

协同过滤推荐算法的改进与研究

孙华艳,李业丽,字云飞,韩旭
(北京印刷学院 信息工程学院,北京 102600)

摘要:针对协同过滤算法中数据稀疏性严重影响推荐精准度的问题,提出了基于因果聚类 and 模糊相似关系的协同过滤推荐算法。基于因果聚类的协同过滤推荐算法通过对用户-项目的评分、收藏、特征值等建立矩阵来计算它们之间的相似度,矩阵的建立是基于项目、用户特征值具有因果关联性和相似性;基于模糊相似关系的协同过滤推荐算法通过项目、用户的加权特征法构建三角模糊数来计算相似度,并根据模糊相似矩阵及最佳阈值法来得到连通图截集,从而实现精准、个性化推荐。实验结果表明,这两种算法都能有效提高相似度计算精度,解决数据稀疏性问题,从而提高推荐的精准度和个性化。且随着最近邻数量的增加,基于模糊相似关系的协同过滤推荐算法的精准度略高于基于因果聚类的协同过滤推荐算法。

关键词:推荐系统;协同过滤;因果聚类;模糊相似关系;最大树;阈值;加权

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2018)10-0044-05

doi:10.3969/j.issn.1673-629X.2018.10.009

Improvement and Research of Collaborative Filtering Recommendation Algorithm

SUN Hua-yan, LI Ye-li, ZI Yun-fei, HAN Xu

(School of Information Engineering, Beijing Institute of Graphic Communication, Beijing 102600, China)

Abstract: Aiming at the problem that data sparsity seriously affects the accuracy of recommendation in collaborative filtering algorithms, we propose a collaborative filtering recommendation algorithm based on causal clustering and fuzzy similarity relation. Collaborative filtering recommendation algorithm based on causal clustering builds matrices to compute their similarity through the user-project scoring, collection, eigenvalues and so on. The establishment of matrix is based on the causality and similarity between the project and the user's eigenvalue. To achieve accurate, personalized recommendations, the collaborative filtering recommendation algorithm based on fuzzy similarity relation computes the similarity through constructing triangular fuzzy number by the weighted feature method of project and user, and gets connected graph truncation set according to the fuzzy similarity matrix and the best threshold method. Experiment shows that both algorithms can effectively improve the accuracy of similarity calculation, solving the problem of sparse data and raising the precision and personalization of the recommendation. As the number of nearest neighbors increases, the collaborative filtering recommendation algorithm based on fuzzy similarity relation is slightly more accurate than the collaborative filtering recommendation algorithm based on causal clustering.

Key words: recommendation systems; collaborative filtering; cause-effect clustering; fuzzy similarity relation; biggest tree; threshold; weighting

0 引言

伴随着互联网的迅速发展,智能化、个性化的服务成为电商产业极大的需求,如何通过存储有海量数据的自媒体为用户提供个性化的推荐成为了智能网络研

究的热点。研究人员提出了以协同过滤推荐算法为主流的多种推荐算法,相比其他算法,协同过滤推荐算法的易实现性和可扩展性使其在电商和自媒体等领域应用广泛。但随着互联网数据呈几何级数增长、数据内

收稿日期:2017-11-20

修回日期:2018-03-06

网络出版时间:2018-05-25

基金项目:国家自然科学基金(11603004);北京市科技创新服务能力协调创新项目(PXM2016_014223_000025);北京市科技创新服务能力建设科研水平提高定额项目(2017-04190117010);北京市自然科学基金项目(1173010)

作者简介:孙华艳(1990-),女,硕士研究生,研究方向为数据挖掘、推荐算法及可视化分析;李业丽,教授,研究方向为数据挖掘、推荐算法、算法研究及可视化分析。

网络出版地址:cnki.net/kcms/detail/61.1450.TP.20180525.1550.018.html

容复杂度不断提高,协同过滤技术也面临着两个瓶颈问题^[1-5]:数据的稀疏性和可扩展性。

数据稀疏性问题是推荐系统面临的主要问题,也是导致系统推荐质量下降的主要原因^[6]。目前,对该问题的解决分为两类,一类为矩阵填充技术,另一类为矩阵降维技术^[6-8]。矩阵填充技术比较理想的方法有BP神经网络、Naive Bayesian分类方法、基于内容的预测^[9-10]。矩阵降维技术比较典型的方法是奇异值分解(singular value decomposition, SVD)技术^[11-12]。针对个性化推荐过程中最近邻准确性的不足,文中提出了基于因果聚类分析的协同过滤推荐算法和基于模糊相似关系的协同过滤推荐算法。前者利用因果聚类对收集的信息进行分析研究,找出其内在联系,实现个性化的精准推荐。后者通过最大树法及确定最佳阈值 λ 以更加合理、精准地计算最近邻^[13-14]。最后将算法应用于美食推荐系统中进行验证^[15]。

1 传统的协同过滤推荐算法

协同过滤(collaborative filtering, CF)推荐算法^[16-18]是通过用户对项目之间的评价、分享、转存等数据的分析,找出基于用户或项目的相似度,实现精准化推荐。协同过滤推荐算法分为:(1)基于用户的协同过滤推荐算法(user-based collaborative filtering),通过用户搜索,对喜欢产品的评价、收藏等数据信息度量有相同行为的用户之间的关系来实现产品的推荐。(2)基于项目的协同过滤推荐算法(item-based collaborative filtering),通过不同用户对于不同产品的评价及喜好程度来度量产品之间的紧密程度,从而实现产品的推荐。

1.1 协同过滤推荐算法的实现

- (1)收集用户—项目数据;
- (2)计算用户—项目的相似度,找出最近邻或相似度最高的 N 个用户或项目;
- (3)把 N 个相似度最高的用户或项目推荐给对应的项目或用户。

1.2 相关性评价计算

(1)皮尔逊相关系数(Pearson correlation coefficient):通过用户对产品的评价数据计算两组数据与某一直线的拟合程度,若两用户相似度较高就会形成对角线。设用户为 x 和 y , \bar{m}_x 为用户 x 对所有产品的评分数据, m_{xk} 为用户 x 对第 k 项目的评分数据,相似度计算如下:

$$\text{sim}(x, y) = \frac{\sum_k (m_{xk} - \bar{m}_x)(m_{yk} - \bar{m}_y)}{\sqrt{\sum_k (m_{xk} - \bar{m}_x)^2 \sum_k (m_{yk} - \bar{m}_y)^2}} \quad (1)$$

万方数据

计算相似度之后寻找最近邻,令 Z_{um} 为最近邻,对项目的 s 预测评分的计算为:

$$m_{us} = \bar{m}_u + \frac{\hat{A}_{u, Z_{um}} \text{sim}(u, x) (m_{xs} - \bar{m}_x)}{\hat{A}_{u, Z_{um}} |\text{sim}(u, x)|} \quad (2)$$

其中, \bar{m}_x 为 Z_{um} 中 x 的所有评分数据的平均分; \bar{m}_u 为目标用户 u 基于所有项目评分数据的均值。

$$\bar{m}_u = \frac{\hat{A}_{u, T_s} m_{uw}}{|T_s|} \quad (3)$$

其中, T_s 为项目总和; m_{uw} 为 u 对产品 w 的评价。

(2)余弦相似度(cosine similarity):通过两个或以上用户对不低于两个产品的评分来计算两用户之间的兴趣相似度。设用户 x 和 y 对产品的评价数组数据为 x 和 y ,相似度计算如下:

$$\text{cossim}(x, y) = \frac{\langle x, y \rangle}{\|x\| \|y\|} \quad (4)$$

2 基于因果聚类的协同过滤推荐算法

事物本身的因果关系就存在着相似性,也就是可以通过对以往的经验 and 数据寻找事物存在的规律。例如:用户 a 因为喜欢《python 数据挖掘》,所以他也可能喜欢《数据挖掘算法》,它们之间存在着这种因果关系。同理对于 a 和 b 两个用户,因为他们之间有共同喜爱的一门课程,那么他们之间也有可能存在共同喜爱的另外一门课程。本节基于因果聚类的协同过滤推荐算法,将以更加精细和准确的关联计算来实现项目的精准可靠推荐。引入项目和用户的因果关系计算相似性更进一步提高度量相似度。

2.1 因果模糊聚类

因果模糊聚类^[19-20]的核心是择近原则。设对要分析处理的项目集 α ,每个项目存在的特征值为 (x_1, x_2, \dots, x_n) ,然后对特征值的因果模糊聚类和模式识别判定、预测它们的归属,从而实现更好的推荐。

2.2 因果模糊聚类实现过程

(1)特征模糊集的建立。

设所有项目集为 α ,对于 α 的聚类通过三元结构 (X, Y, φ) 处理,所以 $X = (X_1 \times X_2 \times \dots \times X_n)$ 是 n 元Descartes累乘, Y 为非空幂集, φ 为 X 到 Y 的映射,如: $\varphi: X_1 \times X_2 \times \dots \times X_n \rightarrow Y$ 。

同时记 $z_i = (x_i, y_i)$,通过模糊等价计算出对应的最佳聚类 $(z_1, z_2, \dots, z_m) = (T_1, T_2, \dots, T_m)$,特征模糊集 $V_i = \{x_{i_1}, x_{i_2}, \dots, x_{i_k}\} (i = 1, 2, \dots, m)$,其中 (x_{i_s}, y_{i_s}) 为项目 T_i 的最佳聚类, V_i 是 T_i 的特征轴投影。计算:

$$\bar{x}_i = \frac{1}{k_i} \sum_{s=1}^{k_i} x_{i_s} = (\bar{x}_i^{(1)}, \bar{x}_i^{(2)}, \dots, \bar{x}_i^{(n)}) \quad (5)$$

$$\sigma_{ij}^2 = \frac{1}{k_i} \sum_{s=1}^{k_i} (x_{i_s}^{(j)} - \bar{x}_{i_s}^{(j)})^2 \quad (6)$$

对于 $x = (x_1, x_2, \dots, x_n)$ 属于项目的特征值, 令

$$P_i = \sum_{j=1}^n \omega_{ij} e^{-\frac{(x_j - \bar{x}_i^{(j)})^2}{9\sigma_{ij}^2}} \quad (7)$$

其中, $i, j = 1, 2, \dots, m$ 为特征取值的范围; $(\omega_1, \omega_2, \dots, \omega_n)$ 为一组确定的权重。

令 $W_i = \{y_{i_1}, y_{i_2}, \dots, y_{i_{k_i}}\}$, 其中 $i = 1, 2, \dots, m$, $(x_{i_s}, y_{i_s}) = z_{i_s}$ 属于项目最佳聚类 T_i , W_i 是 T_i 的预测相似轴投影, 计算:

$$\bar{y}_i = \frac{1}{k_i} \sum_{s=1}^{k_i} y_{i_s} \quad (8)$$

同时求: $\delta_i = \max_{1 \leq s \leq k_i} |y_{i_s} - \bar{y}_i|$, 最后构造 $(\bar{y}_i, 3\delta_i)$ 为参数的三角模糊数。通过三角模糊特征值

$$I(x) = \begin{cases} \frac{1}{\sigma}x + \frac{\sigma - \alpha}{\sigma}, & \sigma - \alpha \leq x \leq \alpha \\ -\frac{1}{\sigma}x + \frac{\sigma - \alpha}{\sigma}, & \alpha \leq x \leq \alpha + \sigma \end{cases}, I_i (i = 1, 2, \dots, m)$$

建立因果模糊聚类 $A \in X$, 对应项目的分类 (T_1, T_2, \dots, T_m) 为:

$$\begin{bmatrix} P_1 & P_2 & \dots & P_m \\ I_1 & I_2 & \dots & I_m \end{bmatrix}$$

(2) 进行因果聚类的判定。

若 s 期特征的状态为确定特征值 $x_s \in X$, 则 x_s 与 $\{P_1, P_2, \dots, P_m\}$ 为最大隶属度原则, $A_j = \max(A_1(u_0), A_2(u_0), \dots, A_n(u_0))$ 。

若 s 期特征的状态为模糊集 $B_s \in X$, 则 B_s 与 $\{P_1, P_2, \dots, P_m\}$ 通过择近原则。

最后, 选取 P_{i_0} , 同时对应的 I_{i_0} 为项目集 α 与 s 期的因果聚类的模糊判定值 ($i_0 \in \{1, 2, \dots, m\}$)。

该方法通过计算某项目影响推荐效果的 n 个因素, 计算并构造相应的三角模糊数, 最后根据最大隶属度原则或择近原则进行模糊预测, 从而找出影响最终推荐的可能性最大的因素及特征值, 再根据这一影响因素实现精准、个性化的推荐。

步骤 1: 对于项目的原始特征及数据, 进行标量化处理, 标定值的大小表示两个特征之间的相似度;

步骤 2: 通过计算将项目原始特征数据转化成模糊矩阵 R ;

步骤 3: 将模糊矩阵 R 进行模糊聚类分析, 并且计算出最佳聚类;

步骤 4: 根据各聚类的特征函数及权值, 构建三角模糊数, 从而选取最佳特征值, 实现精准、个性化推荐。

2.3 最近邻的取值

算法最后通过对项目的评价、分享、收藏等原始数据值最高的 n 个特征进行计算及分析处理, 最佳特征为一

个取值空间的中间值, 所以以中间值为中心的不超过两端取值范围为最佳特征数。

3 基于模糊相似关系的协同过滤推荐算法

用户在网上有目标的搜索是一种显性需求, 然而显性需求的信息数据背后隐藏着隐性需求, 这就是个性化推荐的核心。起初, 网站极少的用户共同评价数据, 很难准确地获得最近邻。例如, 两个非常喜欢计算机科学的用户由于不存在共同评价数据, 其相似度计算结果为零。而与他们不是同一类的书籍爱好者由于存在评价数据, 相似度反而高于零。这就是网站对于新用户评价项目的第一评价以及数据稀疏性问题, 仅以共同评价数据判断相似级别显然不精准。文中通过结合项目所属分类来计算用户的相似程度, 这是由于用户对于统一属性的项目会有相似的喜好程度。

用户喜欢的项目类可能拥有多个特征, 而它们之间又可能存在着特征重叠或相通性, 所以可以通过模糊相似关系的直接聚类法求其项目之间的紧密程度, 也就是相似度。基于模糊相似关系的协同过滤推荐算法就是通过求其最大树法寻找最近邻, 从而实现个性化精准推荐。

基于模糊相似关系的协同过滤推荐算法^[21]是通过项目的标定得到模糊矩阵 R , 然后根据相似矩阵 R 得到一棵权重树, 通过最佳阈值 λ 确定法对相似矩阵进行截集, 应用直接聚类求出相似性, 获得最佳聚类, 最后对用户进行精准的个性化产品推荐。

3.1 相似关系的协同过滤推荐算法的实现过程

(1) 项目特征值的标定。

从项目 u_i 和 u_j 中选取最具有特征性的 m 个特征值, 分别为 $u_i = (x_{i1}, x_{i2}, \dots, x_{im})$ 、 $u_j = (x_{j1}, x_{j2}, \dots, x_{jm})$, 然后计算得到项目模糊相似矩阵。

$$r_{ij} = \frac{\sum_{k=1}^m (x_{ik} - \bar{x}_i)(x_{ik} - \bar{x}_j)}{\sqrt{\sum_{k=1}^m (x_{ik} - \bar{x}_i)^2} \sqrt{\sum_{k=1}^m (x_{ik} - \bar{x}_j)^2}} \quad (9)$$

其中, $\bar{x}_i = \frac{1}{m} \sum_{k=1}^m x_{ik}$, $\bar{x}_j = \frac{1}{m} \sum_{k=1}^m x_{jk}$ 。

通过式 5, 从项目集合中分别计算两个项目之间的紧密性, 得到项目模糊相似矩阵 \bar{R} 。

(2) 构造最大树。

首先, 根据第 1 步对项目特征值的标定得到模糊相似矩阵 R , 画出被分类因素为节点, 以相似矩阵 R 的因素 r_{ij} 为权重的一棵最大树; 其次, 选取最佳阈值 λ ; 最后, 砍断权重低于 λ 的枝, 得到一幅连通图, 相联的分支就构成了 λ 水平上的分类, 这就是获得相似度较高的最佳聚类。

(3) 计算最佳阈值 λ 。

要实现精准、个性化推荐, 阈值 λ 的选取是核心。通过模糊相似矩阵依次取 λ 的最大树, 再利用最大树法得到的截集来实现精准、个性化推荐。

聚类图可以给出各 λ 对应的分类, 形成一种动态聚类。对于阈值 λ 的选取, 由于人为的分类总是带有主观性, 所以可用 F -统计量来选取。

设 $U = \{u_1, u_2, \dots, u_n\}$ 为需要分类的所有项目集, $u_j = (x_{j1}, x_{j2}, \dots, x_{jm})$, x_{jk} 是项目 u_j 的第 k ($k=1, 2, \dots, m$) 个特征数据, c 是对应于 λ 的类数, n_i 为第 i 类项目的个数, 第 i 类项目记作: $u_1^i, u_2^i, \dots, u_{n_i}^i$, 记 $\bar{x}_k^i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{jk}$ ($k=1, 2, \dots, m$) 为第 i 类项目的第 k 个特征值的平均数, 而 $\bar{u}^i = (\bar{x}_1^i, \bar{x}_2^i, \dots, \bar{x}_m^i)$ 为第 i 类项目的聚类中心向量, $\bar{u} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k)$ 是所有项目集合的中心向量, 且 $\bar{x}_k = \frac{1}{n} \sum_{j=1}^n x_{jk}$ ($k=1, 2, \dots, m$)。所以, F -统计量为:

$$F = \frac{\sum_{i=1}^c n_i \|\bar{u}^i - \bar{u}\|}{(c-1)} \quad (10)$$

$$\sum_{i=1}^c \sum_{j=1}^{n_i} \frac{\|u_j^i - \bar{u}^i\|^2}{(n-c)}$$

其中, $\|\bar{u}^i - \bar{u}\| = \sqrt{\sum_{k=1}^m (\bar{x}_k^i - \bar{x}_k)^2}$ 是计算 \bar{u}^i 与 \bar{u} 的距离关系, $\|u_j^i - \bar{u}^i\|$ 是计算第 i 类项目中项目 u_j^i 与中心 \bar{u}^i 的距离关系, 所以 $\sum_{i=1}^c \frac{n_i \|\bar{u}^i - \bar{u}\|}{(c-1)}$ 表示项目与项目之间的相似度, $\sum_{i=1}^c \sum_{j=1}^{n_i} \frac{\|u_j^i - \bar{u}^i\|^2}{(n-c)}$ 表示项目中各特征之间的紧密程度。说明 F 值越大, 聚类、截集越精准。

通过最佳阈值 λ 确定法计算模糊相似矩阵的方法, 可以更精准地把相似度高的特征属性项目聚为同一种推荐项目, 然后运用这一方法对用户进行精准、个性化的产品推荐。

3.2 最近邻的取值

基于对项目的原始特征值的标定, 得到模糊相似矩阵, 然后通过最大树法获得最佳聚类, 最后通过最佳阈值确定法计算出截矩阵 \bar{R}_λ , 所以不小于 \bar{R}_λ 的截集为最近邻。

4 实验结果及分析

4.1 数据集

实验数据集是在合法范围内, 基于 Python 语言的数

据挖掘技术在美团网上抓取的。美团网拥有中国最多的美食信息资源, 也是国内最活跃、最权威的外卖网之一。抓取数据中包含对美食的评分(值为 1~5, 值越大表明用户喜好度越高)、特征和网络社交关系等。在美团网上用户可以对美食进行等级评分、收藏以及给朋友分享推荐, 所以美团网的数据对于美食推荐具有很高的参考价值。其中抓取了 10 万条评分数据, 含有 947 个用户对 1 678 个美食的评判分数, 每个用户不少于对 27 个美食进行评分。

4.2 数据密度计算

利用用户对美食的评分值来表达对美食的喜爱程度。

$$\text{数据密度} = 100\,000 / (947 \times 1\,678) = 0.062\,9$$

(11)

表明数据的稀疏性非常高。

4.3 测试原则

通过对抓取数据随机抽取 70 000 条作为实验训练集, 抽取 30 000 条作为测试集, 分别对传统过滤推荐算法和基于因果聚类和模糊相似关系的协同过滤推荐算法进行实验。

4.4 度量标准

文中应用的统计精度测量中的绝对误差 (MAE) 是度量推荐结果好坏的最有效方法之一。MAE 值越小表明推荐效果越佳。

$$\text{MAE} = \frac{\sum_{i=1}^m |p_i - q_i|}{m} \quad (12)$$

其中, m 为项目集中评分项的量; p_i 为对算法预估的评判分数; q_i 为测试数据的实际分数。

4.5 实验结果

实验基于抓取的数据对传统的协同过滤推荐算法与因果聚类及模糊相似关系的协同过滤推荐算法进行了平均绝对误差值的比较分析, 如图 1 所示。最近邻的数量也将影响推荐的效果, 过少将会降低推荐的精准性, 过多将极大增加算法的复杂度, 实验中选取最近邻的数量为 10~40。

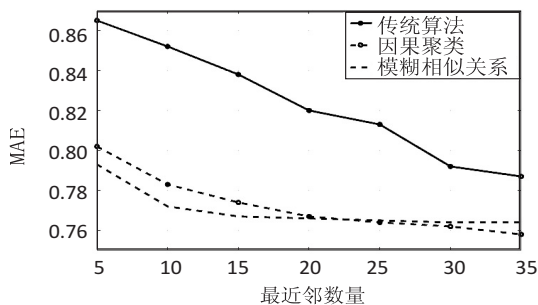


图 1 3 种算法的 MAE 比较

通过图 1 可知, 无论选取多大最近邻数, 提出的基

于因果聚类及模糊相似关系的协同过滤推荐算法的平均 MAE 都低于传统推荐算法,表明该算法具有较高的精度,对推荐系统精度的提高有明显帮助。其中基于因果聚类的协同过滤推荐算法在最近邻数量低于 23 左右时,精度略高于基于模糊相似关系的协同过滤推荐算法。这是因为因果聚类是基于模糊相似关系法中的模糊矩阵加各项目的权重来进行聚类分析推荐,所以它的平均绝对误差略小。当最近邻高于 23 以后,因果聚类法的最终预测值在最大隶属度或择近范围中会扩大,所以精度趋于不变,而基于模糊相似关系的协同过滤推荐算法的精度则继续提高。

5 结束语

各推荐算法中相似度量法的好坏直接影响了推荐项目或用户的精度和个性化。文中提出的基于因果聚类与模糊相似关系的协同过滤推荐算法主要通过增加项目权重的三角模糊数法和计算最大树及最佳阈值 λ 确定法来提高相似度的度量。同时,该算法通过对项目或用户特征值的内外距离计算或加权计算特征模糊集,都能减小稀疏性带来的负面影响。下一步将结合聚类技术解决协同过滤推荐技术的可扩展性问题。

参考文献:

- [1] HUANG Hao, HUANG Jianqing, ZIAVRAS S G, et al. A personalized recommendation algorithm based on Hadoop [C]//5th international conference on electronics information and emergency communication. Beijing, China: IEEE, 2015.
- [2] ADOMAVICUS 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.
- [3] 韦素云, 业宁, 吉根林, 等. 基于项目类别和兴趣度的协同过滤推荐算法[J]. 南京大学学报: 自然科学版, 2013, 49(2): 142-149.
- [4] 李聪, 梁昌勇, 马丽. 基于领域最近邻的协同过滤推荐算法[J]. 计算机研究与发展, 2008, 45(9): 1532-1538.
- [5] 黄创光, 印鉴, 汪静, 等. 不确定近邻的协同过滤推荐算法[J]. 计算机学报, 2010, 33(8): 1369-1377.
- [6] 潘骏驰, 张兴明, 汪欣. 融合用户可信度的改进奇异值分解推荐算法[J]. 小型微型计算机系统, 2016, 37(10): 2171-2176.
- [7] PARK C, KIM D, OH J, et al. Improving top-K recommendation with truster and trustee relationship in user trust network [J]. Information Sciences, 2016, 374(C): 100-114.
- [8] LINDEN G, SMITH B, YORK J. Amazon.com recommendations: item-to-item collaborative filtering [J]. IEEE Internet Computing, 2003, 7(1): 76-80.
- [9] 曾艳, 麦永浩. 基于内容预测和项目评分的协同过滤推荐[J]. 计算机应用, 2004, 24(1): 111-113.
- [10] 王伟, 王洪伟, 孟园. 协同过滤推荐算法研究: 考虑在线评论情感倾向 [J]. 系统工程理论与实践, 2014, 34(12): 3238-3249.
- [11] 张朝恒, 何小卫, 陈勇兵. 基于社交网络信息的协同过滤推荐算法[J]. 计算机技术与发展, 2017, 27(12): 28-34.
- [12] 苏凡军, 唐启桂. SBHCF: 基于奇异值分解的混合协同过滤推荐算法[J]. 电子科技, 2016, 29(1): 44-47.
- [13] 姜书浩, 薛福亮. 一种利用协同过滤预测和模糊相似性改进的基于内容的推荐方法 [J]. 现代图书情报技术, 2014(2): 41-47.
- [14] 曼舒总理可汗. 模糊相似性和空间数据建模及其在多维尺度矢量地图空间中的应用研究 [D]. 北京: 北京理工大学, 2015.
- [15] PERA M S, NG Y K. Analyzing book-related features to recommend books for emergent readers [C]//Proceedings of the 26th ACM conference on hypertext & social media. [s. l.]: ACM, 2015: 221-230.
- [16] 杨芳. 电子商务系统协同过滤推荐算法研究 [D]. 天津: 河北工业大学, 2006.
- [17] 荣辉桂, 火生旭, 胡春华, 等. 基于用户相似度的协同过滤推荐算法[J]. 通信学报, 2014, 35(2): 16-24.
- [18] ZHANG Qinghua, XU Kai, WANG Guoyin. Fuzzy equivalence relation and its multigranulation spaces [J]. Information Sciences, 2016, 346-347: 44-57.
- [19] 麻兴斌, 唐林炜. 基于因果聚类和模式识别的模糊预测方法及其应用 [J]. 山东科技大学学报: 自然科学版, 2004, 23(1): 76-79.
- [20] 李波, 陈广平, 于正兴. 基于因果模糊聚类法预测覆岩裂隙带高度研究 [J]. 中国安全生产科学技术, 2013, 9(11): 25-29.
- [21] 姜永森, 陆媛, 杨慧中. 一种模糊相似关系的基因表达数据聚类方法 [J]. 计算机工程与应用, 2011, 47(8): 236-238.