文章编号: 1007-5321(2014)03-0018-05 DOI: 10.13190/j.jbupt.2014.03.004

基于流形排序的社会化推荐方法

胡 祥[1,2], 王文东[1], 龚向阳[1], 王 柏[3], 阙喜戎[1]
(1. 北京邮电大学 网络与交换技术国家重点实验室, 北京 1000876; 2. 华北电力大学 控制与计算机工程学院, 北京 102206;
3. 北京邮电大学 计算机学院, 北京 100876)

摘要: 提出一种基于流形排序和社会化矩阵分解的推荐方法,采用流形排序方法度量用户间的社会相似度,利用正则化技术构建用于评分矩阵因式分解的目标函数,将用户之间的偏好差异作为目标函数的惩罚项,从而将用户之间的社会相似性融入评分矩阵的低阶矩阵分解过程. 实验结果表明,在大型的数据集上,该方法获得了比当前同类方法更好的推荐精度和更低的评分预测均方根误差/评分预测平均绝对误差(RMSE/MAE)值.

关 键 词: 社会化推荐; 流形排序; 矩阵分解
中图分类号: TN393.4 **文献标志码:** A

Social Recommendation Based on Manifold Ranking

HU Xiang[1,2], WANG Wen-dong[1], GONG Xiang-yang[1], WANG Bai[3], QUE Xi-rong[1]
(1. State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China;
2. School of Control and Computer Engineering, North China Electronic Power University, Beijing 102206, China;
3. School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: A new recommendation method based on manifold ranking and social matrix factorization is proposed, in which the social similarities among users are calculated by means of manifold ranking, the objective function of ratings matrix factorization is constructed via the regularization technique, with the differences among users' preferences as the penalty of objective function, the social similarities are infused into the low-rank matrix factorization. Experiments show that this method achieves higher precisions and lower root mean square error/mean absolute error (RMSE/MAE) value than other that of cognate methods.

Key words: social recommendation; manifold ranking; matrix factorization

推荐技术的研究吸引了来自机器学习、数据挖掘、信息检索等不同领域的研究者,但是,大多数推荐系统都受以下问题的困扰: 1) 数据稀疏性问题, Badrul 等[1]的研究表明,大多数用户只对极少数物品评分,稀疏的评分数据常使得推荐系统无法准确捕捉用户的偏好; 2) 冷启动问题,当新构建的推荐系统没有评分历史数据或者新用户加入时,系统无法学习用户的偏好,因此推荐效果不理想; 3) 传统的推荐技术忽略了用户之间的社会关系这一重要因素. 笔者的研究目标是将用户的社会关系融入推荐系统,解决或减轻推荐系统所面临的数据稀疏性、冷启动等问题,使推荐过程更加符合现实情况,提高推荐系统的性能. 笔者主要基于如下假设: 用户的偏好可以通过对评分矩阵的因式分解得到; 用户的偏

收稿日期: 2013-07-07
基金项目: 高等学校博士学科点专项科研基金资助项目(20130005110011); 北京市高等学校青年英才计划项目(71A1311172); 中央高校基本科研业务费专项项目
作者简介: 胡 祥(1976—),男,博士生,E-mail: xianghu.fox@gmail.com; 王文东(1963—),男,教授.

好不是独立的,会受到来自其朋友的影响; 用户之间的社会关系越亲近,他们的偏好就越相似.

1　相关工作

构建社会化推荐系统的过程可分成紧密相关的两个任务,一个是度量用户的社会相似度,另一个是利用社会关系和评分数据构建社会化推荐系统. 度量用户的社会相似度本质上是图上的连接预测问题,即度量图上结点之间的相似性的问题. 基于共同邻居的度量方法计算量小,但只考虑到图结构的本地一致性; 基于路径的度量方法虽然考虑到图结构的全局一致性,但是往往计算量很大. 而流形排序算法同时考虑了图结构的本地一致性和全局一致性,迭代过程简单,适用于各种类型的图. 针对社会网络的大规模和稀疏性特点,笔者拟用流形排序算法度量用户的社会相似度.

当前,已有研究者提出若干社会化推荐系统,SocialMF 模型[2] 将信任关系融入矩阵分解,但该方法假定用户的偏好完全由其朋友的偏好均值决定. STE 模型[3] 认为用户的偏好由自己和所信任的人共同决定,但训练过程比较复杂. Ma 等[4] 提出了基于正则化技术的 SR1 模型和 SR2 模型,SR1 模型假定用户的偏好取决于其直接朋友偏好的加权平均值,SR2 模型则强调偏好在用户之间的传播以及由此形成的直接/间接朋友之间偏好的一致性,实验结果表明,SR2 模型显著优于 SR1 模型.

针对 SR2 模型强调偏好传播的特点,利用基于得分传播机制的流形排序方法度量用户之间的全局相似度的方法,并提出一种结合 SR2 模型和流形排序方法的新的社会化推荐方法.

2　基于流形排序的社会化推荐方法

2.1　问题定义

推荐系统包含用户集合 $U=\{u_1,u_2,\cdots,u_M\}$ 以及物品集合 $O=\{i_1,i_2,\cdots,i_N\}$,用户对物品的评分为 $\boldsymbol{R}=[R_{u,i}]_{M\times N}$,$R_{u,i}$ 表示用户 u 对物品 i 的评分,分值常常是 1 ~ 5 的整数. 用户的社会关系由邻接矩阵 $\boldsymbol{A}=[A_{u,v}]_{M\times M}$ 表示,$A_{u,v}$ 表示用户 u 和用户 v 之间存在的社会关系,1 表示存在,0 表示不存在. 社会化推荐系统的任务是: 如果用户 u 没有对物品 i 评分,根据评分矩阵 $\boldsymbol{R}$ 和社会关系矩阵 $\boldsymbol{A}$ 来预测用户 u 对物品 i 的评分 $\hat{R}_{u,i}$.

2.2　推荐系统的矩阵分解方法

隐语义模型[5] 认为每个用户的偏好仅由少量的因素决定,物品的特征也是由少数因素决定的. 假设 M,N 分别表示用户和物品数量,D 为特征的个数,且 $D<\min(M,N)$,以矩阵 $\boldsymbol{U}_{M\times D}$ 表示用户的偏好,每个用户的偏好用向量 $\boldsymbol{U}_i$ 表示,即 $\boldsymbol{U}$ 中的一行; 同理,以矩阵 $\boldsymbol{V}_{N\times D}$ 表示物品的特征,每个物品的特征表示为 $\boldsymbol{V}_j$; 用 $\boldsymbol{R}_{M\times N}$ 表示评分矩阵. Thomas[5] 认为评分 $\boldsymbol{R}_{u,i}$ 由用户 u 的偏好和物品 i 的特征共同决定,可以利用低阶矩阵分解方法进行评分预测,$\boldsymbol{R}\approx\boldsymbol{U}^{\mathrm{T}}\boldsymbol{V}$. 该低阶矩阵分解问题可转化为一个最优化问题,$\mathscr{L}_2=\frac{1}{2}\|\boldsymbol{R}-\boldsymbol{U}\boldsymbol{V}^{\mathrm{T}}\|_{\mathrm{F}}$,这里 $\|\cdot\|_{\mathrm{F}}$ 表示矩阵的 Frobenius 范数,$\mathscr{L}_2$ 为评分的预测误差的平方和,通过求解该函数的最小值,可以得到 $\boldsymbol{U}$ 和 $\boldsymbol{V}$. 由于 $\boldsymbol{R}$ 中大多数元素缺失,所以可进一步将该函数的优化问题重写为

$$\min_{\boldsymbol{U},\boldsymbol{V}}\mathscr{L}_2(\boldsymbol{R},\boldsymbol{U},\boldsymbol{V})=\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}\delta_{i,j}(\boldsymbol{R}_{i,j}-\boldsymbol{U}_i^{\mathrm{T}}\boldsymbol{V}_j)^2 \tag{1}$$

其中: δ 为指示函数,如果 $\boldsymbol{R}_{i,j}$ 有值,则 $\delta_{i,j}$ 为 1,如果 $\boldsymbol{R}_{i,j}$ 缺失,则 $\delta_{i,j}$ 为 0,式(1) 中为避免过度拟合的问题,进一步在目标函数加入变量 $\boldsymbol{U},\boldsymbol{V}$ 的正则项

$$\min_{\boldsymbol{U},\boldsymbol{V}}\mathscr{L}_2(\boldsymbol{R},\boldsymbol{U},\boldsymbol{V})=\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}\delta_{i,j}(\boldsymbol{R}_{i,j}-\boldsymbol{U}_i^{\mathrm{T}}\boldsymbol{V}_j)^2+\frac{\lambda_1}{2}\|\boldsymbol{U}\|_{\mathrm{F}}+\frac{\lambda_2}{2}\|\boldsymbol{V}\|_{\mathrm{F}} \tag{2}$$

其中 λ_1、λ_2 为正则化系数. 利用梯度下降算法可以求解上述最优化问题,获得目标函数的局部最小值以及对应的 $\boldsymbol{U}$ 和 $\boldsymbol{V}$. Ruslan 等[6] 利用图模型给出了低阶矩阵分解的概率解释.

2.3　基于流形排序的社会相似度度量

协同过滤方法常使用基于评分的相似度,如 VSS、PCC、Cosine 等,为了将用户的社会关系融入推荐系统,笔者使用用户的社会相似度,提出基于流形排序的社会相似度度量方法. 流形排序的过程可以形象地描述为: 首先根据社会关系构造一个权重图,并且为目标用户赋以一个初始的正得分,其余用户的初始得分设为 0; 然后,所有的节点迭代地将其得分传播给与其相邻的节点,当网络达到全局平衡状态时, 传播过程停止; 网络平衡后各个用户得分可被用来度量他们与目标用户的社会相似度.

将用户的社会关系用邻接矩阵 $\boldsymbol{A}=[A_{u,v}]_{M\times M}$

表示;用户的社会相似度用矩阵 $\boldsymbol{S}=[S_{u,v}]_{M\times M}$ 表示,所有元素的初始值为零,$\boldsymbol{S}$ 矩阵的第 u 列记为向量 $\boldsymbol{S}_u$,$\boldsymbol{I}_{M\times M}$ 为单位矩阵,$\boldsymbol{I}$ 中的第 u 列记为向量 $\boldsymbol{I}_u$. 将基于流形排序的社会相似度度量算法描述如下.

算法1 基于流形排序的社会相似度度量算法

输入: $\boldsymbol{A}_{M\times M}$,$\alpha$

输出: $\boldsymbol{S}_{M\times M}$

//变量初始化 $\boldsymbol{S}$

$\boldsymbol{S}\leftarrow 0_{M\times M}$

//用迭代方式计算社会相似度 $\boldsymbol{S}$

$\boldsymbol{W}\leftarrow \mathrm{diag}(\mathrm{sum}(\boldsymbol{A}))$

$\boldsymbol{P}\leftarrow \boldsymbol{W}^{-\frac{1}{2}}\boldsymbol{A}\boldsymbol{W}^{-\frac{1}{2}}$

for t from 1 to $\mathrm{maxIter}_1$

  for u from 1 to M

    $\boldsymbol{S}_u(t+1)\leftarrow \alpha\boldsymbol{P}\boldsymbol{S}_u(t)+(1-\alpha)\boldsymbol{I}_u$

  end

end

算法首先对 $\boldsymbol{A}$ 进行对称规范化 $\boldsymbol{P}\leftarrow \boldsymbol{W}^{-\frac{1}{2}}\boldsymbol{A}\boldsymbol{W}^{-\frac{1}{2}}$,这里 $\boldsymbol{W}$ 为一个对角矩阵,$W_{u,u}$ 为各个结点的度,对应于 $\boldsymbol{A}$ 矩阵第 u 行元素的和. 然后对 $\boldsymbol{S}_u(t+1)=\alpha\boldsymbol{P}\boldsymbol{S}_u(t)+(1-\alpha)\boldsymbol{I}_u$ 进行迭代,直到收敛. 若用 $\boldsymbol{S}_u^*$ 表示序列$\{\boldsymbol{S}_u(t)\}$的极限,则用户 u 与其他用户之间的社会相似度用 $\boldsymbol{S}_u^*$ 来表示. Zhou 等[7]证明该流形排序的结果最终会收敛到 $\boldsymbol{S}_u^*=(1-\alpha)(\boldsymbol{I}-\alpha\boldsymbol{P})^{-1}\boldsymbol{I}_u$. 由于收敛结果包含求逆运算,大数据集会产生很大计算开销,所以本算法选择迭代方式求解. 参数 α 在 $[0,1)$ 之间取值,参数 $\mathrm{maxIter}_1$ 用于调节流形排序的迭代次数.

2.4 将社会相似度融入矩阵分解

社会网络中,用户与朋友们的偏好是趋于一致的,将用户 u 和用户 f 的偏好分别用 $\boldsymbol{U}_u$ 和 $\boldsymbol{U}_f$ 表示,则用户 u 和用户 f 的偏好的差异可表示为 $\|\boldsymbol{U}_u-\boldsymbol{U}_f\|_{\mathrm{F}}$,用户 u 与朋友们总的偏好差异可以表示为 $\sum_{f\in \mathrm{OF}(u)}\|\boldsymbol{U}_u-\boldsymbol{U}_f\|_{\mathrm{F}}$,式中 $\mathrm{OF}(u)$ 表示用户 u 认识的朋友,根据上文假设,社会关系越亲近的朋友,对用户 u 产生的影响越大,用 $\boldsymbol{S}$ 代表社会相似度,可以得到用户 u 的加权的偏好差异为 $\sum_{f\in \mathrm{OF}(u)}\boldsymbol{S}_{u,f}\|\boldsymbol{U}_u-\boldsymbol{U}_f\|_{\mathrm{F}}$,将所有用户与朋友间的偏好差异作为惩罚项加入到式(2),得到新的融入社会化关系的矩阵分解目标函数,如式(3)所示.

$$\min_{\boldsymbol{U},\boldsymbol{V}}\mathscr{L}_2(\boldsymbol{R},\boldsymbol{U},\boldsymbol{V})=\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}\boldsymbol{I}_{i,j}(\boldsymbol{R}_{i,j}-\boldsymbol{U}_i^{\mathrm{T}}\boldsymbol{V}_j)^2+\frac{\beta}{2}\sum_{u=1}^{m}\sum_{f\in \mathrm{OF}(u)}\boldsymbol{S}_{u,f}\|\boldsymbol{U}_u-\boldsymbol{U}_f\|_{\mathrm{F}}+\frac{\lambda_1}{2}\|\boldsymbol{U}\|_{\mathrm{F}}+\frac{\lambda_2}{2}\|\boldsymbol{V}\|_{\mathrm{F}} \tag{3}$$

其中:β 是正则化系数,与式(2)类似,式(3)可以通过梯度下降搜索目标函数 $\mathscr{L}_2$ 的局部最小值,进而求解 $\boldsymbol{U}$、$\boldsymbol{V}$. 为此,这里推导 $\mathscr{L}_2$ 对 $\boldsymbol{U}$、$\boldsymbol{V}$ 的偏导数如式(4)所示.

$$\frac{\partial\mathscr{L}_2}{\partial\boldsymbol{U}_i}=\sum_{j\in\{j|\boldsymbol{R}_{i,j}>0\}}(\boldsymbol{U}_i^{\mathrm{T}}\boldsymbol{V}_j-\boldsymbol{R}_{i,j})\boldsymbol{V}_j+\lambda_1\boldsymbol{U}_i+\beta\sum_{f_1\in \mathrm{OF}(i)}\boldsymbol{S}_{i,f_1}(\boldsymbol{U}_i-\boldsymbol{U}_{f_1})+\beta\sum_{f_2\in \mathrm{IF}(i)}\boldsymbol{S}_{i,f_2}(\boldsymbol{U}_i-\boldsymbol{U}_{f_2})$$

$$\frac{\partial\mathscr{L}_2}{\partial\boldsymbol{V}_j}=\sum_{i\in\{i|\boldsymbol{R}_{i,j}>0\}}(\boldsymbol{U}_i^{\mathrm{T}}\boldsymbol{V}_j-\boldsymbol{R}_{i,j})\boldsymbol{U}_i+\lambda_2\boldsymbol{V}_j \tag{4}$$

2.5 SMF-MR 推荐算法

根据以上分析,笔者提出基于流形排序和社会化矩阵分解的推荐方法(SMF-MR,social matrix factorization with manifold ranking),先利用流形排序度量用户的社会相似度,然后将社会关系融入评分矩阵的因式分解. 算法2给出 SMF-MR 推荐算法的算法描述.

算法2 SMF-MR 推荐算法

输入: $\boldsymbol{R}_{M\times N}$,$\boldsymbol{A}_{M\times M}$,$\beta$

输出: $\hat{\boldsymbol{R}}_{M\times N}$

//变量初始化

$\boldsymbol{U}_{M\times D}\leftarrow \mathrm{normal}(0,1)$,$\boldsymbol{V}_{N\times D}\leftarrow \mathrm{normal}(0,1)$

由算法1计算社会相似度 $\boldsymbol{S}$

//用梯度下降算法求解 $\boldsymbol{U}$,$\boldsymbol{V}$

for epoch from 1 to $\mathrm{maxIter}_2$

  foreach (i,j) in $\{(i,j)\,|\boldsymbol{R}_{i,j}>0\}$

    根据式(4)计算 $\frac{\partial\mathscr{L}_2}{\partial\boldsymbol{U}_i}$,$\frac{\partial\mathscr{L}_2}{\partial\boldsymbol{V}_j}$

    $\boldsymbol{U}_i\leftarrow\boldsymbol{U}_i-\mathrm{rate}\frac{\partial\mathscr{L}_2}{\partial\boldsymbol{U}_i}$

    $\boldsymbol{V}_j\leftarrow\boldsymbol{V}_j-\mathrm{rate}\frac{\partial\mathscr{L}_2}{\partial\boldsymbol{V}_j}$

  end

end

//计算预测评分 $\hat{\boldsymbol{R}}$

$\hat{\boldsymbol{R}}=\boldsymbol{U}\boldsymbol{V}^{\mathrm{T}}$

算法的输入包括评分数据 $\boldsymbol{R}_{M\times N}$、社会关系数据 $\boldsymbol{A}_{M\times M}$,输出为预测评分 $\hat{\boldsymbol{R}}_{M\times N}$. 除了已经提到的参数

λ_1、λ_2、D、β,为了控制算法训练过程,算法还额外引入一些参数: 包括学习率 rate,社会化分解的迭代次数 maxIter_2. 用户和物品的特征向量用 D 维标准正态分布的随机样本进行初始化.

3　实验结果及评价

3.1　数据集

为了验证 SMF-MR 算法的有效性,实验将基于 Flixster 和 Douban 两个真实的数据集. Flixster 是一个电影社交网站,Flixster 数据集包含大约 100 万个用户、820 万个评分、4.9 万部电影、2 670 万个双向的朋友关系,评分为取值 [0.5,5] 之间的离散数值,共分 10 级. Daban 是中国最大的在线社会网络之一,Douban 数据集包含大学 12.9 万用户、5.9 万部电影、168.3 万个评分、169.2 万个双向朋友关系,其中评分为 1 ~5 之间的数值.

3.2　评估方法

采用了 2 个度量方法进行评估,评分预测的平均绝对误差(MAE,mean alosolute error) 和评分预测的均方根误差(RMSE,root mean square error) . 实验将评分数据 $\boldsymbol{R}$ 按比率分成 $\boldsymbol{R}_1$和 $\boldsymbol{R}_t$ 2 个集合,集合 $\boldsymbol{R}_1$ 用于学习,集合 $\boldsymbol{R}_t$ 用于测试,MAE 可以定义为 $\frac{1}{|\boldsymbol{R}_t|}\sum_{u,i}|\boldsymbol{R}_{u,i}-\hat{\boldsymbol{R}}_{u,i}|$. 其中: $\boldsymbol{R}_{u,i}$表示用户 u 对物品 i 的真实评分,$\hat{\boldsymbol{R}}_{u,i}$表示推荐系统对给定的用户 u 及物品 i 的预测评分. $|\boldsymbol{R}_t|$表示用户测试后的评分的个数. RMSE 定义为 $\sqrt{\frac{1}{|\boldsymbol{R}_t|}\sum_{u,i}(\boldsymbol{R}_{u,i}-\hat{\boldsymbol{R}}_{u,i})^2}$.

可以看出,MAE 和 RMSE 越低预测的精度越高,推荐系统的性能越好.

3.3　实验结果及分析

为了评估笔者提出推荐方法的性能,通过实验与另外 3 种推荐方法进行了比较: 1) 协同过滤方法,使用最广泛的基于内存的推荐方法; 2) 概率矩阵分解方法[8],采用基本的矩阵分解方法,未将用户的社会关系考虑在内; 3) 社会正则化方法(SR2) [9],将社会关系融入到矩阵分解,该方法的用户相似度采用了基于评分的 VSS 相似度和 PCC 本地相似度,而 SMF-MR 模型采用了基于流形排序的全局相似度.

笔者设计了 2 个实验,实验 1 分为 2 组,分别度量各种推荐方法的 MAE 值和 RMSE 值. 第 1 组实验在 Flixster 数据集上进行,由于 Flixster 数据集比较稀疏,实验随机选取 90% 和 80% 的评分数据作为训练集,余下部分作为测试集. 第 2 组实验在 Douban 数据集上进行,实验随机选取 80% 和 60% 的评分数据作为训练集,余下的为测试集. 为了得到稳定的度量结果,实验将每组实验重复 5 次,度量的结果取各次度量结果的平均值. 实验中,正则化参数 λ_1、λ_2、β 的取值均为 10^{-3},维度 D 取值为经验值 10,由于协同过滤不是基于矩阵分解的方法,所以不考虑这几个参数; 流形排序参数 α 取经验值 0.99. 各种推荐方法的性能比较如表 1 所示.

表 1　各种推荐方法的推荐性能比较(维度 $D=10$)

训练集	数据集占比/%	度量	协同过滤	概率矩阵分解	SR2_VSS	SR2_PCC	SMF-MR
Flixster	90	MAE	0.713 0	0.695 1	0.675 8	0.675 6	0.668 4
		RMSE	0.914 2	0.878 2	0.852 9	0.851 7	0.841 9
	80	MAE	0.716 6	0.698 0	0.676 9	0.676 2	0.669 2
		RMSE	0.926 9	0.882 2	0.860 7	0.857 4	0.843 1
Douban	80	MAE	0.576 7	0.569 3	0.554 8	0.554 3	0.554 1
		RMSE	0.723 5	0.720 0	0.699 2	0.698 8	0.698 5
	60	MAE	0.578 3	0.573 7	0.559 8	0.559 3	0.559 3
		RMSE	0.736 0	0.729 0	0.704 6	0.704 2	0.703 5

从表 1 的结果可以看出,笔者提出的推荐方法所得到的 MAE 和 RMSE 结果稳定地低于其他方法. 而 Douban 数据集上的实验结果低于 Flixster 数据集上的实验结果,这说明 Flixster 数据集可能包含较多的噪声.

上述实验中,SMF-MR 中的参数 λ_1、λ_2、D、α 采用了普遍接受的经验值[8],而参数 β 控制着社会关系在系统中的重要性,β 越大则社会关系对系统的影响越大,因此,实验 2 着重研究了参数 β 对 SMF-MR 性能的影响,通过调整 β 的不同取值,观察 β 对

SMF-MR 模型性能的影响，结果如图 1 和图 2 所示. 图 1 显示了参数 β 在不同的取值情况下，SMF-MR 在两个数据集上的评分预测的 RMSE 误差的变化情况，从图 1 可以看出，β 值小于 10^{-4} 或者大于 10^{-2} 时都会引起 RMSE 的值上升(系统性能下降). 图 2 显示了参数 β 对评分预测的 MAE 的影响，β 值小于 10^{-4} 或者大于 10^{-3} 时都会引起 MAE 值上升，因此，实验 1 中 β 的值设为 10^{-3} 是合理的. 上述实验结果同时也说明，适当地考虑社会关系对推荐过程的影响，可以进一步提高传统推荐系统的性能.

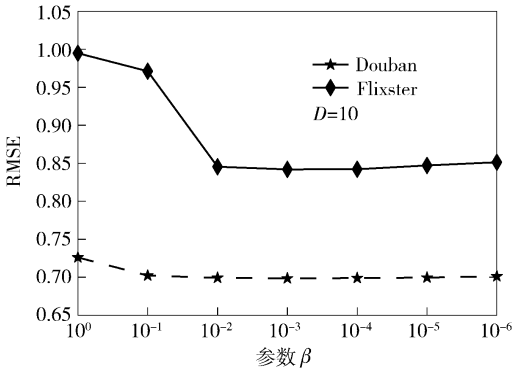

图 1 参数 β 对 RMSE 的影响

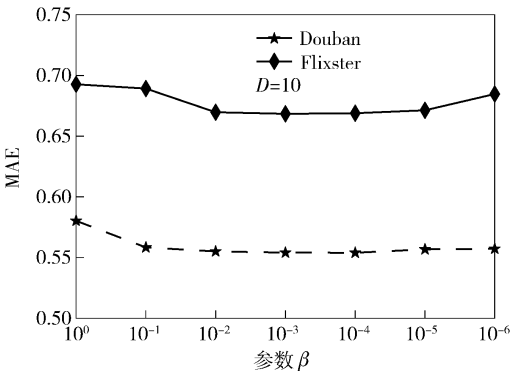

图 2 参数 β 对 MAE 的影响

4 结束语

提出一种基于流形排序和矩阵分解的社会化推荐方法(SMF-MR)，该方法将用户之间的社会相似性融入到评分矩阵的低阶矩阵分解过程，采用流形排序方法度量用户的社会相似度，通过学习得到表示用户偏好和物品特征的低阶隐含语义矩阵，进而利用低阶隐含语义矩阵进行评分预测. 实验表明，该方法在大型数据集上比同类的方法具有更好的性能. 主要贡献在于 3 个方面：1) 将流形排序方法用于度量用户社会关系的亲近程度，并将度量结果用户社会化推荐系统；2) 指出了 SR2 模型在用户相似度度量上存在的不足，在结合 SR2 模型和流形排序的基础上，提出了融入全局社会相似度的 SMF-MR 推荐方法，实验表明，不论与同类的推荐方法还是 SR2 模型相比，SMF-MR 方法均提高了推荐系统的精度；3) 提出的 SMF-MR 社会化推荐方法通过同时使用社会网络和传统评分两种数据，可减轻数据稀疏性问题、冷启动问题对推荐系统的影响.

参考文献：

[1] Badrul S, George K, Joseph K, et al. Item-based collaborative filtering recommendation algorithms [C] // Proceedings of the 10th International Conference on World Wide Web. New York: [s. n.], 2001: 285-295.

[2] Mohsen J, Martin E. A matrix factorization technique with trust propagation for recommendation in social networks [C] // Proceedings of the 4th ACM Conference on Recommender Systems. 2010: 135-142.

[3] Ma Hao, Irwin K, Michael R L. Learning to recommend with social trust ensemble [C] // Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2009: 203-210.

[4] Ma Hao, Zhou Dengyong, Liu Chao, et al. Recommender systems with social regularization [C] // Proceedings of the 4th ACM International Conference on Web Search and Data Mining. 2011: 287-296.

[5] Thomas H. Latent semantic models for collaborative filtering [J]. ACM Transactions on Information Systems (TOIS), 2004, 22(1): 89-115.

[6] Ruslan S, Andriy M. Probabilistic matrix factorization [J]. Advances in Neural Information Processing Systems, 2008(20): 1257-1264.

[7] Zhou Dengyong, Jason W, Arthur G, et al. Ranking on data manifolds [J]. Advances in Neural Information Processing Systems, 2003(16): 169-176.