Web信息处理课程报告——用于学术推荐的异质图神经协同过滤

信息学院 2020000889 赵海源

1. 项目背景

近年来，为了满足用户的个性化需求与解决信息过载的问题，个性化推荐系统变得无处不在，在诸如音乐、图书、电影、电商、社交媒体等不同场景下为用户推荐其潜在喜欢的物品。虽然经典的协同过滤模型（CF, Collaborative Filtering在大部分场景当中已经取得了很好的效果，但是在真实的推荐场景下还存在着一些对推荐能够起到一定辅助作用的信息，而这些信息尚未被传统的协同过滤模型所利用。以学术领域的个性化推荐为例，除了学术文献和用户的直接交互之外，在学术数据当中还存在着大量的高阶交互信息，例如在图1当中，“用户-文章-作者-文章”就形成了一个高阶交互的实例。这表明一个用户在看了某一篇文章之后可能会去看这篇文章的作者所写的其他文章。从过往的研究来看，这种高阶交互信息往往能够捕捉到用户和物品之间的潜在交互。同时，除了用户和文献的交互之外，还包括有文章之间的引用关系、作者之间的合作关系、文章和期刊之间的隶属关系、作者和机构之间的隶属关系等等。这些辅助关系则包含着很多对推荐有用的信息，而这些关系当中所包含的信息则被称为异质信息 (Heterogeneous Information)。

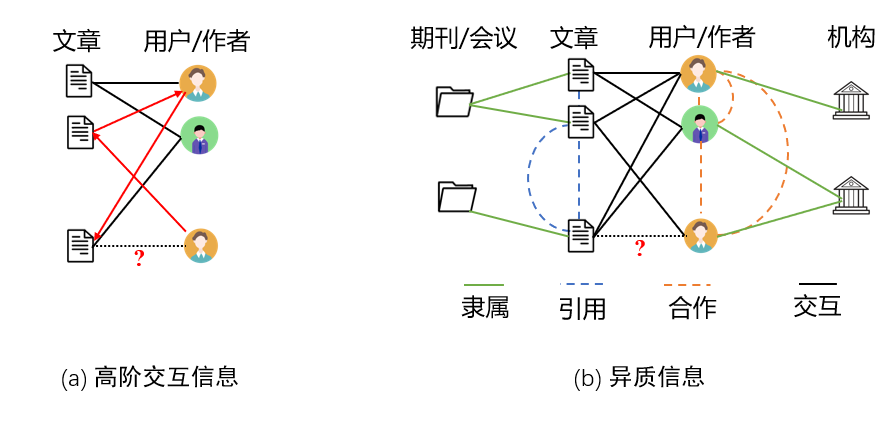


图 1 个性化推荐当中的高阶交互信息和异质信息

由于高阶交互信息和异质信息对于协同过滤模型推荐效果的增益，近年来，如何在协同过滤模型当中引入并利用这两种信息成为了研究者所关注的研究热点。在利用高阶交互信息的研究中，早期一部分研究者在“用户-物品”图上使用随机游走来采样高阶邻居，然后利用标签传播的方式捕获高阶信息，从而增强模型的表现。而在最近的研究当中，一些研究者将图卷积网络（GCN, Graph Convolutional Network）应用到“用户-物品”图上以直接捕获高阶交互信息，并取得了超过随机游走方法的效果。在利用异质信息的研究当中，目前大部分方法都是以元路径（Meta-Path）为基础而构建的。这些方法通过基于元路径的随机游走，为用户之间、物品之间或者一些没有交互的用户与物品之间建立起语义关联，从而使得模型更容易学习到不同用户和物品的相似性，提升推荐的效果。

尽管研究者在以上两个方面的研究取得了巨大的进步，但是当前的研究始终将引入高阶交互信息和引入异质信息这两者分割为两个不同的问题来看待。事实上，这两种信息在提升协同过滤模型效果的内在机制上是一致的。总的来说，传统的协同过滤模型仅仅建模历史的“用户-物品”的直接交互，然后在多个“用户-物品对”的优化当中通过距离的传递性使得某些潜在的物品在向量空间当中的距离足够靠近目标用户。然而，这种间接地距离传递极容易丢失协同信息，即在多个“用户-物品对”之间的距离拉近过程当中很难保证每一个都能够成功地拉近距离。而高阶交互信息和异质信息的引入则很好地解决了传统协同过滤的这一个问题。一方面，通过找到高阶/异质的同类型结点邻居，使得模型在每一次优化当中都可以直接拉近潜在物品表示和目标用户表示之间地距离，从而显式地传递了协同信息。另一方面，通过找到高阶/异质的不同类型结点邻居，使得模型能够在训练时就“看见”一些测试集实例，即在训练时就已经将测试集样本泄漏给了模型，从而提高了模型的性能。综上可见，向协同过滤模型当中引入的高阶交互信息和异质信息起作用的内在机制是一致的。但是，因为两种信息寻找额外邻居结点的方式不同，所以其寻找到的额外邻居结点重复较少，从而为模型带来更多的提升。因此，将两种信息结合起来引入到协同过滤模型当中就显得自然且必要。

1. 模型架构

总的来说，本研究设计了一种用于个性化推荐的异质图神经协同过滤模型(Neural Heterogeneous Graph Collaborative Filtering，NHGCF)。对于每一层NHGCF，首先将图卷积操作直接应用到推荐异质图上进行领域聚合。但是考虑到异质结点之间向量表示的差异，NHGCF针对不同的邻居结点所传递过来的信息使用了关系感知的消息传递方式来代替了原始图卷积网络的消息传递参数共享机制。同时，为了减少运算消耗和信息的损失，本研究将“用户-异质中介结点”和“物品-异质中介结点”这类关系之间的信息传递分成两种有向操作——有向消息传递（Directed Message Passing）和有向消息存储（Directed Message Storage）。当信息从异质中介结点传入到用户或物品结点时使用消息传递操作，而当信息从用户或物品结点传入到异质中介结点中时使用消息存储操作。然后，为了对从不同领域结点所传入的信息进行筛选，本研究设计了关系注意力机制，旨在为通过不同关系所传递过来的信息赋予不同的权重，从而实现筛选。NHGCF将堆叠多层以捕捉不同阶数的高阶交互信息和异质信息，同时将每一层所学习到的用户/物品向量分别拼接起来作为最后的表示。最终，模型使用点积作为相似性度量方法，输出用户向量和物品向量点积后的数值作为模型的打分输出。整个模型架构如图所示。

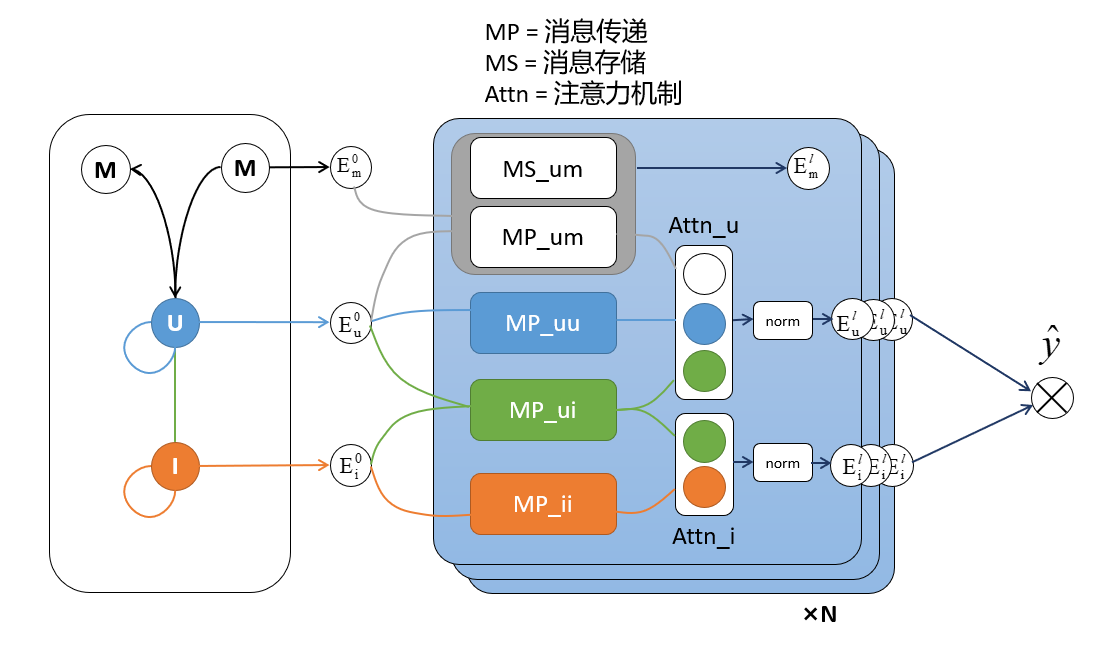


图 2 模型结构

**2.1 嵌入层**

和现有的工作类似，NHGCF首先将用户、物品还有其他中介结点的ID映射为向量表示：，，。其中是向量表示的维度。事实上中介结点m可能有多个类型，为了简化说明，在本节的后面叙述当中只引入一个类型的中介结点。嵌入层的功能是将这些向量表示拼接起来作为查找表，这样每输入一个id就可以得到对应的向量表示作为输出：



(2-1)



(2-2)



(2-3)



在通过嵌入层获得初始化向量表示之后，NHGCF则按照GCN的形式进行消息传递以同时捕获高阶交互信息和异质信息。然而，和GCN所处理的同质图不同的是，NHGCF所处理的异质交互图具有多种结点和多种关系，而不同类型的结点通过不同类型的关系所传递的信息是不同的。因此，本研究提出了关系感知的消息传递机制来处理这种异质交互图。具体来说，本研究认为，不同类型结点的向量表示不应该成为其固有的不变属性，而是应该随着该结点通过不同关系向其他结点传递信息时呈现出不同的效果。因此，关系感知的消息传递对不同的关系使用不同的线性变换参数和不同的归一化参数。

**2.2 关系感知消息传递**

对于一个用户，在一个异质交互图当中，他可能面临三种类型的关系，即：，，。按照关系感知消息传递的思路，每一种关系应该构建不同的消息传递参数。



(2-4)



(2-5)



(2-6)



其中，，，表示不同关系的消息传递函数，，，则表示这些消息传递函数所对应的参数。表示用户的向量表示。，，表示与用户直接相连的其他类型邻居结点的向量表示。这些原始的向量表示在通过消息传递函数的处理之后又经过了激活函数的筛选，最终得到输出向量。，，则表示经过消息传递之后所得到的新的用户向量表示。之所以选用作为激活函数，则是因为其既可以对消息传递变换后的向量表示完整地保留正值又允许少量的负值通过，从而减少了所导致的难以训练的问题。同理，对于物品，也有同样的消息传递方式：



(2-7)



(2-8)



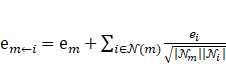
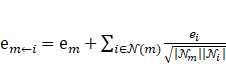
(2-9)



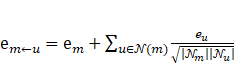
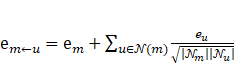
特别地，由于本研究将异质交互图视作无向图，所以和。但是，对于由用户结点、物品结点向中介结点传递信息的过程，即和，学习中介结点的表示和就显得没有十分赘余。从根本上来说，我们的目标是在用户和物品之间传递信息，而去学习中介结点m的表示不仅会使得模型计算量增加，而且中介结点的表示还会干扰从用户和物品传来的信息，使得其他用户和物品在从中介结点接收信息的时候难以间接获得协同信号。针对这一个问题，本研究将与、与的消息传递变成一个有向的过程，其中当进行和时按照公式（3-6）和公式（3-9）进行消息传递。而对于和则定义了一种新的操作，称作“消息存储”（Message Storage）：



(2-10)



(2-11)

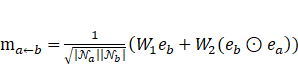
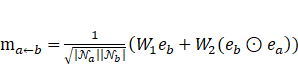


可见，当进行和的信息传递时，NHGCF不对聚合的结点表示做任何线性变换和激活函数的处理，只是简单地将中心结点m表示为其自身表示及周围结点表示的线性组合。这样一来，既能够减少所需要学习的参数，也能够降低经过中介结点传递协同信号的过程当中的信息损失。



对于本研究所提出的NHGCF模型而言，其消息传递函数是可以任意选择的，例如GCN，GAT，GraphSAGE等研究都分别提出了不同的消息传递函数。为了简化研究问题，本研究采用了NGCF所提出的消息传递函数：

(2-12)



(2-13)

(2-14)



对于从b到a的消息传递而言，该消息传递首先通过输入的和构建了消息，其中和是线性变化参数，负责将向量投影到不同的空间当中而改变他们之间的距离分布。特别地，由于在交互图和异质交互图当中a和b往往不是同一类型的结点，所以相比于传统GCN的消息构建方式，NGCF所提出的消息构建方式还考虑了中心结点和周围结点向量表示的交互来使得中心结点从更相似的周围结点当中获取更多信息。这一交互由进行捕获，其中表示逐元素点乘。和是通过拉普拉斯矩阵计算的归一化系数，用来对周围结点的信息进行加权，既可以防止度数大的结点具有更多的信息，也可以根据度数来给重要的结点信息赋予更大的权重。同时，为了防止GCN带来的过平滑问题，NGCF的消息构建方式还考虑了中心结点自身的信息的重要性，通过保留了中心结点自身的信息。最后将中心结点自身的信息，和周围结点的信息都加起来作为总的传入中心结点a的信息。



**2.3 关系注意力机制**

在经过关系感知消息传递的处理后，用户和物品基于不同关系的消息传递得到了多组向量表示。而这些表示对于用户和物品而言，往往有着不同的重要程度。因此，NHGCF设计了一个关系注意力机制为用户和物品基于不同关系的消息传递得到的向量表示提供可训练的权重，从而达到信息筛选的目的。而为了叙述简洁，本节仅以用户为例来说明NHGCF的关系注意力机制，而物品的关系注意力机制与用户类似。

(2-15)



(2-16)



(2-17)



(2-18)



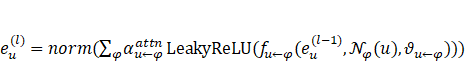
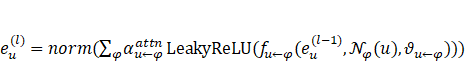
(2-19)



首先，模型将基于不同关系的消息传递得到了多组向量表示通过一个打分函数变换成分数。其中表示线性变换的参数，通过在不同向量当中共享同一个线性变换参数可以将原本处于不同空间的k维向量表示投影成同一个空间当中的a维向量，从而能够更好地衡量他们的得分。是一个注意力向量，用来对投影后的向量表示进行打分。，，是不同关系的消息传递所得的向量表示的得分，它们衡量了这些向量对于用户这一中心结点的重要性。而作为一个打分向量，也是在不同关系的表示间共享的。然后，使用了一个函数对各个得分进行归一化，所得的系数则是不同向量表示的权重。最后，使用这些权重对得到的向量进行加权求和。因为模型最终使用点积来预测打分，所以又对加权求和得到的向量进行了L2的归一化，从而得到本层最终输出的向量。而这一层的输出又可以继续输入到下一层作为输入，经过层数的叠加，NHGCF所能够捕捉到更高阶的信息。公式（2-20）描述了一层NHGCF的用户向量更新过程。其中代表与相连接的不同类型结点。



(2-20)



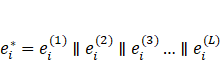
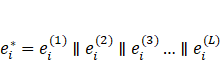
**2.4 模型预测**

由于NHGCF在每一层都会输出用户和物品的向量，而这些向量则聚合了不同阶邻居的信息，因此，参照NGCF，本研究所提出的模型对每一层的输出向量进行了拼接，作为最后输出的向量。

(2-21)



(2-22)



然后模型使用点积作为相似性度量方式来为用户和物品进行打分，作为模型最终输出：

(2-23)

值得注意的是，除了拼接之外，还可以选择组合不同层输出向量的方式，例如求和、求平均、最大池化等等；同时，除了点积之外，也可以在模型最后加入一个多层感知机(MLP)作为非线性的相似性度量函数。具体的改进方式因不同的数据集而异。

**2.5 模型优化**

为了优化所提出的模型，本研究使用贝叶斯个性化排名损失（BPR Loss，Bayesian Personalized Ranking Loss）来优化NHGCF。BPR损失是一种逐对的损失，与常用的逐点损失不同，BPR假设被观测到的样本（正例）应该比未被观测到的样本（负例）获得更多的评分，而模型应该尽量拉大这两者的评分差距。BPR损失的公式如下：

(4-24)



其中，表示成对的训练数据，表示观测到的样本（正例），表示没有被观测到的样本（负例）。表示sigmoid函数。表示模型当中所有可以训练的参数。表示L2正则项。控制了L2正则项的强度来防止过拟合。



1. 实验与分析

**3.1数据集**

为了验证所提出的模型的优势，本研究从ACM Digital Library上抓取了SIGIR，SIGKDD，CIKM，WWW，WSDM这5个会议从2010到2020年这10年间的会议论文集。一共抓取了其论文题目、摘要、作者、参考文献等信息。由于用户点击数据的无法获取，本研究以每一篇文章的参考文献作为文章作者（同时也是用户）点击过的文献，从而构造了一个学术推荐领域的异质交互图。

数据集的异质交互图模式被展示在了图3当中。其中U表示用户（user），P表示文章（paper）。各数据集的统计信息以及划分方式被展示在了表1当中。其中每个数据集的第一个关系表示用户-物品的关系。训练集、验证集和测试集的数量是加入了抽样的负例之后的样本。

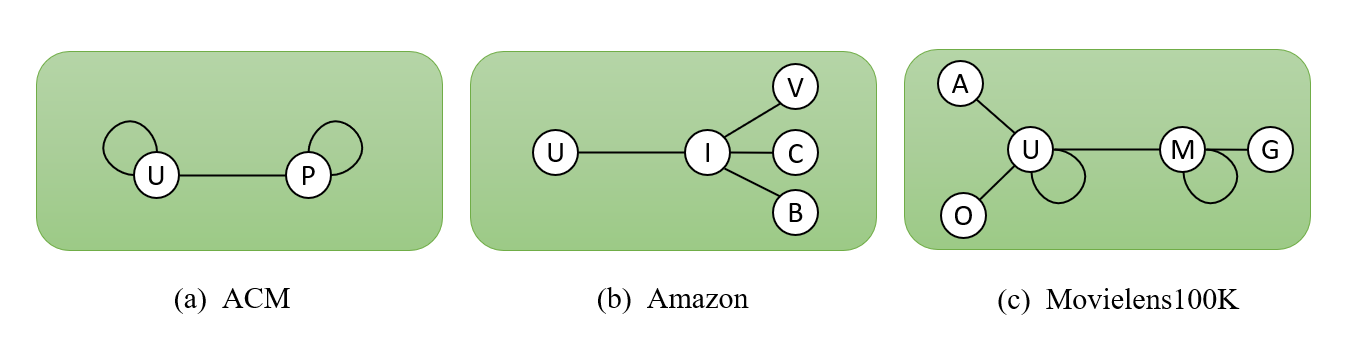


图 3 ACM数据集的异质图交互模式

表 1 ACM数据集的统计信息

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 数据集 | 关系A-B | A的数量 | B的数量 | A-B数量 | 训练集 | 验证集 | 测试集 |
| ACM | User-Paper | 18388 | 10669 | 121566 | 289296 | 203742 | 236376 |
| User-User | 18388 | 18388 | 136952 |
| Paper-Paper | 10669 | 10669 | 35587 |

由于未观测到的样本（负例）相比于观测到的样本（正例）的比例非常大，因此，本研究采用了抽样的方式来抽取未观测到的样本。具体来说，对于每个用户，从他未交互的物品当中按照不同物品出现的频率作为概率来进行不放回的抽样，对每一个正例抽取N个未观测到的样本作为该用户的负例。之所以按照不同物品出现的频率作为概率来进行不放回的抽样，是因为模型偏向于给用户推荐热门商品，而抽取出一些热门商品作为负样本则能够改善这种情况，从而有利于长尾物品的推荐。在本研究当中，N取为5.

最后，对于每一个用户，将抽样后的数据集依次均匀分配在训练集、验证集和测试集上，从而尽量保证每一个用户的交互数据都能在训练集、验证集和测试集当中出现。

**3.2基线模型**

为了和本研究的模型进行参照，从而突出其提升的效果和带来的优势。本研究使用了三个基线的推荐模型，分别代表传统的协同过滤模型、引入高阶信息的协同过滤模型以及引入异质信息的协同过滤模型。

**MF** 该模型通过贝叶斯个性化排名（BPR）损失优化的矩阵分解来进行推荐。仅利用用户-物品直接交互作为交互函数的目标值，并且以简单的点积作为相似性度量函数。本研究以这个模型作为传统协同过滤的代表。

**NGCF** 该模型是当前利用GCN做推荐的一个最先进的模型。NGCF通过利用在交互图上进行的消息传递来捕捉高阶交互信息，从而显式地传递了协同信号，在基于隐式反馈的推荐上取得了极大地提升。本研究将以这个模型作为引入高阶交互信息的协同过滤模型的代表。

**HAN** 该模型是当前基于元路径的表示学习模型当中最先进的模型。HAN通过基于元路径的游走来选择目标结点的元路径邻居，然后用结点间注意力来聚合周围的结点信息，最后通过元路径语义注意力来为不同的元路径下学习到的嵌入进修加权求和。在本研究当中，HAN被应用在异质交互图上，并被改造为可以用作推荐的形式。本研究以这个模型作为引入异质信息的协同过滤模型的代表。

**3.3参数设置**

在本节中，NHGCF和其他基线模型在各个数据集上所使用的训练参数被详细地展示在了表4当中，其中lr表示学习率，λ表示L2正则项的系数。为了公平起见，所有模型的嵌入维度都设定为64，所有的模型都移除了dropout这种提升泛化性能的辅助手段，同时所有的模型都将使用BPR作为损失函数，并用mini-batch Adam进行优化，batch size设定为128。为了得到最优的模型，本研究以验证集上的NDCG@5作为监测指标进行early stopping，当该指标连续20个epoch不上升的时候则停止训练，取之前最优的模型作为进行性能评估的模型。特别地，对于HAN，由于其要求利用所有元路径邻居来进行训练，而其元路径邻居往往数量十分巨大。因此，为了减少算力和时间消耗以及为了公平起见，我们对三个数据集上HAN的每一个元路径下的元路径邻居进行了采样。其中HAN在ACM数据集当中使用的元路径为AP，APP，AAP；在Amazon数据集当中使用的元路径为UI，UICI，UIVI，UIBI；在Movielens100K数据集当中使用的元路径为UM，UUM，UMM，UAUM，UOUM，UMGM。HAN的注意力头数按照其原论文当中所推荐的那样被设置为4，注意力投影维度设为128，而NHGCF的注意力投影维度被设置为64。

表 2 各个模型的训练参数

|  |  |  |
| --- | --- | --- |
| 模型 | ACM | |
| lr | λ |
| MF | 1.00E-02 | 1.00E-04 |
| NGCF | 1.00E-03 | 1.00E-02 |
| HAN | 5.00E-03 | 1.00E-04 |
| NHGCF | 5.00E-03 | 1.00E-04 |

**3.4实验结果**

在ACM数据集上的实验结果被展示在表3中。其中的每一个数据都取的是10次实验后的平均值。其中下划线表示在基线模型当中最好的结果，加粗表示在所提出模型当中最好的结果。特别地，对于本研究所提出的模型，表3报告了其在堆叠不同层数下的效果。可以看到我们所提出的方法在堆叠两层的时候可以取得很好的效果。

表 3 各个模型在ACM数据集上的表现

|  |  |  |
| --- | --- | --- |
| 模型 | ACM | |
| HR@5 | NDCG@5 |
| MF | 0.5423 | 0.7136 |
| NGCF | 0.5571 | 0.7147 |
| HAN | 0.5515 | 0.7097 |
| NHGCF  (layer=1) | 0.5692 | 0.7303 |
| NHGCF  (layer=2) | **0.5715** | **0.7314** |
| NHGCF  (layer=3) | 0.5699 | 0.7301 |
| NHGCF  (layer=4) | 0.5621 | 0.7108 |

**3.5 模型分析**

尽管本研究所提出的NHGCF在ACM数据集上取得了较好的效果，然而到现在为止，仍然缺乏对NHGCF的内在机制缺乏了解。因此在这一章当中，本研究将进一步对NHGCF的参数灵敏度进行进一步地研究。

从3.4节的实验结果来看，NHGCF的模型效果会受到层数的影响。而这也符号之前的研究得到的结论：图网络由于受到过平滑问题的影响，当层数堆叠过高之后，会导致图上的所有点的表示趋向于接近，从而导致其性能下降。由于不同的数据集上最优的模型层数是不同的，因此本研究建议在具体的场景当中需要对模型的层数这一参数进行仔细地选择，以保证最好的效果。如图4所示，本研究将NHGCF在不同层数下的性能变化展现了出来。

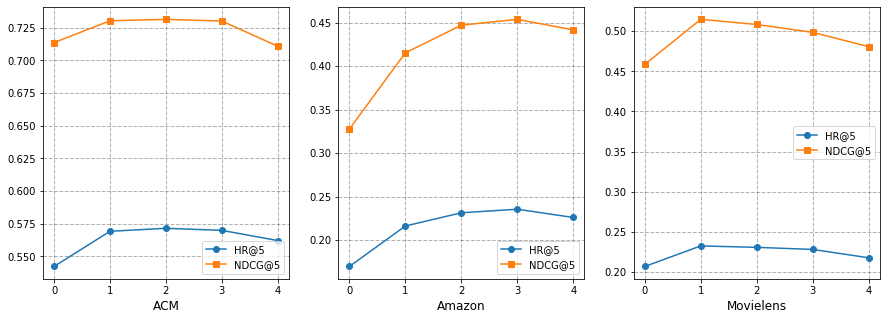


图 4 模型层数对NHGCF的影响

尽管本研究将所有的嵌入维度都固定在了64维，然而，嵌入维度对模型的影响却不可以忽略。从表示学习的角度来说，嵌入的维度大小意味着其参数空间的容量大小，从而也就意味着其中的向量能够表达的信息有多少。在图10当中，本研究展示了NHGCF在8,16,32,64,128这5个维度上的不同效果。可以看到，当模型嵌入维度刚刚开始增加时，模型的性能有了大幅度的提升，而这种增益随着维度的增长而放缓。而三个数据集当中，最优的维度都在64~128维左右。值得注意的是，虽然维度上升能够带来性能上的增益，然而随之而来的是参数量的上升。

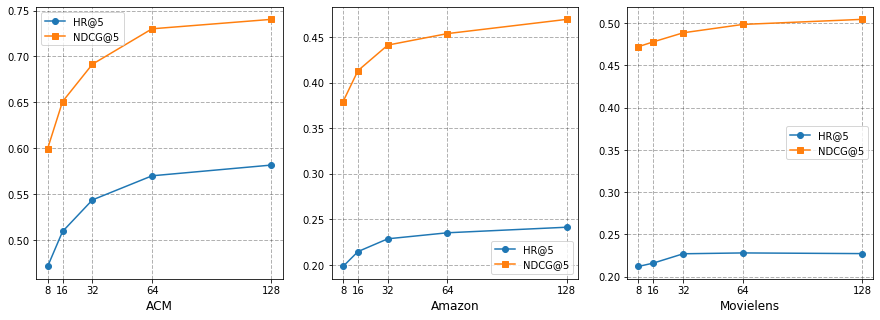


图 5 嵌入纬度对NHGCF的影响

1. 结论

总的来说，本研究提出了一种用于个性化推荐的异质图神经协同过滤模型(Neural Heterogeneous Graph Collaborative Filtering，NHGCF)。整个模型由嵌入层、关系感知消息传递、关系注意力机制和预测函数组成。相比于现有的协同过滤模型，该模型同时引入了高阶交互信息和异质信息提升了推荐效果，并且在模型结构层解决了基于元路径的模型的不足。本研究将所提出的模型在ACM的数据集上与现有的最先进模型进行了比较，并对其参数灵敏度进行了细致地分析。结果表明，我们的模型不仅在多个领域的推荐效果优于当前最先进的模型，而且模型内部的组件能够解决现有模型存在的问题。

参考文献

[1] 曾子明,金鹏.智慧图书馆个性化推荐服务体系及模式研究[J].图书馆杂志,2015,34(12):16-22.

[2] 李瑞敏,林鸿飞,闫俊.基于用户-标签-项目语义挖掘的个性化音乐推荐[J].计算机研究与发展,2014,51(10):2270-2276.

[3] 张富国.基于社交网络的个性化推荐技术[J].小型微型计算机系统,2014,35(07):1470-1476.

[4] Sarwar B, Karypis G, Konstan J, et al. Item-based collaborative filtering recommendation algorithms[C]// Proceedings of the 10th international conference on World Wide Web. New York, NY: ACM Press, 2001: 285–295.

[5] Lee C H，Kim Y H，Rhee P K． Web personalization expert with combining collaborative filtering and association rule mining technique[J]. Expert System with Applications,2001,21( 3) : 131 －137．

[6] Christoffel F, Paudel B, Newell C, et al. Blockbusters and Wallflowers: Accurate, Diverse, and Scalable Recommendations with Random Walks[C]//Proceedings of the 9th ACM Conference on Recommender Systems. New York, NY: ACM Press, 2015: 163–170.

[7] Schafer J B, Frankowski D, Herlocker J, et al.: Collaborative Filtering Recommender Systems, Brusilovsky P, Kobsa A, Nejdl W, editor[J]. The Adaptive Web, 2007: 291-324.

[8] Shi C, Li Y, Zhang J, et al. A Survey of Heterogeneous Information Network Analysis[J]. IEEE Transaction on Knowledge and Data Engineering , 2017, 29(1): 17–37.

[9] Cooper C, Lee S H, Radzik T, et al. Random walks in recommender systems: exact computation and simulations[C]//Proceedings of the 23rd International Conference on World Wide Web. New York, NY: ACM Press, 2014: 811–816.

[10] Francois Fouss, Alain Pirotte, Saerens M. A Novel Way of Computing Similarities between Nodes of a Graph, with Application to Collaborative Recommendation[J]. Web Intelligence, 2005: 550-556.

[11] Yang J-H, Chen C-M, Wang C-J, et al. HOP-rec: high-order proximity for implicit recommendation[C]//Proceedings of the 12th ACM Conference on Recommender Systems. New York, NY: ACM Press, 2018: 140–144.

[12] He X, Gao M, Kan M-Y, et al. BiRank: Towards Ranking on Bipartite Graphs[J]. IEEE Transaction on Knowledge Data Engineering, 2017, 29: 57-71.

[13] Kipf T N, Welling M. Semi-Supervised Classification with Graph Convolutional Networks[C]//Proceedings of the 5th International Conference on Learning Representations. ICLR, 2017.

[14] Ying R, He R, Chen K, et al. Graph Convolutional Neural Networks for Web-Scale Recommender Systems[C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York, NY: ACM Press, 2018: 974–983.

[15] Berg R V D, Kipf T N, Welling M. Graph Convolutional Matrix Completion[C]//Proceedings of the 23th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York, NY: ACM Press, 2017.

[16] Wang X, He X, Wang M, et al. Neural Graph Collaborative Filtering[C]//Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, NY: ACM Press, 2019: 165–174.

[17] Sun Y, Han J, Yan X, et al. PathSim: Meta Path-Based Top-K Similarity Search in Heterogeneous Information Networks[C]//Proceedings of 37th Very Large Data Bases. Trondheim, Norway: VLDB Endowment, 2011:992-1003

[18] Yu X, Ren X, Gu Q, et al. Collaborative filtering with entity similarity regularization in heterogeneous information networks[C]//Proceedings of 21st International Joint Conference of Artificial Intelligence (Heterogeneous Information Network workshop). Menlo Park, CA: AAAI Press, 2012.

[19] Shi C, Hu B, Zhao W X, et al. Heterogeneous Information Network Embedding for Recommendation[J]. IEEE Transactions on Knowledge and Data Engineering, 2019, 31(2): 357-370.