

# 基于动态超图小波神经网络的半监督超图节点分类\*

庞俊<sup>1,2</sup>, 程俊澳<sup>1</sup>, 林晓丽<sup>1,2</sup>, 王蒙湘<sup>3†</sup>

(1. 武汉科技大学 计算机科学与技术学院, 武汉 430065; 2. 智能信息处理与实时工业系统湖北省重点实验室, 武汉 430065; 3. 中国标准化研究院, 北京 100088)

**摘要:** 半监督超图节点分类任务旨在利用少量节点的标签信息, 预测未标记节点的标签, 在复杂网络分析和应用中具有重要意义。现有半监督超图节点分类模型主要存在以下两点不足: a) 神经网络层数增加后, 忽视隐藏的高阶关系; b) 模型时间复杂度高。因此, 提出了一种新的动态超图小波神经网络(dynamic hypergraph wavelet neural network, DHGWNN), 解决半监督超图节点分类问题。一方面, 提出了一种基于 $k$ -NN、 $k$ -Hop和注意力机制的超图动态构建方法, 挖掘节点之间隐藏的高阶关系, 进一步丰富超图特征嵌入; 另一方面, 使用超图上的小波变换, 构建了一种超图小波卷积网络, 并重新推导了卷积更新公式, 以解决模型时间复杂度高的问题。四个引文网络真实数据集上的实验结果表明, 该方法的分类性能优于基线模型。

**关键词:** 超图; 半监督学习; 节点分类; 小波变换

**中图分类号:** TP391

**文献标志码:** A

**文章编号:** 1001-3695(2024)12-028-3735-07

doi: 10.19734/j.issn.1001-3695.2024.05.0129

## Semi-supervised hypergraph node classification based on dynamic hypergraph wavelet neural network

Pang Jun<sup>1,2</sup>, Cheng Jun'ao<sup>1</sup>, Lin Xiaoli<sup>1,2</sup>, Wang Mengxiang<sup>3†</sup>

(1. College of Computer Science & Technology, Wuhan University of Science & Technology, Wuhan 430065, China; 2. Hubei Province Key Laboratory of Intelligent Information Processing & Real-time Industrial System, Wuhan 430065, China; 3. China Institute of Standardization, Beijing 100088, China)

**Abstract:** The semi-supervised hypergraph node classification task aims to predict the labels of unlabeled nodes using the labeling information of a small number of nodes, which is important in complex network analysis and applications. Existing semi-supervised hypergraph node classification models have the following two main shortcomings: a) neglecting hidden higher-order relations with the number of neural network layers increasing. b) high model time complexity. Therefore, this paper proposed a new dynamic hypergraph wavelet neural network (DHGWNN) to solve the semi-supervised hypergraph node classification problem. On the one hand, it proposed a hypergraph dynamic construction method based on  $k$ -NN,  $k$ -Hop and attention mechanism to mine the hidden higher-order relationships between nodes, which further enriched the hypergraph feature embedding. On the other hand, it constructed a hypergraph wavelet convolutional network using wavelet transform on hypergraph and rederived the convolutional update formula to solve the problem of high time complexity of the model. Experimental results on four real datasets of citation networks show that the classification performance of the proposed method is better than the baseline models.

**Key words:** hypergraph; semi-supervised learning; node classification; wavelet transform

## 0 引言

图数据由节点和边组成, 可表示各种关系和网络结构, 如社交网络、知识图谱和分子结构等。半监督图节点分类任务旨在利用少量标记节点的标签信息, 预测未标记节点的标签, 使得模型能够将节点分配到正确的类别中。该任务具有广泛的应用领域, 例如社交媒体分析<sup>[1]</sup>、推荐系统<sup>[2,3]</sup>、生物信息学<sup>[4]</sup>、计算机视觉<sup>[5]</sup>等。

现有半监督节点分类方法多面向普通图(即简单二元关系图, 参见图1(b)), 主要采用图神经网络(graph neural networks, GNN)处理非欧几里德数据, 以实现节点分类。例如, Kipf等人<sup>[6]</sup>在谱理论的基础上结合深度学习, 提出了一种图卷

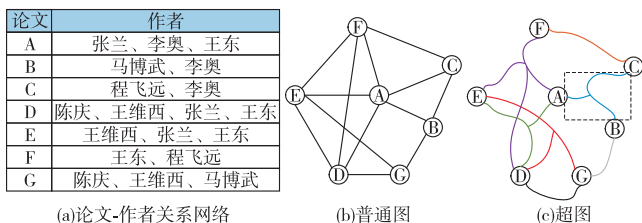
积模型GCN并将其应用于半监督节点分类任务。Velickovic等人<sup>[7]</sup>关注节点邻域特征, 为邻域中不同节点指定不同权重来更新节点特征, 并利用自注意层来解决图卷积只关注邻居节点而无法捕捉图结构中更远距离的节点关系的不足。此外, GraphSAGE<sup>[8]</sup>模型使用采样方法, GraphGAN<sup>[9]</sup>模型加入生成对抗网络。

这些方法针对普通图, 聚焦成对关系。然而现实世界中许多对象之间的关系并非简单的成对关系。例如, 多位作者合作完成一篇论文, 多位用户讨论同一话题等。虽然这些高阶关系可转换为成对关系进行表示, 但是存在信息丢失或信息冗余问题<sup>[10]</sup>。因此, 可用超图直接建模这些高阶关系<sup>[11]</sup>。超图是由节点和超边组成的图, 超边可以连接任意数量的节点。下文给

**收稿日期:** 2024-05-11; **修回日期:** 2024-06-27 **基金项目:** 新一代人工智能国家科技重大专项资助项目(2020AAA0108503); 国家自然科学基金资助项目(62372342); 武汉科技大学“十四五”湖北省优势特色学科(群)资助项目(2023D0301)

**作者简介:** 庞俊(1983—), 男, 湖北咸宁人, 副教授, 博导, 博士, CCF会员, 主要研究方向为图数据管理、图挖掘、大数据管理和分析; 程俊澳(1999—), 男, 湖北随州人, 硕士研究生, 主要研究方向为图挖掘; 林晓丽(1980—), 女, 河南信阳人, 副教授, 硕导, 博士, 主要研究方向为大数据发现、机器学习、计算生物学、生物信息学; 王蒙湘(1991—), 女(通信作者), 内蒙古赤峰人, 助理研究员, 博士, 主要研究方向为图数据挖掘、地理信息系统、城市计算、国际标准化(wangmx@cnis.ac.cn)。

出一个用普通图或超图表示引文网络的示例。



(a)论文-作者关系网络

(b)普通图

(c)超图

图1 普通图和超图的示例

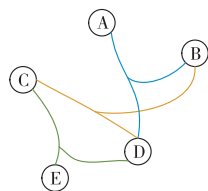
Fig. 1 Examples of graph and hypergraph

例1 如图1(a)所示,已知作者合作发表论文的相关信息(张兰、李奥、王东合作发表了论文A等),这些信息的普通成对图和超图表示分别如图1(b)(c)所示。节点表示论文,边表示所关联的论文有共同的作者。图1(c)虚线框起来的超边表示论文A、B、C有共同作者。显然从图1(b)很难判断多篇论文是否有共同作者(如只能判断论文A、B、C两两分别有共同作者,而不能判别这三篇论文有共同的作者),而从图1(c)却可清晰地辨识。

采用超图直接建模高阶关系后,半监督图节点分类转换为半监督超图节点分类。半监督超图节点分类具有广泛的应用。例如:a)图1中引文网络的例子,通过对论文分类可以识别出不同的研究领域和学科的论文,在实际的应用中这有助于了解各领域的发展状况和趋势;b)图2中社交网络的例子,通过对用户节点进行分类,可以更好地了解用户的兴趣和行为,从而提供个性化的内容推荐。已有不少学者研究了半监督超图节点分类方法,根据处理方式的不同可分为两类:基于谱分析的方法<sup>[10,12]</sup>和基于神经网络的方法<sup>[13~20]</sup>。基于谱分析的方法是一类以谱理论为基础的分析方法。例如:Agarwal等人<sup>[10]</sup>在谱理论的基础上提出星式展开和团式展开,将超图转换为普通图,从而将超图问题转换为普通图问题,进而采用普通图方法进行解决,该方法存在信息丢失、信息冗余的问题。Zhou等人<sup>[12]</sup>定义了超图上的拉普拉斯算子,首次提出超图学习,并将超图应用于半监督节点分类,但该方法有存在孤立节点时拉普拉斯算子失效的问题。该类方法通常有着严密的推导过程,同时相关的理论基础对其他类型算法的设计有着重要的指导意义。

| 用户 | 经常访问的内容  |
|----|----------|
| A  | 体育       |
| B  | 体育、经济    |
| C  | 经济、文娱    |
| D  | 体育、经济、文娱 |
| E  | 文娱       |

(a)用户-内容关系网络



(b)用户-内容关系网络对应超图

图2 社交网络示例

Fig. 2 Examples of social networks

基于神经网络的方法首先根据已知信息进行超图建模,接着将超图表示为关联矩阵,并采用卷积、注意力等技术来提取节点的特征表示,最后基于上一步获取的节点嵌入表示训练一个分类器来预测未分类节点的类别<sup>[13,14]</sup>。例如:Feng等人<sup>[13]</sup>受图卷积网络的启发,将Zhou等人<sup>[12]</sup>提出的拉普拉斯矩阵引入图卷积网络,从而设计超图卷积运算并将其应用于半监督节点分类任务。Wu等人<sup>[14]</sup>提出一种节点-超边网络的超图卷积来对节点和超边进行过滤,并使用交叉熵损失和超图重构损失进行反向传播。该类方法主要存在以下两点不足:a)当前大部分神经网络模型都采用静态超图结构,在神经网络层数增加的过程中,忽视了隐藏的高阶关系;b)当前普遍使用文献<sup>[13]</sup>中提出的超图卷积网络,该卷积网络由于正反傅里叶变换及大量矩阵分解操作,模型训练的时间复杂度过高。本文研

究第二类方法。

针对第一个问题,本文提出了一种针对半监督超图节点分类的超图动态构建方法,通过在每一层神经网络后添加超图动态构建模块,使用由超图小波卷积网络和注意力机制提取的节点特征,重新构建新的超图结构。该方法结合 $k$ -NN和 $k$ -Hop算法以及注意力机制带来的先验信息,迭代学习超图中节点之间更深层次的隐藏高阶关系信息。针对第二个问题,本文提出了一种超图上的小波变换,使用超图小波变换和超图拉普拉斯矩阵设计出了一种超图小波卷积网络,并推导出新的卷积更新公式,通过超图小波变换稀疏且易于计算的特点降低模型时间复杂度。基于上述工作,本文提出了一种新的动态超图小波神经网络解决半监督超图节点分类问题。与DHGNN<sup>[19]</sup>相比,本文使用 $k$ -NN和 $k$ -Hop算法构建超边集,通过等量融合方法形成超图。 $k$ -NN算法利用注意力权重选择相关性高的邻居节点,提供局部信息,减少噪声。 $k$ -Hop算法通过聚合目标节点 $k$ 跳内的所有节点,提供全局信息。与小波卷积网络相比,本文利用超图小波变换构建了超图小波卷积网络,推导了新的卷积公式,并引入注意力机制增强节点特征表达能力。

本文的主要贡献包括:

a)提出了一种新的超图动态构建方法,使用 $k$ -NN和 $k$ -Hop算法以及注意力机制带来的先验信息构建新的超图,作为下一层神经网络的输入,迭代学习深层隐藏高阶关系,进一步丰富节点嵌入特征。

b)提出了一种基于小波变换的超图小波卷积网络,使用超图小波变换重新推导了新的卷积更新公式,利用小波变换稀疏的特点以降低模型训练的时间复杂度;为进一步提升模型表达能力,采用注意力权重矩阵聚合节点特征。

c)在四个公开的引文网络数据集上进行实验,验证了本文方法的节点分类准确率优于所有baseline方法。

## 1 相关工作

### 1.1 面向普通图的半监督节点分类方法

最早期的简单图嵌入方法,如DeepWalk<sup>[21]</sup>、node2vec<sup>[22]</sup>等侧重于捕获节点之间的邻域邻近性。近些年,图神经网络因其强大的表示学习能力,在一些任务上取得了优异的成绩,例如节点分类任务、社交网络分析任务等。现有基于图神经网络的半监督节点分类方法可分为两类<sup>[23]</sup>:基于谱方法的半监督节点分类方法和基于空间方法的半监督节点分类方法。第一类方法通过图的傅里叶变换在谱域中定义卷积过程。由于该方法涉及拉普拉斯矩阵分解的操作,导致算法复杂度升高。为了降低计算复杂度,Defferrard等人<sup>[24]</sup>通过图拉普拉斯的切比雪夫展开来参数化卷积滤波器。为了进一步提高图卷积性能,Kipf等人<sup>[6]</sup>在卷积的每一层采用切比雪夫一阶多项式,在有效性和效率上取得良好的性能。第二类方法利用空间采样器和聚合器来获取邻居的特征表示。Hamilton等人<sup>[8]</sup>提出了GraphSAGE方法,通过采样选择相邻节点,然后将其馈送到聚合函数。注意力机制可以改进卷积方法,Velickovic等人<sup>[7]</sup>将注意力机制引入到神经网络对节点信息聚合的过程中,提出了图注意力网络GAT。陈东洋等人<sup>[15]</sup>提出了SC-HGANN模型,使用单纯形提取网络高阶结构,并结合注意力网络完成半监督节点分类任务。然而,这些方法是针对成对关系而设计,直接应用于非成对的高阶关系时,因未充分考虑多节点之间的组合关系,造成信息丢失;这些方法还会使图规模急剧增大,引发信息冗余问题。

## 1.2 面向超图的半监督节点分类方法

根据使用技术的不同,该方法可分为基于谱分析的方法和基于神经网络的方法两类。

第一类是谱分析的方法。Zhou 等人<sup>[12]</sup>将普通图上的矩阵分解方法推广到超图上提出了 N-cut 算法,进一步推动了超图上谱分析方法的发展,为后续的其他方法提供了一个良好的工具。Agarwal 等人<sup>[10]</sup>提出了星式展开和团式展开两种经典的展开方法,将超图分解为普通图,然后使用普通图的方法完成半监督节点分类任务,但展开存在信息丢失或者信息冗余的问题。然而,由于理论的限制,这类方法通常泛化性较差,且不够灵活,神经网络类方法能够很好地弥补谱分析方法的缺陷。

第二类是神经网络的方法。Feng 等人<sup>[13]</sup>提出了第一个超图神经网络 HGNN,将图卷积神经网络的思想自然地扩展到了超图上,设计了超图上的卷积操作,然而矩阵分解操作和傅里叶变换使得模型复杂度升高。同年,Yadati 等人<sup>[16]</sup>认为HGNN在信息融合过程中有太多噪声,使用采样展开的方式提出了 HyperGCN 并将其应用到超图半监督学习中。之后,陆续出现在超图神经网络的基础上进行神经网络模型改进的研究。Bai 等人<sup>[17]</sup>在 HGNN<sup>[13]</sup>中的卷积网络基础上加入了注意力机制,通过利用自注意力层进一步增强了表示学习的能力。吴越等人<sup>[18]</sup>将 HGNN<sup>[13]</sup>的应用范围扩充至异质网络,使用网络模体来构造超边,将异质网络转换为超图。该模型通过超图卷积网络和注意力机制聚合超图内部不同类型节点的信息,完成半监督超图节点分类任务。Jiang 等人<sup>[19]</sup>认为构造的初始超图可能不适合最终的数据表示,随后首次提出了一种动态超图神经网络 DHGNN 用于进行半监督超图节点分类等任务。Wu 等人<sup>[14]</sup>受 DHGNN<sup>[19]</sup>的启发,使用节点特征重构超图结构的方法,加强模型发掘信息的能力。同时,引入超图重建损失函数和交叉熵损失函数,加强反向传播的效果,从而提升半监督超图节点分类的性能。Gao 等人<sup>[20]</sup>在 HGNN<sup>[13]</sup>的工作基础上,将卷积进一步扩展到空间域中,提出 HGNN<sup>+</sup>。Li 等人<sup>[25]</sup>提出 HGTN 模型,利用不同类型的节点和超边之间的通信能力,学习高阶关系,并使用多尺度的注意力模块聚合节点嵌入。此外,Gu 等人<sup>[26]</sup>研究了超图的并行加速处理。

综上所述,现有神经网络类方法忽视了节点间隐藏的高阶关系,存在卷积网络时间复杂度高的问题。本文通过超图动态构建和超图小波卷积网络解决了这些问题。

## 2 问题形式化定义

**定义 1** 超图。它由节点和超边组成,记作  $\mathcal{G}=(\mathcal{V},\mathcal{E},\mathbf{W})$ ,其中, $\mathcal{V}=\{v_1,\cdots,v_n\}$ 表示超图中的  $n$  个节点集合, $\mathcal{E}=\{e_1,\cdots,e_m\}$ 表示超图中的  $m$  个超边集合。 $\mathbf{W}$  矩阵是超边权重的对角矩阵。超图  $\mathcal{G}$  可以用一个  $|\mathcal{V}| \times |\mathcal{E}|$  的关联矩阵  $\mathbf{H}$  来表示,其中,如果节点  $v_i$  与超边  $e_j$  相连接,那么  $H_{ij}=1$ 。节点  $v \in \mathcal{V}$  的度定义为  $d(v)=\sum_{e \in \mathcal{E}} w(e)h(v,e)$ ,其中, $w(e)$  表示的是超边  $e$  的权重。超边  $e \in \mathcal{E}$  的度定义为  $\delta(e)=\sum_{v \in \mathcal{V}} h(v,e)$ 。此外  $\mathbf{D}_v$  和  $\mathbf{D}_e$  分别表示节点度和超边度的对角矩阵。

**定义 2** 超图谱卷积。其依据超图拉普拉斯矩阵的定义,将 GCN 类比到超图上,定义了超图上的卷积运算  $g * x = \Phi(g(\Lambda)\Phi^T x)$ 。为了加速计算,使用了截断的切比雪夫多项式求得  $g * x$  的近似表达,最终超图谱卷积如式(1)所示。

$$\mathbf{X}_{i+1} = \sigma(\mathbf{D}_v^{-1/2} \mathbf{H} \mathbf{W} \mathbf{D}_e^{-1} \mathbf{H}^T \mathbf{D}_v^{-1/2} \mathbf{X}_i \Theta) \quad (1)$$

其中: $\mathbf{X}_{i+1}$ 、 $\mathbf{X}_i$  分别是卷积第  $i+1$  和  $i$  层的特征矩阵; $\Theta$  是在训练过程中需要学习的参数; $\sigma$  是非线性激活函数。

**定义 3** 半监督超图节点分类。假设给定一个超图  $\mathcal{G}=(\mathcal{V},\mathcal{E})$ ,其中  $\mathcal{V}$  表示节点的集合, $\mathcal{E}$  表示边的集合。对于每个节点,可以用特征向量  $\mathbf{X}_i$  表示节点的特征。半监督超图节点分类的目标是,通过少量已知的节点标签信息和超图中的高阶关系,学习一个函数  $f: \mathcal{V} \rightarrow \mathcal{C}$ ,将每个节点  $v$  映射到其所属的类别或标签。

## 3 DHGWN 模型

### 3.1 DHGWN 框架

DHGWN 模型的主要思想是:首先,通过节点特征构建超图,并获得超图关联矩阵;然后,将关联矩阵和初始节点特征传入超图小波卷积网络,得到卷积后的节点特征;接着,使用注意力权重对节点特征加权聚合,采用聚合后的节点特征和超图动态构建和重建超图结构,并重复上述步骤直至达到停止条件(设定的模型层数);最后将最终的节点特征送入分类器中训练,并对未分类节点进行类别预测。模型整体框架结构如图3所示。DHGWN 模型主要包括超图小波卷积网络、注意力加权聚合、超图动态构建三个模块。下文分别展开介绍这三个模块。

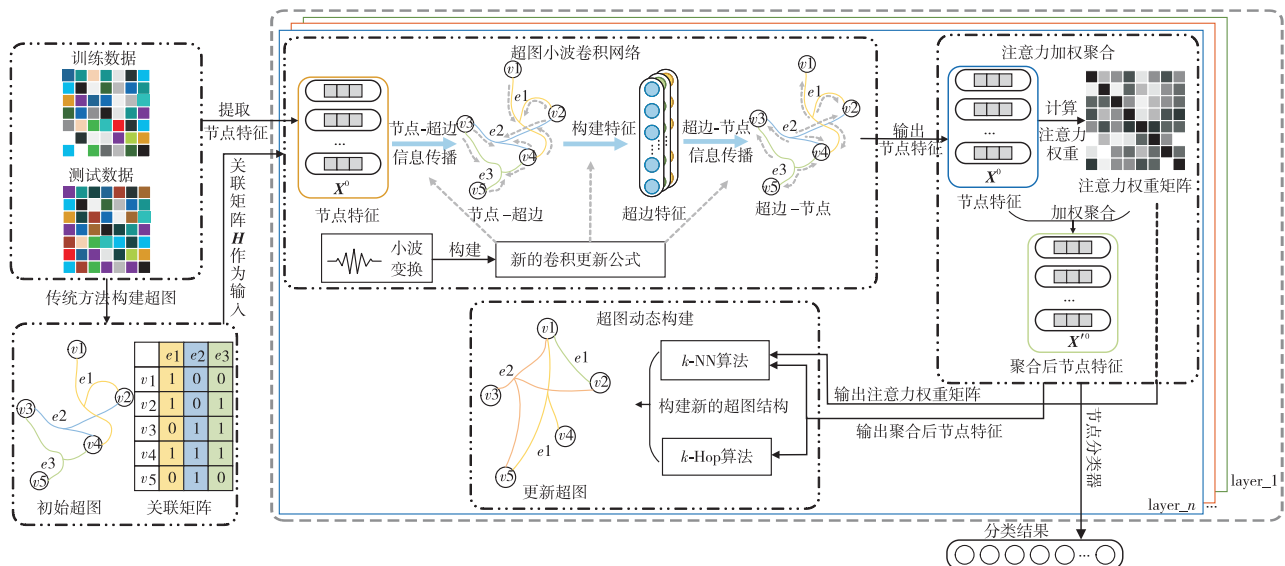


图3 总体框架  
Fig.3 Overall framework



### 3.2 超图小波卷积网络

本节提出一种新的超图小波卷积网络。该卷积网络与第 2 章中介绍的超图谱卷积不同,其使用超图上的小波变换重新构建了网络,并推导了新的卷积更新公式。尺度  $s > 0$  的超图小波基定义如式(2)所示。

$$\Psi_s = (\psi_{s,1} | \psi_{s,2} | \cdots | \psi_{s,N}) = \mathbf{U} \mathbf{G}_s \mathbf{U}^T \quad (2)$$

其中:  $\mathbf{G}_s = \text{diag}(g(s\lambda_1), g(s\lambda_2), \cdots, g(s\lambda_N))$  是在图拉普拉斯矩阵上定义的拉伸滤波器。通过将超图小波基应用到单个节点上,则节点  $v_i$  上尺度为  $s$  的超图小波基定义如式(3)所示。

$$\psi_{s,i} = \mathbf{U} \mathbf{G}_s \mathbf{U}^T \delta_i \quad (3)$$

其中:  $\delta_i$  是一个向量,其第  $i$  项为 1,其他项为 0。信号  $x$  的超图小波系数  $C_x(s, i)$  是由超图小波基与信号  $x$  的内积而产生的。计算过程如式(4)所示。

$$C_x(s, i) = \langle \psi_{s,i}, x \rangle = \psi_{s,i}^T x \quad (4)$$

然而,通过式(1)定义的小波变换需要计算超图拉普拉斯矩阵的完整特征值和特征向量,计算成本太高使处理速度大幅下降。为了解决该问题,使用切比雪夫多项式来近似超图小波系数。那么尺度  $s$  节点  $v_i$  上的超图小波系数可近似为

$$C_x(s, i) \approx \frac{1}{2} c_{s,0} x + \sum_{m=1}^M c_{s,m} T_m(\Delta) x \quad (5)$$

其中:  $T_m(\Delta)$  为切比雪夫多项式;  $c_{s,m}$  为切比雪夫系数。基于式(5)中定义的超图小波变换,尺度  $s$  上第  $l+1$  层卷积更新如式(6)所示。

$$\mathbf{X}^{l+1} = \sigma \left( \left( \frac{1}{2} c_{s,0} \mathbf{I} + \sum_{m=1}^M c_{s,m} T_m(\Delta) \right) \mathbf{X}^l \boldsymbol{\Theta} \right) \quad (6)$$

其中:  $\mathbf{X}^{l+1}$  和  $\mathbf{X}^l$  分别为卷积第  $l+1$  层和卷积第  $l$  层的节点特征矩阵;  $\boldsymbol{\Theta}$  是在训练过程中需要学习的参数矩阵;  $\mathbf{I}$  是单位矩阵;  $\sigma$  是非线性激活函数。为了方便计算机运算,将其改写为矩阵形式,同时使用切比雪夫多项式令  $M=1, \lambda_{\max}=1, \Delta = \mathbf{I} - \mathbf{D}_e^{-1/2} \mathbf{H} \mathbf{W} \mathbf{D}_e^{-1/2}$ , 将其代入式(6)中,得到矩阵形式的卷积更新公式,如式(7)所示。

$$\mathbf{X}^{l+1} = \sigma \left( \left( \frac{1}{2} c_{s,0} + c_{s,1} \right) \mathbf{I} - 2c_{s,1} \mathbf{D}_e^{-1/2} \mathbf{H} \mathbf{W} \mathbf{D}_e^{-1/2} \right) \mathbf{X}^l \boldsymbol{\Theta} \quad (7)$$

如图 3 中超图小波卷积网络所示,超图小波卷积层可以进行“节点-超边-节点”变换,完成节点到超边再到节点的信息传播过程,这样可以更好地细化节点特征。首先通过可学习的参数矩阵  $\boldsymbol{\Theta}$  对初始的节点特征进行处理,然后通过节点特征聚合到超边中形成超边特征,即通过乘以关联矩阵  $\mathbf{H}$  的转置来实现。最后通过聚合超边特征生成这层最终输出的节点特征,即通过乘以关联矩阵  $\mathbf{H}$  来实现。

### 3.3 注意力加权聚合

本节介绍了注意力权重的计算、注意力权重使用和节点特征融合的过程。因为与目标节点相距太远的节点提供信息有限,所以本文将节点注意力权重计算的范围限制在 2 跳内的所有邻居节点。

为了通过注意力机制获得足够的表达能力,将输入的节点特征转换为更高级别的特征,本文遵循 GAT<sup>[7]</sup> 中的设置。注意力权重的计算过程如式(8)所示。

$$e_{ij} = a(\mathbf{W} \mathbf{x}_i, \mathbf{W} \mathbf{x}_j) \quad (8)$$

其中:  $e_{ij}$  表示节点  $j$  的特征对节点  $i$  的重要性;  $\mathbf{W}$  为权重矩阵;  $a$  为一种共享注意力机制。本文设置只需要计算节点  $j \in \mathcal{N}_i$  的注意力权重  $e_{ij}$ , 其中:  $\mathcal{N}_i$  是节点  $i$  的某个邻域。为了使不同节点之间的注意力权重易于比较,使用 softmax 函数对节点  $i$  邻域内的所有节点  $j$  进行标准化,标准化过程如式(9)所示。

$$a_{ij} = \text{softmax}(e_{ij}) = \frac{\exp(e_{ij})}{\sum_{k \in \mathcal{N}_i} \exp(e_{ik})} \quad (9)$$

获得归一化的注意力权重后,将其用于计算对应特征的线性组合,作为每个节点的输出特征,计算过程如式(10)所示。

$$\mathbf{x}'_i = \sigma \left( \sum_{j \in \mathcal{N}_i} a_{ij} \mathbf{W} \mathbf{x}_j \right) \quad (10)$$

其中:  $\sigma$  为非线性激活函数。

通过对超图节点特征执行上述超图小波卷积网络和注意力机制后,获得了最终的节点嵌入表示  $\mathbf{X}'$ , 接着将其用于学习基于标记节点分类器。本文使用 softmax 函数作为激活函数,并且同时评估所有样本的交叉熵损失函数,计算公式如式(11)(12)所示。

$$\hat{\mathbf{X}}_i = \text{softmax}(\mathbf{x}_i + \mathbf{b}) \quad (11)$$

$$\mathcal{L}_{\text{cross-entropy}} = - \sum_{l \in \mathcal{L}_L} \sum_{k=1}^K C_{lk} \ln \hat{X}_{lk} \quad (12)$$

其中:  $\mathbf{x}_i$  为节点特征向量;  $\mathbf{b}$  为偏置项;  $\mathcal{L}_L$  为具有标签节点的索引集。

### 3.4 超图动态构建

本节介绍提出的超图动态构建方法。超图构建方法基于注意力加权聚合后的节点特征,重新获得超边,使得超图结构可以获得更多的隐藏高阶关系。受到 DHGNN<sup>[26]</sup> 超图动态构建的启发,本文使用两种不同的超边集构建策略来构建超边,还将注意力权重作为  $k$ -NN 方法的先验信息,最后通过等量融合方法构成超图。具体如算法 1 所示。

算法 1 超图动态构建

输入: 经过神经网络处理后的节点特征  $\mathbf{X}$ 、 $k$ -NN 方法中的  $k$  值  $k_1$ 、 $k$ -Hop 方法中的  $k$  值  $k_2$ 。

输出: 超图关联矩阵  $\mathbf{H}$ 。

函数:  $k$  近邻算法  $k$ -NN、 $k$  跳邻居算法  $k$ -Hop、拼接算法 concat、获取关联矩阵算法 GETH。

```

1 for  $u$  in range(len( $\mathbf{X}$ )) do //循环开始
2   Eknn =  $k$ -NN( $\mathbf{X}[u]$ ,  $\mathbf{X}$ ,  $k_1$ ) // 通过  $k$ -NN 函数选择目标节点  $x$  的  $k_1$  个邻居形成超边 Eknn */
3   KNNE.insert(Eknn) // 将超边 Eknn 插入到超边集 KNNE 中
4   Ekhop =  $k$ -Hop( $\mathbf{X}[u]$ ,  $\mathbf{X}$ ,  $k_2$ ) // 通过  $k$ -Hop 函数选择目标节点  $x$  的  $k_2$  跳内的所有节点形成超边 Ekhop */
5   KHOPE.insert(Ekhop) // 将超边 Ekhop 插入到超边集 KHOPE 中 */
6 end for //循环结束
7  $\mathbf{H}_1$  = GETH(KNNE) // 通过 GETH 函数获得超边集 KNNE 的关联矩阵  $\mathbf{H}_1$  */
8  $\mathbf{H}_2$  = GETH(KHOPE) // 通过 GETH 函数获得超边集 KHOPE 的关联矩阵  $\mathbf{H}_2$  */
9  $\mathbf{H}$  = concat( $\mathbf{H}_1$ ,  $\mathbf{H}_2$ ) // 通过 concat 函数拼接  $\mathbf{H}_1$  和  $\mathbf{H}_2$  得到超图的关联矩阵  $\mathbf{H}$  */
10 return  $\mathbf{H}$  //返回关联矩阵  $\mathbf{H}$ 

```

算法结束

算法 1 描述了超图动态构建流程。第 1~6 行循环执行直到遍历完所有节点。首先,通过  $k$ -NN 算法选择目标节点的  $k_1$  个邻居节点形成超边,并插入到超边集 KNNE 中(第 2,3 行)。然后,通过  $k$ -Hop 算法选择目标节点的  $k_2$  跳内的所有节点形成超边,并插入到超边集 KHOPE 中(第 4,5 行)。接着,根据两个超边集分别获得  $\mathbf{H}_1$  和  $\mathbf{H}_2$  关联矩阵(第 7,8 行)。最后通过 concat 操作获得最终的超图关联矩阵(第 9 行)。构建出的具有更多隐藏高阶关系的新超图结构可以作为下一层神经网络的输入。

接下来,给出一个根据上述两种不同策略构建超边的示例,如图 4 所示。

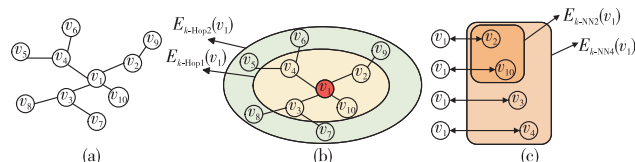


图 4 超图构建方法

Fig. 4 Hypergraph construction method

例 1 已知图 4(a) 所示的普通图,分别根据  $k$ -Hop 方法和  $k$ -NN 方法构建超边组。使用  $k$ -Hop 方法来生成超边组如图 4(b) 所示,图中选择  $v_1$  作为中心节点。当  $k$  值设置为 1 时,将节点  $v_1$  和其 1 跳范围内的所有节点连接起来形成超边  $E_{k\text{-Hop}1}(v_1)$  插入超边集,如图中内圈椭圆所示。当  $k$  值设置为 2 时,将外圈椭圆内所有节点连接起来形成超边  $E_{k\text{-Hop}2}(v_1)$  插入超边集。这类超边可以捕捉到节点之间更远距离的相关性。使用  $k$ -NN 算法生成超边组如图 4(c) 所示,图中选择  $v_1$  作为中心节点。当  $k$  值设置为 2 时,将中心节点及其 2 个邻居节点连接起来形成超边  $E_{k\text{-NN}2}(v_1)$  插入超边集。当  $k$  值设置为 4 时,将中心节点及其 4 个邻居节点连接起来形成超边  $E_{k\text{-NN}4}(v_1)$  插入超边集。由于中心节点每个邻居对中心节点的贡献度是不同的,所以采用注意力机制选择中心节点的  $k$  个邻居,以达到更好的效果。

### 3.5 训练过程

算法 2 描述了 DHGWNN 模型的训练过程。第 2~9 行不断迭代执行,迭代次数由模型层数  $Item$  决定,其处理过程如下:首先根据节点特征和关联矩阵等信息,使用卷积公式计算下一层的节点特征(第 3 行),然后根据节点特征计算注意力权重,并使用注意力权重对节点特征进行加权聚合得到新的节点特征(第 4、5 行),接着计算交叉熵损失进行反向传播(第 6、7 行),最后根据新的节点特征,使用超图动态构建模块得到新的超图继续迭代(第 8 行)。

#### 算法 2 DHGWNN 模型训练过程

输入:节点特征  $X$ 、模型层数  $Item$ 。

输出:节点特征  $X'$ 、节点类别  $C$ 。

根据普通图构建初始超图,获得超图关联矩阵  $H$

for  $i$  range( $Item$ ) do

由式(7)计算  $i+1$  层节点特征  $X^{i+1}$

由式(8)计算注意力权重矩阵  $E$

由式(11)计算融合注意力权重后的节点特征  $X'^{i+1}$

根据节点特征  $X'^{i+1}$  使用超图动态构建模块构建出新的超图结构并获得超图关联矩阵  $H'$

end for

根据神经网络最后一层聚合后的节点特征  $X'$  训练一个分类器 classifier

使用分类器对未分类节点预测得到节点类别  $C$

由式(13)计算交叉熵损失函数  $L_{c-e}$

进行反向传播

return  $X'$  and  $C$

算法结束

## 4 实验

本章在四个广泛使用的公开数据集上进行了实验,对比了本文 DHGWNN 模型和 baseline 模型的节点分类准确率、 $\text{macro-}F_1$ 、 $\text{micro-}F_1$  和模型运行时间。此外,还报道和分析了相关消融实验、参数分析和案例实验。下文分别介绍数据集、实验设置、基线模型、相关实验结果和分析。

### 4.1 数据集

本文实验采用四个用于节点分类任务的引文网络数据集: Cora<sup>[13]</sup>、Pumbed<sup>[13]</sup>、Citeseer<sup>[15]</sup> 和 DBLP<sup>[15]</sup>。

表 1 展示了这些数据集的统计信息,其中, nodes、edges、feature 和 class 分别表示数据集中节点数量、边的数量、特征数和节点类别数。

表 1 引文分类数据集的统计信息

Tab. 1 Statistical information of citation classification dataset

| dataset  | nodes  | edges  | feature | class |
|----------|--------|--------|---------|-------|
| Cora     | 2 708  | 5 429  | 1 433   | 7     |
| Pumbed   | 19 717 | 44 338 | 500     | 3     |
| Citeseer | 3 327  | 4 732  | 3 703   | 6     |
| DBLP     | 41 302 | 22 363 | 1 425   | 6     |

### 4.2 实验设置

本文实验使用的操作系统为 Ubuntu, CPU 为 IntelCore i5-10400F, GPU 为 NVIDIA GeForce RTX 3070, 在 PyCharm 平台上进行开发和调试。使用 PyTorch 框架实现 DHGWNN 模型。

对于所有数据集,从每个类别随机选择 10 个样本作为训练集,剩余样本 50% 作为测试集,另外 50% 作为验证集。例如, Cora 数据集中节点有 7 种类型,在每个类型的节点中随机选择 10 个节点作为训练集,剩余节点分别作为测试集和验证集。隐藏层的特征维数被设置为 16,并且使用 dropout drop 率  $p=0.5$  以避免过拟合。选择 ReLU 作为非线性激活函数。在训练过程中,使用 Adam 优化器使交叉熵损失函数最小化,学习率设置为  $l=0.001$ 。

### 4.3 基线模型和评价指标

本文采用目前流行的节点分类模型作为基线模型,包括:

a) 基于普通图的模型 GCN<sup>[6]</sup>、GraphSAGE<sup>[8]</sup>、GAT<sup>[7]</sup>、SC-HGANN<sup>[15]</sup>。

b) 基于超图的模型 HyperGCN<sup>[16]</sup>、Hyper-Atten<sup>[17]</sup>、HGNN<sup>[13]</sup>、HCNH<sup>[14]</sup>、HGTN<sup>[25]</sup>、HGNN+<sup>[20]</sup>。

评价指标选择节点分类常用的评价标准,即节点分类准确率(accuracy)<sup>[13]</sup>、 $\text{macro-}F_1$ <sup>[14]</sup> 和  $\text{micro-}F_1$ <sup>[14]</sup>。这三者值越大代表模型性能越好。

### 4.4 实验结果分析

本节报道并分析了本文提出模型与基线模型的节点分类准确率、 $\text{macro-}F_1$ 、 $\text{micro-}F_1$ 、模型训练时间。

#### 4.4.1 节点分类准确率

本节分别在 Cora、Pumbed、Citeseer 和 DBLP 四个数据集上对比本文 DHGWNN 模型与 Baseline 方法的节点分类准确率,实验结果如表 2 所示。其中,加粗的数据为每项评价指标的最优结果。

表 2 分类准确率结果对比

Tab. 2 Comparison of accuracy results of different models

| model       | Cora           | Pumbed         | Citeseer       | DBLP           |
|-------------|----------------|----------------|----------------|----------------|
| GCN         | 0.763 2        | 0.730 5        | 0.638 3        | 0.721 2        |
| GraphSAGE   | 0.732 5        | 0.723 0        | 0.593 2        | 0.698 5        |
| GAT         | 0.7670         | 0.737 8        | 0.647 2        | 0.702 8        |
| SC-HGANN    | 0.748 6        | 0.724 7        | 0.617 5        | 0.675 9        |
| HyperGCN    | 0.758 6        | 0.730 9        | 0.641 1        | 0.713 0        |
| Hyper-Atten | 0.768 4        | 0.733 6        | 0.639 8        | 0.758 2        |
| HGNN        | 0.771 8        | 0.728 7        | 0.639 9        | 0.776 5        |
| HCNH        | 0.760 1        | 0.714 5        | 0.622 5        | 0.762 3        |
| HGTN        | 0.770 1        | 0.731 4        | 0.668 4        | 0.782 1        |
| HGNN+       | 0.767 1        | 0.740 8        | 0.664 3        | 0.795 3        |
| DHGWNN      | <b>0.798 9</b> | <b>0.761 2</b> | <b>0.680 2</b> | <b>0.821 1</b> |

从表 2 的实验结果中可以看到, DHGWNN 模型在四个数据集上的节点分类准确率均优于其他基线模型。其中,相较于每个数据集中准确率最高的方法, DHGWNN 模型在四个数据集上的节点分类准确率分别提升了 2.71、2.04、1.59、2.58 百分点。一方面,本文 DHGWNN 模型使用了超图动态构建,能够更好地适应神经网络运行过程中节点特征的变化,构建出更好的超图结构;同时在使用  $k$ -NN 策略构建超边时,使用上一层模型中的注意力权重作为先验知识,提高了节点对邻居节点特征的聚合能力。另一方面,使用超图上的小波变换构建了新的超图小波卷积网络,使得卷积网络能够更有效地聚合节点的邻居信息,并将信息传播到更远的节点,从而获取更全局的超图结构信息。

综上所述,在公开数据集 Cora、Pumbed、Citeseer 和 DBLP 上的实验结果表明,相较于基线模型, DHGWNN 模型在节点分类准确性中取得了最优结果,验证了该模型的有效性。

#### 4.4.2 $F_1$ 得分与模型训练时间

为了更好地评估 DHGWNN 模型的性能,还与 HGNN 模型对比了  $F_1$  得分及模型的运行时间。

$F_1$  得分结果如表 3 所示,从表中可以看出,DHGWNN 模型在各个数据集上的 macro- $F_1$  和 micro- $F_1$  得分都要比 HGNN 模型高。因为在超图动态构建中加入的  $k$ -Hop 方法可以同时获得邻居的信息、邻居的邻居信息,使得超图不仅更全面地描述了节点之间的关系,还能够获得更多的邻居节点信息,与  $k$ -NN 方法获得的邻居信息互补,有利于发现节点的相似性。表 3 的实验结果也验证了 DHGWNN 模型捕获节点信息的能力更强,节点分类的效果更好。

表 3  $F_1$  得分结果对比

Tab. 3 Comparison of  $F_1$  score results

| metric       | model  | Cora    | Pumbed  | Citeseer | DBLP    |
|--------------|--------|---------|---------|----------|---------|
| macro- $F_1$ | HGNN   | 0.714 1 | 0.705 3 | 0.548 6  | 0.758 9 |
|              | DHGWNN | 0.721 6 | 0.714 5 | 0.605 8  | 0.796 8 |
| micro- $F_1$ | HGNN   | 0.717 7 | 0.711 2 | 0.587 1  | 0.765 4 |
|              | DHGWNN | 0.728 9 | 0.721 0 | 0.611 2  | 0.802 1 |

图 5 展示了 DHGWNN 和 DHGNN 模型分别在四个实验数据集上的训练时间。DHGNN 模型为将本文 DHGWNN 模型中的超图小波卷积网络替换为文献[13]中提出的超图卷积网络,其他模块与本文提出的模型保持一致。训练过程中两种模型的神经网络层数都设置为 2。从图中可知,DHGWNN 模型在四个数据集上的训练时间均少于 DHGNN 模型,验证了 DHGWNN 模型使用的超图小波卷积网络能够更快地完成卷积操作。通过使用超图上的小波变换构建的超图小波卷积网络,可通过快速算法获得,不需要进行正反傅里叶变换,节省时间。此外,小波变换是稀疏的计算,计算效率比傅里叶变换高。

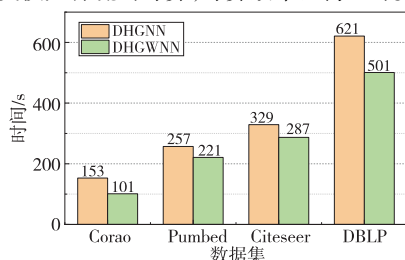


图 5 模型训练时间对比

Fig. 5 Comparison of model training time

综上,DHGWNN 模型的节点分类准确率、 $F_1$  得分、模型训练时间均优于其他基线模型。

#### 4.5 消融实验

本节将通过消融实验来验证 DHGWNN 超图动态构建模块和超图小波卷积网络的有效性,其结果如表 4 所示。w/o DHG 表示只使用超图小波卷积网络在数据集上进行分类,w/o HGWNN 表示去除超图小波卷积网络只使用动态构建模块进行节点分类实验。

表 4 消融实验结果

Tab. 4 Results of ablation experiment

| model     | Cora    | Pumbed  | Citeseer | DBLP    |
|-----------|---------|---------|----------|---------|
| w/o DHG   | 0.731 5 | 0.734 3 | 0.648 9  | 0.774 2 |
| w/o HGWNN | 0.752 6 | 0.740 9 | 0.662 5  | 0.799 1 |
| DHGWNN    | 0.798 9 | 0.761 2 | 0.680 2  | 0.821 1 |

从表 4 可见,DHGWNN 模型的半监督超图节点分类性能优于单独使用超图动态构建模块或超图小波卷积网络的模型。这验证了使用超图动态构建模块可以发现更多的隐藏高阶关系,且使用超图小波卷积网络能够对对象的局部特征进行更好地捕捉。将超图动态构建方法和超图小波卷积网络相结合,两者相辅相成、相互促进,可挖掘出更多隐藏高阶关系,提高了半

监督超图节点分类的准确率。

#### 4.6 参数分析

为了探究动态超图小波神经网络的层数对超图半监督节点分类效果的影响,在 Pumbed 和 Citeseer 数据集上进行实验,实验结果如图 6 所示。

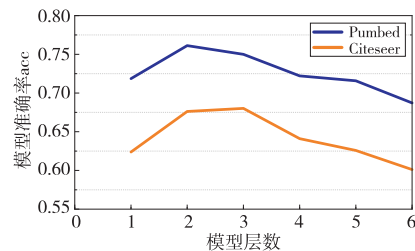


图 6 模型层数对分类结果的影响

Fig. 6 Effect of model layers on classification results

从图 6 可见,随着超图小波卷积网络层数的增加,本文模型的分类型准确率先增加后下降。因为当动态超图小波神经网络层数较少时,增加模型层数能够有效提取节点之间的高阶关系,从而提高节点嵌入质量,但持续增加模型层数,可能会出现梯度消失或者是过拟合的情况,使得模型难以得到优化和调整,从而导致分类准确率开始出现不同程度的下降。

#### 4.7 案例分析

本节进行案例实验与分析。从数据集 Cora 中选取多个经典案例进行实验评估,实验结果如表 5 所示。表 5 中第 1 列是各种测试案例;第 2、3 列中数字表示对于不同模型,在测试案例中模型预测论文正确类别的概率排名结果,其直观反映出模型预测论文类别的准确率。已知 11 148、35 922、30 901 号论文的 ground truth 类标签分别是 rule\_learning、probabilistic\_methods、theory。

表 5 示例案例及预测结果

Tab. 5 Example cases and predicted results

| 案例  | DHGWNN | HGNN <sup>+</sup> |
|---|--------|-------------------|
| (11148, belong to, ?, 25413, 193354, 1131719)→rule_learning           | 1      | 5                 |
| (35922, belong to, ?, 50336, 3229, 4335, 28957)→probabilistic_methods | 1      | 3                 |
| (30901, belong to, ?, 1272, 3243, 11339)→theory                       | 4      | 2                 |

以(11148, belong to, ?, 25413, 193354, 1131719)案例为例,其中编号为 11148 的论文 ground truth 类标签为 rule\_learning,问号代表待模型预测的论文类别,问号后的数字表示 11148 号论文引用的论文编号。该案例在 DHGWNN 模型中,正确论文类别“rule\_learning”的预测排名是第 1;而在 HGNN<sup>+</sup> 模型中,其排名为第 5。实验结果表明,对于未出现在训练样本中的测试论文,DHGWNN 模型节点分类的准确性高于其他基线模型,这再次验证了 DHGWNN 模型的有效性。

#### 5 结束语

本文提出了一种新的动态超图小波神经网络 DHGWNN,解决半监督超图节点分类问题。与目前最优方法相比,DHGWNN 模型提出了一种新的超图动态构建方法,根据神经网络处理后的节点特征动态调整超图结构,挖掘节点之间更多隐藏的高阶关系。为了降低模型训练的时间复杂度,本文提出了一种超图小波卷积网络,依靠小波变换相较于傅里叶变换稀疏易于计算的优势,可以减少卷积网络的时间复杂度,进而减少模型的训练时间。在四个公开数据集上的实验结果表明,本文提出的半监督节点分类模型优于其他基线模型。未来的工

作将研究在不影响模型性能的前提下,通过对注意力机制进行优化,进一步降低模型训练的时间复杂度和加快超图动态构建方法构建新超图的时间。

#### 参考文献:

- [1] Antonio P, Gianvito P, Domenico R, *et al.* SAIRUS: spatially-aware identification of risky users in social networks [J]. *Information Fusion*, 2023, 92: 435-449.
- [2] 吴相帅, 孙福振, 张文龙, 等. 基于图注意力的异构图社交推荐网络 [J]. *计算机应用研究*, 2023, 40(10): 3076-3081, 3106. (Wu Xiangshuai, Sun Fuzhen, Zhang Wenlong, *et al.* GAT based heterogeneous graph neural network for social recommendation [J]. *Application Research of Computers*, 2023, 40(10): 3076-3081, 3106.)
- [3] Su Yuhang, Gu Yu, Wang Zhigang, *et al.* Efficient subhypergraph matching based on hyperedge features [J]. *IEEE Trans on Knowledge and Data Engineering*, 2023, 35(6): 5808-5822.
- [4] Huang Yufei, Li Siyan, Wu Lirong, *et al.* Protein 3D graph structure learning for robust structure-based protein property prediction [C]// Proc of the 38th AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2024: 12662-12670.
- [5] Wei Zhou, Xia Zhiwu, Dou Peng, *et al.* Double attention based on graph attention network for image multi-label classification [J]. *ACM Trans on Multimedia Computing Communications and Applications*, 2023, 19(1): 1-28.
- [6] Kipf N T, Welling M. Semi-supervised classification with graph convolutional networks [EB/OL] (2016) [2017-02-22]. <https://arxiv.org/abs/1609.02907>.
- [7] Velickovic P, Cucurull G, Casanova A, *et al.* Graph attention networks [EB/OL] (2017) [2018-02-04]. <https://arxiv.org/abs/1710.10903>.
- [8] Hamilton W, Ying R, Leskovec J. Inductive representation learning on large graphs [C]// Proc of the 31st Neural Information Processing Systems. San Diego: NIPS Foundation, 2017: 1024-1034.
- [9] Wang Hangwei, Wang Jia, Wang Jialin, *et al.* GraphGAN: graph representation learning with generative adversarial nets [C]// Proc of the 32nd AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2018: 2508-2515.
- [10] Agarwal S, Branson K, Belongie S. Higher order learning with graphs [C]// Proc of the 23rd International Conference on Machine Learning. New York: ICML Press, 2006: 17-24.
- [11] Berge C. Hypergraphs: combinatorics of finite sets [M]. New York: Elsevier Science & Technology Books, 1984.
- [12] Zhou Dengyong, Huang Jiayuan, Schölkopf B. Learning with hypergraphs: clustering, classification, and embedding [C]// Proc of the 20th Neural Information Processing Systems. San Diego: NIPS Foundation, 2006: 1601-1608.
- [13] Feng Yifan, You Haoxuan, Zhang Zizhao, *et al.* Hypergraph neural networks [C]// Proc of the 33rd AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2019: 3558-3565.
- [14] Wu Hanrui, NG M K. Hypergraph convolution on nodes-hyperedges network for semi-supervised node classification [J]. *ACM Trans on Knowledge Discovery from Data*, 2022, 16(4): 1-19.
- [15] 陈东洋, 郭进利. 基于图注意力的高阶网络节点分类方法 [J]. *计算机应用研究*, 2023, 40(4): 1095-1100, 1136. (Chen Dongyang, Guo Jinli. Node classification method of higher order network based on graph attention [J]. *Application Research of Computers*, 2023, 40(4): 1095-1100, 1136.)
- [16] Yadati N, Nimishakavi M, Yadav P, *et al.* HyperGCN: a new method of training graph convolutional networks on hypergraphs [C]// Proc of the 33rd Neural Information Processing Systems. San Diego: NIPS Foundation, 2019: 1511-1522.
- [17] Bai Song, Zhang Feihu, Torr P H S. Hypergraph convolution and hypergraph attention [J]. *Pattern Recognition*, 2021, 110: 107637.
- [18] 吴越, 王英, 王鑫, 等. 基于超图卷积的异质网络半监督节点分类 [J]. *计算机学报*, 2021, 44(11): 2248-2260. (Wu Yue, Wang Ying, Wang Xin, *et al.* Motif-based hypergraph convolution network for semi-supervised node classification on heterogeneous graph [J]. *Chinese Journal of Computers*, 2021, 44(11): 2248-2260.)
- [19] Jiang Jianwen, Wei Yuxuan, Feng Yifan, *et al.* Dynamic hypergraph neural networks [C]// Proc of the 28th International Joint Conference on Artificial Intelligence. San Francisco: Morgan Kaufmann, 2019: 2635-2641.
- [20] Gao Yue, Feng Yifan, Ji Shuyi, *et al.* HGNN+: general hypergraph neural networks [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2023, 45(3): 3181-3199.
- [21] Perozzi B, Al-Rfou R, Skiena S. DeepWalk: online learning of social representations [C]// Proc of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2014: 701-710.
- [22] Grover A, Leskovec J. node2vec: scalable feature learning for networks [C]// Proc of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2016: 855-864.
- [23] 马帅, 刘建伟, 左信. 图神经网络综述 [J]. *计算机研究与发展*, 2022, 59(1): 47-80. (Ma Shuai, Liu Jianwei, Zuo Xin. Survey on graph neural network [J]. *Journal of Computer Research and Development*, 2022, 59(1): 47-80.)
- [24] Defferrard M, Bresson X, Vandergheynst P. Convolutional neural networks on graphs with fast localized spectral filtering [C]// Proc of the 30th Neural Information Processing Systems. San Diego: NIPS Foundation, 2016: 3844-3852.
- [25] Li Mengran, Zhang Yong, Li Xiaoyong, *et al.* Hypergraph transformer neural networks [J]. *ACM Trans on Knowledge Discovery from Data*, 2023, 17(5): 1-22.
- [26] Gu Yu, Yu Laiqiang, Song Zhen, *et al.* Distributed hypergraph processing using intersection graphs [J]. *IEEE Trans on Knowledge and Data Engineering*, 2022, 34(7): 3182-3195.