

# 基于相似云与复合因素度量的个性化推荐算法

孙光明<sup>1</sup> 王 硕<sup>2</sup> 李伟生<sup>1</sup>

(北京交通大学计算机与信息技术学院 北京 100004)<sup>1</sup>

(河北科技大学信息科学与工程学院 石家庄 050035)<sup>2</sup>

**摘 要** 针对相似计算中评分数据的稀疏性、属性严格匹配与单因素度量的偶然性导致的近邻不准问题,提出基于相似云与复合因素度量的个性化推荐算法。该算法定义按类的项目评分云来预填充评分,提出融合类别、评分均值、评分频度、访问频度等多因素度量的项目兴趣度向量,通过云模型计算项目相似度,以按类预测其评分,并基于新的加权平均方法计算其最终评分值。实验结果表明,所提算法产生的近邻更准,推荐质量更高。

**关键词** 相似云,复合因素度量,相似度,加权平均,个性化算法

中图分类号 TP311 文献标识码 A DOI 10.11896/j.issn.1002-137X.2016.8.034

## Personalized Recommendation Algorithm Based on Similar Cloud and Measurement by Multifactor

SUN Guang-ming<sup>1</sup> WANG Shuo<sup>2</sup> LI Wei-sheng<sup>1</sup>

(School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100004, China)<sup>1</sup>

(School of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang 050035, China)<sup>2</sup>

**Abstract** Aiming at the problems of sparse data, occasionality that caused by attribute's strict matching and measuring with single rating in similarity computation, this paper presented a personalized recommendation algorithm based on similar cloud and measurement by multifactor. The algorithm defines a marking cloud classified by item to fill sparse matrix, and puts forward a feature vector of item's interest degree consisting of item category, mean score, frequency of rating and accessing, which contributes to calculating the item's similarity with cloud model. On this basis, this algorithm predicts item's score in different categories according to the nearest neighbors, and gets the item's finally score based on a new weighted average computing method. Experimental results show that the algorithm produces more accurate neighbors and better recommendation quality.

**Keywords** Similar cloud, Measurement by multifactor, Similarity degree, Weighted average, Personalized algorithm

## 1 引言

Web 2.0 以用户为中心,在满足大规模用户需求时“主动化、个性化、自适应”服务的理念改变了 Web 应用架构和人们使用互联网的习惯,信息服务也由被动搜索转变为主动推送,因此个性化推荐技术应运而生。尤其是当前网络世界“信息过载”、商品信息自适应用户偏好的网购新模式,使个性化推荐成为研究的热点<sup>[1]</sup>。

个性化信息服务(Personalized Information Service, PIS)以用户为中心,通过研究用户的行为、兴趣、爱好、习惯和环境等个性化特征,自动为用户聚合满足其个性化偏好的信息并主动推荐,达成信息“按需自适应服务”的目标。其关键是依靠用户或项目的偏好、属性等因素计算相似性,找到兴趣最近邻,以预测目标用户对项目的好恶。协同过滤算法是迄今为止研究和应用最为成功的个性化信息服务技术,它以用户对项目的评分矩阵为数学模型,根据用户对项目的评分,采用余弦相似性、Pearson 相关系数<sup>[2]</sup>计算出用户或项目的兴趣最近

邻。在假设兴趣相似用户对同一项目往往有相似评分的前提下,利用兴趣近邻预测目标用户对该项目的兴趣值。由于用户主观上不愿评分,使得评分矩阵极其稀疏,导致产生的最近邻与实际往往不符。为此,研究者们提出了大量解决评分稀疏问题的方法:文献[3,4]利用项目类别缩小小评分范围来提高数据稠密度,以克服数据稀疏导致相似性计算不准的问题;文献[5-7]利用概率矩阵分解模型对评分矩阵降维以减少数据稀疏性,更有效地预测用户实际评分,提升推荐精度;文献[8]通过分析用户行为,利用其隐式兴趣特征来预测其对项目的评分,以避免用户不愿评分导致的数据稀疏。此外,也有研究者提出通过聚类缩小相似用户或项目的搜索范围来解决数据稀疏问题的方法<sup>[9,10]</sup>。

这些研究虽然在一定程度上减小了评分数据的稀疏性,提高了相似性计算的准确度,但是它们都从用户对项目评分整体或局部的单个向量出发,通过严格匹配评分值属性,采用没有包含用户评分统计特征的余弦相似性或 Pearson 相关性方法计算相似度,产生兴趣最近邻。因此,这些方法的最近邻

到稿日期:2015-06-29 返修日期:2015-08-29 本文受河北省高等学校科学技术研究重点项目(ZD2014061)资助。

孙光明(1979-),男,硕士,讲师,主要研究方向为数据挖掘、分布式计算,E-mail: sungmez@gmail.com;王 硕(1979-),女,硕士,讲师,主要研究方向为数据挖掘;李伟生(1945-),男,教授,主要研究方向为分布并行计算。

往往与实际情况不一致,甚至完全相反<sup>[8]</sup>。为此,文献[12-14]提出了利用云模型及项目类别从知识层面粗粒度计算相似性的方法,克服了传统算法中相似性计算严格匹配评分属性导致与客观情况不符的问题;文献[11]在相似性计算前利用云模型对用户评分矩阵进行填充以减小评分数据的稀疏度,同时通过项目属性的重合度计算其相似度,但仍然采用传统相似性计算方法。这类基于云模型在项目类别层面粗粒度计算相似性的方法考虑了用户或项目的整体信息,较好地克服了传统相似性计算中严格匹配评分属性的不足,同时充分利用了评分数据的统计信息,避免了传统相似性计算中侧重相关性而非相似性的弱点,在一定程度上提高了推荐质量。但是,文献[11]通过 Jaccard 相似系数<sup>[12]</sup>计算同类项目属性间的相似度,忽略了同类项目的属性几乎完全相同的事实,即属性的交集与并集几近相同,使得该方法计算出的属性相似度几乎无法使用;文献[12-14]则未考虑同一用户兴趣近邻在不同项目类别往往不同的事实,同时未被评分值全部预设为0,忽略了项目的个性化特征,导致不能从本质上提高算法的推荐质量。

为此,本文提出了一种基于相似云与复合因素度量的个性化推荐算法。首先,定义按类的项目评分云,利用逆向云算法从项目类别的知识层面粗粒度计算项目间的相似度,通过最近邻预填充未评分项;在此基础上,提出基于项目类别、评分均值、评分频度、访问频度等多重统计与隐性因素复合度量的项目兴趣度特征向量,计算项目在不同分类中对用户的吸引程度(兴趣度),利用逆向云算法计算项目间的兴趣度相似性,以分类预测项目的兴趣最近邻,并提出一种新的基于加权平均策略的计算方法预测其最终评分。该算法较好地解决了评分数据稀疏、相似性度量单一和严格匹配项目评分属性导致的近邻计算的偶然性等弊端,使推荐结果与实际情况一致。

## 2 云模型与相似性计算分析

### 2.1 云模型及其在协同过滤算法中的应用与分析

客观世界中人类知识或事物中概念的不确定性主要表现为模糊性和随机性,为了完全集成这两者以构成定性定量间的映射,李德毅院士提出了定性概念与其定量数值表示之间的不确定性转换模型——云模型<sup>[15]</sup>。云模型被广泛应用在不确定人工智能领域,包括智能控制、模糊评测等众多场景。

**定义 1(云和云滴)** 设  $U$  是一个用数值表示的定量论域,  $C$  是  $U$  上的定性概念,若定量值  $x \in U$  是定性概念  $C$  的一次随机实现,  $x$  对  $C$  的确定度  $\mu(x) \in [0, 1]$  是有稳定倾向的随机数;  $\mu: U \rightarrow [0, 1], \forall x \in U; x \rightarrow \mu(x)$ , 则  $x$  在论域  $U$  上的分布称为云(Cloud), 记为云  $C(x)$ ; 每一个  $x$  称为一个云滴(Drop)<sup>[4, 11, 12]</sup>。

定义中,每次概率意义下随机实现的样本具有一个随机的确定度,它是一个概率分布,也是模糊集意义下的隶属度,体现了随机性和模糊性的关联性。云描述了概念的整体特性,而云滴则是众多概念个体中的某一个。云模型中,正态云模型使用一组相互独立的参数共同表达云的数字特征,具有良好的数学性质,便于描述大量事物的不确定性,得到了广泛的应用。在正态云模型中,使用期望  $Ex$ 、熵  $En$ 、超熵  $He$  3 个数字特征来描述云,  $C(Ex, En, He)$  称为云特征向量,各自含义如下。

**定义 2( $Ex, En, He$ )**  $Ex$  为云滴在论域空间  $U$  上分布的期望,是概念量化的最典型样本点和最能代表该定性概念的点;熵  $En$  反映了某个定性概念的可度量程度和不确定性,表示在论域空间可被定性接受的取值范围,其值越大,概念越宏观;超熵  $He$  反映代表定性概念值的样本出现的随机性,是熵的不确定性度量,揭示了模糊性和随机性的关联<sup>[12, 13, 16]</sup>。

将云模型应用于协同过滤算法时,每个用户对所有项目的评分向量被定义为一“朵”云,某个用户对某个项目的评分项被定义为一个云滴。于是,实现了云特征向量  $C(Ex, En, He)$  在知识层面从整体上表述用户评分,进而通过云相似计算取代传统的相关性计算,克服了传统协同过滤算法中严格匹配评分属性的弊端,充分利用了用户、项目及评分的统计特征,使得相似性计算更符合客观情况。

在云模型中,通过正向云算法将代表定性概念整体的云特征向量  $C(Ex, En, He)$  转换为表示个体定量数值的云滴,实现概念空间到数值空间的转换;通过逆向云算法将一组定量数值——云滴,转换为以数字特征  $\{Ex, En, He\}$  表示的定性概念——云,以实现定量值到定性概念的转换。

本文主要利用逆向云算法将代表个体用户或项目的评分转换为表示用户评分整体特征的云特征向量,为此不再赘述正向云算法,逆向云算法计算过程如下。

Step1: 对  $N$  个云滴  $\{x_1, x_2, \dots, x_i, \dots, x_N\}$ , 按式(1)、式(2)分别计算其样本均值  $\bar{X}$ 、样本均值  $S^2$ 。

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N x_i \quad (1)$$

$$S^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2 \quad (2)$$

Step2: 计算  $Ex, En, He$  的估计值  $\hat{Ex}, \hat{He}, \hat{En}$ , 如式(3)~式(5)所示。

$$\hat{Ex} = \bar{X} \quad (3)$$

$$\hat{He} = \sqrt{\frac{\pi}{2}} \times \frac{1}{N} \sum_{i=1}^N |x_i - \bar{X}| \quad (4)$$

$$\hat{En} = \sqrt{S^2 - \frac{1}{3} \hat{He}^2} \quad (5)$$

基于逆向云算法就能将用户对一组项目的评分云滴转换为具有整体评分特性的用户评分云特征向量,接下来就可以基于用户评分云特征向量从知识层面计算两个用户间的相似性。

### 2.2 相似性计算与分析

在传统的协同过滤算法中,用户或项目的相似性计算方式主要包括余弦相似性、修正的余弦相似性和 Pearson 相关系数 3 种,它们把用户评分看作  $n$  维项目空间的向量,用户间的相似性分别通过式(6)~式(8)来计算。

$$sim(u_1, u_2)_{Cosine} = \cos(\bar{u}_1, \bar{u}_2) = \frac{\bar{u}_1 \cdot \bar{u}_2}{\|\bar{u}_1\| \cdot \|\bar{u}_2\|} \quad (6)$$

$$sim(u_1, u_2)_{Adjusted-Cosine} = \frac{\sum_{i \in I_{u_1 u_2}} (R_{u_1, i} - \bar{R}_{u_1})(R_{u_2, i} - \bar{R}_{u_2})}{\sqrt{\sum_{i \in I_{u_1}} (R_{u_1, i} - \bar{R}_{u_1})^2} \sqrt{\sum_{i \in I_{u_2}} (R_{u_2, i} - \bar{R}_{u_2})^2}} \quad (7)$$

$$sim(u_1, u_2)_{Correlation} = \frac{\sum_{i \in I_{u_1 u_2}} (R_{u_1, i} - \bar{R}_{u_1})(R_{u_2, i} - \bar{R}_{u_2})}{\sqrt{\sum_{i \in I_{u_1 u_2}} (R_{u_1, i} - \bar{R}_{u_1})^2} \sqrt{\sum_{i \in I_{u_1 u_2}} (R_{u_2, i} - \bar{R}_{u_2})^2}} \quad (8)$$

其中,  $u_1, u_2$  分别表示两个任意用户;  $\bar{u}_1, \bar{u}_2$  分别表示用户  $u_1, u_2$  的评分向量;  $I_{u_1}, I_{u_2}$  和  $I_{u_1 u_2}$  分别表示用户  $u_1, u_2$  的评分项目集合和他们共同的评分项目集;  $R_{u_1, i}, R_{u_2, i}$  分别表示用户  $u_1, u_2$  对项目  $i$  的评分;  $\bar{R}_{u_1}, \bar{R}_{u_2}$  分别表示用户  $u_1, u_2$  对所有项目的评分均值。

分析传统相似性计算方法可知,三者均通过严格匹配评分属性进行相似度计算,不同之处在于:余弦相似性没有体现用户评分的统计特征,尤其是不同用户的评分尺度不一致;修正余弦相似性通过减去评分均值以弥补余弦相似性的不足,但是它更多体现的是用户间的相关性而非相似性,这是两个截然不同的概念:相关性体现组合的特征,反映两个概念互相关联的程度;相似性表示聚合的特征,反映两个概念互相相似的程度。在相关相似中,由于评分数据的稀疏,不同用户拥有的共同评分更是少之又少,使得其几乎不能用于实际计算。为此,本文通过提出相似云来计算用户和项目间的相似度,以提高其计算准确度。

**定义 3(相似云)** 设  $C_i, C_j$  为给定的两个云,  $\overline{VC}_i, \overline{VC}_j$  分别为其对应的云特征向量,则  $\overline{VC}_i$  与  $\overline{VC}_j$  之间夹角的余弦值  $\cos(\overline{VC}_i, \overline{VC}_j)$  称为云  $C_i$  和  $C_j$  的相似度  $sim(C_i, C_j)$ ; 当  $\cos(\overline{VC}_i, \overline{VC}_j)$  不小于某个给定的阈值  $\delta$  时,称云  $C_i, C_j$  为相似云。

$$sim(C_i, C_j)_{Yun} = \cos(\overline{VC}_i, \overline{VC}_j) = \frac{\overline{VC}_i \cdot \overline{VC}_j}{\|\overline{VC}_i\| \times \|\overline{VC}_j\|} \quad (9)$$

其中,  $\overline{VC}_i = C_i(Ex_i, En_i, He_i), \overline{VC}_j = C_j(Ex_j, En_j, He_j)$ ;  $\delta$  的值由项目最近邻个数  $K$  确定。

可见,云特征向量  $C(Ex, En, He)$  中的期望  $Ex$  反映了用户的偏好水平,熵  $En$  反映了评分偏好的离散度,超熵  $He$  反映了评分偏好离散性的稳定度。基于云特征向量的余弦相似度计算综合了用户评分的统计特征,克服了余弦相似性不能反映评分统计特征的不足,避免了传统相似性计算严格匹配评分属性特征导致的误推荐问题,实现了从知识层面的粗粒度相似性计算,提高了相似度计算的准确性,能够获得更高的推荐质量。

### 3 基于相似云与复合因素度量的个性化推荐算法

为了提高项目间相似性计算的准确度,本文提出了一种基于相似云与复合因素度量的个性化推荐算法。针对评分数据稀疏的问题,定义了基于类别的项目评分云,一方面,利用项目类别在一定程度上滤除未评分项以提高数据稠密度;另一方面,通过逆向云算法从包含项目统计特征的知识层面粗粒度计算相似度,并利用产生的项目最近邻对未评分项目分值进行预填充。针对已有算法的单一评分因素相似性度量方法导致计算出的用户偏好度与项目兴趣度具有偶然性、不客观性的问题,提出了一种基于项目类别、评分均值、评分频度、访问频度等多重统计与隐性因素复合度量的项目兴趣度特征向量,以计算项目在不同分类中的兴趣度;并利用逆向云算法按类别计算项目间的兴趣度相似性,产生项目在不同类别中的兴趣最近邻并预测其评分。最后,综合目标用户及项目在不同类别中的偏好和兴趣度,提出一种基于加权平均策略的计算方法来计算目标用户对目标项目的最终综合偏好值。

接下来本文将围绕上述研究内容进行阐述。

#### 3.1 基于类别的项目评分云及评分预填充方法

同一个项目可以分属多个不同的项目类别,且在不同的项目分类中对用户的吸引程度及兴趣近邻往往是不同的。如果通过整个用户评分矩阵来预填充未评分项<sup>[1]</sup>,对单个项目而言,扩大了相似性计算的数据范围,在评分矩阵本身较稀疏的前提下实际上是进一步加剧了数据的稀疏程度,同时稀释了用户对项目在不同分类中的兴趣度,削弱了项目在不同分类中的个性化特征,使得计算出的兴趣近邻间相似性不明显。

另一方面,对于评分矩阵中已获评分的项目,使用其评分参与近邻计算来作为评分预填充的依据具有一定的正确性,但是对于未评分的项目,使用统一预设值 0 或评分频度作为评分值参与近邻计算,例如文献<sup>[3, 4, 11-13]</sup>,在项目没有被评分的条件下其做法都是不妥的,因为不同项目的偏好值不同,且没有评分是无法计算出评分频度的。

为此,本文定义了一种基于评分类别、评分值、评分次数与类别评分均值的按类项目评分云,以预填充用户评分矩阵,解决用于相似性计算的评分数据稀疏性问题。

**定义 4(基于类别的评分云)** 给定  $k$  个项目  $I_{i,1}, I_{i,2}, \dots, I_{i,k}$ , 其所属项目类别为  $CI_i$ , 项目  $I_{i,j} (1 \leq j \leq k)$  的评分值为  $PCI_{i,j}$ , 则称向量  $YCI_i = (PCI_{i,1}, PCI_{i,2}, \dots, PCI_{i,j}, \dots, PCI_{i,k})$  为项目类别  $CI_i$  的评分云,任意一个项目  $I_{i,j} (1 \leq j \leq k)$  的评分值  $PCI_{i,j}$  为一个评分云滴,  $C_i \{Ex_i, En_i, He_i\}$  为云  $YCI_i$  的云特征向量。其中,  $PCI_{i,j} (1 \leq j \leq k)$  的计算方法如式(10)所示。

$$PCI_{i,j} = \begin{cases} r_{i,j}, & \text{if the } I_{i,j} \text{ has been rated} \\ \frac{T_{i,j}}{T_i} \bar{R}_i, & \text{if the } I_{i,j} \text{ has been accessed, but not rated} \end{cases} \quad (10)$$

其中,  $i$  为项目的类别编号,  $k$  为每个项目类别中包含的项目个数,  $I_{i,j} (1 \leq j \leq k)$  表示第  $i$  类项目中的第  $j$  个项目,  $Ex_i, En_i, He_i$  分别为云  $YCI_i$  的期望、熵和超熵,  $r_{i,j}$  为项目  $I_{i,j}$  的评分,  $T_{i,j}$  为项目  $I_{i,j}$  在项目类别  $CI_i$  中的被访问次数,  $T_i$  为项目类别  $CI_i$  中所有项目的总评分次数,  $\bar{R}_i$  为项目类别  $CI_i$  中所有项目评分均值。

利用式(10)计算出基于类别的评分云后,即可使用算法 1 对用户评分矩阵进行预填充。

#### 算法 1 基于分类项目评分云的分值预填充算法

输入: 用户评分矩阵  $R_{m,n}$ 、项目类别矩阵  $CI = (C_1, C_2, \dots, C_i, \dots, C_l)^T$   
输出: 预填充的项目类别评分矩阵  $CI'$

步骤:

- 1) 由  $CI$  获得基于类别的项目评分云  $YCI_i (1 \leq i \leq l)$ ,  $l$  为项目类别数。
- 2) 依据式(10)确定  $YCI_i$  中各项目评分值  $PCI_{i,j} (1 \leq j \leq k)$ ,  $k$  为项目类别中的项目个数。
- 3) 依据式(3)一式(5)计算基于类别的项目评分云  $YCI_i (1 \leq i \leq l)$  的云特征向量  $C_i \{Ex_i, En_i, He_i\}$ 。
- 4) 依据式(9)计算项目评分云  $YCI_i (1 \leq i \leq l)$  间的相似度,产生项目评分云的最近邻。
- 5) 基于产生的评分云最近邻,利用评分预测式(11),采用对项目评分并集填充<sup>[3]</sup>的方式对项目类别评分矩阵进行预填充。

$$PR_{u_i, j} = \bar{R}_i + \frac{\sum_{q \in S(i)} [sim(i, q)_{Yun} \times (R_{u_i, q} - \bar{R}_q)]}{\sum_{q \in S(i)} sim(i, q)_{Yun}} \quad (11)$$

其中,  $PR_{u,i}$  为用户  $u$  对项目类别  $i$  中目标项目  $j$  的预测分;  $S(i)$  为项目评分云  $YCI_i$  的最近邻集合;  $R_{u,q}$  为用户  $u$  对项目类别  $q$  中目标项目  $j$  的预测分。

对项目评分并集填充是指对于云  $YCI_i, YCI_q$ , 若其评分用户集合分别为  $U_i, U_q$ , 则只填充用户  $U_i \cup U_q$  中没有对项目  $I_{i,j}$  或  $I_{q,j}$  评分的用户的评分, 而不考虑对  $I_{i,j}$  和  $I_{q,j}$  均未评分的用户。

6) 重复步骤 5), 产生预填充的项目类别评分矩阵  $CI'$ 。

不难看出, 本文提出的预填充算法结合同一类别中不同项目的被关注度(访问次数)与同一项目在不同类别中的评分两个维度, 对未评分项目的分值进行预测, 并充分利用了云模型在知识层面粗粒度地计算相似性, 避免了利用稀疏矩阵单一分值填充未评分数据的不切实际做法, 较好地克服了数据稀疏的问题。

### 3.2 复合因素度量的项目兴趣度特征向量

用户对项目的评分不能完全反映用户的真实兴趣, 具有一定的偶然性, 例如受用户心情、环境氛围等因素的影响。为此, 在基于项目评分云对稀疏的用户评分矩阵进行填充生成项目类别评分矩阵  $CI'$  后, 则不能仅从评分的角度来度量项目的相似度量。为此, 文献[2]提出了一种基于项目类别、评分值与次数的项目兴趣度特征向量, 来度量项目对用户的吸引程度, 以避免单一评分相似度量中存在的偶然性因素, 其定义如式(12)所示。

$$IC_{i,j} = \frac{\sum_{u=1}^m (r_{u,i} - \bar{S}_j)}{N_j} \quad (12)$$

其中,  $IC_{i,j}$  表示项目类别  $j$  中项目  $i$  的兴趣度,  $m$  为对项目类别  $j$  评过分的用户数,  $\bar{S}_j$  为项目类别  $j$  中所有项目的平均评分,  $N_j$  为项目类别  $j$  中所有项目的评分次数。分析  $IC_{i,j}$  可知, 文献[2]定义的项目兴趣度特征值只反映了项目对已评分用户的兴趣度, 而没有反映出对访问了项目而未评分用户的兴趣度, 未能从本质上消除评分的偶然性。

为此, 本文从项目评分均值、项目类别评分均值、评分频度、访问频度 4 个角度出发, 提出融合项目类别、评分值、评分次数、访问次数等多重因素的复合度量项目兴趣度的方法。

**定义 5(项目评分频度)** 设  $TPI_{i,j}$  为项目  $i$  在项目类别  $j$  中的被评分次数,  $TAI_{i,j}$  为项目  $i$  在项目类别  $j$  中的访问次数, 则称  $\frac{TPI_{i,j}}{TAI_{i,j}}$  为项目  $i$  在类别  $j$  中的评分频度。

**定义 6(项目类别评分频度)** 设  $TPCI_j$  为项目类别  $j$  所有项目被评分次数,  $TACI_j$  为项目类别  $j$  所有项目被访问次数, 则称  $\frac{TPCI_j}{TACI_j}$  为项目类别  $j$  的评分频度。

**定义 7(访问频度)** 项目  $i$  在项目类别  $j$  中的访问次数  $TAI_{i,j}$  与项目类别  $j$  所有项目被访问次数  $TACI_j$  的比值  $\frac{TAI_{i,j}}{TACI_j}$  称为项目  $i$  在项目类别  $j$  中的访问频度。

因此, 基于项目评分均值、项目类别评分均值、评分频度、访问频度等多因素复合度量的项目兴趣度特征值计算方法如式(13)所示。

$$IC_{i,j} = (\bar{S}_{i,j} - \bar{S}_j) \times \frac{TPI_{i,j}}{TAI_{i,j}} \times \frac{TPCI_j}{TACI_j} \times \frac{TAI_{i,j}}{TACI_j} \\ = \frac{(\bar{S}_{i,j} - \bar{S}_j) \times TPI_{i,j} \times TPCI_j}{TACI_j^2} \quad (13)$$

其中,  $\bar{S}_{i,j} = \frac{1}{TPI_{i,j}} \sum_{u=1}^m r_{u,i}$ ,  $\bar{S}_j = \frac{1}{TPCI_j} \sum_{i=1}^k r_{i,j}$  分别为项目  $i$  在

项目类别  $j$  中的平均评分及项目类别  $j$  中所有项目的平均评分。

于是, 可得到类别  $j$  的项目兴趣度特征向量  $ICI_j = (ICI_{1,j}, ICI_{2,j}, \dots, ICI_{i,j}, \dots, ICI_{k,j})$ , 其中  $k$  为项目类别  $j$  中的项目个数,  $1 \leq i \leq k$ 。

### 3.3 基于加权平均策略的评分计算方法

通过逆向云算法产生基于类别的项目兴趣度相似矩阵后, 即可计算出目标项目  $I$  在各类别项目中的兴趣度最近邻集合  $Top-K_{l,c}$  ( $1 \leq c \leq m$ )。此时, 使用式(11), 依据目标用户  $U$  对  $I$  在各项目类别兴趣最近邻集合  $Top-K_{l,c}$  相关项目的评分, 即可计算出  $U$  对  $I$  在不同类别中的预测分值  $PR_{U,l_1}, PR_{U,l_2}, \dots, PR_{U,l_i}, \dots, PR_{U,l_l}$ , 其中,  $1 \leq i \leq l$ 。

为了预测目标用户  $U$  对目标项目  $I$  的偏好程度以决定是否产生推荐, 本文提出了基于类别分值平均加权的项目最终预测分值计算方法, 如式(14)所示。

$$PR_{U,I} = \lambda_{l_1} \cdot P_{U,l_1} + \lambda_{l_2} \cdot P_{U,l_2} + \dots + \lambda_{l_i} \cdot P_{U,l_i} + \dots + \lambda_{l_l} \cdot P_{U,l_l} \quad (14)$$

其中,  $PR_{U,I}$  为目标用户  $U$  对目标项目  $I$  的最终评分,  $\lambda_{l_i}$  为  $U$  对类别  $i$  中项目  $I$  评分值的平均加权因子, 采用式(15)计算其值。

$$\lambda_{l_i} = \frac{\sum_{q=1}^{N_i} sim(i, I, q)_{Y_{im}}}{\sum_{i=1}^l \sum_{q=1}^{N_i} sim(i, I, q)_{Y_{im}}} \quad (15)$$

其中,  $N_i$  为第  $i$  个项目类别中项目的个数,  $l$  为项目类别数,  $sim(i, I, q)_{Y_{im}}$  为目标项目  $I$  与类别  $i$  中项目  $q$  的相似度, 其计算方法如式(16)所示。

$$sim(i, I, q) = \cos(\overline{VC}_{i,I}, \overline{VC}_{I,q}) = \frac{\overline{VC}_{i,I} \times \overline{VC}_{I,q}}{\|\overline{VC}_{i,I}\| \cdot \|\overline{VC}_{I,q}\|} \quad (16)$$

其中,  $\overline{VC}_{i,I} = (Ex_{i,I}, En_{i,I}, He_{i,I})$ ,  $\overline{VC}_{I,q} = (Ex_{I,q}, En_{I,q}, He_{I,q})$  分别为项目  $I, q$  对同一项目类别  $i$  的评分云特征向量。

### 3.4 基于相似云与复合因素度量的个性化推荐算法与分析

基于上述稀疏数据填充方法、复合因素度量项目兴趣度及采用了加权平均策略的项目评分预测方法, 提出了基于相似云与复合因素度量的个性化推荐算法。

**算法 2** 基于相似云与复合因素度量的个性化推荐算法  
输入: 用户评分矩阵  $R_{m,n}$ 、项目类别矩阵  $CI = (C_1, C_2, \dots, C_i, \dots, C_l)^T$ 、目标用户  $U$ 、目标项目  $I$

输出: 用户  $U$  对  $I$  的评分  $PR_{U,I}$

1) 由算法 1 计算出  $CI'$ 。

2) 由式(13)计算项目兴趣度特征向量  $ICI_j = (ICI_{1,j}, ICI_{2,j}, \dots, ICI_{i,j}, \dots, ICI_{k,j})$ , 得到项目兴趣度矩阵  $MCI$ , 其定义如下:

$$MCI = \begin{bmatrix} ICI_{1,1} & ICI_{2,1} & \dots & ICI_{i,1} \\ ICI_{1,2} & ICI_{2,2} & \dots & ICI_{i,2} \\ \vdots & \vdots & \vdots & \vdots \\ ICI_{1,j} & ICI_{2,j} & \dots & ICI_{i,j} \end{bmatrix} \quad (17)$$

其中, 行为项目, 列为项目类别,  $ICI_{i,j}$  表示第  $j$  类中第  $i$  个项目的兴趣度特征值。

3) 由逆向云算法及式(16)对  $MCI$  计算出基于类别的项目兴趣度相似矩阵  $MSCI$ 。它是一个三维矩阵, 每个元素  $sim(i, p, q)_{Y_{im}}$  为类别  $i$  中项目  $p$  与项目  $q$  的兴趣相似度, 其形式为:

$$MSCI = \begin{bmatrix} \cdot & \vdots & \cdot \\ \dots & sim(i, p, q)_{Y_{im}} & \dots \\ \cdot & \vdots & \cdot \end{bmatrix} \quad (18)$$

- 4) 由 MSCI 计算出目标项目 I 在各类别中的兴趣最近邻集合 Top-K<sub>I,j</sub> (1 ≤ j ≤ I)。
- 5) 依据 Top-K<sub>I,j</sub> 和式(11), 计算 U 对 I 在各类别中的评分 PR<sub>U,I</sub>。
- 6) 对 PR<sub>U,I</sub>, 由式(14)、式(15)计算评分平均加权因子 λ<sub>I</sub> 及最终评分值 PR<sub>U,I</sub>。
- 7) 依据 PR<sub>U,I</sub>, 决定是否对 U 推荐 I。

本文算法充分利用了云模型实现定性定量概念间转换的优势, 通过评分信息的统计特征及项目类别从知识层面粗粒度计算相似性, 避免了严格匹配评分属性导致的极端性和偶然性, 一定程度上克服了单一稀疏评分预填充未评分项的弊端; 提出的基于类别、评分均值、评分次数的按类项目评分云, 较好地避免了基于整体稀疏评分预填充未评分项时进一步稀释数据、削弱项目个性化特征等不切实际的做法, 较好解决了相似性计算数据的稀疏性问题。

另一方面, 提出融合项目评分均值、项目类别评分均值、评分频度、访问频度等多重因素来复合度量项目兴趣度的方法, 充分利用了用户访问项目而不愿评分的隐性偏好行为, 消除了相似性计算中的偶然性; 同时, 基于不同类别预测目标用户对目标项目的评分, 反映了不同项目在不同类别的兴趣最近邻不同的事实, 使计算出的最近邻吻合实际情况。

因此, 本文算法能从本质上提高个性化推荐算法中相似性计算的准确度, 提升推荐质量。

## 4 实验结果及分析

### 4.1 实验数据集

本文实验采用研究型推荐系统 MovieLens 站点 (<http://movielens.umn.edu>) 提供的 100k 公开数据集。它将电影(项目)分为 19 个类别(0 类为未知类, 属少数异常数据), 要求每个用户至少对 20 部电影评分, 共包含了 943 个用户对 1682 部电影的 10 万条评分记录, 数据稀疏等级<sup>[12]</sup>为  $1 - \frac{100000}{943 \times 1682} = 0.9730$ 。

为了明显反映算法的特性与效果, 本文将数据集按照 1:4 的比例分为测试数据集与训练数据集, 并分 5 次随机抽取 1-18 类中 10、20、30、40、50 个用户近邻产生评分矩阵以进行实验。

### 4.2 推荐质量评价标准

本文对算法质量的评价标准采用推荐系统使用最为广泛的平均绝对偏差 (Mean Absolute Error, MAE), 即算法预测的用户评分与实际用户评分之间的偏离程度。MAE 值越小, 算法的推荐质量越高。

假设算法预测的用户评分集合为  $\{p_1, p_2, \dots, p_N\}$ , 相应的实际评分集合为  $\{q_1, q_2, \dots, q_N\}$ , 则 MAE 的定义如式(19)所示。

$$MAE = \frac{\sum_{i=1}^N |p_i - q_i|}{N} \quad (19)$$

### 4.3 实验设计与结果分析

为了验证本文算法的推荐质量, 通过本文算法的 MAE 随用户个数及项目个数变化情况, 及与文献[4, 11]中基于云模型与项目分类推荐算法的 MAE 的对比来进行说明。

实验 1 本文算法的 MAE 随用户个数及项目个数的变化规律

该实验从用户近邻固定时 MAE 随项目邻居个数的变化

情况及项目近邻固定时 MAE 随用户邻居个数的变化情况两个方面展开, 实验结果如图 1 所示。

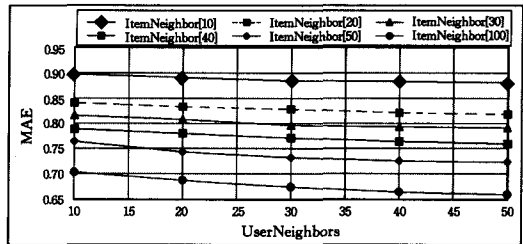


图 1 本文算法 MAE 随用户及项目近邻个数的变化情况

图 1 中, 每条曲线横向为项目近邻分别固定为 10, 20, 30, 40, 50, 100 时, 本文算法的 MAE 随用户近邻的增加而减小的过程。在实际情况下, 用户近邻越多, 相似性计算数据越充分, 对目标项目的预测评分与实际分值越接近, MAE 值就越小; 同理, 纵向为用户近邻固定时, 本文算法 MAE 随项目近邻个数增加而减小的过程, 与实际评分情况也保持一致。另一方面, 每条曲线在用户或项目近邻个数达到一定规模后, MAE 变化趋于稳定, 也反映出只需要一定数目的近邻就基本推算出用户兴趣的事实, 同时也说明本文算法能够在用户评分稀疏的情况下准确计算相似度。

实验 2 验证相似云与复合因素度量方法对推荐质量的影响

该实验中项目近邻个数固定为 100, 比较本文算法与文献[4, 11]中算法的 MAE 随用户近邻个数变化的情况, 其结果如图 2 所示。

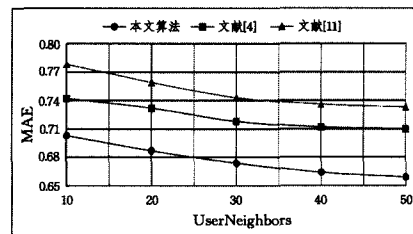


图 2 本文算法与文献[4, 11]算法的 MAE 随用户近邻变化的情况

文献[4]通过填充同类项目的评分及计算同类项目间相似度来解决数据稀疏问题, 以提高推荐质量; 文献[11]则通过计算项目相同属性间的评分相似度产生推荐。文献[4, 11]都基于单一评分依据计算相似性。从图 2 可以看出, 3 种算法的 MAE 值都随用户近邻的增加而减小, 并在用户近邻个数到达一定规模后趋于稳定, 说明 3 种算法正确反映了用户近邻越多, 相似性计算就越准确, 预测评分就越与实际情况一致的事实。

但是, 在相同用户和项目近邻个数的条件下, 文献[4]的 MAE 比文献[11]的小, 而本文的 MAE 又要比文献[4]的小。前者是由于文献[11]忽视了项目类别及评分统计特征等因素对相似性的作用, 仅仅通过严格匹配项目属性评分值, 而不同类的项目间相同属性较少, 导致相似性计算的偶然因素和极端性降低了推荐质量; 而本文算法使用了云模型预填充评分, 通过基于项目类别、类别评分均值、评分频度、访问频度等多因素复合度量的按类项目兴趣特征向量及相似云从知识层面粗粒度计算最近邻, 充分利用了隐性评分行为和统计特征, 减小了评分数据的稀疏性, 克服了文献[4, 11]中单一评分相似性及严格匹配评分的弊端。另一方面, 本文的加权平均策略

基于项目分类预测分值来计算最终评分的方法,反映了同一项目在不同类别中兴趣近邻不同的事实。

### 实验3 余弦相似性度量方法高效性的验证

文献[17]提出一种基于云滴间的距离来计算云相似性的方法,文献[18]提出了期望曲线和最大边界曲线两种正态云模型相似性度量方法。为了验证本文中余弦相似性度量方法的高效性,下面从计算时间和 MAE 两个方面将本文算法与文献[17,18]的方法做比较。

图3为4种算法在项目近邻固定为100时,MAE随用户近邻个数变化的情况。

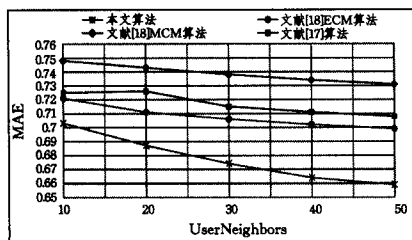


图3 本文算法与文献[17,18]算法的 MAE 随用户近邻变化的情况

图3表明4种算法的MAE随用户近邻个数的增加而减小,推荐结果越准确,与实际情况相符合。文献[17]算法需要对表示评分的云滴进行筛选,表现出一定的不稳定性;文献[18]中MCM算法通过最大边界曲线来逼近评分云模型,虽然能使表示评分云特征的期望、熵、超熵3个数字特征全部用于相似性计算,但是它仅仅反映了云边界的评分分布情况,不能准确地反映整个评分的特征,因而推荐质量较低。文献[18]中的ECM算法和本文算法突出了期望数字特征对相似性度量的影响,一定程度上忽视了熵、超熵的作用,但是由于本文算法中引入了项目类别、类别评分均值、评分频度、访问频度等多因素来复合度量相似性,一定程度上弥补了表示评分偏好的离散度及其稳定度的熵与超熵,因此MAE更小,推荐质量更高。

图4为4种算法在项目近邻固定为100时,推荐时间随用户近邻个数变化的情况。

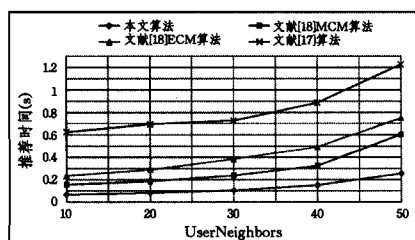


图4 本文及文献[17,18]算法的计算时间随用户近邻变化的情况

由图4可知,在相同的条件下,本文算法的推荐效率最高。原因在于:为了确保计算精度,文献[17]算法需要在 $[Ex-En, Ex+En]$ 范围内产生足够的云滴,并对它们组合排序,因此计算量巨大,计算效率低下;文献[18]的两种算法与本文算法较接近,但由于ECM算法中积分运算和MCM算法中指数运算的计算复杂度较高,导致这两种算法的推荐效率地本文算法低。

**结束语** 本文提出了一种基于相似云与复合因素度量的个性化推荐算法。通过基于项目类别、评分均值、访问次数定义的评分云及云模型预填充评分矩阵,较好地解决了评分数据稀疏问题;利用基于项目类别、评分均值、评分频度及访问

频度多因素度量的项目兴趣度特征向量及逆向云算法消除了传统相似度计算中的偶然性和严格匹配评分属性产生的极端性,提供了相似性计算的准确度;提出基于项目在不同类别中兴趣近邻的加权平均方法来计算目标项目的最终评分,反映了不同项目在不同类别兴趣近邻不同的事实,提高了算法的推荐质量。在当前大数据环境下,如何提高算法计算效率是下一步的研究目标。

### 参考文献

- [1] Hu X, Meng X W, Zhang Y J, et al. Recommendation algorithm combing item features and trust relation of mobile users[J]. Journal of Software, 2014, 24(8):1817-1830(in Chinese)  
胡勋,孟祥武,张玉洁,等.一种融合项目特征和移动用户信任关系的推荐算法[J].软件学报,2014,25(8):1817-1830
- [2] Sun G M, Wang S. Novel collaboration filtering algorithm based on interest degree of item[J]. Application Research of Computers, 2013, 30(12):3618-3621(in Chinese)  
孙光明,王硕.基于项目兴趣度的协同过滤新算法[J].计算机应用研究,2013,30(12):3618-3621
- [3] Deng A L, Zhu Y Y, Shi B L. Accollaborative filtering recommendation algorithm based on item rating prediction[J]. Journal of Software, 2003, 14(9):1621-1628(in Chinese)  
邓爱玲,朱扬勇,施伯乐.基于项目评分预测的系统过滤推荐算法[J].软件学报,2003,14(9):1621-1628
- [4] Xiong Z Y, Liu Q, Zhang Y F, et al. Collaborative filtering recommendation algorithm based on item classification and cloud model[J]. Application Research of Computers, 2012, 29(2):493-496(in Chinese)  
熊忠阳,刘芹,张玉芳,等.基于项目分类的协同过滤改进算法[J].计算机应用研究,2012,29(2):493-496
- [5] Sun G F, Wu L, Liu Q, et al. Recommendation based on collaborative filtering by exploiting sequential behaviors[J]. Journal of Software, 2013, 24(11):2721-2733(in Chinese)  
孙光福,吴乐,刘淇,等.基于时序行为的协同过滤推荐算法[J].软件学报,2013,24(11):2721-2733
- [6] Yin H, Sun Y, Cui B, et al. LCARS: A location-content-aware recommender system[C]// Proc. of the KDD. New York: ACM Press, 2013:221-229
- [7] Lian D, Zhao C, Xie X, et al. GeoMF: Joiny geographical modeling and matrix factorization for point-of-interest recommendation[C]// Proc. of the KDD. New York: ACM Press, 2014:831-840
- [8] Zheng L, Cui S, Yue D, et al. User interest modeling based on browsing behavior[C]// International Conference on Advanced Computer Theory & Engineering. 2010,5:V5-455-V5-458
- [9] Chen K H, Han P P, Wu J. User clustering based social network recommendation[J]. Chinese Journal of Computers, 2013, 36(2):349-359(in Chinese)  
陈克寒,韩盼盼,吴健.基于用户聚类的异构社交网络推荐算法[J].计算机学报,2013,36(2):349-359
- [10] Leng Y J, Lu Q, Zhang J L. Using bipartite network for enhancement of collaborative filtering [J]. Computer Science, 2015, 42(3):256-260(in Chinese)  
冷亚军,陆青,张俊岭.使用二分图网络提高协同推荐的准确性[J].计算机科学,2015,42(3):256-260

(下转第198页)

的测试结果的准确率要高于梯度上升算法和 EM 算法的准确率;特别是对于处于概念模糊区域的样本,EMC 算法的识别准确率有显著提高,可以看出提出的 EMC 参数学习算法能有效处理概念划分过程中概念知识的模糊性和随机性,在参数学习过程中能够考虑到知识的不确定性,对概念模糊区域的参数学习精度也有显著提升,从而提高了网络推理的准确性。

**结束语** 在贝叶斯网络的参数学习过程中需将连续节点离散化,传统的参数学习算法未考虑节点定性概念存在的模糊性和随机性,对概念的划分是一种非此即彼的硬划分,在概念交接处,容易导致样本归属概念选择错误。对此,本文提出了 EMC 参数学习算法,在连续节点概念划分时充分考虑了知识表达存在的模糊性和随机性。仿真实验结果表明,所提的 EMC 参数学习算法能够有效地提高参数学习的精度,特别在概念模糊区域识别精度有明显提高,从而有效提高了网络推理的准确性。由于 EMC 算法增加了节点概念的表达的复杂性,使得参数学习所耗时间增加,如何进一步优化学习及缩短学习时间,是下一步需要继续研究的问题。

### 参考文献

- [1] 肖秦琨,高嵩. 贝叶斯网络在智能信息处理中的应用[M]. 北京:国防工业出版社,2012
- [2] Meng Guang-lei, Gong Guang-hong. Threat assessment of aerial targets based on hybrid Bayesian network[J]. Systems Engineering and Electronics, 2013, 32(11): 2398-2401 (in Chinese)  
孟光磊, 龚光红. 基于混合贝叶斯网的空域目标威胁评估方法[J]. 系统工程与电子技术, 2013, 32(11): 2398-2401
- [3] Fei Zhi-gen, Xu Xiao-jie, Dong Xi-yan. The MLE Algorithm on the Bayes Network for Machine Fault Diagnosis[J]. Henan Science, 2007, 25(1): 98-100 (in Chinese)  
费致根, 徐小洁, 董喜燕. 用于机械故障诊断 Bayes 网络的 MLE 学习方法[J]. 河南科学, 2007, 25(1): 98-100
- [4] Yan Dai-wei, Gu Liang-xian, Pan Lei. A Study of New Method for Weapon System Effectiveness Evaluation Based on Bayesian Network[J]. Journal of China Ordnance, 2008, 3(3): 209-213
- [5] Li Hong, Emmanuel A, Li Ping, et al. Imputation algorithm of

missing values based on EM and Bayesian network[J]. Computer Engineering and Applications, 2010, 46(5): 123-125 (in Chinese)

李宏, 阿玛尼, 李平, 等. 基于 EM 和贝叶斯网络的丢失数据填充算法[J]. 计算机工程与应用, 2010, 46(5): 123-125

- [6] 李德毅, 杜鹞. 不确定性人工智能[M]. 北京: 国防工业出版社, 2005
- [7] Liu Yu-chao. Study on calculating method based on cloud model [D]. Beijing: Tsinghua University, 2012 (in Chinese)  
刘玉超. 基于云模型的粒计算方法研究[D]. 北京: 清华大学, 2012
- [8] Zhang Yin-yan, Li Bi-cheng, Cui Jia-wei. Method of Target Threat Assessment Based on Cloudy Bayesian Network[J]. Computer Science, 2013, 40(10): 127-131 (in Chinese)  
张银燕, 李弼程, 崔家玮. 基于云贝叶斯网络的目标威胁评估方法[J]. 计算机科学, 2013, 40(10): 127-131
- [9] 茆诗松, 程依明, 濮晓龙. 概率论与数理统计[M]. 北京: 高等教育出版社, 2011
- [10] 苏良军. 高等数理统计[M]. 北京: 北京大学出版社, 2007
- [11] Gao Jian-guo, Cui Ye-qin. A New Discretization Method for Continuous Attributes Based on Information Entropy[J]. Microelectronics & Computer, 2011, 28(7): 187-194 (in Chinese)  
高建国, 崔业勤. 基于信息熵理论的连续属性离散化方法[J]. 微电子学与计算机, 2011, 28(7): 187-194
- [12] Zheng Zhong-zhi, Duan Peng, He Xuan-jun, et al. Improvement of Bayesian classification based on the cloud model theory[J]. Journal of Yunnan University of Nationalities (Natural Sciences Edition), 2013, 22(3): 226-229 (in Chinese)  
郑钟志, 段鹏, 贺宣军, 等. 基于云模型理论改进的贝叶斯分类算法[J]. 云南民族大学学报(自然科学版), 2013, 22(3): 226-229
- [13] Zhou Xuan, Wang Lei, Zhu Yan-guang, et al. A Discretization Method of Continuous Variable in Bayesian Network Parameter Learning[J]. Computer Simulation, 2012, 26(10): 136-139 (in Chinese)  
周旋, 王磊, 朱延广, 等. 贝叶斯网络参数学习中连续变量离散化方法研究[J]. 计算机仿真, 2012, 26(10): 136-139

(上接第 170 页)

- [11] Sun J G, Ai L R. Collaborative filtering recommendation algorithm based on item attribute and cloud model filling[J]. Journal of Computer Application, 2012, 32(3): 658-660, 668 (in Chinese)  
孙金刚, 艾丽蓉. 基于项目属性和云填充的协同过滤推荐算法[J]. 计算机应用, 2012, 32(3): 658-660, 668
- [12] Zhang G W, Kang J C, Li H S, et al. Context based on Collaborative filtering recommendation algorithm[J]. Journal of System Simulation, 2006, 18(Suppl. 2): 595-601 (in Chinese)  
张光卫, 康建初, 李鹤松, 等. 面向场景的协同过滤推荐算法[J]. 系统仿真学报, 2006, 18(Suppl. 2): 595-601
- [13] Zhang G W, Li D Y, Li P, et al. A collaborative filtering recommendation algorithm based on cloud model[J]. Journal of Software, 2007, 18(10): 2403-2411 (in Chinese)  
张光卫, 李德毅, 李鹏, 等. 基于云模型的协同过滤推荐算法[J]. 软件学报, 2007, 18(10): 2403-2411
- [14] Li S. Collaborative filtering recommendation algorithm based on cloud clustering of multi-indicators item evaluation[C] // 2011 International Conference on Business Computing and Global Information. Shanghai, China: IEEE CPS, 2011: 654-648

[15] Li X H. Personalized recommendation algorithm research based on Jaccard item category similarity [D]. Changsha: Central South University, 2010: 31-32 (in Chinese)

李小慧. 基于 Jaccard 项目类别相似性的个性化推荐算法研究[D]. 长沙: 中南大学, 2010: 31-32

[16] Li D Y, Liu C Y. Study on the universality of the normal cloud model[J]. Chinese Engineering Science, 2004, 6(8): 28-34 (in Chinese)

李德毅, 刘常昱. 论正态云模型的普适性[J]. 中国工程科学, 2004, 6(8): 28-34

[17] Zhang Y, Zhao D N, Li D Y. The similar cloud and measurement method[J]. Information and Control, 2004, 33(2): 129-132 (in Chinese)

张勇, 赵东宁, 李德毅. 相似云及其度量分析方法[J]. 信息与控制, 2004, 33(2): 129-132

[18] Li H L, Guo C H, Qiu W R. Similarity measurement between normal cloud models[J]. Acta Electronica Sinica, 2011, 39(11): 2561-2567 (in Chinese)

李海林, 郭崇慧, 邱望仁. 正态云模型相似度计算方法[J]. 电子学报, 2011, 39(11): 2561-2567