

# 一种基于增强图卷积神经网络的协同推荐模型

王 磊 熊于宁 李云鹏 刘媛媛

(西南财经大学经济信息工程学院 成都 610074)

(dr.wangl@163.com)

## A Collaborative Recommendation Model Based on Enhanced Graph Convolutional Neural Network

Wang Lei, Xiong Yuning, Li Yunpeng, and Liu Yuanyuan

(School of Economics Information Engineering, Southwestern University of Finance and Economics, Chengdu 610074)

**Abstract** Graph convolutional network is a deep learning model for graph structure data, and it has become a very hot approach in the research of recommendation system due to its powerful ability in feature extraction and representation of data. This paper focuses on the rating prediction tasks in recommendation system, and points out two deficiencies of existing graph convolutional network based recommendation models by detailed analysis, including making use of first-order collaborative signals only and ignoring the differences of user opinions. For solving them, an end-to-end enhanced graph convolutional network based collaborative recommendation model is proposed. It adopts an enhanced graph convolutional layer to take full advantage of collaborative signals to learn embeddings of users and items on graph, which aggregates second-order collaborative signals and incorporates the influence of different user opinions. And it also stacks several graph convolutional layers to iteratively refine the embeddings and finally uses a nonlinear multilayer perceptron network to make rating prediction. The experiments on 5 benchmark recommendation datasets show that the proposed model achieves lower prediction errors compared with several state-of-the-art graph convolutional network based recommendation models.

**Key words** collaborative recommendation; graph convolutional neural network; collaborative signal; user opinion; rating prediction

**摘要** 图卷积神经网络是一种针对图结构数据的深度学习模型,由于具有强大的特征提取和表示学习能力,它也成为当前推荐系统研究的热门方法。以推荐系统中的评分预测为研究对象,通过分析指出了现有的基于图卷积神经网络的推荐模型存在2个方面的不足:图卷积层仅仅利用了1阶协同信号和未考虑用户观点的差异。为此,提出一种端到端的、基于增强图卷积神经网络的协同推荐模型。它采用一种增强的图卷积层,不仅聚合了2阶协同信号而且融合用户观点的影响,从而更合理地利用协同信号学习实体节点的嵌入表示,并通过堆叠多个图卷积层对其进行精化;最后,采用了非线性的多层次感知机实现评分预测。基于5种推荐数据集上的实验结果表明:新模型的预测误差相比于几种主流的推荐模型具有明显的降低。

收稿日期:2020-08-14;修回日期:2020-12-24

基金项目:教育部人文社会科学研究规划基金项目(16XJAZH002);中央高校基本科研业务费专项资金(JBK2102049)

This work was supported by the Research Planning Fund of Humanities and Social Sciences of the Ministry of Education (16XJAZH002) and the Fundamental Research Funds for the Central Universities (JBK2102049).

关键词 协同推荐;图卷积神经网络;协同信号;用户观点;评分预测

中图法分类号 TP391

协同过滤(collaborative filtering, CF)是推荐系统中的主要方法之一<sup>[1]</sup>.根据是否采用了机器学习模型,它可以分为:基于内存的协同过滤<sup>[2-3]</sup>和基于模型的协同过滤<sup>[4-9]</sup>.相比于前者存在的依赖于相似度函数、内存占用度高、难以发现实体之间的非线性和隐含联系等不足,基于模型的方法利用矩阵隐空间分解<sup>[4-5]</sup>、Markov 链<sup>[6]</sup>、深度神经网络<sup>[7-8]</sup>等模型学习实体的嵌入表示,能够从实体交互操作及辅助信息中更有效地捕获实体间的复杂联系,是目前研究最广泛的推荐方法.

近几年,图卷积神经网络(graph convolutional neural network, GCN)的研究逐渐兴起<sup>[9-14]</sup>,由于具有在图结构上学习嵌入表示的强大能力,它成为当前推荐系统中最热门的研究方法<sup>[10-11,15-18]</sup>.GCN利用谱图理论定义了图卷积操作,在拓扑图上传播和聚合邻居节点的消息,具有同时利用实体特征和图结构特征学习实体嵌入表示的优势.实质上,它是将传统卷积神经网络(convolutional neural network, CNN)<sup>[17]</sup>优良的特征学习和表示能力扩展到非规则的图数据上.

推荐场景中的实体及其交互关系可以很自然地用拓扑图表示,因而非常适于采用 GCN 模型实现基于图的推荐任务.近 3 年来,一些基于图卷积神经网络的推荐模型相继被提出.例如,Berg 等人<sup>[10]</sup>使用 GCN 作为编码器在图自编码框架下实现了 GCMC (graph convolutional matrix completion) 推荐模型,它利用单个图卷积层在“用户-项目”二分图上聚合 1 阶邻居的协同信号,实现了端到端的评分预测.Wang 等人<sup>[15]</sup>提出了一种 NGCF (neural graph collaborative filtering) 模型,它利用 GCN 的消息传递机制堆叠多个图卷积层,实现了二分图上高阶邻居的协同信号隐式传递,最后合并所有层的节点嵌入进行推荐,一定程度弥补了 GCMC 模型只能利用 1 阶协同信号的弱点.此外,一些学者从提升图卷积操作效率<sup>[16-17]</sup>、利用注意力机制<sup>[19-20]</sup>、引入辅助信息<sup>[21]</sup>、提高扩展性<sup>[11,22]</sup>、克服冷启动<sup>[23]</sup>等方面改进了基于 GCN 的推荐模型.这些研究均表明:得益于强大的特征表示学习能力,基于 GCN 的推荐模型能够更好地在实体交互图上捕获隐含、非线性的协同信号,通常获得了比传统协同推荐模型更优的准确率.

尽管图卷积神经网络已被公认在推荐任务中具有显著优势,但已提出的基于 GCN 的推荐模型通常忽视了 2 个方面的问题,尚需进一步完善.

1) 大多数模型的图卷积层没有充分利用高阶邻居的协同信号.例如,GCMC 只利用了 1 阶邻居的信号更新节点嵌入.而文献[24]的研究表明高阶邻居的协同信号同样对学习节点嵌入非常重要.此后的多个模型在 GCN 框架下堆叠多个图卷积层隐含地传递高阶协同信号<sup>[15-16,18,23]</sup>,但在他们的图卷积层内部仍然只用 1 阶邻居更新节点嵌入.该方式不仅很难控制聚合高阶信号时的强度,而且容易受到噪声的影响(详见 1.3 节的分析).2019 年提出的 GraphRec 模型直接利用社交关系聚合了好友的协同信号<sup>[20]</sup>,但该方法并没有利用项目的高阶信号且需要额外的辅助信息.

2) 现有模型的图卷积层在传播和聚合协同信号时很少重视用户的不同观点对推荐结果的影响.在推荐场景中,评分代表了用户的观点.例如,若用户对项目  $v_1$  和  $v_2$  分别给予的评分为 2 和 5,则显然他对  $v_2$  表达了更“喜欢”的观点,或者该项目更符合用户的偏好.然而,现有的大多数模型在聚合 1 阶协同信号时忽略了用户观点的差异<sup>[15-19,23]</sup>.我们注意到,GCMC 和文献[25]提出的模型试图在不同评分等级上分别聚合协同信号,但是他们的方法在合并多个等级的信号时仍然没有考虑用户观点的差别.

针对上述不足,本文以推荐系统中的评分预测任务为对象,研究利用 2 阶邻居的协同信号和用户观点来提升图卷积神经网络在学习节点嵌入时的准确性,进而提出一种增强的基于图卷积神经网络的协同过滤推荐模型——EGCN-CF (enhanced graph convolutional neural network based collaborative filtering recommendation model).

本文的主要贡献有 3 个方面:

- 1) 提出一种直接从邻接矩阵计算 2 阶协同信号及其强度的方法,用于更新图卷积层的节点嵌入;
- 2) 给出一种简便方法将用户观点融合入 1 阶协同信号聚合过程,以便利用用户偏好更准确地学习节点的嵌入表示;
- 3) 结合 1)2)方法,提出了一种增强的基于图卷积神经网络的协同推荐模型,并在多个推荐数据集上验证了其性能优势.

## 1 利用图卷积神经网络实现协同推荐

### 1.1 推荐系统中的评分预测问题的定义

在推荐场景中,假设已知稀疏的评分矩阵  $\mathbf{R} \in \mathbb{R}^{m \times n}$ ,它的少量非 0 元素  $r_{u,v} \in \{1, 2, \dots, T\}$  代表用户  $u$  对项目  $v$  的评分结果,共有  $T$  个等级;矩阵  $\mathbf{X}_u \in \mathbb{R}^{m \times d_u}$  和  $\mathbf{X}_v \in \mathbb{R}^{n \times d_v}$  分别表示用户和项目的实体特征,其中,  $m$  和  $n$  分别是用户集  $\mathcal{U}$  和项目集  $\mathcal{V}$  的规模,  $d_u$  和  $d_v$  分别是用户特征向量和项目特征向量的维度。评分预测的目标是:建立预测模型对大量尚未评分的项目给予准确地评分,进而实现项目推荐。

在基于图神经网络的模型中,通常将用户、项目及其评分关系视为无向二分图  $\mathcal{G} = \langle E, A \rangle$ ,其中,  $E = \mathcal{U} \cup \mathcal{V}$  是图上的节点集,邻接矩阵  $A \in \mathbb{R}^{(m+n) \times (m+n)}$  代表任意节点之间是否存在交互关系,可以直接从评分矩阵  $\mathbf{R}$  计算获得。此时,评分预测问题可以视为对二分图  $\mathcal{G}$  上未知链接的预测。

### 1.2 基于 GCN 的通用推荐模型

基于图卷积神经网络的推荐模型利用消息传递的思想在二分图上传播和聚合协同信号,学习节点的嵌入表示并用于推荐预测<sup>[10, 15-16, 18, 26]</sup>。它们的模型结构通常由 3 个部分构成:输入层、图卷积层和预测层。

1) 输入层,对图  $\mathcal{G}$  上的实体节点进行编码,获得它们的初始低维嵌入向量  $\mathbf{e}^0$ ,如

$$\mathbf{e}^0 = Enc(\mathbf{x}, \Theta^{in}), \quad (1)$$

其中,  $\mathbf{x}$  代表实体节点的信息,可以采用节点的实体特征或者基于节点序号的独热编码向量;  $\mathbf{e}^0$  是其编码后的初始嵌入向量,可细分为用户嵌入  $\mathbf{e}_u^0$  和项目嵌入  $\mathbf{e}_v^0$ ;  $\Theta^{in}$  是编码函数  $Enc(\cdot)$  的待学习参数集合。

2) 图卷积层,又称为“嵌入传播层”,它利用二分图的结构特征在图上传播和聚合邻居节点的协同信号,实现节点的嵌入更新。其主要操作包括:协同信号构造和节点嵌入更新。

以用户  $u$  对项目  $v$  的一次评分为例,该操作反映了用户对项目的一定程度的偏好。因此,可以将此次操作传播的 1 阶协同信号构造为

$$s_{u \leftarrow v} = f(\mathbf{e}_u, \mathbf{e}_v, p_{u,v}, \Theta^f), \quad (2)$$

其中,  $p_{u,v} = 1 / \sqrt{|\mathcal{N}_u| \cdot |\mathcal{N}_v|}$  是信号强度系数;在二分图上与用户  $u$  和项目  $v$  直接相邻的 1 阶邻居集合分别用  $\mathcal{N}_u$  和  $\mathcal{N}_v$  表示,  $|\mathcal{N}_u|$  和  $|\mathcal{N}_v|$  是它们的节点度;  $f(\cdot)$  是协同信号编码函数,  $\Theta^f$  是该函数的待

学习参数集合,不同的模型对该函数有不同形式的定义<sup>[10-11, 15-16, 20]</sup>。

嵌入更新操作在二分图上汇集所有 1 阶邻居的协同信号,用于更新用户  $u$  的嵌入表示:

$$\mathbf{e}_u^{\text{update}} = g(\mathbf{e}_u, \{s_{u \leftarrow v} | v \in \mathcal{N}_u\}), \quad (3)$$

其中,聚合函数  $g(\cdot)$  可以采用均值、最大池化、LSTM 网络等方式实现<sup>[13, 18]</sup>。

类似地,当项目  $v$  获得用户  $u$  的评分,表明项目的品质与用户偏好一定程度的相符。因此,可以利用用户的协同信号对项目实体的嵌入进行更新,即

$$s_{v \leftarrow u} = f(\mathbf{e}_v, \mathbf{e}_u, p_{u,v}, \Theta^f), \quad (4)$$

$$\mathbf{e}_v^{\text{update}} = g(\mathbf{e}_v, \{s_{v \leftarrow u} | u \in \mathcal{N}_v\}). \quad (5)$$

这样,图卷积层利用 1 阶协同信号在二分图上实现了实体节点的嵌入表示的一次更新。

此外,许多基于 GCN 的推荐模型还堆叠了  $L$  个图卷积层迭代地精化节点嵌入表示<sup>[11, 15-16, 18]</sup>,也即,依次获得  $\mathbf{e}_u^1, \mathbf{e}_u^2, \dots, \mathbf{e}_u^L$  和  $\mathbf{e}_v^1, \mathbf{e}_v^2, \dots, \mathbf{e}_v^L$ 。这种方式也隐含地实现了高阶的协同信号的逐层传播。

3) 预测层,使用最终的节点嵌入  $\mathbf{e}_u^L$  和  $\mathbf{e}_v^L$ ,对任意用户  $u$  和项目  $v$  之间的未知评分进行预测,如

$$\hat{r}_{u,v} = h(\mathbf{e}_u^L, \mathbf{e}_v^L, \Theta^h), \quad (6)$$

其中,预测函数  $h(\cdot)$  通常采用双线性解码函数<sup>[10, 25]</sup>或多层感知机网络<sup>[5, 20]</sup>实现;  $\Theta^h$  代表其参数集合。

### 1.3 问题分析

一方面,如 1.2 节所述,图卷积层采用的协同信号聚合式(3)(5)中,只使用了来自直接交互的 1 阶邻居和自身的协同信号,忽略了卷积层内的高阶信号。然而,根据协同过滤的基本思想<sup>[1]</sup>,用户行为或偏好还受到与之相似的其他用户的影响,他们在二分图上可视为相似的高阶邻居。因此,在图卷积层聚合高阶信号是非常有益的。近期,文献[27]也利用随机游走模型指出了高阶信号对于优化节点嵌入的价值。

一些基于 GCN 的推荐模型虽然通过堆叠多个图卷积层隐含地传播了高阶协同信号<sup>[15-16, 18, 23]</sup>,但这种方式难以控制高阶信号的强度,容易受到噪声的干扰。例如,假设在二分图上存在一条连通路径  $u_0 \rightarrow v_1 \rightarrow u_2 \rightarrow v_3 \rightarrow u_4$ ,其中,用户  $u_4$  由于偶然原因评分了项目  $v_3$ ,尽管它与目标用户  $u_0$  的偏好并不相似,但这次评分操作仍将导致  $u_4$  的协同信号隐含地沿着堆叠的 4 个图卷积层逆向传递到  $u_0$ ,而不是被视为噪声信号受到抑制或消除。

另一方面,在图卷积层的聚合式(3)(5)中,所有的 1 阶邻居信号被同等对待,没有考虑用户观点(即评分值)的差别。显然,高评分项目的品味比低评分

项目更能够反映用户的偏好<sup>[20,28]</sup>,应该在信号聚合操作中起到更大的作用.

## 2 提出的 EGCN-CF 协同推荐模型

本节将针对基于 GCN 的推荐模型存在的如 1.3

节所述的缺陷,提出一种端到端的 EGCN-CF 模型.它通过在图卷积层内引入 2 阶协同信号和用户观点,更有效地利用协同信号学习节点的嵌入表示,最后利用多层的感知机网络(multi-layer perceptron, MLP)实现对未知评分的预测.其基本结构如图 1 所示,包括输入层、增强的图卷积层和预测层 3 个部分.

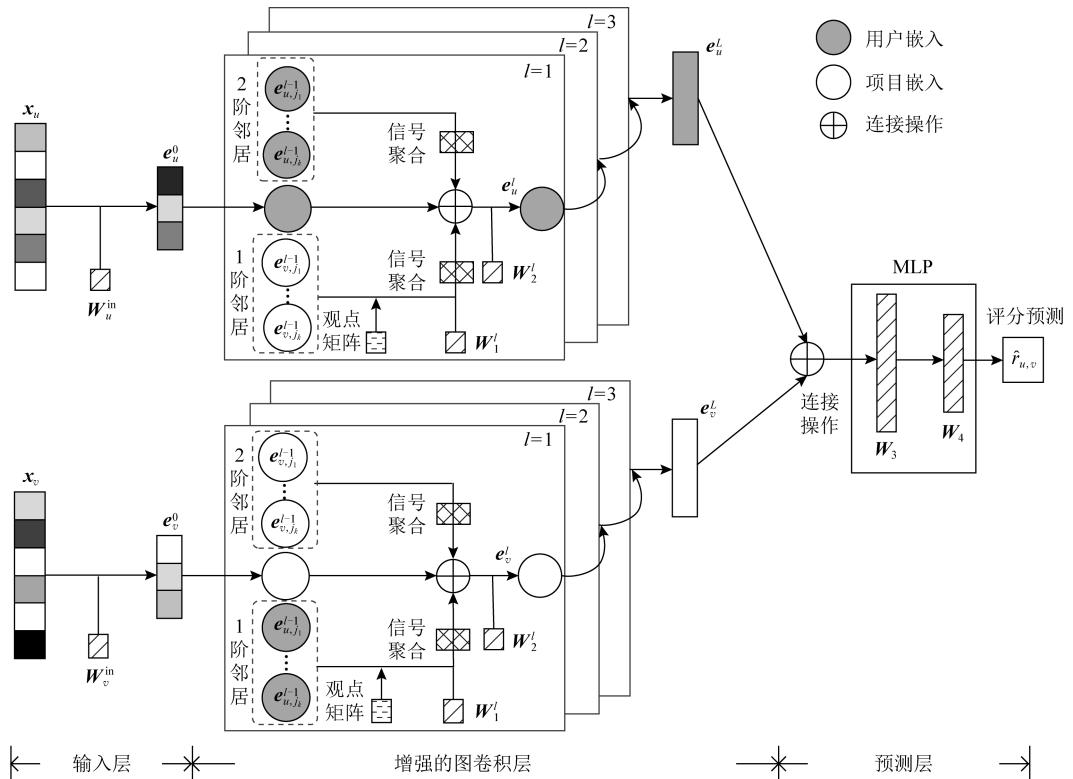


Fig. 1 The general architecture of EGCN-CF model

图 1 EGCN-CF 模型的基本结构

### 2.1 输入层

EGCN-CF 模型的输入层负责对用户或项目的实体特征进行编码,生成的低维嵌入向量作为图卷积层的输入.它采用的编码函数为

$$\begin{cases} e_u^0 = W_u^{\text{in}} \cdot x_u, \\ e_v^0 = W_v^{\text{in}} \cdot x_v, \end{cases} \quad (7)$$

其中,\$W\_u^{\text{in}} \in \mathbb{R}^{h \times d\_u}\$ 和 \$W\_v^{\text{in}} \in \mathbb{R}^{h \times d\_v}\$ 分别是用户和项目的待学习的编码矩阵.

值得注意的是,式(7)采用了直接对用户或项目的原始实体特征进行编码的方法,获得它们的初始嵌入向量.相比于 NGCF 等模型采用的基于节点序号的独热编码方法<sup>[15-16]</sup>,该方法可以将语义更丰富的实体特征或辅助信息编码后输入到图卷积层,更有利于准确地学习节点的嵌入表示;并且,由于只需要优化编码矩阵,使得模型的优化参数更少.实验部分将具体描述对实体特征进行编码的细节.

### 2.2 增强的图卷积层

同 NGCF 模型类似,本文模型通过堆叠 \$L\$ 个图卷积层实现节点嵌入表示的逐层精化.但我们针对现有模型的图卷积层的缺陷,通过融合 2 阶协同信号和用户观点进行了增强.本节以第 \$l \in \{1, 2, \dots, L\}\$ 图卷积层为例介绍其主要操作.

#### 2.2.1 构造的 2 阶协同信号

令 \$u\_i \rightarrow v\_k \rightarrow u\_j \rightarrow \dots\$ 是二分图 \$\mathcal{G}\$ 上的任意一条从用户 \$u\_i\$ 出发的连通路径,下标 \$i, k, j, \dots\$ 表示节点编号.我们按照路径上距离用户 \$u\_i\$ 的跳数,称 \$v\_k\$ 为 \$u\_i\$ 的 1 阶邻居, \$u\_j\$ 为 \$u\_i\$ 的 2 阶邻居,以此类推.

如 1.3 节的讨论,在图卷积层中聚合高阶协同信号有益于节点嵌入表示的学习.但是,本文模型只使用了所有高阶邻居中的 2 阶邻居的协同信号,主要原因有 2 点:1)根据 Xu 等人<sup>[27]</sup>的研究,连通路径上高阶邻居的协同信号的强度随着阶数增加而迅速

衰减,并且噪声信号容易在路径上积累,实际上该论断也一定程度解释了许多基于GCN的模型只堆叠了3或4个图卷积层的原因<sup>[15-16]</sup>;2)2阶邻居是与目标用户(项目)有共同交互项目(用户)的同质实体,按照三元闭包理论的解释<sup>[29]</sup>,2阶邻居相对于更高阶邻居而言,它们与目标实体的偏好更为相似,并且,与目标实体共同的交互实体越多,相似度越高。此外,第3节的实验结果也验证了增加同质的4阶邻居的协同信号不能明显提高推荐模型的性能。

同时,根据文献[30]在推荐实体相似性方面的研究:越相似的邻居所传递的信号越多,共同交互实体越多的邻居所传递的信号越多。受其启发,我们给出一种对2阶协同信号及其强度进行计算的方法。令 $u_j$ 是目标用户 $u_i$ 在图 $\mathcal{G}$ 上的2阶邻居, $u_i \rightarrow v_k \rightarrow u_j$ 是它们之间长度为2的任意连通路径,则从 $u_j$ 传播到 $u_i$ 的2阶协同信号定义为

$$\mathbf{s}_{i \leftarrow j}^l = \sum_{k \in \mathcal{N}_i \cap \mathcal{N}_j} p_{i,k,j} (\mathbf{e}_{u,j}^{l-1} + \mathbf{e}_{u,i}^{l-1} \odot \mathbf{e}_{u,j}^{l-1}), \quad (8)$$

其中, $\mathbf{e}_{u,j}^{l-1}$ 是用户 $u_j$ 在第 $l-1$ 层的嵌入; $\odot$ 是按元素的Hadamard乘积; $\mathcal{N}_i \cap \mathcal{N}_j$ 表示 $u_i$ 和 $u_j$ 共同交互的项目集合; $p_{i,k,j}$ 是图卷积神经网络的2阶协同信号强度系数,与路径上的节点度有关,本文取值为 $p_{i,k,j} = 1 / (\sqrt{|\mathcal{N}_i|} \cdot |\mathcal{N}_k| \cdot \sqrt{|\mathcal{N}_j|})$ 。需要指出,乘积项 $\mathbf{e}_{u,i}^{l-1} \odot \mathbf{e}_{u,j}^{l-1}$ 也起到了调节邻居节点的协同信号强度的作用,使得相似邻居传播更多的2阶协同信号。例如,当邻居 $u_j$ 与目标 $u_i$ 的嵌入表示相似时,由于向量方向一致,乘积项将增强其协同信号;反之,则影响微弱。

分析可知,式(8)对噪声信号有很好的抑制作用。假如用户 $u_j$ 与目标用户 $u_i$ 的偏好并不相似,只是通过偶然的一次评价操作成为了 $u_i$ 的2阶邻居。但是,由于他们的节点嵌入不相似,导致Hadamard乘积项的信号增强作用微弱,再考虑到他们的共同交互项目 $\mathcal{N}_i \cap \mathcal{N}_j$ 数量非常少,根据式(8)的定义, $u_j$ 传播到目标用户 $u_i$ 的噪声协同信号也非常有限。

我们令 $\mathcal{N}_i^{(2)}$ 表示目标用户 $u_i$ 的所有2阶邻居,则所有汇聚到 $u_i$ 的2阶协同信号为

$$\mathbf{s}_{u,i}^{l(2)} = \sum_{j \in \mathcal{N}_i^{(2)}} \mathbf{s}_{i \leftarrow j}^l = \sum_{j \in \mathcal{U}, j \neq i} \sum_{k \in \mathcal{N}_i \cap \mathcal{N}_j} \frac{\mathbf{e}_{u,j}^{l-1} + \mathbf{e}_{u,i}^{l-1} \odot \mathbf{e}_{u,j}^{l-1}}{\sqrt{|\mathcal{N}_i|} \cdot |\mathcal{N}_k| \cdot \sqrt{|\mathcal{N}_j|}}. \quad (9)$$

类似地,容易推出第 $l$ 层汇聚到项目 $v_j$ 的2阶协同信号 $\mathbf{s}_{v,i}^{l(2)}$ ,我们不再赘述。

令 $\mathbf{E}^{l-1} \in \mathbb{R}^{(m+n) \times h}$ 是第 $l-1$ 层所有嵌入向量组成的矩阵; $\mathbf{S}^{l(2)} \in \mathbb{R}^{(m+n) \times h}$ 表示第 $l$ 层所有节点汇

集的2阶协同信号矩阵,其行向量由 $\mathbf{s}_{u,i}^{l(2)}$ 或 $\mathbf{s}_{v,i}^{l(2)}$ 构成。根据谱图理论知识<sup>[31]</sup>,可以整理出2阶协同信号矩阵 $\mathbf{S}^{l(2)}$ 的简洁形式:

$$\mathbf{S}^{l(2)} = \bar{\mathcal{L}}^{(2)} \mathbf{E}^{l-1} + (\bar{\mathcal{L}}^{(2)} \mathbf{E}^{l-1}) \odot \mathbf{E}^{l-1}, \quad (10)$$

其中, $\bar{\mathcal{L}}^{(2)}$ 是去对角的规范2阶邻接矩阵,即 $\bar{\mathcal{L}}^{(2)} = (\mathcal{L}^{(2)} - \text{diag}(\mathcal{L}^{(2)}))$ ;而 $\mathcal{L}^{(2)}$ 是规范化的2阶邻接矩阵,它度量了基于节点度的2阶信号强度,是信号强度系数 $p_{i,k,j}$ 矩阵形式,即

$$\mathcal{L}^{(2)} = (\mathbf{D}^{-\frac{1}{2}} \cdot \mathbf{A} \cdot \mathbf{D}^{-\frac{1}{2}})^2. \quad (11)$$

这里, $\mathbf{D}$ 是由节点度组成的对角矩阵; $\mathbf{A}$ 是图 $\mathcal{G}$ 的邻接矩阵,本文取为 $\mathbf{A} = \begin{pmatrix} \mathbf{0} & z(\mathbf{R}) \\ z(\mathbf{R}^T) & \mathbf{0} \end{pmatrix}$ ,其中 $z(\cdot)$ 是将矩阵中非0元素重置为1的简单函数。

## 2.2.2 带用户观点的1阶协同信号

用户对项目的评分不仅说明他们之间存在交互关系,更重要的是评分具体表达了用户的观点。例如,在电影推荐中,用5分、3分、1分表达了用户的“非常喜欢”“一般喜欢”“不喜欢”3种不同的观点。根据文献[28]的解释,获得了正向观点的高评分项目更能反映用户的真实偏好。因此,在汇聚1阶协同信号时融入用户观点,能够更准确地学习用户嵌入。

我们借鉴GCN的消息传递机制,给出一种带用户观点的1阶协同信号构造方法。它把汇聚到用户 $u_i$ 的1阶协同信号定义为

$$\mathbf{s}_{u,i}^{l(1)} = \sum_{j \in \mathcal{N}_i} c_{i,j} p_{i,j} \mathbf{W}_1^l \mathbf{e}_{v,j}^{l-1}, \quad (12)$$

其中,观点系数 $c_{i,j}$ 表示用户 $u_i$ 对项目 $v_j$ 的观点,本文取值为归一化的评分,即 $c_{i,j} = r_{u_i, v_j} / \max(\mathbf{R})$ ;常数 $p_{i,j}$ 是1阶信号强度,与式(2)中的 $p_{u,v}$ 定义一致; $\mathbf{W}_1^l \in \mathbb{R}^{h \times h}$ 是第 $l$ 层待学习的权重矩阵。

显然,高评分项目的观点系数取较大值,其传递的协同信号更多,反之亦然。因此,这种带观点的方式更合理地聚合了1阶协同信号。

类似地,容易推出第 $l$ 层汇聚到项目 $v_j$ 的带观点的1阶协同信号 $\mathbf{s}_{v,i}^{l(1)}$ ,我们不再赘述。

同样,我们根据谱图理论整理出带观点的1阶协同信号的简洁矩阵形式,即

$$\mathbf{S}^{l(1)} = (\mathbf{C} \odot \mathcal{L}^{(1)}) \mathbf{E}^{l-1} \mathbf{W}_1^l, \quad (13)$$

其中, $\mathbf{C}$ 是由 $c_{i,j}$ 构成的归一化的观点矩阵,定义为

$$\mathbf{C} = \begin{pmatrix} \mathbf{0} & \frac{\mathbf{R}}{\max(\mathbf{R})} \\ \frac{\mathbf{R}^T}{\max(\mathbf{R})} & \mathbf{0} \end{pmatrix}; \quad \mathcal{L}^{(1)} \text{是规范化的1阶邻接矩阵,即 } \mathcal{L}^{(1)} = (\mathbf{D}^{-\frac{1}{2}} \cdot \mathbf{A} \cdot \mathbf{D}^{-\frac{1}{2}}).$$

### 2.2.3 节点的嵌入更新

我们汇集 1 阶、2 阶协同信号以及节点前一层的嵌入,对第  $l$  层的节点嵌入向量进行更新,如

$$\mathbf{E}^l = \text{ReLU}((\mathbf{E}^{l-1} \oplus \mathbf{S}^{l(1)} \oplus \mathbf{S}^{l(2)}) \mathbf{W}_2^l), \quad (14)$$

其中,  $\text{ReLU}(\cdot)$  是激活函数; 运算符  $\oplus$  表示矩阵连接操作;  $\mathbf{W}_2^l \in \mathbb{R}^{3h \times h}$  是待学习的权重矩阵。注意到, 式(14)将前一层的嵌入  $\mathbf{E}^{l-1}$  也连接到协同信号中, 根据文献[16]的解释, 它相当于二分图上的自连接信号, 有助于避免学习节点嵌入时陷入“过度光滑”。

此后, EGCN-CF 模型可以堆叠  $L$  个图卷积层, 按照式(14)实现嵌入表示的逐层精化, 也即, 依次获得  $\mathbf{E}^1, \mathbf{E}^2, \dots, \mathbf{E}^L$ 。

### 2.3 非线性预测层

将图卷积层最终学习的节点嵌入  $\mathbf{e}_u^L$  和  $\mathbf{e}_v^L$  连接起来, 输入到 2 层的 MLP 网络中, 对用户  $u$  和项目  $v$  之间的未知评分  $\hat{r}_{u,v}$  进行预测, 即

$$\hat{r}_{u,v} = \mathbf{W}_4(\text{ReLU}(\mathbf{W}_3(\mathbf{e}_u^L \oplus \mathbf{e}_v^L))), \quad (15)$$

其中,  $\mathbf{W}_3 \in \mathbb{R}^{2h \times d'}$ ,  $\mathbf{W}_4 \in \mathbb{R}^{d' \times 1}$  是待学习的 MLP 权重矩阵;  $d'$  是 MLP 第 1 层(隐层)神经元的数量。

需要指出的是, 相比于一些文献采用的线性预测模型<sup>[10, 15-16, 18]</sup>, 本文采用非线性的 MLP 模型具有对节点嵌入向量之间复杂关系的更强的拟合能力, 更适用于评分预测问题。

### 2.4 模型训练

EGCN-CF 模型是一种端到端的推荐模型。为了训练其模型参数, 我们给出一种基于均方误差的优化目标函数:

$$\text{obj} = \frac{1}{2|O|} \sum_{(u,v) \in O} (r_{u,v} - \hat{r}_{u,v})^2 + \lambda \|\boldsymbol{\theta}\|_2^2, \quad (16)$$

其中,  $(u, v)$  是二分图  $\mathcal{G}$  上的已获评分  $r_{u,v}$  的边,  $O$  是它们的集合;  $\boldsymbol{\theta} = \{\mathbf{W}_u^{\text{in}}, \mathbf{W}_v^{\text{in}}, \{\mathbf{W}_1^l, \mathbf{W}_2^l\}_{l=1}^L, \mathbf{W}_3, \mathbf{W}_4\}$  是模型的所有待训练的参数;  $\|\boldsymbol{\theta}\|_2^2$  代表参数的  $L_2$  正则项, 用于防止模型过拟合,  $\lambda$  是其正则化系数。

此外, 在模型训练时还借鉴文献[15]采用的节点 dropout 技术防止过拟合, 其 dropout 率为  $\eta$ 。

最后, 我们采用 Adam 梯度下降优化器对式(16)进行优化<sup>[32]</sup>, 并采用小批量(minibatch)技术提高模型训练效率。

### 2.5 讨论

在 EGCN-CF 模型的图卷积层中, 显式地增加了对同质的 2 阶邻居的协同信号的聚合。从理论上讲, 它可以聚合更高阶的协同信号。例如, 在二分图

上同时聚合 2 阶和 4 阶的同质协同信号时, 只需简单地将式(11)替换为式(17)即可:

$$\begin{aligned} \mathcal{L}^{(2,4)} &= (\mathbf{D}^{-\frac{1}{2}} \cdot \mathbf{A} \cdot \mathbf{D}^{-\frac{1}{2}})^2 + \\ &(\mathbf{D}^{-\frac{1}{2}} \cdot \mathbf{A} \cdot \mathbf{D}^{-\frac{1}{2}})^4. \end{aligned} \quad (17)$$

另外, 在文献[20]的推荐模型中, 将社交网络中好友的协同信号聚合到图卷积层。社交好友相当于二分图上具有高相似性的同质高阶邻居, 聚合其协同信号显然有利于学习用户节点的嵌入表示。与之相比, 本文模型通过矩阵运算对 2 阶邻居及其信号强度进行了合理的计算, 无需额外的辅助信息。并且, 当数据集中存在关于同质节点关系的辅助信息时, 比如, 社交关系矩阵  $\mathbf{Q}_u \in \mathbb{R}^{m \times m}$ , 项目相似性矩阵  $\mathbf{Q}_v \in \mathbb{R}^{n \times n}$ , 通过修改式(11)很容易把它们融入本文模型, 如

$$\mathcal{L}^{(2)} = (\mathbf{D}^{-\frac{1}{2}} \cdot \mathbf{A} \cdot \mathbf{D}^{-\frac{1}{2}})^2 + \alpha \begin{pmatrix} \mathbf{Q}_u & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_v \end{pmatrix}, \quad (18)$$

其中,  $\alpha$  是融合参数。

最后, EGCN-CF 的 3 层结构与文献[15]的 NGCF 模型有相似之处, 但它们在输入层、2 阶协同信号利用、用户观点结合、预测层设计等方面具有显著差异。

## 3 实验与分析

为了验证提出的 EGCN-CF 推荐模型在评分预测任务上的性能, 我们在 5 个通用的推荐数据集上进行了实验, 并将它与主流的推荐模型进行性能对比。

### 3.1 数据集

表 1 概括了实验用数据集的情况。其中, Flixster 和 Douban 直接使用文献[33]提供的预处理后的数据集; ML-100K, ML-1M, ML-10 M 是来自于 MovieLens 网站的不同规模的电影推荐数据集<sup>①</sup>。

在预处理阶段, 我们在二分图上直接利用辅助信息和图结构特征对用户的原始实体特征进行编码。对于 Flixster 和 Douban 数据集, 实体特征由 3 个部分连接获得: 节点度  $\mathbf{F}_1$ 、评分特征  $\mathbf{F}_2$ 、邻接关系特征  $\mathbf{F}_3$ 。其中, 评分特征  $\mathbf{F}_2$  统计了用户(或项目)给予(或获得)的不同等级评分的比例, 邻接关系特征  $\mathbf{F}_3$  体现了同质节点之间的邻接关系, 本文将其设置为用 Node2vec 模型<sup>[34]</sup>在用户或项目邻接矩阵上生成的 256 维向量。对于 MovieLens 数据集,

<sup>①</sup> <https://grouplens.org/datasets/movielens/>

项目的实体特征由节点度  $F_1$ 、评分特征  $F_2$ 、语义特征  $F_3$  三部分组成。其中，语义特征  $F_3$  采用文献 [28] 中的实验方法，对电影类别、名称、发行年编码获得，用户的实体特征由节点度  $F_1$ 、评分特征  $F_2$ 、

人口统计特征  $F_3$ 、交互项目的平均语义特征  $F_4$  四部分组成。其中，人口统计特征  $F_3$  是由年龄、性别、职业编码的 23 维向量。表 1 最后 2 列给出了编码后的用户和项目的实体特征维度。

Table 1 Statistics of Experimental Datasets

表 1 实验数据集统计

数据集	用户数	项目数	评分数量	评分密度	评分等级	用户特征数	项目特征数
Flixster	3 000	3 000	26 173	0.0029	0.5, 1, ..., 5	263	263
Douban	3 000	3 000	136 891	0.0152	1, 2, ..., 5	262	262
ML-100K	943	1 682	100 000	0.0630	1, 2, ..., 5	329	326
ML-1M	6 040	3 706	1 000 209	0.0447	1, 2, ..., 5	329	326
ML-10M	69 878	10 677	10 000 054	0.0134	1, 2, ..., 5	306	327

### 3.2 实验设置

实验中，我们选取 4 种主流推荐模型作为对比算法，包括：NeuCF<sup>[5]</sup>，GCMC<sup>[10]</sup>，GraphRec<sup>[20]</sup>，NGCF<sup>[15]</sup>。其中，NeuCF 是近几年提出的基于矩阵分解的协同推荐模型，它采用深度神经网络代替嵌入向量的内积运算，实验中我们采用式(16)替换其原优化目标，使之能够适用于评分预测任务。其他 3 种模型都是目前主流的基于 GCN 的协同过滤推荐模型，其中，对于 GraphRec 模型，我们采用用户的 2 阶邻居模拟其在 MovieLens 数据集上的社交关系；对于 NGCF 模型，我们也采用式(16)代替其原有的成对 BPR(Bayesian personalized ranking)损失目标函数，以便进行评分预测。

我们用 Tensorflow 实现了本文的 EGCN-CF 模型，其主要参数配置为：节点嵌入维度  $h = 64$ ，图

卷积层的数量  $L = 3$ ，正则系数  $\lambda$  通过交叉验证的方法在  $\{0.01, 0.05, 0.1, 0.5, 1, 5, 10\}$  范围内选取，MLP 的隐层参数  $d' = 16$ ；训练时 Adam 优化器的学习速率设置为  $10^{-3}$ ，最大迭代次数为 1 000 次，每批训练数据规模设置为 1 024，节点 dropout 概率  $\eta = 0.3$ 。对比算法的参数基本按照原文设置。

另外，将数据集按 0.8, 0.1, 0.1 的比例划分为训练集、验证集和测试集。在验证集上实现模型参数的选取，并采用均方根误差 (root mean square error, RMSE) 指标在测试集上评价各模型的预测误差。

### 3.3 对比实验

我们在各数据集上比较 EGCN-CF 模型与对比算法的性能，它们在测试集上预测时的均方根误差如表 2 所示。这里，除 GCMC 的结果来自于文献 [10] 外，其他结果都是 3 次重复实验的平均值。

Table 2 Comparison of Prediction Errors (RMSE) of Different Recommendation Models

表 2 比较不同推荐模型的预测误差 (RMSE)

模型	Flixster	Douban	ML-100K	ML-1M	ML-10M
NeuCF	1.301	0.822	0.916	0.848	0.787
GCMC	0.917	0.734	0.910	0.832	0.777
GraphRec	0.892	<b>0.720</b>	0.901	0.829	0.777
NGCF	0.898	0.731	0.895	0.823	0.779
EGCN-CF	<b>0.876</b>	0.721	<b>0.884</b>	<b>0.817</b>	<b>0.763</b>

注：黑体值为最优值。

观察表 2 的实验结果，可以得出 4 个结论：

1) 基于图卷积神经网络的 4 种推荐模型的预测误差明显优于基于矩阵分解的 NeuCF 模型。这说明得益于在拓扑图上融合了节点的实体特征和图结构特征，前者能够更加准确地学习节点的嵌入表示，从而获得更好的预测结果。

2) GCMC 模型由于采用 1 个图卷积层，只能聚合 1 阶邻居的协同信号，对节点嵌入的学习能力不足，因而它在 4 种基于 GCN 的推荐模型中表现最差，特别是在非常稀疏的 Flixster 数据集上。

3) GraphRec 和 NGCF 模型都利用了高阶协同信号，均取得了优于 GCMC 的预测结果。不同的是，

GraphRec 直接聚合社交好友的协同信号,他们相当于二分图中的相似的高阶邻居;NGCF 则通过堆叠多个图卷积层隐式地传播和聚合高阶协同信号。这说明了聚合高阶协同信号有助于提高推荐模型的性能。

4) 本文的 EGCN-CF 模型在绝大部分数据集上均取得了最低的预测误差,仅仅在 Douban 数据集上稍微高于 GraphRec 模型。并且,我们通过采用“A/B 测试”进行显著性检验,发现本文的 EGCN-CF 模型的实验结果在大多数情况下以 95% 置信度显著优于其他模型(仅在 3 种情况下无明显优势,包括:在 Douban, ML-1M 数据集上与 GraphRec 比较、在 ML-1M 数据集上与 NGCF 比较)。本文模型的优异性能可以归结于它在图卷积层中增加了对更高阶协同信号的利用和将用户观点引入信号聚合过

程。与之相比,GraphRec 只是从社交关系中直接获取了用户的高阶协同信号,并未考虑项目的高阶信号;NGCF 模型并未在图卷积层内部显式地利用高阶协同信号,也未考虑用户观点的影响。

### 3.4 EGCN-CF 模型的进一步分析

为进一步分析 EGCN-CF 模型取得优异性能的原因,我们给出它的 4 种变体模型,并比较它们在推荐数据集上的性能,包括:EGCN-CF-a4, EGCN-CF-do, EGCN-CF-oo, EGCN-CF-l。与基本模型相比,EGCN-CF-a4 增加了 4 阶协同信号;EGCN-CF-do 去掉了式(13)中的用户观点;EGCN-CF-oo 去掉了式(14)的 2 阶协同信号;EGCN-CF-l 直接采用节点嵌入向量乘积  $(\mathbf{e}_u^L)^T \cdot \mathbf{e}_v^L$  代替式(15)中的非线性神经网络。实验参数设置与 3.2 节的描述一致,表 3 汇总了 3 次重复实验的平均预测误差。

Table 3 Comparison of Prediction Errors (RMSE) of Different Variant Models

表 3 比较不同变体模型的预测误差(RMSE)

模型	Flixster	Douban	ML-100K	ML-1M	ML-10M
EGCN-CF-a4	<b>0.874</b>	0.723	0.888	<b>0.817</b>	0.766
EGCN-CF-do	0.880	0.723	0.890	0.819	0.771
EGCN-CF-oo	0.894	0.729	0.895	0.826	0.775
EGCN-CF-1	0.884	0.725	0.888	0.819	0.770
EGCN-CF	0.876	<b>0.721</b>	<b>0.884</b>	<b>0.817</b>	<b>0.763</b>

注:黑体值为最优值。

观察表 3 的实验结果,可以发现:

1) 尽管 EGCN-CF-a4 模型增加了 4 阶协同信号,但其取得的预测误差与基本模型相当。这说明由于更高阶的协同信号强度较弱且可能带有噪声,无助于明显提高推荐模型的性能,这也是本文模型在图卷积层中只增加同质的 2 阶协同信号的原因。

2) 去掉用户观点的 EGCN-CF-do 模型在各数据集的预测误差相比于基本模型均有一定程度的增大。这说明结合用户观点有利于提高推荐模型的性能。

3) 去掉 2 阶协同信号的 EGCN-CF-oo 模型在各数据集上表现最差。例如,Flixster 数据集上的误差相比基本模型增大了 2.1%。这表明在图卷积层中增加 2 阶协同信号有利于提升评分预测的准确率。

4) 采用线性预测层的 EGCN-CF-1 模型也取得了较低的预测误差,稍差于标准模型。这表明:一方面图卷积层已经能够学习获得准确的节点嵌入表示,即使利用简单的线性预测方法也能获得较好的准确率;另一方面,采用非线性的预测层能进一步提高模型的预测准确率。

### 3.5 图卷积层数量的影响

考虑到 NGCF 和 EGCN-CF 都堆叠了多个图卷积层,实现对节点嵌入的逐层精化。我们通过实验考察 2 种模型取不同数量的图卷积层时的预测误差情况。出于篇幅原因,我们只给出在 Flixster 和 ML-100K 数据集上的对比结果,如图 2、图 3 所示:

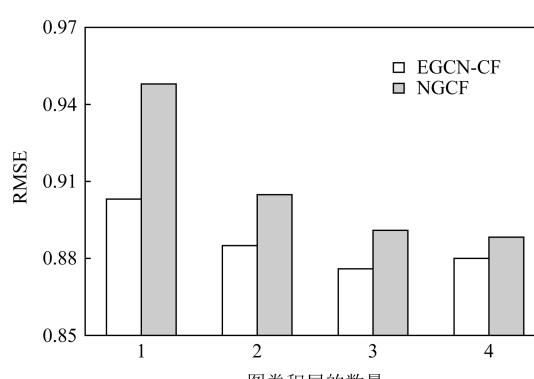


Fig. 2 Effect of graph convolutional layer numbers on Flixster

图 2 在 Flixster 数据集上图卷积层数量的影响

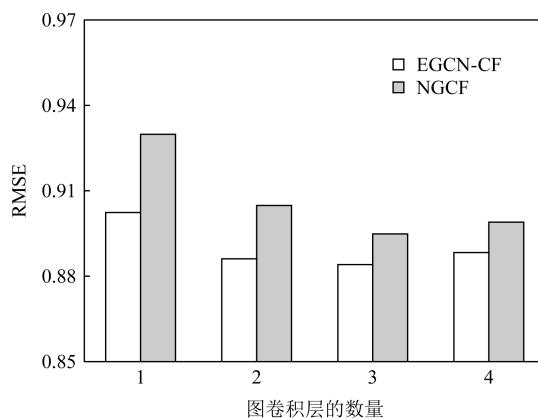


Fig. 3 Effect of graph convolutional layer numbers on ML-100K

图3 在ML-100K数据集上图卷积层数量的影响

分析图2,3中的结果,可以观察到:

1) 2种模型的预测误差均明显受到图卷积层数量的影响,在2个数据集上呈现相似的变化趋势:预测误差先是随着层数增加而明显降低,但在3层以后,不再明显降低并出现小幅增加的现象。这说明堆叠过多的图卷积层容易产生过拟合问题。因此,本文模型的最大层数设置为3。

2) 在各种情况下,本文模型的预测误差均明显优于NGCF。特别是当图卷积层的数量较少时,其优势更加明显。该现象得益于EGCN-CF采用的增强图卷积层的优势,它通过在层内聚合2阶协同信号和融入用户观点,更合理地聚合了协同信号,提高了节点嵌入学习的准确性,因而在图卷积层较少的情况下也能获得优异的预测准确率。

## 4 总 结

本文以推荐系统中的评分预测问题为对象,提出一种端到端的、基于增强图卷积神经网络的协同推荐模型。它通过增加2阶协同信号和融入用户观点,得到一种增强的图卷积层,并堆叠多个层次更合理地捕获二分图上的协同信号,提高对节点嵌入学习的准确性,最后利用非线性MLP网络实现评分预测。在多个推荐数据集上的实验结果表明,本文模型的预测准确率优于现有的推荐模型。

本文模型在运行时需要存储二分图的邻接矩阵,在超大规模图上可能遭遇存储瓶颈。在未来的研究工作中,我们将研究借鉴文献[11]的基于随机游走的节点采样技术,对本文模型进行改进,使其能够适应超大规模的推荐应用。

## 参 考 文 献

- [1] Nilashi M, Bagherifard K, Ibrahim O, et al. Collaborative filtering recommender systems [J]. Research Journal of Applied Sciences, Engineering and Technology, 2013, 5(16): 4168-4182
- [2] Sarwar B, Karypis G, Konstan J, et al. Item-based collaborative filtering recommendation algorithms [C] //Proc of the 10th Int Conf on World Wide Web. New York: ACM, 2001: 285-295
- [3] Belkin N J, Croft W B. Information filtering and information retrieval: Two sides of the same coin? [J]. Communications of the ACM, 1992, 35(12): 29-38
- [4] Koren Y, Bell R, Volinsky C. Matrix factorization techniques for recommender systems [J]. Computer, 2009, 42(8): 30-37
- [5] He Xiangnan, Liao Lizi, Zhang Hanwang, et al. Neural collaborative filtering [C] //Proc of the 26th Int Conf on World Wide Web. New York: ACM, 2017: 173-182
- [6] Shani G, Heckerman D, Brafman R I, et al. An MDP-based recommender system [J]. Journal of Machine Learning Research, 2005, 6(1): 1265-1295
- [7] Zheng Lei, Noroozi V, Yu P S. Joint deep modeling of users and items using reviews for recommendation [C] //Proc of the 10th ACM Int Conf on Web Search and Data Mining. New York: ACM, 2017: 425-434
- [8] Zheng Yin, Tang Baosheng, Ding Wenkui, et al. A neural autoregressive approach to collaborative filtering [C] //Proc of the 33rd Int Conf on Machine Learning. New York: ACM, 2016: 764-773
- [9] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks [C/OL] //Proc of Int Conf on Learning Representations. 2017 [2020-08-05]. <https://arxiv.org/abs/1609.02907>
- [10] Berg R, Kipf T N, Welling M. Graph convolutional matrix completion [C/OL] //Proc of the KDD Deep Learning Day. 2018 [2020-08-05]. <https://arxiv.org/abs/1706.02263>
- [11] Ying R, He Ruining, Chen Kaifeng, et al. Graph convolutional neural networks for web-scale recommender systems [C] // Proc of the 24th ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining. New York: ACM, 2018: 974-983
- [12] Xu Bingbing, Cen Keting, Huang Junjie, et al. A survey on graph convolutional neural network [J]. Chinese Journal of Computers, 2020, 43(5): 755-780 (in Chinese)  
(徐冰冰, 岑科廷, 黄俊杰, 等. 图卷积神经网络综述[J]. 计算机学报, 2020, 43(5): 755-780)
- [13] Hamilton W L, Ying R, Leskovec J. Inductive representation learning on large graphs [C] //Proc of the 31st Int Conf on Neural Information Processing Systems. Cambridge, MA: MIT Press, 2017: 1025-1035
- [14] Wu F, Zhang Tianyi, Souza A H, et al. Simplifying graph convolutional networks [C] //Proc of the 36th Int Conf on Machine Learning. New York: ACM, 2019: 6861-6871

- [15] Wang Xiang, He Xiangnan, Wang Meng, et al. Neural graph collaborative filtering [C] //Proc of the 42nd Int ACM SIGIR Conf on Research & Development in Information Retrieval. New York: ACM, 2019: 165–174
- [16] Chen Lei, Wu Le, Hong Richang, et al. Revisiting graph based collaborative filtering: A linear residual graph convolutional network approach [C] //Proc of the 34th AAAI Conf on Artificial Intelligence. Menlo Park, CA: AAAI Press, 2020: 27–34
- [17] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278–2324
- [18] He Xiangnan, Deng Kuan, Wang Xiang, et al. LightGCN: Simplifying and powering graph convolution network for recommendation [C] //Proc of the 43rd Int ACM SIGIR Conf on Research & Development in Information Retrieval. New York: ACM, 2020: 639–648
- [19] Song Weiping, Xiao Zhiping, Wang Yifan, et al. Session-based social recommendation via dynamic graph attention networks [C] //Proc of the 12th ACM Int Conf on Web Search and Data Mining. New York: ACM, 2019: 555–563
- [20] Fan Wenqi, Ma Yao, Li Qing, et al. Graph neural networks for social recommendation [C] //Proc of the 28th Int Conf on World Wide Web. New York: ACM, 2019: 417–426
- [21] Wang Xiang, He Xiangnan, Cao Yixin, et al. KGAT: Knowledge graph attention network for recommendation [C] //Proc of the 25th ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining. New York, ACM, 2019: 950–958
- [22] Wang Jizhe, Huang Pipei, Zhao Huan, et al. Billion-scale commodity embedding for E-commerce recommendation in Alibaba [C] //Proc of the 24th ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining. New York: ACM, 2018: 839–848
- [23] Zhang Jiani, Shi Xingjian, Zhao Shengli, et al. STAR-GCN: Stacked and reconstructed graph convolutional networks for recommender systems [C] //Proc of the 28th Int Joint Conf on Artificial Intelligence. Menlo Park, CA: AAAI Press, 2019: 4264–4270
- [24] Yang J H, Chen C M, Wang C J, et al. HOP-rec: High-order proximity for implicit recommendation [C] //Proc of the 12th ACM Conf on Recommender Systems. New York: ACM, 2018: 140–144
- [25] Ge Yao, Chen Songchan. Graph convolutional network for recommender systems [J]. Journal of Software, 2020, 31(4): 1101–1112 (in Chinese)  
(葛尧, 陈松灿. 面向推荐系统的图卷积网络[J]. 软件学报, 2020, 31(4): 1101–1112)
- [26] Gilmer J, Schoenholz S S, Riley P F, et al. Neural message passing for quantum chemistry [C] //Proc of the 34th Int Conf on Machine Learning. New York: ACM, 2017: 1263–1272
- [27] Xu K, Li Chengtao, Tian Yonglong, et al. Representation learning on graphs with jumping knowledge networks [C] //Proc of the 35th Int Conf on Machine Learning. New York: ACM, 2018: 5449–5458
- [28] Xiang Rongjing, Neville J, Rogati M. Modeling relationship strength in online social networks [C] //Proc of the 19th Int Conf on World Wide Web. New York: ACM, 2010: 981–990
- [29] Easley D, Kleinberg J. Networks, Crowds, and Markets: Reasoning About a Highly Connected World [M]. Cambridge, UK: Cambridge University Press, 2010
- [30] Wang Jun, De Vries A P, Reinders M J. Unifying user-based and item-based collaborative filtering approaches by similarity fusion [C] //Proc of the 29th Int ACM SIGIR Conf on Research & Development in Information Retrieval. New York: ACM, 2006: 501–508
- [31] Bruna J, Zaremba W, Szlam A, et al. Spectral networks and locally connected networks on graphs [C/OL] //Proc of Int Conf on Learning Representations. 2014 [2020-08-05]. <https://arxiv.org/abs/1312.6203>
- [32] Kingma D P, Ba J. Adam: A method for stochastic optimization [C] //Proc of Int Conf on Learning Representations. 2015 [2020-08-05]. <https://arxiv.org/abs/1412.6980>
- [33] Monti F, Bronstein M, Bresson X. Geometric matrix completion with recurrent multi-graph neural networks [C] //Proc of the 31st Int Conf on Neural Information Processing Systems. Cambridge, MA: MIT Press, 2017: 3700–3710
- [34] Grover A, Leskovec J. Node2vec: Scalable feature learning for networks [C] //Proc of the 22nd ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining. New York: ACM, 2016: 855–864



**Wang Lei**, born in 1978. PhD, associate professor. Member of CCF. His main research interests include machine learning, recommendation system, data mining.

王磊,1978年生.博士,副教授,CCF会员.主要研究方向为机器学习、推荐系统、数据挖掘.



**Xiong Yunying**, born in 1979. PhD candidate, associate professor. Her main research interests include machine learning, intelligence business data mining.

熊于宁,1979年生.博士研究生,副教授.主要研究方向为机器学习、智能商务数据挖掘.



**Li Yunpeng**, born in 1996. Master candidate. His main research interests include data mining, deep learning.

李云鹏,1996年生.硕士研究生.主要研究方向为数据挖掘、深度学习.



**Liu Yuanyuan**, born in 1997. Master candidate. Her main research interests include deep learning and anomaly detection.

刘媛媛,1997年生.硕士研究生.主要研究方向为深度学习和异常检测.