# 一种改进的兴趣相似度个性化推荐算法

李 浩[1],梁京章[1],潘 莹[2*]
(1. 广西大学 电气工程学院,广西 南宁 530004;
2. 广西大学 信息网络中心,广西 南宁 530004)

**摘 要**:传统的协同过滤推荐算法在进行相似度计算时主要考虑用户对物品的评分,通过评分获取用户之间的相似度,缺少对用户兴趣相似度的考虑,同时在进行相似度计算时未考虑用户自身属性的影响,其相似度计算存在一定的失真性。针对这一问题,提出一种改进的兴趣相似度个性化推荐算法,根据不同的用户对物品的兴趣会因用户的自身属性不同而存在差别,设计一种改进的兴趣相似度计算方法,在进行兴趣相似度计算时引入用户的自身属性因素,如年龄、性别等属性因素;根据用户对物品的兴趣会受到物品的热门程度的影响,提出物品热点影响率与物品属性满意度的概念,并根据物品的热点影响率与物品属性满意度在计算相似度时赋予物品不同的权重关系;根据用户的兴趣会随着时间的变化而发生改变,将时间因素加入到推荐过程中,最终通过融合时间因子的影响做出最终的评分预测。在 MovieLens 数据集上的实验结果表明,该算法的平均绝对误差(MAE)和均方根误差(RMSE)值更低,推荐效果更优。

**关键词**:推荐算法;协同过滤;兴趣相似度;物品热点影响率;物品属性满意度;时间因子
**中图分类号**:TP391 **文献标识码**:A **文章编号**:1673-629X(2022)12-0001-06
doi:10.3969/j.issn.1673-629X.2022.12.001

# An Improved Personalized Recommendation Algorithm Based on Interest Similarity

LI Hao[1],LIANG Jing-zhang[1],PAN Ying[2*]
(1. School of Electrical Engineering,Guangxi University,Nanning 530004,China;
2. Information Network Center,Guangxi University,Nanning 530004,China)

**Abstract**:The traditional collaborative filtering recommendation algorithm mainly considers the users´ rating of the item and obtains the similarity between users through the rating. It lacks the consideration of the users´ interest similarity,and does not consider the influence of the users´ own attributes in the similarity calculation,so its similarity calculation has certain inauthenticity. To address this problem,we propose an improved interest similarity personalized recommendation algorithm,and design an improved interest similarity calculation method based on the fact that different users´ interest in items may differ according to their own attributes,and introduce users´ own attributes such as age and gender when calculating interest similarity. Based on the fact that users´ interest in items is affected by the popularity of items,we propose the concepts of item hotspot influence rate and item attribute satisfaction,and assign different weights to items when calculating similarity based on their hotspot influence rate and item attribute satisfaction. Based on the fact that users´ interest changes with time,we add the time factor to the recommendation process,and finally make the final rating prediction by integrating the influence of the time factor. Experimental results on MovieLens dataset show that the average absolute error (MAE) and root mean square error (RMSE) of the algorithm are lower,and the recommendation effect is better.

**Key words**:recommendation algorithm;collaborative filtering;interest similarity;hot spot influence of articles;item attribute satisfaction;time factor

## 1 概 述

在当今数据化时代,海量信息在不断产生和传播的同时,也使人们面临着严重的“信息过载[1]”问题。如在电子商务领域,在面对海量商品的情况下,一方面商家希望能将自己的物品及时推荐给需要的用户,另一方面用户希望能够准确快速地找到自己感兴趣的商

**收稿日期**:2021-11-18 **修回日期**:2022-03-22
**基金项目**:赛尔网络下一代互联网技术创新项目(NGII20190108)
**作者简介**:李 浩(1997-),男,硕士生,研究方向为智能信息处理;通信作者:潘 莹(1980-),女,副研究员,博士,研究方向为模式识别与人工智能。

品。推荐系统的出现为解决这一问题提供了方法。在各种推荐技术中,协同过滤推荐算法是最成功的网络个性化推荐算法[2]。协同过滤算法中应用最广的是基于记忆的协同过滤推荐算法。基于记忆的协同过滤推荐算法的思想是对用户-物品评分矩阵进行相似度计算,通过找到相似的用户或者物品来产生推荐结果[3],其算法中最重要的部分便是相似度的计算。

传统推荐算法中的相似度计算方法主要有余弦相似度、皮尔逊相似度、欧氏距离[4-5]等几种,但由于用户-物品评分矩阵存在数据稀疏等问题[6],导致以上相似度计算方法存在较大的偏差,无法准确找出相似用户或物品,从而导致推荐效果变差,影响用户的体验。

对于相似度计算方法的改进,无数学者做了众多尝试。例如,文献[7-8]对多种传统相似度计算方法使用线性组合的方法来提高推荐精度;文献[9]结合时间衰减效应,着眼于改进少数人的相似度计算方式,提高了推荐精度;文献[10]将物品进行分类后再进行相似度计算,在相似度计算时充分考虑物品的类别信息;文献[11-12]分别对时间因素进行了不同的考虑,采用不同的融合方法将时间因素加入到相似度的计算过程中;文献[13]将用户属性和邻居信任度加入到相似度的计算中,更准确地判断了用户的兴趣偏好;文献[14]设计了一种新的推荐模型,并基于 Kullback-Leibler(KL)散度设计了一个项目相似性度量方法,充分考虑了用户的兴趣偏好,有效提高了推荐精度和推荐效果;文献[15]利用大数据处理框架 MapReduce 对用户-物品矩阵进行细致聚类,在簇类进行多次相似度计算,提高了推荐精度和推荐效率;文献[16]将用户进行物品评分的轨迹信息加入到相似度计算中,并结合多种传统相似度计算方法,使得推荐效果更加精确。

上述改进的相似度计算方法都在一定程度上提高了推荐效果,但是大部分算法在进行相似度的计算时都没有考虑到用户的属性,物品的热门程度都对用户的选择存在着影响,而且大部分算法在相似度计算时以评分的高低判定用户对物品的兴趣程度,这种判定方法是不够准确的,评分高低应代表用户对物品是否满意,而不是用户对物品是否感兴趣。因此,该文提出一种新的融合用户的属性、物品的热门程度以及用户对物品的满意程度的兴趣相似度计算方法,并在评分预测时加入时间因素,以提高推荐效果。

## 2 问题定义

### 2.1 热点影响率

在实际生活中,用户在选择物品时,不仅仅是出自自身的兴趣,还会受到一些社会因素的影响,如商家采取一些营销手段,使得物品获得大量人群的关注,而用户在选择这些物品的时候,往往是出自好奇,而不是用户本身对这些物品存在兴趣。因此,提出热点影响率的概念。

定义1(热点影响率):推荐系统中,物品存在热点的特性,即物品在某个时间段可能引起大量用户的关注,热点的值越大说明物品受用户的关注程度越大。在推荐系统中,物品 $i$ 的热点计算方法如式(1)所示:

$$\mathrm{Hot}(i)=\mathrm{count}(\mathrm{rating}(u,i)>0,u\in U,i\in I) \tag{1}$$

其中,count( )为统计计算方法,rating $(u,i)$ 为用户 $u$ 对物品 $i$ 的评分,$U$ 为用户集合,$I$ 为物品集合。

热点影响率是对热点在推荐系统中兴趣相似度计算的影响的一种度量方式,热点影响率的值在0~1之间,物品 $i$ 的热点影响率计算方法如式(2)所示:

$$\mathrm{n\_Hot}(i)=\frac{\max(\mathrm{Hot})-\mathrm{Hot}(i)}{\max(\mathrm{Hot})-\min(\mathrm{Hot})} \tag{2}$$

其中,max( )和 min( )分别为计算得到的所有物品中热点的最大值和最小值。

物品越受欢迎,说明用户对该物品的评分行为出自自身兴趣的程度越低,因此热点影响率的值越低。

### 2.2 物品属性满意度

在推荐系统中,用户对物品的评分存在高或低两种可能性,传统协同过滤推荐算法中使用用户的评分值来度量用户对物品的兴趣程度,即低评分行为看作用户对该物品的兴趣程度低,高评分行为看作用户对该物品的兴趣程度高。这种度量方法在某些程度上是不合理的,用户对某一个物品存在评分行为,意味着用户对该物品是感兴趣的,用户对物品的评分值低代表该物品在某些属性未能达到用户期望,评分值高代表物品超出用户的期望 。

定义2(物品属性满意度):推荐系统中,物品是存在自身属性的,如在电影推荐系统中,电影存在不同的类型,用户对物品某些属性的期望为用户对与该物品具有相同属性的其他物品的平均评分。用户 $u$ 对物品 $i$ 的属性 $j$ 满意度计算方式如式(3)所示:

$$\mathrm{Sat}_u(ij)=\frac{\mathrm{rating}(u,i)}{\mathrm{average}(u,j)} \tag{3}$$

其中,$\mathrm{average}(u,j)$ 为用户对与物品 $i$ 具有相同属性 $j$ 的物品的平均评分,即用户对物品 $i$ 的属性 $j$ 的期望。

## 3 改进的兴趣相似度计算方法

### 3.1 兴趣相似度

传统兴趣相似度计算方法在进行相似度计算时对影响用户兴趣的因素考虑的不够全面,因此设计一种

新的用户兴趣相似度计算方法。

(1)用户自身属性对用户兴趣的影响。

用户在进行物品选择时,用户自身的属性对用户的兴趣倾向会有一定的影响,如 13 岁的用户在选择时会和 35 岁的用户有差异,男性和女性在选择时也会出现较大的差异。因此,可以看出用户的兴趣受到自身属性的影响。

年龄差距越小对用户兴趣的影响越小,根据文献[17],将用户的年龄差阈值设置为 5,年龄差小于 5 的用户认为其在年龄属性上是相同的。

设用户 $u$ 的年龄为 $A_u$,性别为 $G_u$;用户 $v$ 的年龄为 $A_v$,性别为 $G_v$;则年龄对用户兴趣相似度的计算如式(4)所示:

$$\mathrm{Sim\_age}(u,v)=\begin{cases}1 & |A_u-A_v|\leqslant 5\\ \dfrac{5}{|A_u-A_v|} & |A_u-A_v|>5\end{cases} \tag{4}$$

性别对用户兴趣相似度的计算如式(5)所示:

$$\mathrm{Sim\_g}(u,v)=\begin{cases}\lambda, & G_u=G_v\\ 1-\lambda, & G_u\neq G_v\end{cases} \tag{5}$$

其中,$0<\lambda<1$。

同时,对于冷启动用户,用户自身的属性可以作为最初对用户进行推荐的依据,减少冷启动用户的影响。

(2)物品热点影响率对用户兴趣的影响。

用户在选择热点高的物品时,其出发点可能不是出自本身的兴趣,而是受到商家的宣传营销手段影响,或出于对热点的好奇。因此在对用户兴趣相似度计算时,应该降低热点高的物品对兴趣相似度影响的权重,即热点高的物品热点影响率低。

在计算两个用户的兴趣相似度时,两个用户之间需要存在对某些物品有过共同评分行为,两者之间的共同评分项越多,意味着两个用户的兴趣倾向越相似,用户之间的共同评分项越多且评分方差越小,则用户之间的兴趣相似度越大。

设用户 $u$ 有过评分行为的物品集为 $I_u$,用户 $v$ 有过评分行为的物品集为 $I_v$;则物品热点率对用户兴趣相似度的计算如式(6)所示:

$$\mathrm{Sim\_hot}(u,v)=\frac{|I_u\cap I_v|}{|I_u\cup I_v|}\times\left(1-\sqrt{\frac{\sum\limits_{i\in(I_u\cap I_v)}\left[(\mathrm{rating}(u,i)-\mathrm{rating}(v,i))^2\times n\_\mathrm{Hot}(i)\right]}{|I_u\cap I_v|}}\right) \tag{6}$$

(3)物品属性满意度对用户兴趣的影响。

用户在选择物品时,都带有用户本身的兴趣倾向,用户在对物品打分时,都带有用户本身对该物品的某些属性是否满足自己期望的判断。例如,喜欢喜剧的用户对一部喜剧电影打了低分,这并不意味着用户不喜欢喜剧电影,而是这部喜剧电影可能不够搞笑,未能满足用户对喜剧电影的期望。

设 $S$ 为用户 $u$ 和用户 $v$ 的共同评分项中有着同样属性的物品集,物品集 $S$ 越大,意味着两个用户对于物品属性的选择倾向上越相似,物品集 $S$ 越大,且用户之间的物品属性满意度相差越小,用户之间的兴趣相似度越大。具体计算方式如式(7)所示:

$$\mathrm{Sim\_sat}(u,v)=\frac{|S|}{|I_u\cap I_v|}\times\left(1-\sqrt{\frac{\sum\limits_{i\in S,j\in H_i}(\mathrm{Sat}_u(ij)-\mathrm{Sat}_v(ij))^2}{|S|}}\right) \tag{7}$$

其中,$H_i$ 为物品 $i$ 的属性集。

综上所述,用户的兴趣相似度在计算时应将用户本身的属性、热点影响率、物品属性满意度都考虑在内,因此,用户的兴趣相似度计算公式如式(8)所示:

$$\mathrm{Sim}(u,v)=\mathrm{Sim\_age}(u,v)\times\mathrm{Sim\_g}(u,v)\times\mathrm{Sim\_hot}(u,v)\times\mathrm{Sim\_sat}(u,v) \tag{8}$$

### 3.2　相似用户集

通过兴趣相似度计算公式可以计算出与目标用户与其他用户的兴趣相似度,从而找到目标用户的最相似用户集。但如上文所提到的,在计算两个用户的兴趣相似度时,两个用户之间需要存在对某些物品有过共同评分的行为,但是用户-物品评分矩阵具有稀疏性的问题,某用户可能与其他用户之间存在较少的共同评分项或者没有共同评分项,当用户之间没有共同评分项时,用户之间的兴趣相似度没有意义,当用户之间的共同评分项较少时,用户之间的兴趣相似度存在一些不确定性,使得推荐效果较差。因此需要设置一个阈值 $\beta$,使得用户 $u$ 和用户 $v$ 之间的共同评分项满足式(9):

$$\widetilde{\mathrm{Sim}}(u,v)=|I_u\cap I_v|>\beta \tag{9}$$

当 $\widetilde{\mathrm{Sim}}(u,v)$ 的值即用户之间的共同评分项小于 $\beta$ 时,用户的兴趣相似度为 0。

通过式(8)和式(9)可以找到与目标用户最相似的用户集 $Q$,但由于数据的稀疏性,会出现一部分用户存在较多的相似用户,一部分用户的相似用户很少甚至没有相似用户。因此需要设置相似用户数阈值 $K$,并对于这一现象分类处理:

(1)对于相似用户数> $K$ 的用户,直接选取相似度排名前 $K$ 位用户作为目标用户的相似用户集。

(2)对于相似用户数< $K$,但是>0 的用户,由于相似用户数较少,如果直接使用已有的相似用户集进行推荐,可能会导致“信息茧房”现象的出现,因此需要对该类用户的相似用户进行扩充,由于与目标相似的

用户的相似用户也可能与目标用户相似[18],因此可以使用相似用户集内用户的相似用户对其进行填充至满足第一类情况。填充的用户 $w$ 与目标用户 $u$ 的兴趣相似度计算如式(10)所示:

$$\mathrm{Sim}(u,w)=\mathrm{Sim}(u,v)\times\mathrm{Sim}(v,w) \tag{10}$$

(3)对于相似用户数为0的用户,应在兴趣相似度计算中依次减去物品属性满意度,热点影响率的影响,最终可以同冷启动用户相同处理。

### 3.3 评分预测并推荐

用户的兴趣不是一成不变的,而是随着时间的推移逐渐发生变化,这是人的自然遗忘过程,符合艾宾浩斯遗忘曲线[16]的规律。时间衰减模型的计算方法如公式(11)所示:

$$\mathrm{time}(t_{\mathrm{now}}-t_{ui})=\frac{1}{1+\alpha|t_{\mathrm{now}}-t_{ui}|} \tag{11}$$

其中,$t_{\mathrm{now}}$ 为当前时间,$t_{ui}$ 为用户 $u$ 对物品 $i$ 评分的时间,$\alpha$ 为时间衰减因子。

在对目标用户进行物品推荐时,需要考虑该物品的属性用户现在是否还存在一定的兴趣,用户以前感兴趣的物品现在不一定还存在兴趣,因此时间因子的影响计算公式如公式(12)所示:

$$\mathrm{TIME}(u,x)=\frac{\sum_{y\in T}\mathrm{time}(t_{\mathrm{now}}-t_{uy})}{|T|} \tag{12}$$

其中,$x$ 为待推荐物品,$T$ 为用户 $u$ 所评分过的物品集中与物品 $x$ 具有相同属性的物品合集。

最终评分预测公式为:

$$\mathrm{pre}(u,x)=\mathrm{rating}_u'+\frac{\sum_{q\in Q}[\mathrm{sim}(u,q)\times\mathrm{TIME}(u,x)\times(\mathrm{rating}(q,x)-\mathrm{rating}_u')]}{|Q|} \tag{13}$$

其中,$\mathrm{rating}_u'$ 为用户 $u$ 对所有物品的平均评分。选择最相似用户集 $Q$ 中用户有过评分行为但目标用户未评分过的物品,通过式(13)计算目标用户对其的预测评分,并按预测评分高低排序,选择前 $N$ 个物品进行推荐。

推荐算法流程如图1所示。

推荐算法伪代码如下:

```
算法1
输入:用户集合 U,物品集合 I,共同评分项阈值 β,相似用户数 K。
1. begin
2. uSimv(u,v)={}
3. Simlist(u)={} //用户 u 的相似用户集
4. for each u in U:
5.    for each v in U:
6.       if  u ≠ v:
7.          for each i in I::
8.             if  rating(u,i)>0 & rating(v,i)>0
9.                uSimv(u,v).add(i)
10.       if  Len(Sim~(u,v))>β
11.          Simlist(u).add([v,Sim(u,v)])
12.       if  Len(Simlist(u)) ≥ K:
13.          Simlist(u)=Sort(Simlist(u))  // 按相似度大小排序
14.          elif  Len(Simlist(u))<K & len(Simlist(u))>0:
15.          Simlist(u).add([x,Sim(v,x)])
16. recommendList={}
17. for each w in Simlist(u):
18.    if rating(u,i)=0 & rating(u,i)>0:
19.       recommendList.add(Sort(pre(u,i)))
20. end
```

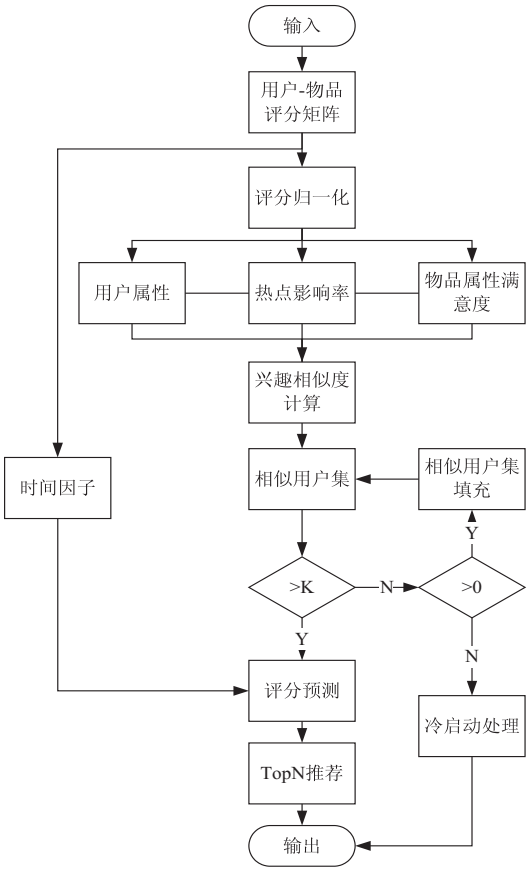


图1 推荐算法流程

## 4 实验与结果分析

### 4.1 实验数据集

实验数据使用的是美国 Minnesota 大学创办的 MovieLens 数据集,数据集的具体描述如表1所示。

表1 数据集描述

| 数据集 | 用户数 | 物品数 | 评分数 | 稀疏度/% |
|---|---|---|---|---|
| M1-100k | 943 | 1 682 | 10 000 | 99.3 |

### 4.2　实验配置

硬件环境:

(1)Windows10;

(2)CPU:Intel7 代 6 700k;

(3)内存 16 G。

软件环境:Pycharm。

### 4.3　评价标准

采用平均绝对误差(MAE)和均方根误差(RMSE[19])对实验结果进行评测:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_{ij} - y_{ij}| \tag{14}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_{ij} - y_{ij})^2} \tag{15}$$

### 4.4　实验安排

由于算法设计时间的影响,将数据集按时间排序,并将排序后的数据集以 8∶2 的比例分为训练集和测试集。为验证提出的相似度计算方式在提升推荐效果方面的有效性,将该算法与使用传统相似度计算的协同过滤算法和几种改进相似度计算方式的协同过滤算法在 MovieLens 数据集进行对比实验,具体涉及算法如下:

(1)文中算法 new_CF;

(2)使用余弦相似度的协同过滤算法 cor_CF;

(3)使用 Person 相似度的协同过滤算法 P_CF;

(4)使用 Jaccard 相似度的协同过滤算法 J_CF;

(5)文献[9]中提出的 min-CF 算法;

(6)文献[11]融入时间因子的协同过滤算法 T_CF;

(7)文献[19]提出的融合兴趣偏好和加权的 Jaccard 相似度的 ED-JM 算法。

### 4.5　结果分析

实验一:$\lambda$ 和 $\alpha$ 对推荐结果的影响。

选择 $K=30$;$\beta=5$, $N=20$ 进行实验,求得 $\lambda$ 和 $\alpha$ 的最佳取值;

逐步递增 $\lambda$ 的值,在不同的 $\alpha$ 取值下,所获得最佳实验结果如图 2 所示。

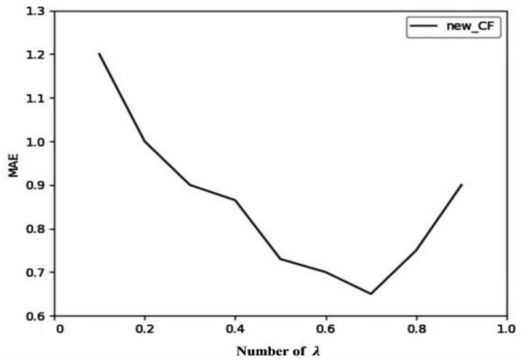


图 2　不同 $\lambda$ 对应的最佳 MAE

逐步递增 $\alpha$ 的值,在不同的 $\lambda$ 取值下,所获得的最佳实验结果如图 3 所示。

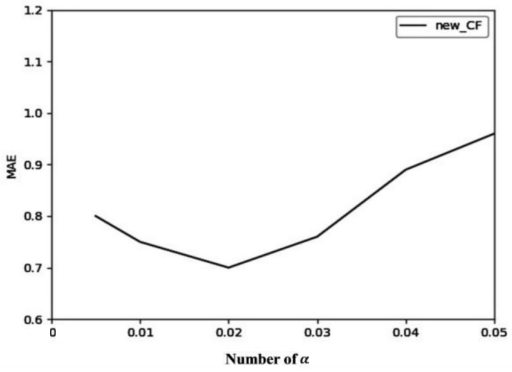


图 3　不同 $\alpha$ 对应的最佳 MAE

从如图 2 和图 3 可以看出,$\lambda$ 的最佳取值约为 0.7,$\alpha$ 的最佳取值约为 0.02,当 $\lambda$ 和 $\alpha$ 的取值从最佳取值处增大或减小时,都会引起 MAE 的剧烈变化。

实验二:最相似用户集数 $K$ 对推荐结果的影响。

从 $K=10$ 开始取值,每次增加 10,直至 $K=70$,实验结果如图 4 和图 5 所示。

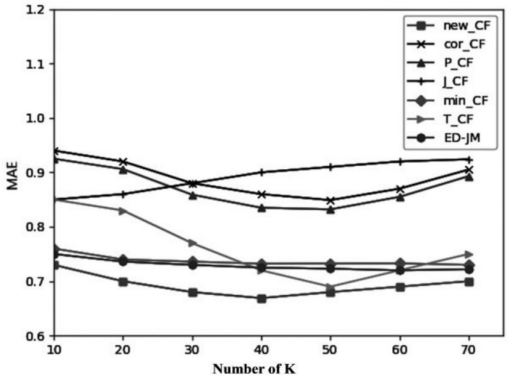


图 4　$K$ 值对各算法的 MAE 的影响

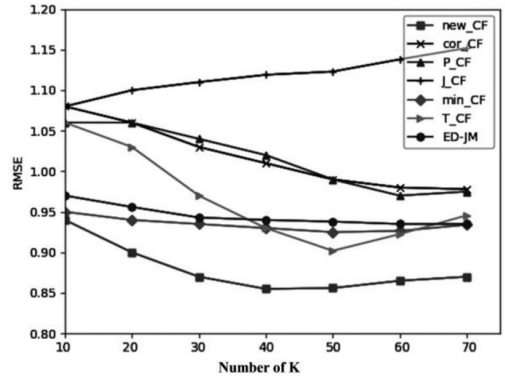


图 5　$K$ 值对各算法的 RMSE 值的影响

从图 4 和图 5 可以看出,传统的推荐算法的误差最大,这表明只使用传统相似度计算方法已经无法满足推荐精度的需要,改进相似度计算方式是有必要的。随着 $K$ 值的增加,除 J_CF 算法外,其他算法模型的 RMSE、MAE 的值都在降低,推荐效果逐渐提升,这可能是因为初始随着 $K$ 值的增加,相似用户数增多,那些

单独近邻所造成的影响被降低,推荐效果得到提升;在这个过程中,相较于传统相似度计算方式,在相似度计算方式上引入其他因素的 J_CF、min_CF、T_CF 算法均提高了推荐效果;该文提出的相似度计算方式在对相似用户的选择上,充分考虑用户自身的属性因素,减少了选取错误最近邻情况的发生,同时考虑到物品流行度的影响,减少了因热点问题导致选取错误近邻的情况;因此各项误差均处于最低,推荐效果最好,当 $K=40$ 时,该算法的 MAE 值取得最小,达到最佳效果。

随着 $K$ 值的继续增大,所有算法均出现推荐效果减弱,误差增大的情况,这是因为一些不重要不相似的用户被加入最近邻集合,导致偏差增大,推荐效果减弱,其中文中算法、J_CF、min_CF 受 $K$ 值增大的影响最小。因为这三类算法在相似度的计算方式上考虑的因素较多,不单纯依赖用户物品评分矩阵,因此在相似用户的选择上更加有效可靠,减少了不相似用户加入到最近邻集合中的情况出现;文中算法充分考虑了用户属性、物品热门程度、时间因素的影响,对相似用户的选择上更准确,推荐效果更好。

从以上可以看出,new_CF 相对于使用传统相似度计算方法的协同过滤算法在推荐效果上有显著的提升,且对比其他几种改进相似度计算方式的协同过滤算法也存在一定的推荐优势,这充分证明提出的改进的兴趣相似度个性化推荐算法能有效提高系统的推荐性能。

## 5 结束语

针对传统相似度计算方法在计算时存在一定的失真性,提出融合用户的自身属性、物品的热点影响率与物品属性满意度的新型兴趣相似度计算方法,并融合时间因子求取最终评分预测。通过实验证明,提出的改进的兴趣相似度个性化推荐算法在 MAE 和 RMSE 上均表现良好,有效提高了推荐效果。下一步将针对推荐系统中的冷启动和数据稀疏性问题,挖掘更多的兴趣相似度影响因素来提高推荐效果。

**参考文献:**


[1] ZHANG J, WANG Y, YUAN Z, et al. Personalized real-time movie recommendation system: practical prototype and evaluation[J]. Tsinghua Science and Technology, 2019, 25(2): 180-191.

[2] DENG J, GUO J, WANG Y. A novel K-medoids clustering recommendation algorithm based on probability distribution for collaborative filtering[J]. Knowledge-Based Systems, 2019, 175: 96-106.

[3] 张 宏,王 慧. 基于用户评分和共同评分项的协同过滤算法研究[J]. 计算机应用研究,2019,36(1):77-80.

[4] JAIN G, MAHARA T, TRIPATHI K N. A survey of similarity measures for collaborative filtering-based recommender system[M]//Soft computing: theories and applications. Singapore: Springer, 2020: 343-352.

[5] POLATO M, AIOLLI F. Exploiting sparsity to build efficient kernel based collaborative filtering for top-N item recommendation[J]. Neurocomputing, 2016, 268: 17-26.

[6] HE Y, YANG S, JIAO C. A hybrid collaborative filtering recommendation algorithm for solving the data sparsity [C]//2011 international symposium on computer science and society. Malaysia: IEEE, 2011: 118-121.

[7] SURYAKANT, TRIPTI M. A new similarity measure based on mean measure of divergence for collaborative filtering in sparse environment[J]. Procedia Computer Science, 2016, 89: 450-456.

[8] 王建芳,韩鹏飞,苗艳玲,等. 一种基于用户兴趣联合相似度的协同过滤算法[J]. 河南理工大学学报:自然科学版,2019,38(5):118-123.

[9] 张 维,黄 勃,张 娟,等. 面向少数类用户兴趣演化的推荐算法[J]. 南京理工大学学报:自然科学版,2021,45(2):214-222.

[10] 何 明,肖 润,刘伟世,等. 融合类别信息和用户兴趣度的协同过滤推荐算法[J]. 计算机科学,2017,44(8):230-235.

[11] 胡安明,陈惠娥. 基于时间因子改进个性化推荐模型[J]. 软件工程,2021,24(7):60-62.

[12] 高茂庭,王 吉. 融合社交关系与时间因素的主题模型推荐算法[J]. 计算机工程,2020,46(3):66-72.

[13] 金志刚,朱 琦,刘晓辉. 基于属性偏好和邻居信任度的协同过滤推荐算法[J]. 南开大学学报:自然科学版,2021,54(3):25-30.

[14] WANG Y, DENG J, GAO J, et al. A hybrid user similarity model for collaborative filtering[J]. Information Sciences, 2017, 418-419: 102-118.

[15] KIM S, KIM H, MIN J K. An efficient parallel similarity matrix construction on MapReduce for collaborative filtering [J]. Journal of Supercomputing, 2019, 75(1): 123-141.

[16] CAO L, DENG M, ZHU J, et al. An improved slope one recommendation algorithm based on user preferences[C]// 2018 IEEE international conference on information and automation (ICIA). Wuyishan: IEEE, 2018: 1323-1328.

[17] 柯翔敏,陈 江,罗光华. 一种改进的基于兴趣相似度推荐算法[J]. 计算机工程,2020,46(8):78-84.

[18] 印桂生,崔晓晖,马志强. 遗忘曲线的协同过滤推荐模型[J]. 哈尔滨工程大学学报,2012,33(1):85-90.

[19] HYNDMAN R J, KOEHLER A B. Another look at measures of forecast accuracy[J]. International Journal of Forecasting, 2006, 22(4): 679-688.