动碳排放情况, 根据智研咨询发布的报告[23], 2019 年全年海南省接待游客数量 8311 万人次。根据 3.2 中的公式(3), 可以估算各类旅游活动的碳排放量如表 7, 其中“碳排放量 1”未考虑情感倾向系数, “碳排放量 2”增加了情感倾向系数的影响, “碳排放量排名 1”是按照“碳排放量 1”中的数据降序排列, “碳排放量排名 2”是按照“碳排放量 3”中的数据降序排列。

**Table 7.** Carbon emissions from tourism activities in Hainan Province  
**表 7.** 海南省旅游活动的碳排放量

大类	旅游活动占比(%)	情感倾向系数	碳排放量 1 (千克)	碳排放量 2 (千克)	碳排放量排名 1	碳排放量排名 2
自然观光	29.1	0.74	$1.01 \times 10^7$	$7.46 \times 10^6$	3	3
文艺观赏	15.1	0.66	$2.16 \times 10^6$	$1.42 \times 10^6$	4	4
普通运动	43.8	0.87	$8.15 \times 10^7$	$7.09 \times 10^7$	2	1
高端活动	8.1	0.56	$1.03 \times 10^8$	$5.77 \times 10^7$	1	2

#### 4.4. 结果分析

根据 4.2 节中游客个人的碳足迹结果, 碳排放量最高的高端活动在游客活动中占比最小且情感积极

程度最低,可以看出该游客对此类活动的兴趣度较低,未来可以进一步减少此类活动。而碳排放量最低的文艺欣赏类活动是游客的兴趣偏好所在,未来可以进一步增大此类活动占比,从游客个体出发减少旅游活动碳排放,整体实现碳中和旅游。

地区旅游活动碳排放模型即大量旅游个体在地区的碳排放情况。根据 4.3 节中海南省的旅游活动碳足迹来看,不考虑情感倾向时,游艇、高尔夫等高端活动的碳排放量最高,其次是潜水、游泳、冲浪等普通运动类。而加入情感倾向系数后,普通运动类的碳排放量超过了高端运动类。这是由于游客对普通运动类的情感更加积极,游客对该类活动的满意度更高,未来预期将会有更多人参与其中。实验表明,融入情感倾向系数能够动态预测未来的碳排放趋势。海南省旅游管理部门在限制游艇游轮等活动时,也要控制潜水冲浪等活动的人数。

## 5. 总结

在碳中和的背景下,减少旅游碳排放成为旅游产业发展的重要问题,政府需要进行相关政策的调整,指引“碳中和旅游”的发展。对旅游活动碳足迹的研究,前人的文献主要基于相关数据调查和统计。本文以旅游游记为基础,提出了结合词性规则和依存关系的旅游活动特征-情感词对的抽取方法用以分析旅游活动,改进了旅游活动碳足迹模型,并以海南省和游客个人为例,分析了地区和个人的碳排放情况,对旅游政策调整和个人旅游行为具有指导作用。

## 基金项目

中央高校基本科研业务费专项资金资助项目(2232018H-07)。

## 参考文献

- [1] Weidema, B.P., Thrane, M., Christensen, P., Schmidt, J. and Lokke, S. (2008) Carbon Footprint: A Catalyst for Life Cycle Assessment. *Journal of Industrial Ecology*, **12**, 3-6. <https://doi.org/10.1111/j.1530-9290.2008.00005.x>
- [2] 王志民, 张新林, 邱小樱. 基于碳中和的旅游景区净碳排放估算与低碳景区建设——以镇江“焦山”风景区为例[J]. *南京师大学报(自然科学版)*, 2016, 39(4): 107-113.
- [3] 姚治国, 陈田. 基于碳足迹模型旅游碳排放实证研究——以海南省为案例[J]. *经济管理*, 2016, 38(2): 151-159.
- [4] 孙燕燕. 上海市旅游碳排放估算及其效应分解[J]. *地域研究与开发*, 2020, 39(1): 122-126.
- [5] De Choudhury, M., Feldman, M., Amer-Yahia, S., et al. (2010) Automatic Construction of Travel Itineraries Using Social Breadcrumbs. In: *Proceedings of the 21st ACM Conference on Hypertext and Hypermedia*, ACM, New York, 35-44. <https://doi.org/10.1145/1810617.1810626>
- [6] Arase, Y., Xie, X., Hara, T., et al. (2010) Mining People's Trips from Large Scale Geo-Tagged Photos. In: *Proceedings of the 18th ACM International Conference on Multimedia*, ACM, New York, 133-142. <https://doi.org/10.1145/1873951.1873971>
- [7] Yin, H., Cui, B., Huang, Z., et al. (2015) Joint Modeling of Users' Interests and Mobility Patterns for Point-of-Interest Recommendation. In: *Proceedings of the 23rd Annual ACM Conference on Multimedia*, ACM, New York, 819-822. <https://doi.org/10.1145/2733373.2806339>
- [8] Yu, Z., Xu, H., Yang, Z., et al. (2016) Personalized Travel Package with Multi-Point-of-Interest Recommendation Based on Crowd Sourced User Footprints. *IEEE Transactions on Human Machine Systems*, **46**, 151-158. <https://doi.org/10.1109/THMS.2015.2446953>
- [9] 吴清霞, 周娅, 文锦尧, 贺正红. 基于用户兴趣和兴趣点流行度的个性化旅游路线推荐[J]. *计算机应用*, 2016, 36(6): 1762-1766.
- [10] Chen, X.H., Zhang, Y., Ma, P.F., et al. (2015) A Package Generation and Recommendation Framework Based on Travelogues. *IEEE Computer Software & Applications Conference*, Taichung, 1-5 July 2015. <https://doi.org/10.1109/COMPSAC.2015.28>
- [11] Zhu, Z., Shou, L. and Chen, K. (2016) Get into the Spirit of a Location by Mining User-Generated Travelogues. *Neurocomputing*, **204**, 61-69. <https://doi.org/10.1016/j.neucom.2015.04.129>

- [12] 孙文平, 常亮, 宾辰忠, 等. 基于知识图谱和频繁序列挖掘的旅游路线推荐[J]. 计算机科学, 2019, 46(2): 56-61.
- [13] Hu, M. and Liu, B. (2004) Mining and Summarizing Customer Reviews. *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Seattle, 22-25 August 2004, 168-177.  
<https://doi.org/10.1145/1014052.1014073>
- [14] 廖祥文, 陈兴俊, 魏晶晶, 陈国龙, 程学旗. 基于多层关系图模型的中文评价对象与评价词抽取方法[J]. 自动化学报, 2017, 43(3): 462-471.
- [15] 王素格, 吴苏红. 基于依存关系的旅游景点评论的特征——观点对抽取[J]. 中文信息学报, 2012, 26(3): 116-121.
- [16] 张凌, 冯欣. 基于词性模板与依存分析的中文微博情感要素抽取[J]. 计算机科学, 2015, 42(z1): 474-478.
- [17] 张建华, 肖中正. 结合词性规则和依存句法分析的评价对象抽取方法[J]. 计算机与现代化, 2016(4): 16-20.
- [18] 邱云飞, 陈艺方, 王伟, 等. 基于词性特征与句法分析的商品评价对象提取[J]. 计算机工程, 2016(7): 173-180.
- [19] 江腾蛟, 万常选, 刘德喜, 等. 基于语义分析的评价对象——情感词对抽取[J]. 计算机学报, 2017, 40(3): 617-633.
- [20] 王治敏, 朱学锋, 俞士汶. 基于现代汉语语法信息词典的词语情感评价研究[J]. 中文计算语言学, 2005, 10(4): 581-592.
- [21] Che, W., Li, Z. and Liu, T. (2010) LTP: A Chinese Language Technology Platform. *COLING 2010, 23rd International Conference on Computational Linguistics, Demonstrations Volume*, Beijing, 23-27 August 2010.
- [22] 王振宇, 吴泽衡, 胡方涛. 基于 HowNet 和 PMI 的词语情感极性计算[J]. 计算机工程, 2012, 38(15): 187-189+193.
- [23] 智研咨询. 2020-2026 年中国海南省旅游行业发展现状调研及未来前景分析报告[Z].