利用多图卷积网络来感知综合症的草药推荐方法

Published in: 2020 IEEE 36th International Conference on Data Engineering (ICDE)

Date of Conference: 20-24 April 2020

INSPEC Accession Number: 19760493

Date Added to IEEE *Xplore*: 27 May 2020

DOI: 10.1109/ICDE48307.2020.00020

ISBN Information:

Publisher: IEEE

ISSN Information:

Conference Location: Dallas, TX, USA, USA

摘要

草药推荐在中药（TCM）的治疗过程中起着至关重要的作用，TCM的目的是推荐一组可以治疗患者症状的草药。虽然已经开发了几种机器学习方法来推荐草药，但它们仅限于对草药和症状之间的相互作用进行建模，而忽略了综合症诱导的中间过程。进行中医诊断时，经验丰富的医生通常会根据患者的症状推断出综合症，然后根据诱发的综合症建议使用草药。因此，我们认为诱发的症状（可以看做是对症状的一种总体描述）对于草药推荐很重要，应适当处理。然而，由于综合症归纳过程的含混性和复杂性，大多数处方都缺乏综合症的明确事实依据。

在本文中，我们提出了一种新的方法，该方法将综合征的隐性诱导过程纳入草药推荐的考虑范围。具体来说，给定一组要治疗的症状，我们旨在通过有效融合所有症状在该组中的嵌入来生成总体的综合征表现，以模仿医生诱导出该综合征的过程。为了进行症状嵌入学习，我们还从输入处方中构造了一个症状-症状图，以捕获症状之间的关系（共发模式）。然后，我们在症状-症状和症状-草药图上建立图卷积网络（GCN），以学习症状嵌入。同样，我们构建了一个草药-草药图，并在草药和症状-草药图上构建了GCN，以学习草药嵌入，最后将其与症候表示交互用来预测草药的得分。这种多图GCN架构的优点是可以获得症状和草药的更全面的表示。我们在公共TCM数据集上进行了广泛的实验，与最先进的草药推荐方法相比，实验结果显示我们有了重大的进步。进一步的研究证明了我们的综合症表示和多图设计的有效性。

索引词-草药推荐，症状草药图，图神经网络，表示学习

I介绍

作为建立了数千年的古老而全面的治疗系统，中医在中国社会中起着至关重要的作用[1]。中医理论的基础是整体思想，它强调人体的完整性及其与自然环境的相互关系[2]。图1以经典的归脾汤处方为例，说明中药的三步治疗过程：（1）症状的征集。医生检查病人的症状。症状组sc包含“盗汗”，“苍白的舌头（舌淡）”，“小而微弱的脉搏（脉细无力）”和“健忘症”。 （2）综合症的诱导。在对症状进行全面分析后可以确定相应的综合征。在这种情况下，主要症候群是实圈圈出来的“脾脏和血液不足（脾血两虚）”。因为“脾脏不能治血”时也会出现“苍白的舌头”和“小而微弱的脉搏”，因此还存在图中虚线圈出来的一种可选综合症，称为“脾脏不能治血”。 （3）治疗方案的确定。医生选择一套草药作为治疗综合症的药物。在此步骤中同时考虑了草药的相

容性。此处的草药集由“人参”，“龙眼假种皮”，“当归”和“茯苓（食用菌核的一种）”组成。我们可以看到，综合症归纳中的第二步，系统地总结了症状，对于草药的最终推荐非常关键。然而，如以上示例所示，**同一症状可以在各种综合症中出现，这使得综合症的诱导变得模糊不清和复杂[3]。**实际上，对于某种症状集合，不同的中医医生可能会给出不同的综合症集合（如图1所示），因此不存在标准的基本事实。

在TCM处方语料库中，每个数据实例包含两个部分——一组症状和一组草药，这意味着该草药组可以很好地治愈症状组。为了推广到未知的症状集，**草药推荐任务着重于对症状和草药之间的交互进行建模，这类似于对用户与项目之间的交互进行建模的传统推荐任务[4]。**值得注意的是，一个主要区别是，**在传统推荐中，预测大多是在单用户级**别上进行的，而在草药推荐中，我们需要共同考虑一组症状来进行预测。由于缺乏综合症，有关草药推荐[5] – [8]的现有努力将综合症概念视为潜在主题。但是，他们仅在给定单一症状下的主题范围内，进行潜在综合征的学习。**特别是，他们专注于模拟一种症状和一种草药之间的相互作用，然后汇总来自多种症状的相互作用再对草药进行排名。**因此，症状集的集合信息被忽略了。

在本文中，我们建议将综合征的隐式诱导过程纳入草药推荐中。这符合我们从图1中获得的直觉，即综合征诱导是总结症状然后进行有效草药推荐的关键步骤。具体而言，给定一组要治疗的症状，我们的目标是在与草药进行交互以产生推荐之前，**根据症状的组成，来学习一种整体的隐式综合征表现形式。**通过这种方式，医生的处方行为能够被模仿。

为此，我们提出了一种名为“利用多图卷积网络（SMGCN）感知综合症”的新方法。该方法是一种多层神经网络模型，可对综合症和草药进行交互建模以完成草药推荐任务。在交互建模组件（SMGCN的顶层）中，我们首先**通过多层感知器（MLP）融合目标症状集中的症状嵌入**，以直接获得一个整体的隐式综合症表示形式，然后与草药嵌入进行交互输出预测分数。在嵌入学习组件（SMGCN的底层）中，我们**通过GCN在多个图上学习症状嵌入和草药嵌入**。具体来说，除了输入的症状-草药图，我们还根据处方条目中同时出现的症状（草药）来**构建症状-症状和草药-草药图**。直观上，一些症状会经常同时出现在身上患者（例如，恶心和呕吐），这种建模有利于症状表示学习；类似地，**草药-草药图证明了经常共存的草药，可用于编码它们的相容性**。我们在公共TCM数据集上进行了实验[5]，证明了我们SMGCN方法的整体有效性，并验证了每个单一意图设计的合理性。

这项工作的主要贡献如下：

•我们强调了综合症表示、对综合症与草药之间的相互作用进行建模以推荐草药的重要性。

•我们提出了SMGCN，它将MLP在融合建模（即将症状嵌入融合到整体的隐式综合征嵌入中）和GCN在关系数据学习（即学习症状和草药之间的嵌入）中的优势统一起来，以推荐草药。

•我们建立了草药-草药图和症状-症状图，以丰富草药和症状的关系，并将GCN扩展到多个图，以提高它们的表示学习质量。

本文的其余部分安排如下：第2节描述了问题的定义。第3节介绍了我们的总体框架。第4节介绍了我们提出的方法。第5节评估了我们的方法。第6节概述了相关工作。最后，第7节提供了一些总结性说明。

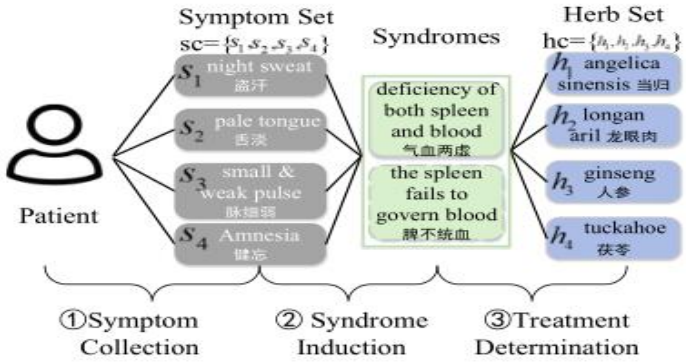


Fig. 1. An example of the therapeutic process in TCM.

Ⅱ问题定义

草药推荐的任务是针对特定症状集，产生具有相应治愈效果的草药集。草药推荐系统通常会从大型处方语料库中学习。令S = {s1，s2，...，sM}和H = {h1，h2，...，hN}分别表示所有症状和草药。每个处方由症状组和草药组组成，例如p = <{s1，s2，...}，{h1，h2，...} >。在综合症归纳过程中，我们需要为每个症状集诱导出一个总体的综合症表现形式，后续我们将使用它们生成适当的草药集。本段以后，我们将分别用sc = {s1，s2，...}和hc = {h1，h2，...}表示症状集和草药集。以这种方式，每个处方由p＝<sc，hc>表示。

给定一个症状集sc，我们的任务是计算一个N维概率向量，其中维i的值表示草药i可以治愈sc的概率。这是通过学习到的预测函数 $\hat{y}_{sc} = g(sc, H; \theta)$ 来实现的，其中 $\hat{y}_{sc}$ 表示概率矢量，θ 表示函数 g 的可训练参数。输入和输出的定义如下：

•输入：草药H，症状S，处方P。

•输出：学习到的函数 $g(sc, H; \theta)$，该函数为给定症状集sc下，对应于H中所有草药的一个概率矢量

Ⅲ方法概述：

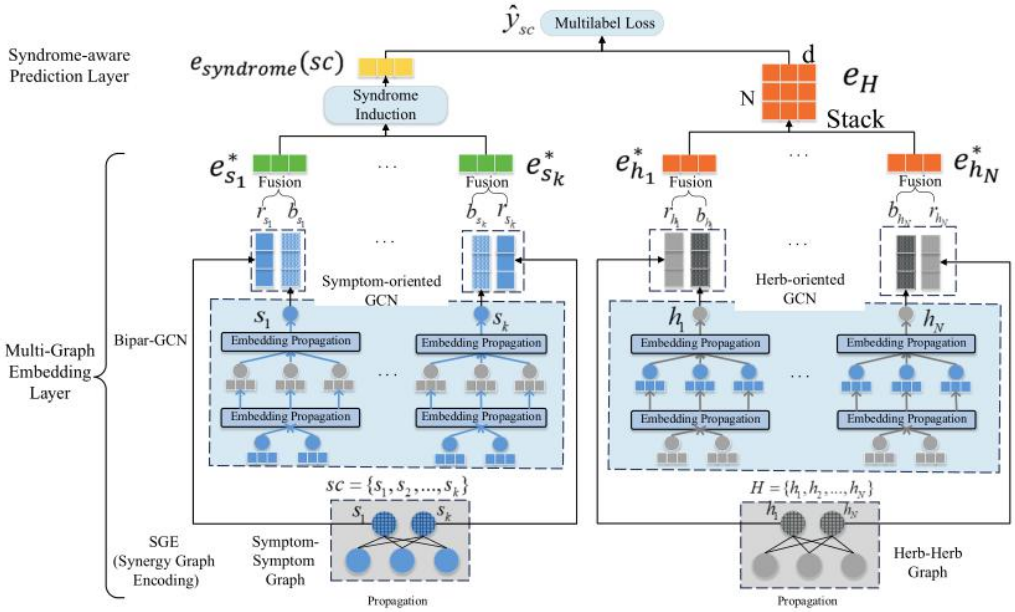


图2.我们提出的模型的总体架构（包括**Bipar-GCN，协同图编码（SGE）和综合症归纳（SI）**）。症状节点为蓝色，草药节点为灰色。显然，具有斜线的节点是目标节点。

在本节中，我们将详细讨论所提出的“可以感知综合症的多图卷积网络”框架，如图2所示。我们提出的模型采用一个单一症状集sc = {s1，s2，...，sk}和 所有草药H = {h1，...，hN}作为输入，并输出预测的维度为$|H|$的概率矢量$\hat{y}_{sc}$。在$\hat{y}_{sc}$中，位置i的值表示hi适用于治疗sc的概率。

为了完成此任务，它主要**由两层组成：多图嵌入层和综合症感知预测层**。

多图嵌入层：**该层旨在获取S的所有症状和H的所有草药的良好表达形式。**考虑到中医中的症状和草药之间的复杂相互关系，我们首先开发了二部图卷积神经网络（Bipar-GCN）以处理二部症状-草药图。为了捕获症状和草药之间的内在差异，Bipar-GCN分别对目标症状节点执行面向症状的嵌入传播，对目标草药节点进行针对草药的嵌入传播。通过这种方式，症状嵌入b_s和草药嵌入b_h被学习到了。其次，引入了协同图编码（SGE）组件来编码症状对和草药对的协同信息。对于症状对，它根据症状对同时出现的频率构造了一个症状-症状图，并在症状-症状图上执行图卷积以学习症状嵌入r_s。类似地，它也从草药-草药图学习草药嵌入r_h。第三，对于每种症状（草药），将Bipar-GCN和SGE中的两种类型的嵌入b和r融合在一起，以形成集成嵌入e *。

综合征感知预测层：在这一层中，牢记综合症诱导过程的重要性，**综合症诱导（SI）组件将症状集sc中所有症状的嵌入内容馈送到MLP中，以生成整体的综合症表示**$e_{syndrome}$**（SC）。其次，将所有草药嵌入堆叠到**e_H**中**，即N×d矩阵，其中d是每个草药嵌入的维度。综合症表示$e_{syndrome}$（sc）与e_H相互作用产生针对H中所有草药的概率得

分向量 $\hat{y}_{sc}$。

考虑到一组草药会被作为一个整体进行推荐，我们利用**多标签损失函数来优化我们提出的模型。** Tab I中汇总了本文中使用的所有符号。

TABLE I
SUMMARY OF ALL NOTATIONS

e_h, e_s	initial embeddings for herbs, symptoms
S, H	symptom collection and herb collection
SC, HC	collection of symptom sets and collection of herb sets
P	prescription collection
N_s, N_h	neighborhood of symptom, herb on the bipartite graph
SH	symptom-herb graph
SS, HH	symptom-symptom graph and herb-herb graph
x_s, x_h	threshold for constructing SS and HH
N_s^{SS}, N_h^{HH}	neighborhood of symptom on SS neighborhood of herb on HH
T_s^k, T_h^k	message construction function for symptom, herb at k-th Bipar-GCN layer
W_s^k, W_h^k	message aggregation function for symptom, herb at k-th Bipar-GCN layer
V_s, V_h	aggregation function for symptom on SS aggregation function for herb on HH
$b_{N_s}^k, b_{N_h}^k$	symptom, herb neighborhood embedding at k-th Bipar-GCN layer
b_s^k, b_h^k	symptom, herb output embeddings at k-th Bipar-GCN layer
r_s, r_h	symptom output embeddings on SS herb output embeddings on HH
e_h^*, e_s^*	herb, symptom final embedding after fusion
W^{mlp}, b^{mlp}	the MLP weight matrix and bias parameter used in Syndrome Induction
W^{att}, z	the attention network parameters in HeteGCN
$e_{syndrome}(sc)$	the induced syndrome embedding for symptom set sc
$\hat{y}(sc)$	the predicted probability vector for sc in dimension $\|H\|$

Ⅳ方法学：

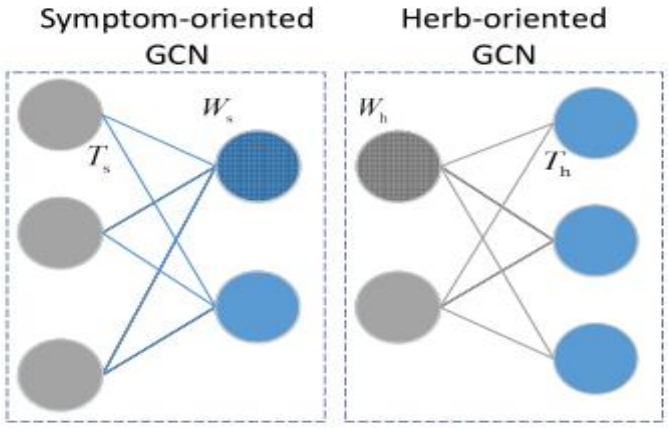


图3.二分GCN。蓝色边缘和灰色边缘表示不同的图卷积函数。带有斜线的节点是目标节

点。

A.二分图卷积网络[4]

最近的工作如[4]已经证明了在推荐系统中对用户-项目图执行图卷积的令人信服的性能。尽管它们有效，但我们认为它们忽略了二分图中两种类型的节点(用户和项目)之间的内在差异，并在图中使用了共享的聚合和转换函数，这可能会限制信息传播的灵活性，并在一定程度上影响嵌入的表达性。**为了对草药和症状之间的内在差异建模**，我们利用了图3所示的Bipar-GCN方法。当目标节点的类型为“症状”时，左侧面向症状的GCN将用于获取该目标节点的表示。否则，右侧面向草药的GCN将被用来学习结点的嵌入。**这两个部分共享同一个症状-草药的拓扑结构，但采用不同的聚合和转换函数。**通过这种方式，不同类型的节点可以**开发自己的传播灵活性，从而学习到更有效的表达**。接下来我们将详细介绍Bipar-GCN。

1）症状-草药图的构建:

拿一个中药处方p=<sc={s1，s2，...，sk}，hc={h1，h2，...，hm}>举例，同一个方子里的症状和草药是相互关联的。因此，{(s1,h1)、...,(s1,hm),...,(sk,h1),...,(sk，hm)}构成图的边。我们将症状-草药图视为无向图，其公式如下:

$$SH_{s,h}, SH_{h,s} = \begin{cases} 1, if\ (s,h)\ co-occur\ in\ prescriptions; \\ 0, otherwise \end{cases}$$

其中SH表示症状-草药图。

2）消息构造：

为了将信息从每个邻居节点传播到目标节点，需要定义两个操作：如何生成每个节点传输到目标节点的信息以及如何将多个邻居消息合并在一起

对于症状s，它的单跳邻居草药h传送给它的消息被定义为m_{h}，

$$\mathrm{m}^{0}{}_{\mathrm{h}} = e_h \bullet T_s^1(1)$$

e_h是草药h的初始嵌入，$T^1{}_{\mathrm{s}}$是第一层（症状）的转换权重矩阵。在收集到所有邻居的信息后，我们使用平均操作去合并它们，定义如下：

$$B^0{}_{N_s} = \tanh(\frac{1}{|N_{\mathrm{s}}|}\sum_{h\in N_{\mathrm{s}}} m^0{}_{\mathrm{h}})$$

其中N_{s}是s的一跳邻居集，然后我们使用tanh作为激活函数。类似地，对于草药h,其合并的一跳邻居信息可以表示为：

$$B^0{}_{N_{\mathrm{h}}} = \tanh(\frac{1}{|N_{\mathrm{h}}|}\sum_{\mathrm{s}\in N_h} m^0{}_s)$$

其中$\mathrm{m}^0{}_{\mathrm{s}} = e_s \bullet T^1{}_h$

3）消息聚合：

在接受到合并的邻居表示后，下一步是更新目标结点的嵌入。在这里我们采用【9】中提取出来的GraphSAGE聚合器，它将两个表示形式连接起来，然后进行非线性激活操作。第一层的症状表示$b^1{}_s$和草药表示$b^1{}_h$可以定义如下：

$$b^1{}_s = \tanh(w^1{}_s \bullet (e_s \,\|\, b^0{}_{N_s})) \quad (4)$$

$$b^1{}_h = \tanh(w^1{}_h \bullet (e_h \,\|\, b^0{}_{N_h})) \quad (5)$$

其中||表示两个向量的连接操作，Ws和Wh分别表示症状和草药的聚合权重矩阵。

4）高阶传播：

我们可以进一步将单条传播规则拓展到多层。具体来说，在k层中，

我们递归地将草药h表示为：

$$b^k{}_h = \tanh(w^k{}_h \bullet (b^{k-1}{}_h \,\|\, b^{k-1}{}_{N_h})) \quad (6)$$

其中对于h其来自第k层的邻居的消息定义如下：

$$b^{k-1}{}_{N_h} = \tanh(\frac{1}{|N_h|}\sum_{s\in N_h} b^{k-1}{}_s T^k{}_h) \quad (7)$$

对于症状s,公式是类似的：

$$b^k{}_s = \tanh(w^k{}_s \bullet (b^{k-1}{}_s \,\|\, b^{k-1}{}_{N_s})), \quad (8)$$

其中对于s其在第k层传播的信息可以定义如下：

$$b^{k-1}{}_{N_s} = \tanh(\frac{1}{|N_s|}\sum_{h\in N_s} b^{k-1}{}_h \bullet T^k{}_s). \quad (9)$$

B. 协同图编码层

除了症状和药物的关系，症状和药物之间也还存在其他的一些协同模式。给定一个处方p=<sc,hc>, 症状集sc中的症状不是相互独立却是相互关联的，hc中的草药也相互影响并且形成一个完整的组合物。因此，这些关系可以分别用来构造症状和草药的协同图。值得注意的是，**虽然症状-草药图上的二阶信息传播可以捕获症状和草药的同质关系，但是二阶的症状-症状和草药-草药的链接不等同于处方中的并发对**。例如，在处方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。另一方面，很明显，二部的症状草药图不能直接从同质协同图中导出。因此，我们得出结论，症状-草药图和协同图都有它们自己的特点，并且可以相互补充去学习更精确的结点表示。

1)协同图的构建：

通常，草药和症状间的协同模式可以从它们的高共生频率反映出来。以构造草药-草药图为例，我们首先计算处方中所有草药-草药对的频率：如果草药h_m和h_n出现在同一个草药集合hc中，则（h_m,h_n)对的频率增加一。在获取草药-草药的频率矩阵之后，我们

可以手动地设置阈值x_h来过滤条目。对于同时出现次数超过x_h的对，将相应的条目设置为1，不然设置为0.可以公式化如下：

$$HH_{h_m,h_n,}HH_{h_n,h_m}=\begin{cases}1,if\ \ frequency\ (h_m,h_n)>x_h;\\0,\text{other}wise.\end{cases}$$

其中HH表示为草药-草药图，xh是草药-草药对的阈值。

通过参考上述过程，也可以构造症状-症状图。

2）信息传播：

给定构造好的草药-草药图HH和症状-症状图SS，我们使用一层的图卷积网络来产生症状嵌入和草药嵌入：

$$r_s=\tanh(\sum_{k\in N^{SS}{}_s}e_k.V_s)$$

$$r_h=\tanh(\sum_{q\in N^{HH}{}_h}e_q.V_h)\qquad(10)$$

其中e_k和e_q分别是症状k和草药q的初始嵌入，$N^{SS}{}_s$表示SS中s的邻居集，$N^{HH}{}_h$表示HH中h的邻居集。 Vs和Vh权重参数分别用于SS和HH。**通过我们的局部计算，节点度的平均值表明症状-草药图比协同图更密集，并且标准偏差证明协同图的度数分布比症状-草药图更平滑。**考虑到我们最近需要将b和r融合在一起，我们**在协同图上采用了求和聚合器**以使这两个部分更加平衡，这在一定程度上有利于训练过程。

从草药推荐任务的角度来看，SS和HH编码了中药的协同作用模式，这进一步改善了症状和草药的表示质量。此外，引入更多信息有助于在某种程度上**缓解中药处方的数据稀疏性问题[7]。**

C.信息融合

到目前为止，我们已经从二分GCN图和每个节点的协同图中获得了两种类型的嵌入。我们使用简单的加法运算来合并这些嵌入，

$$e_s^*=r_s+b_s$$

$$e_h^*=r_h+b_h\qquad(11)$$

其中$e^*{}_s$and $e^*{}_h$分别是症状s和药草h的合并嵌入。

综上所述，以上过程阐明了提出的多图嵌入层。它是一种通用体系结构，可以在其他情况下用于对两种类型的对象之间的交互进行建模。例如，在推荐方案中，可以利用Bipar-GCN来捕获用户和项目之间的内在差异。附加的用户-用户图可以是用户之间的社交关系图。项目-项目图可以是通过项目的内容属性链接的项目关系。

D.综合症的诱导

如上所述，综合症的诱导在中医临床实践中起着至关重要的作用。考虑到综合症归纳的含糊性和复杂性，在这项工作中，我们提出了一种基于MLP的方法来考虑综合症的隐式归纳过程，该过程可以描述症状之间的非线性相互作用并生成整体的隐式症状表现形式。

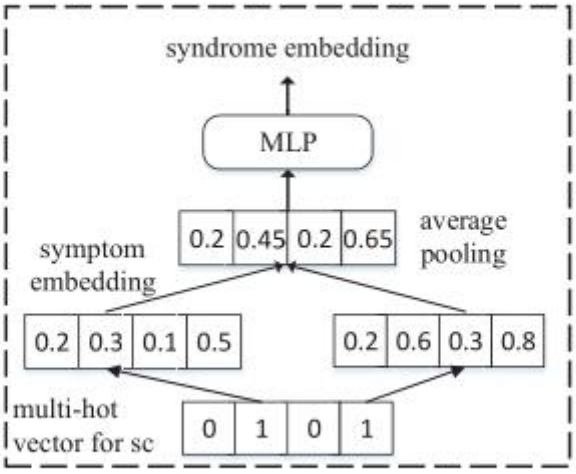


Fig. 4. The MLP-based method for syndrome induction.

如图4所示，我们将症状集中的所有症状嵌入都馈送到MLP中，以诱导总体隐式综合症表示形式。给定症状集sc，首先我们用多重热向量表示它。在此向量中，如果sc包含症状s，则相应的条目设置为1，否则设置为0。其次，我们在sc中查找每个症状s的嵌入e^*_s，并将这些向量堆叠以构建矩阵$e_{sc} \in R^{|sc|*d}$，其中d是单个症状嵌入的维数。第三，为了从e_{sc}得出总体表示，我们使用了平均值池化层。此外，考虑到综合症归纳的复杂性，我们应用单层MLP来变换均值向量，它借用了MLP中的非线性强度来学习更具表现力的综合症表示。以上计算过程如下：

$$e_{syndrome}(sc) = \mathrm{Re}LU(W^{mlp} \bullet Mean(e_{sc}) + b^{mlp}), \quad (12)$$

其中$e_{syndrome}(sc)$是针对sc的诱导综合症嵌入。

E.训练和推理

在草药推荐场景中，给定一个症状集，生成一个草药集来治疗这些症状。对于每一个处方，我们都需要评估推荐草药集和有事实依据的草药集之间的距离，这类似于多标签分类任务。如图5所示，**不同药物在处方中出现的频率是不平衡的。因此，我们需要解决标签不平衡的问题。**

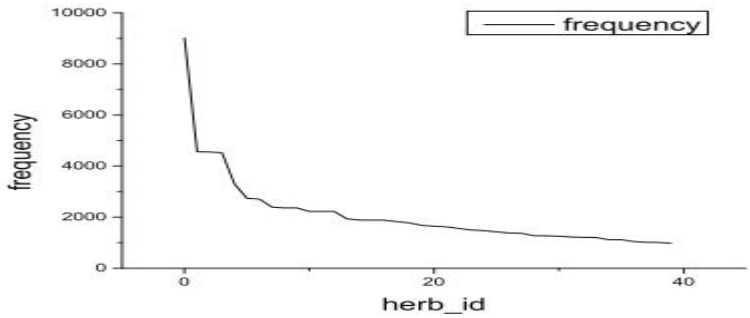


Fig. 5. Frequency distribution of the top 40 most frequent herbs.

这里，我们使用以下目标函数(13)来描述草药推荐场景中的上述特征，其中e_H是学习到的关于草药集合H的嵌入矩阵。

$$Loss = \arg\min_{\theta} \sum_{(sc,hc')\in P} WMSE(hc', g(sc,H)) + \lambda_\Theta ||\Theta||_2^2$$
$$g(sc,H) = e_{syndrome}(sc) \cdot e_H^T \quad (13)$$

在给定输入sc的情况下，有事实依据的药草集hc被表示为维度为|H|中的多热向量hc'。g(sc，H)是所有草药的输出概率向量。λ_θ控制L_2防止过拟合的正则化强度。

$$WMSE(hc', g(sc,H)) = \sum_{i=1}^{|H|} w_i (hc_i' - g(sc,H)_i)^2 \quad (14)$$

WMSE [10]是hc'和g(sc，H)之间的加权均方损失，定义如下:

hc'和g(sc，H)的维度都是|H|. hc'$_i$和g(sc，H)$_i$分别代表向量的第i个条目。Wi是草药i的权重，

$$w_i = \frac{max_k freq(k)}{freq(i)}$$

其中freq（i）是i在处方中出现的频率。自适应权重设置是为了平衡各种频率下草药的贡献。可以看到，我出现的次数越多，其重量就越低。我们**采用Adam[11]来优化预测模型并以小批量的方式更新模型参数。**

一些研究认为，不同标签之间存在某些模式，可以利用这些模式来提高多标签分类的性能。张等[12]引入一个正则化项以最大化属于一个集合的标签和不属于该集合的标签之间的概率间隔。但是，在我们的情况下，**成对的边距是不合理的**。详细的讨论在实验部分。

推论：遵循[10]中的设置，我们还采用贪婪策略来生成推荐的草药集。具体来说，我们选择g（sc，H）中概率最高的前k种草药作为sc的推荐草药。

Ⅴ实验：

在本节中，我们在基准中医数据集上评估我们提出的SMGCN [5]。有几个重要的问题要回答：

RQ1：我们提出的模型是否能胜过最先进的草药推荐方法？

RQ2：我们提出的模型是否能胜过基于最先进的图神经网络的推荐方法？

RQ3：我们建议的组件（BiparGCN，协同图编码（SGE）和综合症诱导（SI））的效果如何？

RQ4：我们的模型性能如何对不同的超参数设置（例如，隐藏层维度，GCN层深度和正则化强度）做出反应？

RQ5：我们提议的SMGCN是否可以提供合理的草药推荐？

我们首先介绍TCM数据集，基线，指标和实验设置。然后详细说明了实验结果。最后，我们将讨论几个关键超参数的影响。

A. 数据集

为了与工作[13]保持一致，我们对基准中医数据集[5]进行了实验。中医数据集包含98,334份原始医疗处方和33,765份经过处理的医疗处方（仅由症状和草药组成）。如图6所示，每个处方都包含几种症状和相应的草药。 Wang [13]等人在33,765例已处理的医疗案件中进一步选择26,360张处方。 26,360个医疗案例被分为22,917个培训案例和3,443个测试案例。表II中汇总了实验数据集的统计信息。

Prescription	Symptoms	Herbs
归脾汤 (Guipi Decoction)	心悸(palpitation), 健忘(amnesia), 失眠(insomnia), 舌淡苔薄(pale tongue), 盗汗(night sweat), 体倦(physical lassitude), 脉细弱(small and weak pulse)	人参(ginseng) 黄芪(astragalus mongholicus) 当归(angelica sinensis) 龙眼肉(longan aril) 甘草(glycyrrhiza uralensis) 茯苓(tuckahoe)

Fig. 6. The prescription example.

TABLE II
STATISTICS OF THE EVALUATION DATA SETS

Dataset	#prescriptions	#symptoms	#herbs
All	26,360	360	753
Train	22,917	360	753
Test	3,443	254	558

B. 评估

给定症状集，我们提出的模型产生了缓解症状的草药集。为了评估我们方法的性能，我们采用了推荐系统中常用的以下三种措施。对于测试数据集中的所有处方（sc，hc），它们的定义如下：

$$Precision@K = \frac{|Top(sc, K) \bigcap hc|}{K} \tag{16}$$

$$Recall@K = \frac{|Top(sc, K) \bigcap hc|}{|hc|} \tag{17}$$

$$NDCG@K = \frac{DCG@K}{IDCG@K} \tag{18}$$

其中，Top（sc，K）是在给定sc的情况下具有最高预测得分的前K个草药。 P recision @ K得分表示作为真正草药的前K个草药的命中率。在实验中，我们使用P recision @ 5来确定最佳参数。 Recall @ K描述了根据top-K建议得出的真实草药的覆盖范围。 N DCG @ K（归一化贴现累积增益）说明了命中草药在推荐列表中的位置。如果命中草药在列表中排名更高，则得分会更高。我们将这三种量度的排名均截断在20位，并报告测试集中所有处方的平均量度。

C. 基线

我们采用以下方法进行比较。

1）主题模型

•HC-KGETM [13]：它将从中医知识获图谱得的TransE [14]嵌入集成到主题模型中，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。

2）基于图神经网络的模型：

•GC-MC [15]：该模型利用GCN [16]获得用户和项目的表示。为了与原始工作保持一致，我们在实验中设置了一个图卷积层，并且隐藏层的维数与嵌入相同。

•PinSage [17]：PinSage是GraphSAGE [9]在项目-项目图上的工业应用。在我们的设置中，我们将其应用于症状-草药相互作用图。具体来说，我们在[17]之后采用两个图卷积层，并且隐藏层的维数与嵌入相同。

•NGCF [4]：NGCF是基于图的最新协作过滤方法。它显式构造了一个二部用户-项目图，以对高阶连通性进行建模，并为用户和项目获得更具表现力的表示形式。

3）我们提出的模型

•HeteGCN：这是我们提出的基线，它是基于基于异构图的GCN构建的[18]。我们将症状-草药图，草药-草药图和症状-症状图集成到一个异构图中。对于每个节点，有两种类型的邻居，即症状邻居和草药邻居。我们应用基于类型的注意力机制来执行消息构建。对于症状s，单跳邻居消息位于（19）和（20）中，其中tp = {症状，草药}表示邻居类型集，m在（1）中定义，并且||指示串联操作。 W^{att}和z是注意力网络参数。信息传播与（4）相同。请注意，症状和草药节点共享相同的网络参数。草药节点的公式相似。 HeteGCN采用平均池进行综合症诱导，与（13）相似，定义了多标签损失函数。 GCN的深度设置为1，隐藏尺寸为128，以获得更好的性能。

$$b_{N_s}^0 = tanh(\sum_{t \in tp} \alpha^t \frac{1}{|N_s^t|} \sum_{n \in N_s^t} m_n^0) \tag{19}$$

$$\alpha^t = \frac{exp(z^T ReLU(W^{att} \cdot (e_s || \frac{1}{|N_s^t|} \sum_{n \in N_s^t} m_n^0)))}{\sum_{t' \in tp} exp(z^T ReLU(W^{att} \cdot (e_s || \frac{1}{|N_s^{t'}|} \sum_{n \in N_s^{t'}} m_n^0)))} \tag{20}$$

TABLE III
OPTIMAL PARAMETERS OF COMPARATIVE MODELS

Approaches	Best parameter settings
HC-KGETM	α = 0.05 $\beta_s = \beta_h$ = 0.01 γ = 1
GC-MC	lr = 9e-4 dropout = 0.0 λ = 1e-6
PinSage	lr = 9e-4 dropout = 0.0 λ = 1e-3
NGCF	lr = 3e-3 dropout = 0.0 λ = 1e-5
HeteGCN	lr = 3e-3 dropout = 0.0 λ = 1e-3 x_s =5 x_h=40
SMGCN	lr = 2e-4 dropout = 0.0 λ = 7e-3 x_s =5 x_h=40

•SMGCN：提出的方法可以学习多个图（即症状-草药二分图，症状症状图和草药-草药图），并对它们执行图卷积以描述中医中的症状与草药之间的复杂关系。在预测层中，我们设计了一种基于MLP的方法来为每个症状集引入总体隐式综合症表示形式。因此，

这与现有的草药推荐方法相比有很大不同。

D. 参数设置

我们使用Tensorflow来实现我们的方法和比较方法。对于主题模型HC-KGETM，我们遵循[13]中的参数设置。我们采用网格搜索以搜索最佳学习率lr，正则化系数λΘ和丢弃率。具体而言，lr在{10^{-5},10^{-4},10^{-3}}中变化，λΘ在{0,10^{-6},10^{-5},10^{-4},10^{-3}}中调谐，丢弃率在{0,0.1，...，0.8}中查找。我们使用Xavier初始化程序[19]和Adam优化程序[11]来训练批量大小为1024的模型。

对于图神经网络基线，嵌入大小和潜在维数都设置为64。对于我们建议的SMGCN和HeteGCN，嵌入大小固定为64，并且第一个输出层的尺寸为128。在{64,128,256,512}中搜索最后一层的尺寸。 GCN层深度在{1,2,3}中进行调整。选项卡中汇总了最佳参数设置。III如果没有规范，我们接下来的SMGCN模型将是2个GCN层，其最后一层的尺寸为256。

E.性能比较

在这一部分中，我们首先展示了不同方法的总体结果，以及它们的最优参数设置。接下来，我们进行一些消融分析，以验证不同模型组件的有效性。然后我们详细讨论了超参数的影响。

1）总体结果:(RQ1&RQ2)表。

TABLE IV
THE OVERALL PERFORMANCE COMPARISON. HC-KGETM UTILIZES LOG-LOSS BUT WITHOUT SI. HETEGCN UTILIZES MULTI-LABEL LOSS BUT WITHOUT SI. THE OTHER MODELS ARE WITH SI AND ADOPT MULTI-LABEL LOSS. THE SECOND BEST RESULTS ARE UNDERLINED. P@K AND R@K ARE SHORT FOR PRECISION@K AND RECALL@K

Approaches	p@5	p@10	p@20	r@5	r@10	r@20	ndcg@5	ndcg@10	ndcg@20
HC-KGETM	0.2783	0.2197	0.1626	0.1959	0.3072	0.4523	0.3717	0.4491	0.5501
GC-MC	0.2788	0.2223	0.1647	0.1933	0.3100	0.4553	0.3765	0.4568	0.5610
PinSage	0.2841	0.2236	0.1650	0.1995	0.3135	0.4567	0.3841	0.4613	0.5647
NGCF	0.2787	0.2219	0.1634	0.1933	0.3085	0.4505	0.3790	0.4571	0.5599
HeteGCN	0.2864	0.2268	0.1676	0.2018	0.3192	0.4667	0.3837	0.4620	0.5665
SMGCN	**0.2928**	**0.2295**	**0.1683**	**0.2076**	**0.3245**	**0.4689**	**0.3923**	**0.4687**	**0.5716**
%Improv. by HC-KGETM	5.22%	4.44%	3.52%	5.95%	5.63%	3.67%	5.55%	4.36%	3.90%
%Improv. by PinSage	3.09%	2.61%	2.02%	4.02%	3.49%	2.68%	2.13%	1.60%	1.23%
%Improv. by HeteGCN	2.24%	1.17%	0.44%	2.87%	1.66%	0.46%	2.24%	1.45%	0.90%

Tab IV展示了整体的性能表现。需要注意的是，原始的基于图的神经网络基线并没有使用综合症归纳(SI)和多标签损失函数。为了公平比较，我们对GC-MC、PinSage和NGCF进行了修改，添加了SI部分并使用(13)中定义的多标签损失函数。我们可以看到:

•具体而言，我们提议的SMGCN在比较方法中表现最好。具体来说，SMGCN在p@5方面优于主题模型HC-KGETM 5.22%，r@5优于5.95%，ndcg@5优于5.55%。此外，对于最强的基线HeteGCN，SMGCN在p@5方面优于它2.24%，r@5优于它2.87%，ndcg@5优于它2.24%。对于第二好的基线PinSage，SMGCN在p@5方面超过了它3.09%，r@5超过了4.02%，ndcg@5超过了2.13%。

HC-KGETM几乎是所有指标中表现最差的。其原因可能包含两个方面:1)在交互建模阶段，只基于每个单一症状对候选药物进行排序，忽略症状集信息；2)在嵌入学习步骤中，采用TereE[14]来捕获中医知识图谱中的信息。与基于翻译的图嵌入方法相比，图神经网络在显式利用高阶连通度方面具有优势。

•在GC-MC，PinSage和NGCF中，NGCF表现最差，而PinSage表现最好。将GCMC与NGCF进行比较，GC-MC的性能略优于NGCF。考虑到GC-MC仅利用一阶邻居，因此NGCF的多个图卷积层可能会导致过拟合并损害性能。此外，PinSage，GC-MC和NGCF

具有各种传播功能：PinSage将目标节点和邻居节点的表示连接在一起，GC-MC将这两种表示相加，并且NGCF在构造消息时额外集成了目标节点和邻居节点的元素积部分。与按元素的乘积或求和操作相比，级联操作似乎在捕获处方中的丰富关系方面更为有效。

•HeteGCN的性能优于PinSage，这表明额外整合草药-草药和症状-症状并发关系可以引入更多信息。但是，SMGCN仍然优于HeteGCN，这证明与基于统一异构图的GCN相比，多图GCN框架可以在某种程度上学习更灵活和更具表达性的模型，并且MLP的选择适合描述复杂的综合症诱导过程。

2）消融分析：（RQ3）

为了更好地理解我们提出的SMGCN模型，我们将整个模型分为三个部分：Bipar-GCN，SGE和综合症诱导（SI），以分别评估它们对统一草药推荐系统的贡献。注意，在Bipar-GCN中，我们仅使用平均池对每种症状集进行综合症归纳。在SI部分，我们采用平均池化，然后进行MLP转换。Tab.V显示了消融分析的性能。此处，输出嵌入大小设置为256，图卷积层设置为2。在子模型中，Bipar-GCN和Bipar-GCN w / SI不包含协同图。因此，放弃使用基于异构图的HeteGCN，我们将更简单的基线PinSage与所有子模型进行比较。我们有以下观察结果：

•从整体上看，我们提出的模型的所有三个组成部分，即Bipar-GCN，SGE和SI，都被证实是有效的，因为它们在比较中展现更好的性能。

•比较Bipar-GCN与Bipar-GCN w / SI，发现MLP的选择优于仅使用平均池，这证明了MLP中的非线性变换有助于对症状之间的复杂关系建模，并进一步生成高质量的隐式综合症表示。

•对于具有SI的Bipar-GCN和Bipar-GCN，集成Synergy Graph Encoding（SGE）带来了进一步的改进，这表明嵌入学习层中的多个图的体系结构不仅有益于学习更具表达性的表示形式，而且也能协助草药的预测。

•SMGCN（Bipar-GCN，SGE和SI的组合）达到最佳性能，表明在草药推荐方案中，对综合症诱导过程中的非线性进行建模并通过多图统一复杂关系是有效的。

3）超参数的影响：（RQ4）

在这一部分中，我们将详细讨论关键因素。

•层数的影响

TABLE VI
EFFECT OF LAYER NUMBERS ON BIPAR-GCN W/ SI

depth	p@5	p@20	r@5	r@20	ndcg@5	ndcg@20
1	0.2898	0.1688	0.2044	**0.4702**	0.3864	0.5684
2	**0.2914**	**0.1690**	**0.2060**	0.4695	**0.3885**	**0.5699**
3	0.2882	0.1684	0.2030	0.4684	0.3869	0.5693

为了探究我们提出的模型是否可以从大量的嵌入传播层中受益，我们调整了Bipar-GCN w / SI子模型上GCN层的数目，其变化范围为{1,2,3} 。最后一层的尺寸设置为256。我们从TabVI中得到以下观察结果：

-我们提出的Bipar-GCN w / SI对传播层的深度不是很敏感。与一层的性能相比，两层模型的性能略高。

-将层数进一步增加到三层时，与一层相比，性能似乎有所下降。原因可能是由于较大的传播深度而导致过度拟合。

-当改变传播层的深度时，我们的BiparGCN w / SI始终优于最强的基线HeteGCN。它再次验证了SI部分的有效性，从经验上证明了MLP的非线性可以帮助描述复杂的综合症诱导过程。

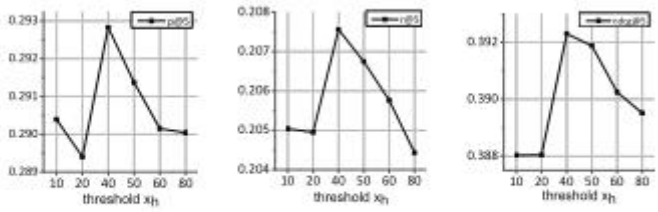
•最终嵌入尺寸的影响

TABLE VII
EFFECT OF LAST LAYER DIMENSIONS ON SMGCN

dimension	p@5	p@20	r@5	r@20	ndcg@5	ndcg@20
64	0.2857	0.1651	0.1999	0.4554	0.3847	0.5627
128	0.2882	0.1670	0.2018	0.4631	0.3853	0.5660
256	**0.2928**	**0.1683**	**0.2076**	**0.4689**	0.3923	**0.5716**
512	0.2922	0.1673	0.2068	0.4632	**0.3930**	0.5700

嵌入层的尺寸会极大地影响性能。我们对提出的SMGCN方法进行了实验，并且嵌入传播的深度设置为2。 TabVII显示了根据最后输出层的各种尺寸的实验结果。随着输出尺寸的增加，嵌入尺寸会不断改善，直到尺寸为256。当尺寸增加到512时，性能会略有下降，但仍优于第二强的基准PinSage。但是，当维度下降到64时，我们的模型在r @ 20和ndcg @ 20方面的表现不及PinSage。该观察表明，我们提出的模型依赖于合理的大型嵌入尺寸，以具有足够的灵活性来构造有用的嵌入。

•频率阈值在协同图中的影响

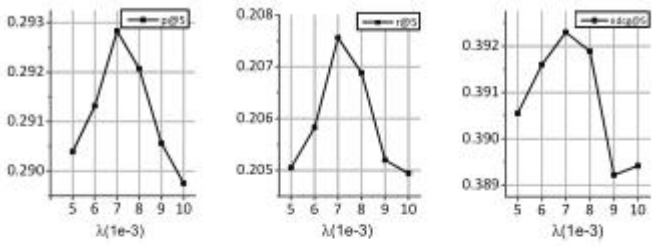


(a) precision@5 (b) recall@5 (c) ndcg@5

Fig. 7. Performance for different thresholds on SMGCN.

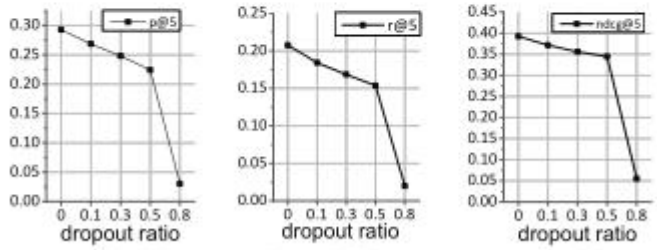
包含症状症状图和药草图的协同图编码（SGE）组件为我们提出的SMGCN做出了贡献。这两个图用于反映草药-草药对和症状-症状对之间的并发模式，这在中医理论中起着重要的作用。有两个控制协同图构建的超参数，草药-草药共存阈值x_h和症状-症状共存阈值x_s。例如，如果症状-症状对（s_m，s_n）在处方中出现的次数超过x_s次，则将边缘（s_m，s_n）添加到症状-症状图中。我们修复x_s为5，并在{10,20,40,50,60,80}中调整x_h。图7显示了不同阈值的实验结果。我们展示了topk = 5时的度量，并且在x_h = 40时可获得更好一些的性能。当阈值较低时，草药-草药图相对密集，但可能包含一些噪声。随着阈值的增加，图表变得稀疏，并且一些有用的信息可能会被过滤掉。因此，找到合适的阈值似乎会影响协同图的构建。

•正则化的影响



(a) precision@5 (b) recall@5 (c) ndcg@5

Fig. 8. Performance for different λ on SMGCN.



(a) precision@5 (b) recall@5 (c) ndcg@5

Fig. 9. Performance for different dropout ratios on SMGCN.

由于神经网络的强大表达能力，很容易过拟合训练数据。防止过度拟合的典型方法包括正则项和神经元脱落。在我们的设置中，λ控制参数的正则化强度，而丢弃率则控制训练过程中神经元去除的比例。图8展示了λ的影响，图9展示了丢弃率的影响，其中尺寸设置为256，深度设置为2。从图8中，我们观察到当λ等于7e-3时我们的模型获得了更好的性能。较大的λ可能会导致拟合不足并损害性能。较小的λ可能很弱，无法防止训练过程中的过度拟合趋势。至于丢弃技术，我们不是以一定的概率完全丢弃某些节点，而是仅对聚合的邻域嵌入应用消息丢弃，从而使我们的模型对存在或不存在单个边缘更加稳健。可以观察到，性能随着丢弃率的增加而下降，这表明上面的正则项足以防止过度拟合趋势。

•损失函数的影响

TABLE VIII
COMPARISON OF DIFFERENT LOSS FUNCTIONS

Approaches	p@5	p@20	r@5	r@20	ndcg@5	ndcg@20
NGCF w/ SI BPR	0.2760	0.1606	0.1953	0.4472	0.3825	0.5624
Bipar-GCN w/ SI BPR	0.2774	0.1623	0.1951	0.4479	0.3762	0.5565
NGCF w/ SI multi-label	0.2787	0.1634	0.1933	0.4505	0.3790	0.5599
Bipar-GCN w/ SI multi-label	**0.2914**	**0.1690**	**0.2060**	**0.4695**	**0.3885**	**0.5699**

在选项卡TabIV，我们通过添加SI分量并对其采用多标签损失，将基于GNN的基线（即GCMC，PinSage和NGCF）与我们建议的SMGCN对齐。因此，性能比较仅验证了我们模型中嵌入学习层的有效性。我们也对不同的嵌入层和预测层组合的有效性感到好奇。我们从基线中选择NGCF作为代表方法，比较损失函数是常用的成对BPR [20]。实

验结果总结在TabVIII中。至于BPR损失，Bipar-GCN w / SI表现更好。对于多标签损失，Bipar-GCN w / SI在所有指标中均优于。它验证了分别学习症状和草药表示可以帮助获得更具表达力的嵌入。此外，多标签丢失也胜过BPR丢失，这表明多标签丢失比BPR丢失更适合草药推荐任务。我们将详细说明原因。在中医处方中，草药集是根据草药相容性规则生成的，该规则在很大程度上取决于中医的个人经验。对于相同的症状集，可能有多种草药集可以作为治疗方法。因此，当草药A出现在处方中时，并不意味着A比每种缺失的草药B更合适。这仅表示由于某些草药相容性规则，草药B不适合加入当前的草药组。与BPR不同，多标签损失计算的是推荐药草集与有事实依据的药草集之间的距离，后者可从药草集视图评估结果。这也是我们不把正负标签边距约束加入（13）中损失函数的原因。

4）案例研究：（RQ5）

Symptom Set	Herb Set	
	SMGCN	Ground Truth
面色青(greenish complexion) 肠鸣(borborygmus) 恶心(nausea) 恶寒(aversion to cold) 四肢疼痛(pain in the limbs) 恶风(adversion to wind) 头痛(headache) 腹痛(abdominal pain)	甘草(glycyrrhiza uralensis) 人参(ginsen) 白术(largehead atractylodes) 茯苓(tuckahoe) 当归(chinese angelica) 木香(costusroot) 附子(aconitum carmichaelii) 半夏(ternate pinellia) 干姜(dried ginger) 川芎(szechwan lovage rhizome)	白术(largehead atractylodes) 甘草(glycyrrhiza uralensis) 藿香(agastache rugosa) 人参(ginsen) 草果(lanxangia tsaoko) 茯苓(tuckahoe) 附子(aconitum carmichaelii) 厚朴(magnolia officinalis) 半夏(ternate pinellia)
p@10=0.6 r@10=0.67 ndcg@10=0.88		
心悸(palpitation) 但热不寒(chill without fever) 舌干(dry tongue) 口苦(bitter taste) 小便黄赤(deep-colored urine)	甘草(glycyrrhiza uralensis) 人参(ginsen) 柴胡(bupleuri radix) 茯苓(tuckahoe) 白术(largehead atractylodes) 当归(chinese angelica) 藿香(agastache rugosa) 黄芩(scutellaria baicalensis) 白芍(paeonia lactiflora) 麦门冬(ophiopogon japonicus)	白术(largehead atractylodes) 甘草(glycyrrhiza uralensis) 青皮(citrus reticulata blanco peel) 人参(ginsen) 黄芩(scutellaria baicalensis) 草果(lanxangia tsaoko) 茯苓(tuckahoe) 厚朴(magnolia officinalis) 半夏(ternate pinellia) 柴胡(bupleuri radix)
p@10=0.6 r@10=0.6 ndcg@10=0.988		

Fig. 10. The herb recommendation cases.

在这一部分中，我们进行了案例研究，以验证我们提出的草药推荐方法的合理性。图10显示了草药推荐方案中的两个真实示例。给定症状集，我们建议的SMGCN会生成一种草药集，以治愈具有所列症状的综合症。在“草药集”列中，红色粗体表示SMGCN推荐的草药集和与有事实依据的草药之间的常用草药。根据草药知识，缺少的草药实际上与有事实依据的草药具有相似的功能，并且可以在临床实践中作为替代方法。通过以上比较分析，我们可以发现我们提出的SMGCN有能力提供合理的草药建议。

Ⅵ相关工作

A. 草药推荐

处方在中医临床经验和实践的传承中起着至关重要的作用。中医处方挖掘的发展历史包括三个阶段：1）传统的频率统计和数据挖掘技术，主要包括关联分析，简单聚类和分类方法。 2）主题模型。现有的研究[5]，[6]，[13]，[21] – [25]计算出并发的症状和草药词的条件概率，以捕获症状和草药之间的关系。 3）基于图模型的方法。研究[7]，[8]，[26]，[27]将中医处方组织成图表以捕获复杂的规律性。由于第一类方法仅适用于

单一疾病，因此我们主要关注第二和第三类方法。

基于主题模型的草药推荐：主题模型适用于以自然语言处理处方，其中中医处方是包含草药和症状（如单词）的文档。背后的动机是，同一主题下发生的草药和症状相似。Ma等[21]提出了一个“症状综合症”模型来挖掘症状和潜在综合症主题之间的相关性。 Ji等[6]认为“发病机理”是将症状和草药联系起来的潜在主题。 Lin等[22]通过主题模型共同对处方中的症状，草药，诊断和治疗进行建模。 Wang等[23]设计一个不对称概率生成模型，以同时对症状，草药和疾病进行建模。 Y ao等[5]将“综合症”，“治疗”和“草药作用”等中医概念整合到主题建模中，以更好地刻画处方的生成过程。 Chen等[24]和Wang等[13]将中医领域知识引入主题模型，以捕获草药相容性规律。

不幸的是，标准主题模型对短文本不是很友好。因此，处方的稀疏性[7]将在一定程度上限制主题模型在大规模处方上的性能。此外，他们无法全面分析各个实体之间的复杂相互关系。

基于图的草药推荐：图形是建模复杂关系数据的有效工具。基于图表示学习的草药推荐是当今研究的热点，其重点是获取中药实体的低维表示，然后基于嵌入来推荐草药。一些研究已经将深度学习技术引入了基于图的处方挖掘中。Li等[26]利用注意力Seq2Seq [28]设计一种多标签分类方法，以便自动生成处方。Li等[27]采用BRNN [29]对中医文献中的草药词进行文本表示学习，以完成补充治疗任务。 [7]，[8]将自动编码器模型与元路径集成，以挖掘TCM异构信息网络。

上述基于图的模型的弱点在于，所应用的深度学习技术最初是为欧几里德空间数据设计的，并且缺乏非欧几里德空间图数据的可解释性和推理能力。

B. 基于图神经网络的推荐系统

图神经网络（GNN）是图数据上神经网络的扩展，可以同时处理图的节点特征和边缘结构。由于其令人信服的性能和高解释性，GNN最近在推荐系统中得到了广泛应用。GNN应用于以下各种图：1）用户项交互图：Berg等[15]提出了一种基于自动编码器框架的图卷积矩阵完成模型。 Wang等[4]在基于GNN的嵌入过程中对协作信号进行编码，可以充分捕获协作滤波效果。 2）知识图：Wang等[30]提出了“涟漪网络”，该网络沿知识图中的边缘迭代扩展用户的潜在兴趣，以刺激用户偏好的传播Wang等[31]提出了一种知识图注意力网络，该网络递归地传播节点邻居的嵌入以获得节点嵌入，并采用注意力机制来区分邻居的重要性。 3）用户社交网络：Wu等[32]和范等[33]应用GCN来捕获社交网络中社交传播过程如何影响用户的偏好； 4）用户顺序行为图：Wu等[34]和Wang等[35]通过捕获用户行为序列中项目之间的复杂转换关系，将GNN应用于基于会话的推荐。

Ⅶ结论和未来的工作

在本文中，我们从新颖的角度考虑隐含的综合征归纳来研究草药推荐任务。我们开发了一系列GCN，以从症状-药草，症状-症状和药草-药草图中同时学习症状嵌入和草药嵌入。要了解整体隐式综合症的嵌入，我们将多个症状嵌入送入MLP，然后将其与草药嵌入集成以产生草药推荐。在公共中医数据集上进行的广泛实验证明了所提出模型的优越性，验证了模仿有经验的医生诱导综合征的有效性。在以后的工作中，为了进行嵌入学习，我们将通过采用注意力机制等先进技术来提高中药实体的嵌入质量。对于图的构建，我们将在TCM图中引入更多的TCM领域特定知识，包括草药的剂量和禁忌症。