

一种基于社交网络的非负矩阵分解算法

谢海迪¹, 周云², 李彤岩¹

(1. 成都信息工程大学通信工程学院, 四川 成都 610225; 2. 78111 部队, 四川 成都 610200)

摘要: 基于社交网络的推荐算法主要是将用户社交关系和评分信息相结合, 有效解决因缺乏评分数据而引起的冷启动问题。但基于社交网络的推荐算法只针对用户之间的相关性进行分析, 事实上用户之间的关系水平也会对推荐结果产生一定程度的影响。因此提出一种基于社交网络的非负矩阵分解算法 CTSVD。CTSVD 算法通过用户的社交网络进行信任和信任的亲密度计算, 更新用户之间信任值和信任值, 校正社交关系对预测结果的影响。通过在实际数据集 Epinions 的实验, 验证 CTSVD 方法的准确性, 并能较好地解决传统的冷启动问题。

关键词: 推荐系统; 社交网络; 信任关系; 亲密度; 可靠度

中图分类号: TP391

文献标志码: A

doi: 10.16836/j.cnki.jcuit.2024.01.009

0 引言

近几年互联网技术快速发展, 人类制造的数字信息量飞速增长, 同时也出现了信息过载的问题^[1]。推荐系统(recommended system, RS)是一种能够帮助使用者在大量资讯中发现自己感兴趣的资讯筛选工具。

基于用户的协同过滤推荐、基于物品的协同过滤推荐和基于模型的协同过滤推荐是协同过滤系统主要组成部分^[2]。其中最常用的矩阵分解算法^[3]是基于模型的协同过滤推荐中最重要的一种。矩阵分解包括的经典算法有奇异值分解算法(SVD++)^[4], 加权非负矩阵分解算法(WNMF)^[5], 概率矩阵分解算法(PMF)^[6]等。

矩阵分解算法最重要的问题是数据稀疏性, 会对推荐的结果产生不好的影响。常用的方法就是结合用户的社交信息进行推荐。Ma 等^[7]在概率矩阵分解的基础上, 提出一种将评分矩阵与用户社交矩阵相结合的推荐算法 SoRec。Yang 等^[8]提出一种通过对网络信息进行详细分解, 从信任用户和被信任用户两个角度考虑的混合推荐模型 TrustMF。Jamali 等^[9]提出将社交网络和矩阵分解相结合的推荐算法 SocialMF, 从而提高推荐的准确性。

本文提出一种基于社交网络的非负矩阵分解算法 CTSVD, 根据用户的社交网络与电路的相似度来确定用户之间的信任值和信任值。利用计算后的信任值和信任值修正用户的社交关系, 在增加信任用户对预测评分的影响同时减少不信任用户对预测评分的影响。通过实验, 验证了 CTSVD 方法可以很好地提高推荐结果的准确性, 缓解数据稀少对推荐结果的影响。

1 推荐算法

1.1 TrustSVD 算法

TrustSVD^[10]算法在 SVD++模型基础上, 融合评分信息和信任信息, 可以更好地应对冷启动问题。在这里, 评分矩阵用矩阵 $Y = [y_{u,i}]_{m \times n}$ 表示, $y_{u,i}$ 代表用户 u 对物品 i 的评分结果。并且, 通过对用户特征矩阵 P 与项目特征矩阵 Q 的乘积, 也就是 $Y \approx P^T Q$ 。 $T = [t_{u,v}]_{m \times m}$ 表示信任矩阵, 信任矩阵 T 由信任用户特征矩阵 P 和被信任用户特征矩阵 V 的乘积获得, 即 $T \approx P^T V$, 并且该潜在的特征矩阵 P 由评分矩阵和信任矩阵所共有的。用户 u 对物品 i 的预测评分为

$$\hat{y}_{u,i} = s_u + s_i + \mu + q_i^T (p_u + |I_u|^{-\frac{1}{2}} \sum_{i \in I_u} x_i + |T_u|^{-\frac{1}{2}} \sum_{v \in T_u} v_v)$$

式中 s_u 和 s_i 分别代表用户偏好和物品偏好, μ 代表所有用户评分的平均分, x_i 代表被评分的物品对评分结果的可能影响, v_v 表示被信任的用户的潜在特征矢量, $q_j^T v_v$ 代表用户对评分结果的影响, 最小化的目标函数如下:

$$L_{\text{TrustSVD}} = \frac{1}{2} \sum_u \sum_{j \in I_u} (\hat{y}_{u,j} - y_{u,j})^2 + \frac{\lambda}{2} \sum_u \sum_{v \in T_u} (\hat{t}_{u,v} - t_{u,v})^2 + \frac{\lambda}{2} \sum_u \left(\frac{\lambda}{2} |I_u|^{-\frac{1}{2}} + \frac{\lambda}{2} |T_u|^{-\frac{1}{2}} \right) \|p_u\|_F^2 + \frac{\lambda}{2} \sum_u |I_u|^{-\frac{1}{2}} b_u^2 + \frac{\lambda}{2} \sum_u |U_j|^{-\frac{1}{2}} b_j^2 + \frac{\lambda}{2} \sum_j |U_j|^{-\frac{1}{2}} \|q_j\|_F^2 + \frac{\lambda}{2} \sum_u |U_i|^{-\frac{1}{2}} \|x_i\|_F^2 + \frac{\lambda}{2} |T_v^+|^{-\frac{1}{2}} \|v_v\|_F^2$$

式中 U_i, U_j 分别表示对物品 i 和物品 j 进行评分的用户集合; T_v^+ 表示一组用户信任用户 v 。

1.2 信任度和不信任度计算

1.2.1 亲密度

张紫茵等^[11]提出一种融合直接和间接不对称信任关系相结合的推荐算法,本文用户亲密度指用户之间的信任程度,与电路拥有类似的特性,特别是串联和并联电路能很好地实现信任的传递和多路径。在此基础上,把社会网络看作是一张电路图,把社会中的每一位用户视为一个结点,并把社会联系看作电阻。如图 1 所示。

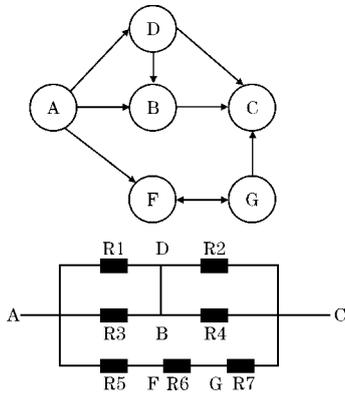


图 1 社交关系转化成电路

根据欧姆定律 $I = \frac{U}{R}$, 电流的幅值和电阻幅值成反

比,所以社交网络中,用户之间串联的节点越多,信任度就越小,反之并联的节点越多,信任度就越大。Abrams 等^[12]指出,与单向行为相比,交互的资讯传送效率更高。所以,在设置两个使用者间的电阻值时,应视使用者间的交互程度而定。

用户亲密度计算步骤:

步骤 1 对于任意一个用户 $v (v \in N_u^d)$, 根据用户社交网络得到用户 u 到用户 v 的所有路径,用 $\mathbf{P}(u \sim v)$ 表示。

步骤 2 构造一个路径矩阵 \mathbf{H} , 若 $\mathbf{P}(u \sim v)$ 中用户个数为 m , 则路径矩阵 \mathbf{H} 的大小为 $m \times m$ 。

步骤 3 填充路径矩阵 \mathbf{H} , 将开始用户 u 设为矩阵 \mathbf{H} 中的第一行第一列, 将结束用户 v 设为矩阵 \mathbf{H} 中的最后一行的末列; $\mathbf{P}^{\sim}(u, v)$ 表示 $\mathbf{P}(u \sim v)$ 中的某一条路径; \mathbf{P}_{xy}^{\sim} 属于路径 $\mathbf{P}^{\sim}(u, v)$ 中用户 x 和用户 y 的关系, 若两个用户之间相互作用, 则两个用户之间的电阻值更小, 设其对应的电导值 $G_2^* = 1$, 反之两个用户之间单向作用, 设其对应的电导值 $G_1^* = 0.6$, 电导值 $G^* = \frac{1}{R}$, 然后将两个用户的电导值负数填入路径矩阵 \mathbf{H} 对应的位置。

步骤 4 根据公式 $\mathbf{H}(i, i) = -\sum_{j=1}^m \mathbf{H}(i, j)$ 计算矩阵 \mathbf{H} 对角线的值, 将用户 v 作为零电压结点, 并将矩阵 \mathbf{H} 中 v 所对应的行和列去掉, 从而获得一个大小为 $(m-1) \times (m-1)$ 的新路径矩阵 \mathbf{H} 。

步骤 5 根据公式 $\mathbf{H} \times \mathbf{U}^* = \mathbf{I}^*$, 计算获得结点电

压向量 $\mathbf{U}^* = (u_1^*, u_2^*, \dots, u_{m-1}^*)$ 。

步骤 6 用户 u 和 v 之间的电压差为 $\Delta u^* = u_1^* - u_0^*$, 通过公式 $\mathbf{I}^*(u, v) = \Delta u^* / r_0$ 计算得到用户 u 和 v 之间的电流值, 也就是用户 u 和 v 的亲密度, 为便于计算设定 u_0^*, r_0 设为 1。

研究表明, 信任信息和不信任信息都会影响推荐质量^[13]。根据六度分割理论可知信任传递最大路径长度为 6^[14], 依据刘智捷等^[15]的研究设置信任传递最大路径长度 d 值为 3, 得到用户之间的传递路径, 根据用户亲密度计算步骤, 计算出用户之间的信任亲密度 $t(u, v)$ 和不信任亲密度 $d(u, v)$ 。

在考虑不信任关系后的将预测评分定义为

$$\hat{y}_{u,i} = s_u + s_i + \mu + q_i^T (p_u + |\mathbf{I}_u|^{-\frac{1}{2}} \sum_{j \in I_u} x_j + |\mathbf{T}_u|^{-\frac{1}{2}} \sum_{v \in T_u} v_v - |\mathbf{D}_u|^{-\frac{1}{2}} \sum_{v \in D_u} w_v)$$

式中 $\mathbf{D} = [d_{u,j}]_{m \times n}$ 表示不信任矩阵, 将不信任矩阵分解成两个低阶矩阵的乘积, \mathbf{P} 表示不信任者特征矩阵, \mathbf{W} 表示不被信任者特征矩阵即 $\mathbf{D} \approx \mathbf{P}^T \mathbf{W}$, 并且评分矩阵, 信任矩阵和不信任矩阵具有同一基本属性 \mathbf{P} 。 w_v 表示不被信任用户的潜在因子向量, $q_j^T w_v$ 表示不被信任用户对评分的影响, 因为不信任用户对预测评分具有消极影响, 所以在预测评分时, 通过减去不信任用户的消极影响得到最终的预测评分。损失函数变化为

$$\begin{aligned} L_{\text{CTSVD}} = & \frac{1}{2} \sum_u \sum_{j \in I_u} (\hat{y}_{u,i} - y_{u,j})^2 + \frac{\lambda_t}{2} \sum_u \sum_{v \in T_u} (\hat{t}_{u,v} - t_{u,v})^2 + \\ & \frac{\lambda_d}{2} \sum_u \sum_{v \in D_u} (\hat{d}_{u,v} - d_{u,v})^2 + \\ & \frac{\lambda}{2} \sum_u \left(\frac{1}{2} |\mathbf{I}_u|^{-\frac{1}{2}} + \frac{\lambda_t}{2} |\mathbf{T}_u|^{-\frac{1}{2}} \right) \|\mathbf{p}_u\|_F^2 + \frac{\lambda}{2} \sum_u |\mathbf{I}_u|^{-\frac{1}{2}} b_u^2 + \\ & \frac{\lambda}{2} \sum_u |\mathbf{U}_j|^{-\frac{1}{2}} b_j^2 + \frac{\lambda}{2} \sum_j |\mathbf{U}_j|^{-\frac{1}{2}} \|\mathbf{q}_j\|_F^2 + \\ & \frac{\lambda}{2} \sum_u |\mathbf{U}_i|^{-\frac{1}{2}} \|\mathbf{x}_i\|_F^2 + \frac{\lambda}{2} |\mathbf{T}_v^+|^{-\frac{1}{2}} \|\mathbf{v}_v\|_F^2 + \\ & \frac{\lambda}{2} |\mathbf{D}_v^-|^{-\frac{1}{2}} \|\mathbf{w}_v\|_F^2 \end{aligned}$$

1.2.2 用户可靠度计算

用户可靠度是指用户在社交网络中的影响力, 在社交网络中信任该用户的越多, 则其在社交网络中的信誉度就越高, 该用户更有可能得到其他用户的信任。根据 eBay's Feedback Forum^[16], 用户可靠度可以使用较为简单的方法进行计算, 用户可靠度为积极评价影响减去消极影响。因此提出用户可靠度的计算公式, 如下所示:

$$\text{re}(v) = \begin{cases} \text{re}_t(v) = \frac{2}{1 + e^{-(T_v - D_v)}}, & T_v - D_v \geq 0 \\ \text{re}_d(v) = \frac{2}{1 + e^{-(D_v - T_v)}}, & T_v - D_v < 0 \end{cases}$$

式中, $\text{re}(v)$ 代表用户 v 的可靠度, T_v 代表信任用户 v 的用户数目, D_v 代表不信任用户 v 的用户数目。在考

虑用户可靠度后的损失函数定义为

$$L_{\text{CTSVD}} = \frac{1}{2} \sum_u \sum_{j \in I_u} (\hat{y}_{u,j} - y_{u,j})^2 + \frac{\lambda_t}{2} \sum_u \sum_{v \in T_u} \text{re}_t(v) (\hat{t}_{u,v} - t_{u,v})^2 + \frac{\lambda_d}{2} \sum_u \sum_{v \in D_u} \text{re}_d(v) (\hat{d}_{u,v} - d_{u,v})^2 + \frac{\lambda}{2} \sum_u \left(\frac{1}{2} |I_u|^{-\frac{1}{2}} + \frac{\lambda_t}{2} |T_u|^{-\frac{1}{2}} \right) \|p_u\|_F^2 + \frac{\lambda}{2} \sum_u |I_u|^{-\frac{1}{2}} b_u^2 + \frac{\lambda}{2} \sum_u |U_j|^{-\frac{1}{2}} b_j^2 + \frac{\lambda}{2} \sum_j |U_j|^{-\frac{1}{2}} \|q_j\|_F^2 + \frac{\lambda}{2} \sum_u |U_i|^{-\frac{1}{2}} \|x_i\|_F^2 + \frac{\lambda}{2} |T_v^+|^{-\frac{1}{2}} \|v_v\|_F^2 + \frac{\lambda}{2} |D_v^-|^{-\frac{1}{2}} \|w_v\|_F^2$$

在原始数据集中,用户评分过的物品数量和用户社交关系数量都很少,评分矩阵和社交矩阵都非常稀疏。通过计算用户亲密度可以获得用户之间更多的社交关系,完善社交矩阵,可以得到更准确的用户特征向量。算法流程图如图2所示。

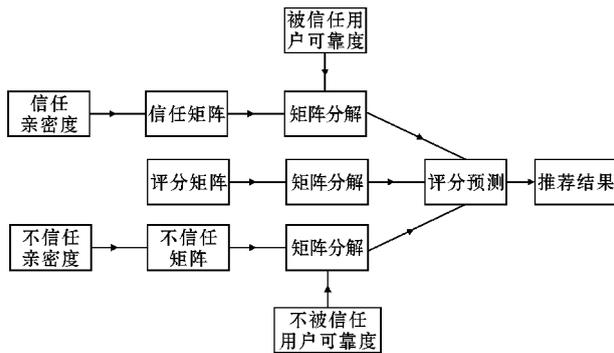


图2 CTSVD算法流程图

2 实验结果与分析

2.1 数据集

本文将唯一包含不信任信息的公开数据集 Epinions 作为实验数据集。该数据集中包含 120492 个用户,755760 个物品,13668320 条评分数据,715760 条信任数据,123705 条不信任数据。由于计算机的内存有限和数据集的缺乏,将 Epinions 数据集拆分为两个不相交的数据集 D1, D2。D1 数据集包含 19569 个用户,79097 个用户,646710 个评分,92602 个信任值声明,10562 个不信任值声明,评分密度为 0.00042,社交密度为 0.00027。D2 数据集包含 35404 个用户,103024 个用户,1118861 个评分,189139 个信任值声明,25229 个不信任值声明,评分密度为 0.00042,社交密度为 0.00002。

对数据集进行 80% ~ 20% 的分割。其中 80% 是训练集合,20% 是测试集合。实验以 python3.8 为主要开发语言,在 64 位的 windows10,2.4GHzCPU 和 16 GB 内存的电脑上进行。

2.2 实验指标

本文采用平均绝对误差 MAE 和均方根误差 RMSE 这两个评价指标^[17],MAE 和 RMSE 的计算方法如下:

$$\text{MAE} = \frac{\sum_{r_{u,j} \in R} |y_{u,j} - \hat{y}_{u,j}|}{|R|}$$

$$\text{RMSE} = \frac{\sqrt{\sum_{r_{u,j} \in R} (y_{u,j} - \hat{y}_{u,j})^2}}{|R|}$$

在以上的评估指标的公式中,|R|代表数据集中评分的数量, $y_{u,j}$ 代表用户 u 在物品 i 上的真实评分。MAE 和 RMSE 的数值愈低,说明该算法准确性愈高。

2.3 对比算法和参数设置

CTSVD 算法通过与 4 种典型算法的对比,以证明 CTSVD 方法的正确性和可行性。

SVD++^[4]:一种在 SVD 基础上加入用户的隐式反馈的改进算法。

SoRec^[7]:利用概率矩阵分解方法,将用户的评分矩阵和用户的社交网络进行整合。

TrustSVD^[13]:基于 SVD++,除了增加用户的信任关系,还考虑信任用户之间的显示反馈和隐式反馈。

TDSVD^[18]:一种不仅考虑用户的信任关系还考虑用户的不信任关系的推荐算法。

本文的实验参数设置为:D1 数据集: $\lambda=0.1, \lambda_t=2.5, \lambda_d=1.5, u_0^*=1, r_0=1$,学习率 $\eta=0.007$,迭代次数 $n=100$;D2 数据集: $\lambda=0.1, \lambda_t=1.4, \lambda_d=1.0, u_0^*=1, r_0=1$,学习率 $\eta=0.003$,迭代次数 $n=100$ 。

2.4 实验结果

实验结果如表 1 ~ 2 所示,CTSVD 算法与其他算法相比,结果更加优异。在 D1 数据集中 CTSVD 算法的 MAE 值为 0.84091, RMSE 值为 0.92003, CTSVD 算法的误差与已有算法的误差相比得到了很大程度的下降,MAE 值和 RMSE 值下降了 11% 左右;在 D2 数据集中 CTSVD 算法的 MAE 值为 0.87235, RMSE 值为 0.93772, MAE 值和 RMSE 值下降了 11.4% 左右。CTSVD 算法在 D2 数据集上的效果更好,表明 CTSVD 算法在数据稀疏的数据集上的提升效果更好。

表1 D1数据集上算法对比实验结果

参数	SVD++	SoRec	TrustSVD	TDSVD	CTSVD
MAE	0.85394	0.93732	0.84801	0.84586	0.84091
RMSE	0.93422	1.00036	0.92592	0.92399	0.92003

表2 D2数据集上算法对比实验结果

参数	SVD++	SoRec	TrustSVD	TDSVD	CTSVD
MAE	0.89628	0.97218	0.88392	0.88363	0.87235
RMSE	0.97540	1.03082	0.95010	0.94863	0.93772

图 3、图 4 分别表示 D1、D2 数据集上不同 $\eta, \lambda, \lambda_t, \lambda_d$ 对 MAE 值的影响。随着学习率和 λ 值的不断变大, MAE 值先减小后变大, 随着 λ_t 和 λ_d 值的逐渐增加, MAE 值逐渐收敛。在 D1 数据集上, 当学习率 η

$= 0.007, \lambda = 0.1, \lambda_t = 2.5, \lambda_d = 1.5$ 时, CTSVD 的 MAE 为最小值, 获得最佳推荐效果。在 D2 数据集上, 当学习率 $\eta = 0.003, \lambda = 0.1, \lambda_t = 1.4, \lambda_d = 1.0$ 时, CTSVD 的 MAE 为最小值, 获得最佳推荐效果。

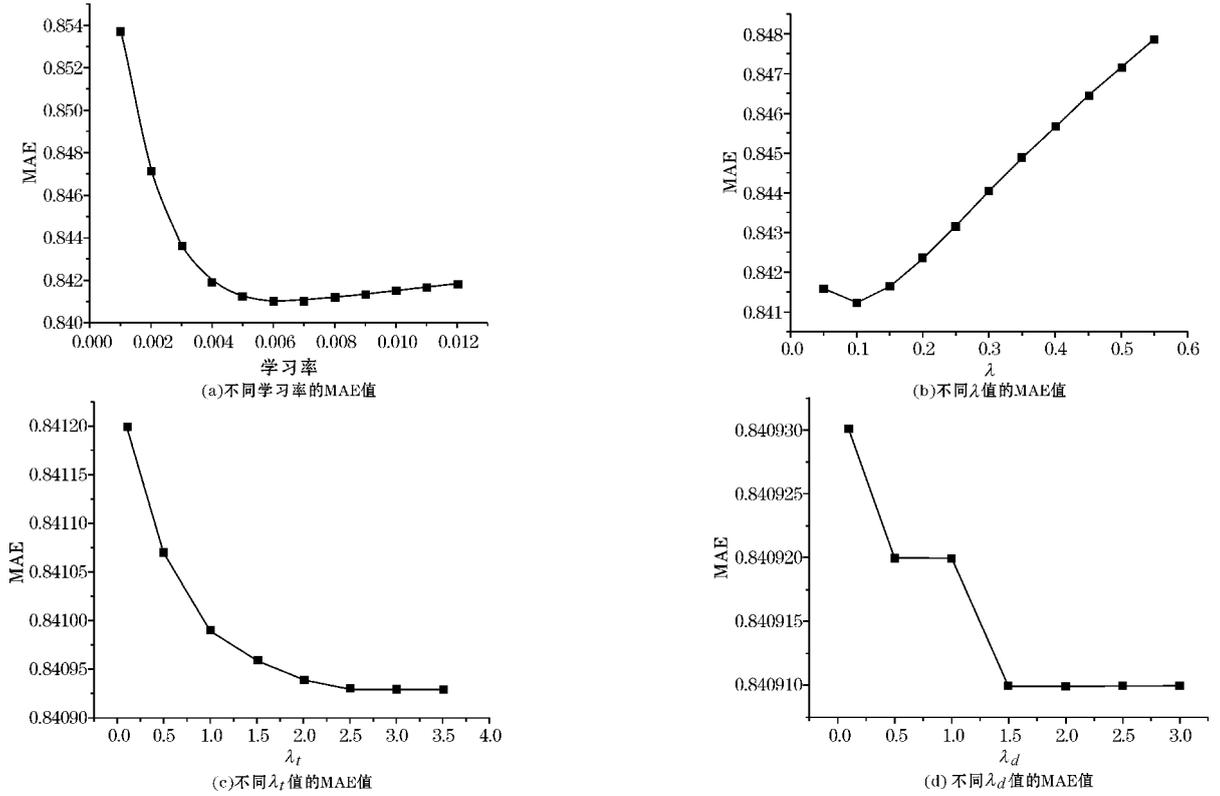


图 3 D1 数据集上不同 $\eta, \lambda, \lambda_t, \lambda_d$ 对 MAE 值的影响

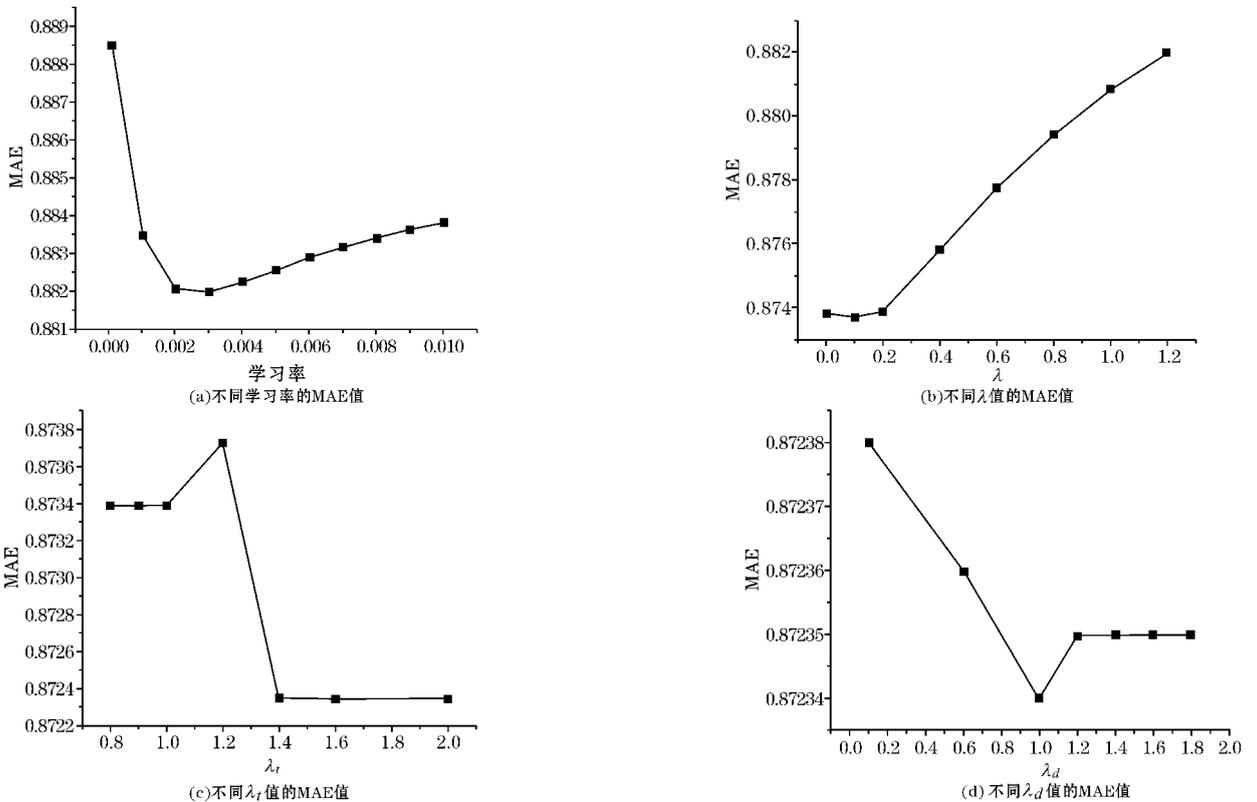


图 4 D2 数据集上不同 $\eta, \lambda, \lambda_t, \lambda_d$ 对 MAE 值的影响

鲁棒性是评测推荐系统抗作弊能力的一个指标,为验证 CTSVD 算法的鲁棒性,在训练过程中对评分数据、信任数据、不信任数据加入了高斯白噪声,即

$$r_{u,i} \leftarrow r_{u,i} + \varepsilon \quad \varepsilon \sim N(0, \delta^2)$$

$$t_{u,v} \leftarrow t_{u,v} + \varepsilon \quad \varepsilon \sim N(0, \delta^2)$$

$$d_{u,v} \leftarrow d_{u,v} + \varepsilon \quad \varepsilon \sim N(0, \delta^2)$$

用含有高斯噪声的数据训练 CTSVD 模型,通过不同噪声强度下模型的 MAE 值或 RMSE 值来验证模型的鲁棒性。

图 5 和图 6 分别表示 D1、D2 数据集中不同噪声强度下的 MAE 的变化情况。其中 x 轴表示噪声强度, y 轴表示 MAE 值, MAE 值越小表示模型预测的准确度越高。从图 5 和图 6 可以看出,随着噪声强度变化,误差值也随之发生改变,但变化幅度都很小。这说明模型具有较强的鲁棒性。

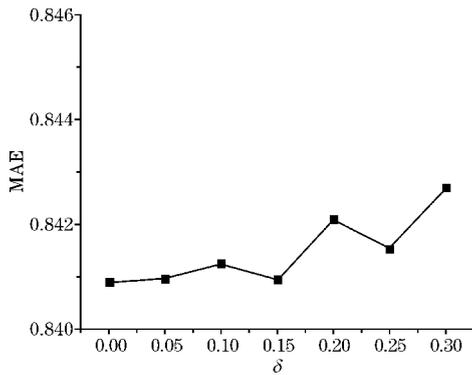


图 5 D1 数据集上不同噪声强度下的 CTSVD 算法的 MAE 值

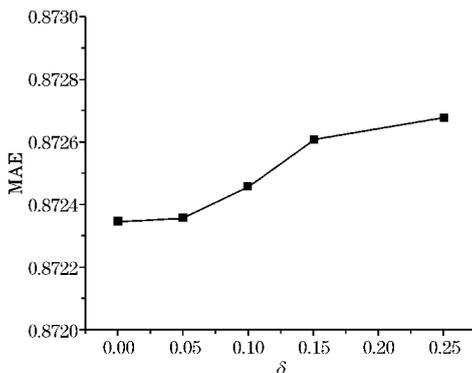


图 6 D2 数据集上不同噪声强度下的 CTSVD 算法的 MAE 值

表 3 和表 4 给出了不同算法的收敛时间,可以看出收敛时间最短的是 SVD++ 算法, CTSVD 算法的收敛时间最长。因为 SVD++ 算法是最简单的模型,它仅对评分数据进行矩阵分解,而 TrustSVD 算法则包含信任关系, TDSVD 算法包含不信任关系,两者都不仅在评分数据上进行矩阵分解,还对社交数据进行矩阵分解,因此它们比 SVD++ 算法的收敛性要好得多。

表 3 D1 数据集上不同算法的收敛时间比较 单位:s

算法	SVD++	SoRec	TrustSVD	TDSVD	CTSVD
Time	64	163	457	575	833

表 4 D2 数据集上不同算法的收敛时间比较 单位:s

算法	SVD++	SoRec	TrustSVD	TDSVD	CTSVD
Time	164	590	2000	2717	5267

本文的 CTSVD 算法不仅考虑评分信息、信任信息和不信任信息,还考虑用户社交之间的直接和间接关系,所以算法的收敛时间最长。虽然与其它算法相比 CTSVD 算法需要花费更多的时间,但是 CTSVD 算法是最佳的推荐方法,因此实际应用中要结合实际情况来选择合适的方法。

3 结束语

一些关于社交网络推荐的算法,不能很好地体现用户之间信任和不信任程度,本文在此基础上,给出一个基于社交网络的非负矩阵分解算法。CTSVD 方法相对于传统的社会网络推荐算法,利用社交网络来确定用户的信任度和不信任度,填充社交矩阵,表现社交程度。最后在实验结果上也表明 CTSVD 算法的推荐准确率优于其他算法。

将来可以在考虑时间的影响下,对 CTSVD 的算法作进一步完善。由于用户的社会关系网及其兴趣项目都是随时间而变化的,因此可以在一定程度上考虑时间因素来提高算法的鲁棒性。

参考文献:

- [1] Sahu A K, Dwivedi P J A I. User profile as a bridge in cross-domain recommender systems for sparsity reduction[J]. Applied Intelligence, 2019, 49(7): 2461-2481.
- [2] 项亮. 推荐系统实践[M]. 北京:人民邮电出版社, 2012:197.
- [3] Sardanios C, Ballas Papadatos G, Varlamis I. Optimizing parallel collaborative filtering approaches for improving recommendation systems performance [J]. Information, 2019, 10(5): 155.
- [4] Koren Y. Factorization meets the neighborhood: a multifaceted collaborative filtering model[C]. KDD '08: Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and

- data mining,2008:426–434.
- [5] 杨志君,叶东毅.基于加权的不完备非负矩阵分解算法[J].计算机应用,2010,30(5):1280–1283.
- [6] Cui Z, Xu X. Personalized recommendation system based on collaborative filtering for IoT scenarios [J]. IEEE Transactions on Services Computing, 2020, 13(4):685–695.
- [7] Ma H, Yang H, Lyu M R, et al. Sorec: social recommendation using probabilistic matrix factorization [C]. CIKM '08: Proceedings of the 17th ACM conference on Information and knowledge management, 2008:931–940.
- [8] Yang B, Lei Y, Liu J, et al. Social collaborative filtering by trust[J]. transactions on pattern analysis, 2016, 39(8):1633–1647.
- [9] Jamali M, ESTER M. A matrix factorization technique with trust propagation for recommendation in social networks [C]. RecSys '10: Proceedings of the fourth ACM conference on Recommender systems, 2010:135–142.
- [10] Guo G, Zhang J, Yorke-Smith N. TrustSVD: Collaborative filtering with both the explicit and implicit influence of user trust and of item ratings [C]. Twenty-Ninth AAAI Conference on Artificial Intelligence, 2015.
- [11] 张紫茵,张恒汝,徐媛媛,等.用户非对称信任关系的推荐算法[J].计算机科学,2018,45(10):37–42.
- [12] Zhang S, Zhu J. Reliable Potential Friends Identification Based on Trust Circuit for Social Recommendation [C]. WASA 2020: Wireless Algorithms, Systems, and Applications, 2020:716–729.
- [13] 张琦,柳玲,文俊浩.一种基于领域信任及不信任的奇异值分解推荐算法[J].计算机科学,2019,46(10):27–31.
- [14] 郑鹏,王应明,梁薇.基于信任和矩阵分解的协同过滤推荐算法[J].计算机工程与应用,2018,54(13):34–40.
- [15] 刘智捷.基于融合信任关系的协同过滤推荐算法[J].杭州电子科技大学学报(自然科学版),2018,38(3).
- [16] Resnick P, Zeckhauser R. Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system [M]. The Economics of the Internet and E-commerce, Emerald Group Publishing Limited, 2002.
- [17] 张大鹏,张伟.基于综合信任的奇异值分解推荐算法研究[J].高技术通讯,2021,31(1):102–112.
- [18] 乔猛,魏国亮,吴超异.一种基于信任和不信任的矩阵分解推荐算法[J].小型微型计算机系统,2021,44(1):56–62.

A Non-negative Matrix Decomposition Algorithm based on Social Networks

XIE Haidi¹, ZHOU Yun², LI Tongyan¹

(1. School of Communication Engineering, Chengdu University of Information Technology, Chengdu 610225, China; 2. Unit 78111, Chengdu 610200, China)

Abstract: The current social network-based recommendation algorithm mainly combines user social relations and scoring information, so as to effectively solve the cold start problem caused by the lack of scoring data. However, the current social network-based recommendation algorithms only analyze the correlation between users, which can have an impact on the recommendation results. Therefore, this paper proposes a non-negative matrix factorization algorithm CTSVD based on social network. The CTSVD algorithm calculates the intimacy of trust and distrust through the user's social network, updates the trust value and distrust value between users, and corrects the influence of social relations on the prediction results. The results of experiments using the actual dataset Epinions shows that the accuracy of the CTSVD method is verified, and the traditional cold start problem can be solved well.

Keywords: recommendation system; social networking; a relationship of mistrust; intimacy; reliability;