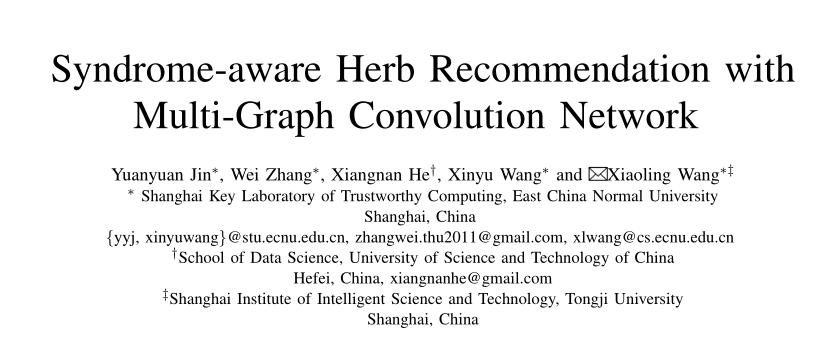
利用多图卷积网络来感知综合症的草药推荐方法

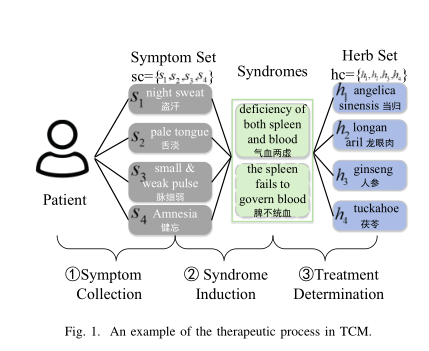




1. 论文简介

草药推荐问题，创新点受到传统中医治疗过程的启发，将由多个症状推导出综合症的过程引入进来，研究进而变为：

症状和草药的单一交互关系----->综合症与草药的交互关系

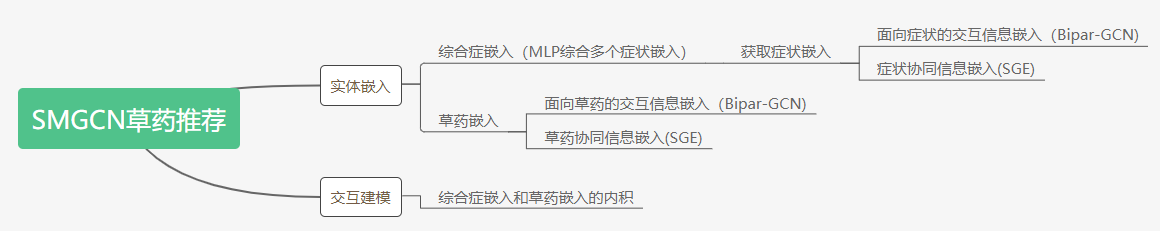


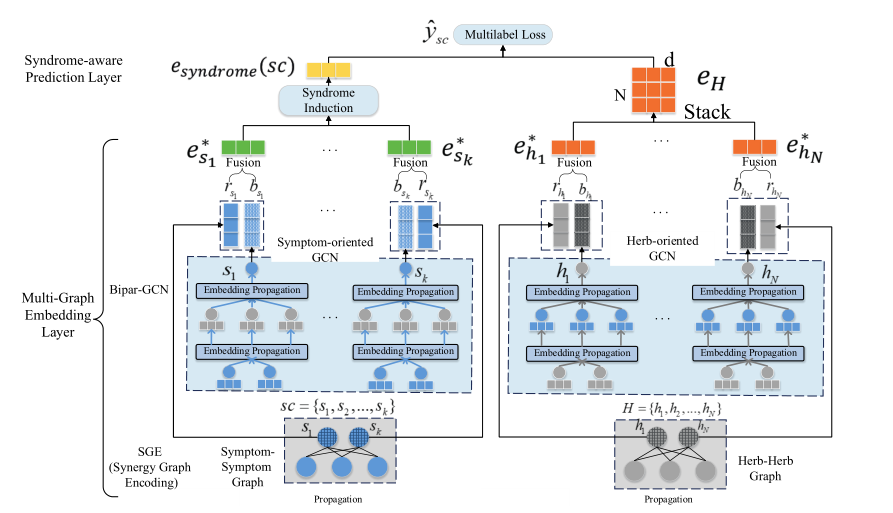
1. 论文具体思路

推荐任务可以概括为：

|  |  |  |  |
| --- | --- | --- | --- |
| 步骤 | 含义 | 领域范畴 | 具体技术 |
| 获取实体嵌入 | 获取能反映实体属性和实体间交互关系的嵌入表示 | 图表示学习 | 同质图的结点表示  异质图的结点表示 |
| 对交互关系建模 | 即构建交互函数 |  | 简单内积、  结合潜在因子分解和MLP（论文Neural Collaborative Filtering) |

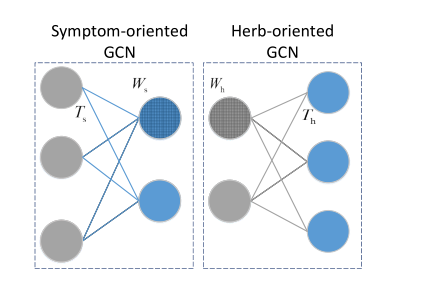
本论文的实验思路：





1. Bipar-GCN

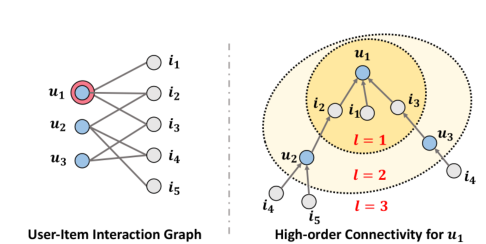
构建此组件的原因：建模草药和症状两种实体间的关系，

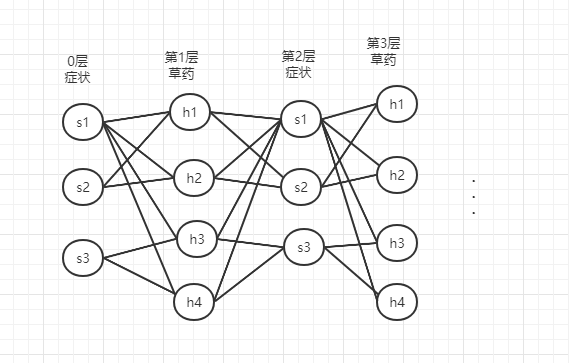
组件： 

说明：a.图卷积通常有两类，一类是谱域图卷积，即利用图上的傅里叶变化将空间域的卷积转化到谱域上，如SCNN、ChebNet、GCN；另一类是空域图卷积，即利用结点在空间中的邻居结点聚集信息，与矩阵的卷积很类似，如GraphSAGE、GAT、PGC方法；这里使用的是后一类

b.具体的卷积过程：选择邻居结点、构造消息、聚合消息、高阶传播







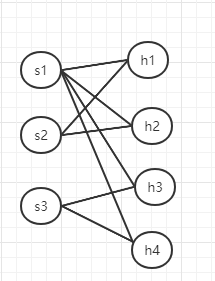
c.草药嵌入和症状嵌入的方法略有不同，区别在于第0层即初始嵌入层的对象不同；

1. SGE协同图编码

构建此组件的原因：考虑到推荐的草药集在药理上应该是相互兼容的，症状间存在继发、相互影响等联系，我们需要建模同种实体间的关系；

同时引入更多信息有助于缓解中药处方的数据稀疏性问题

不使用“二部图的2阶近邻”建模同质关系的原因：



p1=<{s1,s2},{h1,h2}>和p2=<{s1,s3},{h3,h4}>，

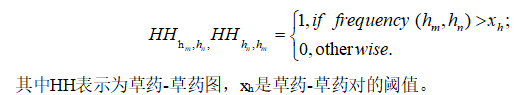
{h2,h3,h4}通过连接s1成为h1的2阶邻居。因此在协同图中，h1和h3、h1和h4没有边。

组件： 草药-草药图，症状-症状图

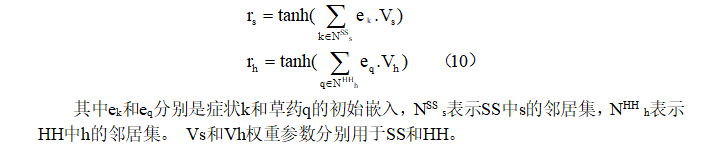
说明：a. 草药-草药图、症状-症状图实际上是同质图，不过在逻辑上可以看成是同类实体构成的特殊二部图；

b.将处方中草药间和症状间同时出现的次数作为衡量协同效果的强度，设置相应的阈值，超过阈值才被视作具有协同效果；

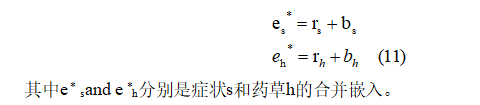
c.论文中使用邻接矩阵来表示图，其中具有协同效应记作1，而不是使用其次数n，我觉得使用次数更好；



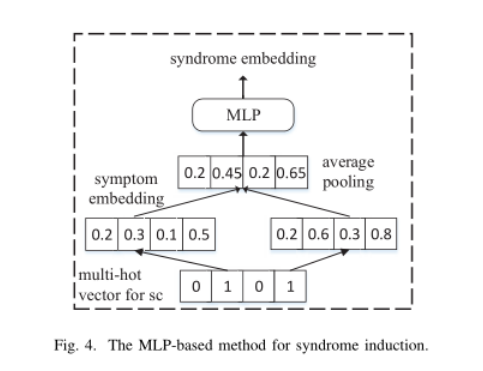
d.对协同图进行一层图卷积以期获得嵌入：



1. 信息融合：

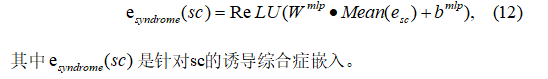


1. 综合症感知

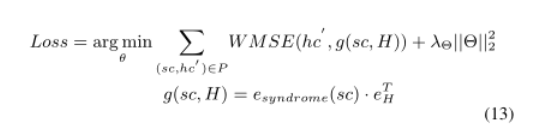


输入：所有症状的合并嵌入

处理过程： 一层平均池化层、一层MLP



1. 模型训练：多标签损失函数

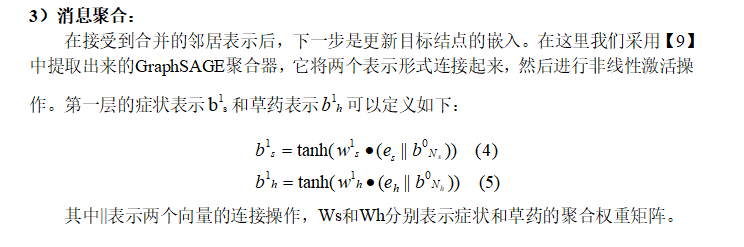


1. 论文实验部分

基线：

|  |  |  |
| --- | --- | --- |
| 基线种类 | 包含 | 含义 |
| 主题模型 | HC-KGETM | 不仅要考虑中医处方中的共现信息，还要考虑知识图谱中症状和草药的全面语义相关性 |
| 基于图神经网络的模型 | GC-MC  PinSage  NGCF | 分别利用GCN 获得用户和项目的表示、GraphSAGE 在项目-项目图上的工业应用、显式构造了二部用户-项目图以对高阶连通性进行建模 |
| 作者提出的基线 | HeteGNN | 基于内容的异质图表示学习问题，异质包含异质图结构信息和异质节点内容（比如文本、图像）两部分信息；  主要过程包括重定位的随机游走采样邻居节点、聚合异质的节点内容嵌入、聚合同一类型的邻居节点、聚合不同类型的邻居节点 |

1. 思考和发现
2. 论文中Bipar-GCN部分提到了GraphSAGE来聚合消息，公式中却只用了一个连接，但是GraphSAGE是对不同阶的邻居采用不同的聚合函数，并不用于嵌入连接，不知道其是否指的是GraphSAGE中使用的聚合方法；



1. 还没有推导Bipar-GCN的矩阵表示形式，论文中只给出的是单一结点的，代码中需要使用矩阵
2. 本文中Bipar-GCN所述方法引用了何向南团队2019年的一篇论文“Neural Graph collaborative filtering”，后找到今年的一篇论文也引用了并对这个框架做了一点泛化，后面可以试试使用它改进的Bipar-GCN:

