一种基于系统主题挖掘的协同过滤算法

李 雪 高心丹

(东北林业大学 信息与计算机工程学院 哈尔滨 150040)

E-mail: gaoxd@ nefu. edu. cn

摘 要:针对传统推荐算法在近邻寻找时忽略了系统自身群组特性的问题。依据项目内容与分类标签间的高关联性,对项目类别标签进行聚类。提出一种系统主题生成算法。基于同一时间段内相同主题的项目具有较高相似性的思想。在计算类内项目相似度时考虑了项目的评分和时间相似性。对于跨主题分布的项目。在评分预测阶段引入了主题偏重系数对类内评分进行加权计算。根据上述理论思想对传统的协同过滤推荐算法进行改进。提出一种基于系统主题挖掘的协同过滤算法。实验结果表明。该算法对传统算法中存在的问题进行了改善且推荐精度有所提高。

关键词: 推荐系统;协同过滤;聚类;主题挖掘;相似度

中图分类号: TP391 文献标识码: A 文章 编号: 1000-1220(2018) 04-0664-04

Collaborative Filtering Recommendation Algorithm Based on Theme Mining

LI Xue ,GAO Xin-dan

(School of Computer and Information Engineering Northeast Forestry University ,Harbin 150040 ,China)

Abstract: Concerning the problem that the traditional recommendation algorithms ignore the organization of the system in neighbor-finding according to that the contents of an item have high similarity with its genres Affinity Propagation Clustering was introduced when calculating new classification based on genres and a method for generating system themes was proposed. The algorithm respectively based on score and time similarity to measure item similarity in the class. For multiple-theme items the theme weight factor was introduced in the prediction. Based on the above points an improved algorithm based on theme mining was proposed. The experimental results show that the disadvantages existing in traditional algorithms get improved and the improved algorithm has better recommendation effects.

Key words: recommendation system; collaborative filtering; clustering; theme mining; similarity

1 引言

随着互联网信息与网络服务(如:电影、音乐)的爆炸式 增长 信息过载成为一大难题 个性化推荐系统应运而生 吃 的出现有效地提高了信息处理效率和服务质量. 当前在电子 商务[1]、餐饮[2]、交通运输[3]等领域都存在着各种形式的推 荐系统. 协同过滤算法^[3] 是众多推荐系统中应用最广泛、最 成功的推荐技术之一 在众多领域得到了应用. 然而随着系统 数据多维化、稀疏化情况的加剧,近邻寻找的复杂度急剧增 加 推荐系统性能显著下降. 为了解决这些问题 ,很多学者采 用聚类算法来提高推荐效果[4 67]. 由于项目分类标签与项目 内容有较高的相关性 独立于评分数据的系统本身就存在一 种自然的群组 基于信息缺失的稀疏矩阵的聚类虽然能够提 高推荐精度 但存在聚类结果解释性较差、聚类数量的选择带 有主观性、聚类空间较大等问题. 本文通过更新聚类算法、改 变聚类对象 生成可解释性较高的系统主题分类以改善传统 方法的不足 同时兼顾项目的时间和评分属性寻找近邻. 在解 决项目的主题跨度问题方面,本文引入一种主题偏重系数来 衡量系统内项目的内容倾向 最后得到相应主题下用户对未 评分目标项目的偏好 并依据偏好度排序得到推荐结果.

2 相关工作

2.1 项目属性对推荐的影响

传统推荐算法是基于用户项目评分矩阵计算项目相似度 或用户相似度 由于用户项目评分数据是一种有缺失值的、稀 疏的数据 在传统相似度计算过程中会产生一定量的误差 ,这 种误差会降低推荐系统的性能. 为了提高推荐性能一些研究 者采用矩阵填充的方法来改善相似度计算的数据基础,如邓 爱林等[8] 通过对用户评分并集中的未评分项目进行预测评 分达到局部填充的效果,在此基础上进行相似度计算.张锋 等^[9] 利用 BP 神经网络算法对用户未评分项目进行预测评 分 增强数据密度 基于填充矩阵进行相似度计算. 然而填充 过程存在预测误差并且计算复杂度较高. 也有一些研究者通 过降维技术增加数据紧密度 如 Sarwar B M 等[11] 通过奇异 值分解减少项目空间维数 在降维后的紧密数据上计算用户 关联度 但降维会导致信息丢失 对推荐系统的影响不一定是 正面的. 上述文献中相似度计算是以应用填充或降维技术处 理后的用户评分矩阵为基础 在项目或用户关系中仅仅考虑 了用户评分的影响,计算得到的相似度是一种单一相似度,在 刻画项目关联性时具有片面性、局限性 ,所以又有一些学者在

收稿日期: 2017-03-29 收修改稿日期: 2017-04-14 基金项目: 中央高校基本科研业务费专项基金 E 类项目(2572014EB05) 资助; 国家自然科学基金项目(61300098) 资助. 作者简介: 李 雪 ,男 ,1991 年生 ,硕士研究生 ,CCF 会员 ,研究方向为社交网络、数据挖掘; 高心丹(通信作者) ,女 ,1971 年生 ,博士 副教授 ,研究方向为计算机应用技术、3S 技术及理论.

相似度计算过程中增加了项目属性 如刘健等[10] 引入项目标签 分别计算用户与标签、标签与项目的关联度 构建用户兴趣模型. 张新猛等[5] 在相似度计算过程中考虑了项目的评分相似性与类别相似性 提升了项目间的关联度. 上述改进方案增加了项目关联的紧密度、弱化了评分数据稀疏性的影响 提升了推荐质量 ,可见项目属性对于提升关系描述准确度具有积极的作用. 本文在相似度计算过程中融入了项目的评分与时间属性 通过平衡因子协调项目评分相似度与项目时间相似度.

2.2 基干聚类的推荐算法

聚类算法的应用有效的缩小了近邻搜索空间、提升了计算效率、增强了推荐性能,如 Rashid 等[12] 在评分矩阵的基础上进行用户聚类 聚类中心作为聚类的代理用户,然后基于目标用户的最相似代理用户进行最近邻推荐. 邓爱林等[6] 采用 K-means 算法对评分矩阵进行项目的聚类,得到 k 个项目分类,然后在与目标项目最相似的几个类中通过计算评分矩阵的相似度来搜索近邻并排序产生推荐. George 等[13] 使用联合聚类方法基于评分矩阵对项目与用户同时聚类.

上述方法基于评分数据分别对项目、用户进行聚类 然而系统项目评分数据是一种高维的、有缺失的、信息稀疏的数据 影响聚类的准确性;同时,上述聚类算法应用时需要事先人为确定分类数量,存在一定的主观因素,并且产生的聚类中心并非真实存在数据.为了改善上述问题,本文改变了聚类对象,引入一种低维的项目-类别矩阵,降低了聚类空间,并且在聚类阶段应用了一种较新的高效聚类算法—近邻传播聚类算法(AP聚类),该算法不需要事先指定聚类个数,聚类中心是原始数据中真实存在的值,且多次执行结果完全一样.本文利用 AP 算法对项目-类别矩阵进行聚类,从而生成新的分类,聚类中心作为新分类的主题描述.

3 改进的推荐算法

3.1 基于分类标签的 AP 聚类

3.1.1 项目分类标签相似度

分类标签指系统中实际存在的、描述性的类别标注 如果两个分类标签(如喜剧、爱情)在项目分布上趋向一致 ,说明这两个标签在内容表达上具有一定的联动性.

表 1 项目-类别标签矩阵

Table 1 Matrix of item-genre

项目	电影	电影	电影	电影	电影
类型	A	В	C	D	E
喜剧	1	1	0	1	0
爱情	1	1	0	1	0
动作	0	0	1	1	0
战争	0	0	1	0	0
纪录片	0	0	0	0	1

基于上述思想 在系统中构建与表 1 同样的项目-类别标签矩阵. 可以使用 Jaccard 系数计算标签之间的相似性 ,设分类标签 c 和 d 标注的项目集为 I_c 和 I_d 则 c 和 d 之间的相似性 sim(c,d) 如公式(1) 所示:

$$sim(c \mid d) = \frac{|I_c \cap I_d|}{|I_c \cup I_d|}$$
 (1)

通过相似度计算得到的分类标签相似度矩阵作为 AP 聚

类的输入.

3.1.2 AP 聚类

AP 聚类的基本思想^[7]: 给定 N 个数据点,计算数据点两两之间的相似度构造矩阵 S. 在相似度矩阵 S 基础上,依靠两点之间的消息传递迭代更新,即计算吸引度 $R(i \ k)$ 和归属度 $A(i \ k)$ $R(i \ k)$ 和 $A(i \ k)$ 值越大,点 k 作为聚类中心和 i 属于以 k 为聚类中心的簇的可能性越大. 当迭代次数达到上限或收敛程度符合要求时,停止迭代,产生 n 个符合要求的聚类中心,再根据聚类中心与普通点的相似度,将普通点划分到不同的聚类中,完成整个聚类过程.

R(i,k) 和 A(i,k) 的计算公式如下:

$$R(i \ k) = S(i \ k) - \max\{A(i \ k) + S(i \ j)\}$$
 (2)

$$A(i \ k) = \min\{0 \ R(k \ k) + \sum\{\max\{0 \ R(j \ k) \}\}$$
 (3)

$$R(k|k) = P(k) - \max\{A(k|j) + S(k|j)\}$$
 (4)

其中 $j \in \{1 \ 2 \ 3 \ , \cdots \ N \ j \neq k\}$

引入阻尼系数 λ 控制迭代的收敛 ,每次迭代都需要考虑 上一次的吸引度和归属度 ,加权更新如下:

$$R_i = (1 - \lambda) \times R_i + \lambda \times R_{i-1} \tag{5}$$

$$A_i = (1 - \lambda) \times A_i + \lambda \times A_{i-1}$$
 (6)

其中 $\lambda \in [0.5,1)$.

下面给出聚类的具体流程:

Step 1. 计算分类标签相似度矩阵 S S 作为 AP 聚类的输入.

Step 2. 选取参考度阀值 P(P) 为点 i 作为聚类中心的标准 本文取相似度矩阵 S 的中值).

Step 3. 计算吸引度和归属度 根据 R(k,k) + A(k,k) 的 值来判断 k 是否为聚类中心.

Step 4. 若迭代次数达到上限或者聚类中心已经不在发生变化 转 Step 5 否则 返回 Step 3.

Step 5. 将其余点根据相似度划分到各个聚类中,生成新类别.

Step 6. 根据项目初始标签所属的新类别进行项目类别划分,允许跨类分布,聚类结束.

3.2 类内项目相似度

在计算项目相似度时考虑了项目的时间、评分属性 即认为同一主题下同时代分布的项目更加相似.

在计算项目年代相似性时,以年代作为项目时间分布的单位,由于项目具有固定的生产或发布时间。在时间分布上具有唯一性,所以在计算相似度时,相同年代项目间相似度为 1 不同年代项目间的相似度为 0 建立项目—时间相似度矩阵 T(n,n) T_n 值为 0 或 1.

在计算评分相似性时以主题为单位进行类内计算,分别采用余弦相似度、修正余弦相似度、皮尔森相似度系数进行对比实验,没有主题集合 $M = \{M_1, M_2 \cdots M_n\}$ $M_k \in M$ $k \in (1, 2 \cdots n)$ 则项目 i 和项目 j 在主题 M_k 内被共同评分的用户集合为 U^k_{ij} , $U^k_{ij} = U_i \cap U_j \cap M_k$,则 i 和 j 在主题 M_k 下的评分相似性 $sim_r(i\ j)$ 为:

余弦相似度:

$$sim_{r}(i j) = \frac{\sum_{u \in U_{ij}^{k}} (r_{ui} * r_{uj})}{\sqrt{\sum_{u \in U_{ij}^{k}} (r_{ui})^{2}} \sqrt{\sum_{u \in U_{ij}^{k}} (r_{uj})^{2}}}$$
(7)

修正余弦相似度:

$$sim_{r}(i j) = \frac{\sum_{u \in U_{ij}^{k}} (r_{ui} - \bar{r}_{u}) (r_{uj} - \bar{r}_{u})}{\sqrt{\sum_{u \in U_{ij}^{k}} (r_{ui} - \bar{r}_{u})^{2}} \sqrt{\sum_{u \in U_{ij}^{k}} (r_{uj} - \bar{r}_{u})^{2}}}$$
(8)

皮尔森相似度

$$sim_{r}(i j) = \frac{\sum_{u \in U_{ij}^{k}} (r_{ui} - \overline{r}_{i}) (r_{uj} - \overline{r}_{j})}{\sqrt{\sum_{u \in U_{ij}^{k}} (r_{ui} - \overline{r}_{i})^{2}} \sqrt{\sum_{u \in U_{ij}^{k}} (r_{uj} - \overline{r}_{j})^{2}}}$$
(9)

式中 r_i 和 r_j 表示项目在系统内的平均评分 $r_{ui} \cdot r_{uj}$ 表示用户对项目的评分 r_u 表示用户评分的平均值.

本文引入了平衡因子 α 协调项目时间相似性和评分相似性对项目相似性的影响,由此定义的项目间相似性度量公式作为项目i 和项目j 之间类内相似性度量的标准,如公式(10)所示.

$$\operatorname{sim}_{k}(i \mid j) = \alpha \operatorname{sim}_{r}(i \mid j) + (1 - \alpha)T(i \mid j) \tag{10}$$

3.3 产生推荐

3.3.1 主题偏重系数

由于存在跨主题分布项目,所以类内评分不能作为项目最终的评分,所以引入了主题偏重系数 W 作为项目所跨主题的系统偏重。设项目 h 横跨主题集合为 $M=\{M_1,M_2\cdots M_n\}$, $M_k\in M$ k=1 2 $;\cdots n$ 对应的项目集合为 $I_M=I_{M_1}\cup I_{M_2}\cup\cdots\cup I_{M_n}$ 则项目所跨主题对应的偏重系数为:

$$W(M_k) = \frac{|I_{M_k} \cap C_M I_{M_k}|}{|I_{M_k}|}$$
 (11)

式中 $C_M I_{M_k}$ 为 I_{M_k} 在集合 M 中的补集

3.3.2 预测推荐

通过本文提出的相似度计算方法得到目标项目的类内最近邻居后,下一步需要产生相应的推荐. 设目标项目 x 的最近邻居集合用 NN_x = $\{NN_1,NN_2,\cdots,NN_n\}$ 表示,则用户 u 对项目 x 的类内预测评分 $P_{(u,x)}^{M_k}$ 可以通过用户 u 对最近邻居集合 NN_x 中项目的评分得到.

由于系统存在单一主题项目和多主题项目,所以分两种 方式进行计算.

对于单一主题项目,其最后评分与类内预测评分 $P^{M_k}_{(u,x)}$ 相等:

$$P_{(u,x)} = P_{(u,x)}^{M_k} = \bar{R}_x + \frac{\sum_{n \in NN_x} sim(x,n) * (R_{u,n} - \bar{R}_n)}{\sum_{n \in NN_x} (|sim(x,n)|)} (12)$$

式中 M 为项目 x 对应的主题分类 R_x 为项目 x 的平均分 R_n 为近邻项目 n 的平均分. 对于多主题项目 引入主题偏重系数 对项目所跨主题的类内评分进行加权计算:

$$P_{(u,x)} = \frac{\sum_{M_k \in M} W(M_k) * P_{(u,x)}^{M_k}}{\sum_{M_k \in M} W(M_k)}$$
(13)

 $P_{(u,v)}^{M_k}$ 为类内预测评分 $W(M_k)$ 为主题 k 对应的偏重系数.

4 实验结果及分析

4.1 实验数据集及度量标准

实验数据采用 GroupLens 站点(http://www.grouplens.

org/) 提供的 MovieLens 数据集 GroupLens 项目组提供了不同大小的数据集 本文采用 ml-100k 数据集 该数据集包含 943 位用户对 1628 部电影的 1000 000 个评分 且评分范围为 1-5 每个用户至少对 20 部以上的电影进行了评分 数据稀疏度为 93.7%.

实验采用的数据集包含 1682 部电影、19 种电影类型标签 忽略掉仅含 2 部电影且信息缺失的 unknown 类型 数据集更新为 1680 部电影 18 种类型. 18 个电影类型依次对应 0、1、…、17 ,一部电影对应有 1-7 个类型标签. 电影发行时间的时间轴为 1922-1998 年 横跨 8 个年代 ,以年代为单位对电影进行划分 ,年代类型依次对应 0、1、…、7. 具体实验取数据集中u. data 的评分子集: 随机选取 400 名用户作为实验数据集 ,随机选取 80% 数据作为训练集 ,20% 数据作为测试集 ,训练集稀疏度为 94.0%.

本文使用 MAE 方法^[14] 度量预测准确度 ,主要通过计算测试集中用户的预测评分与实际评分之间的平均偏差来度量预测算法的准确性 ,可以直观地度量推荐质量 ,MAE 越小 ,推荐质量越高. MAE 定义为:

$$MAE = \frac{\sum_{i=1}^{N} |p_i - q_i|}{N}$$
 (14)

式中 p_i 为预测评分、 q_i 为真实评分、N 为测试集数量.

4.2 实验分析

算法中有 2 个参数需要选择: 聚类算法中的参考度 P、相似度平衡因子 α 参考度 P 取项目标签相似度的中值 平衡因子 α 通过实验进行选取.

表 2 聚类结果 Table 2 Result of clustering

类别	组 成	描述	
1	聚类中心: Action		
	标签: Action、War、Horror、Western、Sci-	偏打斗场景	
	Fi, Thriller, Adventure		
2	そ类中心: Children´s		
	标 签: Children´s、Animation、Fantasy、	适合儿童	
	Musical		
3	聚类中心: Documentary	纪录片	
	标签: Documentary	-03071	
4	聚类中心: Film-Noir	社会阴暗面	
	标签: Film-Noir、Crime、Mystery		
5	聚类中心: Romance	浪漫、轻松	
	标签: Romance、Drama、Comedy		

AP 聚类算法对项目分类标签进行聚类 ,产生 5 个聚类 集 ,如表 2 所示.

实验结果表明 融合评分相似度和时间相似度的算法优于单一相似度算法($\alpha = 0.0 \ 1.0$)的预测效果. 随着 α 的增加 MAE 呈递减趋势,评分越来越接近实际评分. 在实验对比阶段,选择对应相似度下的最优 α 值进行实验.

4.3 实验对比

为验证本文算法的有效性,设计实验与传统协同过滤算法的预测性能进行对比.最早在MovieLens数据集上应用传统协同过滤算法的是GroupLens研究小组的Sarwar教授等人.

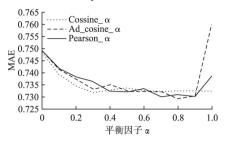


图 1 平衡因子对 MAE 的影响

Fig. 1 Influence of α to MAE

在文献[14]中,Sarwar 教授等人通过实验得到 Item-based 方法的推荐效果要略好于 User-based 方法的结论,并且对传统 Item-based 方法在 MovieLens 数据集上的性能进行了评估 如图 2 所示.

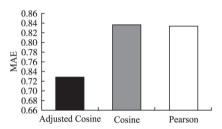


图 2 传统协同过滤的预测效果

Fig. 2 Effect of traditional collaborative filtering

传统 Item-based 方法在不同相似度下预测效果不同,其中基于修正余弦相似度(AdjustedCosine) Item-based 方法表现最佳. 本文将与文献[14]中传统协同过滤在 MovieLens 数据集上的最佳预测效果即基于修正余弦相似度 Item-based 方法的 MAE 值进行对比实验 如图 3 所示.

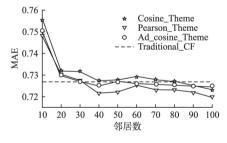


图 3 预测精度比较

Fig. 3 Accuracy comparison of algorithm

从预测精度的比较可以看出,本文提出的方法在 Cosine 度量标准下,邻居数量大于 80 ,MAE 值较优;在 Adjusted-Cosine 度量标准下,邻居数量大于 30 ,MAE 值较优;在 Pearson 度量标准下,邻居数量大于 30 ,MAE 值较优,且整体预测效果最优.从对比实验可以看出 本文提出的基于系统主题挖掘的协同过滤算法对推荐效果的准确性有明显的提升作用.

5 结 论

针对传统推荐算法在近邻寻找时忽略了系统本身的群组

特性的问题 本文引入了系统主题模型 统分挖掘系统项目本身的特点 改善了传统算法对评分矩阵的依赖性 缓解了新物品的初始群组问题.实验表明本文提出的方法可以显著提高推荐算法的准确性.本文方法虽然一定程度上缓解了数据稀疏问题对推荐精度的影响,但仍存在冷启动、邻居稀疏等问题 因而未来将会重点研究结合矩阵填充、图论等知识的推荐算法以降低推荐算法的局限性.

References:

- [1] Ai Dan-xiang Zuo Hui , Yang Jun. C2C e-commerce recommender system based on three-dimensional [J]. Computer Engineering and Design 2013 34(2):702-706.
- [2] Xiong Cong-cong Deng Ying Shi Yan-cui et al. Food recommendation algorithm based on collaborative filtering [J]. Application Research of Computers 2017 34(7):1-5.
- [3] Shao Kuo-yi Ban Xiao-juan Wang Xiao-kun et al. Geographic information recommender system optimized by transportation network data [J]. Chinese Journal of Engineering 2015 37(12):1651-1657.
- [4] Li Tao ,Wang Jian-dong ,Ye Fei-yue. Collaborative filtering recommendation algorithm based on clustering basal users [J]. System Engineering and Electronics 2007 29(7):1177-1178.
- [5] Zhang Xin-meng Li Song. Collaborative filtering recommendation algorithm considering data weight and item attributes [J]. Automation and Instrumentation 2016 36(9):69-73.
- [6] Deng Ai-lin , Zuo Zi-ye , Zhu Yang-yong. Collaborative filtering recommendation algorithm based on clustering [J]. Mini-Micro Systems 2004 25(9):1665-1670.
- [7] Hu Ji-hua Cheng Zhi-feng Zhan Cheng-zhi. AP algorithm for public transportation station clustering [J]. Computer Engineering , 2011 37(S1): 223-225 + 232.
- [8] Deng Ai-lin ,Zhu Yang-yong ,Shi Bo-le. A collaborative filtering recommendation algorithm based on item rating prediction [J]. Journal of Software 2003 ,14(9):1621-1628.
- [9] Zhang Feng Chang Hui-you. Employing BP neural networks to alleviate the sparsity issue in collaborative filtering recommendation algorithms [J]. Journal of Computer Research and Development, 2006 49(4):667-672.
- [10] Liu Jian Zhang Kun Chen Xuan. Personalized recommendation algorithm based on tags and collaborative filtering [J]. Computer and Modernization 2016 32(2):62-65+71.
- [11] Sarwar B ,Karypis G ,Konstan J ,et al. Application of dimensionality reduction in recommender system-a case study [R]. Minnesota Univ ,Minneapolis Dept. of Computer Science 2000.
- [12] Rashid A M Lam S K Karypis G et al. ClustKNN: a highly scalable hybrid model—& memory-based CF algorithm [C]. Proceeding of Web Knowledge Discovery and Data Mining (Web KDD) 2006.
- [13] George T ,Merugu S. A scalable collaborative filtering framework based on co-clustering [C]. IEEE 5th International Conference on Data Mining (ICDM) 2005: 625-628.
- [14] Sarwar B ,Karypis G ,Konstan J ,et al. Item-based collaborative filtering recommendation algorithms [C]. Proceedings of the 10th International World Wide Web Conference (WWW) 2001: 285-295.

附中文参考文献:

- [1] 艾丹祥 左 晖 杨 君. 基于三维协同过滤的 C2C 电子商务推 荐系统 [J]. 计算机工程与设计 2013 34(2):702-706.
- [2] 熊聪聪,邓 滢,史艳翠,等. 基于协同过滤的美食推荐算法[J]. 计算机应用研究 2017 34(7):1-5.
- [3] 邵阔义 班晓娟 汪笑琨 筹.基于交通网络数据优化的地理信息 推荐系统[J]. 工程科学学报 2015 37(12):1651-1657.
- [4] 李 涛 汪建东,叶飞跃,一种基于用户聚类的协同过滤推荐算法[J]. 系统工程与电子技术 2007 29 (7):1177-1178.
- [5] 张新猛 李 松.基于项目属性与数据权重的协同过滤推荐算法 [J].自动化与仪表 2016 36(9):69-73.
- [6] 邓爱林,左子叶,朱扬勇.基于项目聚类的协同过滤推荐算法 [J]. 小型微型计算机系统 2004 25(9):1665-1670.
- [7] 胡继华,程智锋,詹承志. 一种用于公交站点聚类的 AP 算法 [J]. 计算机程 2011 37(S1):223-225+232.
- [8] 邓爱林 朱扬勇 施伯乐.基于项目评分预测的协同过滤推荐算法[J].软件学报 2003 ,14(9):1621-1628.
- [9] 张 锋 常会友. 使用 BP 神经网络缓解协同过滤推荐算法的稀疏性问题[J]. 计算机研究与发展 2006 49(4):667-672.
- [10] 刘 健 涨 琨 陈 旋 基于标签和协同过滤的个性化推荐算 法[J]. 计算机与现代化 2016 32(2):62-65+71.