



## 结合矩阵补全的宽度协同过滤推荐算法

史加荣, 何攀

引用本文:

史加荣,何攀. 结合矩阵补全的宽度协同过滤推荐算法[J]. 智能系统学报, 2024, 19(2): 299–306.

SHI Jiarong, HE Pan. Broad collaborative filtering recommendation algorithm combined with matrix completion[J]. *CAAI Transactions on Intelligent Systems*, 2024, 19(2): 299–306.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202209005>

## 您可能感兴趣的其他文章

### 联邦推荐系统的协同过滤冷启动解决方法

Cold starts in collaborative filtering for federated recommender systems

智能系统学报. 2021, 16(1): 178–185 <https://dx.doi.org/10.11992/tis.202009032>

### 融合用户特征优化聚类的协同过滤算法

Collaborative filtering algorithm combining user features and preferences in optimized clustering

智能系统学报. 2020, 15(6): 1091–1096 <https://dx.doi.org/10.11992/tis.201710024>

### 基于图游走的并行协同过滤推荐算法

Parallel collaborative filtering recommendation algorithm based on graph walk

智能系统学报. 2019, 14(4): 743–751 <https://dx.doi.org/10.11992/tis.201806002>

### 融合协同过滤与用户偏好的旅游组推荐方法

A tourist group recommendation method combining collaborative filtering and user preferences

智能系统学报. 2018, 13(6): 999–1005 <https://dx.doi.org/10.11992/tis.201802011>

### 面对智能导诊的个性化推荐算法

A personalized recommendation algorithm for intelligent guidance

智能系统学报. 2018, 13(3): 352–358 <https://dx.doi.org/10.11992/tis.201711036>

### 个性化信息推荐方法研究

Research on the recommendation method of personalized information

智能系统学报. 2018, 13(2): 189–195 <https://dx.doi.org/10.11992/tis.201701002>

DOI: 10.11992/tis.202209005

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20231117.1651.006>

# 结合矩阵补全的宽度协同过滤推荐算法

史加荣, 何攀

(西安建筑科技大学 理学院, 陕西 西安 710055)

**摘要:** 协同过滤是推荐系统中最经典的方法之一, 能够满足人们对个性化推荐任务的需求, 但许多协同过滤算法在面对评分数据稀疏性问题时推荐效果不佳。为解决此问题, 提出一种结合矩阵补全的宽度协同过滤推荐算法。先使用矩阵补全技术对用户项目评分矩阵进行补全, 再利用补全后的矩阵对已评分的用户和项目分别寻找其近邻项, 进而构造用户与项目的评分协同向量, 最后使用宽度学习系统来构建用户项目与评分之间的复杂的非线性关系。在 MovieLens 和 filmtrust 数据集上对所提出算法的有效性进行检验。试验结果表明, 与当前最先进的方法相比, 该方法能够有效地缓解数据稀疏性问题, 具有较低的计算复杂度, 在一定程度上提升了推荐系统的性能。

**关键词:** 推荐系统; 宽度学习系统; 矩阵补全; 宽度协同过滤; 协同过滤; 深度矩阵分解; 数据稀疏性; 深度学习  
**中图分类号:** TP181 **文献标志码:** A **文章编号:** 1673-4785(2024)02-0299-08

中文引用格式: 史加荣, 何攀. 结合矩阵补全的宽度协同过滤推荐算法 [J]. 智能系统学报, 2024, 19(2): 299-306.

英文引用格式: SHI Jiarong, HE Pan. Broad collaborative filtering recommendation algorithm combined with matrix completion[J]. CAAI transactions on intelligent systems, 2024, 19(2): 299-306.

## Broad collaborative filtering recommendation algorithm combined with matrix completion

SHI Jiarong, HE Pan

(School of Science, Xi'an University of Architecture &amp; Technology, Xi'an 710055, China)

**Abstract:** Collaborative filtering is a classic method used in recommendation systems, designed to cater to the need for personalized recommendations. However, many collaborative filtering algorithms struggle when confronted with sparse rating data. To address this issue, we propose a broad collaborative filtering recommendation algorithm that integrates matrix completion. Initially, a matrix completion technique is employed to recover the user-item rating matrix. Subsequently, this completed rating matrix is utilized to identify respective neighbors for a given user-item pair, which in turn helps create the user-item rating collaboration vector. Finally, a broad learning system is employed to establish the complex nonlinear relationship between user-items and ratings. The effectiveness of the proposed algorithm has been validated through tests on MovieLens and Filmtrust data sets. The experimental results show that, compared with state-of-the-art collaborative filtering methods, the proposed method can effectively alleviate the data sparsity problem. It also possesses lower computational complexity and enhances recommendation performance to a certain extent.

**Keywords:** recommendation system; broad learning system; matrix completion; broad collaborative filtering; collaborative filtering; deep matrix factorization; data sparsity; deep learning

现今, 人们生活在数据爆炸不断加剧的时代<sup>[1]</sup>。所接收到的巨大的数据量已远超出人们的

信息处理的能力范围, 且很难从海量数据中挖掘出所需要的信息。因此, 从大量数据中过滤出可用数据的技术是非常必要的。推荐系统是挖掘数据以及解决信息过载的有效途径<sup>[2]</sup>, 在音乐、短视

收稿日期: 2022-09-05. 网络出版日期: 2023-11-20.

基金项目: 陕西省自然科学基金项目 (2021JM-378).

通信作者: 史加荣. E-mail: [shijiarong@xauat.edu.cn](mailto:shijiarong@xauat.edu.cn).

©《智能系统学报》编辑部版权所有

频和电影推荐等方面应用广泛<sup>[3-4]</sup>。

协同过滤(collaborative filtering, CF)是众多个性化推荐方法中最经典且最受欢迎的方法之一<sup>[4]</sup>,传统的CF推荐算法可分为2类:基于模型和基于邻域。第1类CF方法以矩阵分解(matrix factorization, MF)为主,即将用户项目评分矩阵分解为用户潜在低秩矩阵和项目潜在低秩矩阵的乘积<sup>[5]</sup>。第2类通过计算用户或项目之间的相似度来帮助其找到若干个合适的近邻。但用户对项目评分数据的长尾分布导致评分矩阵的稀疏性问题,进而严重影响CF算法的推荐性能。

随着深度学习<sup>[6]</sup>在多个领域中的成熟和发展,人们开发出许多基于深度神经网络(deep neural network, DNN)的CF推荐算法来提高推荐性能<sup>[7-9]</sup>。He等<sup>[10]</sup>将MF模型中的内积运算替换为深度神经网络结构,提出了神经协同过滤(neural collaborative filtering, NeuCF)。Xue等<sup>[11]</sup>将深度神经网络用于学习用户和项目在公共低维空间的表示,提出了深度矩阵分解(deep matrix factorization, DMF)。但由于数据存在稀疏性问题<sup>[12]</sup>,导致复杂的DNN模型的推荐性能不佳。

宽度协同过滤(broad collaborative filtering, BCF)<sup>[13]</sup>是一种基于神经网络的CF算法,其采用相似性度量来生成低维的用户与项目联合特征,并基于宽度学习系统(broad learning system, BLS)<sup>[14]</sup>来建立映射函数,以学习低维用户项目联合特征向量与评分之间的复杂非线性关系。与深度学习和传统的神经网络相比,BCF具有闭形式解,无需使用梯度下降法更新权值,从而具有非常低的计算复杂度。但BCF仍然存在评分数据稀疏性问题<sup>[15]</sup>。

为了解决前述问题并充分利用用户和项目的近邻信息,本研究提出一种结合矩阵补全的宽度协同过滤(broad collaborative filtering combined with matrix completion, MC+BCF)推荐算法。该算法先利用矩阵补全技术将原始的用户项目评分矩阵补全,再对已评分元素在补全后的矩阵中寻找用户及项目的近邻项,进而构造用户项目评分协同向量。最后将所构建的协同向量作为输入特征、原始用户评分的独热编码作为标签输入到宽度学习系统中,获取用户项目特征与评分之间的复杂非线性关系。

## 1 相关工作

为了克服原始评分矩阵的稀疏性问题,使用矩阵补全算法恢复原始评分矩阵;介绍了宽度学

习系统,以用于推荐算法。

### 1.1 矩阵补全

矩阵补全(matrix completion, MC)<sup>[16-17]</sup>是指矩阵存在缺失元素情况下,根据数据矩阵的低复杂度将其恢复。MC广泛地应用在计算机视觉<sup>[18]</sup>、推荐系统<sup>[19]</sup>等领域。通常假设待恢复的矩阵是低秩的,通过低秩矩阵重构来实现缺失元素的恢复。此时,矩阵 $\mathbf{R} = (r_{ij})_{m \times n}$ 的补全问题等价于秩最小化模型:

$$\min_{\mathbf{D}} \text{rank}(\mathbf{D}), \quad \text{s.t.} \quad d_{ij} = r_{ij}, \quad \forall (i, j) \in \Omega \quad (1)$$

式中:rank( $\cdot$ )表示矩阵的秩; $\mathbf{R}$ 为原始的用户项目评分矩阵; $\mathbf{D}$ 为补全后的评分矩阵; $\Omega$ 为用户已评分项目的用户项目对的下标集; $m$ 和 $n$ 分别为用户数和项目数。

上述优化模型是一个非凸的NP难问题,故难以求解。因此,用矩阵核范数将模型(1)凸松弛

$$\min_{\mathbf{D}} \|\mathbf{D}\|_*, \quad \text{s.t.} \quad d_{ij} = r_{ij}, \quad \forall (i, j) \in \Omega \quad (2)$$

其中 $\|\cdot\|_*$ 为矩阵核范数。

使用非精确的增广拉格朗日乘子(inexact augmented lagrange multiplier, IALM)算法<sup>[20]</sup>求解问题(2)。为此,将该问题重新表示为

$$\min_{\mathbf{D}, \mathbf{E}} \|\mathbf{D}\|_*, \quad \text{s.t.} \quad \mathbf{D} + \mathbf{E} = \mathbf{R}, \quad \pi_{\Omega}(\mathbf{E}) = 0 \quad (3)$$

这里 $\pi_{\Omega}: \mathbf{R}^{m \times n} \rightarrow \mathbf{R}^{m \times n}$ 是一个正交投影算子,用来保持在 $\Omega$ 中的条目不变,并将 $\Omega$ 之外的其他元素设置为0。构造问题(3)的部分增广拉格朗日函数

$$L(\mathbf{D}, \mathbf{E}, \mathbf{Y}, \mu) = \|\mathbf{D}\|_* + \langle \mathbf{Y}, \mathbf{R} - \mathbf{D} - \mathbf{E} \rangle + \frac{\mu}{2} \|\mathbf{R} - \mathbf{D} - \mathbf{E}\|_F^2 \quad (4)$$

式中: $\mathbf{Y}$ 为拉格朗日乘子矩阵; $\mu > 0$ 是惩罚项系数; $\|\cdot\|_F$ 为矩阵的Frobenius范数; $\langle \cdot, \cdot \rangle$ 表示两矩阵的内积。使用IALM求解问题(3)的算法流程如下。

**输入** 评分矩阵 $\mathbf{R} \in \mathbf{R}^{m \times n}$ ,  $r_{ij} = 0 (\forall (i, j) \notin \Omega)$ 。

1)初始化:  $\mathbf{Y}_0 = 0, \mathbf{E}_0 = 0, \mu_0 > 0, \rho > 1, k = 0$ ;

2)while 未收敛 do:

3)  $(\mathbf{U}, \mathbf{S}, \mathbf{V}) = \text{SVD}(\mathbf{R} - \mathbf{E}_k + \mu_k^{-1} \mathbf{Y}_k)$ ;

4)  $\mathbf{D}_{k+1} = \mathbf{U} \mathbf{S}_{\mu_k^{-1}}[\mathbf{S}] \mathbf{V}^T$ ;

5)  $\mathbf{E}_{k+1} = \pi_{\Omega}(\mathbf{R} - \mathbf{D}_{k+1} + \mu_k^{-1} \mathbf{Y}_k)$ ;

6)  $\mathbf{Y}_{k+1} = \mathbf{Y}_k + \mu_k(\mathbf{R} - \mathbf{D}_{k+1} - \mathbf{E}_{k+1})$ ;

7)  $\mu_{k+1} = \min(\rho \mu_k, \bar{\mu})$ ;

8)  $k := k + 1$ ;

9)end while

**输出**  $(\mathbf{D}_k, \mathbf{E}_k)$ 。

在上述迭代过程中,算子SVD表示对矩阵进行奇异值分解, $\mathbf{S}_\lambda[\cdot]$ 是矩阵的绝对值阈值算子<sup>[21]</sup>, $\bar{\Omega}$ 表示 $\Omega$ 的补集, $\bar{\mu}$ 为事先给定的正数。

## 1.2 宽度学习系统

宽度学习系统(BLS)<sup>[14]</sup>是基于随机向量函数连接神经网络<sup>[22]</sup>的一类轻量级模型。在BLS建模过程中,还可以对新的输入样本采用增量学习算法<sup>[23]</sup>来更新整个网络结构,且无需重复整个训练过程,从而快速完成模型的重建。BLS主要包括:映射特征层、增强特征层和输出层,其结构如图1所示。

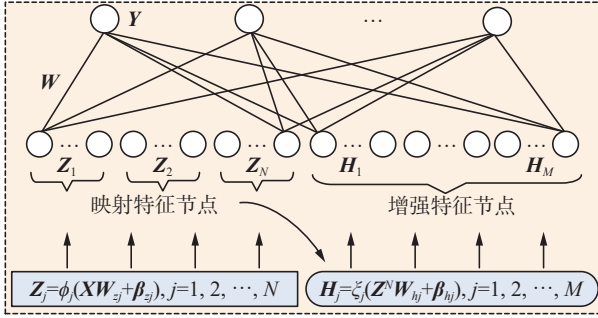


图1 宽度学习系统网络结构

Fig. 1 Network structure of broad learning system

将输入数据集组成矩阵 $X$ ,其中每行表示一个样本。在映射特征层,先随机生成权重矩阵 $W_{zj}$ 和偏置矩阵 $\beta_{zj}$ ,再建立 $N$ 个映射

$$Z_j = \phi_j(XW_{zj} + \beta_{zj}), \quad j = 1, 2, \dots, N \quad (5)$$

式中: $\phi_j$ 为第 $j$ 个映射的激活函数。对于所得到的 $N$ 个映射特征节点 $\{Z_j\}_{j=1}^N$ ,在特征维上作连接

$$Z^N = [Z_1 | Z_2 | \dots | Z_N] \quad (6)$$

类似地,在增强特征层采用 $M$ 个函数将 $Z^N$ 映射为增强特征节点

$$H_j = \xi_j(Z^N W_{hj} + \beta_{hj}), \quad j = 1, 2, \dots, M \quad (7)$$

式中: $W_{hj}$ 和 $\beta_{hj}$ 分别是第 $j$ 个映射随机生成的权重矩阵和偏置矩阵; $\xi_j$ 为激活函数。同理,将 $M$ 个增强特征节点 $\{H_j\}_{j=1}^M$ 横向连接起来

$$H^M = [H_1 | H_2 | \dots | H_M] \quad (8)$$

最后,在输出层对映射特征节点 $Z^N$ 和增强特征节点 $H^M$ 作线性映射

$$Y = [Z^N | H^M] W \quad (9)$$

式中: $W$ 表示待求的连接权重矩阵; $Y$ 为所有训练集的输出组成的矩阵。在模型训练过程中,根据最小二乘法求解 $[Z^N | H^M]$ 的伪逆得最优权重 $W$ 。

## 2 基于矩阵补全的宽度协同过滤

先给出协同过滤中输入特征与输出特征的构造方式,再建立宽度协同过滤模型。

### 2.1 特征构造

#### 2.1.1 基于矩阵补全的用户项目评分协同向量

经过矩阵补全技术,所恢复的评分矩阵 $D$ 不再是稀疏的。针对已评分的用户项目对 $(u, i)$ ,通

过近邻规则分别寻找用户 $u$ 的 $k$ 近邻用户和项目 $i$ 的 $l$ 近邻项目,即得到用户低维特征和项目低维特征。将这些近邻项对应的特征进行连接,就得到用户项目的评分协同向量,这种构造方式有效地克服了数据稀疏性问题。图2给出了用户项目评分协同向量的构造示意图。

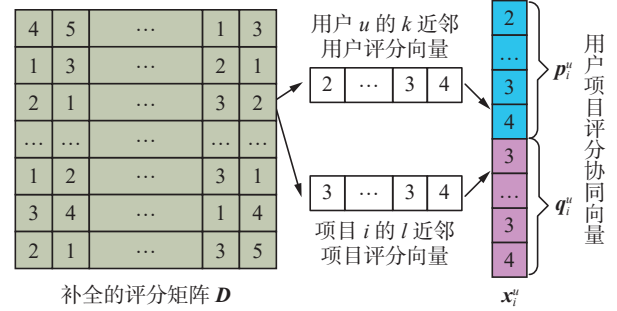


图2 用户项目评分协同向量的构造

Fig. 2 Construction of user-item rating collaborative vector

对于固定的用户 $u$ ,根据 $D_u$ 和 $D_j$ 的夹角余弦得到用户 $u$ 与用户 $j$ 的相似度,其中 $D_u$ 表示矩阵 $D$ 的第 $u$ 行。选择用户 $u$ 的 $k$ 个最近邻用户,其索引构成的向量记为 $\mathbf{n}^u = (n_1^u, n_2^u, \dots, n_k^u)$ 。考虑用户 $u$ 对项目 $i$ 评分时,用户 $u$ 的 $k$ 近邻用户的评分向量为

$$\mathbf{p}_i^u = (d_{n_1^u, i}, d_{n_2^u, i}, \dots, d_{n_k^u, i}) \quad (10)$$

类似地,根据 $D_{\cdot, i}$ 和 $D_{\cdot, j}$ 可得项目 $i$ 与项目 $j$ 的相似度,从而计算出项目 $i$ 的 $l$ 个最近邻项目,其索引构成的向量记为 $\mathbf{n}^i = (n_1^i, n_2^i, \dots, n_l^i)$ ,其中 $D_{\cdot, i}$ 表示矩阵 $D$ 的第 $i$ 列。项目 $i$ 的 $l$ 近邻项目的评分向量为

$$\mathbf{q}_i^u = (d_{u, n_1^i}, d_{u, n_2^i}, \dots, d_{u, n_l^i}) \quad (11)$$

最后,将 $\mathbf{p}_i^u$ 和 $\mathbf{q}_i^u$ 连接起来,得到用户项目评分协同向量

$$\mathbf{x}_i^u = [\mathbf{p}_i^u | \mathbf{q}_i^u] \in \mathbf{R}^{1 \times (k+l)} \quad (12)$$

将 $\mathbf{x}_i^u$ 作为评分值 $r_{ui}$ 的输入特征。

#### 2.1.2 用户项目评分的独热向量

考虑评分取值为1到 $d_y$ 的整数。对原始评分矩阵中 $r_{ui} \neq 0$ 的元素,采用独热(one-hot)编码将其转换为 $\mathbf{y}_i^u \in \mathbf{R}^{1 \times d_y}$ ,其第 $j$ 个分量定义如下

$$y_i^u[j] = \begin{cases} 1, & j = r_{ui} \\ 0, & \text{其他} \end{cases} \quad (13)$$

独热编码的优势是将一维的输出变换成多维输出,从而有益于网络模型的构建。

将用户项目评分协同向量 $\mathbf{x}_i^u$ 和用户项目评分的独热向量 $\mathbf{y}_i^u$ 分别作为训练数据集的输入特征和输出标签,即得到MC+BCF方法的样本集

$$T = \{(\mathbf{x}_i^u, \mathbf{y}_i^u) | (u, i) \in \Omega\} \quad (14)$$



## 2.2 宽度协同过滤

将样本集  $T$  划分成训练集  $T_1$  和测试集  $T_2$ 。对于训练集  $T_1$ , BLS 模型的输入矩阵和输出矩阵分别为

$$X = \begin{bmatrix} \mathbf{x}_{i_1}^{u_1} \\ \mathbf{x}_{i_2}^{u_2} \\ \vdots \\ \mathbf{x}_{i_{|T_1|}}^{u_{|T_1|}} \end{bmatrix} \in \mathbf{R}^{|T_1| \times (k+l)}, \quad Y = \begin{bmatrix} \mathbf{y}_{i_1}^{u_1} \\ \mathbf{y}_{i_2}^{u_2} \\ \vdots \\ \mathbf{y}_{i_{|T_1|}}^{u_{|T_1|}} \end{bmatrix} \in \mathbf{R}^{|T_1| \times d_y} \quad (15)$$

式中,  $|T_1|$  是训练样本的数量, 即对应用户项目对  $(u_1, i_1), (u_2, i_2), \dots, (u_{|T_1|}, i_{|T_1|})$ 。

将  $X$  和  $Y$  输入到 BLS 网络中, 建立宽度协同过滤模型

$$\begin{cases} \mathbf{Z}_j = \phi_j(X\mathbf{W}_{zj} + \beta_{zj}), & j = 1, 2, \dots, N \\ \mathbf{Z}^N = [\mathbf{Z}_1 | \mathbf{Z}_2 \cdots | \mathbf{Z}_N] \\ \mathbf{H}_j = \xi_j(\mathbf{Z}^N \mathbf{W}_{hj} + \beta_{hj}), & j = 1, 2, \dots, M \\ \mathbf{H}^M = [\mathbf{H}_1 | \mathbf{H}_2 \cdots | \mathbf{H}_M] \\ \mathbf{W} = [\mathbf{Z}^N | \mathbf{H}^M]^\dagger Y \end{cases} \quad (16)$$

式中:  $N, M$  分别表示映射特征组和增强特征组的数量;  $\phi_j, \xi_j$  为非线性激活函数, 在实验中选择 ReLU;  $\mathbf{W}_{zj} \in \mathbf{R}^{(k+l) \times d_z}$ 、 $\mathbf{W}_{hj} \in \mathbf{R}^{Nd_z \times d_h}$  为随机权重矩阵,  $\beta_{zj} \in \mathbf{R}^{|T_1| \times d_z}$ 、 $\beta_{hj} \in \mathbf{R}^{|T_1| \times d_h}$  为偏置矩阵;  $\mathbf{W} \in \mathbf{R}^{(Nd_z + Md_h) \times d_y}$  为可训练权重矩阵;  $d_z, d_h$  分别表示映射特征组和增强特征的维数。

假设  $\max(d_y, k+l) < Nd_z + Md_h$ 。模型 (16) 的主要计算量集中在求伪逆  $[\mathbf{Z}^N | \mathbf{H}^M]^\dagger$ , 其复杂度为  $O((Nd_z + Md_h)^2 \times (Nd_z + Md_h + |T_1|))$ 。所提出的 MC+BCF 模型的算法求解步骤如下。

1) 利用矩阵补全 (MC) 技术, 得到补全后的评分矩阵  $\mathbf{D}$ 。

2) 对于已知的每个评分元素, 充分利用邻域信息, 基于近邻规则构造相应的用户项目评分协同向量; 对输出元素执行独热编码。

3) 将用户项目评分协同向量输入到宽度协同过滤网络中, 捕获用户项目联合特征与输出编码之间的复杂函数关系。

## 2.3 评分预测

对于训练集, MC+BCF 算法可以确定宽度学习网络的所有权重矩阵和偏置矩阵。在对用户  $u$  在项目  $i$  的评分进行预测时, 先构造用户项目对  $(u, i)$  的评分协同向量  $\mathbf{X}_i^u \in \mathbf{R}^{1 \times (k+l)}$ ; 再将其输入到 BLS 模型中, 最终得到  $\hat{\mathbf{y}}_i^u = [\hat{\mathbf{Z}}^N | \hat{\mathbf{H}}^M] \mathbf{W} \in \mathbf{R}^{1 \times d_y}$ , 其中  $\hat{\mathbf{Z}}^N$  和  $\hat{\mathbf{H}}^M$  均为  $\mathbf{X}_i^u$  对应的行特征向量。记向量  $\hat{\mathbf{y}}_i^u$  的最大值所对应的下标索引为  $b$ 。考虑评分取值向量  $\mathbf{L} = (1, 2, \dots, d_y)$  时, 评分预测值为  $\hat{r}_{ui} = \mathbf{L}[b]$ , 其中  $\mathbf{L}[b]$  表示向量  $\mathbf{L}$  的第  $b$  个分量。

## 3 数值试验

先将所提出的 MC+BCF 在 3 个基准数据集上与其他 5 种基线模型的推荐性能进行比较, 这些模型分别为矩阵分解 (MF)<sup>[5]</sup>、概率矩阵分解 (probabilistic matrix factorization, PMF)<sup>[24]</sup>、神经协同过滤 (NeuCF)<sup>[10]</sup>、深度矩阵分解 (DMF)<sup>[11]</sup> 以及宽度协同过滤 (BCF)<sup>[13]</sup>。然后探讨了 MC+BCF 模型超参数对推荐性能的影响。

### 3.1 数据集简介与模型参数设置

数值试验使用公开可获取的 3 个基准数据集: MovieLens 的 ml-latest (ml-la) 数据集、ml-100k 数据集 (<https://grouplens.org/datasets/movielens>) 以及 filmtrust 数据集 (<https://www.librec.net/datasets.html>)。表 1 列出了这些数据集的相关统计信息, 其中前 2 个数据集的评分为 1~5 的整数, 第 3 个数据集的评分集合为  $\{0.5|i| = 1, 2, \dots, 10\}$ 。矩阵补全算法使用 MATLAB2017b 进行编程, 其余试验代码均使用 Python。所有试验均在 Intel Core i7-4710MQ CPU 2.50 GHz、GeForce GTX 950 M、RAM 8 GB 服务器上进行。

表 1 3 个基准数据集的相关统计信息

Table 1 Statistical information of three benchmark datasets

数据集	用户/个	项目/个	评分/条	稀疏度	评分范围/分
ml-la	610	9724	100836	0.0170	1~5
ml-100k	943	1682	100000	0.0630	1~5
filmtrust	1508	2071	35497	0.0114	0.5~5

在 MF 和 PMF 2 个模型中, 用户和项目潜在因子数均设置为 30, 且 PMF 的用户和项目的正则化参数分别设置为 0.01 和 0.001。对于 NeuCF 模型, 广义矩阵分解部分网络各层的特征维数分别为 64、32、16、8, 用户和项目的嵌入特征维数均设为 16。在 DMF 模型中, 用户和项目的隐向量维数均设置为 64, 网络层数设置为 2。对于 NeuCF 和 DMF 这 2 种深度学习模型, 批量大小均设置为 256, 学习率设为 0.0001, 选择 Adam 优化器。对于 BCF 和 MC+BCF, 将近邻用户和近邻项目的数量设置为 5, 映射特征组和增强特征组的数量均为 25, 映射特征维度和增强特征维度分别设为 10 和 15。

### 3.2 评估指标

使用均方根误差 (RMSE) 和均值绝对误差 (MAE) 来评价各模型的性能, MAE 或 RMSE 越小, 表明模型推荐性能越好。对于测试集  $T_2$ , RMSE (式中记为  $E_{\text{RMS}}$ ) 和 MAE (式中记为  $E_{\text{MA}}$ )

的计算公式如下

$$E_{\text{RMS}} = \sqrt{\frac{\sum_{(i,j) \in \Omega_2} (r_{ij} - \hat{r}_{ij})^2}{|T_2|}} \quad (17)$$

$$E_{\text{MA}} = \frac{\sum_{(i,j) \in \Omega_2} |r_{ij} - \hat{r}_{ij}|}{|T_2|} \quad (18)$$

式中:  $\Omega_2$  表示测试集中已评分的下标集合;  $\hat{r}_{ij}$  表示评分  $r_{ij}$  的预测值。

### 3.3 推荐性能比较

#### 3.3.1 误差比较

在 ml-1a、ml-100k 和 filmtrust 数据集上比较下列 3 类模型的推荐性能: 浅层矩阵分解 (MF、PMF)、深度学习 (NeuCF、DMF) 和宽度学习 (BCF、MC+BCF)。对于每个数据集, 按照 7:3 的

比例随机划分为训练集和测试集, 且各基线模型其余超参数按默认值设置。

表 2 列出了 MC+BCF 与其他 5 种基线模型在 ml-1a、ml-100k 和 filmtrust 数据集上的 MAE 和 RMSE 误差。由表 2 可以看出, 与 MF、PMF、NeuCF、DMF 和 BCF 相比, MC+BCF 的 RMSE 都有显著地降低; MC+BCF 在数据集 ml-100k 上的 MAE 稍逊于 BCF, 但明显优于 MF、PMF、NeuCF 和 DMF。与 MF、PMF、NeuCF 和 DMF 相比, BCF 和 MC+BCF 具有较好的推荐效果, 且 MC+BCF 总体上优于 BCF。这主要是因为 MC+BCF 在特征构造阶段, 先使用矩阵补全来恢复缺失数据, 再使用近邻规则构造训练样本的联合特征, 从而有效减小了缺失评分对推荐性能的影响。

表 2 MC+BCF 与基线模型的误差比较

Table 2 Error comparison of MC+BCF and baseline models

数据集	评价误差	MF	PMF	NeuCF	DMF	BCF	MC+BCF
ml-1a	MAE	0.7663	0.8015	0.7328	0.7704	0.5122	<b>0.4763</b>
	RMSE	1.0281	1.0692	0.9559	1.0130	0.8089	<b>0.7615</b>
ml-100k	MAE	0.8293	0.7547	0.8140	0.7789	0.6920	<b>0.6546</b>
	RMSE	1.0540	0.9623	1.0401	0.9919	1.0354	<b>0.9371</b>
filmtrust	MAE	0.7850	0.8375	0.7325	0.7888	<b>0.5135</b>	0.5156
	RMSE	1.0568	1.1402	0.9586	0.9954	0.8081	<b>0.7363</b>

对于表 2, 再比较 MC+BCF 和 BCF 的预测误差。对于 ml-1a 和 ml-100k 数据集, MC+BCF 的 MAE 比 BCF 分别下降了 7.01%、5.40%, RMSE 下降了 5.86%、9.49%。而在 filmtrust 数据集上, MC+BCF 的推荐误差 RMSE 也降低了 8.89%, 但其 MAE 比 BCF 增加了 0.0021。在 3 个基准数据集上, MC+BCF 的推荐性能总体上优于 BCF, 这是由于原始数据评分进行矩阵补全后增加了数据中用户和项目的信息量。综上, 前述试验结果验证了 MC+BCF 能够有效地缓解数据稀疏性问题。

#### 3.3.2 运行时间比较

表 3 列出了各算法的运行时间, 其中 MC+BCF 包括 MC 和 BCF 两部分时间之和。由表 3 可以看出, 与 BCF 相比, MC 的运行时间较短, 即

MC+BCF 在计算复杂度上稍逊于 BCF。在 ml-1a 和 filmtrust 数据集上, MC+BCF 的运行时间明显优于 MF; 在 filmtrust 数据集上, MC+BCF 的运行时间与 MF 相当。在 ml-1a 和 filmtrust 数据集上, PMF 的运行时间超过 MC+BCF 的 1.6 倍, 但在 ml-100k 数据集上 PMF 的时间更快一些。与 DMF 和 NeuCF 等深度学习的推荐算法相比, 在 filmtrust 数据集上 MC+BCF 的运行时间与 NeuCF 相当, MC+BCF 的运行时间总体上有显著的优势。

总之, 与基于矩阵分解和基于深度神经网络的模型相比, MC+BCF 往往具有较低的计算复杂度, 这是由于 BCF 算法使用基于岭回归的伪逆来求解可训练权值, 从而避开了梯度下降法或迭代更新。因此, MC+BCF 的运行效率比 MF、PMF、DMF 和 NeuCF 等模型存在显著优势。

表 3 运行时间比较

Table 3 Comparison of running time

模型	MF	PMF	DMF	NeuCF	BCF	MC+BCF
ml-1a	91.4234	123.9498	2295.0268	97.5944	45.0166	60.5441
ml-100k	67.4944	43.3328	412.8960	75.3897	32.5498	50.6041
filmtrust	32.0436	54.2636	359.2771	32.5269	22.5786	33.1259

### 3.4 超参数分析

MC+BCF 模型的超参数主要包括用户近邻数量  $k$ 、项目近邻数量  $l$ 、映射特征组数  $N$ 、映射特征

组维数  $d_e$ 、增强特征组数  $M$  和增强特征组维数  $d_h$ 。在 ml-1a、ml-100k 和 filmtrust 数据集上探讨这 6 个超参数的取值对推荐性能的影响。

### 3.4.1 用户和项目近邻数量

用户近邻数量  $k$  和项目近邻数量  $l$  是构造用户项目评分协同向量的关键。设置  $k \in \{3, 4, 5, 6\}$  和  $l \in \{3, 4, 5, 6\}$ , 考虑 MC+BCF 在  $k$  和  $l$  的 16 种组合下的测试误差。图 3 和 4 分别给出了 3 个数据集的 MAE 和 RMSE 热力图。由这 2 个图可以看出, 对于 3 个基准数据集, RMSE 和 MAE 值一般会随

着  $k$  或  $l$  的增加而小幅增大, 这可能是由于相对较大的用户和项目近邻数量会引入更多的噪声。在 ml-100k 上, 当  $k \in \{3, 4\}$  且  $l = 3$  时, RMSE 和 MAE 值较大, 这可能是由该数据集的稀疏度相对较低所导致的, 即邻域信息引入不足可能会影响预测的精度。因此, 在后面的试验中将  $k$  和  $l$  都设置为 5。

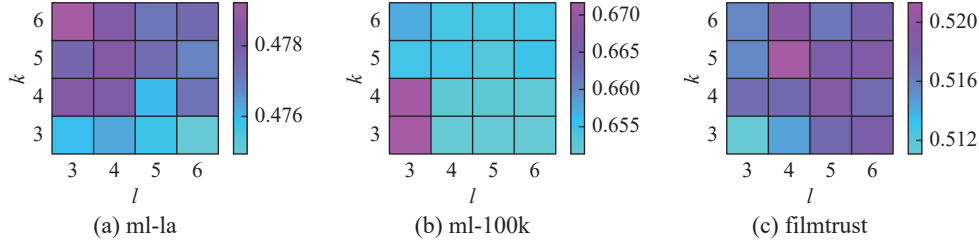


图 3 不同用户和项目近邻数量下的 MAE 热力图

Fig. 3 MAE heatmap for different number of users and item neighbors

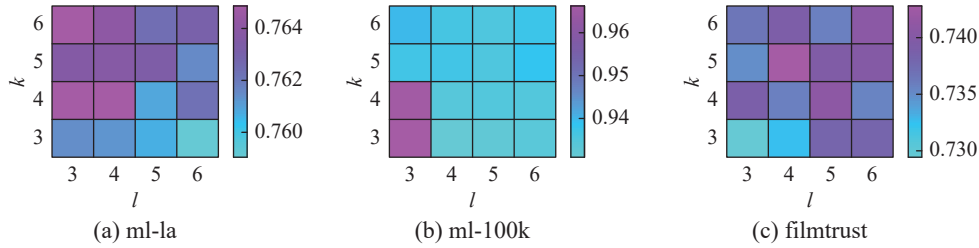


图 4 不同用户和项目近邻数量下的 RMSE 热力图

Fig. 4 RMSE heatmap for different number of users and item neighbors

### 3.4.2 映射特征组数量和映射特征组维数

特征映射模块是捕获用户项目协同向量和评分之间的非线性关系的第 1 步。设置映射特征组数量  $N \in \{15, 20, 25, 30, 35\}$  和映射特征组维数  $d_z \in \{5, 10, 15, 20, 25\}$ 。图 5 给出了不同映射特征组数量  $N$  下 3 个数据集对应的 MAE 和 RMSE。由图 5 可以看出, RMSE 和 MAE 有相同的变化趋势; 映射特征组数量  $N \geq 20$  时, 推荐误差趋于稳定, 且波动幅度均小于 0.01。对于 3 个数据集, 不同映射特征组维数  $d_z$  下的 MAE 和 RMSE 如图 6 所示。由图 6 可以发现, 随着  $d_z$  的增加, RMSE 和 MAE 的变化相对平稳, 且改变幅度非常小。图 5 和图 6 说明 MC+BCF 对超参数  $N$  和  $d_z$  相对不敏感。在试验中将  $N$  和  $d_z$  分别设置为 25 和 10。

### 3.4.3 增强特征组数量和增强特征组维数

特征增强模块是捕获用户项目协同向量和评分之间的非线性关系的第 2 步。设置增强特征组数量  $M \in \{15, 20, 25, 30, 35\}$  和增强特征组维数  $d_h \in \{5, 10, 15, 20, 25\}$ 。图 7 给出了不同增强特征组数量  $M$  下 3 个数据集对应的 MAE 和 RMSE。由图 7 可以发现, 随着  $M$  的增加, RMSE 和 MAE 的变化相对平稳, 且波动幅度均小于 0.01。对于 3 个数据集, 不同增强特征组维数  $d_h$  下的 MAE 和 RMSE 如图 8 所示。由图 8 可以看出, RMSE 和 MAE 有相同的变化趋势, 增强特征维数  $d_h \geq 10$  时, 推荐误差趋于稳定, 且波动幅度均小于 0.01。图 7 和图 8 说明 MC+BCF 对超参数  $M$  和  $d_h$  相对不敏感。在试验中将  $M$  和  $d_h$  分别设置为 25 和 10。

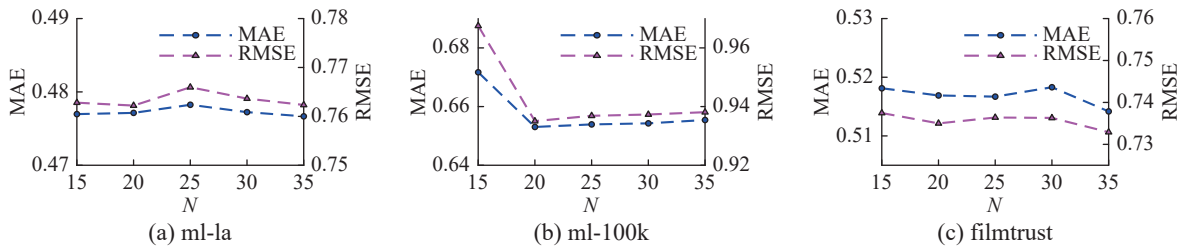


图 5 不同映射特征组数量下的 MAE 和 RMSE

Fig. 5 MAE and RMSE for different number of mapped feature groups

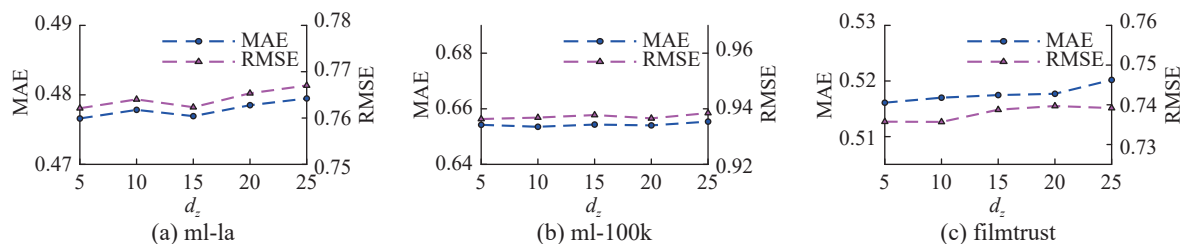


图6 不同映射特征组维数下的 MAE 和 RMSE

Fig. 6 MAE and RMSE of different dimension in mapped feature groups

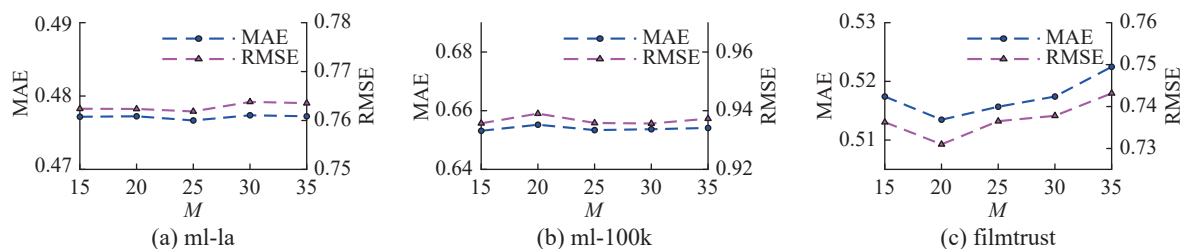


图7 不同增强特征组数量下的 MAE 和 RMSE

Fig. 7 MAE and RMSE for different number of enhanced feature groups

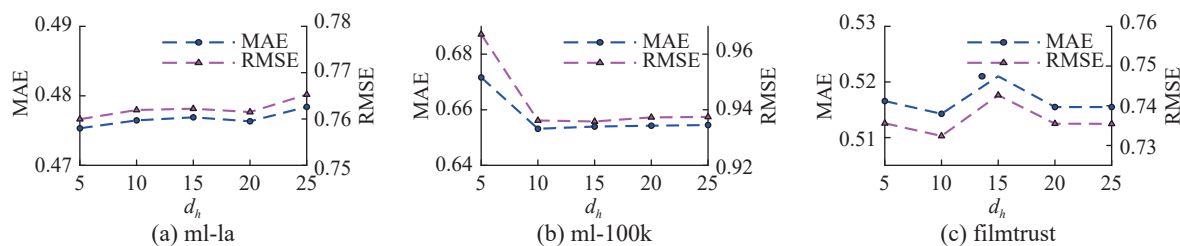


图8 不同增强特征组维数下的 MAE 和 RMSE

Fig. 8 MAE and RMSE of different dimension in enhanced feature groups

## 4 结束语

本研究提出了一种结合矩阵补全和宽度协同过滤的推荐算法。先对原始评分矩阵中的未知数据进行补全,以丰富评分数据的用户和项目的信息量。再使用近邻法则构造用户项目评分协同向量。所提出的算法充分利用了用户和项目补全后的邻域信息,进而有效缓解了数据稀疏性问题。在算法效率方面,MC+BCF也比基于深度学习和矩阵分解的推荐算法的运算效率高。在3个基准数据集上对MC+BCF的推荐性能和算法运算效率进行了验证,试验结果表明,MC+BCF显著地提升推荐性能,且算法效率也更优。

所提算法只对评分数据进行了挖掘,而评分数据的相关辅助信息(如:用户属性、历史访问时间等)对于推荐性能也至关重要。因此,在之后的研究中可以引入相关辅助信息,来提升推荐精度。

## 参考文献:

- [1] SINGH N, SINGH D P, PANT B. Big data knowledge discovery as a service: recent trends and challenges[J]. *Wireless personal communications*, 2022, 123(2): 1789–1807.
- [2] GUPTA S, DAVE M. An overview of recommendation system: methods and techniques[M]//*Advances in Computing and Intelligent Systems*. Singapore: Springer Singapore, 2020: 231–237.
- [3] ASSUNCAO W G, PICCOLO L S G, ZAINA L A M. Considering emotions and contextual factors in music recommendation: a systematic literature review[J]. *Multimedia tools and applications*, 2022, 81(6): 8367–8407.
- [4] LI Xiangpo. Research on the application of collaborative filtering algorithm in mobile E-commerce recommendation system[C]//2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers. Dalian: IEEE, 2021: 924–926.
- [5] KOREN Y, BELL R, VOLINSKY C. Matrix factorization techniques for recommender systems[J]. *Computer*, 2009, 42(8): 30–37.
- [6] 胡越, 罗东阳, 花奎, 等. 关于深度学习的综述与讨论[J]. *智能系统学报*, 2019, 14(1): 1–19.

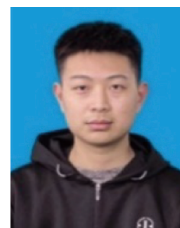


- HU Yue, LUO Dongyang, HUA Kui, et al. Overview on deep learning[J]. CAAI transactions on intelligent systems, 2019, 14(1): 1–19.
- [7] WANG Changdong, XI Wudong, HUANG Ling, et al. A BP neural network based recommender framework with attention mechanism[J]. IEEE transactions on knowledge and data engineering, 2022, 34(7): 3029–3043.
- [8] XI Wudong, HUANG Ling, WANG Changdong, et al. BPAM: recommendation based on BP neural network with attention mechanism[C]//Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence. Macao: International Joint Conferences on Artificial Intelligence Organization, 2019: 3905–3911.
- [9] DENG Zhihong, HUANG Ling, WANG Changdong, et al. DeepCF: a unified framework of representation learning and matching function learning in recommender system[EB/OL]. (2019–01–15)[2022–01–01]. <http://arxiv.org/abs/1901.04704v1>.
- [10] HE Xiangnan, LIAO Lizi, ZHANG Hanwang, et al. Neural collaborative filtering[C]//Proceedings of the 26th International Conference on World Wide Web. New York: ACM, 2017: 173–182.
- [11] XUE Hongjian, DAI Xinyu, ZHANG Jianbing, et al. Deep matrix factorization models for recommender systems[C]//Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. Melbourne: International Joint Conferences on Artificial Intelligence Organization, 2017: 3203–3209.
- [12] YIN Hongzhi, WANG Qinyong, ZHENG Kai, et al. Overcoming data sparsity in group recommendation[J]. IEEE transactions on knowledge and data engineering, 2022, 34(7): 3447–3460.
- [13] HUANG Ling, GUAN Canrong, HUANG Zhengwei, et al. Broad recommender system: an efficient nonlinear collaborative filtering approach[EB/OL]. (2022–04–20)[2022–08–29]. <https://arxiv.org/pdf/2204.11602.pdf>.
- [14] PHILIP CHEN C L, LIU Zhulin. Broad learning system: an effective and efficient incremental learning system without the need for deep architecture[J]. IEEE transactions on neural networks and learning systems, 2018, 29(1): 10–24.
- [15] FLETCHER K K. A method for dealing with data sparsity and cold-start limitations in service recommendation using personalized preferences[C]//2017 IEEE International Conference on Cognitive Computing. Honolulu: IEEE, 2017: 72–79.
- [16] CANDLES E J, RECHT B. Exact low-rank matrix completion via convex optimization[C]//2008 46th Annual Allerton Conference on Communication, Control, and Computing. Monticello: IEEE, 2009: 806–812.
- [17] 陈蕾, 陈松灿. 矩阵补全模型及其算法研究综述[J]. 软件学报, 2017, 28(6): 1547–1564.
- CHEN Lei, CHEN Songcan. Survey on matrix completion models and algorithms[J]. Journal of software, 2017, 28(6): 1547–1564.
- [18] WANG Weigang, SONG Wei, WANG Guangyuan, et al. Image recovery and recognition: a combining method of matrix norm regularisation[J]. IET image processing, 2019, 13(8): 1246–1253.
- [19] CHEN Zhaoliang, WANG Shiping. A review on matrix completion for recommender systems[J]. Knowledge and information systems, 2022, 64(1): 1–34.
- [20] LIN Zhouchen, CHEN Minming, MA Yi. The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices[EB/OL]. (2010–09–26)[2022–08–29]. <https://arxiv.org/abs/1009.5055.pdf>.
- [21] BIOUCAS-DIAS J M, FIGUEIREDO M A T. A new twist: two-step iterative shrinkage/thresholding algorithms for image restoration[J]. IEEE transactions on image processing, 2007, 16(12): 2992–3004.
- [22] PAO Y H, PARK G H, SOBAJIC D J. Learning and generalization characteristics of the random vector functional-link net[J]. Neurocomputing, 1994, 6(2): 163–180.
- [23] YANG Qing, GU Yudi, WU Dongsheng. Survey of incremental learning[C]//2019 Chinese Control and Decision Conference. Nanchang: IEEE, 2019: 399–404.
- [24] SALAKHUTDINOV R, MNH A. Probabilistic matrix factorization[C]//Proceedings of the 20th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2008: 1257–1264.

#### 作者简介:



史加荣, 教授, 博士, 主要研究方向为机器学习与深度学习。承担过国家自然科学基金、中国博士后科学基金各 1 项。发表学术论文 40 余篇。E-mail: shijiarong@xauat.edu.cn。



何攀, 硕士研究生, 主要研究方向为机器学习、数据挖掘。E-mail: 1377713953@qq.com。