

基于二分图卷积表示的推荐算法



熊旭东¹ 杜圣东^{1,2,3} 夏琬钧¹ 李天瑞^{1,2,3}

1 西南交通大学信息科学与技术学院 成都 611756

2 西南交通大学人工智能研究院 成都 611756

3 综合交通大数据应用技术国家工程实验室 成都 611756

(asia123dong@126.com)

摘要 随着数据驱动智能技术的快速发展,个性化推荐算法及相关应用成为了研究热点。推荐可视为将用户与物品进行匹配的问题,但用户与物品之间存在的语义差距不便于两者之间的直接匹配。现有的许多基于深度学习的推荐算法采用的思路都是将不同空间中的实体映射到统一潜在语义空间,利用其嵌入表示来进行匹配度计算。随着网络表示学习方法的出现,由于用户和物品的交互可构成二分图,用户和物品的嵌入表示可被视作二分图节点表示,许多基于二分图节点表示的推荐算法被提出,但现有算法仍难以对高阶交互信息进行有效提取。针对这一问题,文中提出了一种基于二分图卷积表示学习的推荐算法BGCRRA(Bipartite Graph Convolution Representation-based Recommendation Algorithm)。该算法首先将用户和物品交互视作二分图,然后通过实现自适应融合多阶、多层次的图卷积模型来对节点进行嵌入表示,最后计算用户和物品的匹配度,并实现推荐。文中在3个公开的数据集上进行对比实验,通过将该算法与当前表现优异的算法进行HR和NDCG(Normalized Discounted Cumulative Gain)指标的比较分析,验证了所提推荐算法的有效性。

关键词: 推荐算法;嵌入方法;图卷积;二分图

中图法分类号 TP391

Recommendation Algorithm Based on Bipartite Graph Convolution Representation

XIONG Xu-dong¹, DU Sheng-dong^{1,2,3}, XIA Wan-jun¹ and LI Tian-rui^{1,2,3}

1 School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China

2 Institute of Artificial Intelligence, Southwest Jiaotong University, Chengdu 611756, China

3 National Engineering Laboratory of Integrated Transportation Big Data Application Technology, Chengdu 611756, China

Abstract With the rapid development of data-driven intelligent technology, personalized intelligent recommendation algorithms and related applications become research hotspots. Recommendation can be regarded as a matching problem between users and items. As the semantic gap between users and items, it's inconvenient to match them directly. Many existing recommendation methods based on deep learning use the idea of mapping entities from different spaces into a unified semantic space to calculate the matching degree by embedding representation. With the emergence of network representation learning, the bipartite graph can be formed between users and items and the embedding representation in the recommendation algorithm can also be regarded as nodes representation in bipartite graph. Many recommendation algorithms based on bipartite graph nodes representation have been proposed. However, existing algorithms are still hard to extract high-order interactive information effectively. To solve this problem, a bipartite graph convolution representation-based recommendation algorithm(BGCRRA) is developed in this paper. Interactions between users and items are regarded as a bipartite graph at first, nodes are represented by adaptively fusing multi-order and multi-level graphs secondly, and finally the matching degree of users and items is calculated and the recommendation is realized. Comparative experiments are carried out on 3 open datasets and the effectiveness of the proposed algorithm is verified by comparing HR and NDCG(Normalized Discounted Cumulative Gain) indexes of our algorithm and the state-of-the-art algorithms.

Keywords Recommendation algorithms, Embedding methods, Graph convolution, Bipartite graph

1 引言

在大数据时代,互联网中的数据呈爆炸式增长,不仅用户

面临信息过载(Information Overload)的问题,传统的信息检索技术也面临着一定的瓶颈^[1-2]。因此,个性化推荐作为一种优化方案,被广泛研究并应用于诸多领域^[2],包括国内外各大

到稿日期:2020-04-07 返修日期:2020-09-20 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家重点研发计划(2017YFB1401400)

This work was supported by the National Key R&D Program of China(2017YFB1401400).

通信作者:李天瑞(trli@swjtu.edu.cn)

电商平台(如京东、阿里巴巴、Amazon等)、社交网络平台(如腾讯、微博、Twitter等)、新闻推送平台(如今日头条、一点资讯、GroupLens等^[3])。

推荐指利用现有知识,学习出特定规则,主动向用户提供物品或信息建议的过程^[2]。自明尼苏达大学 GroupLens 研究组提出基于协同过滤的新闻推荐系统后,协同过滤作为一种主要的推荐技术,被广泛应用于各大领域^[4],其基本思想是通过用户和物品之间的交互信息来挖掘出用户的潜在偏好以进行推荐^[3]。推荐系统中用户与物品之间的交互联系可以被看作是一种二分图,如图 1 所示。基于二分图思想,Zhou 等^[5-6]提出了基于复杂网络理论的二分图的推荐算法,该算法相比传统的协同过滤算法具有更高的推荐精度。此后,基于二分图的物质扩散和热传导的推荐模型及其优化算法被提出^[6],Zhou 等^[4]证明了协同过滤可被视为二分图推荐的一种特例,并论述了基于二分图推荐算法的有效性。

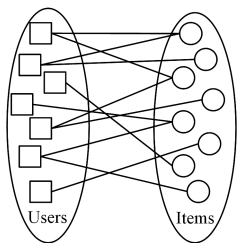


图 1 用户物品二分图示意^[11]

Fig. 1 Bipartite graph of users and items^[11]

随着深度学习(Deep Learning)技术的快速发展,基于深度神经网络的推荐算法被陆续提出,其由于具有自动学习特征和提取信息的能力,成为了推荐算法的研究热点^[3]。卷积神经网络(Convolutional Neural Networks,CNN)被广泛应用于图像处理、语义分析、机器翻译等领域,且随着复杂网络相关理论的发展,提取非欧氏空间结构信息的图卷积网络(Graph Convolutional Network,GCN)被提出^[7-9]。由于传统基于二分图的推荐算法只针对具体场景做出优化^[6],而图卷积网络能直接从二分图中提取用户和物品的交互信息,得到用户和物品的嵌入表达,进而进行智能匹配,因此图卷积网络被广泛地应用于推荐系统^[10-16]。

现有的基于图卷积的推荐算法因图卷积网络的特点,存在难以对高阶交互信息进行有效提取的问题^[9-11],因此本文提出了一种基于二分图卷积表示的推荐算法 BGCRRA,用于对用户物品交互数据构建的二分图进行多阶、多层次的卷积操作,以得到用户和物品的嵌入表达,进而计算匹配度以完成推荐。本文的主要贡献如下:1)给出了 GCN 在二分图上的两部分交替计算的形式,并利用神经网络来实现;2)在进行嵌入表达时,自适应融合用户与物品之间的高阶隐藏交互;3)通过多层次的交互信息融合增强了模型的表达能力;4)在 LastFM^[17]、MovieLens-1m^[17-18]和 Pinterest^[18]这 3 个公开数据集上进行对比实验,实验结果验证了所提算法的有效性。

2 相关工作

2.1 基于深度学习的推荐算法

当前基于深度学习的推荐算法大都通过神经网络来学习用户和物品的嵌入表达,再进行匹配推荐^[10],其模型结构如

图 2 所示。基于深度学习的推荐算法有两类典型的研究思路。

(1)利用用户和物品的自身属性信息来获取各自的嵌入表示。如 Elkahky 等^[19]考虑到了在传统的推荐系统中用户特征难以获取的问题,通过分析用户行为特征,提出了一种利用多视角的深度神经网络模型(Multi-View Deep Neural Network,MVDNN)来获取用户嵌入表示并进行匹配推荐的算法。Covington 等^[20]利用用户信息、情景信息等多源异构信息,提出了一种基于嵌入表示的用于 YouTube 视频推荐的深度神经网络模型。Xu 等^[21]分别通过用户标签和目标对象特征潜在空间表示来产生推荐。Cheng 等^[22]根据用户特征和目标特征,结合线性模型和深度模型,提出了基于嵌入表示的 Wide&Deep 推荐模型。

(2)利用用户和物品的交互信息来获取各自的嵌入表示。如 Xue 等^[23]提出了利用多层感知机来代替传统矩阵分解以获取交互信息,进而得到用户和物品嵌入的算法 DMF。He 等^[24]于 2017 年提出了结合矩阵分解和多层感知机得到嵌入表示的 NeuMF。Xue 等^[18]提出了通过交互信息来得到物品嵌入并间接表示用户的 DeepICF 算法。Deng 等^[17]提出了通过综合交互信息表示和匹配函数学习来进行推荐的 DeepCF 算法。

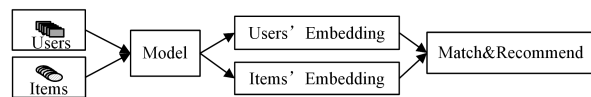


图 2 基于深度学习推荐模型的结构示意图

Fig. 2 Structure of recommendation model based on deep learning

2.2 基于图卷积的推荐算法

利用图卷积可以有效地获取图结构的节点嵌入表示。图卷积一般有两种,一种是根据图的拓扑结构,聚合相邻节点属性信息的空域图卷积^[12,25],另一种是利用傅里叶变换与拉普拉斯特征映射得到谱信息,在谱域上进行运算的谱域图卷积^[8-9]。本文基于谱域图卷积来开展推荐算法的研究。

基于图卷积的推荐算法的主要思路是,通过图卷积来获取图中用户和物品节点的嵌入表达以进行匹配推荐。如 SpectralCF^[12]通过图卷积方法来学习用户物品连接的谱域信息,从而进行推荐。SocialGCN^[14]利用图卷积来学习用户与其他用户的偏好,经过社交系统的扩散过程并建模来进行推荐。DiffNet^[15]通过图卷积来学习用户和物品的迭代影响以进行推荐。PinSage^[16]通过图卷积来对邻居节点进行采样,并通过信息聚合的方式实现针对大规模数据的商业推荐。

此外,针对用户与物品属性信息不足的情况,基于图卷积的推荐算法主要探究如何对交互信息本身进行充分提取。Burg 等^[11]提出的 GCMC 首先通过图卷积方法来提取交互代表的边信息,然后进行用户物品的嵌入表达和推荐。Wang 等^[10]提出基于图卷积的 NGCF 模型来提取并拼接用户和物品高阶交互信息,以获得用户和物品嵌入表达。

由图 2 可知,本文提出的基于二分图卷积表示学习的推荐算法 BGCRRA,利用图卷积来获取用户物品交互信息以得到嵌入表达,进而进行推荐。针对 NGCF^[10]和 GCMC^[11]没有考虑到针对高阶交互信息进行有效提取和融合的问题,BGCRRA 考虑到了不同阶交互信息对用户和物品表达的影

响不同,提出了自适应结构来学习不同阶卷积表达的权重。此外,BGCRRA参考注意力机制中的多头机制^[26-27],融合多层次卷积信息来扩展模型宽度,进一步增强模型的表达能力。

3 算法介绍

为对用户和物品进行有效嵌入表示,本文提出的算法BGCRRA首先将用户与物品交互构成的二分图通过优化的图卷积操作得到节点表达,然后进行匹配计算,最后完成推荐任务。

3.1 用户和物品的二分图定义

对于推荐中的数据,可将用户和物品之间的隐式交互定义为一个 $m \times n$ 的矩阵 \mathbf{M} ,其中 m 表示用户数量, n 表示物品数量,其元素满足:

$$y_{u,i} = \begin{cases} 1, & \text{用户 } u \text{ 和物品 } i \text{ 存在交互} \\ 0, & \text{用户 } u \text{ 和物品 } i \text{ 不存在交互} \end{cases} \quad (1)$$

为定义统一语义下的邻接矩阵以便于计算,本文参照文献^[11]中的方法,将交互矩阵定义为一个无向图,并以扩展邻接矩阵的形式进行表达,即将 $\mathbf{M} \in \mathbb{R}^{m \times n}$ 转化为扩展的邻接矩阵 $\mathbf{A} \in \mathbb{R}^{(m+n) \times (m+n)}$,具体如式(2)所示:

$$\mathbf{A} = \begin{bmatrix} \mathbf{0}_{m \times m}, & \mathbf{M} \\ \mathbf{M}^T, & \mathbf{0}_{n \times n} \end{bmatrix} \quad (2)$$

矩阵 \mathbf{A} 可被视为一个由交互信息构成的图。由于这个图中的节点具有用户和物品这两种不同的性质,且仅考虑用户和物品之间的交互,故该图为一个二分图。由式(1)的定义可知,图 \mathbf{A} 中有连接的节点才表示有交互,这与一般图的定义一致,因此可直接使用图神经网络相关计算方法来对其进行嵌入表示等操作。使用图卷积对交互矩阵节点进行表达可得到有效的结果^[10,12,14-16],因此BGCRRA对二分图进行图卷积操作来实现用户和物品的嵌入表达。

由式(2)的定义可知,若直接进行图卷积操作,用户和物品的自身信息的表达则不够充分,会限制交互信息的提取从而导致结果不理想。因此,本文在二分图中参照同构图卷积,加入自连接信息,令二分图的邻接矩阵 \mathbf{A} 的对角线值为1,即:

$$\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I} \quad (3)$$

其中, $\mathbf{I} \in \mathbb{R}^{(m+n) \times (m+n)}$ 为单位矩阵, $\tilde{\mathbf{A}}$ 为具有节点自相连的二分图邻接矩阵,与同构图的邻接矩阵形式一致,可直接进行图卷积计算。

3.2 二分图上的图卷积定义

Defferrard等^[8]提出了基于切比雪夫公式的谱域图卷积操作,如式(4)所示:

$$\mathbf{X}_{l+1} = \sigma \left(\sum_{k=0}^{K-1} \alpha_k \mathbf{L}^k \mathbf{X}_l \right) \quad (4)$$

其中, \mathbf{X}_l 表示第 l 层的输出, α_k 表示可训练系数, $\mathbf{L} = \mathbf{I} - \mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}}$ 表示图的拉普拉斯矩阵, \mathbf{L}^k 是图的拉普拉斯矩阵的 k 次方, $\sigma(\cdot)$ 是激活函数。Kipf等^[9]利用近似切比雪夫公式的方式进一步简化了基于谱域的卷积方法,并将其定义为:

$$\mathbf{H}^{(l+1)} = \sigma \left(\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{H}^{(l)} \mathbf{W}^{(l)} \right) \quad (5)$$

其中, $\mathbf{H}^{(l+1)}$ 表示第 $l+1$ 层输出的特征矩阵, $\tilde{\mathbf{A}}$ 表示网络结构的邻接矩阵 \mathbf{A} 所对应的自连接矩阵,即式(3)中的定义,此处

\mathbf{A} 的元素与第3.1节中式(2)所示的二分图中的元素具有相同的含义, $\tilde{\mathbf{D}}$ 表示 $\tilde{\mathbf{A}}$ 对应的度矩阵, $\mathbf{W}^{(l)}$ 表示第 l 层对应的权重矩阵, $\sigma(\cdot)$ 表示激活函数。

若按照式(5)进行卷积操作,由式(2)和式(3)仅得到一个隐藏表达 $\mathbf{H}^{(l+1)}$,这不便于后续优化,且随着用户和推荐物品数量的增多,会出现参数量非线性增加,进而限制计算速度和实用性。

因此,将式(5)进行变换,由式(2)可知, \mathbf{A} 可以看作是由矩阵 \mathbf{M} 及其转置 \mathbf{M}^T 和两个单位对角矩阵构成的,故将式(5)进行如下变换:

$$\begin{aligned} \mathbf{H}^{(l+1)} &= \sigma \left(\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{H}^{(l)} \mathbf{W}^{(l)} \right) \\ &= \sigma \left(\tilde{\mathbf{D}}^{-\frac{1}{2}} \left(\begin{bmatrix} \mathbf{0}_{m \times m}, & \mathbf{M} \\ \mathbf{M}^T, & \mathbf{0}_{n \times n} \end{bmatrix} + \mathbf{I} \right) \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{H}^{(l)} \mathbf{W}^{(l)} \right) \\ &= \sigma \left(\text{normal} \left(\begin{bmatrix} \mathbf{I}_{m \times m}, & \mathbf{M} \\ \mathbf{M}^T, & \mathbf{I}_{n \times n} \end{bmatrix} \right) \left(\begin{bmatrix} \mathbf{H}_u^{(l)} \\ \mathbf{H}_i^{(l)} \end{bmatrix} \right) \mathbf{W}^{(l)} \right) \\ &= \sigma \left(\begin{bmatrix} \text{normal}(\mathbf{M} \mathbf{H}_i^{(l)} + \mathbf{H}_u^{(l)}) \\ \text{normal}(\mathbf{M} \mathbf{H}_u^{(l)} + \mathbf{H}_i^{(l)}) \end{bmatrix} \right) \mathbf{W}^{(l)} \end{aligned} \quad (6)$$

其中, $\mathbf{W}^{(l)}$ 为可训练的参数,由于矩阵 $\begin{bmatrix} \mathbf{0}_{m \times m}, & \mathbf{M} \\ \mathbf{M}^T, & \mathbf{0}_{n \times n} \end{bmatrix}$ 是维度为 $(m+n) \times (m+n)$ 的一个稀疏矩阵,将其进行拆分。为了增加模型的表达能力,将 $\mathbf{W}^{(l)}$ 拆分到两个不同的参数矩阵中,用户的隐藏层矩阵表示为:

$$\mathbf{H}_u^{(l+1)} = \sigma \left((\tilde{\mathbf{M}} \mathbf{H}_u^{(l)} + \mathbf{H}_i^{(l)}) \mathbf{W}_u + \mathbf{b}_u \right) \quad (7)$$

同理,物品的隐藏层矩阵表示为:

$$\mathbf{H}_i^{(l+1)} = \sigma \left((\tilde{\mathbf{M}}^T \mathbf{H}_u^{(l)} + \mathbf{H}_i^{(l)}) \mathbf{W}_i + \mathbf{b}_i \right) \quad (8)$$

其中, $\tilde{\mathbf{M}}$ 表示归一化之后的交互矩阵, $\mathbf{W}_u, \mathbf{W}_i, \mathbf{b}_u, \mathbf{b}_i$ 分别表示用户输入 u 和物品输入 i 对应的权重和偏置。

3.3 模型结构

根据式(7)和式(8)的隐藏层表示,构造一个堆叠的基于二分图的多层图卷积神经网络模型,如式(9)所示:

$$\begin{aligned} \left(\begin{aligned} \mathbf{H}_u^{(0)} &= \sigma \left((\tilde{\mathbf{M}} \mathbf{X}_u + \mathbf{X}_u) \mathbf{W}_u^{(0)} + \mathbf{b}_u^{(0)} \right) \\ \mathbf{H}_i^{(0)} &= \sigma \left((\tilde{\mathbf{M}}^T \mathbf{X}_u + \mathbf{X}_i) \mathbf{W}_i^{(0)} + \mathbf{b}_i^{(0)} \right) \\ &\dots \end{aligned} \right) \\ \left(\begin{aligned} \mathbf{H}_u^{(l)} &= \sigma \left((\tilde{\mathbf{M}} \mathbf{H}_i^{(l-1)} + \mathbf{H}_u^{(l-1)}) \mathbf{W}_u^{(l)} + \mathbf{b}_u^{(l)} \right) \\ \mathbf{H}_i^{(l)} &= \sigma \left((\tilde{\mathbf{M}}^T \mathbf{H}_u^{(l-1)} + \mathbf{H}_i^{(l-1)}) \mathbf{W}_i^{(l)} + \mathbf{b}_i^{(l)} \right) \end{aligned} \right) \\ \mathbf{Y} = \text{match}(\mathbf{H}_u^{(l)}, \mathbf{H}_i^{(l)}) \end{aligned} \quad (9)$$

其中, $\mathbf{X}_u, \mathbf{X}_i$ 分别表示用户和物品的特征矩阵。由式(9)可知,此处要求用户特征维度和物品数量维度一致,若在真实数据中数据输入的维度不一致,则需要通过一层或多层线性或者非线性变换来实现。函数 $\text{match}(\cdot)$ 表示匹配函数,当前基于深度学习的推荐方案通常通过学习潜在语义空间表示向量的相似度来进行匹配,如式(10)所示,也可以通过多个全连接神经网络来实现^[17],即:

$$\begin{aligned} \mathbf{Y} &= \text{match}(\mathbf{H}_u^{(l)}, \mathbf{H}_i^{(l)}) \\ &= \sigma \left((\mathbf{H}_u^{(l)} \odot \mathbf{H}_i^{(l)}) \mathbf{W}_{\text{output}} + \mathbf{b}_{\text{output}} \right) \\ &= \text{sigmoid} \left((\mathbf{H}_u^{(l)} \odot \mathbf{H}_i^{(l)}) \mathbf{W}_{\text{output}} + \mathbf{b}_{\text{output}} \right) \end{aligned} \quad (10)$$

其中, \odot 表示矩阵或者向量对应元素相乘, $\mathbf{W}_{\text{output}}$ 和 $\mathbf{b}_{\text{output}}$ 分别表示输出层的权重和偏置。

对于主要包含交互信息的推荐数据集,用户和物品不一

定都存在属性特征,若用户和物品其中有一个不存在或者都不存在属性特征,则一般采用在式(9)中分别用单位对角矩阵 $\mathbf{I}_i, \mathbf{I}_u$ 来代替对应的当前用户和物品的特征 $\mathbf{X}_i, \mathbf{X}_u$ 。

3.3.1 多阶交互信息融合

GCN本身可以被看作是一种信息传递获取的模型^[27]。若在式(10)中直接使用 $\mathbf{H}_u^{(l)}, \mathbf{H}_i^{(l)}$ 作为输出,则最终结果容易被图中第 l 阶传递的信息所覆盖,或模型更加侧重于学到图中的连接信息在经过 l 次传递之后的结果。通常, n 阶图卷积获取的是,在整个图中每个节点信息经过 n 次传递后产生的对其他节点的影响信息。不同次数的传递信息表达的意义不同,因此为了更好地获取高阶信息,本文提出了一种自适应学习权重并获取多阶信息的方案,其方法如式(11)所示:

$$\mathbf{H}^{\text{combine}} = \sum_{\text{by_column}} \left(\frac{\exp \left(\begin{bmatrix} \mathbf{H}^{(0)} \\ \vdots \\ \mathbf{H}^{(l)} \end{bmatrix} \mathbf{W}_s \right)}{\sum_{\text{by_row}} \left(\exp \left(\begin{bmatrix} \mathbf{H}^{(0)} \\ \vdots \\ \mathbf{H}^{(l)} \end{bmatrix} \mathbf{W}_s \right) \right)} \right) \cdot \begin{bmatrix} \mathbf{H}^{(0)} \\ \vdots \\ \mathbf{H}^{(l)} \end{bmatrix}^T \quad (11)$$

其中, \mathbf{W}_s 是一个可学习的参数矩阵,函数 $\exp(\cdot)$ 表示 e 为

底的指数函数, $\text{sum}_{\text{by_column}}(\cdot)$ 表示对矩阵按列进行求和, $\text{sum}_{\text{by_row}}(\cdot)$ 表示对矩阵按行进行求和。

本文方案通过一个可自适应学习的权重来整合结构,学习不同阶图卷积权重,并根据权重融合不同阶图卷积信息来得到最终结果。利用此方法分别对用户和物品的所有隐藏层矩阵进行式(11)所示的操作,可得到新的 $\mathbf{H}_u^{\text{combine}}$ 和 $\mathbf{H}_i^{\text{combine}}$, 然后可将其代入式(10)。

3.3.2 多层次信息融合

第3.3.1节提到融合多阶卷积信息可以提高模型提取交互信息的能力,但在深度学习中基于随机梯度下降的优化方法有时会出现局部最优的情况。因此,为了提高模型的泛化能力,本文参考多头注意力机制的思想^[26],引入类似注意力机制的多头模型,即输出多个 H_u^{combine} 和 H_i^{combine} 的多层次信息,然后通过拼接融合的方式来将其整合,如式(12)所示:

$$\begin{aligned} \mathbf{H}_u^{\text{combine}} &= \text{concatenate}([\mathbf{H}_{u,1}^{\text{combine}}, \mathbf{H}_{u,2}^{\text{combine}}, \dots, \mathbf{H}_{u,n}^{\text{combine}}]) \\ \mathbf{H}_i^{\text{combine}} &= \text{concatenate}([\mathbf{H}_{i,1}^{\text{combine}}, \mathbf{H}_{i,2}^{\text{combine}}, \dots, \mathbf{H}_{i,n}^{\text{combine}}]) \end{aligned} \quad (12)$$

其中, $\text{concatenate}(\cdot)$ 为向量拼接函数。

最终算法对应的模型结构如图3所示,首先通过交互数据构建二分图,然后经过自适应融合多阶、多层次卷积结果,得到用户和物品的嵌入表达,最后通过函数进行匹配预测。

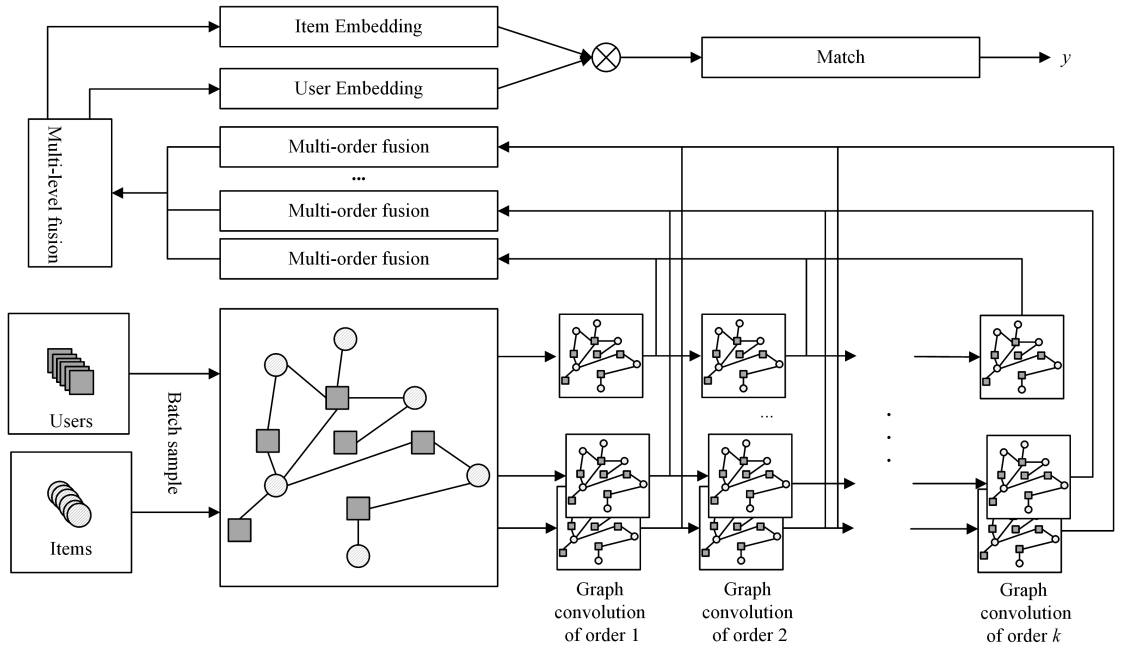


图3 本文模型的主要结构

Fig. 3 Main structure of proposed model

4 实验评测与分析

为了验证 BGCRRA 的有效性,本文基于 Keras 2.2.4 框架完成实验,使用的语言为 Python3.6,操作系统为 Ubuntu 16.04.1。服务器硬件配置为:11 GB 显存 NVIDIA GeForce RTX 2080Ti GPU, Intel(R) Core(TM) i9-7920X CPU, 32 GB 内存。

4.1 数据集介绍

本文实验使用了3个公开数据集对算法进行评估,其分别为音乐社区数据集 Last-fm¹⁾、社交网站数据集 Pinterest²⁾、电影评分数据集 MovieLens-1m³⁾,数据集的统计情况如表1所列。3个数据集均按照式(1)的定义使用隐式交互作为交互信息。

¹⁾ <http://www.dtic.upf.edu/~ocelma/MusicRecommendationDataset/>

²⁾ <https://sites.google.com/site/xueatalpha/academic-projects>

³⁾ <https://grouplens.org/datasets/movielens/>

表1 数据集统计

Table 1 Statistics of datasets

Dataset	Lastfm	MovieLens-1m	Pinterest
Users	1740	6040	9916
Items	2665	3706	55187
Ratings	69149	1000209	1500809
Sparsity	0.9851	0.9553	0.9973

4.2 基准对比算法

本文针对以下几种算法分别进行了实验,通过对比来验证本文所提算法的效果。

(1)eALS^[28]:将所有未观察到交互的用户物品对作为负实例,并按物品受欢迎程度对它们进行非均匀加权。

(2)DMF^[29]:基于矩阵分解的算法,利用深度学习来进行矩阵分解。

(3)NeuMF^[24]:结合隐函数的匹配矩阵分解算法,通过结合深度模型 MLP 和浅层模型 GMF 来进行推荐。

(4)DeepICF^[18]:结合深度学习和协同过滤的算法,利用物品嵌入向量的组合来表示用户嵌入向量并进行推荐。

(5)DeepCF^[17]:结合深度匹配和深度嵌入表达的推荐算法。

4.3 实验参数及评价指标

本文对比所用的神经网络类的基准算法,其超参数(如神经元数量、网络层数、学习率和最大迭代次数)均使用原文献中的默认值。本文算法使用 Adam 作为模型优化器,设定学习率为 0.0001,使用的 batch size 为 8192。

本实验可被定义为一个二分类问题实验,因此最小化目标即损失函数为:

$$\text{loss} = -[y_i \cdot \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i)] \quad (13)$$

其中, y_i 表示给定用户和物品之间是否真实存在交互,取值范围为 $\{0, 1\}$, \hat{y}_i 表示模型预测值。

为测试本文算法对用户与候选推荐物品的交互概率,即匹配度的预测表现,实验设定每一个真实用户和物品的交互为正样本,参照文献[17-18],每个正样本采集 99 个负样本。然后根据算法输出评分进行对比实验,实验选用了两个评价指标,即 $HR@K$ 和 $NDCG@K$,其定义如式(14)和式(15)所示。

$$HR@K = \frac{|\text{hits}@k|}{|\text{Samples}|} \quad (14)$$

其中,分子表示在所有排序中前 K 个结果命中真实结果的个数,分母表示测试样本的总个数。

$$NDCG@K = \frac{DCG@K}{IDCG@K} \quad (15)$$

$NDCG@K$ 为标准化的排序评价指标,表示对预测结果的 $DCG@K$ 和理想结果的 $DCG@K$ 的比值。其中 $DCG@K$ 的定义如式(16)所示:

$$DCG@K = \sum_{i=1}^k \frac{2^{rel_i} - 1}{\log(i+1)} \quad (16)$$

其中, rel_i 表示当前排序为 i 的输出结果。

4.4 实验结果对比分析

4.4.1 只包含交互信息实验

本节进行了只包含交互信息的实验,即针对只包含交互信息的场景进行了实验。除 eALS 外,其他对比算法均参照

文献[17-18]的设置,在训练集中针对每个正样本默认采样 5 个负样本。本实验采用留一法进行验证, $HR@10$ 越大表明针对每个测试用户,在 100 个测试物品样本中,正样本被预测的交互概率排在前 10 的能力越强; $NDCG@10$ 越大表明正样本被预测的交互概率的排名越靠前。实验结果如表 2 所列。

表2 只含有交互信息的实验结果

Table 2 Experimental results with only interactive information included

Dataset	Lastfm		MovieLens-1m		Pinterest	
	HR@10	NDCG@10	HR@10	NDCG@10	HR@10	NDCG@10
eALS	0.8268	0.5201	0.6906	0.4195	0.8659	0.5429
DMF	0.8848	0.5836	0.6721	0.3756	0.8639	0.5407
NeuMF	0.8862	0.6012	0.7230	0.4375	0.8736	0.5512
DeepICF	0.8850	0.6008	0.6879	0.4135	0.8794	0.5631
DeepCF	0.8904	0.6136	0.7225	0.4356	0.8675	0.5638
BGCRRRA	0.8921	0.6203	0.7249	0.4313	0.8781	0.5678

相比 NeuMF, BGCRRRA 在 MovieLens-1m 数据集上的 $NDCG@10$ 指标相对较差。NeuMF 主要通过结合浅层矩阵分解模型和深层 DNN 模型来对用户和物品之间的交互进行建模,因此可以较好地平衡用户和物品的高低阶交互信息。而对于高阶图卷积,其在学习高阶交互的过程中可能会学到额外的交互信息,对于正样本的排序有影响。

DeepICF 通过聚合物品嵌入向量来表示用户嵌入向量,在用户物品交互较少的数据集(如 Pinterest,其稀疏度为 0.9973,如表 1 所列)上对用户物品的匹配计算相比其他依赖交互传递的算法更有优势。在实验中,DeepICF 的 $HR@10$ 取得了最好的结果,这也说明本文方法对于交互信息较稀疏的二分图信息获取存在局限性。

eALS 是一种经典的基于矩阵因子分解的算法,相比基于神经网络的算法,其交互预测的能力相对较弱。DMF 方法虽基于神经网络,但其核心是利用神经网络来代替矩阵分解,并且依赖于低阶交互关系,对高阶交互信息的提取能力不足。

DeepCF 是在 NeuMF 的基础上,将其中融合的矩阵进行分解,并与 DNN 模型进行泛化,用神经网络来代替矩阵分解,因此其表现较好。

综上所述, BGCRRRA 考虑到了用户和物品交互信息所构成的二分图,基于同构图的 GCN 得到二分图卷积,从而获取用户和物品之间的交互信息。同时,其还考虑到了用户和物品之间本身存在的隐性高阶交互,通过自适应来融合多阶交互信息。实验结果表明, BGCRRRA 在 3 个数据集上都取得了超过或者接近 State-of-the-art 算法的效果,这表明了 BGCRRRA 的有效性。

4.4.2 有无属性特征对比实验

若用户和物品包含属性特征,则式(9)中矩阵 \mathbf{X}_u 表示用户的特征矩阵(如用户年龄、性别等), \mathbf{X}_i 表示物品的特征矩阵(如物品类别、用途等)。图卷积网络将得到包含用户与物品交互信息以及各自的属性特征的嵌入表达,相比只有交互信息的嵌入表达,该表达包含了更多的信息,得到了更好的推荐表现。本文在 MovieLens-1m 数据集上,针对无特征交互数据和包含用户物品特征的交互数据进行了对比实验。用户

特征包含用户性别、年龄、职业,物品特征包含电影类别,使用 One-hot 进行编码。实验中,参数和训练目标与第 4.4.1 节相同,实验结果对比如表 3 所列。

表 3 在 MovieLens-1m 数据集上有无特征对比实验结果

Table 3 Comparisons of experimental results on MovieLens-1m dataset with and without features

Measures	MovieLens-1m without features	MovieLens-1m with features
HR@10	0.7249	0.7336
NDCG@10	0.4313	0.4410

由表 3 可知,包含属性特征后,算法的表现更优。这表明本文提出的 BGCRRRA 不仅能在仅包含交互信息的场景中进行有效推荐,而且在包含交互信息和属性特征的场景中也能得到更好的表现,这说明 BGCRRRA 具有一定的扩展性。

4.4.3 参数设置影响分析

由于超参数对算法的表现有影响,本节对相关参数设置进行了实验比较,其包括训练时的负样本比例、多阶信息融合时采用的图卷积阶数和多层次融合的数量。实验以数据集 Last-fm 为例,仅使用交互信息,实验结果如图 4—图 6 所示。

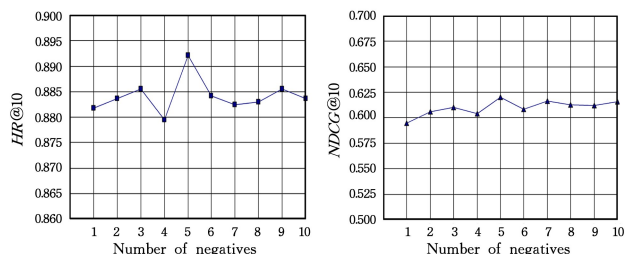


图 4 负样本采样比例探究实验结果

Fig. 4 Experimental results for exploring negative sample ratio

在探究负样本采样比例时,控制变量的设置如下:层次融合数为 3,图卷积阶数为 3。由图 4 可知,在负样本比例逐渐增加的过程中,实验效果也随之提升;当负样本比例为 5 时,实验效果达到最好;但当负样本比例超过 5 后,实验效果降低。这说明随着负样本比例的增加,正负样本数量越来越不平衡,模型更加偏重于学到负样本所包含的信息。

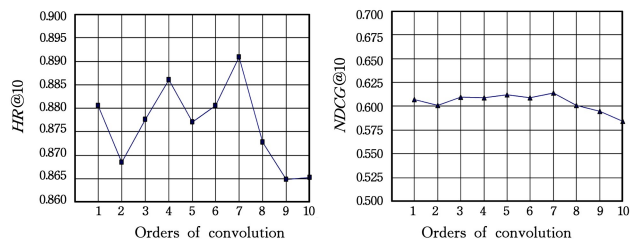


图 5 图卷积阶数探究实验结果

Fig. 5 Experimental results for exploring graph convolution orders

在探究卷积阶数对结果的影响时,控制变量的设置如下:层次融合数为 3,负样本比例为 5。由图 5 可知,图卷积阶数并非越多越好。若原始 GCN 超过两次卷积,则对信息提取效果的提升不大甚至没有提升^[9]。但本文充分考虑了不同阶信息对结果的影响不同,并进行了融合,因此由实验可知:当等于 7 阶时,结果最好;当小于 7 阶时,学到的信息不足;当大于 7 阶时,会过拟合额外的交互信息,从而导致最终效果变差。

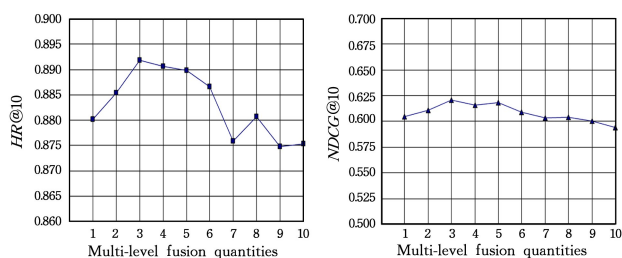


图 6 多层次融合数量探究实验结果

Fig. 6 Experimental results for exploring multi-level fusion quantities

在探究多层次融合数量时,控制变量的负样本比例为 5,图卷积阶数为 3。由图 6 可知,当本实验模型的融合层次小于 3 时,通过增加模型融合层次可提高模型的拟合能力;但随着层次的进一步增加,实验效果降低。这一定程度上反映出深度学习模型无论是在其深度还是宽度上,都应该根据实际情况进行设置。

结束语 推荐问题通常可以看作是一个匹配问题。本文从匹配问题中关于用户和物品的嵌入表达入手,首先参照同构图卷积的定义,根据用户和物品构成的二分图给出二分图上的两部分交替计算的图卷积形式,并利用神经网络来实现;然后自适应融合多阶、多层次卷积操作以实现模型构建;最后在 3 个公开的数据集上与多种基准算法进行对比实验,其结果表明,相比基准算法推荐,本文提出的 BGCRRRA 算法的效果得到了一定的提升,验证了其有效性。同时,本文还针对 BGCRRRA 融合用户物品属性的能力进行了实验,结果表明其具有较好的扩展性。此外,本文通过多角度超参数探究实验,对算法的表现进行了进一步的分析,为算法应用提供了参考。在未来的工作中,将进一步研究如何更有效地表示并融合二分图中的节点属性信息,并结合 Attention 等机制,提升推荐表现。

参考文献

- [1] CHENG X Q, JIN X L, WANG Y Z, et al. Survey on Big Data System and Analytic Technology[J]. Journal of Software, 2014, 25(9): 1889-1908.
- [2] WANG G X, LIU H P. Survey of personalized recommendation system [J]. Computer Engineering and Applications, 2012, 48(7): 66-76.
- [3] HUANG L W, JIANG B T, LV S Y, et al. Survey on Deep Learning Based Recommender System[J]. Journal of Computer Science and Technology, 2018, 41(7): 1619-1647.
- [4] ZHOU B. Research on Relationship Between Bipartite Network Recommendation Algorithm and Collaborative Filtering Algorithm[J]. Computer Science, 2019, 46(S2): 163-166, 177.
- [5] ZHOU T, REN J, MEDO M, et al. Bipartite network projection and personal recommendation [J]. Physical Review E, 2007, 76(4): 046115.
- [6] ZHOU B, YANG C F. The Research Progress of Recommendation Algorithm Based on Bipartite Network[J]. Technology Intelligence Engineering, 2016, 2(1): 77-90.
- [7] ZHOU J, CUI G, ZHANG Z, et al. Graph Neural Networks: A

- Review of Methods and Applications[J]. arXiv:1812.08434, 2018.
- [8] DEFFERRARD M, BRESSON X, VANDERGHEYNST P, et al. Convolutional neural networks on graphs with fast localized spectral filtering[C]//Neural Information Processing Systems. 2016:3844-3852.
- [9] KIPF T, WELLING M. Semi-Supervised Classification with Graph Convolutional Networks[J]. arXiv:1609.02907, 2016.
- [10] WANG X, HE X, WANG M, et al. Neural Graph Collaborative Filtering[C]//Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval. 2019:165-174.
- [11] BERG R, KIPF T N, WELLING M. Graph convolutional matrix completion[J]. arXiv:1706.02263, 2017.
- [12] ZHENG L, LU C, JIANG F, et al. Spectral collaborative filtering [C]//Proceedings of the 12th ACM Conference on Recommender Systems. 2018:311-319.
- [13] HAMILTON W L, YING Z, LESKOVEC J, et al. Inductive Representation Learning on Large Graphs[C]//Neural Information Processing Systems. 2017:1024-1034.
- [14] WU L, SUN P, HONG R, et al. SocialGCN: An Efficient Graph Convolutional Network based Model for Social Recommendation [J]. arXiv:1811.02815, 2018.
- [15] WU L, SUN P, FU Y, et al. A Neural Influence Diffusion Model for Social Recommendation[C]//Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval. 2019:235-244.
- [16] YING R, HE R, CHEN K, et al. Graph Convolutional Neural Networks for Web-Scale Recommender Systems[C]//Proceedings of the Knowledge Discovery and Data Mining. 2018:974-983.
- [17] DENG Z H, HUANG L, WANG C D, et al. Deepcf: A unified framework of representation learning and matching function learning in recommender system[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2019:61-68.
- [18] XUE F, HE X, WANG X, et al. Deep Item-based Collaborative Filtering for Top-N Recommendation[J]. ACM Transactions on Information Systems, 2019, 37(3):1-25.
- [19] ELKAHKY A, SONG Y, HE X, et al. A Multi-View Deep Learning Approach for Cross Domain User Modeling in Recommendation Systems[C]//Proceedings of the Web Conference. 2015:278-288.
- [20] COVINGTON P, ADAMS J, SARGIN E. Deep neural networks for Youtube recommendations [C] // Proceedings of the 10th ACM Conference on Recommender Systems. ACM, 2016:191-198.
- [21] XU Z, CHEN C, LUKASIEWICZ T, et al. Tag-aware personalized recommendation using a deep-semantic similarity model with negative sampling[C]//Proceedings of the 25th ACM Conference on Information and Knowledge Management. ACM, 2016:1921-1924.
- [22] CHENG H T, KOC L, HARMSSEN J, et al. Wide & deep learning for recommender systems [C] // Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. ACM, 2016:7-10.
- [23] XUE H, DAI X, ZHANG J, et al. Deep matrix factorization models for recommender systems[C]//Proceedings of the International Joint Conference on Artificial Intelligence. 2017:3203-3209.
- [24] HE X, LIAO L, ZHANG H, et al. Neural Collaborative Filtering [C]//Proceedings of the Web Conference. 2017:173-182.
- [25] VELICKOVIC P, CUCURULL G, CASANOVA A, et al. Graph Attention Networks[C]//Proceedings of the International Conference on Learning Representations. 2018:1-12.
- [26] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C] // Neural Information Processing Systems. 2017:6000-6010.
- [27] GILMER J, SCHOENHOLZ S, RILEY P, et al. Neural message passing for Quantum chemistry[C]//Proceedings of the International Conference on Machine Learning. 2017:1263-1272.
- [28] HE X, ZHANG H, KAN M, et al. Fast Matrix Factorization for Online Recommendation with Implicit Feedback[C]//Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval. 2016:549-558.



XIONG Xu-dong, born in 1995, post-graduate. His main research interests include recommendation algorithms and network representation learning.



LI Tian-ru, born in 1969, Ph.D, professor, Ph.D supervisor, is a distinguished member of China Computer Federation. His main research interests include big data intelligence, rough sets and granular computing.