

基于嵌入表示和加权矩阵分解的线路推荐

乔永卫

中国民航大学工程技术训练中心, 天津

Email: qiaoyongwei76@126.com

收稿日期: 2021年3月20日; 录用日期: 2021年4月14日; 发布日期: 2021年4月22日

摘要

为了提升游客体验和增加旅游公司收入, 旅游线路推荐成为重要手段。目前已有方法利用了游客与线路的交互历史, 提升了推荐性能, 但忽略了喜欢相同线路的共现游客对信息、被很多游客喜欢/不喜欢的线路对信息, 这些对建模旅客偏好至关重要。因此本文提出了基于嵌入表示和加权矩阵分解旅游线路推荐算法。首先根据游客与线路的交互历史, 抽取共现游客对、共现喜欢/不喜欢线路对, 建立其词向量模型; 接着利用加权矩阵分解模型对游客与线路的交互历史、共现游客对、喜欢/不喜欢线路对信息进行联合分解完成线路预测; 最后在厦航航空公司某旅游公司的真实数据集上评估了算法的有效性。

关键词

词向量, 共现游客对, 加权矩阵分解, 喜欢/不喜欢线路对

Route Recommendation Based on Embedding and Weighted Matrix Factorization

Yongwei Qiao

Engineering Technology Training Center, Civil Aviation University of China, Tianjin

Email: qiaoyongwei76@126.com

Received: Mar. 20th, 2021; accepted: Apr. 14th, 2021; published: Apr. 22nd, 2021

Abstract

To improve the tourist experience and increase the income of tourism companies, tourism route recommendation has become an important method. Many existing methods train and learn the interaction history between tourists and routes to improve the recommendation performance.

However, the user co-occurrence information, co-disliked routes and co-liked routes are very important for modeling tourist preference. Therefore, this paper proposes a route recommendation algorithm based on embedding representation and weighted matrix factorization. Firstly, it extracts user's co-occurrence information, co-disliked routes and co-liked routes based on the interaction history between tourists and routes to establish word vector model by extracting co-occurrence tourists' likes/dislikes of routes. Then it uses weighted matrix factorization to decompose the interaction history, the co-occurrence traveler information, co-disliked routes and co-liked routes to complete the route recommendation. Finally, the effectiveness of our method is evaluated on a real data set of a travel company of Xiamen Airlines.

Keywords

Word Embedding, Co-Occurrence Traveler, Weighted Matrix Factorization, Co-Disliked Routes, Co-Liked Routes

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

随着人们生活水平的提高, 人们有越来越多的出门旅游的机会。据统计近几年旅游的人数和收入以10%以上的速度在增长。为了争夺客源及增加收入, 旅游公司需了解游客的需求, 制定普受欢迎的线路; 同时游客需要从大量已有的线路中选择符合自己兴趣的线路。推荐系统能分析用户的历史行为, 建模用户的偏好, 主动推送符合用户兴趣的个性化产品。将推荐系统应用旅游线路的推荐, 不仅能更好地掌握游客的偏好, 为其推荐符合其兴趣的线路, 提升他们的体验, 而且能帮助旅游公司制定更好的线路, 增加旅游公司的收入。

目前推荐系统贯穿着整个的旅游过程, 比如旅游前的线路推荐[1], 旅游中景点、个性化服务推荐[2]及旅游结束后信息的收集和反馈[3]。张艳根据游客的地理位置信息, 生成游客轨迹, 并根据景点实时容量统计, 为游客推荐景点[4]。韩艳等[5]引入热门链路概念, 以游客满意度最大、景点间行程最短、景点拥挤度最低为目标, 利用粒子群优化算法为游客推荐景点。陈雄等[6]利用蚂蚁算法在考虑时间关系的基础上为游客推荐其感兴趣的景点。Duan 等分析游客的签到行为, 利用 Convolutional Neural Network (CNN) 抽取游客签到地区特征, 为游客推荐其感兴趣的景点[7]。Zhu 等[8]根据游客的签到数据进行了时间敏感性的个性化线路推荐。这些方法在一定程度上提高了线路推荐的精度, 但由于相较于电影、音乐等数据, 游客参与过的线路相对更少, 因此旅游数据集中“旅客-线路”关联矩阵稀疏度更高。

为了丰富旅游信息, 缓解数据稀疏带来的影响, 研究者将线路主题[9] [10]、旅客和线路属性[11] [12] [13]、社交关系[14]、评论[15] [16]等信息融入到推荐过程, 提高了线路推荐准确度。朱桂祥[9]等从线路描述中挖掘主题信息来捕捉用户行为模式, 从访问序列数据中挖掘频繁序列模式, 利用马尔科夫模型完成用户实时点击流与模式库匹配计算进行线路推荐。李雅美等[11]用景点特征标签描述旅客的兴趣, 并利用基于用户的协同过滤进行推荐。丁恒等[12]利用线路属性特征从多个维度计算相似性。孙文平等[13]融合多源旅游数据构建知识图谱来丰富旅游数据集。李晓旭等[14]采用优化聚类挖掘社交关系中 Coterie 组模式, 并提出基于语义的距离敏感度和基于语义的从中性两种推荐策略。崔春生等[15]根据在线评论信息计算旅客对各景点属性的权重并进行排序。潘禄生[16]分析游客对线路评论的情感, 结合游客的喜好, 为

游客推荐喜欢的线路。这些方法利用线路主题、社交关系、属性特征及评论等信息，缓解了旅游数据的高稀疏性，提高了线路推荐性能，但这些额外信息的表示困难。

基于表示学习的推荐算法利用项目/用户的属性信息、游客和线路的交互行为训练、学习用户和项目的向量表示，它们能更好地表示了游客的偏好和项目特征，提高推荐性能。贾中浩等[17]使用表示学习得到线路的低维向量表示，利用门控制单元(GRU)建模用户长短期偏好，完成线路推荐，但它们只利用了游客和线路的交互历史。在实际过程中两个游客可能总是喜欢相同的线路，某些线路对出现在很多人喜欢或不喜欢的列表内。这些信息对于为旅客精准推荐符合其兴趣的线路都至关重要。

为了解决上述问题，本文提出了基于嵌入表示和加权矩阵分解的旅游线路推荐算法。首先根据旅客与线路的交互行为，抽取共现游客对、共同喜欢线路对、共同不喜欢线路对。接着利用词向量模型分别学习游客和线路的嵌入向量表示。最后利用加权矩阵分解模型将游客线路交互矩阵，游客共现矩阵，游客喜欢/不喜欢线路共现矩阵联合分解完成评分预测并进行线路推荐。

2. 基于词向量的游客和线路的嵌入表示

假设游客集合 $U = \{u_1, u_2, \dots, u_m\}$ ，旅游线路的集合 $L = \{l_1, l_2, \dots, l_n\}$ ，游客与线路的交互历史矩阵为 $M^{m \times n} = \{r_{ul} | u \in U, l \in L\}$ ，如果游客 u 参加了线路 l ，则 $r_{ul} = 1$ ，否则为 0。模型的任务是为目标旅客推荐其可能感兴趣的线路集合。

2.1. 共现线路对和共现游客对集合的抽取

在旅游线路推荐系统，如果一个游客 u 参与过某条线路 l ，则认为游客 u 喜欢线路 l 。从游客与线路的交互历史矩阵中，可得每位游客可参加多条线路，多名游客可共同选择相同的线路 l_1 和 l_2 ， l_1 和 l_2 被看作是游客共同喜欢的线路对。当某个游客选择过 l_1 ，还没有选择过 l_2 时，此时可以将 l_2 推荐给他。相似地，当游客 u_i 和 u_j 共同参加过很多共同线路时，认为他们具有似的线路偏好，当游客 u_i 参加过某线路 l_i ，而 u_j 并没有参加过，可以将 l_i 推荐给 u_j 。游客 u_i 和 u_j 被认为是具有相似偏好的共现游客对。

如果给定交互矩阵中 $M^{m \times n}$ ，可从中抽取三元组集合 $L_{like-co} = \{(l_i, l_j, t_{ij}) | l_i, l_j \in L, r_{ul_i} = 1, r_{ul_j} = 1\}$ 和 $U_{co} = \{(u_m, u_n, t_{mn}) | l_k \in L, r_{u_m l_k} = 1, r_{u_n l_k} = 1\}$ ，它们分别表示被游客共同喜欢的共现线路对集合、参加过相同线路的具有相似偏好的共现游客对集合及它们出现次数。 $L_{like-co}$ 表示了线路被旅客喜欢的程度，反映线路的受欢迎程度， t_{ij} 越大，表示它们越受欢迎。 U_{co} 表示游客的共同偏好， t_{mn} 越大，表示游客 u_m 和 u_n 的偏好越相似。

2.2. 不喜欢共现线路对集合的抽取

旅游公司为游客推荐线路时，如果能了解游客不喜欢的线路特点，就能避免向游客推荐其根本不可能有兴趣的线路。对于电影等具有评分等显示反馈的推荐系统，抽取喜欢或不喜欢的的项目都是相对容易的。游客与线路的交互属于隐式反馈，对于大量 $r_{ul} = 0$ 的线路，游客没有参加，有可能是喜欢，也可能是不了解此线路造成的。因此对于旅游数据集最大的挑战在于如何推测游客不喜欢的线路。

为了抽取不喜欢共现线路对集合来提高推荐精度，假设游客没有参与过的线路是有很大概率不被游客喜欢的。因此对于所有线路，利用面向游客的 EM-like 算法推测游客不喜欢的线路。首先假设在游客的线路排序列表中，被游客喜欢的线路的排序评分要高于具有较高概率不被喜欢的线路。对于游客 u_i 的所有线路排序列表为 s_{u_i} ，则利用 softmax 函数计算游客 u_i 不喜欢线路 l_k 的概率为式(1)

$$P_{u_i}^k = \frac{\exp(-s_{u_i}(k))}{\sum_{q=1}^n \exp(-s_{u_i}(q))} \quad (1)$$

其中 $s_{u_i}(k)$ 表示线路 l_k 在游客 u_i 排序列表中的位置。为了避免游客参加过线路被选择，将游客旅游过的线路排序设置为 $+\infty$ 。由于每一个游客旅游过的线路数量不同，因此抽取的不喜欢线路的数目也应不同。设定超参数 τ 控制每个游客抽取不喜欢线路的数量，游客参与过的线路越多，抽取不喜欢的数目也就越大，将该参数称为负采样率。假设某游客旅游过 10 条线路， $\tau = 0.8$ ，则为其抽取 8 条不喜欢的线路。利用式(1)得到线路的先验概率。为每个游客随机抽取其不喜欢的线路数据，并计算器不喜欢的概率后，进而可得到不喜欢的共现线路对集合，用三元组表示为 $L_{dislike-co} = \left\{ (l_i, l_j, t_{ij}) \mid l_i, l_j \in L, r_{ul_i} = 0, r_{ul_j} = 0 \right\}$ 。

2.3. 基于旅游数据抽取的词向量模型

词向量模型最初应用于文本处理，给定文本序列，它能学习每个词的潜在表示。Word2vector (比如，Skip-gram) 是其最流行的模型。采用负采样机制的 Skip-gram 模型可分解每个词及其上下文组成的词对。因此可以用负采样机制的 Skip-gram 处理在交互序列对、共现游客对、共现线路对和不喜欢的共现线路对集合，得到共现游客对、共现线路对和不喜欢的共现线路对的潜在表示形式。因此假设 D 表示旅游过线路和与它们共现的喜欢线路对集合，则线路 l_i 与它的共现线路 l_j 的共现概率为式(2)

$$PI(i, j) = \log \frac{\#(i, j) * |D|}{\#(i) * \#(j)} \quad (2)$$

$|D|$ 表示 D 中所包含的交互序列对的总数， $\#(i)$ ， $\#(j)$ 分别表示 l_i 和 l_j 出现的次数， $\#(i, j)$ 表示 l_i 和 l_j 共同出现的次数。通过计算 D 中所有线路及其共现线路对的 PI ，可以形成一个 $n \times n$ 的平方矩阵 M^{PI} ，这里的 n 是 D 中独立的线路数。相似地，对于共现游客对集合、不喜欢线路共现对集合都可计算相应的 M^{PI} 。这样就可以利用词向量模型得到共现用户游客对、共同喜欢线路对和不喜欢的共现线路对的潜在表示。

3. 游客和线路的加权矩阵联合分解模型的建立

综合考虑共现游客对、共现线路对和不喜欢的共现线路对，利用词向量模型得到共现游客对、共现线路对和不喜欢的共现线路对潜在表示，随后将其表示引入加权矩阵分解模型进行联合分解，学习并优化游客和线路的向量表示，从而提高推荐性能。

3.1. 游客与线路交互历史的加权矩阵分解

设有游客与线路的交互历史矩阵为 $M^{m \times n} = \{r_{ul} \mid u \in U, l \in L\}$ ，加权矩阵分解将 M 分解为 2 个低阶矩阵 $U^{m \times k}$ 和 $P^{n \times k}$ 分别表示游客和线路的潜在表示，其中 k 是维数， $U^T = (\alpha_1, \alpha_2, \dots, \alpha_m)$ ， α_u 表示游客 u 的向量表示， $P^T = (\beta_1, \beta_2, \dots, \beta_n)$ 表示线路的向量表示， β_l 表示线路 l 的表示。 M 的加权矩阵分解模型为式(3)

$$L_1 = \frac{1}{2} \sum_{u,l} w_{ul} (M_{ul} - \alpha_u \beta_l)^2 + \frac{1}{2} \left(\lambda_u \sum_u \|\alpha_u\|^2 + \lambda_\beta \sum_l \|\beta_l\|^2 \right) \quad (3)$$

其中超参数 w_{ul} 用来衡量游客对线路喜欢程度的置信水平和交互矩阵中非零值和零值的数目。 M_{ul} 为游客与线路的交互历史矩阵。为了避免过拟合，用 λ_u 和 λ_l 调整正则项 $\sum_u \|\alpha_u\|^2$ 和 $\sum_l \|\beta_l\|^2$ 。

3.2. 共同喜欢线路的嵌入表示

将游客参加的线路按交互时间升序排列即可得到旅游线路序列。因此可以利用词嵌入方法来学习线路的潜在表示。给定每个游客的旅游线路序列，即可得到旅客喜欢的共现线路对集合 X 。因此学习共同喜欢线路对嵌入表示的目标函数为式(4)

$$L_2 = \frac{1}{2} \sum_{X_{lm} \neq 0} w_{lm}^{(+p)} (X_{lm} - \beta_l \gamma_m - b_l - c_m)^2 + \frac{1}{2} \lambda_m \sum_m \|\gamma_m\|^2 \quad (4)$$

其中 β_l 表示线路的潜在表示, γ_m 共同喜欢线路潜在表示向量, b_l 是喜欢线路的偏差, c_m 是共同喜欢线路的偏差, $w_{lm}^{(+p)}$ 为超参数, λ_m 为正则项调节系数。

3.3. 共同不喜欢线路嵌入表示

如果许多用户同时不喜欢两条线路时, 这两条线路形成不喜欢共现线路对集合。如果推荐系统了解到这种不喜欢的共现模式, 就可以避免为游客推荐不喜欢的线路, 可降低推荐系统的误报率。因此, 与共同喜欢线路嵌入的类似, 利用游客共同不喜欢的线路对矩阵 Y , 学习共同不喜欢线路的表示。不喜欢共现线路嵌入表示的目标函数可以表示为式(5)

$$L_3 = \frac{1}{2} \sum_{Y_{lm1} \neq 0} w_{lm1}^{(-p)} (X_{lm1} - \beta_l \delta_{m1} - d_l - e_{m1})^2 + \frac{1}{2} \lambda_{m1} \sum_{m1} \|\delta_{m1}\|^2 \quad (5)$$

其中 β_l 表示线路的潜在表示, δ_{m1} 共同不喜欢线路潜在表示向量, d_l 是不喜欢线路的偏差, e_{m1} 是共同不喜欢线路的偏差, $w_{lm1}^{(-p)}$ 为超参数, λ_{m1} 为正则项调节系数。

3.4. 共现游客嵌入表示

当游客 u_i 和 u_j 去过相同的旅游线路时, 认为他们有相似的兴趣。因此, 如果 u_i 去过某线路但游客 u_j 未去过的线路时, 可以向 u_j 推荐该线路。从游客线路交互矩阵 M , 可以得到由游客-游客共现矩阵 Z 及概率。共现游客嵌入的目标函数可以表示为式(6)

$$L_4 = \frac{1}{2} \sum_{Z_{ui} \neq 0} w_{lm1}^{(u)} (Z_{ui} - \alpha_u \theta_i - f_u - g_i)^2 + \frac{1}{2} \lambda_l \sum_l \|\theta_l\|^2 \quad (6)$$

其中 α_u 表示游客的潜在表示, θ_l 表示共现游客的潜在表示, f_u , g_l 表示游客和共现游客的偏差, $w_{lm1}^{(u)}$ 为超参数, λ_l 为正则项调节系数。

3.5. 加权矩阵联合分解模型

综合加权矩阵分解、共同喜欢线路、共同不喜欢线路及共现游客的嵌入表示, 得到最后的联合分解目标函数表示为式(7)

$$L = L_1 + L_2 + L_3 + L_4 \quad (7)$$

其中, 令正则化系数 $\lambda_u = \lambda_\beta = \lambda_m = \lambda_{m1} = \lambda_l = \lambda$ 。利用梯度下降法对公式(7)进行优化, 学习可得到游客和线路的表示, 从而进行线路推荐。

4. 实验结果及分析

4.1. 数据集

旅游数据来源于厦门航空旗下的某旅游公司, 包括 4737 个游客对 1436 条旅游线路的 25717 条旅行记录。每名游客至少包含 3 条参加过的线路, 每条记录包括旅游团号、游客姓名、性别、游客身份证号、线路出发时间、线路价格、景点的详细介绍。它是一个隐式反馈数据集, 只要游客参加过的某线路, 则认为该游客喜欢这条线路, 没有游客对线路明确的喜好。对数据集以 7:2:1 的比例拆分成训练集(train)、验证集(validation)和测试集(test)。训练集用来训练模型, 验证集用来参数选取, 测试集则用来评估模型。

4.2. 评估指标

本文采用召回率(Recall)、归一化折损累计增益(Normalized Discounted Cumulative Gain, NDCG)和平均精准度(Mean Average Precision, MAP)作为评估标准。Recall 描述推荐系统推荐给用户的旅游线路占用户真正感兴趣的线路的比例。NDCG 和 MAP 则表示命中项目在推荐列表中排序位置的情况。

4.3. 参数训练

4.3.1. 负采样率 τ 的影响

负采样率表示为每个旅客抽取多少个不喜欢的项目。由于每个游客实际参与过的线路数不同,因此在固定的负采样率下,为每个游客抽取的不喜欢线路数是不同的。对同一个游客来说,负采样率 τ 越大,抽取的不喜欢的线路越多,优化过程中负样本较多。如果负采样率 τ 的值越小,抽取的不喜欢的线路数越少,负样本会较少。因此在试验中讨论了为每个游客抽取多少个负样本,能使模型性能达到最好。试验中采用网格搜索的方式,根据算法性能来选择负采样率。在固定正则化参数和潜在因子的情况下, τ 的搜索范围为{0.1, 0.2, 0.4, 0.6, 0.8, 1.0},性能结果如图 1 所示。

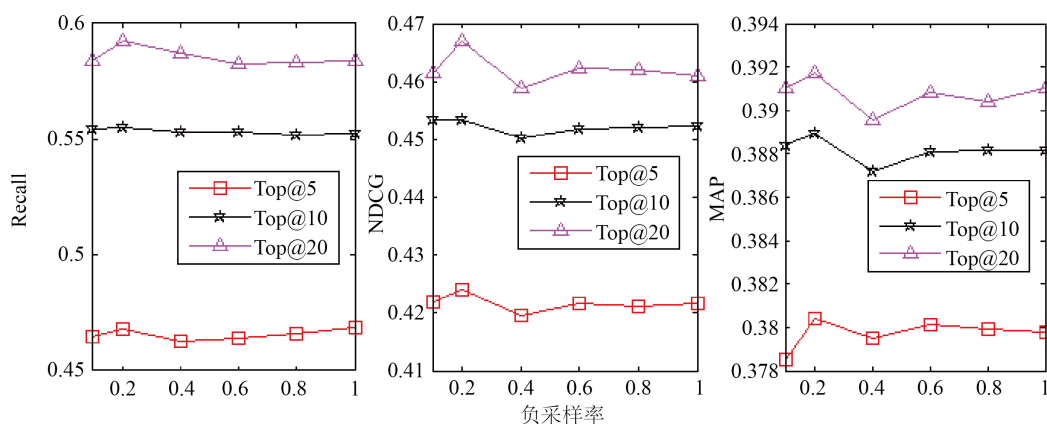


Figure 1. Effect of negative sampling rate
图 1. 负采样率对性能的影响

从图 1 可知,随着 τ 增大,在 Top@5、Top@10 和 Top@20 的三种评估指标都是先增大后减小,之后趋于不变。结果证实了随着负样本数量的增加,模型性能先增加后减少,而负样本数目越大,模型的训练也会越慢,因此为了平衡模型性能和效率,负采样率取为 0.2。

4.3.2. 向量维度 k 的影响

向量维度表示游客和线路潜在表示的长度。向量维度越大,游客和线路的信息表示越丰富,但计算量也越大。向量维度越小,向量表示不够精确,但学习速度快。因此在平衡精度与计算量的情况下,一般通过实验来设定较合适的维度大小。设定向量维度 k 搜索范围为在{50, 100, 150, 200, 250, 300},实验结果如图 2 所示。从图 2 可知,随着向量维度的增大,NDCG 和 MAP 指标快速增长,当 k 达到 150 以上的时候,增速变缓,算法性能趋于稳定。Recall 指标的增长不如 NDCG 和 MAP 变化剧烈,但仍然是先增大后减小,到达 150 之后增速趋缓,因此向量维度取为 150。

4.3.3. 正则化系数 λ 的敏感性分析

正则化系数 λ 控制模型的参数,避免模型的过拟合。正则化系数的敏感性分析在{0.1, 0.5, 1, 1.5, 2}中进行,实验结果如图 3 所示。随着 λ 的增大,NDCG 和 MAP 指标均快速增长后减小,之后趋近稳定。

Recall 虽然也是先增大后减小, 但变化幅度比 NDCG 和 MAP 小, 因此本模型中的正则化参数 λ 取 1。

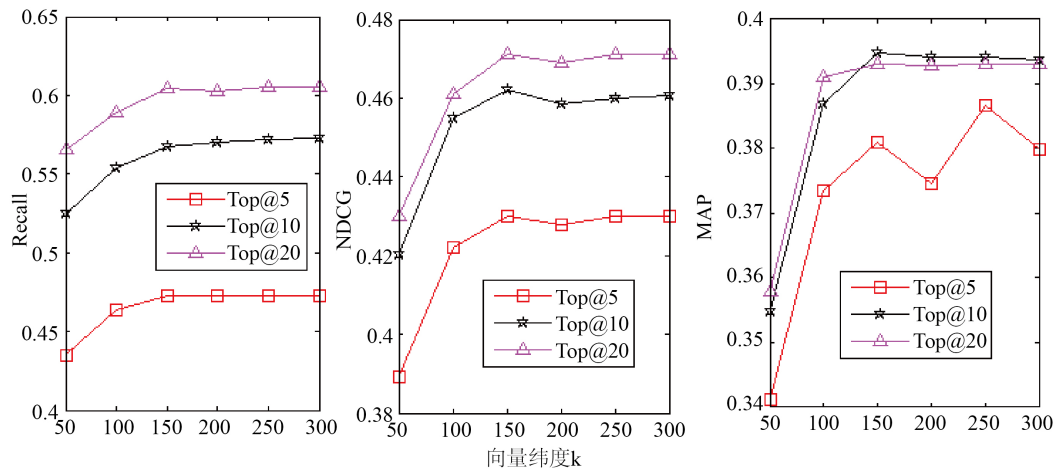


Figure 2. Effect of dimension size
图 2. 向量维度对性能影响

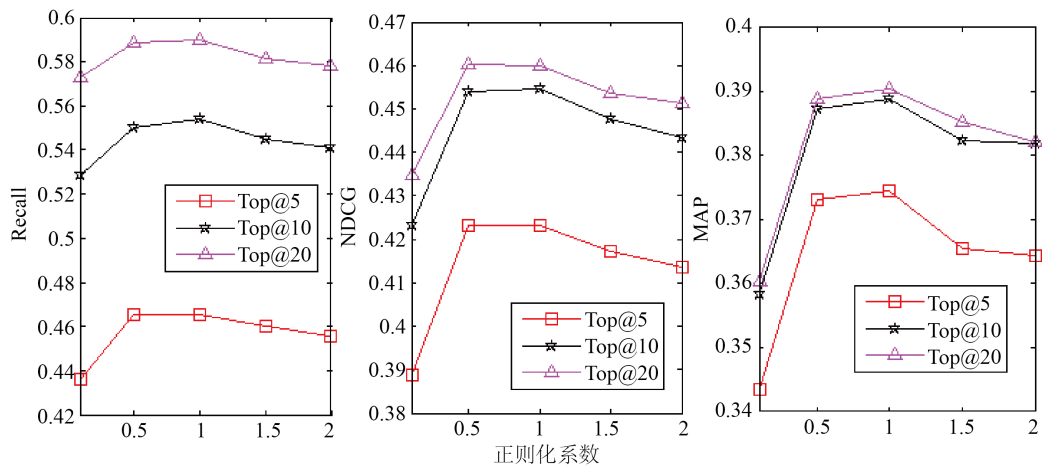


Figure 3. Sensitivity analysis of regularization coefficient
图 3. 正则化系数的敏感性分析

4.4. 实验结果与对比分析

为了更好地评估模型的性能, 将本文的基于嵌入表示和加权矩阵分解模型(Eembedding Weighted Matrix Factorization, EWMF)与以下方法进行了对比: (1)基本加权矩阵分解模型(Weighted Matrix Factorization, WMF): 它是矩阵分解方法的代表, 利用加权矩阵分解模型对用户交互历史矩阵进行分解, 得到旅客和线路的向量表示完成线路推荐; (2) 基于标签的协同过滤推荐(Label based Collaborative Filtering, LCF):它是一种利用丰富额外信息的方法, 它用地域、时间、主题、类型标签丰富旅游数据, 并利用协同过滤进行推荐; (3) 基于知识图谱和长短期偏好的景点推荐(KG-ULSP): 它是表示学习的代表方法, 它用网络表示学习的方法学习景点的特征向量表示, 利用门控循环单元(GRU)学习旅客长短期偏好。

表 1 给出了四种模型在 Recall@N、NDCG@N、MAP@N 线路推荐结果。从表 1 可以看出, 随着 N 的增大, Recall@N、NDCG@N、MAP@N 都增大, 其中 Recall@N 的增长幅度尤为明显。这样的结果验

证了现实中这样的—个事实：即为用户提供越多的线路选择，其中满足用户兴趣的线路就会越多。但是推荐的线路多，用户最感兴趣的线路不一定都能排在推荐列表的前面的最前面。基于网络表示学习的两种方法 EWMF 和 KG-ULSP 的结果要好于 WMF 和 LCF，这是因为利用网络表示学习的方法学习线路和旅客向量表示，更好地利用了数据信息，得到了好的统一表示，而 WMF 和 LCF 融合并充分利用这些信息是困难的。EWMF 要好于 KG-ULSP 说明利用词向量方式加入用户共现、共同喜欢线路对、共同不喜欢线路对的信息，能够得到更好地线路和旅客的向量表示，而 KG-ULSP 只利用了景点的特征信息来学习向量表示。LCF 的性能要优于 WMF，是因为 LCF 很好地利用额外数据信息，而 WMF 只利用了游客与线路的交互矩阵。

Table 1. Comparison EWMF with other methods

表 1. EWMF 与其他方法对比

Method	Recall			NDCG			MAP		
	Top@5	Top@10	Top@20	Top@5	Top@10	Top@20	Top@5	Top@10	Top@20
EWMF	0.509	0.563	0.599	0.468	0.492	0.503	0.401	0.414	0.418
KG-ULSP	0.492	0.543	0.589	0.457	0.482	0.497	0.387	0.404	0.409
LCF	0.479	0.537	0.586	0.441	0.467	0.476	0.384	0.397	0.403
WMF	0.465	0.521	0.554	0.422	0.443	0.453	0.374	0.385	0.388

5. 结语

本文提出了基于嵌入表示和加权矩阵分解的旅游线路推荐算法。通过游客与线路的交互历史，抽取共现游客对、共现喜欢/不喜欢线路对信息，并利用词向量模型得到游客和线路的向量表示，在一定程度上缓解了数据稀疏的问题；接着将共现游客对、喜欢/不喜欢线路对及游客与线路的交互历史利用加权矩阵分解模型进行联合分解，最后预测游客对线路的偏好，丰富了可利用信息。通过在厦航航空公司某旅游公司的真实数据集上的实验结果表明提高了推荐算法的性能，推荐的线路更符合游客的偏好。

基金项目

国家自然科学基金资助项目(U1533104)。

参考文献

- [1] Devasanthiva, C., Vigneshwari, S. and Viverk, J. (2016) An Enhanced Tourism Recommendation System with Relevancy Feedback Mechanism and Ontological Specifications. *Advances in Intelligent Systems and Computing*, **398**, 281-289. https://doi.org/10.1007/978-81-322-2674-1_28
- [2] Cui, G. and Wang, X. (2017) MaP2R: A Personalized Maximum Probability Route Recommendation Using Collaborative Filtering Based on GPS Trajectories. *Processing of 21st Pacific Asia Conference on Knowledge Discovery and Data Mining*, Jeju, 23-26 May 2017, 168-180. https://doi.org/10.1007/978-3-319-57529-2_14
- [3] Su, H., Zheng, K., Huan, J., Liu, T., Wang, H. and Zhou, X. (2014) A Crowd-Based Route Recommendation System-Crowd Planner. *Proceedings of 2014 IEEE 30th International Conference on Data Engineering*, Chicago, 31 March-4 April 2014, 1178-1181. <https://doi.org/10.1109/ICDE.2014.6816735>
- [4] 张艳. 基于空间容量统计的生态旅游路线自动推荐方法[J]. 辽东学院学报(自然科学版), 2020, 27(1): 32-36.
- [5] 韩艳, 杨光, 武鑫森, 等. 考虑游客拥挤感知的旅游线路优化设计[J]. 北京工业大学学报, 2018, 44(12): 71-80.
- [6] 陈雄, 李爱光, 陈冰, 等. 一种基于改进蚁群算法和顾及时间上下文的个性化旅游线路规划方法[J]. 测绘科学技术学报, 2019, 36(2): 215-220.
- [7] Duan, Z., Gao, Y., Feng, J., et al. (2020) Personalized Tourism Route Recommendation Based on User's Active Inter-

ests. *Processing of 21st IEEE International Conference on Mobile Data Management (MDM)*, Versailles, 30 June-3 July 2020, 322-327. <https://doi.org/10.1109/MDM48529.2020.00071>

- [8] Zhu, X.Y., Hao, R.P., Chi, H.T., *et al.* (2017) FineRoute: Personalized and Time-Aware Route Recommendation Based on Check-Ins. *IEEE Transactions on Vehicular Technology*, **66**, 10461-10469. <https://doi.org/10.1109/TVT.2017.2764999>
- [9] 朱桂祥, 曹杰. 基于主题序列模式的旅游产品推荐引擎[J]. 计算机研究与发展, 2018, 55(5): 920-932.
- [10] 肖春景, 夏克文, 乔永卫, 张宇翔. 基于动态聚类的旅游线路推荐[J]. 计算机应用, 2017, 37(8): 2395-2400.
- [11] 李雅美, 王昌栋. 基于标签的个性化旅游推荐[J]. 中国科学技术大学学报, 2017, 47(7): 547-555.
- [12] 丁恒, 黄全舟. 基于属性特征的个性化旅游推荐算法研究[J]. 智能计算机与应用, 2020, 10(1): 193-196.
- [13] 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.
- [14] 李晓旭, 于亚新, 张文超, 王磊. Coteries 轨迹模式挖掘及个性化旅游路线推荐[J]. 软件学报, 2018, 29(3): 587-598.
- [15] 崔春生, 王雪, 李文龙. 基于用户在线评论的旅游景点推荐算法研究[J]. 系统科学与数学, 2020, 40(6): 1103-1116.
- [16] 潘禄生. 基于 LBS 的个性化旅游线路推荐算法研究[J]. 电子测试, 2018, 402(21): 59-61.
- [17] 贾中浩, 宾辰忠, 古天龙, 等. 基于知识图谱和用户长短期偏好的个性化景点推荐[J]. 智能系统学报, 2020, 15(5): 990-997.