

联合正则化的矩阵分解推荐算法^{*}

吴 宾, 娄铮铮, 叶阳东

(郑州大学 信息工程学院, 河南 郑州 450052)

通讯作者: 叶阳东, E-mail: ieydye@zzu.edu.cn



摘 要: 推荐系统已成为一种解决信息过载和帮助用户决策的有效工具. 当前的研究表明, 结合社会关系的推荐模型能够提升推荐的性能. 然而, 已有的社会化推荐模型大都忽略了物品之间的关联关系对推荐性能的影响. 针对此问题, 提出一种度量物品之间关联程度的方法, 并将其用于获取物品之间的关联关系. 然后, 将关联关系与社会关系相结合, 提出一种基于联合正则化的矩阵分解推荐模型, 并证明了联合正则化是一种加权的原子范数. 最后, 根据提出的模型构建了一种推荐算法 CRMF. 在 4 个真实数据集上的实验结果表明: 与主流的推荐算法相比, 该算法不仅可以缓解用户的冷启动问题, 而且更能有效地预测不同类型用户的实际评分.

关键词: 矩阵分解; 联合正则化; 推荐系统; 协同过滤; 社交网络

中图法分类号: TP18

中文引用格式: 吴宾, 娄铮铮, 叶阳东. 联合正则化的矩阵分解推荐算法. 软件学报, 2018, 29(9): 2681–2696. <http://www.jos.org.cn/1000-9825/5274.htm>

英文引用格式: Wu B, Lou ZZ, Ye YD. Co-Regularized matrix factorization recommendation algorithm. Ruan Jian Xue Bao/ Journal of Software, 2018, 29(9): 2681–2696 (in Chinese). <http://www.jos.org.cn/1000-9825/5274.htm>

Co-Regularized Matrix Factorization Recommendation Algorithm

WU Bin, LOU Zheng-Zheng, YE Yang-Dong

(School of Information Engineering, Zhengzhou University, Zhengzhou 450052, China)

Abstract: Recommender systems have been successfully adopted as an effective tool to alleviate information overload and assist users to make decisions. Recently, it has been demonstrated that incorporating social relationships into recommender models can enhance recommendation performance. Despite its remarkable progress, a majority of social recommendation models have overlooked the item relations—a key factor that can also significantly influence recommendation performance. In this paper, a approach is first proposed to acquire item relations by measuring correlations among items. Then, a co-regularized recommendation model is put forward to integrate the item relations with social relationships by introducing co-regularization term in the matrix factorization model. Meanwhile, that the co-regularization term is a case of weighted atomic norm is illustrated. Finally, based on the proposed model a recommendation algorithm named CRMF is constructed. CRMF is compared with existing state-of-the-art recommendation algorithms based on the evaluations over four real-world data sets. The experimental results demonstrate that CRMF is able to not only effectively alleviate the user cold-start problem, but also help obtain more accurate rating predictions of various users.

Key words: matrix factorization; co-regularization; recommender system; collaborative filtering; social network

随着信息技术和互联网的发展, 人们逐渐从信息匮乏的时代走入了信息过载的时代. 如何快速而精准地从浩瀚的数据海洋中帮助用户获取感兴趣的信息, 已成为亟需解决的问题. 推荐系统作为一种有效的信息过滤技术, 是解决这一问题的重要手段^[1–4]. 推荐系统通过收集和分析用户的历史行为来学习用户的喜好, 从而向用户

* 基金项目: 国家自然科学基金(61502434, 61772475, 61672469)

Foundation item: National Natural Science Foundation of China (61502434, 61772475, 61672469)

收稿时间: 2016-09-23; 修改时间: 2016-11-29; 采用时间: 2017-03-05; jos 在线出版时间: 2017-03-31

CNKI 网络优先出版: 2017-03-31 21:54:46, <http://kns.cnki.net/kcms/detail/11.2560.TP.20170331.2154.008.html>

推荐可能感兴趣的信息和服务.目前,推荐系统已在很多领域得到了应用,并为企业带来了可观的商业价值^[5-8].

在推荐系统领域,矩阵分解模型因具有较好的理论基础^[9-13]、良好的扩展性^[14-16]等优点,受到了学术界和工业界的广泛关注^[16-19].传统的矩阵分解模型通过分解用户-物品的评分矩阵来获得用户和物品的潜在特征矩阵,然后根据两者的乘积预测目标用户对特定物品的喜好程度.Salakhutdinov 等人从概率角度解释了传统矩阵分解模型的合理性,并提出概率矩阵分解模型(probabilistic matrix factorization,简称 PMF)^[20]和贝叶斯概率矩阵分解模型(Bayesian probabilistic matrix factorization,简称 BPMF)^[21],使得矩阵分解模型可高效地处理大规模数据,并取得较理想的推荐效果.在 Netflix 竞赛期间,Koren 等人^[22]在 PMF 的基础上,综合用户的偏置、物品的偏置及用户的隐式反馈信息,提出 SVD++ 算法,从而将矩阵分解模型在推荐系统中的应用推向了一个新的高度.

社交网络的普及,为物品推荐应用中用户之间社会关系的获取提供了便利条件,这种关系丰富了单个用户的信息,从而能够更好地对用户建模.如何将用户之间的社会关系融入传统矩阵分解模型中,是当前推荐系统研究的热点之一.研究者们^[23-28]以此为出发点,提出多种社会化推荐算法,并取得了良好的效果.目前,社会关系与传统矩阵分解模型结合的方式主要有两种^[29].

- 其一,通过同时分解社会关系矩阵和评分矩阵并共享用户潜在特征矩阵的方式来对用户和物品进行建模^[23-25].例如:Ma 等人^[23]将用户的社会关系信息与评分信息相结合,提出 SoRec 算法;Guo 等人^[25]将用户之间的信任关系强度融入 PMF,提出信任关系强度敏感的推荐算法 StrengthMF;
- 其二,将社会关系以正则化的方式来约束用户潜在特征矩阵的学习^[26-28].例如,Ma 等人^[26]使用用户的社会关系信息设计了社会正则化,并将其用于约束 PMF 的目标函数,提出 SoReg 算法;Jamali 等人^[27]将社交网络中的信任传播结构与 PMF 结合,提出 SocialMF 算法,该算法考虑了直接和间接信任的用户对目标用户评分的影响.

在物品推荐应用中,除用户之间的社会关系外,物品之间也存在一定的关联关系,例如手机与手机配件之间的关联关系、衣服和鞋子之间的关联关系.物品之间的这种关联关系通常也会影响用户是否购买某一物品,如何有效地获取物品之间的关联关系并将其应用到推荐系统中,仍是当前研究的难点问题,主要体现在:

- (1) 物品之间关联关系是隐式的:用户之间的社会关系可通过社交网络直接获取,而物品之间的关联关系却是隐式的,无直接获取的途径;
- (2) 物品之间的关联关系应是有向的:向购买手机的用户推荐手机配件成功的概率通常会比较高,反向推荐则不一定成功.因此在获取物品之间关联关系时,考虑有向性更为合理;
- (3) 物品之间相关程度计算方法的合理性:在计算物品之间的相关程度时,既要从小观的角度考虑两个物品个体之间的相关程度,又要从全局角度分析两个物品相关程度的强弱;
- (4) 如何将有向的关联关系融入到推荐模型中:在物品推荐过程中,考虑的因素越多,推荐模型越复杂,求解空间越大,因此需要一种有效的机制将关联关系融入到推荐模型中.

在已有的工作中:Sun 等人^[30]通过评分的时序信息来建立用户-物品的消费网络图,并构建用户之间和物品之间的近邻关系,然后将两者融入 PMF 中,提出基于时序行为的协同过滤推荐算法;Wu 等人^[31]通过标签信息来构造用户和物品的相似性关系,并将两者融入 PMF 中,提出基于近邻的概率矩阵分解模型;Gu 等人^[32]通过使用物品的属性信息来获取物品的相似性关系,并将其与社会关系相结合,以图正则化的方式融入非负矩阵分解中,提出基于图正则化权重的非负矩阵分解模型;Liang 等人^[35]在解决推荐系统的物品排序任务时,通过物品之间共现矩阵的点对互信息来计算物品之间的相关程度,并与权重矩阵分解模型相结合,提出联合分解模型 CoFactor.这些工作通过不同的途径来获取物品之间的关联关系并将其应用到推荐领域中,在一定程度上虽提高了传统推荐算法的性能,但仍存在一定的问题.例如:文献[30-34]需借助外部信息(评分时间信息^[30]、标签信息^[31]、物品属性信息^[32-34])来获取物品之间的关联关系,然而现实应用中,由于物品的外部信息通常不易大量获取,这在一定程度上限制了算法的应用范围;文献[31-37]获取的关联关系是无向的,无法更好地反映物品之间真实的关联关系.

针对物品之间关联关系获取这一难点问题,本文在不借助外部信息的前提下,综合物品之间的支持度与置

信度,提出一种从全局-局部角度计算物品之间相关程度的方法,且由此所得关联关系是有向的.进一步,综合考虑物品的关联关系、用户的社会关系,将它们融入到传统矩阵分解模型中,提出一种基于联合正则化的矩阵分解模型,并给出相应的推荐算法 CRMF(co-regularized matrix factorization).本文主要贡献有:

- (1) 给出一种仅从用户-物品的交互矩阵中获取物品之间关联程度的计算方法,它通过支持度与置信度,分别从全局和局部角度来考虑物品之间的相关程度,且由此所获取物品之间的关联关系具有有向性;
- (2) 提出一种基于联合正则化的矩阵分解模型,证明了联合正则化是一种加权原子范数,理论上保证了该模型更加低秩,从而使得在预测评分时采用较少的模型参数便可获得较为理想的推荐效果;
- (3) 联合正则化矩阵分解模型因考虑自适应正则化,使得构建的 CRMF 推荐算法应用于大规模非平衡数据集时更易搜索到合适的正则化超参数.

在 4 个真实数据集上的实验表明:CRMF 与主流的推荐算法相比,不仅可以缓解用户的冷启动问题,而且能够更有效地预测不同类型用户的实际评分.

1 相关知识

在本文中,我们使用 Euclid Math One 字体 \mathcal{P} 表示集合,大写字母 \mathbf{P} 表示矩阵, \mathbf{P}_{ij} 表示矩阵 \mathbf{P} 的第 i 行第 j 列的值, \mathbf{p}_i 表示矩阵 \mathbf{P} 的第 i 行向量, \mathbf{p}_i^j 表示向量 \mathbf{p}_i 的第 j 个值.本文使用的有关向量和矩阵的范数表示如下:

- $\|\mathbf{p}_i\|_2 = \sqrt{\sum_{j=1}^d \mathbf{p}_i^j \mathbf{p}_i^j}$, 称为向量的 2 范数;
- $\|\mathbf{P}\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^d \mathbf{P}_{ij}^2}$, 称为矩阵的 F 范数.

此外,我们使用 \mathbf{P}^T 表示矩阵的转置, $\text{tr}(\mathbf{P})$ 表示矩阵的迹, $\mathbf{I} = \text{diag}(\mathbf{I}_{11}, \mathbf{I}_{22}, \dots, \mathbf{I}_{nn})$ 表示一个对角矩阵, $|\mathcal{P}|$ 表示集合 \mathcal{P} 中的数量.

1.1 传统的矩阵分解

传统矩阵分解模型的基本思想^[14,20]:假设在推荐系统中存在 m 个用户和 n 个物品,对于给定的评分矩阵 $\mathbf{R} \in \mathbb{R}^{m \times n}$,寻求两个潜在特征矩阵 \mathbf{P} 和 \mathbf{Q} ,使得两者乘积近似拟合 \mathbf{R} .使用随机梯度下降(stochastic gradient descent, 简称 SGD)或交替最小二乘法(alternating least squares, 简称 ALS),可求出公式(1)的局部最优解 \mathbf{P} 和 \mathbf{Q} :

$$f = \frac{1}{2} \sum_{u=1}^m \sum_{j \in \mathcal{P}_u} (\hat{\mathbf{R}}_{uj} - \mathbf{R}_{uj})^2 + \frac{\lambda}{2} (\|\mathbf{P}\|_F^2 + \|\mathbf{Q}\|_F^2) \quad (1)$$

其中, $\mathbf{P} \in \mathbb{R}^{m \times d}$ 表示用户的潜在特征矩阵, \mathbf{p}_u 表示用户 u 的潜在特征向量, $\mathbf{Q} \in \mathbb{R}^{n \times d}$ 表示物品的潜在特征矩阵, \mathbf{q}_j 表示物品 j 的潜在特征向量, $d < \min(m, n)$ 表示特征向量维度.公式(1)中, $\mathbf{R}_{uj} > 0$ 表示用户 u 对物品 j 的实际评分值, $\hat{\mathbf{R}}_{uj} = \mathbf{p}_u^T \mathbf{q}_j$ 表示用户 u 对物品 j 的预测评分值, \mathcal{P}_u 表示用户 u 评分过的物品集合.超参数 λ 用于控制正则化的程度.传统的矩阵分解模型仅考虑了评分信息对用户和物品的潜在特征矩阵进行估计,而忽略了与用户密切相关的社会关系信息,有效地结合这些信息可提高推荐的精度.

1.2 社会正则化的矩阵分解

在社交网络中,社会相关性理论指出,人们比较喜欢与自己兴趣相投的用户建立连接,因此,目标用户决策时容易受到信任用户的影响,且用户之间的信任度越高,受影响程度越大.Ma 等人^[26]认为,社会关系是用户之间兴趣相似的先验知识,并提出用于约束传统矩阵分解目标函数的社会正则化,如公式(2)所示:

$$\frac{1}{2} \sum_{u=1}^m \sum_{v \in \mathcal{S}_u^+} S_{uv} \|\mathbf{p}_u - \mathbf{p}_v\|_2^2 = \text{tr}[\mathbf{P}^T (\mathbf{D}^p - \mathbf{S}) \mathbf{P}] = \text{tr}(\mathbf{P}^T \mathbf{L}^p \mathbf{P}) \quad (2)$$

其中,

- \mathbf{S} 表示社会关系矩阵;

- S_{uv} 表示用户 u 与用户 v 的相似度,由公式(3)计算得到:

$$S_{uv} = \frac{\sum_{j \in (\mathcal{P}_u \cap \mathcal{P}_v)} (\mathbf{R}_{uj} - \bar{\mathbf{R}}_u)(\mathbf{R}_{vj} - \bar{\mathbf{R}}_v)}{\sqrt{\sum_{j \in (\mathcal{P}_u \cap \mathcal{P}_v)} (\mathbf{R}_{uj} - \bar{\mathbf{R}}_u)^2} \sqrt{\sum_{j \in (\mathcal{P}_u \cap \mathcal{P}_v)} (\mathbf{R}_{vj} - \bar{\mathbf{R}}_v)^2}} \quad (3)$$

- \mathcal{S}_u^+ 表示用户 u 拥有的社会关系集合;
- 矩阵 \mathbf{D}^p 是与用户相关的对角矩阵,且 $\mathbf{D}_{uu}^p = \sum_{v=1}^m S_{uv}$;
- 矩阵 $\mathbf{L}^p = \mathbf{D}^p - \mathbf{S}$ 是与用户相关的拉普拉斯矩阵;
- $\mathcal{P}_u, \mathcal{P}_v$ 分别表示训练集中用户 u, v 的评分集合;
- $\bar{\mathbf{R}}_u, \bar{\mathbf{R}}_v$ 分别表示训练集中用户 u, v 的平均评分.

公式(2)与公式(1)结合,可得社会正则化的矩阵分解模型的目标函数,如公式(4)所示:

$$f = \frac{1}{2} \sum_{u=1}^m \sum_{j \in \mathcal{P}_u} (\hat{\mathbf{R}}_{uj} - \mathbf{R}_{uj})^2 + \frac{\lambda}{2} (\|\mathbf{P}\|_F^2 + \|\mathbf{Q}\|_F^2) + \frac{\beta}{2} \text{tr}(\mathbf{P}^T \mathbf{L}^p \mathbf{P}) \quad (4)$$

其中, β 控制着社会关系在用户决策时所占的比重.上述目标函数通过社会正则化的方式约束用户潜在特征向量的学习,使得当目标用户为冷启动用户时,其潜在特征向量更多地表现为社交网络中近邻用户的特征因子.但该模型潜在假设物品之间是相互独立的,忽略了现实世界中物品之间普遍存在着关联关系的事实.

2 联合正则化的矩阵分解

2.1 物品之间关联关系的获取

物品之间的关联关系在推荐应用中是影响用户决策的重要因素之一,如何从用户-物品之间的交互矩阵中有效地量化物品之间的关联程度并获取它们之间有向的关联关系,是推荐过程中核心步骤之一.文献[34]指出,数据挖掘中关联规则的支持度(Support)与置信度(Confidence)可在一定程度上反映物品 j 与物品 i 之间的相关程度,计算方法分别如公式(5)和公式(6)所示:

$$\text{Support}(j, i) = |\mathcal{Q}_i \cap \mathcal{Q}_j| / |\mathcal{R}^r| \quad (5)$$

$$\text{Confidence}(j \rightarrow i) = |\mathcal{Q}_i \cap \mathcal{Q}_j| / |\mathcal{Q}_j| \quad (6)$$

其中, $|\mathcal{R}^r|$ 表示在用户-物品交互矩阵中对物品做过评价的总人数,即评分矩阵 \mathbf{R} 中非零元素个数; $\mathcal{Q}_i, \mathcal{Q}_j$ 分别表示对物品 i 和 j 做过评价的用户集合.从公式(5)可得出,支持度从全局的角度反映了物品之间的相关程度,由此获得的关联关系是无向的;从公式(6)可得出,置信度则从局部的两个物品来计算物品之间的相关程度,由此获得的关联关系是有向的.如果单独考虑两者中的任意一个作为物品之间相关程度度量方法,则无法兼顾物品间关联关系的有向性及全局-局部特性.此外,若以数据挖掘中的关联规则来获取物品之间的关联关系,通常需要支持度和置信度大于相应的阈值时才有效^[34].然而在推荐应用中,支持度阈值和置信度阈值往往不易选取,过高的阈值将会减少关联关系的数量,过低的阈值会导致考虑较多的关联关系,使得模型训练需要较多的时间.

为解决上述获取物品之间关联关系的不足之处,在综合考虑物品之间支持度与置信度的基础之上,本文定义了物品之间关联程度的计算方法,如公式(7)所示:

$$\mathbf{W}_{ji} = \frac{\text{Support}(j, i)}{\text{Support}(j, i) + \varphi} \times \text{Confidence}(j \rightarrow i) = \frac{|\mathcal{Q}_i \cap \mathcal{Q}_j|}{|\mathcal{Q}_i \cap \mathcal{Q}_j| + \rho} \times \frac{|\mathcal{Q}_i \cap \mathcal{Q}_j|}{|\mathcal{Q}_j|} \quad (7)$$

其中, \mathbf{W} 表示物品之间的关联关系矩阵, \mathbf{W}_{ji} 表示物品 j 与物品 i 之间的关联程度, $\varphi \geq 0$ 是一个适用于所有物品的超参数, $\rho = \varphi |\mathcal{R}^r|$.在公式(7)中,第2项 $\text{Confidence}(j \rightarrow i)$ 从局部的角度考虑了两个物品个体之间的相关程度,第1项融合 $\text{Support}(j, i)$, 用于从全局角度调节两个物品之间的相关程度.当超参数 $\varphi=0$ (即 $\rho=0$) 时, $\mathbf{W}_{ji} = \text{Confidence}(j \rightarrow i)$, 此时,两个物品之间的关联程度等于关联规则中的置信度.当超参数 φ 为大于0的固定值时,若反映物品 j 与物品 i 之间全局相关程度的 $\text{Support}(j, i)$ 越大,则第1项的值也越大;若 $\text{Support}(j, i)$ 的值很小,即便物

品 j 到物品 i 的 $Confidence(j \rightarrow i)$ 比较高,则也可通过第 1 项相应地降低它们之间的关联程度 \mathbf{W}_{ji} 的值.

由公式(7)可知:在不借助外部信息的前提下,仅从用户-物品之间的交互矩阵中便可获取物品之间的关联关系.由于 $Confidence(i \rightarrow j) \neq Confidence(j \rightarrow i)$,则 $\mathbf{W}_{ij} \neq \mathbf{W}_{ji}$,因此,本文所定义的公式(7)可在一定程度上反映物品之间关联关系的有向性.

在下文中,使用 \mathcal{M} 表示物品之间的关联关系集合, \mathcal{M}_j^+ 表示物品 j 相关联的物品集合.在选择物品 j 的关联关系时,首先根据 \mathbf{W}_{ji} 对关联程度进行排序,然后对目标物品 j 选择关联程度最大的 σ 个物品作为近邻集 \mathcal{M}_j^+ ,进一步简化关联关系集合 \mathcal{M} .

2.2 模型定义

获取到物品的关联关系之后,我们可将其应用到已有的矩阵分解模型.此时,物品的潜在特征向量将受到存在关联关系的物品的影响.与社会正则化相似,由物品的关联关系可得出关联正则化,如公式(8)所示:

$$\frac{1}{2} \sum_{j=1}^n \sum_{i \in \mathcal{M}_j^+} \mathbf{W}_{ji} \|\mathbf{q}_j - \mathbf{q}_i\|_2^2 = \text{tr}[\mathbf{Q}^T (\mathbf{D}^q - \mathbf{W}) \mathbf{Q}] = \text{tr}(\mathbf{Q}^T \mathbf{L}^q \mathbf{Q}) \quad (8)$$

其中,矩阵 \mathbf{D}^q 是与物品相关的对角矩阵且 $\mathbf{D}_{jj}^q = \sum_{i=1}^n \mathbf{W}_{ji}$, 矩阵 $\mathbf{L}^q = \mathbf{D}^q - \mathbf{W}$ 是与物品相关的拉普拉斯矩阵.我们将关联正则化融入基于社会正则化的矩阵分解模型可得改进后的目标函数如公式(9)所示:

$$f = \frac{1}{2} \sum_{u=1}^m \sum_{j \in \mathcal{P}_u} (\hat{\mathbf{R}}_{uj} - \mathbf{R}_{uj})^2 + \frac{\lambda}{2} (\|\mathbf{P}\|_F^2 + \|\mathbf{Q}\|_F^2) + \frac{\beta}{2} \text{tr}(\mathbf{P}^T \mathbf{L}^p \mathbf{P}) + \frac{\alpha}{2} \text{tr}(\mathbf{Q}^T \mathbf{L}^q \mathbf{Q}) \quad (9)$$

其中,超参数 $\alpha > 0$ 用于控制物品的关联关系在推荐过程中所占的比重.

上述目标函数在求解模型参数时,用于避免过拟合的正则化对于每个用户(物品)使用相同的超参数,使得学习与系统交互频繁的用户(物品)的潜在特征向量时可能出现欠拟合,在学习与系统交互较少的用户(物品)的潜在特征向量时可能出现过拟合.Salakhutdinov 等人^[39]认为:在学习用户(物品)的潜在特征向量时,根据用户(物品)的评分数量进行不同程度的惩罚会比较合理,由此可知需考虑自适应正则化.由文献[39]分析可知,矩阵分解模型结合本文的自适应正则化后有如下优点:其一,评分的分布可为非均匀分布;其二,只需较少的评分样本即可获得较高的评分预测精度;其三,对于大规模非平衡数据集,更易搜索到合适的正则化参数.我们将其与公式(9)结合,可得基于联合正则化的矩阵分解的目标函数,如公式(10)所示:

$$f = \underbrace{\frac{\lambda}{2} \left(\sum_{u=1}^m |\mathcal{P}_u|^{-\frac{1}{2}} \|\mathbf{p}_u\|_2^2 + \sum_{j=1}^n |\mathcal{Q}_j|^{-\frac{1}{2}} \|\mathbf{q}_j\|_2^2 \right)}_{\text{自适应正则化}} + \frac{1}{2} \sum_{u=1}^m \sum_{j \in \mathcal{P}_u} (\hat{\mathbf{R}}_{uj} - \mathbf{R}_{uj})^2 + \frac{\beta}{2} \text{tr}(\mathbf{P}^T \mathbf{L}^p \mathbf{P}) + \frac{\alpha}{2} \text{tr}(\mathbf{Q}^T \mathbf{L}^q \mathbf{Q}) \quad (10)$$

$$= \frac{1}{2} \sum_{u=1}^m \sum_{j \in \mathcal{P}_u} (\hat{\mathbf{R}}_{uj} - \mathbf{R}_{uj})^2 + \underbrace{\frac{1}{2} \text{tr}(\mathbf{P}^T \tilde{\mathbf{L}}^p \mathbf{P}) + \frac{1}{2} \text{tr}(\mathbf{Q}^T \tilde{\mathbf{L}}^q \mathbf{Q})}_{\text{联合正则化}}$$

其中, $\tilde{\mathbf{L}}^p = \lambda \mathbf{I}^p + \beta \mathbf{L}^p$, $\mathbf{I}^p = \text{diag} \left(|\mathcal{P}_1|^{-\frac{1}{2}}, \dots, |\mathcal{P}_u|^{-\frac{1}{2}}, \dots, |\mathcal{P}_m|^{-\frac{1}{2}} \right)$, $\mathbf{I}^q = \text{diag} \left(|\mathcal{Q}_1|^{-\frac{1}{2}}, \dots, |\mathcal{Q}_j|^{-\frac{1}{2}}, \dots, |\mathcal{Q}_n|^{-\frac{1}{2}} \right)$, $\tilde{\mathbf{L}}^q = \lambda \mathbf{I}^q + \alpha \mathbf{L}^q$.

为了说明联合正则化的有效性,我们以下给出详细的理论分析.根据文献[40]可知:核范数是基于原子集 $\mathcal{F}_* = \{\mathbf{x}_i \mathbf{z}_i^T : \|\mathbf{x}_i\|_2 = \|\mathbf{z}_i\|_2 = 1\}$ 的 *gauge* 函数, \mathcal{F}_* 中每个秩为 1 的矩阵都有单位 F 范数.假设 $\mathbf{U} \in \mathbb{R}^{m \times m}$ 是由一组基向量组成的矩阵, $\mathbf{S}_U^{-1/2}$ 是一个对角矩阵且 $(\mathbf{S}_U^{-1/2})_{ii} \geq 0$ 表示第 i 个基向量所占的权重.

同样可假设 $\mathbf{V} \in \mathbb{R}^{n \times n}$, $(\mathbf{S}_V^{-1/2})_{ii} \geq 0$.令 $\mathbf{A} = \mathbf{U} \mathbf{S}_U^{-1/2}$, $\mathbf{B} = \mathbf{V} \mathbf{S}_V^{-1/2}$, 由此可得加权的原子集如公式(11)所示:

$$\mathcal{F} = \{a_i = \mathbf{p}_i \mathbf{q}_i^T : \mathbf{p}_i = \mathbf{A} \mathbf{x}_i, \mathbf{q}_i = \mathbf{B} \mathbf{z}_i, \|\mathbf{x}_i\|_2 = \|\mathbf{z}_i\|_2 = 1\} \quad (11)$$

其中,加权原子集 \mathcal{F} 中每个原子 a 都有非单位 F 范数.基于加权的原子集,可定义加权原子范数如公式(12)所示:

$$\|\hat{\mathbf{R}}\|_{\mathcal{F}} = \inf \sum_{a_i \in \mathcal{F}} |c_i|, \quad \text{s.t.} \quad \hat{\mathbf{R}} = \sum_{a_i \in \mathcal{F}} c_i a_i \quad (12)$$

定理 1. 对于任意的 $\mathbf{A} = \mathbf{U} \mathbf{S}_U^{-1/2}$, $\mathbf{B} = \mathbf{V} \mathbf{S}_V^{-1/2}$ 和加权的原子集 \mathcal{F} , 可得出:

$$\|\hat{\mathbf{R}}\|_{\mathcal{F}} = \inf_{\mathbf{P}, \mathbf{Q}} \frac{1}{2} \{ \|\mathbf{A}^{-1}\mathbf{P}\|_F^2 + \|\mathbf{B}^{-1}\mathbf{Q}\|_F^2 \}, \text{ s.t. } \hat{\mathbf{R}} = \mathbf{P}\mathbf{Q}^T \quad (13)$$

证明:对于满足 $\|\hat{\mathbf{R}}\|_{\mathcal{F}} = \sum_{a_i \in \mathcal{F}} |c_i|$, $a_i = \mathbf{A}\mathbf{x}_i \mathbf{z}_i^T \mathbf{B}^T$ 的所有 $\hat{\mathbf{R}} = \sum_{a_i \in \mathcal{F}} c_i a_i$, 我们可以构造 \mathbf{P} 和 \mathbf{Q} 矩阵的第 i 行分别为

$$\mathbf{p}_i = \sqrt{|c_i|} \mathbf{A}\mathbf{x}_i, \mathbf{q}_i = \sqrt{|c_i|} \mathbf{B}\mathbf{z}_i, \|\mathbf{A}^{-1}\mathbf{P}\|_F^2 = \sum_i \|\mathbf{A}^{-1}\mathbf{p}_i\|_2^2 = \sum_i \|\mathbf{A}^{-1}\sqrt{|c_i|} \mathbf{A}\mathbf{x}_i\|_2^2 = \sum_i \sqrt{|c_i|} \|\mathbf{A}^{-1}\mathbf{A}\mathbf{x}_i\|_2^2 = \sum_i |c_i|.$$

相似地,我们可得出 $\|\mathbf{B}^{-1}\mathbf{Q}\|_F^2 = \sum_i |c_i|$.

由 $\|\hat{\mathbf{R}}\|_{\mathcal{F}} = \sum_{a_i \in \mathcal{F}} |c_i|$ 可知, $\|\hat{\mathbf{R}}\|_{\mathcal{F}} > \frac{1}{2} (\|\mathbf{A}^{-1}\mathbf{P}\|_F^2 + \|\mathbf{B}^{-1}\mathbf{Q}\|_F^2)$ 不成立.

由 $\hat{\mathbf{R}} = \mathbf{P}\mathbf{Q}^T$ 可构造 $\mathbf{x}_i = \mathbf{A}^{-1}\mathbf{p}_i / \|\mathbf{A}^{-1}\mathbf{p}_i\|_2, \mathbf{z}_i = \mathbf{B}^{-1}\mathbf{q}_i / \|\mathbf{B}^{-1}\mathbf{q}_i\|_2, c_i = \|\mathbf{A}^{-1}\mathbf{p}_i\|_2 \|\mathbf{B}^{-1}\mathbf{q}_i\|_2$.

由此可得 $\mathbf{p}_i \mathbf{q}_i^T = c_i \mathbf{A}\mathbf{x}_i \mathbf{z}_i^T \mathbf{B}^T, \hat{\mathbf{R}} = \sum_i c_i \mathbf{A}\mathbf{x}_i \mathbf{z}_i^T \mathbf{B}^T$.

由均值不等式可得出 $|c_i| = \|\mathbf{A}^{-1}\mathbf{p}_i\|_2 \|\mathbf{B}^{-1}\mathbf{q}_i\|_2 \leq \frac{1}{2} (\|\mathbf{A}^{-1}\mathbf{p}_i\|_2^2 + \|\mathbf{B}^{-1}\mathbf{q}_i\|_2^2)$.

进一步得出 $\sum_i |c_i| \leq \frac{1}{2} (\sum_i \|\mathbf{A}^{-1}\mathbf{p}_i\|_2^2 + \sum_i \|\mathbf{B}^{-1}\mathbf{q}_i\|_2^2)$.

因此可得 $\sum_i |c_i| \leq \frac{1}{2} (\sum_i \|\mathbf{A}^{-1}\mathbf{p}_i\|_F^2 + \sum_i \|\mathbf{B}^{-1}\mathbf{q}_i\|_F^2)$.

又因 $\|\hat{\mathbf{R}}\|_{\mathcal{F}} = \inf_{a_i \in \mathcal{F}} \sum |c_i|$, 最终可得定理成立.

推论 1. 联合正则化是一种加权的原子范数.

证明:由上述定理可得 $\|\mathbf{A}^{-1}\mathbf{P}\|_F^2 = \text{tr}[\mathbf{P}^T (\mathbf{A}^{-1})^T \mathbf{A}^{-1} \mathbf{P}]$, $\|\mathbf{B}^{-1}\mathbf{Q}\|_F^2 = \text{tr}[\mathbf{Q}^T (\mathbf{B}^{-1})^T \mathbf{B}^{-1} \mathbf{Q}]$.

对公式(10)中的矩阵 $\tilde{\mathbf{L}}^p$ 和 $\tilde{\mathbf{L}}^q$ 分别进行特征分解,可得 $\tilde{\mathbf{L}}^p = \mathbf{U}_p \mathbf{S}_p \mathbf{U}_p^T, \tilde{\mathbf{L}}^q = \mathbf{U}_q \mathbf{S}_q \mathbf{U}_q^T$.

将其带入 $\text{tr}(\mathbf{P}^T \tilde{\mathbf{L}}^p \mathbf{P}), \text{tr}(\mathbf{Q}^T \tilde{\mathbf{L}}^q \mathbf{Q})$, 可得 $\text{tr}(\mathbf{P}^T \tilde{\mathbf{L}}^p \mathbf{P}) = \text{tr}(\mathbf{P}^T \mathbf{U}_p \mathbf{S}_p \mathbf{U}_p^T \mathbf{P}), \text{tr}(\mathbf{Q}^T \tilde{\mathbf{L}}^q \mathbf{Q}) = \text{tr}(\mathbf{Q}^T \mathbf{U}_q \mathbf{S}_q \mathbf{U}_q^T \mathbf{Q})$.

令 $\mathbf{A} = \mathbf{U}_p \mathbf{S}_p^{-1/2}, \mathbf{B} = \mathbf{U}_q \mathbf{S}_q^{-1/2}$, 将其分别带入 $\|\mathbf{A}^{-1}\mathbf{P}\|_F^2, \|\mathbf{B}^{-1}\mathbf{Q}\|_F^2$, 可得:

$$\|\mathbf{A}^{-1}\mathbf{P}\|_F^2 = \text{tr}(\mathbf{P}^T \mathbf{U}_p \mathbf{S}_p \mathbf{U}_p^T \mathbf{P}), \|\mathbf{B}^{-1}\mathbf{Q}\|_F^2 = \text{tr}(\mathbf{Q}^T \mathbf{U}_q \mathbf{S}_q \mathbf{U}_q^T \mathbf{Q}).$$

因此可得出 $\text{tr}(\mathbf{P}^T \tilde{\mathbf{L}}^p \mathbf{P}) = \|\mathbf{A}^{-1}\mathbf{P}\|_F^2, \text{tr}(\mathbf{Q}^T \tilde{\mathbf{L}}^q \mathbf{Q}) = \|\mathbf{B}^{-1}\mathbf{Q}\|_F^2$.

根据定理 1 可得出,本文所提出的联合正则化是一种加权原子范数.

在矩阵分解求解模型参数时,根据文献[41,42]可知:考虑加权原子范数,能使得模型更加低秩且能获得较高的预测精度.因此,证明联合正则化是一种加权原子范数,从理论上能够说明考虑联合正则化的矩阵分解在预测评分时采用较少的模型参数便可获得较为理想的推荐效果.

文献[22]的研究表明,考虑用户的偏置、物品的偏置以及用户的隐式反馈信息可提高推荐算法的精度.我们将其与基于联合正则化的矩阵分解模型结合,可得改进后的目标函数,如公式(14)所示:

$$f = \frac{1}{2} \sum_{u=1}^m \sum_{j \in \mathcal{P}_u} (\hat{\mathbf{R}}_{uj} - \mathbf{R}_{uj})^2 + \frac{1}{2} \{ \text{tr}(\mathbf{P}^T \tilde{\mathbf{L}}^p \mathbf{P}) + \text{tr}(\mathbf{Q}^T \tilde{\mathbf{L}}^q \mathbf{Q}) \} + \frac{\lambda}{2} \left(\sum_{u=1}^m |\mathcal{P}_u|^{\frac{1}{2}} \mathbf{b}_p^u \mathbf{b}_p^u + \sum_{j=1}^n |\mathcal{Q}_j|^{\frac{1}{2}} \mathbf{b}_q^j \mathbf{b}_q^j + \sum_{k \in \mathcal{P}_u} |\mathcal{Q}_k|^{\frac{1}{2}} \|\mathbf{y}_k\|_2^2 \right) \quad (14)$$

其中, $\mathbf{b}_p, \mathbf{b}_q$ 分别表示用户和物品的偏置向量, $\mathbf{b}_p^u, \mathbf{b}_q^j$ 分别表示用户 u 和物品 j 的偏置项.

每个物品 k 关联的一个隐式反馈向量 $\mathbf{y}_k \in \mathbb{R}^d$ 用于建模隐式反馈信息对用户的潜在特征向量的影响,此时, $\hat{\mathbf{R}}_{uj}$ 由公式(15)可得:

$$\hat{\mathbf{R}}_{uj} = \mathbf{b}_p^u + \mathbf{b}_q^j + \mu + (\mathbf{p}_u + |\mathcal{P}_u|^{\frac{1}{2}} \sum_{k \in \mathcal{P}_u} \mathbf{y}_k)^T \mathbf{q}_j \quad (15)$$

其中, μ 表示全局平均评分.

图 1 详细展示了社交网络中基于联合正则化的矩阵分解推荐过程.

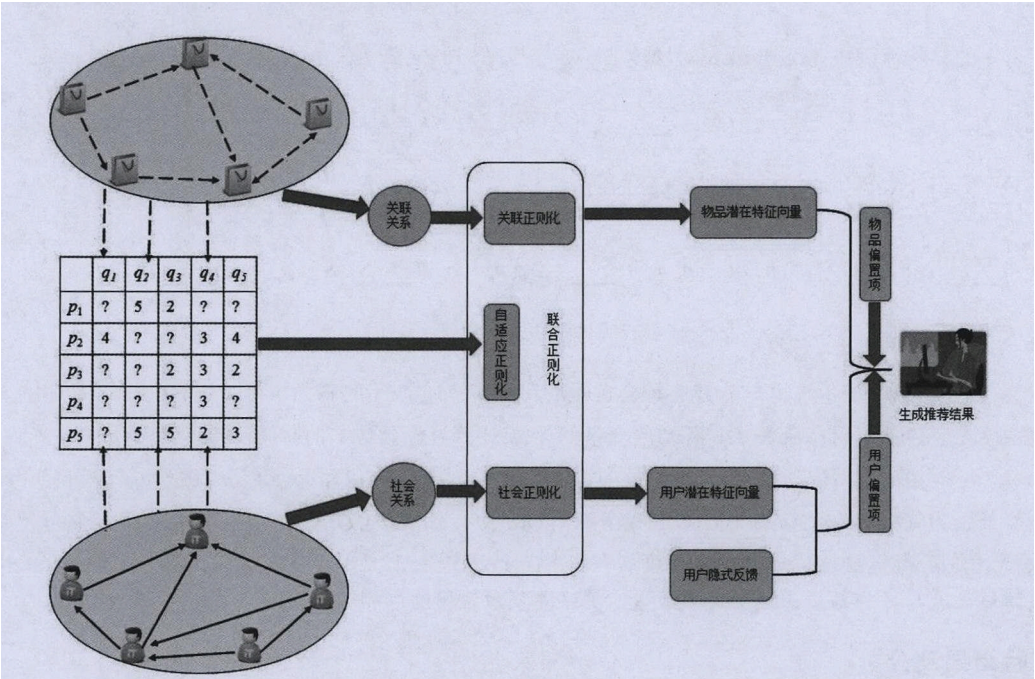


Fig.1 An illustration of co-regularized matrix factorization recommendation

图 1 基于联合正则化的矩阵分解推荐示意图

2.3 CRMF算法

本文使用 SGD 求解联合正则化的矩阵分解模型的参数,由此可得出 CRMF 算法的详细流程见表 1.

Table 1 CRMF algorithm
表 1 CRMF 算法

输入:社会关系集合 \mathcal{S} ;关联关系集合 \mathcal{W} ;训练集 \mathcal{R} ;潜在特征向量的维度 d ;	
超参数 α, β, λ ;学习速率 η ;最大迭代次数 $MaxIter$.	
输出: b_p, b_q, P, Q, Y .	
1	随机初始化 b_p, b_q, P, Q, Y
2	for $t \leftarrow 1$ to $MaxIter$ do
3	令 P', Q' 为临时变量
4	$\hat{R}_{ij} \leftarrow \text{Eq.}(14)$
5	$b_p^u \leftarrow b_p^u - \eta \cdot \partial f / \partial b_p^u$
6	$b_q^j \leftarrow b_q^j - \eta \cdot \partial f / \partial b_q^j$
7	$p'_u \leftarrow p'_u + \partial f_1 / \partial p_u$
8	$q'_j \leftarrow q'_j + \partial f_1 / \partial q_j$
9	for $k \in \mathcal{R}_t$ do
10	$y_k \leftarrow y_k - \eta \cdot \partial f / \partial y_k$
11	end
12	for $u \leftarrow 1$ to m do
13	$p'_u \leftarrow p'_u + \partial f_2 / \partial p_u$
14	end
15	for $j \leftarrow 1$ to n do
16	$q'_j \leftarrow q'_j + \partial f_2 / \partial q_j$
17	end
18	$P \leftarrow P - \eta P'$
19	$Q \leftarrow Q - \eta Q'$
20	$Y \leftarrow Y - \eta Y'$
21	end
22	return b_p, b_q, P, Q, Y

在表 1 的算法流程中,核心步骤是对梯度 $\partial f / \partial \mathbf{b}_p^u, \partial f / \partial \mathbf{b}_q^j, \partial f / \partial \mathbf{y}_k, \partial f_1 / \partial \mathbf{p}_u, \partial f_2 / \partial \mathbf{p}_u, \partial f_1 / \partial \mathbf{q}_j, \partial f_2 / \partial \mathbf{q}_j$ 的计算:

$$\begin{aligned} \partial f / \partial \mathbf{b}_p^u &= \sum_{j \in \mathcal{P}_u} e_{uj} + \lambda |\mathcal{P}_u|^{-\frac{1}{2}} \mathbf{b}_p^u, \partial f / \partial \mathbf{b}_q^j = \sum_{u \in \mathcal{Q}_j} e_{uj} + \lambda |\mathcal{Q}_j|^{-\frac{1}{2}} \mathbf{b}_q^j, \forall k \in \mathcal{P}_u, \partial f / \partial \mathbf{y}_k = \sum_{k \in \mathcal{P}_u} e_{uj} |\mathcal{P}_u|^{-\frac{1}{2}} \mathbf{q}_j + \lambda |\mathcal{Q}_k|^{-\frac{1}{2}} \mathbf{y}_k, \\ \text{令 } f_1 &= \frac{1}{2} \sum_{u=1}^m \sum_{j \in \mathcal{P}_u} e_{uj}^2 + \frac{\lambda}{2} \left(\sum_{u=1}^m |\mathcal{P}_u|^{-\frac{1}{2}} \|\mathbf{p}_u\|_2^2 + \sum_{j=1}^n |\mathcal{Q}_j|^{-\frac{1}{2}} \|\mathbf{q}_j\|_2^2 \right), f_2 = \frac{\beta}{2} \text{tr}(\mathbf{P}^T \mathbf{L}^p \mathbf{P}) + \frac{\alpha}{2} \text{tr}(\mathbf{Q}^T \mathbf{L}^q \mathbf{Q}), \\ \partial f_1 / \partial \mathbf{p}_u &= \sum_{j \in \mathcal{P}_u} e_{uj} \mathbf{q}_j + \lambda |\mathcal{P}_u|^{-\frac{1}{2}} \mathbf{p}_u, \partial f_2 / \partial \mathbf{p}_u = \frac{\beta}{2} (\sum_{v \in \mathcal{S}_u^+} S_{uv} (\mathbf{p}_u - \mathbf{p}_v) + \sum_{g \in \mathcal{S}_u^-} S_{ug} (\mathbf{p}_u - \mathbf{p}_g)), \\ \partial f_1 / \partial \mathbf{q}_j &= \sum_{u \in \mathcal{Q}_j} e_{uj} (\mathbf{p}_u + \sum_{k \in \mathcal{P}_u} \mathbf{y}_k) + \lambda |\mathcal{Q}_j|^{-\frac{1}{2}} \mathbf{q}_j, \partial f_2 / \partial \mathbf{q}_j = \frac{\alpha}{2} (\sum_{i \in \mathcal{W}_j^+} \mathbf{W}_{ji} (\mathbf{q}_j - \mathbf{q}_i) + \sum_{o \in \mathcal{W}_j^-} \mathbf{W}_{oj} (\mathbf{q}_j - \mathbf{q}_o)), \end{aligned}$$

其中, $e_{uj} = \hat{\mathbf{R}}_{uj} - \mathbf{R}_{uj}, \mathcal{S}_u^+$ 表示与用户 u 建立社会关系的用户集, \mathcal{W}_j^+ 表示与物品 j 相关联的物品集. CRMF 算法的时间复杂度主要包括对目标函数 f 和各梯度变量的计算. 计算目标函数 f 的时间复杂度为 $O(d|\mathcal{R}^r| + d|\mathcal{S}| + d|\mathcal{W}|)$, $|\mathcal{S}|$ 表示社会关系数量, $|\mathcal{W}|$ 表示关联关系数量. 计算梯度 $\partial f / \partial \mathbf{b}_p^u, \partial f / \partial \mathbf{b}_q^j, \partial f_1 / \partial \mathbf{p}_u, \partial f_2 / \partial \mathbf{p}_u, \partial f_1 / \partial \mathbf{q}_j, \partial f_2 / \partial \mathbf{q}_j, \partial f / \partial \mathbf{y}_k$ 的时间复杂度分别为 $O(d|\mathcal{R}^r|), O(d|\mathcal{R}^r|), O(d|\mathcal{R}^r|), O(d|\mathcal{S}|), O(d|\mathcal{R}^r|), O(d|\mathcal{W}|), O(d|\mathcal{R}^r| \bar{r})$, 其中, \bar{r} 表示物品平均被评分的数量. 由此可得算法迭代一次的复杂度 $O(d|\mathcal{R}^r| \bar{r} + d|\mathcal{S}| + d|\mathcal{W}|)$. 由于 $\bar{r} \ll \min(|\mathcal{R}^r|, |\mathcal{S}|, |\mathcal{W}|)$, 因此算法的整体复杂度与评分数量、社会关系数量、关联关系数量呈线性相关, 可用于处理大规模数据.

3 实验结果及分析

3.1 数据集及实验环境

为了对比不同信息对推荐结果的影响, 本文使用 4 个含有评分信息和社会关系的数据集对提出的算法进行验证, 分别为 Flixster(<http://www.cs.ubc.ca/~jamalim/datasets/>), Epinions(<http://www.trustlet.org/epinions.html>), Ciao(<http://www.cse.msu.edu/~tangjili/trust.html>), FilmTrust(<http://www.librec.net/datasets.html>). Flixster 是基于社交网络的电影评论网站, 用户可对电影做出 [0.5, 5.5] 范围内的评分, 步长是 0.5. Ciao 是物品评论网站, 用户可对物品做出 [1, 5] 范围内的评分. FilmTrust 是基于信任关系的电影推荐网站, 用户可根据自己的喜好对电影做出 [0.5, 4] 范围内的评分, 其步长是 0.5. Epinions 是物品评论数据集, 包含 40 163 个用户和 139 738 个物品上的 664 824 个评分, 该数据集非常接近实际应用的数据量, 因此它可用于考察模型的实际性能. 实验中所使用的 4 个数据集的统计特性见表 2.

Table 2 Statistics of the four datasets

表 2 4 个数据集的统计特性

信息	Epinions	Flixster	Ciao	FilmTrust
用户数量	40 163	53 213	17 615	1 508
物品数量	139 738	18 197	16 121	2 071
评分记录	664 824	409 803	72 665	35 497
社会关系	487 183	655 054	40 133	1 853

本文将数据集分成训练集和测试集, 训练集用于学习推荐算法中的参数, 测试集用于评估推荐算法的准确性. 在实验中, 按照 4:1 的比例将 4 个数据集随机地分成训练集和测试集. 本文的实验是在 Windows 7 操作系统, 64G 内存, Intel(R) Xeon(R) CPU E5-2620 0@2.00GHz 的机器上进行, 实验程序基于 Java 语言和集成开发工具 Eclipse4.5 实现.

3.2 评价指标

为了评估推荐算法的性能, 本文使用平均绝对误差(mean absolute error, 简称 MAE)和均方根误差(root mean squared error, 简称 RMSE)作为实验结果的评价指标, 两种评价指标的计算方式如公式(16)和公式(17). 通过计算预测评分与真实评分的差别来衡量推荐结果的准确性, 其值越小, 推荐准确性越高.

$$MAE = \frac{1}{|\mathcal{R}^{te}|} \sum_{(u,j) \in \mathcal{R}^{te}} |R_{ij} - \hat{R}_{ij}| \tag{16}$$

$$RMSE = \sqrt{\frac{1}{|\mathcal{R}^{te}|} \sum_{(u,j) \in \mathcal{R}^{te}} (R_{ij} - \hat{R}_{ij})^2} \tag{17}$$

其中, \mathcal{R}^{te} 表示测试集, $|\mathcal{R}^{te}|$ 表示测试集中评分的数量.

3.3 对比算法及超参数设定

在实验中,我们选择基于 Librec(<http://www.librec.net>)平台来实现并验证 CRMF 算法.为了验证社会关系和关联关系在推荐过程中所起的作用,本文选取了 6 种主流的推荐算法与 CRMF 进行了详细对比,并在表 3 中给出了 7 种推荐算法中的超参数在 4 个数据集上的最佳取值.此外,实验 2~实验 6 中设置潜在特征向量维度 $d=10$.

- (1) PMF:由 Salakhutdinov 等人^[20]提出的一种基于概率的矩阵分解推荐算法,该算法仅考虑了用户对物品的评分信息;
- (2) SoReg:由 Ma 等人^[26]提出的一种基于社会关系的推荐算法,该算法将社会关系以社会正则化的方式约束传统矩阵分解时用户潜在特征向量的学习,使得存在社会关系的两个用户其潜在特征向量尽可能的相似;
- (3) SoRec:由 Ma 等人^[23]提出的一种基于社会关系的推荐算法,该算法通过同时分解评分矩阵和用户的社会关系矩阵并且以共享用户潜在特征矩阵的方式,学习出用户和物品的潜在特征矩阵;
- (4) SocialMF:由 Jamali 等人^[27]提出的一种将社交网络中的信任传播结构与 PMF 相结合的推荐算法,该算法有效地考虑了直接和间接信任的用户对目标用户评分的影响;
- (5) TrustMF:由 Yang 等人^[24]提出的一种基于信任关系的推荐算法,该算法合理地反映了信任用户之间的相互关系对于推荐过程中评分预测的影响;
- (6) SVD++:由 Koren 等人^[22]提出的一种同时考虑用户偏置、物品偏置以及用户隐式反馈信息的推荐算法,该算法是矩阵分解中推荐精度较高的算法,并且是当年 Netflix 大赛获奖者所使用的关键算法.

Table 3 Hyper-Parameter settings of compared algorithms

表 3 对比算法的超参数设置

Algorithms	Epinions	Flixster	Ciao	FilmTrust
PMF	$\lambda_U=\lambda_V=0.01$	$\lambda_U=\lambda_V=0.1$	$\lambda_U=\lambda_V=0.1$	$\lambda_U=\lambda_V=0.1$
SoRec	$\lambda_U=\lambda_V=0.001, \lambda_C=1$	$\lambda_U=\lambda_V=0.001, \lambda_C=0.001$	$\lambda_U=\lambda_V=0.001, \lambda_C=0.01$	$\lambda_U=\lambda_V=0.001, \lambda_C=0.1$
SoReg	$\lambda_U=\lambda_V=0.001, \beta=0.1$	$\lambda_U=\lambda_V=0.001, \beta=0.01$	$\lambda_U=\lambda_V=0.001, \beta=0.01$	$\lambda_U=\lambda_V=0.001, \beta=0.1$
SocialMF	$\lambda_U=\lambda_V=0.001, \lambda_T=5$	$\lambda_U=\lambda_V=0.001, \lambda_T=1$	$\lambda_U=\lambda_V=0.001, \lambda_T=1$	$\lambda_U=\lambda_V=0.001, \lambda_T=1$
TrustMF	$\lambda_U=\lambda_V=0.001, \lambda_T=1$	$\lambda_U=\lambda_V=0.001, \lambda_T=1$	$\lambda_U=\lambda_V=0.001, \lambda_T=1$	$\lambda_U=\lambda_V=0.001, \lambda_T=1$
SVD++	$\lambda_b=\lambda_U=\lambda_V=0.35$	$\lambda_b=\lambda_U=\lambda_V=0.03$	$\lambda_b=\lambda_U=\lambda_V=0.1$	$\lambda_b=\lambda_U=\lambda_V=0.1$
CRMF	$\lambda=0.4, \rho=80, \alpha=0.1, \beta=0.1$	$\lambda=0.8, \rho=50, \alpha=0.1, \beta=0.01$	$\lambda=0.4, \rho=100, \alpha=0.01, \beta=0.01$	$\lambda=1.2, \rho=20, \alpha=0.01, \beta=0.1$

3.4 实验结果与分析

实验 1:不同特征向量维度下的实验结果对比.

根据文献[27],我们将训练集中评分数量少于 5 个的用户所构成的集合称为冷启动用户集.实验 1 比较了各种推荐算法在全体用户集和冷启动用户集上不同特征向量维度下的结果,文中分别设定潜在特征向量的维度为 $d=5,10$.表 4 给出了各推荐算法在全体用户集上 MAE 和 RMSE 的比较结果,表 5 给出了各推荐算法在冷启动用户集上 MAE 和 RMSE 的比较结果.从两个表中的实验结果可看出,

- (1) 在全体用户集上,SVD++ 与 PMF 相比,其推荐精度有了较大的提升,这说明用户的偏置因素、物品的偏置因素以及用户的隐式反馈信息可提高算法的评分预测精度;
- (2) 在全体用户集和冷启动用户集上,SoReg,SoRec,SocialMF,TrustMF 与 PMF 相比,其推荐精度均有了较大的提高.这说明用户的社会关系有助于提高传统矩阵分解模型的预测精度,并有助于缓解用户的冷启动问题;

- (3) 在冷启动用户集上,CRMF 相比于 6 种对比算法其推荐精度有明显的提升.这是因为当用户为冷启动时,其社会关系的数量往往也较少,此时仅依赖社会关系对推荐算法精度的提升是有限的;但当目标物品为非冷启动时,考虑物品之间的关联关系可提升推荐算法的评分预测精度.

Table 4 Performance comparison on all users
表 4 全体用户集上的性能对比

All		Metrics	PMF	SoRec	SoReg	SociaMF	TrustMF	SVD++	CRMF	Improve (%)
Epinions	d=5	MAE	0.979	0.882	0.946	0.825	0.818	0.818	0.799	2.32
		RMSE	1.290	1.114	1.245	1.070	1.069	1.057	1.043	1.32
	d=10	MAE	0.909	0.884	0.896	0.826	0.819	0.818	0.797	2.56
		RMSE	1.197	1.142	1.167	1.082	1.095	1.057	1.038	1.79
Flixster	d=5	MAE	0.814	0.750	0.774	0.770	0.790	0.794	0.728	5.45
		RMSE	1.076	0.974	1.017	0.994	1.046	1.062	0.951	2.36
	d=10	MAE	0.806	0.785	0.785	0.784	0.794	0.791	0.718	8.42
		RMSE	1.063	1.018	1.034	1.009	1.061	1.048	0.947	6.14
Ciao	d=5	MAE	1.105	0.786	0.920	0.762	0.764	0.735	0.718	2.31
		RMSE	1.448	0.997	1.239	0.977	0.987	0.974	0.952	2.25
	d=10	MAE	0.903	0.776	0.854	0.749	0.767	0.733	0.717	2.18
		RMSE	1.179	1.026	1.167	0.997	1.024	0.967	0.952	1.55
FilmTrust	d=5	MAE	0.714	0.628	0.674	0.638	0.631	0.613	0.608	0.82
		RMSE	0.949	0.810	0.878	0.837	0.810	0.804	0.788	2.00
	d=10	MAE	0.735	0.638	0.668	0.643	0.631	0.611	0.607	0.65
		RMSE	0.968	0.831	0.875	0.844	0.819	0.802	0.787	1.87

Table 5 Performance comparison on cold start users
表 5 冷启动用户集上的性能对比

Cold Start		Metrics	PMF	SoRec	SoReg	SociaMF	TrustMF	SVD++	CRMF	Improve (%)
Epinions	d=5	MAE	1.451	0.892	1.398	0.864	0.891	0.889	0.848	1.85
		RMSE	1.770	1.138	1.735	1.133	1.125	1.162	1.078	4.18
	d=10	MAE	1.153	0.846	0.986	0.857	0.853	0.891	0.847	0.70
		RMSE	1.432	1.180	1.217	1.152	1.176	1.166	1.073	6.85
Flixster	d=5	MAE	1.097	0.872	1.058	0.881	0.871	0.868	0.843	2.88
		RMSE	1.390	1.096	1.358	1.103	1.093	1.122	1.053	3.65
	d=10	MAE	0.951	0.892	0.946	0.884	0.876	0.869	0.843	3.00
		RMSE	1.206	1.144	1.201	1.112	1.102	1.112	1.052	4.54
Ciao	d=5	MAE	1.406	0.775	0.905	0.757	0.750	0.719	0.696	3.20
		RMSE	1.763	0.983	1.106	0.964	0.955	0.958	0.923	3.35
	d=10	MAE	1.047	0.720	0.757	0.732	0.722	0.717	0.695	3.07
		RMSE	1.340	1.072	1.079	0.961	0.958	0.956	0.922	3.56
FilmTrust	d=5	MAE	0.814	0.670	0.783	0.697	0.674	0.677	0.649	3.13
		RMSE	1.079	0.857	0.980	0.916	0.867	0.897	0.838	2.27
	d=10	MAE	0.767	0.668	0.761	0.680	0.687	0.680	0.645	3.44
		RMSE	1.009	0.853	0.958	0.907	0.900	0.905	0.841	1.41

实验 2:不同评分数量下的实验结果对比.

为了对比不同评分数量下各算法的效果,本文依据训练集中用户的评分数量将用户划分为[0:10],[11:30],[31:50],[51:100],100 以上这 5 组,图 2 显示了 4 个数据集中每组用户所占的比例,从图中可看出:4 个数据集中,评分数量少于 10 个的用户占大部分.依据用户的评分数量对训练集中的用户进行划分之后,使用训练集学习相应的推荐算法,然后在 5 组用户上分别计算 RMSE,结果如图 3 所示.从 4 个数据集的实验结果可看出,

- (1) 在数据极度稀疏时(评分数量小于 10),CRMF 算法明显优于其他推荐算法.这因为评分数据稀疏时,用户的社会关系往往也比较稀疏,同时考虑物品的关联关系可提高算法的推荐精度;
- (2) 随着用户评分数量的增加,CRMF 及对比算法所得出的 RMSE 值并没有持续变小.这因为用户的评分数量较少时,数据比较稀疏使得模型不易学习到用户的潜在特征向量.用户的评分数据过多,用户的兴趣会发散,这造成无法准确学习到表示用户喜好的潜在特征向量.但 C RMF 仍优于其他对比算法,这因为基于联合正则化的矩阵分解模型考虑了自适应正则化,它能够根据用户评分的数量对正则化程度进行自适应地调整,因此使得 C RMF 算法可更加有效地学习出模型的参数.

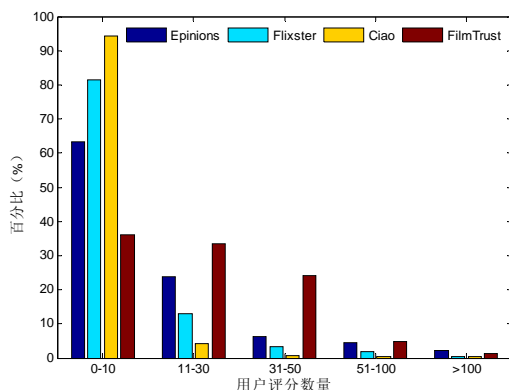


Fig.2 Distribution of users with different ratings

图 2 不同评分数量下用户的分布

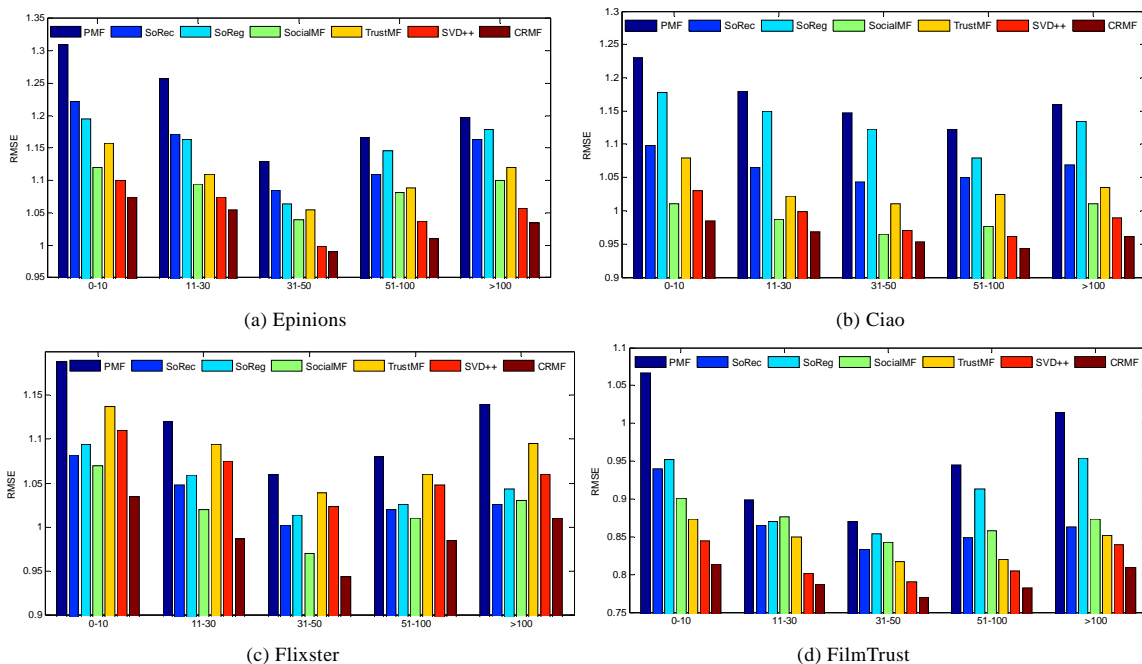


Fig.3 Comparison on users with different ratings

图 3 不同评分数量下用户对比结果

实验 3:不同社会关系下的实验结果对比.

为了评估各算法在不同社会关系下的推荐精度,依据训练集中用户拥有的社会关系数量将用户划分为 [1:5],[6:10],[11:20],[21:40],[41:100],[101:500],500 以上这 7 组,图 4 显示了每组用户所占的比例,从图中可看出,大多数用户所拥有的社会关系数量不多于 5 个.依据用户所拥有的社会关系数量对训练集中的用户进行划分后,使用训练集学习相应的算法.本实验仅对考虑社会关系的 SoRec,SoReg,SocialMF,TrustMF,CRMF 这 5 种推荐算法进行了详细对比,结果如图 5 所示.从 4 个数据集的实验结果可看出,

- (1) 在 FilmTrust 数据集上,各算法的 $RMSE$ 值反而随着用户拥有社会关系数量的增加而增加.根据文献 [43]可知,该数据集将社会关系中真实的信任度转成二值数据时丢失了一些有用的信息,同时产生了噪声数据;
- (2) 在 Epinions,Flixster 及 Ciao 数据集上,随着社会关系数量的增加,5 种推荐算法的 $RMSE$ 值并没有持续

变小.这因为当用户拥有的社会关系数量较少时,仅靠评分数据不能有效地对用户建模而影响了算法的精度.当用户的社会关系数量过多被考虑时,用户的兴趣过度依靠近邻的影响,使得学习的用户的潜在特征向量会比较发散.因此仅当用户的社会关系达到一定数量时,5 种推荐算法才取得较优效果;

(3) 在 4 个数据集的每组用户上,CRMF 均要优于其他推荐算法.这因为基于联合正则化的矩阵分解模型不仅考虑了社会关系与关联关系,而且使用了自适应正则化,使得矩阵分解模型能有效地对用户和物品进行建模.

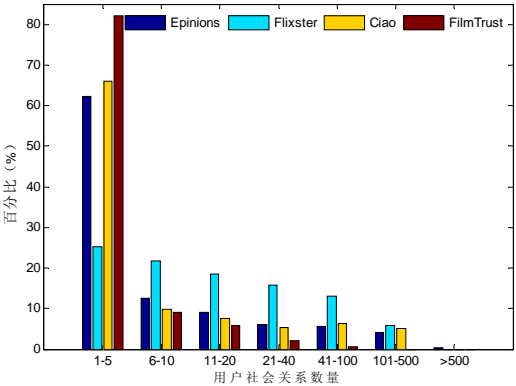


Fig.4 Distribution of users with different social relationships

图 4 不同社会关系下用户的分布

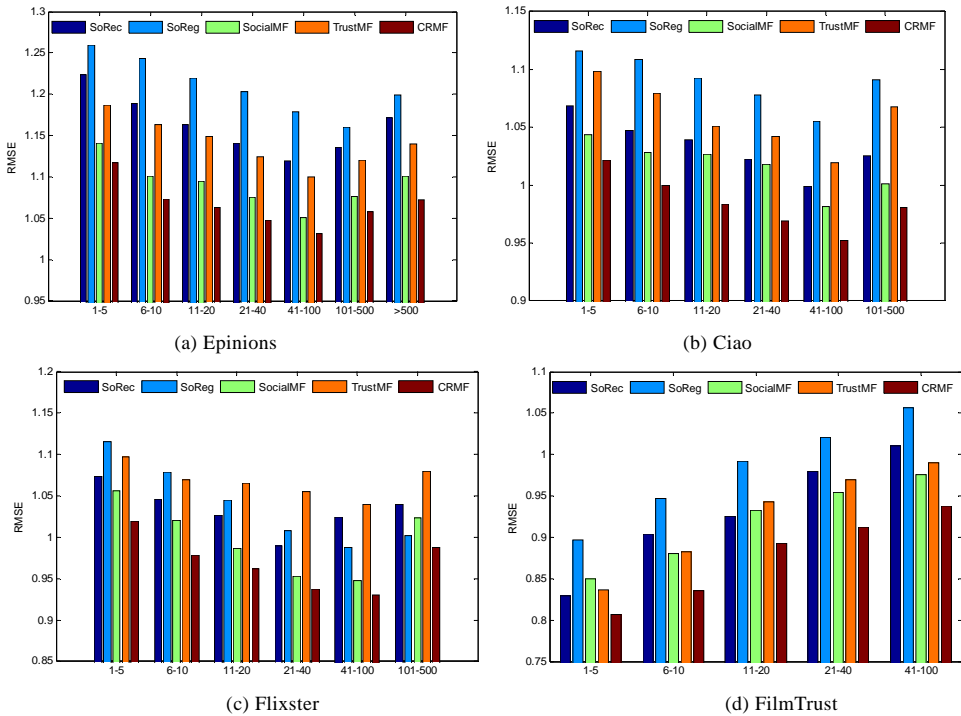


Fig.5 Comparison on users with different social relationships

图 5 不同社会关系下用户的对比结果

实验 4:关联关系获取方法的对比.

本实验是为了评估不同的获取关联关系的方法对推荐算法精度的影响.我们首先根据皮尔逊相关系数计算两两物品之间的相似度;然后,为每个目标物品选择 10 个最相似的物品作为关联关系;最后,将得到的无向关

联关系与 SVD++结合,得到改进的推荐算法,简称 SVD-IS.相应地,基于本文构建的关联程度(公式(7))可为每个物品选择 10 个最相关的物品作为有向的关联关系.同样地,我们将获取的有向关联关系与 SVD++结合得到改进的推荐算法,简称 SVD-IR.SVD-IS,SVD-IR 在相同超参数下的实验结果如表 6 所示.从表中的实验结果可看出,

- (1) 相比于表 4 中的 SVD++算法,考虑了物品之间相似性关系的 SVD-IS 在 4 个数据集上,其推荐精度有不同程度的提升.这是因为考虑物品的近邻关系,能够更好地对物品进行建模,这与文献[36,37]的结论是一致的;
- (2) 相比于 SVD-IS,SVD-IR 在 4 个数据集上,其推荐精度有了较高的提升.这是因为相比于无向的相似性关系,本文方法所获取的有向关联关系更符合现实世界中物品推荐的场景.

Table 6 Impact of item relations on RMSE
表 6 物品的关联关系对 RMSE 的影响

算法	数据集			
	Epinions	Flixster	Ciao	FilmTrust
SVD-IS	1.051	0.995	0.964	0.799
SVD-IR	1.045	0.974	0.960	0.795

实验 5:自适应正则化的有效性验证.

本实验是为了评估自适应正则化对于推荐算法精度的影响.我们将用户的社会关系和物品的关联关系同时与 SVD++算法结合,可得改进的推荐算法,简称 SVD-DR.SVD-DR 与 CRMF 算法在 4 个数据集上的评分预测精度见表 7.从表中的实验结果可看出:考虑了自适应正则化的推荐算法 CRMF 明显优于仅考虑社会关系和关联关系的推荐算法 SVD-DR.这是因为考虑自适应正则化的矩阵分解模型不仅无需评分矩阵中评分的分布为均匀分布.而且对于大规模非平衡数据集更易搜索到合适的正则化超参数.

Table 7 Impact of adaptive regularization on RMSE
表 7 自适应正则化对 RMSE 的影响

算法	数据集			
	Epinions	Flixster	Ciao	FilmTrust
SVD-DR	1.041	0.968	0.957	0.793
CRMF	1.038	0.947	0.952	0.787

实验 6:超参数对算法的影响.

为了说明 CRMF 算法的超参数对推荐结果的影响,本文以数据集 Ciao 为例,详细讨论了各超参数下的实验结果.本文设定 α 和 β 的取值为 0.00001,0.0001,0.001,0.01,0.1,0.3.超参数 C 的取值为 0,10,20,50,100,150,200.超参数 K 的取值为 0,10,20,30,40,50,60,70,80,90,100.从图 6 可看出,

- (1) β 控制着用户的社会关系对推荐结果的影响.图 6(a)显示了超参数 β 的变化对 RMSE 的影响,其值从 0.00001 开始先下降,到达某个阈值后上升.这说明用户的社会关系在 CRMF 算法中所占的比重直接影响推荐结果的好坏;
- (2) α 控制着物品之间的关联关系在推荐过程所占的比重.图 6(b)显示了超参数 α 的变化对 RMSE 值的影响,其值从 0.00001 开始先下降,到达某个阈值后上升.这表明物品之间的关联关系在 CRMF 算法中所占的比重直接影响推荐结果的好坏.在数据集 Ciao 上,由 $\alpha=\beta=0.01$ 算法取得较高精度,说明物品之间的关联关系在用户决策时也是必不可少的;
- (3) $\rho=0$ 表示物品之间的关联程度等于关联规则中的置信度.从图 6(c)可看出:随着 ρ 值的增加,算法的推荐精度先增加后降低.这说明物品之间关联程度的获取,仅当综合考虑支持度与置信度时才使得算法精度较高;
- (4) σ 表示物品拥有的关联关系数量.从图 6(d)可看出:当 $\sigma=0$ 时,表示 CRMF 算法不考虑物品之间的关联关系;随着 σ 值的增加,算法的精度有所提高,这说明本文获取物品之间关联关系的方法的有效性;在 $\sigma=40$ 以后,算法的精度开始下降,这主要是因为 σ 值过大会导致对目标物品引入一些关联程度较低的

关联关系.

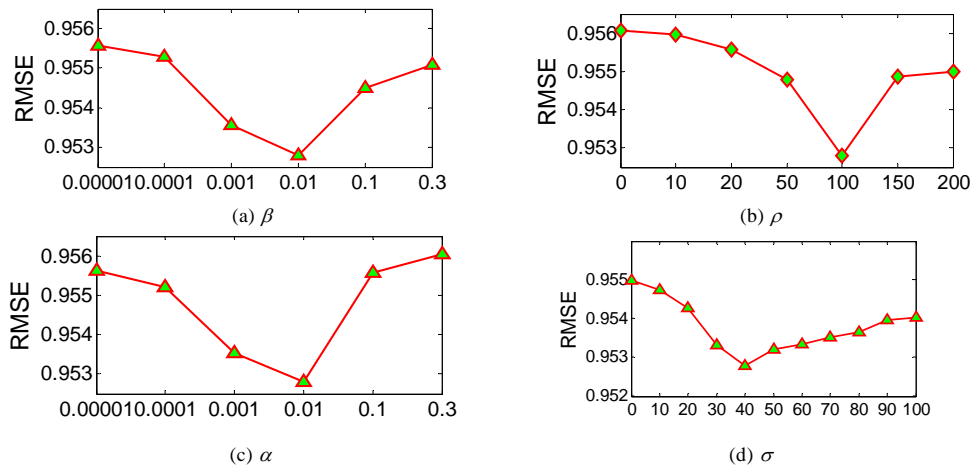


Fig.6 Impact of hyperparameters

图 6 超参数的影响

4 结 论

本文首先给出一种度量物品之间关联程度的方法,通过该方法可获取物品之间的有向关联关系;然后将其以关联正则化的方式与自适应正则化相结合,融入基于社会正则化的矩阵分解模型中,提出了基于联合正则化的矩阵分解模型,并证明了联合正则化是一种加权的原子范数.由于物品之间关联关系的获取无需物品的标签、属性等额外信息,因此,基于联合正则化的矩阵分解模型拥有更广泛的应用范围.在 4 个真实数据集上的实验表明:与主流的推荐算法相比,本文算法不仅可缓解用户的冷启动问题,而且能够更有效地预测不同类型用户的实际评分.

在未来的研究中,我们将把重点放在解决物品冷启动相关的问题上.这种情况下,如何将情境信息(时间、地理等)和物品的视觉特征融入到本文的推荐模型中,来更好地预测目标用户对特定物品的喜好,这也是我们下一步要改进的方向.

References:

- [1] Lü LY, Medo M, Yeung CH, Zhang YC, Zhang ZK, Zhou T. Recommender systems. *Physics Reports*, 2012,519(1):1–49.
- [2] Bobadilla J, Ortega F, Hernando A, Gutiérrez A. Recommender systems survey. *Knowledge-Based Systems*, 2013,46(1):109–132.
- [3] Wu L, Liu Q, Chen E, Yuan NJ, Guo GM, Xie X. Relevance meets coverage: A unified framework to generate diversified recommendations. *ACM Trans. on Intelligent Systems and Technology (TIST)*, 2016,7(3):39.
- [4] Liu Q, Chen E, Xiong H, Ding CH, Chen J. Enhancing collaborative filtering by user interest expansion via personalized ranking. *IEEE Trans. on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2012,42(1):218–233.
- [5] Liu JG, Zhou T, Wang BH. Research progress of personalized recommendation system. *Progress in Natural Science*, 2009,19(1): 1–15 (in Chinese with English abstract).
- [6] Adomavicius G, Tuzhilin A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Trans. on Knowledge and Data Engineering*, 2005,17(16):734–749.
- [7] Sarwar B, Karypis G, Konstan J, Riedl J. Item-Based collaborative filtering recommendation algorithms. In: *Proc. of the 10th Conf. on World Wide Web*. ACM Press, 2001. 285–295.
- [8] Liu Q, Chen E, Xiong H, Ge Y, Li ZM, Wu X. A cocktail approach for travel package recommendation. *IEEE Trans. on Knowledge and Data Engineering*, 2014,26(2):278–293.
- [9] Chiang KY, Hsieh CJ, Dhillon IS. Matrix completion with noisy side information. In: *Proc. of the 28th Conf. on Neural Information Processing Systems*. Curran Associates Press, 2015. 3447–3455.

- [10] Rao N, Yu HF, Ravikumar P, Dhillon IS. Collaborative filtering with graph information: Consistency and scalable methods. In: Proc. of the 28th Conf. on Neural Information Processing Systems. Curran Associates Press, 2015. 2107–2115.
- [11] Li DS, Chen C, Lv Q, Yan JC, Shang L, Chu SM. Low-Rank matrix approximation with stability. In: Proc. of the 33rd Int'l Conf. on Machine Learning. ACM Press, 2016.
- [12] Wang YX, Xu H. Stability of matrix factorization for collaborative filtering. In: Proc. of the 29th Int'l Conf. on Machine Learning. ACM Press, 2012. 417–424.
- [13] Bhaskar SA. Probabilistic low-rank matrix completion from quantized measurements. *Journal of Machine Learning Research*, 2016, 17(60):1–34.
- [14] Koren Y, Bell R, Volinsky C. Matrix factorization techniques for recommender systems. *IEEE Computer*, 2009, 42(8):30–37.
- [15] Koren Y. Collaborative filtering with temporal dynamics. In: Proc. of the 9th Int'l ACM SIGKDD Conf. on Knowledge Discovery and Data Mining. ACM Press, 2009. 89–97.
- [16] Hu Y, Koren Y, Volinsky C. Collaborative filtering for implicit feedback datasets. In: Proc. of the 8th IEEE Int'l Conf. on Data Mining. IEEE, 2008. 263–272.
- [17] Srebro N, Rennie JDM, Jaakkola T. Maximum margin matrix factorization. In: Proc. of the 17th Int'l Conf. on Neural Information Processing Systems. Curran Associates Press, 2004. 1329–1336.
- [18] Lawrence ND, Urtasun R. Non-Linear matrix factorization with Gaussian processes. In: Proc. of the 26th Int'l Conf. on Machine Learning. ACM Press, 2009. 601–608.
- [19] Liu XY, Aggarwal C, Li YF, Kong XG, Sun XY, Sathe S. Kernelized matrix factorization for collaborative filtering. In: Proc. of the SIAM Int'l Conf. on Data Mining. SIAM/Omnipress, 2016. 399–416.
- [20] Salakhutdinov R, Mnih A. Probabilistic matrix factorization. In: Proc. of the 21st Conf. on Neural Information Processing Systems. Curran Associates Press, 2008. 1257–1264.
- [21] Salakhutdinov R, Mnih A. Bayesian probabilistic matrix factorization using Markov chain Monte Carlo. In: Proc. of the 25th Int'l Conf. on Machine Learning. ACM Press, 2008. 880–887.
- [22] Koren Y. Factorization meets the neighborhood: A multifaceted collaborative filtering model. In: Proc. of the 8th Int'l ACM SIGKDD Conf. on Knowledge Discovery and Data Mining. ACM Press, 2008. 426–434.
- [23] Ma H, Yang HX, Lyu MR, King I. SoRec: Social recommendation using probabilistic matrix factorization. In: Proc. of the 17th Int'l Conf. on Information and Knowledge Management. ACM Press, 2008. 978–991.
- [24] Yang B, Lei Y, Liu DY, Liu JM. Social collaborative filtering by trust. In: Proc. of the 23rd Int'l Joint Conf. on Artificial Intelligence. IJCAI/AAAI Press, 2013. 2747–2753.
- [25] Guo L, Ma J, Chen ZM. Trust strength aware social recommendation method. *Journal of Computer Research and Development*, 2015, 50(9):1805–1813 (in Chinese with English abstract).
- [26] Ma H, Zhou D, Liu C, Lyu MR, King I. Recommender systems with social regularization. In: Proc. of the 25th Int'l Conf. on Web Search and Data Mining. ACM Press, 2011. 287–296.
- [27] Jamali M, Ester M. A matrix factorization technique with trust propagation for recommendation in social networks. In: Proc. of the 4th Int'l Conf. on Recommender Systems. ACM Press, 2010. 135–142.
- [28] Ma H, Lyu MR, King I. Learning to recommend with trust and distrust relationships. In: Proc. of the 3rd Int'l Conf. on Recommender Systems. ACM Press, 2009. 189–196.
- [29] Meng XW, Liu SD, Zhang YJ, Hu X. Research on social recommender systems. *Ruan Jian Xue Bao/Journal of Software*, 2015, 26(6):1356–1372 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4831.htm> [doi: 10.13328/j.cnki.jos.004831]
- [30] Sun GF, Wu L, Liu Q, Zhu C, Chen EH. Recommendations based on collaborative filtering by exploiting sequential behaviors. *Ruan Jian Xue Bao/Journal of Software*, 2013, 24(11):2721–2733 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4478.htm> [doi: 10.3724/SP.J.1001.2013.04478]
- [31] Wu L, Chen E, Liu Q, Xu LL, Bao TF, Zhang L. Leveraging tagging for neighborhood-aware probabilistic matrix factorization. In: Proc. of the 21st ACM Int'l Conf. on Information and Knowledge Management. ACM Press, 2012. 1854–1858.
- [32] Gu QQ, Zhou J, Ding C. Collaborative filtering: Weighted nonnegative matrix factorization incorporating user and item graphs. In: Proc. of the SIAM Int'l Conf. on Data Mining. SIAM/Omnipress, 2010. 199–210.
- [33] Yu X, Ren X, Gu Q, Sun YZ, Han JW. Collaborative filtering with entity similarity regularization in heterogeneous information networks. In: Proc. of the 2nd IJCAI Workshop on Heterogeneous Information Network Analysis. 2013.

- [34] Luo C, Pang W, Wang Z. Hete-CF: Social-Based collaborative filtering recommendation using heterogeneous relations. In: Proc. of the 2014 IEEE Int'l Conf. on Data Mining. IEEE, 2014. 917–922.
- [35] Liang D, Altosaar J, Charlin L, Blei DM. Factorization meets the item embedding: Regularizing matrix factorization with item co-occurrence. In: Proc. of the 10th Int'l Conf. on Recommender Systems. ACM Press, 2016. 59–66.
- [36] Ma H. An experimental study on implicit social recommendation. In: Proc. of the 36th Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. ACM Press, 2013. 73–82.
- [37] Zheng J, Liu J, Shi C, Zhuang FZ, Li JZ, Wu B. Dual similarity regularization for recommendation. In: Proc. of the 20th Pacific-Asia Conf. on Knowledge Discovery and Data Mining. Springer Int'l Publishing, 2016. 542–554.
- [38] Charu C. Recommender Systems: The Textbook. New York: Springer-Verlag, 2016. 77–81.
- [39] Srebro N, Salakhutdinov R. Collaborative filtering in a non-uniform world: Learning with the weighted trace norm. In: Proc. of the 23rd Conf. on Neural Information Processing Systems. Curran Associates Press, 2010. 2056–2064.
- [40] Chandrasekaran V, Recht B, Parrilo PA, Willsky AS. The convex geometry of linear inverse problems. Foundations of Computational Mathematics, 2012,12(6):805–849.
- [41] Chen YD, Bhojanapalli S, Sanghavi S, Ward R. Coherent matrix completion. Journal of Machine Learning Research, 2014,32: 674–682.
- [42] Bouchard G, Guo SB, Yin DW. Convex collective matrix factorization. Journal of Machine Learning Research, 2013,31:144–152.
- [43] Fang H, Bao Y, Zhang J. Leveraging decomposed trust in probabilistic matrix factorization for effective recommendation. In: Proc. of the 28th AAAI Conf. on Artificial Intelligence. AAAI Press, 2014. 30–36.

附中文参考文献:

- [5] 刘建国,周涛,汪秉宏.个性化推荐系统的研究进展.自然科学研究进展,2009,19(1):1–15.
- [25] 郭磊,马军,陈竹敏.一种信任关系强度敏感的社会化推荐算法.计算机研究与发展,2015,50(9):1805–1813.
- [29] 孟祥武,刘树栋,张玉洁,胡勋.社会化推荐系统研究.软件学报,2015,26(6):1356–1372. <http://www.jos.org.cn/1000-9825/4831.htm> [doi: 10.13328/j.cnki.jos.004831]
- [30] 孙光福,吴乐,刘淇,朱琛,陈恩红.基于时序行为的协同过滤推荐算法.软件学报,2013,24(11):2721–2733. <http://www.jos.org.cn/1000-9825/4478.htm> [doi: 10.3724/SP.J.1001.2013.04478]



吴宾(1991—),男,河南柘城人,硕士,CCF 学生会员,主要研究领域为机器学习,推荐系统.



叶阳东(1962—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为智能系统,机器学习,数据库.



娄铮铮(1984—),男,博士,讲师,CCF 专业会员,主要研究领域为机器学习,模式识别,推荐系统.