

# 基于柯西分布量子粒子群的混合推荐算法

王桐, 曲桂雪

(哈尔滨工程大学 信息与通信工程学院, 黑龙江 哈尔滨, 150001)

**摘要:** 协同过滤推荐算法是最经典、应用最成功的推荐算法之一, 但该算法在数据稀疏性、冷启动和时间因素等方面还存在一定问题, 于是, 提出一种基于柯西分布量子粒子群的混合推荐算法。该算法首先构建基于时间因子的混合推荐模型, 再利用柯西分布量子粒子群算法搜索模型中的最优参数组合, 其中, 混合推荐模型通过把用户和项目的属性信息添加到协同过滤推荐算法中, 并引入能够代表用户兴趣迁移特性的时间因子构建而成。最后, 与人工蜂群算法(ABC)以及基本粒子群算法(PSO)进行比较。研究结果表明: 在提高推荐准确度、缓解数据稀疏性以及冷启动等方面, 本文提出的算法优于其他算法。

**关键词:** 推荐算法; 柯西分布量子粒子群; 数据稀疏; 冷启动; 时间因子

中图分类号: TP301.6

文献标志码: A

文章编号: 1672-7207(2015)08-2898-08

## Hybrid recommendation algorithm based on Cauchy quantum-behaved particle swarm optimization

WANG Tong, QU Guixue

(College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China)

**Abstract:** Collaborative filtering recommendation algorithm is one of the most typical and successful technologies, about exist problems such as data sparsity, cold start and time factor. Therefore, a hybrid recommendation algorithm based on Cauchy quantum-behaved particle swarm optimization was proposed. According to the algorithm, the hybrid recommendation model was constructed based on time factor, and then, the Cauchy quantum-behaved particle swarm optimization algorithm was applied for searching the optimal parameters of the model. The hybrid recommender model was built by adding the features of the users and items to the traditional collaborative filtering algorithm and introducing a time factor represented the change of users' interests. The algorithm proposed in this paper was compared with the artificial bee colony (ABC) and particle swarm optimization (PSO). The results show that in increasing recommendation accuracy and alleviating the data sparsity and cold start, the proposed algorithm is better than other algorithms.

**Key words:** recommendation algorithm; Cauchy quantum-behaved particle swarm optimization; data sparseness; cold start; time factor

互联网的发展将我们带入了一个大数据信息化的时代, 给人们带来方便的同时, 也使人们在选择所需

信息时显得力不从心, 于是, 推荐系统应运而生, 并得到广泛关注和应用<sup>[1]</sup>。目前, 主流的推荐算法有基

收稿日期: 2014-08-25; 修回日期: 2014-11-17

**基金项目(Foundation item):** 国家自然科学基金资助项目(61102105); 黑龙江省博士后启动基金资助项目(LBH-Q12117); 教育部博士点基金资助项目(20102304120014); 黑龙江省自然科学基金资助项目(F201029)(Project (61102105) supported by the National Natural Science Foundation of China; Project (LBH-Q12117) supported by the Postdoctoral Start Foundation of Heilongjiang Province; Project (20102304120014) supported by the Doctoral Fund of Ministry of Education; Project (F201029) supported by the Natural Science Foundation of Heilongjiang Province)

**通信作者:** 王桐, 博士, 副教授, 从事物联网、WSN、智能计算等研究; E-mail: wangtong@hrbeu.edu.cn

于规则、内容、协同过滤以及混合推荐等,这些推荐算法通常存在如忽略用户兴趣的时效性、数据稀疏性以及冷启动等缺陷<sup>[2]</sup>。专家学者不断探索,提出不少改进算法,如 Jeong 等<sup>[3]</sup>提出依据用户可信度进行推荐的方法,并证明该方法在解决冷启动问题上具有良好的效果;邓爱林等<sup>[4]</sup>将聚类算法用于推荐系统中,可缓解数据稀疏性问题,但聚类算法存在没有成熟的聚类策略及更新模型训练结果成本较高的问题;孟宪福等<sup>[5]</sup>采用贝叶斯网络算法,可解决推荐的时效性问题,但模型建立的时间复杂度很高;黄正<sup>[6]</sup>提出时间函数的概念,考虑到用户兴趣随时间变化,有效提高推荐的准确度,但对于时间函数的设计仍不够理想;赵丽嫚<sup>[7]</sup>采用智能算法中的遗传算法搜索推荐过程中的最优参数组合,取得了良好的效果,但遗传算法存在局部搜索能力差、易陷入局部极值及迭代速度慢等问题,对于搜索最优参数组合,遗传算法不是最优的选择。针对上述问题,本文提出一种基于柯西分布量子粒子群混合推荐算法(hybrid recommendation based on Cauchy quantum-behaved particle swarm optimization, CQPSO-HR),主要包括:1) 建立基于时间因子的混合推荐模型,可提高用户相似度计算的准确性,从而提高推荐的准确度;2) 提出 CQPSO-HR 算法,该算法基于混合推荐模型,采用柯西分布量子粒子群算法在问题空间搜索最优参数组合。与传统进化算法如粒子群算法(particle swarm optimization, PSO)和人工蜂群算法(artificial bee colony, ABC)<sup>[8-9]</sup>相比可知,CQPSO-HR 比 PSO 和 ABC 的搜索能力更强,推荐准确度更高。

## 1 基于时间因子的混合推荐模型

传统的协同过滤推荐算法虽然应用广泛,但在以下几方面还不够完善。

1) 只根据用户历史评分数据来推荐。这不但全面,而且会导致冷启动问题,造成新项目或新用户无法推荐,同时还会使数据稀疏性问题更加严重,影响推荐效果。因此,要综合考虑用户和项目的评分及内容信息,这不仅可以在新用户或新项目加入时,利用内容信息计算相似度,解决冷启动的问题,而且能在一定程度上缓解数据稀疏性问题<sup>[10]</sup>。

2) 忽略用户兴趣随时间变化的特性。例如,某用户曾经很喜欢看动作片,一段时间后,他的兴趣发生了改变,开始喜欢剧情片。与此同时,另一用户开始喜欢动作片。这时,他们会被认为是相似邻居用户,

可这不合理,因为在当前时间内,这2个用户并没有共同的喜好。实际上,只有在同一时间内爱好相似的用户才应该被称为相似用户<sup>[11]</sup>。举例说明,如表1所示。

表1 用户随时间的兴趣程度

Table 1 Users' interests with time

用户	时间	爱好1	爱好2	爱好3
$U_1$	$T_A$		4	2
$U_2$	$T_B$	5		
$U_3$	$T_A$		4	4
	$T_B$	5		

表1中有3个用户  $U_1$ ,  $U_2$  和  $U_3$ , 2个不同的时间  $T_A$  和  $T_B$ , 表中数字1~5代表用户兴趣程度,数字越大表示兴趣程度越高。根据传统的相似度计算方法,  $U_1$  和  $U_2$  都是  $U_3$  的最近邻居。但实际上,在  $T_A$  时,用户  $U_3$  的最近邻居只有  $U_1$ , 到  $T_B$  时,最近邻居变为只有  $U_2$ 。可见,时间因子是决定推荐质量的重要因素,引入时间因子能够提高推荐系统的推荐准确度。

因此,本文提出基于时间因子的混合推荐模型。该模型主要分以下3步实现:构建用户-项目评分矩阵、寻找用户和项目近邻和产生推荐<sup>[12]</sup>。

1) 构建用户-项目评分矩阵。评价信息可以有多种表示方式,可用二进制0和1表示用户喜欢或不喜欢的项,也可用整数1~5表示用户对项目的评分级别。

2) 构建用户、项目的相似度矩阵,寻找用户和项目的最近邻居集合。首先,分别计算用户评分相似度矩阵  $U_{Rsim}$ 、用户内容相似度矩阵  $U_{Csim}$ 、项目评分相似度矩阵  $I_{Rsim}$  以及项目内容相似度矩阵  $I_{Csim}$ ,从而获得用户综合相似度矩阵和项目综合相似度矩阵;其次,依据具体的选择策略选择用户和项目的最近邻居。

当前,研究人员针对用户兴趣的变化问题进行了大量的研究。邢春晓等<sup>[13]</sup>提出一种非递减的线性时间函数<sup>[13]</sup>。该函数依照用户评价项目时间的不同,为每个用户分配不同的权重,事实上,用户的兴趣变化是非线性的。因此,引入时间因子  $T(u,i)$ :

$$T(u,i) = 1 - \frac{|t_{ui} - t_{vi}|}{\max(|t_{ki} - t_{vi}|)}; \quad u, v, k \in U_i \quad (1)$$

其中:  $t_{ui}$  和  $t_{vi}$  分别为用户  $u$  和目标用户  $v$  对项目  $i$  的评分时间;集合  $U_i$  由对项目  $i$  评过分的的所有用户共同组成;  $T(u,i)$  为用户  $u$  对项目  $i$  评分的时间因子,其取

值范围为(0,1), 并且当用户  $u$  的评分时间与目标用户  $v$  的评分时间越接近时获得的权重越大。

在基于时间因子的混合推荐模型中, 用户评分相似度的计算式为

$$S(u, v) = \frac{\sum_{c \in C_{uv}} (R_{ui} - \bar{R}_u)(R_{vi} - \bar{R}_v) \times T(u, i)}{\sqrt{\sum_{c \in C_u} (R_{ui} - \bar{R}_u)^2} \sqrt{\sum_{c \in C_v} (R_{vi} - \bar{R}_v)^2}} \quad (2)$$

其中:  $C_{uv}$  表示用户  $u$  和  $v$  各自评分项目集合的并集;  $C_u$  和  $C_v$  代表用户  $u$  和  $v$  各自评分过的项目集合;  $\bar{R}_u$  和  $\bar{R}_v$  分别代表用户  $u$  和  $v$  对所有评价过的项目评分的均值;  $R_{ui}$  和  $R_{vi}$  分别为用户  $u$  和  $v$  对项目  $i$  的评分。

用户和项目综合相似度矩阵  $U_{sim}$  和  $I_{sim}$  计算方法分别为

$$U_{sim} = w_1 \times U_{Rsim} + (1 - w_1) \times U_{Csim} \quad (3)$$

$$I_{sim} = w_2 \times I_{Rsim} + (1 - w_2) \times I_{Csim} \quad (4)$$

其中:  $w_1$  为用户评分与内容的权重;  $w_2$  为项目评分与内容的权重。

本文设定用户最近邻居阈值为  $w_3$ , 项目最近邻居阈值为  $w_4$ 。

3) 产生推荐。得到用户和项目最近邻居矩阵后, 利用推荐式(5)和(6)分别计算基于用户推荐的项目预测评分  $U_{Pui}$  和基于项目推荐的项目预测评分  $I_{Pui}$ <sup>[14]</sup>, 最后, 利用式(7)得最终的混合推荐结果。

$$U_{Pui} = \bar{R}_u + \frac{\sum_{c \in C_{uv}} U_{S(u, v)} \times (R_{vi} - \bar{R}_v)}{\sum_{c \in C_{uv}} |U_{S(u, v)}|} \quad (5)$$

$$I_{Pui} = \bar{R}_i + \frac{\sum_{c \in I_{uv}} I_{S(u, v)} \times (R_{vi} - \bar{R}_i)}{\sum_{c \in I_{uv}} |I_{S(u, v)}|} \quad (6)$$

$$P_{ui} = w_5 \times U_{Pui} + (1 - w_5) \times I_{Pui} \quad (7)$$

其中:  $U_{S(u, v)} \geq w_3$ ,  $I_{S(u, v)} \geq w_4$ , 表示只有相似度大于  $w_3$  的用户才能成为目标用户的最近邻居, 只有相似度大于  $w_4$  的项目才能成为目标项目的最近邻居。  $w_5$  为混合推荐权重。  $P_{ui}$  为用户  $u$  对项目  $i$  的预测评分值。

推荐系统的评价标准较多, 本文采用预测准确度来衡量推荐算法的性能<sup>[15]</sup>。评分预测的准确度一般通过平均绝对误差  $E_{MA}$  计算。对于测试集  $T$  中的用户  $u$  和项目  $i$ , 令  $R_{ui}$  为用户对项目  $i$  的实际评分, 而  $\hat{R}_{ui}$  为推荐算法给出的预测评分, 则平均绝对误差  $E_{MA}$  的定义为

$$E_{MA} = \frac{\sum_{u, i \in T} |R_{ui} - \hat{R}_{ui}|}{|T|} \quad (8)$$

基于时间因子的混合推荐模型如图 1 所示。

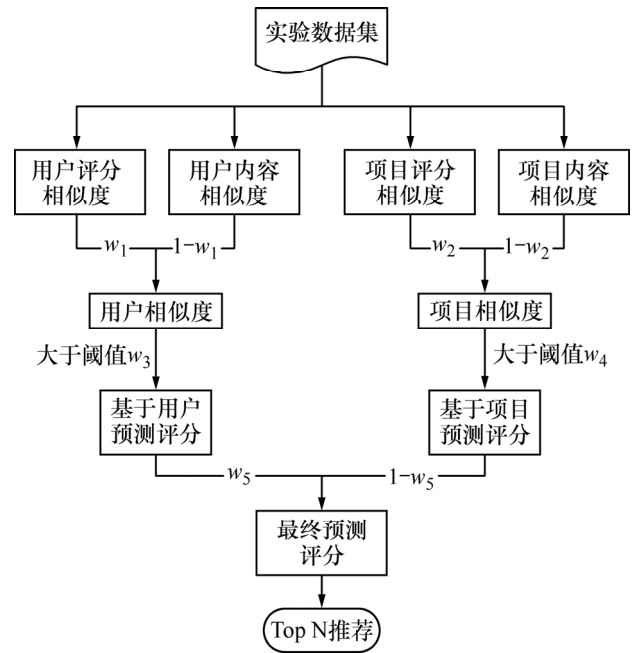


图 1 基于时间因子的混合推荐模型流程图

Fig. 1 Flow chart of hybrid recommendation model based on time factor

## 2 基于柯西分布量子粒子群的混合推荐算法(CQPSO-HR)

### 2.1 基于柯西分布量子粒子群算法

在标准粒子群优化算法中, 粒子的位置和速度共同决定了其运行轨迹, 另外, 由于速度有限, 轨迹固定, 所以搜索空间有限, 导致全局搜索能力差<sup>[16]</sup>。为克服由确定性所导致的问题, 考虑在 PSO 中引入不确定性(即概率), 将粒子的搜索空间从经典空间向量子空间转化。2004 年 Sun 等<sup>[17]</sup>提出了基于 Delta 势阱的量子粒子群算法, 该算法描述如下: 在  $D$  维搜索空间中有  $n$  个粒子, 其中, 第  $i$  个粒子的位置  $X_i = (X_{i1}, X_{i2}, \dots, X_{iD})$  ( $i=1, 2, \dots, n$ ), 将代入目标模型即可得到其适应值; 记第  $i$  个粒子搜索到的最优位置即个体最优值为  $p_i = (p_{i1}, p_{i2}, \dots, p_{iD})$  ( $i=1, 2, \dots, n$ ); 整个粒子群搜索到的最优位置即全局最优值为  $p_g = (p_{g1}, p_{g2}, \dots, p_{gD})$ ,  $g \in \{1, 2, \dots, n\}$ 。粒子位置更新方程为

$$X_{id}(t+1) = p_{id}(t) \pm \frac{L}{2} \ln\left(\frac{1}{\mu}\right) \quad (9)$$

式中:  $\mu$  服从  $[0, 1]$  上的均匀分布;  $L$  由

$$L(t+1) = 2\beta |m_{\text{best}}(t) - X_{id}(t)| \quad (10)$$

确定;  $m_{\text{best}}(t)$  为第  $t$  次迭代时的粒子平均值, 表示群体中平均最佳位置, 定义为

$$m_{\text{best}}(t) = \frac{1}{n} \sum_{i=1}^n p_i(t) = \left[ \frac{1}{n} \sum_{i=1}^n p_{i1}(t), \frac{1}{n} \sum_{i=1}^n p_{i2}(t), \dots, \frac{1}{n} \sum_{i=1}^n p_{iD}(t) \right] \quad (11)$$

基于 Delta 势阱的量子粒子群更新公式为

$$X_{id}(t+1) = p_{id}(t) \pm \beta |m_{\text{best}}(t) - X_{id}(t)| \cdot \ln\left(\frac{1}{\mu}\right) \quad (12)$$

其中:  $i=1, 2, \dots, n$ ;  $d=1, 2, \dots, D$ ;  $t$  为当前迭代次数;  $\beta$  为收缩扩张系数, 是量子粒子群优化算法重要参数, 它的取值可固定, 也可按递增或递减的方式变化。在多数情况下通过

$$\beta = a - (a-b) \times \frac{T_{\max} - t}{T_{\max} + 0.5} \quad (13)$$

确定;  $T_{\max}$  为迭代的最大次数; 常量  $a=1$ ; 常量  $b=0.5$ 。

虽然基于 Delta 势阱的量子粒子群算法全局搜索能力远远优于一般的粒子群优化算法, 但是由于粒子在搜索过程中根据学习自身的历史最优位置来实现下一步搜索, 没有考虑是否有陷入局部最优的危险, 在这一过程中, 种群多样性不可避免地会减少, 从而出现早熟现象。鉴于以上问题, 本文在更新粒子位置时同时考虑个体最优位置和全局最优位置, 采用基于柯西分布量子粒子群算法(Cauchy quantum-behaved particle swarm optimization)。柯西分布最大的特点是它的两翼概率特性, 可有效扩大粒子的全局搜索范围, 使它比高斯分布产生的随机数分布区域更大, 因而, 柯西分布有更大的概率快速跳出局部最优点<sup>[18]</sup>。采用标准的柯西分布  $C(0,1)$  分布, 将随机数  $\mu$  由原来的服从  $[0,1]$  上的均匀分布改为服从  $C(0,1)$ , 则粒子位置更新方程为

$$X_{id}(t+1) = q_{id}(t) \pm \beta |m_{\text{best}}(t) - X_{id}(t)| \ln\left(\frac{1}{\mu}\right) \quad (14)$$

且

$$q_{id}(t) = (c_1 p_{id}(t) + c_2 p_{gd}(t)) / (c_1 + c_2), \quad c_1 \sim C(0,1), \quad c_2 \sim C(0,1) \quad (15)$$

## 2.2 CQPSO-HR 算法流程

所描述的推荐模型中涉及 3 个权值参数和 2 个阈值参数, 在 CQPSO-HR 算法中, 种群中粒子搜索空间为 5 维, 分别是推荐模型中项目评分和内容权值、用户和项目推荐权值、项目近邻阈值、用户评分和内容权值以及用户近邻阈值。CQPSO-HR 算法流程如图 2

所示。

## 3 实验测试

本实验运行环境为 CPU(Intel(R)Pentium (R2.13 GHz, 内存 2 GB)及 Windows7 操作系统, 算法程序采用 Matlab 语言实现。

### 3.1 实验数据集

选取美国 Minnesota 大学 Group Lens 研究小组提供的 Movie Lens 数据集<sup>[19]</sup>评估 CQPSO-HR 算法的性能。Movie Lens 是 1 个著名的电影评分数据集, 包含 943 个用户对 1 682 部电影共 10 万条的电影评分记录。除此之外, Movie Lens 数据集还提供用户和电影的内容信息数据以及用户对电影的评分时间。

数据集有 1 个特别重要的特征指标, 即该数据集中数据的稀疏程度, 简称为稀疏度( $S_p$ ), 它定义为用户-项目评分矩阵中未评分项目所占的百分比, 如下所示:

$$S_p = 1 - \frac{D}{N_u \times N_i} \quad (16)$$

其中:  $D$  为用户已评分数据;  $N_u$  为用户数;  $N_i$  为项目数。本文所采用的数据集的稀疏度为  $S_p=93.7\%$ 。

MovieLens 数据集中包含用户评分数据、用户特征数据以及项目特征数据 3 部分信息。

用户评分数据包括用户编号、项目编号及用户评分。其中最大用户编号为 943, 最大项目编号为 1 682, 用户评分的范围为 1~5(其中, 1 表示最低评分, 5 表示最高评分)。

用户特征数据中包括用户编号、年龄、性别、职业、地址、邮编以及用户对电影的评分时间戳。其中用户年龄分成如下 5 段: 0~20 岁, 21~30 岁, 31~40 岁, 41~50 岁, 51 岁及以上, 分别用 1, 2, 3, 4 和 5 表示。用户性别表示为: 男(1), 女(0)。用户职业分为教师、学生、医生、律师、工程师等 21 类, 分别用整数 1, 2, 3, ..., 21 表示, 其中不属于任何类型的职业定义为其他, 令其他归为一类职业, 否则推荐算法难以实施。所在地按用户邮编首字母编码, 首字母相同表示处于同一地区。用户时间戳记录了从系统启用对用户电影进行评价的时间间隔, 本文中该信息用于计算用户的时间因子, 以反映用户兴趣随时间变化的特性。

项目特征数据中包括电影的编号和种类。电影种类包含动作片、剧情片、恐怖片、喜剧片等 19 类, 有些影片同时属于多个类, 所以, 实验中某电影属于当

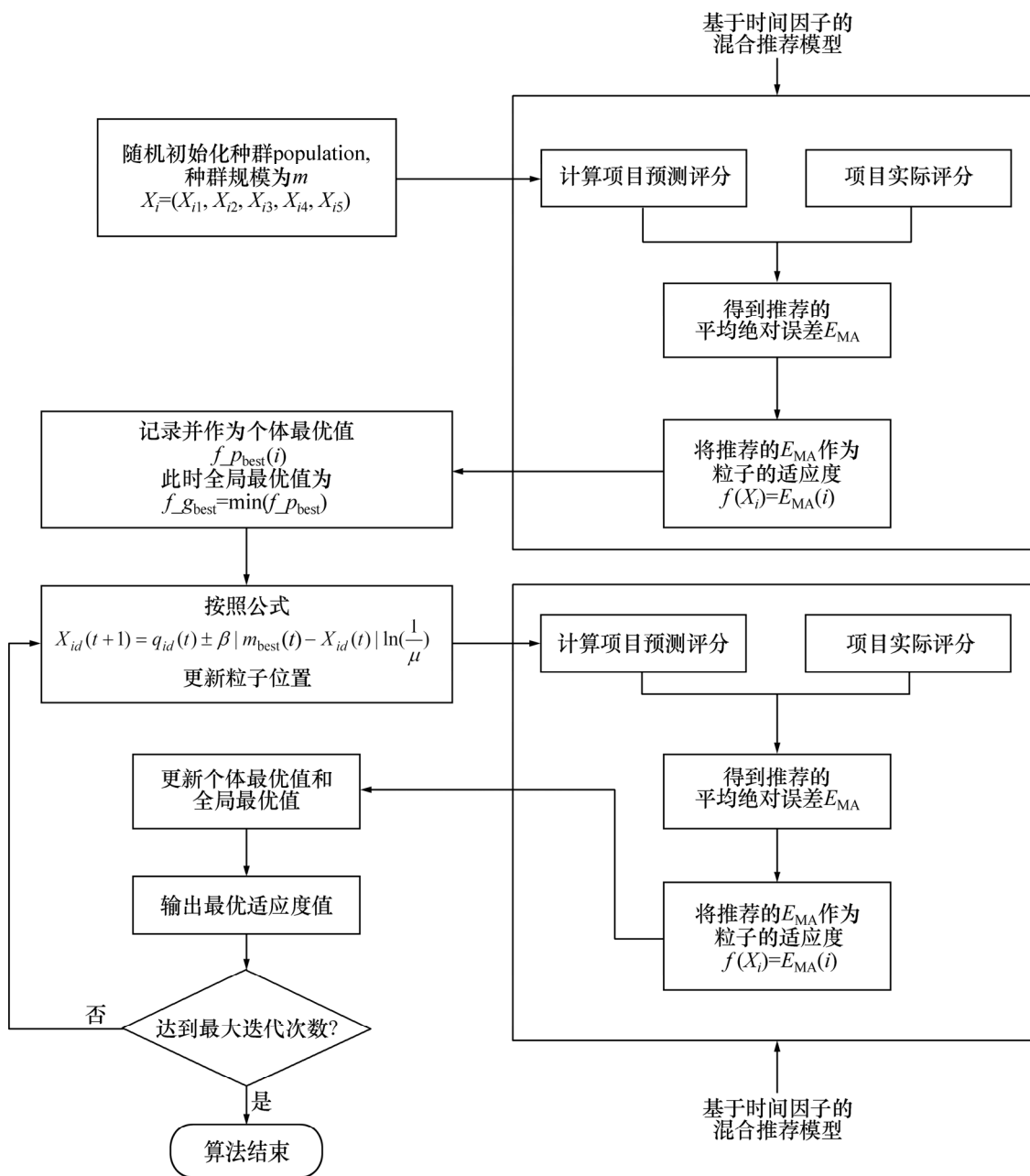


图 2 CQPSO-HR 算法流程图

Fig. 2 Flow chart of CQPSO-HR algorithm

前类别时对应单元格置为 1, 否则置为 0。

### 3.2 实验结果分析

#### 3.2.1 测试时间因子的性能

由于混合推荐算法中涉及很多参数的设定, 不利于进行实验, 因此, 这里采用传统的 CF 算法与加入时间因子后的 CF 算法(T-CF)进行对比来测试本文改进的时间因子的性能, 对比结果如图 3 所示。

图 3 中横坐标代表用户近邻阈值, 纵坐标为平均绝对误差  $E_{MA}$ , 从图 3 可以看出: 随着用户近邻阈值的提高, 推荐的平均绝对误差  $E_{MA}$  在减小, 其中, T-CF

算法的最小  $E_{MA}$  为 0.588 6, 明显低于 CF 算法的最小  $E_{MA}$  0.629 2, 说明所引入的时间因子可有效提高推荐的准确度。

#### 3.2.2 测试 CQPSO-HR 的性能

为了更直观地看到 CQPSO-HR 的优化效果, 将人工蜂群算法(ABC)和标准粒子群算法(PSO)也引入混合推荐算法中与 CQPSO-HR 进行对比。由于这几种智能优化算法的初始种群都是随机的, 所以, 为了保证算法结果的可靠性, 每种算法都进行 10 次实验, 取其平均值作为最后的结果。其中, 每个粒子的维度为 5,

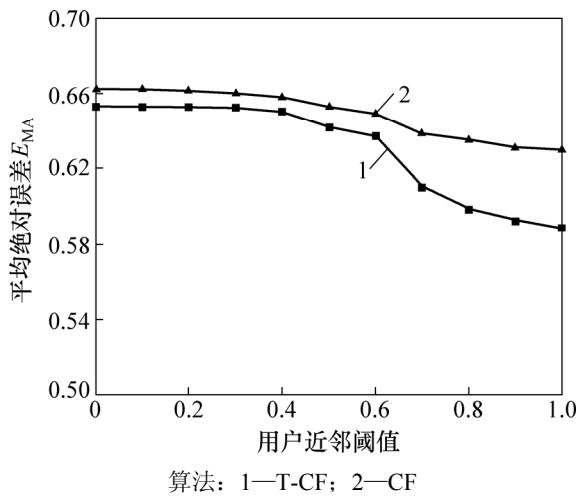


图 3 传统 CF 算法与 T-CF 性能对比图

Fig. 3 Contrast figure of traditional CF and T-CF performance

分别代表项目评分和内容权值、用户和项目推荐权值、项目近邻阈值、用户评分和内容权值以及用户近邻阈值。实验前设定 3 种算法使用同一个规模为 30 的种群作为初始种群, 迭代 30 次。另外, PSO 中粒子速度变化范围为[0, 1], 位置变化范围为[0, 1], 得到的对比结果如图 4 所示。

图 4 中横坐标代表迭代次数, 纵坐标代表每次迭代所得到的平均绝对误差  $E_{MA}$ , 从图 4 可以看出: 随着迭代次数的增加, 3 种算法最优的适应度都在减小, 但是 CQPSO-HR 的优化性能优于 PSO 和 ABC 算法, 能够获得更低的平均绝对误差  $E_{MA}$ 。综上所述, 本文的 CQPSO-HR 算法得到了很好的参数优化效果, 提高

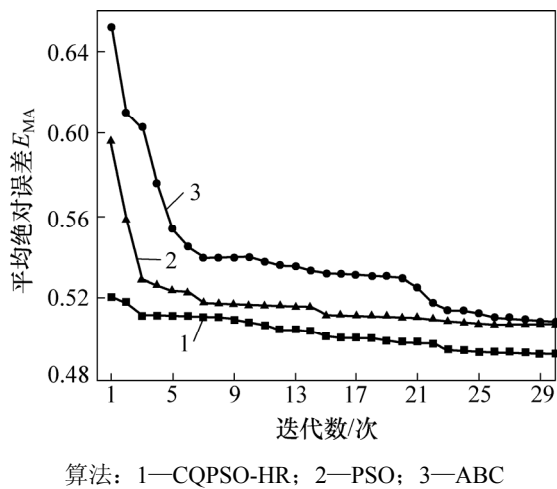


图 4 PSO, ABC 和 CQPSO-HR 性能对比图

Fig. 4 Contrast figure of PSO, ABC and CQPSO-HR performance

了推荐的准确度。在 CQPSO-HR 算法的 10 次实验中, 最理想的  $E_{MA}$  为 0.488 1, 对应的最佳参数组合如表 2 所示。

表 2 最佳参数组合

Table 2 Optimal best parameters	
CQPSO 粒子参数	最佳值
项目评分和内容权值 $w_1$	0.895 6
用户和项目推荐权值 $w_5$	0.212 9
项目近邻阈值 $w_4$	0.830 6
用户评分和内容权值 $w_2$	0.292 6
用户近邻阈值 $w_3$	0.983 8

表 2 中的最佳参数组合表明: 将项目评分和内容相结合, 用户评分和内容相结合, 可以提高评分预测的准确度, 进而提高最终推荐的准确度。

为验证本文的算法在数据稀疏性更差的情况下也能取得很好的效果, 实验中随机抽掉数据集中的若干条评分记录以提高稀疏度, 重复以上实验 4 次, 并与传统的协同过滤算法进行比较, 得如表 3 所示的不同数据稀疏度下的  $E_{MA}$ 。

表 3 不同数据稀疏度下的  $E_{MA}$

Table 3 $E_{MA}$ values under different data sparseness degrees			
数据稀疏度( $S_p$ )	CQPSO-HR	PSO	CF
0.94	0.488 7	0.504 3	0.629 2
0.95	0.495 6	0.532 6	0.657 3
0.96	0.512 2	0.568 8	0.694 6
0.97	0.530 6	0.643 7	0.763 3
0.98	0.564 8	0.734 6	0.850 9

从表 3 可见: 在数据稀疏度逐渐提高的情况下, 虽然本文提出的 CQPSO-HR 算法的  $E_{MA}$  在逐渐提高, 但变化范围较小, 推荐效果仍然优于传统协同过滤算法。

3.2.3 测试 CQPSO-HR 解决冷启动的能力

为验证本文的算法可解决冷启动问题, 将粒子结构做如下约束: 项目评分和内容权值设为 0, 用户评分和内容权值设为 0。在上述条件下用 CQPSO-HR 算法、PSO 算法以及 ABC 算法进行参数优化, 得到最优的参数组合以及最小  $E_{MA}$  分别如表 4~6 所示。

由表 4~6 可见: 在用户和项目冷启动的情况下, 本文提出的 CQPSO-HR 算法依然能够照常推荐, 而且平均绝对误差仅为 0.594 9; 另外, CQPSO-HR 在推荐

表 4 冷启动下 CQPSO-HR 搜索到的最优参数

Table 4 Optimal parameters under cold start using CQPSO-HR

参数	最佳值
项目评分和内容权值	0
用户和项目推荐权值	0.693 8
项目近邻阈值	0.817 0
用户评分和内容权值	0
用户近邻阈值	0.897 7
$E_{MA}$	0.594 9

表 5 冷启动下 PSO 搜索到的最优参数

Table 5 Optimal parameters under cold start using PSO

参数	最佳值
项目评分和内容权值	0
用户和项目推荐权值	0.521 7
项目近邻阈值	0.792 0
用户评分和内容权值	0
用户近邻阈值	0.953 3
$E_{MA}$	0.864 3

表 6 冷启动下 ABC 搜索到的最优参数

Table 6 Optimal parameters under cold start using ABC

参数	最佳值
项目评分和内容权值	0
用户和项目推荐权值	0.834 1
项目近邻阈值	0.887 6
用户评分和内容权值	0
用户近邻阈值	0.854 5
$E_{MA}$	0.966 7

准确度方面高于传统 PSO 与 ABC。

## 4 结论

1) 引入时间因子,能够很好地描述用户兴趣的变化情况,使相似度的计算更可靠。

2) 本文提出的算法能够更快速而准确地找到具有最佳推荐效果的推荐参数组合,不仅在数据稀疏度提高的情况下能获得很好的推荐效果,而且在冷启动的情况下仍具有较高的预测准确度,说明其具有很强的抗数据稀疏性能力和抗冷启动能力。

## 参考文献:

- [1] 孟祥武, 胡勋, 王立才, 等. 移动推荐系统及其应用[J]. 软件学报, 2013, 24(1): 91-108.  
MENG Xiangwu, HU Xun, WANG Licai, et al. Mobile recommender systems and their applications[J]. Journal of Software, 2013, 24(1): 91-108.
- [2] 王立才, 孟祥武, 张玉洁. 上下文感知推荐系统[J]. 软件学报, 2012, 23(1): 1-20.  
WANG Licai, MENG Xiangwu, ZHANG Yujie. Context-aware recommender systems[J]. Journal of Software, 2012, 23(1): 1-20.
- [3] Jeong B, Lee J, Cho H. User credit-based collaborative filtering[J]. Expert Systems with Applications, 2009, 36(3): 7309-7312.
- [4] 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.  
DENG Ailin, ZUO Ziye, ZHU Yangyong. Collaborative filtering recommendation algorithm based on item clustering[J]. Journal of Chinese Computer Systems, 2004, 25(9): 1665-1670.
- [5] 孟宪福, 陈莉. 基于贝叶斯理论的协同过滤推荐算法[J]. 计算机应用, 2009, 29(10): 2733-2735.  
MENG Xianfu, CHEN Li. Collaborative filtering recommendation algorithm based on Bayesian theory[J]. Journal of Computer Applications, 2009, 29(10): 2733-2735.
- [6] 黄正. 面向数据稀疏的协同过滤推荐算法研究与优化[D]. 广州: 华南理工大学计算机科学与工程学院, 2012: 8-37.  
HUANG Zheng. Research and optimization on collaborative filtering recommendation algorithms for data sparsity[D]. Guangzhou: South China University of Technology. School of Computer Science & Engineering, 2012: 8-37.
- [7] 赵丽嫒. 一种新型的协同过滤推荐算法[D]. 南京: 南京邮电大学计算机学院, 2013: 6-46.  
ZHAO Liman. A novel collaborative filtering recommendation algorithm[D]. Nanjing: Nanjing University of Posts and Telecommunications. School of Computer Science & Technology, 2013: 6-46.
- [8] 毕晓君, 王艳娇. 改进人工蜂群算法[J]. 哈尔滨工程大学学报, 2012, 33(1): 117-123.  
BI Xiaojun, WANG Yanjiao. A modified artificial bee colony algorithm and its application[J]. Journal of Harbin Engineering University, 2012, 33(1): 117-123.
- [9] 高卫峰, 刘三阳, 姜飞, 等. 混合人工蜂群算法[J]. 系统工程与电子技术, 2011, 33(5): 1167-1170.  
GAO Weifeng, LIU Sanyang, JIANG Fei, et al. Hybrid artificial bee colony algorithm[J]. Systems Engineering and Electronics, 2011, 33(5): 1167-1170.
- [10] 姜维, 庞秀丽. 面向数据稀疏问题的个性化组合推荐研究[J]. 计算机工程与应用, 2012, 48(21): 21-25.

- JIANG Wei, PANG Xiuli. Research on personal hybrid recommendation overcoming data sparse problem[J]. Computer Engineering and Applications, 2012, 48(21): 21–25.
- [11] Choi S M, Ko S K, Han Y S. A movie recommendation algorithm based on genre correlations[J]. Expert Systems with Applications, 2012, 38(3): 399–405.
- [12] 秦继伟, 郑庆华, 郑德立, 等. 结合评分和信任的协同推荐算法[J]. 西安交通大学学报, 2013, 47(4): 100–104.
- QIN Jiwei, ZHENG Qinghua, ZHENG Deli, et al. A collaborative recommendation algorithm based on ratings and trust[J]. Journal of Xi'an Jiaotong University, 2013, 47(4): 100–104.
- [13] 邢春晓, 高凤荣, 战思南, 等. 适应用户兴趣变化的协同过滤推荐算法[J]. 计算机研究与发展, 2007, 44(2): 296–301.
- XING Chunxiao, GAO Fengrong, ZHAN Sinan, et al. A collaborative filtering recommendation algorithm incorporated with user interest change[J]. Journal of Computer Research and Development, 2007, 44(2): 296–301.
- [14] Wanaskar U H, Vij S R, Mukhopadhyay D. A hybrid web recommendation system based on the improved association rule mining algorithm[J]. Journal of Software Engineering and Applications, 2013, 6(8): 396–404.
- [15] 朱郁筱, 吕琳媛. 推荐系统评价指标综述[J]. 电子科技大学学报, 2012, 41(2): 163–175.
- ZHU Yuxiao, LÜ Linyuan. Evaluation metrics for recommender systems[J]. Journal of University of Electronic Science and Technology of China, 2012, 41(2): 163–175.
- [16] 杨文东, 张强勇, 李术才, 等. 粒子群算法在时效变形参数反演中的应用[J]. 中南大学学报(自然科学版), 2013, 44(1): 282–288.
- YANG Wendong, ZHANG Qiangyong, LI Shucai, et al. Application of particle swarm optimization in time-dependent parameters inversion[J]. Journal of Central South University (Science and Technology), 2013, 44(1): 282–288.
- [17] SUN Jun, FENG Bin, XU Wenbo. Particle swarm optimization with particles having quantum behavior[C]//Evolutionary Computation. Portland, USA, 2004: 325–331.
- [18] 相荣荣. 协同量子粒子群优化及其应用研究[D]. 西安电子科技大学电子工程学院, 2012: 17–30.
- XIANG Rongrong. Cooperative quantum-behaved particle swarm optimization and its applications[D]. Xi'an: Xi'an University of Electronic Science and Technology. School of Electronic Engineering, 2012: 17–30.
- [19] QIAN Wang, YUAN Xianhu, SUN Min. Collaborative filtering recommendation algorithm based on hybrid user model[C]//Proceedings of 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery(FSKD). Yantai, China, 2010: 1985–1990.

(编辑 刘锦伟)