

# 融合共同评分数量和质量的协同过滤个性化推荐算法

蔡依芳, 艾均, 苏湛

(上海理工大学光电信息与计算机工程学院, 上海 200093)

✉worldchinacai@126.com; aijun@outlook.com; suzhan@foxmail.com



**摘要:** 为了改善经典协同过滤个性化推荐算法准确性不高、推荐列表排序性能和偏好分类预测精确性低的缺点, 综合考虑了用户共同评分物品数量及质量, 通过降低共同评分物品数量较少或质量不高的邻居用户权重, 设计了一种用户相似性度量方法进行协同过滤和评分预测。通过实验, 在MovieLens电影评分数据集上与领域内几个典型协同过滤算法对比, 研究发现本文设计的融合用户共同评分数量和质量的协同过滤个性化推荐算法可以将预测误差降低8.41%, 将推荐列表排序性能提高10.21%, 将偏好分类预测准确率提高2.55%。

**关键词:** 推荐算法; 评分预测; 相似性; 共同评分

**中图分类号:** TP182 **文献标识码:** A

## A Collaborative Filtering Personalized Recommendation Algorithm Integrating the Quantity and Quality of Common Rating

CAI Yifang, AI Jun, SU Zhan

(School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

✉worldchinacai@126.com; aijun@outlook.com; suzhan@foxmail.com

**Abstract:** Traditional collaborative filtering personalized recommendation algorithm has the shortcomings of low accuracy, low ranking performance of recommendation list and low accuracy of preference classification prediction. Aiming at these problems, this paper proposes to design a user similarity measurement method for collaborative filtering and rating prediction by reducing the weight of neighbor users with small number of common rating items or low quality of common rating, after comprehensively considering the quantity and quality of users' common rating items. Through experiments, compared with several typical collaborative filtering algorithms in the field on MovieLens rating dataset, it is found that the proposed collaborative filtering personalized recommendation algorithm, which incorporates the quantity and quality of user co-ratings, can reduce the prediction error by 8.41%, improve the performance of recommendation list sorting by 10.21%, and improve the accuracy of preference classification prediction by 2.55%.

**Keywords:** recommendation algorithm; rating prediction; similarity; common rating

### 1 引言(Introduction)

近年来, 信息技术与移动网络的飞速发展使得互联网每天都产生海量信息, 在网络上畅游的用户在寻求满足自身需求时必然面临过多的选择。大量相似的选择和质量未知的信息增加了用户的困惑, 让选择更加困难, 这就导致用户不仅没能从海量数据中获取便利, 还要花费大量的时间进行目标筛选, 这种现象被称为信息过载<sup>[1]</sup>。

推荐系统<sup>[2]</sup>是解决信息过载问题的一个有效方案, 它通过学习用户历史行为偏好主动预测用户对未评分或反馈的物品潜在的反应, 并将那些预测评分较高或偏好正面的物品, 通过推荐列表等形式自动化、个性化地推荐给不同用户, 从而帮助他们更快速地找到需要或喜欢的物品。这种技术的成功应用有效提升了用户对系统的好感, 也可以帮助商家寻找对某物品感兴趣的潜在用户群体, 既解决了用户的困扰, 又

为商家创造了价值。因此，推荐系统目前已被广泛应用于多个领域，例如电子商务平台、在线流媒体网站、电影推荐网站、在线音乐网站或应用、社交网络平台，等等<sup>[3]</sup>。

本文通过分析推荐算法过程中用户共同评分物品的数量和质量问题，设计用户间的相似性度量方法和对应的协同过滤个性化推荐算法，并通过在真实的电影评分数据集上实验以验证算法的有效性。

## 2 推荐算法分类(The classification of recommendation algorithms)

### 2.1 算法分类

个性化推荐算法通常分为三类：基于内容的推荐算法<sup>[4]</sup>、基于内存和模型的协同过滤算法<sup>[5-6]</sup>以及混合推荐算法<sup>[7]</sup>。

三种算法存在不同的使用限制。(1)基于内容的推荐算法通过分析物品的内容信息(新闻、图书等)预测或形成推荐，但其非常依赖内容信息的数据质量，也非常容易给出那些显而易见的推荐，但推荐的准确性和多样性受限。(2)协同过滤算法只要系统提供用户对物品的评分信息集合就可以工作。其中，基于内存的协同过滤利用相似用户打分的习惯，使用 $k$ 近邻算法筛选邻居并预测未知评分；基于模型的协同过滤，例如贝叶斯分类器<sup>[8]</sup>、模糊算法<sup>[9]</sup>和矩阵分解模型<sup>[10]</sup>，使用历史评分数据训练不同的模型用于推荐或预测。(3)混合推荐算法则融合不同类型的算法，试图结合它们各自的优点进行推荐，但其计算复杂性更高，推荐结果可解释性较差。

### 2.2 经典算法及其问题

在各类算法中最知名的是协同过滤算法，它的出现标志个性化推荐系统的诞生。作为当今应用最为广泛的推荐算法，协同过滤算法较其他算法复杂度更低，可解释强，部署方便，可以有效缓解信息过载问题。

所有协同过滤算法的目标都是尽可能地减小预测误差，提高推荐列表多样性，增强推荐列表的排序质量。但是，受到冷启动和稀疏性这两个推荐系统固有问题的影响<sup>[11]</sup>，传统的基于用户或基于物品的协同过滤相似性计算方法(例如，Cosine相似性<sup>[12]</sup>和Pearson相似性<sup>[13]</sup>)总是因冷启动问题使新用户没有邻居参考，或因稀疏性问题导致评分数据不足，选择的邻居不准确，预测结果不理想。对于这些问题，许多研究人员提出了相应的改进方法，例如，NIKOLAOS等人<sup>[14]</sup>将Pearson相似性值进行分级处理以增加相似性高邻居的权值；HE等人利用观点传播的理论<sup>[15]</sup>，通过归一化之后的评分来计算推荐系统用户之间的相互影响，并以此为依据进行预测；基于信息熵的协同过滤<sup>[16]</sup>将熵的概念与传统的相似性度量相结合以提高推荐系统的性能，等等。

这些方法在一定程度上确实提高了预测的准确性，但是，仅考虑用户共同评分数值差异是不全面的，也应同时考

虑利用用户之间共同评分物品集合中物品的数量。因为即便两个用户对相同物品评分差异较大，但只要这两个用户具有一个较大的共同评过分的物品集合，就可以推测这两个用户间的相似性较高。

更进一步来说，还应该考虑这个共同评分物品集合的质量，以区分共同评分物品集合大致相等时，集合中不同物品特征对用户相似性的影响。如果用户的共同评分集合中热门物品较多，其相似性应低于共同评分集合中以冷门物品为主的用户，因为那些喜欢小众物品的用户之间显然更具有偏好上的相似性。

综上所述，本文研究共同评分在总评分数中占比的问题和用户共同评分物品中热门物品的影响度量问题。设计融合共同评分数量和质量的相似性度量，以及在此基础上的个性化推荐算法，以实现提高推荐系统性能的目标。

## 3 算法设计(Algorithm design)

### 3.1 基于共同评分数值度量用户相似性

协同过滤推荐的核心假设是相似的用户具有相似的偏好，在此基础上，可以选择相似性较高的邻居，利用这些邻居对目标物品的评分来预测目标用户对该物品的评分。在这类算法中，计算用户间的相似性是算法学习过程的核心步骤。两个用户相似性越大，该邻居用户在推荐的过程中的影响就越大。一般来说，计算用户相似性的主要依据就是两个用户对同一集合中物品评分的差异。理论上，评分数值差异越小，两个用户的偏好和品味就越相似；反之，数值差异越大，偏好差异也就越大。

本文采用Pearson相关系数来度量两个用户间基于共同评分物品集合的评分数值，以度量相似性的一部分权重。Pearson相关系数是统计学上求两个变量关系的方法，可以用来度量两个用户对相同物品集合的评分差异。具体说来，Pearson相关系数对用户评分的偏差进行加权平均，结果作为用户或物品的相似性，且其取值范围为-1.0至1.0，计算公式如式(1)所示：

$$PCC(u, v) = \frac{\sum_{i \in I_u \cap I_v} (r_{ui} - \bar{r}_u) \cdot (r_{vi} - \bar{r}_v)}{\sqrt{\sum_{i \in I_u \cap I_v} (r_{ui} - \bar{r}_u)^2} \cdot \sqrt{\sum_{i \in I_u \cap I_v} (r_{vi} - \bar{r}_v)^2}} \quad (1)$$

其中， $I_u$ 和 $I_v$ 分别是用户 $u$ 和 $v$ 评分的物品集， $r_{ui}$ 是用户 $u$ 对物品 $i$ 的已知评分， $\bar{r}_u$ 是用户 $u$ 的平均评分。

### 3.2 融合共同评分集合数量和质量的相似性度量

使用Pearson相关系数计算相似性时，考虑了两个用户选择过的物品的交集 $I_u \cap I_v$ 的评分信息计算相似性，但其并没有考虑每两个用户的共同评分集合中元素数量的差异，以及该集合中热门物品的情況特征。

共同评分集合元素数量多的用户间相似性应该大于共同

评分数量少的用户，但如果两个用户是共同评分物品集合中物品较为热门的情况，则他们的相似性应低于相似条件下物品集合中物品较为冷门的情况。本文采用杰卡德相似性来度量两个用户因共同评分物品集合元素数量而决定的相似性权重。

Jaccard系数公式如式(2)所示，它计算了用户 $u$ 和用户 $v$ 都评过分的物品个数与两个用户评过分的物品总个数之比。该值越大，说明这两个用户之间共同评过分的物品集合与两人评过评分物品的总数之间的比例越高，说明在不考虑评分数值的情况下，这两个用户对物品的选择是相似的。

$$Jaccard(u,v) = \frac{|I_u \cap I_v|}{|I_u \cup I_v|} \quad (2)$$

为了进一步度量用户共同评分物品集合中热门物品的特征，即用户共同评分集合的质量，本文设计了如式(3)所示的 $quality(u,v)$ 系数以降低含有热门物品的共同评分集合的权重。其中，共同评分物品集合中每个物品 $i$ 被训练集中所有用户选择的次数记为物品度值 $d_i$ ，给定共同评分物品集合中热门物品特征，由该集合所有物品 $d_i$ 的最大值表示。基于式(3)，如果两个用户共同评分物品集合中存在一个非常流行，即度值很高的物品，则这两个用户之间的共同评分物品集合的质量就相对较低，其权值在相似性计算过程中被降低。

在此基础上，融合共同评分集合数量和质量的相似性由式(4)度量，称该系数为共同评分数量和质量系数(Common Rating Quantity and Quality, CRQQ)。

$$quality(u,v) = \frac{|I_u \cap I_v|}{\max(d_{i \in I_u \cap I_v}) + |I_u \cap I_v|} \quad (3)$$

$$CRQQ(u,v) = \frac{Jaccard(u,v) + quality(u,v)}{2} \quad (4)$$

最终，本文设计的综合考虑共同评分物品数值差异与集合数量和质量的相似性计算公式如式(5)所示，即两个用户间评分差异小，并且共同评分物品集合数量和质量均较高时，两个用户最相似。如共同评分物品数量、共同评分物品质量或共同评分数值差异三者当中只有一项到两项较高时，其权重被降低，以便更加准确地度量不同用户间的特征差异。

$$Sim(u,v) = PCC(u,v) \cdot CRQQ(u,v) \quad (5)$$

### 3.3 未知评分预测方法

基于相似性计算结果，以及按照相似性从高到低选择的 $k$ 个邻居，本文使用式(6)把邻居用户 $v$ 对预测目标物品 $i$ 的评分和该邻居平均评分的差值做加权计算，基于该邻居和目标用户相似性，来实现预测用户对未知物品的评分。

$$\tilde{r}_{ui} = \bar{r}_u + \frac{\sum_{v=1}^k (r_{vi} - \bar{r}_v) Sim(u,v)}{|Sim(u,v)|} \quad (6)$$

## 4 实验设计(Experimental design)

### 4.1 数据集与基准算法

本文使用MovieLens数据集进行对比实验，该数据集有610个用户对9,742部电影的100,836条评分信息。为了保证实验的有效性，实验采用折十交叉验证进行，即数据集被分成10份，每次实验取其中一份作为测试集，另外9份作为训练集，共进行10次实验，最终取10次结果的平均值作为最终实验结果进行比较。

为了验证本文设计的CRQQ算法，用来对比的基准算法有三个，分别是：多层级协同过滤(Multi-Level Collaborative Filtering, MLCF)<sup>[14]</sup>、用户观点传播算法(User Opinion Spreading, UOS)<sup>[15]</sup>以及使用了信息熵的Pearson算法(Pearson Entropy)<sup>[16]</sup>。

### 4.2 评价指标

本文使用平均绝对误差(Mean Absolute Error, MAE)、均方根误差(Root Mean Squared Error, RMSE)衡量算法对未知评分的预测误差，使用半衰期(Half Life Utility, HLU)度量推荐列表排序质量，采用Accuracy和F1比较各个算法对用户偏好分类的预测准确率。

MAE和RMSE公式如式(7)和式(8)所示，其计算由算法预测值和真实值的差值加权构成，因此两个误差度量参数均为越低越好。

$$MAE = \frac{1}{n} \sum_{i=1}^n |\tilde{r}_{ui} - r_{ui}| \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\tilde{r}_{ui} - r_{ui})^2} \quad (8)$$

HLU代表算法的排序能力，算法将用户喜欢的物品排在前面的能力越好，HLU值越大，其计算方法如式(9)所示：

$$HLU = \frac{1}{m} \sum_{i=1}^m \frac{\max(r_{ui} - \bar{r}_u, 0)}{2^{(i-1)}} \quad (9)$$

Accuracy和F1指标是推荐系统分类预测指标，这两个指标的计算公式如式(10)一式(13)所示：

$$Precision = \frac{TP}{TP + FP} \quad (10)$$

$$Recall = \frac{TP}{TP + FN} \quad (11)$$

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (12)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (13)$$

其中，在一次预测中， $TP$ 代表算法推荐了用户喜欢的物品的总体数量， $FP$ 代表算法推荐了用户不喜欢的物品的总体数量， $FN$ 代表系统未推荐用户喜欢的物品的总体数量， $TN$ 代表系统未推荐用户不喜欢的物品的总体数量。这几个参数的数值

越高，算法对用户偏好分类(喜欢或不喜欢的)预测就越准确。

### 5 实验结果与分析(Experimental results and analysis)

图1展示了四种算法随邻居数变化的MAE误差结果。邻居数为1时，CRQQ算法比Pearson Entropy算法误差低6.68%。随着邻居数的增加，所有算法误差都随之降低，但是CRQQ算法总是能在最少的邻居数下最先达到最小误差值，其次是UOS、Pearson Entropy、MLCF，最终CRQQ算法将准确率提高了0.57%—6.68%。结果表明，本文设计的算法能够更快地选取最合适的邻居，从而得到最精准的预测结果。

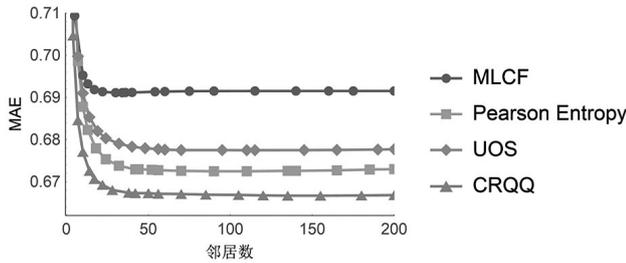


图1 几种算法的MAE结果对比

Fig.1 Comparison of MAE results of different algorithms

图2展示了四种算法的RMSE误差对比。RMSE对大误差加大了惩罚，此时，CRQQ算法的误差依旧是所有算法中最低的，这代表CRQQ算法预测结果很少出现大误差，结果最准确。随着邻居数增加，CRQQ算法误差结果迅速下降，优势更加明显，分别比Pearson Entropy误差降低了0.48%—8.41%，比UOS误差降低了1.40%—3.24%，比MLCF误差降低了0.92%—3.23%。由此可知，CRQQ算法无论是大误差还是小误差都比其他算法少，进一步说明本文在预测准确性上的改善能力。

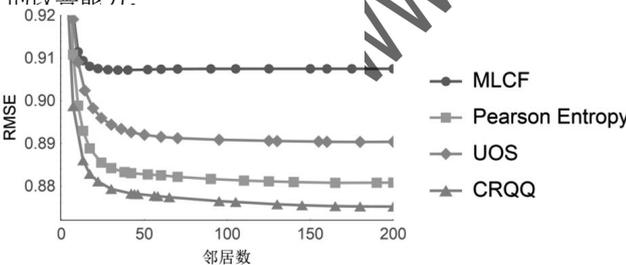


图2 不同算法的RMSE结果对比

Fig.2 Comparison of RMSE results of different algorithms

图3是四种算法的HLU排序能力指标，值越大说明算法排序能力越好。总体来说，所有算法的HLU得分前期增长明显，后期略微下降。但不变的是，CRQQ得分一直是第一，它在邻居数为26时取值最大为1.02。相同邻居数下，UOS值为0.99，Pearson Entropy和MLCF算法为0.96，差别不大，但MLCF后期得分比Pearson Entropy算法更加平稳。CRQQ算

法在排序能力上比其他算法提高了0.63%—10.21%，这说明共同评分数量和质量在推荐算法排序能力上的优势也比其他

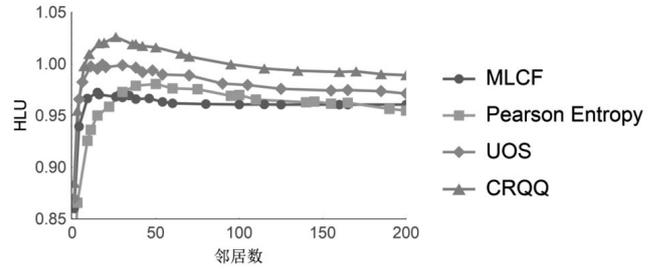


图3 不同算法的HLU结果对比

Fig.3 Comparison of HLU results of different algorithms

图4是算法的Accuracy得分对比。Accuracy代表算法预测用户喜欢与否正确的概率，可以判断系统分类是否有效。CRQQ算法得分稳定在0.704上下，排名第一；Pearson Entropy得分在0.703上下，排名第二；第三、第四分别是MLCF(0.700)和UOS(0.699)。CRQQ比其他算法高了0.14%—0.71%，代表本文算法在对用户喜欢和不喜欢某物品

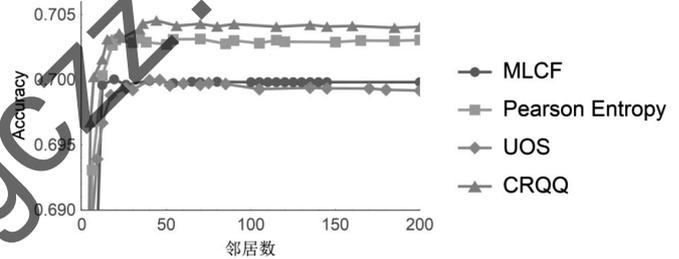


图4 不同算法的Accuracy结果对比

Fig.4 Comparison of Accuracy results of different algorithms

F1代表Precision和Recall的综合得分，得分越高，说明模型或算法的效果越理想。从图5中可以看到，CRQQ算法从邻居数为1时就具有领先的优势，UOS第二，Pearson Entropy第三，MLCF最差。在邻居数为0至50之间，CRQQ算法总是能够取得最高的F1得分。此时，CRQQ算法比其他算法高出了0.24%—1.05%，由此可见，使用共同评分数量和质量

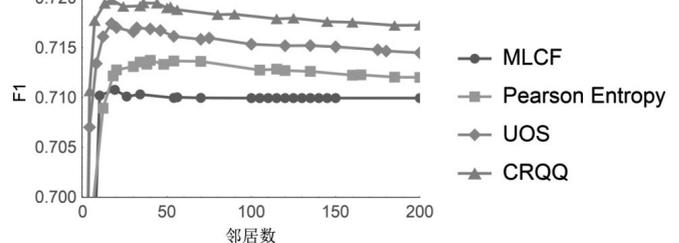


图5 不同算法的F1结果对比

Fig.5 Comparison of F1 results of different algorithms

以上分析了四种算法的整体表现,表1则记录了在相同邻居数下,CRQQ算法与其他算法在各指标的对比中提升的最高百分比。

表1 最高提升百分比  
Tab.1 Best improvement percentage

算法	MAE/%	RMSE/%	HLU/%	Accuracy/%	F1/%
CRQQ	-	-	-	-	-
MLCF	2.48	3.22	5.98	0.63	1.53
UOS	2.84	3.24	2.93	1.37	0.70
Pearson Entropy	6.68	8.41	10.21	2.55	4.44

由表1可知,本文算法在五种指标上分别高出MLCF算法0.63%—5.98%,高出UOS算法0.70%—3.24%,高出Pearson Entropy算法4.44%—10.21%。

总体而言,本文设计的算法CRQQ综合表现领先,Pearson Entropy算法次之,UOS算法第三,多层次协同过滤算法MLCF第四。但当考虑算法在相同的、且较少邻居数下的对比效果时,Pearson Entropy算法表现较差,因其达到最优预测效果,在相同条件下需要更多的邻居。值得注意的是,CRQQ算法不但各个指标领先,且需要的邻居数量也最少,体现出一定的性能优越性。

## 6 结论(Conclusion)

为提高协同过滤个性化推荐算法的预测准确性,提升推荐列表排序质量,增强算法对偏好分类的准确性,本文研究了两个用户间共同评分物品集合的数量和质量对用户相似性度量的影响,设计了融合共同评分数量和质量相似性度量方法和对应的协同过滤个性化推荐算法。并在MovieLens电影评分数据集上与领域内几个典型协同过滤算法对比,研究发现本文设计的融合用户共同评分数量和质量协同过滤个性化推荐算法可以将预测误差降低8.41%,将推荐列表排序性能提高10.21%,将偏好分类预测准确率提高2.55%。实验证明,改良后的算法无论是预测准确性还是分类或排序能力都有较大的提升,充分说明共同评分物品对推荐算法评分预测的影响。在未来的工作中,可以继续挖掘用户公共评分的其他影响,进一步提高算法性能。

## 参考文献(References)

- [1] MATTHES J, KARSAY K, SCHMUCK D, et al. "Too much to handle": Impact of mobile social networking sites on information overload, depressive symptoms, and well-being[J]. Computers in Human Behavior, 2020, 105:106217.
- [2] BOBADILLA J, ORTEGA F, HERNANDO A, et al. Recommender systems survey[J]. Knowledge-Based Systems, 2013, 46:109-132.
- [3] 黄川林,鲁艳霞.基于协同过滤和标签的混合音乐推荐算法研究[J].软件工程,2021,24(4):10-14.
- [4] WANG D, LIANG Y, XU D, et al. A content-based

recommender system for computer science publications[J]. Knowledge-Based Systems, 2018, 157:1-9.

- [5] ADOMAVICIUS G, TUZHILIN A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions[J]. IEEE Transactions on Knowledge and Data Engineering, 2005, 17(6):734-749.
- [6] ZHANG Z, KUDO Y, MURAI T, et al. Improved covering-based collaborative filtering for new users' personalized recommendations[J]. Knowledge and Information Systems, 2020, 62(8):3133-3154.
- [7] XIONG R, WANG J, ZHANG N, et al. Deep hybrid collaborative filtering for web service recommendation[J]. Expert Systems with Applications, 2018, 110:191-205.
- [8] 贾志鹏.基于朴素贝叶斯分类器的校园信息智能推荐算法[J].软件工程,2016,19(12):30-32.
- [9] SU Z, AI J, ZHANG Q, et al. An improved robust finite-time dissipative control for uncertain fuzzy descriptor systems with disturbance[J]. International Journal of Systems Science, 2017, 48(8):1581-1596.
- [10] 蔡崇超,许年虎.矩阵分解技术在电影推荐系统中的应用[J].软件导刊,2021,20(1):174-177.
- [11] 陈金晨.缓解数据稀疏性问题的传统技术研究综述[J].数据通信,2021(2):32-35.
- [12] JONES K S, WILLETT P. Readings in information retrieval[M]. San Francisco, USA: Morgan Kaufmann, 1997:85-92.
- [13] YAN W, LI S, CHENG Y. Multiple similarity collaborative filtering recommendation among users[J]. IOP Conference Series: Materials Science and Engineering, 2020, 768:072010.
- [14] NIKOLAOS P, CHRISTOS K. A multi-level collaborative filtering method that improves recommendations[J]. Expert Systems with Applications, 2016, 48:100-110.
- [15] HE X S, ZHOU M Y, ZHUO Z, et al. Predicting online ratings based on the opinion spreading process[J]. Physica A: Statistical Mechanics and its Applications, 2015, 436:658-664.
- [16] LEE S. Using entropy for similarity measures in collaborative filtering[J]. Journal of Ambient Intelligence and Humanized Computing, 2020, 11(1):363-374.

## 作者简介:

蔡依芳(1996-),女,硕士生.研究领域:推荐系统.

艾均(1980-),男,博士,副教授.研究领域:复杂网络,推荐系统.本文通讯作者.

苏湛(1983-),女,博士,讲师.研究领域:智能控制,推荐系统,复杂网络.