

doi: 10.12068/j.issn.1005-3026.2019.11.026

基于酒店特征信息和在线评价信息的酒店选择

陶玲玲, 尤天慧, 袁媛
(东北大学 工商管理学院, 辽宁 沈阳 110169)

摘 要: 为辅助游客通过旅游网站进行酒店选择,提出了一种基于旅游网站提供的备选酒店特征信息和在线评价信息的酒店选择方法. 首先,基于酒店特征信息和在线评价信息构建备选酒店有向加权图,依据备选酒店特征信息,基于离差最大化法对酒店特征进行客观赋权,并采用简单加权法确定有向加权图结点权重,依据备选酒店间在线评价信息的比较关系确定有向加权图的有向边及有向边权重;然后,基于 PageRank 算法原理给出备选酒店排序值求解算法;最后,以基于缤客网站提供的酒店特征信息和在线评价信息进行酒店选择. 结果说明了提出方法的有效性和可行性.

关 键 词: 酒店特征信息;在线评价信息;有向加权图;PageRank;酒店选择

中图分类号: C 934 **文献标志码:** A **文章编号:** 1005-3026(2019)11-1667-06

Hotel Selection Based on Information of Hotel Features and Online Ratings

TAO Ling-ling, YOU Tian-hui, YUAN Yuan
(School of Business Administration, Northeastern University, Shenyang 110169, China. Corresponding author: YOU Tian-hui, E-mail: thyu@mail.neu.edu.cn)

Abstract: In order to assist tourists to select hotels through tourism websites, a method for selecting desirable hotels is proposed based on features information and online ratings information of alternative hotels provided by tourism websites. Firstly, the directed and weighted graph of alternative hotels is constructed, in which hotel features are weighted objectively by the maximum deviation method based on hotel features information and then the node weights of directed weighted graph are determined by the simple weighting method. The directed edge and its weight are determined by comparing the online ratings information of node hotels. Then, the algorithm for calculating the sorting value of alternative hotels is given based on the principle of PageRank algorithm. Finally, in order to verify the effectiveness and feasibility of the proposed method, a case study on desirable hotel selection is proposed based on the information of hotel features and online ratings from the Booking.com website.

Key words: hotel features information; online ratings information; directed and weighted graph; PageRank; hotel selection

近年来,随着互联网和电子商务的发展,越来越多的游客在出行时会通过相关旅游网站选择并预订酒店. 在酒店选择过程中,游客通常会浏览相关网站上提供的备选酒店的价格、位置和房间面积等信息,同时也会查看网站上关于备选酒店的员工素质、清洁度和舒适度等在线评价信息,并考虑备选酒店之间在线评价信息的差异. 研究表明,

在线评价信息对于游客的酒店选择具有重要影响^[1-2]. 因此,如何基于酒店特征信息和在线评价信息来辅助游客进行酒店选择是一个值得关注的问题.

目前,基于各类属性/特征信息的酒店排序/选择方法的研究已经引起了学者的关注^[1-9]. Kim 等^[3]基于 logit 模型研究了认知属性(如价

格、服务和食品质量、品牌及运动设施)、情感属性(如舒适度和娱乐性)和感官属性(如房间质量和整体氛围)对消费者酒店选择的影响. Akincilar 等^[4]基于酒店网站所提供的面向客户的(如网站的个性化、网站的响应能力等)、面向技术的(如网站可用性、网站所提供的信息质量等)、面向营销的(如网站促销活动的充分性、网站预订能力等)、面向安全的(如网站安全性、用户对网站的信任程度等)和其他方面的(如网站的交互性、用户对网站内容的满意程度等)指标信息,提出了一种混合多指标决策模型对酒店网站进行评估并排序. Agagt 等^[5]基于实证方法研究了顾客忠诚度、对网站信任度和对网上预定酒店的态度对网上预订酒店意愿的影响. Yu 等^[6-7]提出基于区间二型模糊数似然估计的 MABAC 方法和基于 NLP 的扩展 TODIM 方法,并将这两种方法应用在基于在线评论的酒店排序中. Ghose 等^[8]基于酒店的位置信息、服务信息和在线评论信息构建消费者剩余价值函数,并基于此给出了酒店排序方法. Peng 等^[9]将在线评价信息转换成概率语言信息,在此基础上构建云决策支持模型对酒店进行排序. 然而,已有的研究大多没有从游客的角度来研究如何基于酒店特征信息和在线评价信息进行酒店选择的决策问题. 在现实生活中,当在旅游网站上进行酒店预订时,游客会查看相关网站上的酒店特征信息,同时,也会浏览相关酒店的在线评价信息,并会对备选酒店之间的在线评价信息进行比较以进行酒店选择. 基于此,本文研究基于酒店特征信息和在线评价信息,同时考虑备选酒店之间的在线评价信息差异性,给出一种酒店选择方法为游客进行酒店选择提供决策支持.

1 问题描述

考虑一种基于酒店特征信息和在线评价信息的酒店选择问题,相关的数学符号说明如下:

$R = \{r_1, r_2, \dots, r_m\}$: 游客所关注的 m 个备选酒店的集合,其中 r_i 表示游客所关注的第 i 个备选酒店, $i = 1, 2, \dots, m$.

$C^o = \{c_1^o, c_2^o, \dots, c_n^o\}$: 游客所关注的针对备选酒店的 n 个特征的集合,其中 c_j^o 表示游客所关注的备选酒店的第 j 个特征, $j = 1, 2, \dots, n$.

$C^s = \{c_1^s, c_2^s, \dots, c_q^s\}$: 游客所关注的针对备选酒店的 q 个在线评价特征的集合,其中 c_k^s 表示游客所关注的备选酒店的第 k 个在线评价特征, $k = 1, 2, \dots, q$.

$A^o = [a_{ij}^o]_{m \times n}$: 基于备选酒店特征信息的决策矩阵,其中, a_{ij}^o 表示备选酒店 r_i 对应于备选酒店特征 c_j^o 上的一个特征值,不失一般性,假设 $a_{ij}^o > 0$.

$A^s = [a_{ik}^s]_{m \times q}$: 基于备选酒店在线评价特征信息的评价矩阵,其中, a_{ik}^s 表示备选酒店 r_i 对应于备选酒店在线评价特征 c_k^s 上的一个评价结果, a_{ik}^s 值越大,评价越高.

$w = (w_1, w_2, \dots, w_n)$: 备选酒店特征权重向量,其中 w_j 表示特征 c_j^o 权重或重要程度,满足 $\sum_{j=1}^n w_j = 1, w_j \geq 0$.

本文要解决的问题: 针对游客关注的若干备选酒店、备选酒店的若干特征和若干在线评价特征,如何依据相关网站提供的针对备选酒店的特征信息、在线评价特征信息及未知的酒店特征权重信息,确定一个备选酒店排序方法来支持游客选择合适的酒店.

2 原理与方法

为解决备选酒店选择问题,本文给出了一种基于有向加权图的酒店选择方法. 有向加权图既能处理酒店特征信息,同时又能处理两两备选酒店在线评价特征信息之间的比较关系. 该方法分为两部分: 基于酒店特征信息和在线评价特征信息的有向加权图的构建; 基于有向加权图的备选酒店排序方法. 下面给出本文提出方法的具体过程.

2.1 基于酒店特征信息和在线评价特征信息的有向加权图的构建

记 $G = (V, E, w(r_i), w(r_i, r_{i'}))$ 为用 4 元组表示的备选酒店有向加权图,其中 V 为结点集合, E 为有向边集合(即结点有序对集合), $w(r_i)$ 为结点权重, $w(r_i, r_{i'})$ 为有向边权重. 在有向加权图中,结点为备选酒店,结点权重可基于备选酒店的特征信息确定,有向边权重为结点间的比较关系,可基于备选酒店的在线评价特征信息确定.

1) 基于备选酒店特征信息的有向加权图结点权重的确定. 有向加权图中的结点权重表示结点的重要程度,可基于备选酒店特征的决策矩阵 $A^o = [a_{ij}^o]_{m \times n}$ 确定,首先需要消除不同物理量纲对处理结果的影响,本文采用“比重变换法”^[10],将决策矩阵 $A^o = [a_{ij}^o]_{m \times n}$ 规范化为决策矩阵 $B = [b_{ij}]_{m \times n}$,其计算公式为

$$\left. \begin{aligned} b_{ij} &= b_{ij} / \sum_{i=1}^m b_{ij}, \text{ 当 } c_j^o \text{ 为效益型特征时;} \\ b_{ij} &= (1/b_{ij}) / \left(\sum_{i=1}^m 1/b_{ij} \right), \text{ 当 } c_j^o \text{ 为成本型特征时.} \end{aligned} \right\} \quad (1)$$

为了基于 $\mathbf{B} = [b_{ij}]_{m \times n}$ 确定有向加权图中各结点的权重, 需要确定酒店特征的权重 $\mathbf{w} = (w_1, w_2, \dots, w_n)$, 酒店特征权重既可由游客主观赋权, 也可以基于决策矩阵 $\mathbf{B} = [b_{ij}]_{m \times n}$ 客观赋权. 本文基于离差最大化法对酒店特征进行客观赋权. 离差最大化法是根据特征值的离散程度来确定其权重的大小, 即若特征值在所有备选酒店中差异较大, 则说明其对酒店排序起重要作用, 故赋予较大的权重, 反之则赋予较小的权重. 依据文献[11]可构建最优化模型:

$$\begin{aligned} \max F(\mathbf{w}) &= \sum_{j=1}^n \sum_{i=1}^m \sum_{i'=1}^m |b_{ij} - b_{i'j}| w_j, \\ \text{s. t.} \quad \sum_{j=1}^n (w_j)^2 &= 1; w_j \geq 0. \end{aligned} \quad (2)$$

通过求解最优化模型(2), 可得

$$w_j = \frac{\sum_{i=1}^m \sum_{i'=1}^m |b_{ij} - b_{i'j}|}{\sqrt{\sum_{j=1}^n \left(\sum_{i=1}^m \sum_{i'=1}^m |b_{ij} - b_{i'j}| \right)^2}}. \quad (3)$$

进一步地, 对求得的酒店特征权重 w_j 进行归一化处理, 即令

$$w_j^* = w_j / \sum_{j=1}^n w_j. \quad (4)$$

由此可得

$$w_j^* = \frac{\sum_{i=1}^m \sum_{i'=1}^m |b_{ij} - b_{i'j}|}{\sum_{j=1}^n \sum_{i=1}^m \sum_{i'=1}^m |b_{ij} - b_{i'j}|}. \quad (5)$$

基于酒店特征权重 $\mathbf{w}^* = (w_1^*, w_2^*, \dots, w_n^*)$ 和规范化决策矩阵 $\mathbf{B} = [b_{ij}]_{m \times n}$, 采用简单加权法可得到有向加权图中各结点的权重 $w'(r_i)$, 即

$$w'(r_i) = \sum_{j=1}^n b_{ij} w_j^*. \quad (6)$$

为计算方便, 对有向加权图中各结点权重进行归一化, 有

$$w(r_i) = w'(r_i) / \sum_{i=1}^m w'(r_i). \quad (7)$$

2) 基于备选酒店在线评价特征信息的有向边及有向边权重的确定. 有向加权图中的有向边为两结点酒店间比较关系, 即若两节点酒店存在优劣关系, 则存在有向边, 有向边的权重体现两节点酒店优劣关系的程度, 下面给出基于备选酒店在线评价特征信息的评价矩阵 $\mathbf{A}^s = [a_{ik}^s]_{m \times q}$ 对结

点酒店进行比较确定有向边及有向边权重的过程.

首先, 确定每个在线评价特征下的两两节点酒店间比较关系, 针对在线评价特征 c_k^s , 如果有 $a_{ik}^s > a_{i'k}^s, i' \in \{1, 2, \dots, m\}, i \neq i'$, 则针对特征 c_k^s 存在从酒店 $r_{i'}$ 到 r_i 的有向边, 且有向边的权重计算公式为

$$w_k(r_i, r_{i'}) = \frac{(a_{ik}^s - a_{i'k}^s)}{\sum_{i'=1}^m \rho(a_{ik}^s - a_{i'k}^s)}. \quad (8)$$

其中, 若 $a_{ik}^s > a_{i'k}^s$, 对 $\forall i' \in \{1, 2, \dots, m\}, \rho$ 为 1; 否则 ρ 为 0.

然后, 基于式(8)可确定针对在线评价特征 $C^s = \{c_1^s, c_2^s, \dots, c_q^s\}$ 的有向边的综合权重, 计算公式为

$$w(r_i, r_{i'}) = \frac{\sum_{k=1}^q w_k(r_i, r_{i'})}{\sum_{i'=1}^m \sum_{k=1}^q w_k(r_i, r_{i'})}. \quad (9)$$

2.2 基于有向加权图的备选酒店排序方法

针对 2.1 节构建的备选酒店的有向加权图, 本文基于 PageRank 算法对备选酒店进行排序. 在 PageRank 算法中, 将网页作为结点, 将网页之间的超链接作为有向边. 该算法反映了在某一特定网站上, 上网者随机从一个网页开始, 遵循网页的链接访问下一个网页的过程. PageRank 值可解释为具有随机搜索行为的游客访问某一结点的概率, PageRank 值越高, 该网页的中心位置越显著, 即指向该网页的链接越多^[12]. 本文将备选酒店视为网页, 在线评价特征信息之间的比较关系视为网页之间的超链接^[13]. 在备选酒店有向加权图中, 可以将游客看作是一个悠闲上网者, 他会从一个随机的备选酒店开始, 沿着该备选酒店连接到其他备选酒店的有向边搜索下一个备选酒店. 设游客继续搜索其他酒店的概率为 σ , 则游客停止搜索的概率为 $1 - \sigma$. 依据文献[14-15], 酒店 r_i 的 PageRank 值 $\text{PCS}(r_i)$ (即排序值) 为

$$\text{PCS}(r_i) = (1 - \sigma) w(r_i) + \sigma \sum_{i'=1}^m \beta \text{PCS}(r_{i'}) w(r_i, r_{i'}). \quad (10)$$

其中: σ 为阻尼系数; m 为备选酒店的数量. 若从备选酒店 $r_{i'}$ 到 r_i 存在有向边, 即 r_i 优于 $r_{i'}$, β 为 1; 否则 β 为 0.

为了求解式(10), 记 $\mathbf{w} = [w(r_1), w(r_2), \dots, w(r_m)]^T$, $\mathbf{PCS} = [\text{PCS}(r_1), \text{PCS}(r_2), \dots, \text{PCS}(r_m)]^T$, 则式(10)可转化为矩阵形式

$$\mathbf{PCS} = (1 - \sigma) \times \mathbf{w} + \sigma \times \mathbf{M} \times \mathbf{PCS} \quad (11)$$

其中, \mathbf{M} 为比较关系矩阵, 也叫转移矩阵, 表示游客从一个酒店转到另一个具有更高评价值的酒店的可能性(即有向边的权重), \mathbf{M} 的形式为

$$\mathbf{M} = \begin{bmatrix} 0 & w(r_1, r_2) & \cdots & w(r_1, r_m) \\ w(r_2, r_1) & 0 & \cdots & w(r_2, r_m) \\ \vdots & \vdots & & \vdots \\ w(r_m, r_1) & w(r_m, r_2) & \cdots & 0 \end{bmatrix} \quad (12)$$

进一步地, 令 $\sum_{i=1}^n \mathbf{PCS}(r_i) = 1$, 因为

$$\begin{aligned} \mathbf{PCS} &= (1 - \sigma) \times \mathbf{w} + \sigma \times \mathbf{M} \times \mathbf{PCS} = \\ &= (1 - \sigma) \times \mathbf{w} \times \mathbf{e}^T \times \mathbf{PCS} + \sigma \times \mathbf{M} \times \mathbf{PCS} = \\ &= ((1 - \sigma) \times \mathbf{w} \times \mathbf{e}^T + \sigma \times \mathbf{M}) \times \mathbf{PCS} \quad (13) \end{aligned}$$

其中, $\mathbf{e}^T = [1, 1, \cdots, 1]_{1 \times m}$. 为计算方便, 记 $\mathbf{A} = (1 - \sigma) \times \mathbf{w} \times \mathbf{e}^T + \sigma \times \mathbf{M}$, 则式(11)可表示为 $\mathbf{PCS} = \mathbf{A} \times \mathbf{PCS}$, \mathbf{A} 为最终的转移矩阵^[16].

依据文献[9, 16], 求解 \mathbf{PCS} 的过程是一个马尔可夫过程, 且若矩阵 \mathbf{A} 是一个不可约的、非周期的、随机的转移矩阵, \mathbf{PCS}_k 将收敛于一个平稳的概率分布, 则可由迭代法得到 \mathbf{PCS} 的唯一解. 下面给出式(11)可用迭代法计算的说明.

首先, 基于本文有向加权图的构建过程可知, 矩阵 \mathbf{A} 是不可约的;

其次, 因为 $\mathbf{A} = (1 - \sigma) \times \mathbf{w} \times \mathbf{e}^T + \sigma \times \mathbf{M}$, 有 $A_{i'i} = (1 - \sigma) \times w(r_i) + \sigma \times w(r_i, r_{i'})$; 又因为 $w > 0$ 及 $w(r_i, r_{i'}) > 0$, $A_{i'i} = (1 - \sigma) \times w(r_i) + \sigma \times w(r_i, r_{i'}) > 0$; 所以矩阵 \mathbf{A} 是一个正矩阵.

最后, 又因

$$\begin{aligned} \sum_{i=1}^m A_{i'i} &= \sum_{i=1}^m [(1 - \sigma) \times \mathbf{w} \times \mathbf{e}^T + \sigma \times \mathbf{M}]_{i'i} = \\ &= (1 - \sigma) \times \sum_{i=1}^m w(r_i) + \sigma \times \sum_{i=1}^m M_{i'i} = \\ &= (1 - \sigma) \times \sum_{i=1}^m (H_i / \sum_{i=1}^m H_i) + \sigma \times \sum_{i=1}^m M_{i'i} = \\ &= (1 - \sigma) \times \sum_{i=1}^m (H_i / \sum_{i=1}^m H_i) + \sigma \times \sum_{i=1}^m w(r_{i'}, r_i) = \\ &= (1 - \sigma) + \sigma \times 1 = 1. \quad (14) \end{aligned}$$

所以矩阵 \mathbf{A} 中每一行的元素相加和都等于 1, 即 \mathbf{A} 是非周期的、随机的矩阵.

由此可知, \mathbf{PCS}_k 收敛于一个平稳的概率分布. 因此可通过迭代法对式(11)求解, 依据文献[17-18], 给出用迭代法求解式(11)的具体步骤.

- 1) 对 \mathbf{PCS} 赋初始值 \mathbf{PCS}_0 ;
- 2) 将初始值 \mathbf{PCS}_0 代入式(11)中, 得到

$$\mathbf{PCS}_1 = (1 - \sigma) \times \mathbf{w} + \sigma \times \mathbf{M} \times \mathbf{PCS}_0;$$

3) 将 \mathbf{PCS}_1 赋值给 \mathbf{PCS}_0 , 得到 \mathbf{PCS}_2 ;

4) 重复步骤 3), 即不断地将 \mathbf{PCS}_{k+1} 赋值给 \mathbf{PCS}_k , 直到满足条件 $|\mathbf{PCS}_{k+1} - \mathbf{PCS}_k| < \varepsilon$ 后迭代停止, ε 为停止迭代的条件, 该条件可由决策者给出;

5) 得到 \mathbf{PCS} , 依据该值对备选酒店进行排序.

3 实例分析

以游客到上海旅游选择酒店为例来说明所给出的基于酒店特征信息和在线评价信息的酒店选择方法的潜在应用. 假设某游客准备去上海旅游, 通过查阅缤客网站(<https://www.booking.com>), 其关注了上海市中心附近的 5 家四星级酒店王宝和大酒店($\mathbf{PCS}(r_1)$)、上海富豪东亚酒店(r_2)、上海广场长城假日酒店(r_3)、上海海神诺富特大酒店(r_4)和上海浦东丽晟假日酒店(r_5)及关于这 5 家酒店价格(c_1^o)、距市中心距离(c_2^o)、房间面积(c_3^o)特征信息和关于员工素质(c_1^s)、清洁度(c_2^s)、舒适度(c_3^s)的在线评价情况. 从缤客网站获取的这 5 家酒店的特征信息和在线评价特征信息(决策矩阵和评价矩阵)如表 1 所示, 其中, 酒店价格和房间面积选取的是标准大床房的信息, 在线评价特征信息是截止到 2018 年 3 月 8 日缤客网站基于所有在线评价信息给出的针对备选酒店各在线评价特征的综合评价价值. 为了辅助该游客进行酒店选择, 采用前文给出的方法, 给出主要计算过程和结果的描述.

表 1 5 家备选酒店的决策矩阵和评价矩阵						
Table 1 Decision matrix and evaluation matrix of the five alternative hotels						
矩阵	r	r_1	r_2	r_3	r_4	r_5
\mathbf{A}^o	c_1^o	783.0	850.0	675.8	800.0	820.0
	c_2^o	1.5	3.0	2.2	2.5	3.1
	c_3^o	30.0	32.0	28.0	26.0	32.0
\mathbf{A}^s	c_1^s	8.1	7.8	8.2	8.5	7.6
	c_2^s	8.8	7.9	7.7	8.2	8.0
	c_3^s	8.8	7.8	7.5	8.3	7.9

第一步: 备选酒店有向加权图的构建.

首先, 依据式(1)将决策矩阵 \mathbf{A}^o 规范化为决策矩阵 \mathbf{B} , 得

$$B = \begin{bmatrix} 0.199 & 0.306 & 0.203 \\ 0.184 & 0.153 & 0.216 \\ 0.231 & 0.209 & 0.189 \\ 0.195 & 0.184 & 0.176 \\ 0.190 & 0.148 & 0.216 \end{bmatrix}.$$

通过求解模型(2)可得备选酒店特征权重分别为： $w_1=0.178, w_2=0.637, w_3=0.185$ 。依据式(6)和式(7)可得有向加权图各结点权重分别为： $w(r_1)=0.268, w(r_2)=0.170, w(r_3)=0.209, w(r_4)=0.184, w(r_5)=0.168$ 。

其次,基于评价矩阵 A^s ,依据式(8)可判断和计算每个在线评价特征下有向加权图两两节点酒店间是否存在有向边及有向边权重,得员工素质(c_1^s)特征下存在的有向边及权重为

$$\begin{aligned} w_1(r_1, r_2) &= 0.375; w_1(r_1, r_5) = 0.625; \\ w_1(r_2, r_5) &= 1.000; w_1(r_3, r_1) = 0.091; \\ w_1(r_3, r_2) &= 0.364; w_1(r_3, r_5) = 0.545; \\ w_1(r_4, r_1) &= 0.174; w_1(r_4, r_2) &= 0.304; \\ w_1(r_4, r_3) &= 0.130; w_1(r_4, r_5) &= 0.391. \end{aligned}$$

清洁度(c_2^s)特征下存在的有向边及权重为
 $w_2(r_1, r_2)=0.265; w_2(r_1, r_3)=0.324;$

$$\begin{aligned} w_2(r_1, r_4) &= 0.176; w_2(r_1, r_5) = 0.235; \\ w_2(r_2, r_3) &= 1.000; w_2(r_4, r_2) = 0.300; \\ w_2(r_4, r_3) &= 0.500; w_2(r_4, r_5) = 0.154; \\ w_2(r_5, r_2) &= 0.250; w_2(r_5, r_3) = 0.800. \end{aligned}$$

舒适度(c_3^s)特征下存在的有向边及权重为

$$\begin{aligned} w_3(r_1, r_2) &= 0.270; w_3(r_1, r_3) = 0.351; \\ w_3(r_1, r_4) &= 0.135; w_3(r_1, r_5) = 0.243; \\ w_3(r_2, r_3) &= 1.000; w_3(r_4, r_2) = 0.294; \\ w_3(r_4, r_3) &= 0.471; w_3(r_4, r_5) = 0.200; \\ w_3(r_5, r_2) &= 0.200; w_3(r_5, r_3) = 0.800. \end{aligned}$$

依据式(9)可得有向边的综合权重为

$$\begin{aligned} w(r_1, r_2) &= 0.303; w(r_1, r_3) = 0.225; \\ w(r_1, r_4) &= 0.104; w(r_1, r_5) = 0.368; \\ w(r_2, r_3) &= 0.667; w(r_2, r_5) = 0.333; \\ w(r_3, r_1) &= 0.091; w(r_3, r_2) = 0.364; \\ w(r_3, r_5) &= 0.545; w(r_4, r_1) = 0.060; \\ w(r_4, r_2) &= 0.308; w(r_4, r_3) = 0.377; \\ w(r_4, r_5) &= 0.255; w(r_5, r_2) = 0.225; \\ w(r_5, r_3) &= 0.775. \end{aligned}$$

基于以上计算结果,可构建备选酒店的有向加权图如图 1 所示。

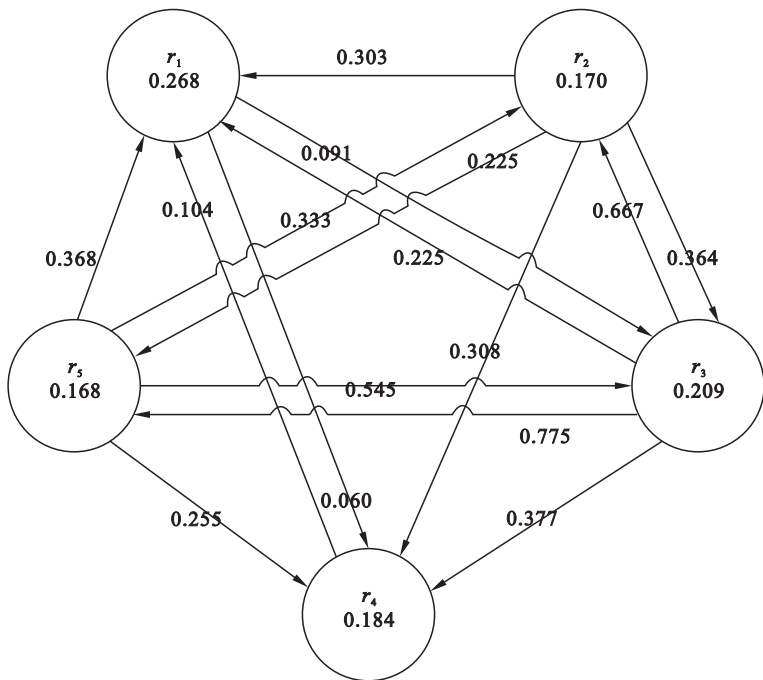


图 1 5 个备选酒店的有向加权图
Fig. 1 Directed weighted graph of the five alternative hotels

第二步:基于有向加权图的备选酒店排序值的计算。

首先,依据图 1,可确定式(11)中各参数的值,即有, $w = [0.268 \ 0.170 \ 0.209 \ 0.184 \ 0.168]^T$;

$$M = \begin{bmatrix} 0 & 0.303 & 0.225 & 0.104 & 0.368 \\ 0 & 0 & 0.667 & 0 & 0.333 \\ 0.091 & 0.364 & 0 & 0 & 0.545 \\ 0.060 & 0.308 & 0.377 & 0 & 0.255 \\ 0 & 0.225 & 0.775 & 0 & 0 \end{bmatrix};$$

另外,依据文献[15],本文取 $\sigma = 0.85$.

然后,应用 Matlab 软件编程进行迭代运算得到备选酒店的 $PCS(r_i)$ 值为

$PCS(r_1) = 0.210\ 2, PCS(r_2) = 0.195\ 5,$

$PCS(r_3) = 0.201\ 4, PCS(r_4) = 0.197\ 6,$

$PCS(r_5) = 0.195\ 2.$

由此可得到 5 家备选酒店的排序结果为:

$r_1 > r_3 > r_4 > r_2 > r_5$, 游客可以选择王宝和大酒店(r_1)进行预定.

4 结 语

本文给出了一种基于酒店特征信息和在线评价信息的酒店选择方法. 依据酒店特征信息和在线评价信息比较关系构建备选酒店有向加权图,在此基础上,依据 PageRank 算法计算备选酒店排序值以辅助游客进行酒店选择. 本文给出的方法具有概念清晰、计算简便、易于软件实现等特点,为解决现实中基于酒店特征信息和在线评价信息的酒店选择问题提供了一种新途径. 另外,本文提出的方法也适用于基于产品特征信息和在线评论信息的各类产品选择问题. 本文局限在于仅考虑在线评价信息,今后研究工作中需要考虑文本形式在线评论信息.

参考文献:

- [1] Ladhari R, Michaud M. eWOM effects on hotel booking intentions, attitudes, trust, and website perceptions [J]. *International Journal of Hospitality Management*, 2015, 46: 36–45.
- [2] Casaló L V, Flavián C, Guinalfú M, et al. Do online hotel rating schemes influence booking behaviors? [J]. *International Journal of Hospitality Management*, 2015, 49: 28–36.
- [3] Kim D, Perdue R R. The effects of cognitive, affective, and sensory attributes on hotel choice [J]. *International Journal of Hospitality Management*, 2013, 35: 246–257.
- [4] Akincilar A, Dagdeviren M. A hybrid multi-criteria decision making model to evaluate hotel websites [J]. *International Journal of Hospitality Management*, 2014, 36: 263–271.
- [5] Agag G, El-Masry A A. Understanding the determinants of hotel booking intentions and moderating role of habit [J]. *International Journal of Hospitality Management*, 2016, 54: 52–67.
- [6] Yu S M, Wang J, Wang J Q. An interval type-2 fuzzy likelihood-based MABAC approach and its application in selecting hotels on a tourism website [J]. *International Journal of Fuzzy Systems*, 2017, 19(1): 47–61.
- [7] Yu S M, Wang J, Wang J Q. An extended TODIM approach with intuitionistic linguistic numbers [J]. *International Transactions in Operational Research*, 2016: 1–25.
- [8] Ghose A, Ipeiritos P G, Li B B. Designing ranking systems for hotels on travel search engines by mining user-generated and crowd sourced content [J]. *Social Science Electronic Publishing*, 2012, 31(3): 493–520.
- [9] Peng H G, Zhang H Y, Wang J Q. Cloud decision support model for selecting hotels on TripAdvisor. com with probabilistic linguistic information [J]. *International Journal of Hospitality Management*, 2018, 68: 124–138.
- [10] Goh C H, Tung Y C A, Cheng C H. A revised weighted sum decision model for robot selection [J]. *Computers & Industrial Engineering*, 1996, 30(2): 193–199.
- [11] 王应明. 运用离差最大化方法进行多指标决策与排序[J]. *系统工程与电子技术*, 1998, 20(7): 24–26.
(Wang Ying-ming. Using the method of maximizing deviations to make decision for multi-indicies [J]. *Journal of System Engineering and Electronics*, 1998, 20(7): 24–26.)
- [12] 景楠, 王建霞, 许皓, 等. 基于用户社会关系的社交网络好友推荐算法研究 [J]. *中国管理科学*, 2017, 25(3): 137–146.
(Jing Nan, Wang Jian-xia, Xu Hao, et al. Friend recommendation algorithm based on user relations in social networks [J]. *Chinese Journal of Management Sciences*, 2017, 25(3): 137–146.)
- [13] 王伟, 王洪伟. 面向竞争力的特征比较网络: 情感分析方法 [J]. *管理科学学报*, 2016, 19(9): 109–126.
(Wang Wei, Wang Hong-wei. Comparative network for product competition in feature-levels through sentiment analysis [J]. *Journal of Management Sciences in China*, 2016, 19(9): 109–126.)
- [14] Yang X, Yang G F, Wu J N. Integrating rich and heterogeneous information to design a ranking system for multiple products [J]. *Decision Support Systems*, 2016, 84: 117–133.
- [15] Scholz M, Pfeiffer J, Rothlauf F. Using Page Rank for non-personalized default rankings in dynamic markets [J]. *European Journal of Operational Research*, 2017, 260(1): 388–401.
- [16] Fotopoulos S B. Probability and random process [M]. Hoboken: Wiley, 1985.
- [17] Li S, Zha Z J, Ming Z Y, et al. Product comparison using comparative relations [C] // International ACM SIGIR Conference on Research & Development in Information Retrieval. Beijing, 2011: 1151–1152.
- [18] Li S, Ming Z Y, Leng Y, et al. Product ranking using hierarchical aspect structures [J]. *Journal of Intelligent Information Systems*, 2017, 48(2): 453–474.