

一种基于领域信任及不信任的奇异值分解推荐算法

张 琦 柳 玲 文俊浩

(重庆大学大数据与软件学院 重庆 401331)

摘 要 传统协同过滤算法存在数据稀疏与冷启动问题,社会化推荐算法虽然能在一定程度上缓解这些问题,但大多数的算法都只从单一的角度来衡量信任关系的影响。为了更准确地度量社交关系对推荐预测的影响,提出了一种基于领域信任及不信任的社会化奇异值分解(Field Trust and Distrust based Singular Value Decomposition,FTDSVD)推荐算法。该算法在 SVD 推荐算法的基础上加入了用户的信任关系与不信任关系,利用不信任关系对社交关系进行修正,并且充分考虑用户的信任领域相关性和全局影响力。在 Epinions 数据集上将 FTDSVD 算法与相关算法进行了对比,结果证实了该算法在提高推荐质量和缓解冷启动问题上效果显著。

关键词 信任推荐,不信任关系,领域相关性,奇异值分解(SVD),推荐系统

中图法分类号 TP391 **文献标识码** A **DOI** 10.11896/jsjcx.190300388

Recommendation Algorithm with Field Trust and Distrust Based on SVD

ZHANG Qi LIU Ling WEN Jun-hao

(School of Big Data and Software Engineering,Chongqing University,Chongqing 401331,China)

Abstract The collaborative filtering algorithms in recommender systems usually suffer from data sparsity or cold-start problems. Although most of the existing social recommendation algorithms can alleviate these problems to a certain extent, they only measure the influence of trust relationship from a single aspect. In order to measure the influence of the social relationship on recommendation prediction more accurately, this paper proposed a novel social recommendation algorithm with field trust and distrust based on singular value decomposition (SVD), named FTDSVD. Based on the SVD algorithm, the trust relationship and distrust relationship information of users is added in order to correct the social relationship, and the global influence of users and the field relevance of trust are considered. Finally, it is compared with the state-of-the-art methods on the Epinions dataset. Experiment results show that the FTDSVD algorithm has obvious effects in improving the recommendation quality and alleviating the cold start problem.

Keywords Trust recommendation, Distrust relationship, Field correlation, Singular value decomposition (SVD), Recommender system

1 引言

互联网的发展使得网络信息量暴增。为帮助人们在信息过载的情况下更好地选择与决策,推荐系统应运而生并获得了长足的发展。传统的基于协同过滤的推荐系统在一定程度上缓解了信息过载的问题,但存在着数据稀疏性和冷启动等问题^[1-2];而基于矩阵分解的推荐系统,特别是奇异值分解(Singular Value Decomposition, SVD)算法^[3],能有效缓解这些问题。SVD 通过奇异值分解技术,不仅能将原始数据映射到低维度空间,还能挖掘用户的潜在因子和项目的潜在因子来估计缺失值,有效缓解了数据稀疏性问题。随着社交网络的迅猛发展,将社交数据中的信任关系应用到推荐系统中也

是近几年研究的热点^[4-5]。Massa^[6]针对信任的属性进行了分析,提出可以从局部和全局两个角度来定义信任。其中,全局影响力指一个用户在整个社交网络中的地位和声誉,在网络中拥有较高全局影响力的权威用户很少^[7],可此类用户的社交影响力却极为显著^[8-9]。目前,大多数对社交网络的研究都集中在信任及信任传递的问题上,忽视了不信任问题,而在实际中,Epinions 网站已允许用户对其他人进行不信任关系的标注。不信任关系对社交网络推荐有着积极的作用,且它的影响力大于信任评价的影响力^[10]。例如,A 信任 B, B 信任 C,如果仅看信任关系的传递性,A 将信任 C;但如果 A 直接有着不信任 C 的声明,那么 A 将不会受到来自 C 的推荐影响。不信任关系不仅能从另一方面解决评分数据及信任关系

到稿日期:2019-03-01 返修日期:2019-05-14 本文受国家自然科学基金(61502062),重庆市基础与前沿研究计划项目(cstc2015jcyjA40049)资助。

张 琦(1993—),女,硕士生,CCF 学生会员,主要研究方向为个性化推荐;柳 玲(1970—),女,博士,副教授,硕士生导师,CCF 会员,主要研究方向为推荐系统、数据仓库与数据挖掘,E-mail:Liuling@cqu.edu.cn(通信作者);文俊浩(1969—),男,博士,教授,博士生导师,CCF 高级会员,主要研究方向为服务计算、数据挖掘。

数据的稀疏问题,还可以修正用户间的信任关系^[9,11],从而减小恶意用户对推荐系统性能的影响^[12-13]。然而,由于不信任关系有着与信任关系不同的特点^[14],因此在推荐系统中融入不信任存在一定的挑战。

作为一个社会范畴的概念,信任在本质上拥有差异性和多面性^[15],如用户可能拥有多方面的兴趣爱好而专家通常仅精通某一领域的技能,对于不同的领域,用户的偏好程度以及对朋友的信任程度都不同,同时领域中的信任与用户间的整体信任程度也存在偏差^[16]。例如,用户 u 对他的家人和同学都存在信任关系,且对家人的信任程度更深,但在书籍推荐领域,用户更倾向于同学的推荐,而在食物口味上却更信任家人。因此,必须结合用户所涉及的领域信息,才能更加精准地描述用户间社交关系对推荐结果的影响。

Guo 等提出的 TrustSVD 算法^[17]在 SVD++ 算法的基础上加入了用户显式信任关系,通过信任关系的影响获得了较好的推荐结果,但并未考虑到不信任关系能缓解信任存在的数据稀疏和信任领域性的问题。Forsati 等^[18]提出的算法,将目标用户及其信任用户和不信任用户划分成三元组,若目标用户的评分特征与不信任用户更为相似,则将减小该社交关系对评分预测的影响。该算法虽然考虑到了信任关系和不信任关系的融合,但并未考虑信任关系只在用户兴趣相似的领域才能发挥作用。Tang 等^[9]提出的 RecSSN 算法也采用了惩罚机制,若目标用户与不信任用户圈更相似,则将强制用户与信任用户更相似。但该算法忽略了只有不信任关系的用户的影响力,且没有考虑不信任关系对全局影响力的作用。

综上分析,为了更好地在社交网络中融入不信任关系,并考虑到全局影响力以及信任关系的领域性问题,本文提出了一种基于领域信任及不信任的 SVD 算法——FTDSVD。该算法利用不信任关系来修正社交关系,并考虑到了信任的领域性问题,将根据用户的评分特征计算出用户间的领域相关性,再利用考虑了不信任关系的全局影响力对目标用户的信任好友进行筛选,从而减少领域相关性低的信任用户和不信任关系较多的用户对预测评分的贡献,有效缓解了数据稀疏和冷启动问题,提高了评分预测的准确性。最后,通过实验证明了所提算法可获得较好的推荐结果。

2 FTDSVD 算法

2.1 SVD 算法

在本文的推荐算法中,我们需要根据用户-项目评分矩阵来预测用户对未知项目的评分情况。设 $\mathbf{R}=[r_{u,i}]_{N \times M}$ 为评分矩阵,其中 N 和 M 分别为用户数及项目数, $r_{u,i}$ 为用户 u 对项目 i 的评分。通过传统的 SVD 算法,可以将该评分矩阵进行奇异值分解,得到两个低秩矩阵 $\mathbf{P} \in R^{d \times N}$, $\mathbf{Q} \in R^{d \times M}$, 即 $\mathbf{R} = \mathbf{P}^T \mathbf{D}$ 。其中, $\mathbf{P}$ 为用户特征矩阵, $\mathbf{P}^T$ 为其转置矩阵, $\mathbf{Q}$ 为项目特征矩阵。根据这两个特征矩阵,可以得到用户 u 和项目 i 的潜在特征向量分别为 $\mathbf{p}_u$ 和 $\mathbf{q}_i$, 则用户 u 对项目 i 的评分预测如下:

$$r_{u,i}^* = \mathbf{q}_i^T \mathbf{p}_u \quad (1)$$

由于在用户评分矩阵中存在着用户个人基准偏好这一因

素,即不同用户的打分标准不同,有些用户的打分比其他用户的高,导致某些物品的分数也偏高。因此,在上述算法的基础上加入基准偏移量 b_u 。对评分预测模型进行更改,具体表达式如下:

$$r_{u,i}^* = \mu + b_u + b_i + \mathbf{q}_i^T \mathbf{p}_u \quad (2)$$

其中, μ 为整个用户群的评分平均值; b_u 和 b_i 分别为用户 u 及物品 i 的评分偏置; $\mathbf{q}_i^T \mathbf{p}_u$ 为用户 u 对项目 i 的预测评分,通过用户 u 的特征向量 $\mathbf{p}_u$ 和项目 i 的特征向量 $\mathbf{q}_i^T$ 内积所得。最优优化损失函数为:

$$\min L = \frac{1}{2} \sum_{u=1}^N \sum_{i=1}^M (r_{u,i}^* - r_{u,i})^2 \quad (3)$$

2.2 领域信任和不信任关系的度量

2.2.1 信任和不信任关系的计算

在推荐系统中,通常认为社交关系与用户之间的评分偏好和历史交互有关。用户会对自己信赖的或者与自己评分相似的用户进行信任标注;而对与自己意见不同的用户,通常会给予不信任的评价。因此,用户与信任用户有着相似的评分习惯,而与不信任用户的评分习惯则有所差别。

根据社交关系,本文将用户分成三元组 (u, k, g) , 其中, k 为用户 u 的信任用户, g 则为不信任用户。 $\mathbf{p}_u, \mathbf{p}_k, \mathbf{p}_g$ 分别表示 3 个用户的评分的潜在特征,代表了用户平时的评分习惯。本文将用欧氏距离公式来衡量用户间评分习惯的不同。

$$D_{u,k} = \|\mathbf{p}_u - \mathbf{p}_k\|_2^2 \quad (4)$$

$$D_{u,g} = \|\mathbf{p}_u - \mathbf{p}_g\|_2^2 \quad (5)$$

式(4)表示用户 u 与它的信任好友用户 k 的潜在评分向量的不同;式(5)表示用户 u 与它不信任的用户 g 之间的潜在评分向量的不同。式(4)和式(5)的值越小,说明它们的评分习惯越相似;值越大,则表示它们的评分习惯越不同。

因为用户 u 信任用户 k , 所以他们在评分习惯上有一定的相似之处,或是用户 u 受到用户 k 的影响,它们之间的潜在评分向量相差不多;但用户 u 不信任用户 g , 所以它们之间的评分习惯、兴趣爱好相差甚远。因此, $\mathbf{p}_u$ 应该与 $\mathbf{p}_k$ 更相似, 且与 $\mathbf{p}_g$ 不同, 故 $D_{u,k}$ 应尽量小, $D_{u,g}$ 应尽量大。本文将用式(6)来衡量 3 个用户间的评分习惯差异:

$$\max(0, D_{u,k} - D_{u,g}) \quad (6)$$

若用户 u 与信任用户 k 更相似, 与不信任用户 g 更不同, 那么 $D_{u,k} - D_{u,g} < 0$, 最后式(6)的值为 0。用户间的信任关系将不会对最终的评分预测带来负面的影响, 因此不必对预测结果进行惩罚;若用户 u 与不信任用户 g 更相似, 而与它信任的用户 k 不那么相似, 则 $D_{u,k} - D_{u,g} > 0$, 那么将惩罚该社交关系对最后评分预测的影响。加入信任和不信任关系后, 最优优化损失函数为:

$$\min L = \frac{1}{2} \sum_{u=1}^N \sum_{i=1}^M (r_{u,i}^* - r_{u,i})^2 + \frac{\beta_1}{2} \sum_{(u,k,g) \in \Omega_s} \max(0, D_{u,k} - D_{u,g}) \quad (7)$$

其中, Ω_s 为根据社交关系形成的三元组集合: $\Omega_s = \{(u, k, g) \in [N] \times [N] \times [N]; S_{uk} = 1, S_{ug} = -1\}$; 参数 β_1 则为用户社交关系的协调因子, 控制着用户间社交关系对最终评分预测的贡献。

2.2.2 用户领域值的计算

虽然用户与信任好友的评分习惯相似,但每个用户的兴趣爱好广泛,平时所涉及的商品项目种类繁多,因此信任也具有领域性。用户信任关系对评分的影响,仅限于用户兴趣的相似领域。例如,用户 A 更愿意信任喜欢电影的用户 B 对电影的评价,虽然 A 更信任 C ,但 C 从未涉及电影,他对电影的评价并不具有说服力。因此,在进行推荐时,需要衡量用户间的领域相关性,同时考虑用户的社交关系和评分习惯。本文认为,若各个用户评分过的相同项目越多,则他们之间所涉及的项目范围越相似,他们的领域相关性也就越强;如果用户评分过的项目交集甚少,则他们的兴趣不同,因此他们间的信任关系对评分习惯并不能产生较大影响。用户间的领域值表示为:

$$f_{u,k} = \frac{|I_u \cap I_k|}{|I_u|} \quad (8)$$

其中, I_u 为用户 u 评分过的项目集, I_k 为用户 k 评分过的项目集, $I_u \cap I_k$ 表示两个用户共同评分过的项目集。共同项目集占有项目集的比例越大,两个用户间的兴趣爱好越相似,领域相关性也就越强,领域值亦越大。为防止用户 u 未评分过任何项目,导致用户共同评分项目集为 0,将式(8)改写为:

$$f_{u,k} = \begin{cases} \frac{|I_u \cap I_k|}{|I_u|}, & |I_u| > 0 \\ 0, & |I_u| = 0 \end{cases} \quad (9)$$

因此,考虑了用户领域值后,用户的社交关系便修正为:

$$\min L = \frac{1}{2} \sum_{u=1}^N \sum_{i=1}^M (r_{u,i}^* - r_{u,i})^2 + \frac{\beta_1}{2} \sum_{(u,k,g) \in \Omega_S} f_{u,k} \max(0, D_{u,k} - D_{u,g}) \quad (10)$$

2.3 用户全局影响力的计算

全局影响力表现为用户在整个社交网络中的声誉、地位和影响力^[6]。计算社交网络中用户全局影响力的算法较多,但大都未考虑用户的不信任关系对全局影响力的作用。若用户被大多数用户不信任,那么它的全局影响力会很小。而那些被大多数人信任的用户则拥有更大的全局影响力,且全局影响力所带来的权威性也使得它所不信任的用户应该拥有极低的全局影响力。根据 eBay's Feedback Forum^[19],全局影响力可以使用较为简单的统计法来计算,用户的全局影响力为正面评价数减去负面评价数。如果目标用户被越多的用户标记为不信任,那么它将会被越多的用户排斥,其全局影响力也就越小。对于不信任关系较多的用户来说,评分习惯反而会引起预测计算的偏差,因此全局影响力定义为:

$$Global_u = \begin{cases} \frac{|Te_u| - |Dte_u|}{N}, & |Te_u| - |Dte_u| > 0 \\ \frac{1}{N}, & |Te_u| - |Dte_u| \leq 0 \end{cases} \quad (11)$$

其中, $|Te_u|$ 为信任用户 u 的用户数, $|Dte_u|$ 为不信任用户 u 的用户数, N 为用户的总人数。用户被信任关系越多时,越多的信任用户会受到它评分习惯的影响,它的全局影响力也就越大;若不相信它的用户多于信任它的用户,则它的全局影响力为 $1/N$ 。

根据式(11)计算社交网络中所有用户的全局影响力值,

并根据该值的大小对所有用户进行排名,值越大,排名越靠前,全局影响力也就越大。设 $r_u \in [1, N]$,当 r_u 等于 1 时,表明排名最高。为了方便计算,通过式(12)将用户排名映射到 $[0, 1]$ 区间,则用户的全局影响力为:

$$G_u = \frac{1}{1 + \log(r_u)} \quad (12)$$

全局影响力越大的用户在推荐预测中的作用越大,即专家用户在推荐中起到引领作用。本文将考虑目标用户 u 及其信任用户 k 和不信任用户 g 三者的全局影响力。若目标用户 u 自身的全局影响力很大,那么它自身的评分是具有影响且真实的,则用户 k 和用户 g 几乎影响不了它的评分习惯,因此将削弱社交关系对评分预测的贡献。对于用户 k 和 g 来说,若信任用户 k 的全局影响力大于不信任用户 g ,则它对目标用户的评分预测的积极贡献应该更大;但如果是不信任用户 g 的全局影响力更大,则表示目标用户的社交关系存在异常,或者说可信度不高,那么它不正常的社交关系不能对它的最终推荐预测做出贡献,因此此社交关系的加入会给评分预测带来很大的偏差。因此,考虑了三者的全局影响力后,最优化损失函数定义为:

$$\min L = \frac{1}{2} \sum_{u=1}^N \sum_{i=1}^M G_u (r_{u,i}^* - r_{u,i})^2 + \frac{\beta_1}{2} \sum_{(u,k,g) \in \Omega_S} f_{u,k} \max(0, D_{u,k} - D_{u,g}) + \frac{\beta_2}{2} \sum_{(u,k,g) \in \Omega_S} M_{G_k, G_g} D_{u,k} \quad (13)$$

其中, M_{G_k, G_g} 表示用户 k 和 g 全局影响力的比较结果,如式(14)所示:

$$M_{G_k, G_g} = \begin{cases} G_k, & G_k \geq G_g \\ 0, & G_k < G_g \end{cases} \quad (14)$$

若不信任用户 g 的全局影响力更大,则表示目标用户的社交关系可信度不高,那么该社交关系对于评分预测的贡献为 0;若信任用户 k 的全局影响力更大,则信任用户对最后预测结果的影响也就越大。

为了防止预测过程中出现过拟合的情况,我们加入正则项,则最后的最优化损失函数为:

$$\min L = \frac{1}{2} \sum_{u=1}^N \sum_{i=1}^M G_u (\mu + b_u + b_i + \mathbf{q}_i^T \mathbf{p}_u - r_{u,i})^2 + \frac{\beta_1}{2} \sum_{(u,k,g) \in \Omega_S} f_{u,k} \max(0, \|\mathbf{p}_u - \mathbf{p}_k\|_2^2 - \|\mathbf{p}_u - \mathbf{p}_g\|_2^2) + \frac{\beta_2}{2} \sum_{(u,k,g) \in \Omega_S} M_{G_k, G_g} \|\mathbf{p}_u - \mathbf{p}_k\|_2^2 + \frac{\lambda}{2} (\|\mathbf{P}\|_F^2 + \|\mathbf{Q}\|_F^2 + b_u^2 + b_i^2) \quad (15)$$

其中, $\lambda (\lambda > 0)$ 是正则化参数; $\|\mathbf{P}\|$ 和 $\|\mathbf{Q}\|$ 表示 Frobenius 范数。为获得推荐模型,可通过梯度下降法对损失函数进行求解,学习 $\mathbf{P}, \mathbf{Q}$ 矩阵,并预测评分矩阵 $\mathbf{R}$ 中的缺失值。FTDSVD 算法的整体流程如算法 1 所示。

算法 1 FTDSVD 算法

输入: 评分矩阵 $\mathbf{R}$, 信任关系 T , 不信任关系 Dt , 隐类数 d , 调节因子 β_1 和 β_2 , 迭代次数 k

输出: 用户特征矩阵 $\mathbf{P}$, 项目特征矩阵 $\mathbf{Q}$

1. 根据信任关系和不信任关系计算出每个用户的全局影响力 $G\{G_{u1}, G_{u2}, \dots\}$;
2. 根据信任关系 T 和项目集 I 计算出信任用户间的领域值 $f_{u,k}$;

3. 初始化矩阵 $\mathbf{P}, \mathbf{Q}$ (随机数填充);
4. 根据式(4)、式(5)计算 $D_{u,k}, D_{u,g}$;
5. 求解式(15)的负梯度;
6. 按照负梯度的方向更新矩阵 $\mathbf{P}, \mathbf{Q}$;
7. 若迭代次数没有超过预设值 k , 则继续计算, 否则停止计算。

3 实验

3.1 实验数据集

本文使用 Epinions 数据集^[13]来验证所提 FTDSVD 算法的性能。该数据集来自一个流行的电子商务网站和客户审查网站 Epinions, 用户可以对电子产品、公司和电影等各种类型的项目进行评分, 并可以对项目评论进行打分。评分值为 1 到 5 的离散值, 1 表示没有帮助, 5 表示最有帮助。除了评分以外, 用户还可以与其他用户建立信任和不信任关系, 1 表示信任, 而-1 表示不信任, 若用户没有对其他用户进行评分, 则他们之间不存在社交关系。基于计算时间和存储空间考虑, 本文将使用 Epinions 数据集的部分数据, 数据集的具体信息如表 1 所列, 按 4:1 的比例将这些评分数据划分为训练集和测试集进行实验。

表 1 Epinions 数据集的信息
Table 1 Information of Epinions dataset

用户数	项目数	评分数	信任关系数	不信任关系数	评分密度	社交密度
1510	18068	53254	1461	728	0.0019	0.0009

3.2 评价指标

本文采用平均绝对值误差 (Mean Absolute Error, MAE) 及均方根误差 (Root Mean Squared Error, RMSE) 对实验结果进行评价, 其计算公式如下:

$$MAE = \frac{\sum_{u,i} |r_{u,i}^* - r_{u,i}|}{|R|}$$

(16)

$$RMSE = \sqrt{\frac{\sum_{u,i} (r_{u,i}^* - r_{u,i})^2}{|R|}}$$

(17)

其中, $|R|$ 为测试数据集中的评分数。对于推荐结果, MAE 及 RMSE 越小, 推荐结果就越准确。

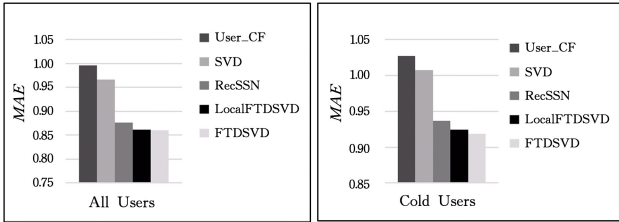
3.3 实验结果及分析

本实验主要将 FTDSVD 与传统基于用户的协同过滤算法 (User_CF)、SVD 算法、RecSNN^[9] 以及不考虑全局影响力的算法 LocalFTDSVD (将算法中的 β_2 设为 0) 的推荐结果进行比较。实验从两个不同的角度来进行对比, 一个是考虑所有用户, 另一个是只考虑冷启动用户, 本实验中冷启动用户为对项目评分低于 5 的用户。实验中, 将学习速率设为 0.01, 迭代次数为 100, 隐类数 d 取值为 10, β_1 和 β_2 取值各为 0.3 和 0.5, 实验结果的对比情况如表 2、图 1 和图 2 所示。

表 2 实验结果的对比

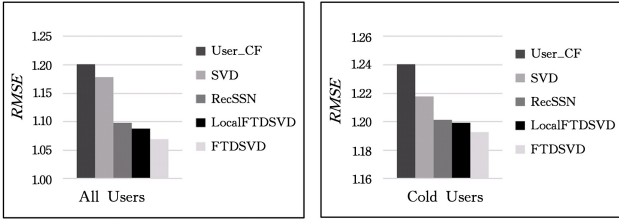
Table 2 Comparison of experiment results

实验条件	Metrics	User_CF	SVD	RecSSN	Local-FTDSVD	FTDSVD
ALL	MAE	0.9955	0.9667	0.8763	0.8617	0.8598
	RMSE	1.2005	1.1780	1.0979	1.0880	1.0697
Cold-Users	MAE	1.0271	1.0074	0.9363	0.9241	0.9189
	RMSE	1.2403	1.2177	1.2013	1.1994	1.1925



(a) 所有用户的 MAE (b) 冷启动用户的 MAE

图 1 MAE 的精度对比
Fig. 1 Comparison of accuracy of MAE



(a) 所有用户的 RMSE (b) 冷启动用户的 RMSE

图 2 RMSE 的精度对比
Fig. 2 Comparison of accuracy of RMSE

由实验结果可以看出, FTDSVD 算法在 MAE 和 RMSE 上都有一定的提高, 相对于传统的基于用户的协同过滤算法而言, 基于矩阵分解的算法在冷启动用户的问题上, 推荐质量有显著提高。而无论是只考虑局部影响力的 LocalFTDSVD 算法, 还是考虑了全局影响力的 FTDSVD 算法, 相较于基本 SVD 算法和 RecSSN 算法, 各个实验结果都有所提高。分析其原因可知, 本文不仅利用了社交关系中的信任关系, 还将不信任关系作为社交信息的补充, 极大程度地缓解了评分数据稀疏的问题。对于冷启动用户, 利用社交关系对矩阵分解的预测结果进行修正, 可提高对其预测评分的准确性。将全局影响力协调因子 β_2 设为 0, 本文算法即为只考虑用户间局部影响力的 LocalFTDSVD 算法。它与 RecSSN 算法^[9] 相似, 前者考虑了信任的领域性问题并加大了不信任关系对预测结果的影响, 它的推荐效果优于后者。但考虑了全局影响力的 FTDSVD 算法在 MAE 及 RMSE 上表现更好, 由此可知全局影响力的加入, 特别是在评分数据稀疏的情况下, 能够提高推荐质量并缓解冷启动问题。因此, 本文提出的 FTDSVD 算法可有效地提高推荐质量, 并且能更好地缓解冷启动问题。

调节因子 β_1 和 β_2 作为社交关系在算法中的权重调节参数, 其值影响着社交关系在推荐预测中所起作用的大小。若两个值较大, 则社交关系在推荐中的占比就更大, 从而削弱了传统评分数据在预测中的贡献; 若值越小, 则社交关系贡献得就越少, 若为 0, 则为传统的 SVD 算法本身。根据实验可得, 当 β_1 为 0.3, β_2 为 0.5 时, 在本文所使用的数据集中 MAE 和 RMSE 的结果更好。根据上述实验结果可知, 用户全局影响力的融入, 能明显地提高推荐精度。

结束语 传统协同过滤算法存在数据稀疏性与冷启动问题, 而社交关系信息的融入可有效地缓解上述问题; 除此之外, 用户之间不信任信息的加入, 缓解了信任信息的数据稀疏问题; 并且对信任关系进行修正, 能更好地提高推荐效率。本文在 SVD 算法的基础上加入了用户的信任关系与不信任关

系,并且考虑了用户的全局影响力,对目标用户的信任好友进行筛选时,有着较多不信任关系的信任好友将对预测评分贡献更少;考虑到信任的领域性问题,根据用户的评分特征计算出用户间的领域相关性,从而减小用户相互信任但用户兴趣却不相似的情况给最后推荐结果带来的影响。实验结果表明,与传统协同过滤、SVD 以及不考虑全局影响力的 Local-FTDSVD 算法相比,本文提出的 FTDSVD 算法进一步提升了推荐性能,因此本文所提出的想法是有效且可行的,可以为社会化推荐的进一步研究提供参考。

随着社会化推荐系统的发展,未来的研究工作可以考虑从以下几个方面开展:1)随着时间的推移,社交关系时刻都在变化,因此在推荐时有必要考虑社交关系的演化;2)不信任关系虽然能起到很大的作用,但它的传递性一直是不确定的,因此如何更进一步地使用不信任关系,也是未来研究需要关注的问题。

参 考 文 献

[1] ABBASI M A, TANG J L, LIU H. Trust-aware recommender systems[M]. USA:Chapman and Hall/CRC Press,2014:11-12.

[2] CAI H, JIA Y B, HUANG C W. Research of collaborative filtering recommendation based on user trust model[J]. CEA,2010, 46(35):148-151. (in Chinese)
蔡浩,贾宇波,黄成伟. 结合用户信任模型的协同过滤推荐方法研究[J]. 计算机工程与应用,2010,46(35):148-151.

[3] FUNK S. Netflix update; try this at home[EB/OL]. [2019-03-01]. <http://www.sifter.org/~simon/journal/20061211.html>.

[4] WU X W, LIU S D, ZHANG Y J, et al. Research on Social Recommender Systems[J]. Journal of Software,2015,26(6):1356-1372. (in Chinese)
孟祥武,刘树栋,张玉洁,等. 社会化推荐系统研究[J]. 软件学报,2015,26(6):1356-1372.

[5] ZHAO S, LIU X M, DUAN Z, et al. A Survey on Social Ties Mining[J]. Chinese Journal of Computers,2017,40(3):535-555.

[6] MASSA P. A survey of trust use and modeling in real online systems[M]. USA:IGIGlobal,2007:51-83.

[7] YUAN Q, CHEN L, ZHAO S. Factorization vs. regularization: Fusing heterogeneous social relationships in top-N recommendation[C]//Proceedings of the 2011 ACM Conference on Recommender Systems. New York:ACM,2011:245-252.

[8] TANG J, HU X, GAO H, et al. Exploiting local and global social context for recommendation[C]//Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence.

Menlo Park,CA:AAAI Press,2013:2712-2718.

[9] TANG J, AGGARWAL C, LIU H. Recommendations in Signed Social Networks[C]//Proceedings of the 25th International Conference on World Wide Web. New York:ACM,2016:31-40.

[10] ABBASSI Z, APERJIS C, HUBERMAN B A. Friends versus the crowd:tradeoffs and dynamics[R]. HP Report, Palo Alto: HP Labs,2013.

[11] RAFAILIDIS D. Modeling trust and distrust information in recommender systems via joint matrix factorization with signed graphs[C]//Proceedings of the 31st Annual ACM Symposium on Applied Computing. New York:ACM,2016:1060-1065.

[12] RAFAILIDIS D, CRESTANI F. Learning to Rank with Trust and Distrust in Recommender Systems[C]//Proceedings of the Eleventh ACM Conference on Recommender Systems. New York:ACM,2017:5-13.

[13] GUHA R, KUMAR R, RAGHAVAN P, et al. Propagation of trust and distrust[C]//Proceedings of the 13th International Conference on World Wide Web. New York:ACM,2004:403-411.

[14] TANG J, HU X, LIU H. Is distrust the negation of trust?: the value of distrust in social media[C]//Proceedings of the 25th ACM Conference on Hypertext and Social Media. New York:ACM,2014:185-199.

[15] TANG J, GAO H, LIU H. mTrust:Discerning Multi-Faceted Trust in a Connected World[C]//Proceedings of the Fifth ACM International Conference on Web Search and Data Mining. New York:ACM,2012:93-102.

[16] LIN T H, GAO C, LI Y. Recommender Systems with Characterized Social Regularization[C]//Proceedings of the 27th ACM International Conference on Information and Knowledge Management. New York:ACM,2018:1767-1770.

[17] GUO G B, ZHANG J, YORKE-SMITH N. TrustSVD: collaborative filtering with both the explicit and implicit influence of user trust and of item ratings[C]//Proceedings of the 29th AAAI Conference on Artificial Intelligence. Menlo Park,CA: AAAI Press,2015:123-129.

[18] FORSATI R, MAHDAVI M, SHAMSFARD M, et al. Matrix Factorization with Explicit Trust and Distrust Side Information for Improved Social Recommendation[J]. ACM Transactions on Information Systems,2014,32(4):1-38.

[19] RESNICK P, ZECKHAUSER R. Trust Among Strangers in Internet Transactions; Empirical Analysis of eBay's Reputation System[C]//Baye M R, ed. Advances in Applied Microeconomics. Amsterdam:Elsevier,2002:127-157.