Journal of Computer Applications
计算机应用,2018,38(10):2886 -2891
ISSN 1001-9081
CODEN JYIIDU
2018-10-10
http://www.joca.cn

文章编号:1001-9081(2018)10-2886-06
DOI:10.11772/j.issn.1001-9081.2018040766

集成用户信任度和品牌认可度的商品推荐方法

冯　勇,韩晓龙,付陈平,王嵘冰*,徐红艳
(辽宁大学 信息学院,沈阳 110036)
(*通信作者电子邮箱 wrb@lnu.edu.cn)

摘　要:针对个性化商品推荐方法中普遍存在的推荐准确率不高的问题,提出一种集成用户信任度和品牌认可度的商品推荐方法(TBCRMI)。该方法通过分析用户的购买行为和评价行为,计算得到用户对商品品牌的认可度和用户自身的活跃度;然后利用 DBSCAN(Density-Based Spatial Clustering of Applications with Noise)算法对用户进行聚类,并在此基础上融合用户信任关系,采用 Top-*K* 方法得到近邻关系;最后,依据近邻关系生成目标用户商品推荐列表。为了验证算法的有效性,使用 Amazon Food 和 Unlocked Mobile phone 两个数据集,选择基于用户的协同过滤算法(UserCF)、融合用户信任的协同过滤推荐算法(SPTUserCF)与合并用户信任的协同过滤算法(MTUserCF),对准确率、召回率和 *F*1 值等指标进行了对比分析。实验结果表明,无论是多品牌综合推荐还是单一品牌推荐,TBCRMI 在各项指标均优于目前常用的个性化商品推荐方法。

关键词:个性化推荐;品牌认可;聚类;用户信任;近邻
中图分类号:TP302.1　　**文献标志码**:A

Commodity recommendation method integrating user trust and brand recognition

FENG Yong, HAN Xiaolong, FU Chenping, WANG Rongbing*, XU Hongyan
(*School of Information, Liaoning University, Shenyang Liaoning* 110036, *China*)

Abstract: Concerning the low recommendation accuracy of personalized commodity recommendation methods, a Commodity Recommendation Method Integrating User Trust and Brand Recognition (TBCRMI) was proposed. By analyzing the user's purchase behavior and evaluation behavior, the user's recognition of brands and the activities of users themselves were calculated. Then Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm was used to cluster the users, based on which the user trust relationship was fused, and the nearest neighbors were obtained by Top-*K* method. Finally, the target user commodity recommendation list was generated based on the nearest neighbors. In order to verify the effectiveness of the algorithm, two datasets (Amazon Food and Unlocked Mobile Phone) were used, User based Collaborative Filtering (UserCF) algorithm, Collaborative Filtering recommendation algorithm with User trust (SPTUserCF) and Merging Trust in Collaborative Filtering (MTUserCF) algorithm were chosen, and the accuracy, recall and *F*1 value were compared and analyzed. The experimental results show that TBCRMI is superior to the commonly used personalized commodity recommendation methods in either multi-brand comprehensive recommendation or single brand recommendation.

Key words: personalized recommendation; brand recognition; clustering; user trust; nearest neighbor

www.joca.cn

0　引言

随着互联网技术的高速发展与广泛应用,电子商务已成为人们日常生活中不可或缺的商品交易模式。在影响商品销售的诸多因素中,信任关系和品牌概念起到至关重要的作用。信任是构建用户间伙伴关系的基石,口碑营销、病毒营销等新型营销模式就是基于信任关系的销售。品牌是商品的标志,企业通过关系营销、数据库营销、整合营销等营销手段提升用户对品牌的认可度,进而增强用户黏度,促进商品销售。这两个关键因素的融合可以使用户在种类繁多的电子商务环境中迅速找到合适的商品,提高用户的购买效率。为此,本文从用户信任和品牌认可两个方面着手,对传统个性化推荐方法进行改进,实现电子商务精准营销。

近年来,众多学者通过研究用户信任关系与个性化商品推荐的关联,或是通过分析用户评分信息来提高推荐的准确率,但却忽视了商品自身最主要的品牌因素对推荐准确率的影响。李良等[1]将信任融合到协同过滤推荐方法中,提出将用户评分信任和偏好信任结合,但该方法在计算信任时忽略了用户间的交互关系。王海艳等[2]通过建立可信联盟的方法来提高推荐服务的精确度,考虑了信任计算中的直接信任和间接信任;但并没有充分考虑信任的非对称性,结果偏离现实信任关系。在用户评分上,Liu 等[3]将用户评分相似性和

收稿日期:2018-04-16;**修回日期**:2018-07-04;**录用日期**:2018-07-06。
基金项目:国家自然科学基金资助项目(71771110);辽宁省博士科研启动基金资助项目(20160199)。
作者简介:冯勇(1973—),男,辽宁沈阳人,教授,博士,CCF 会员,主要研究方向:数据挖掘、个性化推荐;　韩晓龙(1991—),男,山东胶州人,硕士研究生,主要研究方向:数据挖掘、个性化推荐;　付陈平(1992—),女,山东潍坊人,硕士研究生,主要研究方向:深度学习、数据挖掘;王嵘冰(1979—),男,辽宁沈阳人,副教授,博士,CCF 会员,主要研究方向:物联网、大数据、云计算;　徐红艳(1972—),女,辽宁丹东人,副教授,硕士,主要研究方向:Deep Web、个性化推荐、数据挖掘。

用户评论进行综合考虑,这种思想在一定程度上提高了推荐的准确率,却增加了算法时间复杂度。McAuley 等[4] 利用 hidden factors as topics 将评分和评论信息相结合构建特征向量,但这并不能同时对商品和用户两个角度进行充分考虑。FPCA(Fixed Point Continuation with Approximate SVD)算法[5] 和 LmaFit 算法[6] 等通过补全用户评分矩阵来提高推荐算法的准确度;但这些算法大多聚焦于用户的评分信息,而忽视了用户的评论、购买足迹等因素,这些因素能够更细致地描述用户的轮廓,对用户的消费决策产生重要影响。

本文在改进推荐算法中用户信任度计算方式的同时,引入了品牌认可度因素,将品牌认可度与用户信任度融合,提出一种集成用户信任度和品牌认可度的商品推荐方法(Commodity Recommendation Method Integrating User Trust and Brand recognition, TBCRMI)。该方法由品牌认可度、用户活跃度和用户信任度三个核心计算环节构成,其中:品牌认可度充分考虑了用户购买商品时的品牌选择行为;用户活跃度分析了用户的评价行为,突出活跃用户;用户信任度结合了现实世界中信任的双向关系,定义了新的信任度计算方式。

1　TBCRMI 方法

1.1　方法架构

为提升个性化商品推荐的成功率,促进电子商务的开展,本文结合商品营销中的核心要素:信任关系和品牌理念,给出一种集成用户信任度和品牌认可度的个性化商品推荐方法,其架构如图 1 所示。架构图的核心环节包括:品牌认可度计算、用户活跃度计算和用户信任度计算。

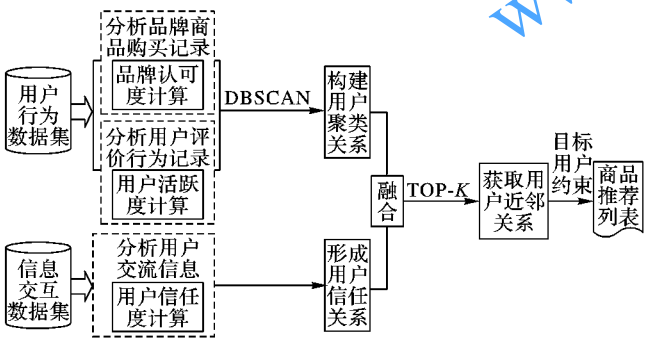


图 1　TBCRMI 方法架构

Fig. 1　Architecture of TBCRMI

1.2　核心算法

设用户集合 $U = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n\}$,品牌集合 $B = \{b_1, b_2, \cdots, b_m\}$,根据用户行为数据集和信息交互数据集,先分析用户的商品购买记录和评价行为记录,计算得到用户品牌认可度和活跃度,然后按品牌构建用户向量。设存在 $b_j \in B, \boldsymbol{u}_i \in U$,用户向量表示为:

$$\boldsymbol{u}_i(Bco_{i,j}, Uac_{i,j})$$

其中:$Bco_{i,j}$ 表示用户 $\boldsymbol{u}_i$ 对品牌 b_j 的认可度;$Uac_{i,j}$ 表示用户 $\boldsymbol{u}_i$ 在品牌 b_j 上的活跃度。

在形成用户信任关系环节,采用 DBSCAN(Density-Based Spatial Clustering of Applications with Noise)算法对用户进行聚类,用户间的距离计算使用欧氏距离。聚类算法如下。

输入　U 为包含 n 个用户的数据集;ε 为半径参数;$MinPts$ 为邻域密度阈值。
输出　基于密度的用户的集合 C。
标记 U 中所有的用户为 unvisited;
do
　{随机选择一个 unvisited 对象 $\boldsymbol{u}_i$;
　标记 $\boldsymbol{u}_i$ 为 visited;
　if $\boldsymbol{u}_i$ 的 ε- 邻域至少有 $MinPts$ 个对象
　　{创建新簇 C,并把 $\boldsymbol{u}_i$ 添到 C;
　　令 N 为 $\boldsymbol{u}_i$ ε- 邻域中对象集合;
　　for N 中每个点 $\boldsymbol{u}'_i$
　　{ if $\boldsymbol{u}'_i$ 是 unvisited
　　标记 $\boldsymbol{u}'_i$ 为 visited
　　if $\boldsymbol{u}'_i$ ε- 邻域至少有 $MinPts$ 个点
　　把它添加到 N;
　　if $\boldsymbol{u}'_i$ 还不是任何簇的成员
　　把 $\boldsymbol{u}'_i$ 添加到 C;}
　　输出 C;}
　Else 标记 $\boldsymbol{u}_i$ 为噪声}
while U 中没有标记为 unvisited 的对象

根据聚类结果,得到关于品牌 b_i 的用户关系矩阵 $\boldsymbol{UR}_{b_i}$:

$$\boldsymbol{UR}_{b_i} = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{bmatrix}$$

其中:数据元素 d_{ij} 是用户 i 和用户 j 间的距离。

在用户信任度计算环节,根据用户间信息交流数据,计算得到用户信任关系矩阵 $\boldsymbol{T}$ 如下:

$$\boldsymbol{T} = \begin{bmatrix} Tr_{11} & Tr_{12} & \cdots & Tr_{1n} \\ Tr_{21} & Tr_{22} & \cdots & Tr_{2n} \\ \vdots & \vdots & & \vdots \\ Tr_{n1} & Tr_{n2} & \cdots & Tr_{nn} \end{bmatrix}$$

其中:数据元素 Tr_{ij} 为用户 $\boldsymbol{u}_i$ 对用户 $\boldsymbol{u}_j$ 的信任度。

在面向目标用户选取近邻关系环节,将用户聚类矩阵与用户信任矩阵融合。设用户 $\boldsymbol{u}_i$ 为目标用户,在矩阵 $\boldsymbol{UR}_{b_i}$ 中第 i 行表示用户 $\boldsymbol{u}_i$ 相对其他用户的距离;同样在矩阵 $\boldsymbol{T}$ 中第 i 行表示用户 $\boldsymbol{u}_i$ 对其他用户的信任度。合并得到推荐用户列表 $\boldsymbol{UR}_n$ 表示如下:

$$\boldsymbol{UR}_n = \alpha \begin{bmatrix} d_{i1} \\ d_{i2} \\ \vdots \\ d_{in} \end{bmatrix} + \beta \begin{bmatrix} Tr_{i1} \\ Tr_{i2} \\ \vdots \\ Tr_{in} \end{bmatrix}$$

最后采用 Top-K 方法选取目标用户的 K 个最近邻,再依据目标用户约束生成商品推荐列表。为了确定 α 和 β 的合理取值,本文针对不同的 α、β 取值进行多次实验并计算相应的准确率 P 和召回率 R。α、β 的测试用例与计算结果如表 1 所示。从表 1 中可以看出当 $\alpha = 0.5$, $\beta = 0.5$ 时推荐效果最佳。

1.3　算法时间复杂度分析

本文算法的时间复杂度主要取决于以下两个过程:1)分析用户行为数据与用户交互数据过程中遍历所有记录的时间;2)对用户进行聚类时使用的 DBSCAN 算法的时间复杂

www.joca.cn

度。由于分析用户品牌商品的购买记录、用户评价行为记录和用户交流信息的过程是并行执行的，所以遍历 n 个用户的 n 条记录所产生的时间复杂度为 $O(n^2)$；而采用 DBSCAN 算法对用户聚类时的基本时间复杂度是 $O(n \times$ 找出 ε- 邻域中的用户所需要的时间)，n 是用户的个数，在最坏情况下其时间复杂度是 $O(n^2)$。综合以上分析，本文的时间复杂度为 $O(n^2)$。

表 1 测试用例及结果

Tab. 1 Test cases and results

α	β	准确率 P	召回率 R
0.1	0.9	0.29	0.24
0.3	0.7	0.28	0.26
0.5	0.5	0.38	0.29
0.6	0.4	0.33	0.26
0.8	0.2	0.28	0.19

2 TBCRMI 方法核心环节

本文所提出的集成用户信任度和品牌认可度的个性化商品推荐方法包含三个核心环节：品牌认可度计算、用户活跃度计算和用户信任度计算。

2.1 品牌认可度

Halim[11] 的研究发现，一个能给消费者带来深刻印象的品牌总能赢得消费者特别的注意力，超值的品牌更可能赢得消费者的认可。在现实生活中，如果用户在体验过某个品牌的商品后对其并不满意，那么再向该用户推荐这个品牌的其他商品时，用户接受的可能性很小。因此，在个性化商品推荐中融合品牌认可度因素，可以提升推荐的准确性。

定义 1 品牌认可度。是指相对一种品牌，用户对其熟悉度和忠诚度的表达。假如在同类商品中用户对某品牌的商品购买的次数明显高于该品牌下所有用户的平均购买次数，且高于用户自身对同类商品的平均购买次数，那就说明该用户对此品牌商品比较偏爱，有较高的认可度。

用户的品牌认可度由式(1)计算得到：

$$Bco_{i,b} = \frac{P_{i,b}}{p_i} \times \log_{P_{b_max}}\left(\frac{P_{i,b}}{\overline{P}_b} + 1\right) \tag{1}$$

其中：$Bco_{i,b}$ 表示用户 $\boldsymbol{U}_i$ 对品牌 b 的认可度；P_i 表示用户 $\boldsymbol{U}_i$ 对某一类商品的购买次数；$P_{i,b}$ 表示用户 $\boldsymbol{U}_i$ 在该类商品中对 b 品牌商品的购买次数；$\overline{P}_b$ 表示所有用户对该品牌商品的平均购买次数；P_{b_max} 表示用户对品牌 b 的最大购买次数。

2.2 用户活跃度

高活跃度用户对品牌商品的评论存在正、负两面，正面评论可以增加品牌的影响力进而吸引更多用户，而负面评论却可以促进企业或商城对商品的持续改进。相对于普通用户，高活跃度用户浏览推荐商品的次数频繁，在浏览过程中发现自己感兴趣商品的概率随之提高，从而提升了商品购买概率。

定义 2 用户活跃度。是指根据用户对推荐列表中商品浏览次数和商品体验评价情况而反映出的用户在社会网络中的活跃程度。

用户活跃度采用式(2)计算得到：

$$Uac_i = \begin{cases} 1, & \lambda \times \dfrac{Rl_i}{Rl_{\max}} + (1-\lambda) \times \log_{C_{\max}}(C_i + 1) \times \log_{\overline{Cw}}(Cw_i + 1) \geqslant 1 \\ \lambda \times \dfrac{Rl_i}{Rl_{\max}} + (1-\lambda) \times \log_{C_{\max}}(C_i + 1) \times \log_{\overline{Cw}}(Cw_i + 1), & \text{其他} \end{cases} \tag{2}$$

其中：Uac_i 表示用户 i 在社交网络上的活跃度；Rl_i 表示用户 i 对推荐商品的浏览次数；$Rl_{\max}$ 表示用户对推荐商品的最大浏览量；$C_{\max}$ 表示社交网络中用户对商品的最大评论次数；Cw_i 表示用户 i 对商品的总评论字数；$\overline{Cw}$ 表示所有用户对商品的平均评论次数。在社交网络中，用户评论比浏览更能体现用户的活跃程度，适当地增加用户评论的比重可以减少随机浏览对用户活跃度的影响，所以依据经验将浏览所占比重 λ 设置为 0.4，评论所占比重为 0.6。

使用式(2)可以对用户在社交网络中的综合活跃度进行计算。当针对单一品牌商品进行推荐时，也可使用式(2)针对用户对某一品牌商品的浏览量、评论次数和字数等，计算用户在单一品牌上的用户活跃度。

2.3 用户信任度

在日常生活，尤其是电子商务中，信任表现出越来越重要的作用，是开展商品交易、交换的基础，成为电子商务网络中用户与用户间沟通的桥梁。推荐系统中融入信任关系，在一定程度上解决了数据稀疏问题，提高了推荐的准确性。

定义 3 用户信任度。是指电商网络中用户之间通过直接、间接联系表现出相对信任的程度，具有不对称性、双向性、传递性等特点。现实生活中，信任是人与人之间的一种相互关系，是双向且不对称的，所以用图论来描述社交网络中用户的信任关系恰到好处，因此在社交网络中，用户关系可用有向带权图 $G(U,E,X)$ 表示。其中：U 是网络中的节点集合，表示所有的用户；E 是网络中节点之间的有向边，表示用户间的交流关系，箭头指向消息的接收方；X 是节点间连线的权值，表示用户之间单向信息交流的数量[12]。用户间信息交流如图 2 所示。

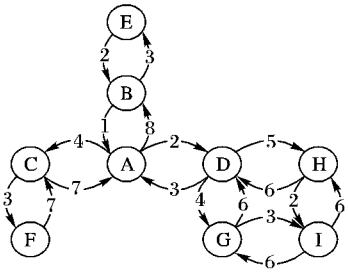


图 2 用户间信息交流示意图

Fig. 2 Schematic diagram of information exchange among users

因为用户间的信息交流是双向且不对称的，所以在计算信任度时不仅考虑信任的双向性，同时还要体现信任的不对称性。以图 2 中 D 节点为例，在计算 D 节点相对于其他节点的信任度时，要以 D 为中心向外扩散，选择由 D 向外指的权值(D→H→I)。扩散范围依据六度分离理论确定，对于因社

www.joca.cn

交网络不能覆盖或缺少关键中间联系人而造成的超出六度分离理论的用户予以保留,但并不作为关系节点参与计算[12]。

用户间信任度由式(3)计算:

$$Tr(D,U) = \sum_{i=1}^{n}\left[\prod_{j=1}^{m}\frac{X(S_{j-1},S_j)}{X_{j_total}}\right] \tag{3}$$

其中:D 是根节点;U 表示与 D 有信任关系的其他节点;i 表示由根节点 D 到目标节点 n 条路径的第 i 条;j 表示该条路径中由根节点 D 到目标节点的 m 层中的第 j 层(层表示由根节点向外,所有直接联系的节点为第一层,隔一个节点间接联系的为第二层,以此类推);$X(S_{j-1},S_j)$ 表示节点 S_{j-1} 与节点 S_j 之间的权值;X_{j_total} 表示第 j 层与根节点 D 相关联的节点的权值之和,图 2 中用户间节点的权值是信任度计算公式的核心参数。

3　实验与分析

为了验证信任关系和品牌概念融合的方法对推荐算法的影响,本文选择基于用户的协同过滤算法(UserCF)[1]、融合用户信任的协同过滤推荐算法(SPTUserCF)[1]、合并用户信任的协同过滤算法(MTUserCF)[17]作为对比算法,采用准确率、召回率和 $F1$ 值等指标与 TBCRMI 方法进行对比分析。

3.1　实验环境配置

本文实验用到的环境配置如下:操作系统 Linux Ubuntu 14.04,CPU i5-4590 3.30 GHz,内存 8 GB 或以上,可用硬盘空间 50 GB 以上,给虚拟机分配 24 GB 空间。算法采用 Java 语言编写。

本文实验处理的两个数据集:Amazon Food(642 MB)和 Unlocked Mobile phone(141 MB),由于数据量很大,为提高数据处理速度和实验运行速度,实验平台采用 Hadoop2.5.0 完全分布式集群,共添加 3 个节点。

算法中的半径参数 ε 根据用户距离计算得出[19],本文为 0.695,邻域密度阈值 $MinPts$ 取值为 3[19]。

3.2　实验数据集与评价指标

为验证本文方法的有效性,将 TBCRMI 与其他推荐算法进行对比,实验所用数据集为 Amazon Food(简称数据集 1)和 Unlocked Mobile phone(简称数据集 2)。两个数据集均是下载于聚数力网站的公共数据集:Amazon Food 是食品评论数据集,包括用户、评论内容、评论食品、食品评分等 9 项共计 50 多万条数据;Unlocked Mobile phone 是无锁手机评论数据集,包括用户、无锁移动手机的价格、用户评分、评论等 8 项共计 40 多万条数据。具体使用的数据信息见表 2,这两个数据集的下载地址均为:http://www.dataju.cn/Dataju/web/dataDescriptionAndDataset。

表 2　数据集信息表

Tab. 2　Dataset information

数据集	用户数	数据项目数	种类数	品牌数
数据集 1	14 410	568 455	1 061	367
数据集 2	9 859	413 705	357	35

本文所给的改进方法主要目的是提高推荐质量,所以在指标选择上选取了准确率、召回率以及 $F1$ 值等常用的推荐算法评价指标来进行对比分析。

1)准确率 P。推荐列表中用户喜欢的物品所占的比例,准确率计算如式(4)所示:

$$P = \sum_{u\in U} L_u \cap T_u \Big/ \sum_{u\in U} T_u \tag{4}$$

2)召回率 R。测试集中有多少用户喜欢的物品出现在推荐列表中,召回率计算如式(5)所示:

$$R = \sum_{u\in U} L_u \cap T_u \Big/ \sum_{u\in U} L_u \tag{5}$$

其中:L_u 表示通过训练集得出的推荐列表;T_u 表示测试集上用户的实际行为列表。

3)$F1$ 值。是信息检索领域常用的一个评价标准,其计算公式为:

$$F1 = 2PR/(P+R) \tag{6}$$

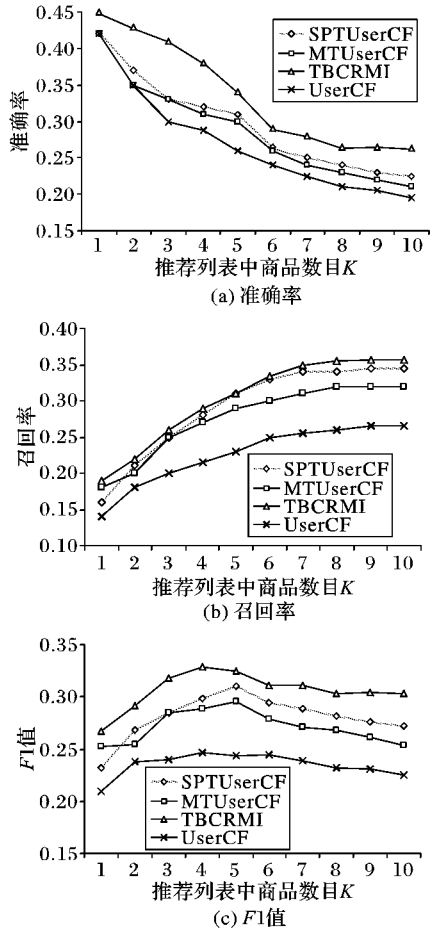
3.3　实验对比分析

在选择品牌效应进行推荐的情况下,推荐列表中商品的数目 K 会对推荐结果产生影响。将商品数目 K 作为变量,在两种数据集中对多种品牌综合推荐效果和单一品牌推荐效果进行多角度验证。

3.3.1　多种品牌综合推荐效果

在 Top-K 的推荐方法中,随着商品推荐数目的不同会产生不同的推荐效果。将商品推荐数目 K 作为自变量,分析不同的推荐算法在相同的数据集下产生的推荐效率。

针对多种不同品牌商品向用户进行综合推荐时,TBCRMI 方法和其他算法在准确率、召回率和 $F1$ 值的对比结果如图 3



(a) 准确率

(b) 召回率

(c) $F1$值

图 3　Amazon Food 数据集中不同 K 值的各指标对比

Fig. 3　Indexes comparison of different K on dataset Amazon Food

www.jocaa.cn

所示。

图 3 显示的是在 Amazon Food 数据集下，针对不同数目的推荐商品所对应的准确率、召回率和 $F1$。从图 3 的实验结果对比可以看出：TBCRMI 拥有最高的推荐准确率，并且明显高于其他方法。此外还可以看出随着推荐商品数目的增加准确率会降低，在 K 值从 1 增加到 10 的过程中：SPTUserCF 准确率下降 46.43%，MTUserCF 准确率下降 50%，UserCF 准确率下降 53.57%，TBCRMI 准确率下降 41.56%，可以看出本文方法在准确率保持方面取得较好效果。当 $K=4$ 时，TBCRMI 的推荐准确率比对比算法中效果最好的 SPTUserCF 还要提升 18.75%；召回率方面随着推荐商品数目的增加而上升，$K=4$ 时，TBCRMI 的召回率较 SPTUserCF 算法提升 3.57%；在 $F1$ 方面，$K=4$ 时，TBCRMI 比 SPTUserCF 提升 10.14%。

图 4 显示了在 Unlocked Mobile phone 数据集下，针对不同数目的商品进行推荐时的准确率、召回率和 $F1$。从实验结果可以看出 TBCRMI 的推荐效果最好，在准确率下降趋势方面，SPTUserCF 准确率下降 46.62%，MTUserCF 准确率下降 39.94%，UserCF 准确率下降 40.81%，TBCRMI 准确率下降 30.36%，依然取得最好的效果。在此次实验中，MTUserCF 是所选择的对比算法中效果最好的，但与本文提出的 TBCRMI 算法相比，在 $K=4$ 时，TBCRMI 的准确率比 MTUserCF 高 29.86%，召回率比其高 40.66%，$F1$ 值比其高 36.45%。

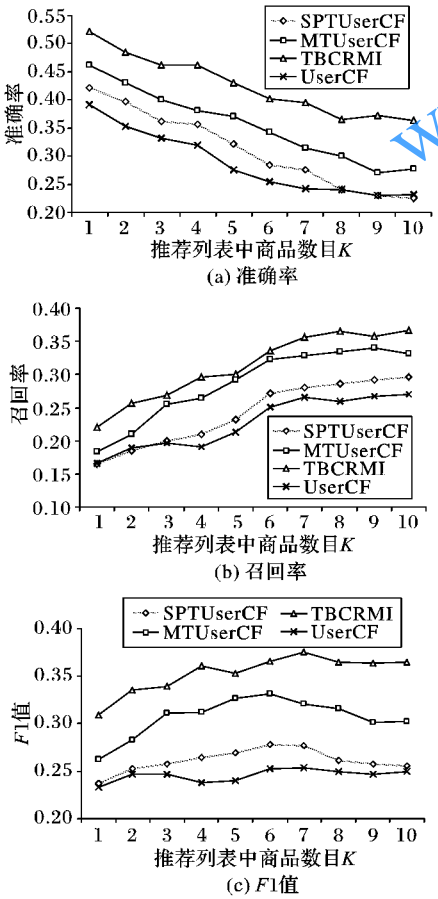


图 4 Unlocked Mobile phone 数据集中不同 K 值的各指标对比

Fig. 4 Indexes comparison of different K on dataset Unlocked Mobile phone

3.3.2 单一品牌推荐效果

图 5 是在 Amazon Food 数据集下，针对用户进行单一品牌商品推荐时的推荐准确率。从图 5 可以看出：TBCRMI 在准确率上取得了比较大的优势，当商品数目 $K=4$ 时，三种品牌平均准确率分别较 SPTUserCF、MTUserCF 和 UserCF 提升了 64.59%、69.9% 和 82.88%。

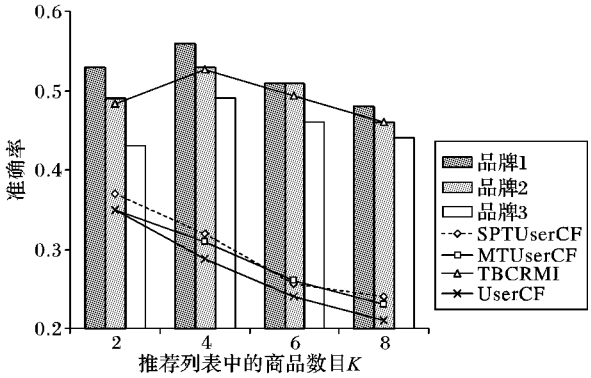


图 5 Amazon Food 数据集中单一品牌不同算法准确率对比

Fig. 5 Accuracy comparison of different algorithms to single brand on dataset Amazon Food

图 6 是在 Unlocked Mobile phone 数据集下，针对用户进行单一品牌商品推荐时的推荐准确率。从图 6 可以看出：TBCRMI 在准确率上的优势依然很明显，当商品数目 $K=4$ 时，三种品牌平均准确率分别较 SPTUserCF、MTUserCF 和 UserCF 提升了 52.23%、46.83% 和 63.85%。

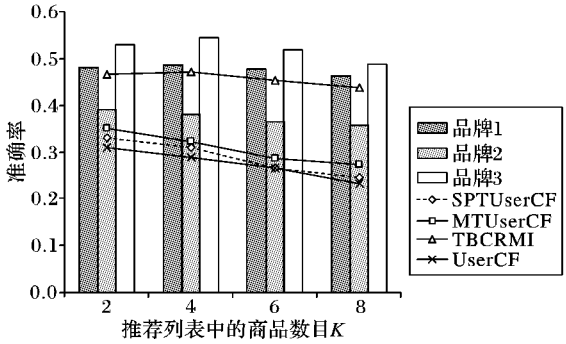


图 6 Unlocked Mobile phone 数据集中单一品牌不同算法准确率对比

Fig. 6 Accuracy comparison of different algorithms to single brand on dataset Unlocked Mobile phone

总体来看无论是多种品牌综合推荐还是单一品牌推荐，TBCRMI 都取得了比较突出的推荐效果，尤其是在单一品牌商品推荐方面，准确率有了较为明显的提高。

4 结语

本文在个性化商品推荐方法中引入了影响商品销售的两个关键因素：用户信任度和品牌认可度，提出了一种集成用户信任度和品牌认可度的商品推荐方法。该方法通过分析用户历史行为得到用户品牌认可度和活跃度，并据此进行用户聚类，融入用户信任度得到近邻关系进行推荐。所给方法充分利用了商品销售中的品牌理念，并且通过对比实验验证了所给方法在商品推荐的准确率、召回率及 $F1$ 值等多方面有所提

www.joca.cn

升,尤其是在单一品牌商品推荐方面准确率提高明显,效果显著。

参考文献(References)

[1] 李良, 董宇欣, 赵春晖, 等. 融合用户信任的协同过滤推荐算法[J]. 小型微型计算机系统, 2017, 38(5): 951 – 955. (LI L, DONG Y X, ZHAO C H, et al. Collaborative filtering recommendation algorithm combined with user trust[J]. Journal of Chinese Computer Systems, 2017, 38(5): 951 – 955.)

[2] 王海艳, 杨文彬, 王随昌, 等. 基于可信联盟的服务推荐方法[J]. 计算机学报, 2014, 37(2): 301 – 311. (WANG H Y, YANG W B, WANG S C, et al. A service recommendation method based on trust worthy community[J]. Chinese Journal of Computers, 2014, 37(2): 301 – 311.)

[3] LIU H, HE J, WANG T, et al. Combining user preferences and user opinions for accurate recommendation[J]. Electronic Commerce Research & Applications, 2013, 12(1): 14 – 23.

[4] McAULEY J, LESKOVEC J. Hidden factors and hidden topics: understanding rating dimensions with review text[C]// Proceedings of the 7th ACM Conference on Recommender Systems. New York: ACM, 2013: 165 – 172.

[5] MA S, COLDFARB D, CHEN L. Fixed point and Bregman iterative methods for matrix rank minimization[J]. Mathematics Programming Series A, 2011, 128(1): 321 – 353.

[6] WEN Z. Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm[J]. Mathematical Programming Computation, 2012, 4(4): 333 – 361.

[7] 陈万志, 张爽, 王德建, 等. 基于近邻模型与概率矩阵分解的高校选课推荐算法[J]. 辽宁工程技术大学学报(自然科学版), 2017, 36(9): 976 – 982. (CHEN W Z, ZHANG S, WANG D J, et al. A recommendation algorithm for elective courses in university based on nearest neighbor model and probability matrix factorization [J]. Journal of Liaoning Technical University, 2017, 36(9): 976 – 982.

[8] WANG D, MA J, LIAN T, et al. Recommendation based on weighted social trusts and item relationships[C]// Proceedings of the 29th Annual ACM Symposium on Applied Computing. New York: ACM, 2014: 254 – 259.

[9] BEDI P, SHARMA R. Trust based recommender system using ant colony for trust computation[J]. Expert System with Applications, 2012, 39(1): 1183 – 1190.

[10] 陈婷, 朱青, 周梦溪, 等. 社交网络环境下基于信任的推荐算法[J]. 软件学报, 2017, 28(3): 721 – 731. (CHEN T, ZHU Q, ZHOU M X, et al. Trust-based recommendation algorithm in social network[J]. Journal of Software, 2017, 28(3): 721 – 731.)

[11] HALIM R E. The effect of the relationship of brand trust and brand affect on brand performance: an analysis from brand loyalty perspective (a case of instant coffee product in indonesia)[EB/OL]. [2006-08-18]. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=925169.

[12] 乔秀全, 杨春, 李晓峰, 等. 社交网络服务中一种基于用户上下文的信任度计算方法[J]. 计算机学报, 2011, 34(12): 2403 – 2411. (QIAO X Q, YANG C, LI X F, et al. A trust calculating algorithm based on social networking service user's context[J]. Chinese Journal of Computers, 2011, 34(12): 2403 – 2411.)

[13] KRZYWICKI A, WOBCKE W, KIM Y S, et al. Collaborative filtering for people-to-people recommendation in online dating: data analysis and user trial[J]. International Journal of Human-Computer Studies, 2015, 76(4): 50 – 66.

[14] JONES K, LEONARD L. Trust in consumer-to-consumer electronic commerce[J]. Information & Mangement, 2008, 45(2): 88 – 95.

[15] FALOUTSOS M, KARAGIANNIS T, MOON S. Online social networks[J]. IEEE Network, 2010, 24(5): 4 – 5.

[16] MA H, KING I, LYU MR. Learning to recommend with social trust ensemble[C]// SIGIR 2009: Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2009: 397 – 406.

[17] GUO G, ZHANG J, THALMANN D. Merging trust in collaborative filtering to alleviate data sparsity and cold start[J]. Knowledge-Based Systems, 2014, 57(2): 57 – 68.

[18] 于金明, 孟军, 吴秋峰. 基于改进相似性度量的项目协同过滤推荐算法[J]. 计算机应用, 2017, 37(5): 1387 – 1391, 1406. (YU J M, MENG J, WU Q F. Item collaborative filtering recommendation algorithm based on improved similarity measure[J]. Journal of Computer Applications, 2017, 37(5): 1387 – 1391, 1406.)

[19] ESTER M, KRIEGEL H P, SANDER J, et al. A density-based algorithm for discovering clusters in large spatial databases with noise [C]// KDD 1996: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. Menlo Park, CA: AAAI Press, 1996: 226 – 231.

This work is partially supported by the National Natural Science Foundation of China (71771110), the Doctoral Scientific Research Foundation of Liaoning Province (20160199).

FENG Yong, born in 1973, Ph. D., professor. His research interests include data mining, personal recommendation.

HAN Xiaolong, born in 1991, M. S. candidate. His research interests include data mining, personal recommendation.

FU Chenping, born in 1992, M. S. candidate. Her research interests include deep learning, data mining.

WANG Rongbing, born in 1979, Ph. D., associate professor. His research interests include Internet of things, big data, could computing.

XU Hongyan, born in 1972, M. S., associate professor. Her research interests include deep Web, personal recommendation, data mining.

www.joca.cn