

基于标签传播增强的多通道图卷积网络

袁立宁, 冯文刚, 刘钊

引用本文

袁立宁, 冯文刚, 刘钊. 基于标签传播增强的多通道图卷积网络[J]. 计算机科学, 2024, 51(8): 304-312.

YUAN Lining, FENG Wengang, LIU Zhao. Multi-channel Graph Convolutional Networks Enhanced by Label Propagation Algorithm [J]. Computer Science, 2024, 51(8): 304-312.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于RoBERTa和加权图卷积网络的中文地质实体关系抽取](#)

Chinese Geological Entity Relation Extraction Based on RoBERTa and Weighted Graph Convolutional Networks

计算机科学, 2024, 51(8): 297-303. <https://doi.org/10.11896/jsjcx.230600231>

[基于深度强化学习的二进制代码模糊测试方法](#)

Fuzz Testing Method of Binary Code Based on Deep Reinforcement Learning

计算机科学, 2024, 51(6A): 230800078-7. <https://doi.org/10.11896/jsjcx.230800078>

[基于相似网络融合算法的癌症亚型预测](#)

Cancer Subtype Prediction Based on Similar Network Fusion Algorithm

计算机科学, 2024, 51(6A): 230500006-7. <https://doi.org/10.11896/jsjcx.230500006>

[基于语义扩充和HDGCN的虚假新闻联合检测技术](#)

Unified Fake News Detection Based on Semantic Expansion and HDGCN

计算机科学, 2024, 51(4): 299-306. <https://doi.org/10.11896/jsjcx.230700170>

[图神经网络节点分类任务基准测试及分析](#)

Benchmarking and Analysis for Graph Neural Network Node Classification Task

计算机科学, 2024, 51(4): 132-150. <https://doi.org/10.11896/jsjcx.230200084>

基于标签传播增强的多通道图卷积网络

袁立宁^{1,2} 冯文刚¹ 刘 钊³

1 中国人民公安大学国家安全学院 北京 100038

2 广西警察学院公安大数据现代产业学院 南宁 530028

3 中国人民公安大学研究生院 北京 100038

(yuanlining@gjxcxy.edu.cn)

摘要 多数图卷积网络(GCN)模型通过设计高效的信息传递和保留方式提升节点分类任务的实验表现,忽略了节点标签信息在拓扑空间和属性空间的传播。针对上述问题,提出了一种基于标签传播算法(LPA)增强的多通道图卷积模型 MGCN-LPA,同时增大同类节点在属性和拓扑空间的关系权重,改善节点间特征和标签信息的传播。首先,计算不同节点的属性相似度值,并采用 k 近邻算法生成属性关系图;然后,利用结合了GCN和LPA的图卷积层GCN-LPA提取属性图和属性关系图的潜在特征,生成拓扑节点表示和属性节点表示;最后,将拓扑和属性表示进行融合,并将生成的最终表示用于节点分类任务。在3个基准图数据集上进行实验,MGCN-LPA的实验表现能够匹配当前较为先进的基线模型,其在Cora和Citeseer数据集上的分类结果相比表现最优的基线模型提升了9.3%和12%。上述实验结果表明,MGCN-LPA能够增大同类节点间路径的权重,从而增强同类节点间的信息传递,提升节点分类任务的实验表现。此外,消融实验结果表明,与仅使用拓扑空间或者属性空间信息的变体相比,融合两类信息的MGCN-LPA能够充分提取和保留原始图中蕴含的潜在特征,提升模型的表征能力和泛化性。

关键词:图卷积网络;标签传播算法;属性图;属性关系图;节点分类

中图分类号 TP183

Multi-channel Graph Convolutional Networks Enhanced by Label Propagation Algorithm

YUAN Lining^{1,2}, FENG Wengang¹ and LIU Zhao³

1 School of National Security, People's Public Security University of China, Beijing 100038, China

2 School of Public Security Big Data Modern Industry, Guangxi Police College, Nanning 530028, China

3 Graduate School, People's Public Security University of China, Beijing 100038, China

Abstract Most graph convolutional networks(GCN) improve the experimental performance of node classification tasks by designing efficient methods for information propagation and preservation, while ignoring the propagation of node label information in the topological and attribute spaces. Aiming at the above problems, the paper proposes a multi-channel graph convolution model MGCN-LPA enhanced by the label propagation algorithm(LPA). The model enhances the propagation of node features and label information by increasing the weights of relationship between nodes of the same class in the attribute space and topology space. Firstly, it calculates the similarity values of different node attributes and generates an attribute relation graph using the k -nearest neighbor algorithm. Then, it combines the GCN and LPA in the graph convolution layer GCN-LPA to extract potential features from the attribute graph and attribute relation graph, generating topological node representations and attribute node representations. Finally, the method combines the topological and attribute representations and utilizes the final representation for node classification tasks. On three benchmark graph datasets, the experimental performance of MGCN-LPA can match the current state-of-the-art baseline models. The classification results on the Cora and Citeseer datasets show improvements of 9.3% and 12% respectively compared to the best-performing baseline. The experimental results demonstrate that MGCN-LPA can increase the weights of paths between nodes of the same class and enhance the propagation of information among nodes of the same class, thereby enhancing the performance of node classification tasks. In addition, the ablation experiments demonstrate that the fusion of both topological space and attribute space information in MGCN-LPA enhances the model's representational capacity and ge-

收到日期:2024-01-17 返修日期:2024-05-07

基金项目:国家重点研发计划(2023YFC3321604);广西哲学社会科学课题(23FTQ005);北京市社会科学基金(22GLB225);广西壮族自治区公安厅专项课题(2023GAQN092)。

This work was supported by the National Key Research and Development Program of China(2023YFC3321604), Social Science Fund of Guangxi(23FTQ005), Social Science Fund of Beijing(22GLB225) and Special Fund of Guangxi Public Security Department(2023GAQN092).

通信作者:冯文刚(fengwengang@ppsuc.edu.cn)

neralization compared to variants using only one type of information. This fusion allows for the full extraction and preservation of latent features present in the original graph.

Keywords Graph convolutional network, Label propagation algorithm, Attribute graph, Attribute relation graph, Node classification

1 引言

图是一种常见的数据形式,用于表示实体和实体间复杂的关系,例如引文网络中的论文和引用关系^[1]、社交网络中的账户和好友关系^[2]等。如何有效利用图中蕴含的属性和关系信息,提升节点分类^[3]、链接预测^[4]等下游图分析任务实验的表现是当下研究的热点方向。传统启发式算法^[5]依赖人工设计的特征提取信息,并且受限于数据的高维非线性,不仅具有较高的计算复杂度和内存需求,在不同数据集上的泛化能力也较弱。图卷积网络(Graph Convolutional Network, GCN)^[6]是一类高效的图深度学习算法,通过节点间信息传递和聚合,提取原始图的拓扑和属性信息,使高维、稀疏且非线性的图数据转换为低维、稠密的节点向量表示,提升下游预测任务的表现。

近年来,基于GCN的图数据分析和挖掘方法受到越来越多研究者的关注,衍生出大量变体来弥补GCN在编码图信息时存在的不足^[7]。当前,GCN的改进方式主要分为3类:1)改善节点间信息传递和聚合的方式;2)引入保留特定信息的损失函数;3)对输入GCN的数据进行增强。例如,GCN在编码过程中使用对称归一化的拓扑邻接矩阵,每个节点在聚合邻域信息时,平等地对待每一个邻居节点,未能有效保留关键节点中蕴含的特征信息,对此Velickovic等提出了引入注意力机制^[8]的图注意力网络(Graph Attention Network, GAT)^[9],为邻居节点分配不同的注意力权重,增强邻域中关键节点信息的传递和聚合。在节点分类任务中,GCN通常使用节点标签和交叉熵损失进行优化,限制了模型对原始图信息的表征能力,对此Li等利用拓扑信息中的社群结构,构建全局结构对比损失^[10],使GCN生成的节点表示能够保留充分的拓扑信息,提升节点分类任务中的实验表现。由于图的稀疏性,拓扑结构中通常存在噪声边以及边缺失问题,使得GCN未能充分传递和聚合邻域信息,对此Xue等提出了一种数据增强算法^[11],将结构相似的节点聚合为超节点,进而扩大节点的接收域,增强GCN对拓扑结构信息的获取能力。

多数GCN模型通过设计高效的信息传递或保留方式来提升模型的实验表现,忽略了节点标签信息在特征传播过程中的重要作用。相反,标签传播算法(Label Propagation Algorithm, LPA)^[12]能够直接借助图的拓扑结构向邻域传递标签信息。因此,Wang等提出了一种融合GCN和LPA的神经网络模型GCN-LPA^[13],利用节点标签增大同类节点在拓扑空间转移路径的权重,改善同类节点间特征信息的传播。然而,分析GCN-LPA在节点分类任务中的实验结果可知,标签传播部分需要已知大量的节点标签信息才能提升模型性能,在标签信息有限的情况下,GCN-LPA的实验表现甚至差于原始的GCN模型。

针对上述问题,本文提出了一种基于LPA的多通道图卷积模型MGCN-LPA,同时增大同类节点在拓扑空间和属性空间中转移路径的权重,改善同类节点间特征和标签信息的传播。首先,计算节点属性相似度值并采用 k 近邻算法^[14]生成属性关系图;然后利用GCN-LPA提取属性图(拓扑空间)和属性关系图(属性空间)的潜在特征,生成拓扑节点表示和属性节点表示;最后,将两种表示进行融合,生成用于节点分类任务的最终表示。此外,相比GCN-LPA,MGCN-LPA在训练过程中保持LPA部分使用的节点标签与模型,用于训练的节点标签相同。在3个基准图数据集上的节点分类实验中,MGCN-LPA能够匹当前较为先进的基线模型,证明了LPA的引入有助于拓扑空间和属性空间的信息传递和聚合,增强模型的表征能力和预测能力。

综上,本文的主要贡献如下:1)提出了一种基于多通道策略的图卷积模型MGCN-LPA,分别对拓扑空间和属性空间信息进行特征提取;2)在训练过程中,通过LPA增大同类节点在拓扑空间和属性空间的关系权重,来增强同类节点间信息的传递和聚合;3)在3个基准图数据集上的节点分类任务中,MGCN-LPA的实验表现始终优于当前较为先进的基线模型;4)相比GCN-LPA,MGCN-LPA能够同时关注标签与属性、拓扑之间的依赖关系,提升模型在不同数据集上的泛化性。

2 相关工作

2.1 图卷积网络

开发高效的特征提取算法,对图数据进行深入分析和挖掘,已成为当前研究的热门领域。Liu等对常见的GCN及其变体进行了较为全面的综述^[15],Zhang等则对节点分类任务中使用的GCN进行系统性分析^[16],从模型和数据两个视角分析现有工作中存在的问题。

近年来,许多新的GCN模型被提出,并应用于下游图分析任务。例如,简化图卷积(Simplified Graph Convolution, SGC)^[17]模型直接使用线性网络对属性矩阵和邻接矩阵进行编码,取消了层间激活函数,降低了模型的计算复杂度;高效图卷积(Efficient Graph Convolution, EGC)^[18]在训练过程中更加关注原始图的各向同性,缓解了GCN内存消耗和延迟问题,其表现也优于关注各向异性的GAT等模型;GATv2^[19]对GAT中使用的静态注意力进行修改,使邻域中节点的注意力排序是动态可变的,在信息传递过程中模型能够自动学习关注更高的节点特征;Guo等提出了一种基于子图增强的图卷积模型TDA-GCN^[20],对子图上节点的一阶邻域关系进行增加或删除操作,将图熵变化最小的子图和原始图分别作为GCN的输入,并通过注意力机制融合子图嵌入和原始图嵌入;Li等将各层GCN生成的特征表示作为对比视图,同时利用少量标记信息和大量无标记信息生成自监督

信息^[21],构建半监督对比损失^[22],增强模型的表征能力;Zhong等对拓扑结构中的互补信息进行深入挖掘,通过柔性损失优化邻接矩阵,增强邻接矩阵的拓扑表示能力,在此基础上利用自监督机制从未标记数据中生成可靠的伪标签,通过上述模块增强GCN在分类任务中的实验表现^[23];Huang等分析了GCN易受对抗性攻击的原因,并对图数据的中频信号进行了深入探索,阐明了对抗性攻击下不同频率信号的特性,提高了中通滤波器图卷积模型Mid-GCN的鲁棒性^[24]。

2.2 标签传播算法

LPA是一类经典的节点分类和局部社区结构发现算法^[25],其基本思想是相邻节点具有同质性,即未知节点的标签信息与邻居节点的标签信息有关,且邻居标签的影响程度由两者的相似度决定。因此,LPA的损失函数可以表示为:

$$\text{loss} = \frac{1}{2} \sum_{i,j \leq n, i \neq j} \gamma_{ij} (y_i - y_j)^2 \quad (1)$$

其中, γ_{ij} 表示节点相似度, n 为节点数。当 γ_{ij} 较大时,节点 v_i 和 v_j 的关系越紧密,标签信息更容易在节点间传递。但是,多数情况下直接使用图的一阶邻域(γ_{ij} 等于邻接矩阵元素 a_{ij})进行标签传播,忽略了节点在属性空间和拓扑空间的相似性。针对上述问题,Wang等将LPA作为正则化项学习边的权重^[13],同时在传播过程中保持已知节点的标签,并将其作为源节点向未标注节点传递标签信息。

3 模型与算法

本章首先对文中使用的符号进行总结(见表1),给出属性图和节点分类的定义;然后介绍LPA和GCN的计算过程,引出融合LPA和GCN的GCN-LPA基本架构,并基于 k 近邻算法和GCN-LPA,构建多通道图卷积模型MGCN-LPA;最后对算法的训练过程和时间复杂度进行分析。

表1 本文中使用的符号

Table 1 Notations used in this paper

符号	含义
G	属性图
G'	属性关系图
V	节点集合
E	边集合
A	邻接矩阵
\hat{A}	LPA优化后的邻接矩阵
A'	属性邻接矩阵
\tilde{A}	LPA优化后的属性邻接矩阵
X	节点属性矩阵
x_i	节点 v_i 的属性向量
S	属性相似度矩阵
Y	标签集合/节点表示
y_i	节点 v_i 的标签向量
Y^T	拓扑节点表示
Y^A	属性节点表示
Y^M	MGCN-LPA生成的节点表示
H	特征矩阵
h	特征向量
D	顶点度矩阵
W	权重矩阵
w	权重向量
n	节点数
F	节点属性数
c	节点标签(类别)数
l	LPA的迭代次数/神经网络层数
k	属性邻接矩阵中相似节点数

3.1 问题定义

属性图通常表示为 $G = (V, E, A, X, Y)$,其中 $V = \{v_1, v_2, \dots, v_n\}$ 表示 n 个节点的集合, $E = \{e_{ij}\}_{i,j=1}^n$ 表示边的集合, $A \in R^{n \times n}$ 表示节点间连接关系的邻接矩阵,当节点 v_i 和 v_j 之间存在边时 a_{ij} 等于边 e_{ij} 的权值(无权图中权值为1), $A \in R^{n \times F}$ 表示具有 F 个不同属性的节点属性矩阵,当节点 v_i 具有属性 j 时 x_{ij} 的值取1,并且 $X = \{x_i\}_{i=1}^n$, x_i 表示节点 v_i 的属性向量, $Y \in R^{n \times c}$ 表示节点标签集合,采用One-hot编码 c 种不同类别的标签,并且 $Y = \{y_i\}_{i=1}^n$, y_i 表示节点 v_i 的标签向量。

节点分类算法利用图的属性矩阵 X 、邻接矩阵 A 以及已知的节点标签 $\{y_1, y_2, \dots, y_m\}$,预测无标签节点的类别 $\{y_{m+1}, y_{m+2}, \dots, y_n\}$ 。因此,节点分类的关键在于找到有效的映射函数 f ,将每个节点 v_i 映射成特定类别的标签。由于多数分类模型在训练过程中同时利用有标签节点和无标签节点的属性及邻接关系,因此节点分类本质上是一种半监督学习任务。

3.2 标签传播与图卷积网络

在属性图中,LPA假设两个相连节点可能具有相同的类别标签,因此利用邻接矩阵 A 中的边信息对节点标签进行迭代传播。首先,初始标签可设置为 $Y^{(0)} = \{y_1^{(0)}, y_2^{(0)}, \dots, y_n^{(0)}\}$,其中节点标签已知的 m 个节点的标签向量(m 小于节点总数 n)对应类别编码为1,其余为0,标签未知的节点用全0向量表示。然后,通过LPA算法对 $Y^{(0)}$ 进行 l 次迭代,生成矩阵 $Y^{(l)}$,其中元素 $y_i^{(l)}$ 表示第 l 次迭代后节点 v_i 的标签概率分布。综上,LPA算法可表示为:

$$Y^{(l+1)} = D^{-1} A Y^{(l)} \quad (2)$$

其中, D 表示 A 的顶点度矩阵, $D^{-1}A$ 表示邻接矩阵中元素的归一化操作,即将节点的边信息转化为转移概率信息。在式(2)中,节点标签通过转移概率传递到邻居节点。此外,为了避免 m 个标签已知节点的区别信息被弱化和覆盖,每次迭代后将其设置为初始值,表达式为:

$$y_i^{(l+1)} = y_i^{(0)}, \forall i \in [1, m] \quad (3)$$

经过无穷次迭代,式(2)中所有节点的标签向量可表示为邻居节点标签的加权平均。

$$y_i^{(\infty)} = \frac{1}{d_{ii}} \sum_{j \in N(i)} a_{ij} y_j^{(\infty)} \quad (4)$$

其中, d_{ii} 为度矩阵 D 的值, $d_{ii} = \sum_j a_{ij}$, $N(i)$ 表示节点 v_i 的一阶邻域。

在属性图中,GCN将属性矩阵 X 和邻接矩阵 A 作为前馈神经网络的输入,对一阶邻域的信息进行传递和聚合,特征的层间传播表达式为:

$$H^{(l+1)} = \sigma(D^{-1/2} A D^{-1/2} H^{(l)} W^{(l)}) \quad (5)$$

其中, W 表示权重矩阵, $H^{(l+1)}$ 表示第 l 层神经网络生成的特征矩阵, $H^{(0)}$ 为 X , $\sigma(\cdot)$ 表示激活函数, $D^{-1/2} A D^{-1/2}$ 表示邻接矩阵的对称归一化操作。为了保持LPA和GCN在信息传递过程中的一致性,使用 $D^{-1}A$ 代替 $D^{-1/2} A D^{-1/2}$,此时GCN的传播表达式修改为^[13]:

$$H^{(l+1)} = \sigma(D^{-1} A H^{(l)} W^{(l)}) \quad (6)$$

通过式(6)能够得出GCN的节点级特征向量表示:

$$h_i^{(l+1)} = \sigma\left(\frac{1}{d_{ii}} \sum_{j \in N(i)} a_{ij} h_j^{(l)} w^{(l)}\right) \quad (7)$$

其中, $\mathbf{h}_i^{(l+1)}$ 表示节点 v_i 在第 l 层生成的特征向量, $\mathbf{w}^{(l)}$ 表示第 l 层的权重向量。对比式(4)和式(7), 假设激活函数 $\sigma(\cdot)$ 为恒等函数, \mathbf{W} 为单位矩阵且迭代无穷次时, GCN 的节点级表示可转化为邻居节点特征向量的加权平均:

$$\mathbf{h}_i^{(\infty)} = \frac{1}{d_{ij}} \sum_{j \in N(i)} a_{ij} \mathbf{h}_j^{(\infty)} \quad (8)$$

3.3 融合标签传播的图卷积网络

对于节点分类任务, 最优的节点表示应该是具有相同标签的节点彼此接近(类内相似性), 具有不同标签的节点彼此远离(类间差异性)。因此, 最直观的节点表示优化方式是在 GCN 训练过程中动态调整节点间边权重 a_{ij} 的值, 进而增大邻域中同类节点特征信息的影响, 但在实际运用过程中, 这种方式的计算复杂度极高, 难以有效应用。针对上述问题, Wang 等将类内节点的边权重优化问题转换为类内节点标签的影响力问题^[13], 即边权重 a_{ij} 如果能够最大化节点 v_i 被 LPA 正确分类的概率, 那么该权重同时最大化了类内标签对 v_i 的影响。基于上述理论, 通过最小化 LPA 预测标签与真实标签的交叉熵损失, 生成优化后的边权重矩阵 $\hat{\mathbf{A}}$:

$$\begin{aligned} \hat{\mathbf{A}} &= \arg \min_{\mathbf{A}} \text{Loss}_{\text{LPA}}(\mathbf{A}) \\ &= \arg \min_{\mathbf{A}} \left[-\frac{1}{m} \sum_{i \leq m} \sum_{c=1}^C \mathbf{y}_{ic} \log(\mathbf{y}_{ic}^{\text{LPA}}) \right] \end{aligned} \quad (9)$$

其中, 边权重 a_{ij} 所代表的邻接矩阵 \mathbf{A} 成为了 LPA 算法可学习的权重 $\hat{\mathbf{A}}$, 即通过最小化优化函数 Loss_{LPA} 改变 \mathbf{A} 中边权重数值 a_{ij} , 获得使 LPA 分类结果最好的权重矩阵 $\hat{\mathbf{A}}$ 。优化后的边权重矩阵 $\hat{\mathbf{A}}$ 与邻接矩阵 \mathbf{A} 中包含的拓扑结构信息相同, 即上述学习过程仅对边的权值进行修改, 不会对原始图的边集 E 进行增加或删除操作。结合式(4)和式(8)中基于邻域信息加权平均的计算方式, $\hat{\mathbf{A}}$ 不仅最大化了节点被 LPA 正确分类的概率, 其元素 \hat{a}_{ij} 也增大了一阶邻域中同类节点特征和标签对节点表示的影响。

为了在 GCN 优化过程中融入 LPA 算法, 使用优化后的边权重矩阵替换邻接矩阵 \mathbf{A} , 得到 GCN-LPA 的层间传播计算式为:

$$\mathbf{H}^{(l+1)} = \sigma(\hat{\mathbf{D}}^{-1} \hat{\mathbf{A}} \mathbf{H}^{(l)} \mathbf{W}^{(l)}) \quad (10)$$

在节点分类任务中, GCN 同样通过最小化交叉熵损失进行优化, 其表达式为:

$$\text{Loss}_{\text{GCN}} = -\frac{1}{m} \sum_{i \leq m} \sum_{c=1}^C \mathbf{y}_{ic} \log(\mathbf{y}_{ic}^{\text{GCN}}) \quad (11)$$

Wang 等将式(9)和式(11)两个优化步骤进行组合, 以端到端的方式训练 GCN-LPA 模型:

$$\text{Loss}_{\text{GCN-LPA}} = \text{Loss}_{\text{GCN}} + \beta \text{Loss}_{\text{LPA}} \quad (12)$$

其中, β 表示控制 LPA 损失比重的超参数。由于 GCN 难以同时学习 \mathbf{W} 和 $\hat{\mathbf{A}}$, 因此 Loss_{LPA} 作为正则化项, 帮助学习边权重 $\hat{\mathbf{A}}$ 。

本质上, GCN-LPA 的边权重就是通过最小化标记节点的硬标签(One-hot 向量, 真实类别值为 1, 其余为 0)和通过多跳邻居重建的软标签(类别的概率分布, 所有元素的和为 1)之间的差异, 增大同类节点间路径的权重, 从而促进同类节点间特征和标签信息的传播。图 1 给出了同类节点间路径权重增强的示例, GCN-LPA 能够识别潜在的同类节点的边并增大其权值。

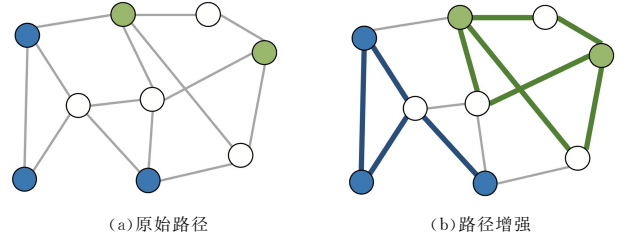


图 1 节点路径权重增强示例

Fig. 1 Example of node path weight enhancement

3.4 多通道图卷积网络

在节点分类任务中, GCN-LPA 在标签数量有限的情况下, 相比原始 GCN 性能提升并不明显。因此, 本文在 GCN-LPA 的基础上, 引入多通道策略, 构建多通道图卷积模型 MGCN-LPA, 分别从拓扑空间和属性空间提取原始图特征, 生成拓扑节点表示 \mathbf{Y}^T 和属性节点表示 \mathbf{Y}^A , 然后将 \mathbf{Y}^T 和 \mathbf{Y}^A 融合为最终表示 \mathbf{Y}^M 。模型总体结构如图 2 所示。

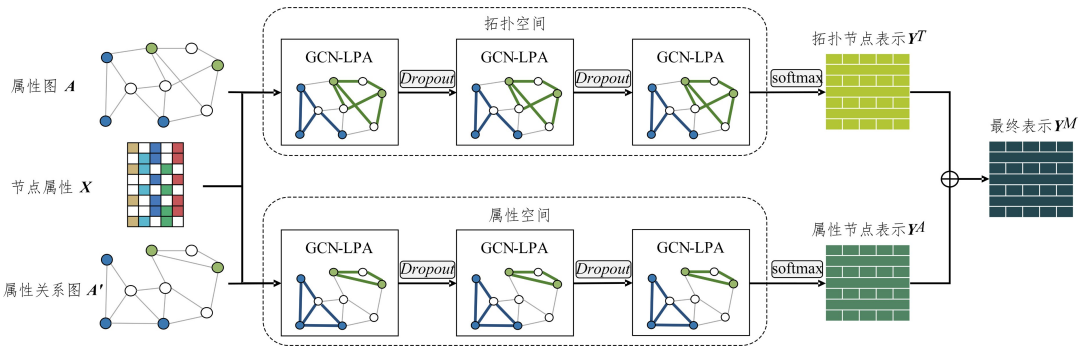


图 2 MGCN-LPA 模型的总体结构

Fig. 2 Overall framework of MGCN-LPA

在拓扑空间中, 直接将属性矩阵 \mathbf{X} 和邻接矩阵 $\hat{\mathbf{A}}$ (优化后拓扑边权重矩阵为 $\hat{\mathbf{A}}$) 作为 GCN-LPA 的输入, 因此拓扑表示 \mathbf{Y}^T 的计算过程可表示为:

$$\begin{aligned} \mathbf{H}_T^{(1)} &= \text{Dropout}(\sigma(\hat{\mathbf{D}}^{-1} \hat{\mathbf{A}} \mathbf{X} \mathbf{W}_T^{(0)})) \\ \mathbf{H}_T^{(2)} &= \text{Dropout}(\hat{\mathbf{D}}^{-1} \hat{\mathbf{A}} \mathbf{H}_T^{(1)} \mathbf{W}_T^{(1)}) \\ \mathbf{Y}^T &= \text{softmax}(\hat{\mathbf{D}}^{-1} \hat{\mathbf{A}} \mathbf{H}_T^{(2)} \mathbf{W}_T^{(2)}) \end{aligned} \quad (13)$$

其中, $\sigma(\cdot)$ 表示 ReLU 激活函数, Dropout 用于防止模型参数过拟合, softmax 表示将实向量转化为概率向量的激活函数。

在属性空间中, 节点之间不存在直接的关联信息, 因此学习属性表示 \mathbf{Y}^A 的前提是构造与邻接矩阵 \mathbf{A} 相似的拓扑关系, 并将其作为 GCN-LPA 的输入。针对上述问题, Wang 等提出了基于相似度和 kNN 算法的构造方式^[26], 生成无权无向的属性关系图 \mathbf{G}' 。本文在上述工作的基础上, 进一步保留不同节点在属性空间关联的差异性, 在生成属性关系图 \mathbf{G}' 的邻接矩阵 \mathbf{A}' 时对相似度值进行归一化。首先, 使用余弦相似性计算节点 v_i 和 v_j 的相似度值 s_{ij} :

$$s_{ij} = \frac{\mathbf{x}_i \cdot \mathbf{x}_j}{|\mathbf{x}_i| |\mathbf{x}_j|} \quad (14)$$

通过节点间相似度计算生成相似度矩阵 \mathbf{S} , 并根据 \mathbf{S} 中元素值 s_{ij} 的大小为每个节点选取 k 个最相似的节点, 生成属性邻接矩阵 \mathbf{A}' (除相似节点保留权值, 其余位置为 0), 此时 \mathbf{A}' 中元素 \mathbf{A}'_{ij} 能够直接表示节点 v_i 和 v_j 在属性空间的拓扑关系。参照式(13), 将属性矩阵 \mathbf{X} 和属性邻接矩阵 \mathbf{A}' (优化后属性边权重矩阵为 $\tilde{\mathbf{A}}$) 作为 GCN-LPA 的输入, 此时属性表示 \mathbf{Y}^A 的计算过程可表示为:

$$\begin{aligned} \mathbf{H}_A^{(1)} &= \text{Dropout}(\sigma(\tilde{\mathbf{D}}^{-1} \tilde{\mathbf{A}} \mathbf{X} \mathbf{W}_A^{(0)})) \\ \mathbf{H}_A^{(2)} &= \text{Dropout}(\tilde{\mathbf{D}}^{-1} \tilde{\mathbf{A}} \mathbf{H}_A^{(1)} \mathbf{W}_A^{(1)}) \end{aligned} \quad (15)$$

$$\mathbf{Y}^A = \text{softmax}(\tilde{\mathbf{D}}^{-1} \tilde{\mathbf{A}} \mathbf{H}_A^{(2)} \mathbf{W}_A^{(2)})$$

MGCN-LPA 通过计算拓扑表示 \mathbf{Y}^T 和属性表示 \mathbf{Y}^A 的平均值生成最终的节点表示 \mathbf{Y}^M :

$$\mathbf{Y}^M = \frac{\mathbf{Y}^T + \mathbf{Y}^A}{2} \quad (16)$$

在节点分类任务中, MGCN-LPA 同样使用交叉熵损失进行优化, 其表达式为:

$$\text{Loss}_{\text{CE}} = -\frac{1}{m} \sum_{i \leq m} \sum_{c=1}^C \mathbf{y}_k \log(\mathbf{y}_k^m) \quad (17)$$

此外, 在模型训练过程中, MGCN-LPA 需要同时优化拓扑边权重矩阵和属性边权重矩阵, 因此将式(12)中的 Loss_{LPA} 修改为:

$$\text{Loss}_{\text{LPA}}^{\text{MGCN}} = \frac{\text{Loss}_{\text{LPA}}^T + \text{Loss}_{\text{LPA}}^A}{2} \quad (18)$$

综上, MGCN-LPA 完整的损失函数为:

$$\text{Loss} = \text{Loss}_{\text{CE}} + \beta \text{Loss}_{\text{LPA}}^{\text{MGCN}} \quad (19)$$

3.5 算法与复杂度分析

MGCN-LPA 的训练过程如算法 1 所示。其中, 节点属性相似度计算、基于 k 近邻算法的属性关系图生成以及节点属性和邻接矩阵的归一化处理均为非参数计算过程, 因此在 MGCN-LPA 中各层神经网络编码过程的计算复杂度与 GCN 相同^[27], 均为 $O(|E|)$ 。此外, 损失函数中的 Loss_{CE} 和 Loss_{LPA} 均采用交叉熵损失, 因此 MGCN-LPA 优化过程中的计算复杂度为 $O(n)$ 。综上, 结合模型训练和优化过程, MGCN-LPA 的时间复杂度与节点数 n 、边数 $|E|$ 成正比。

算法 1 MGCN-LPA

输入: 属性矩阵 \mathbf{X} , 邻接矩阵 \mathbf{A} , 已知标签集合 \mathbf{Y} , 隐藏层维度 nhid1 和 nhid2, 属性近邻个数 k , 损失函数平衡系数 β

输出: 节点表示 \mathbf{Y}^M

1. 计算 \mathbf{X} 的余弦相似度, 生成属性相似度矩阵 \mathbf{S}
2. 每个节点按相似度大小选取 k 个邻居, 构建属性邻接矩阵 \mathbf{A}'
3. 分别初始化属性空间和拓扑空间的可训练参数矩阵 $\mathbf{W}_A^{(1)}$ 和 $\mathbf{W}_T^{(1)}$, $l=1, 2, 3$
4. Repeat
5. 使用 $\mathbf{A}, \mathbf{A}', \mathbf{Y}$ 和式(8)生成拓扑边权重矩阵 $\tilde{\mathbf{A}}$ 和属性边权重矩阵 $\tilde{\mathbf{A}}$
6. 使用 $\mathbf{W}_T^{(l)}, \mathbf{W}_A^{(l)}, \tilde{\mathbf{A}}, \tilde{\mathbf{A}}$, 式(12)和式(14)生成拓扑表示 \mathbf{Y}^T 和属性表示 \mathbf{Y}^A
7. 使用 $\mathbf{Y}^T, \mathbf{Y}^A$ 和式(15)计算节点表示 \mathbf{Y}^M
8. 计算损失函数梯度 ∇Loss , 利用反向传播更新参数矩阵 $\mathbf{W}_A^{(l)}$ 和 $\mathbf{W}_T^{(l)}$
9. Until convergence
10. 输出节点标签概率表示 \mathbf{Y}^M

4 实验与结果分析

4.1 实验设置

4.1.1 数据集

本文使用 Cora, Citeseer 和 Pubmed 这 3 个基准引文网络数据集^[28]来评估基线模型和 MGCN-LPA 在节点分类任务中的实验表现。数据中, 每个节点表示一篇论文, 边表示不同论文之间的引用关系, 节点属性表示描述论文内容的词袋 (Bag-of-Words) 表示, 节点标签表示论文的研究主题。为保证实验的公平性, 各模型采用相同的数据集划分, 每个类别抽取 20 个节点作为训练集, 在剩余节点中抽取 500 个节点作为验证集, 抽取 1000 个节点作为测试集。数据集统计信息如表 2 所列。

表 2 数据集统计信息

数据集	Cora	Citeseer	Pubmed
节点数	2708	3327	19717
边数	5429	4732	44338
节点属性	1433	3703	500
节点标签	7	6	3
训练集	140	120	60
验证集	500	500	500
测试集	1000	1000	1000

4.1.2 基线模型

本文使用当前较为先进和常用的节点分类算法作为基线模型, 主要分为以下 3 类:

1) 机器学习算法

逻辑回归 (Logistic Regression, LR): 用于解决多分类问题机器学习算法, 利用 softmax 函数将线性组合的特征转换为类别概率, 并将概率最高的类别作为分类任务的预测结果。

朴素贝叶斯 (Naive Bayes Model, NBM): 以贝叶斯定理为基础, 假设特征条件之间相互独立, 学习从输入到输出的联合概率分布, 使预测结果的后验概率最大。

支持向量机 (Support Vector Machines, SVM): 寻找最佳超平面, 使每个类别中距离超平面最近的样本到超平面的最小距离最大。

2) 多层感知机方法

MLP:由输入层、隐藏层和输出层组成,相邻层神经元之间采用“全连接”方式进行连接,能够更好地处理非线性特征,泛化性更强。

Graph-MLP^[29]:在 MLP 的基础上,引入保持局部结构的对比损失,使仅使用节点属性作为输入的 MLP 模型能够匹配 GCN 的分类性能。

3) 图神经网络方法

GCN^[27]:将卷积神经网络从欧氏数据推广到非欧氏数据图,利用节点的一阶邻域关系进行信息传递,捕捉节点间的依赖关系。

GCN-LPA^[13]:融合 GCN 和 LPA,利用节点标签增大同类节点在拓扑空间中路径的权重,改善节点间特征信息的传播。

SGC^[17]:直接使用线性网络对属性矩阵和邻接矩阵进行编码,在网络堆叠过程中取消层间激活函数,大幅减少模型参数,降低计算复杂度。

GAT^[9]:在 GCN 的基础上引入注意力机制,为邻居节点分配不同的注意力权重,增强邻域中关键节点信息的传递和聚合。

GAT-DSM^[28]:利用 GAT 生成各节点的基础表征,并计算节点表征的相似度,选出目标节点的相似节点集合,进而增强目标节点的嵌入表示。

DPDNN^[11]:利用数据增强算法,将结构相似的节点聚合为超节点,进而扩大节点接收域,增强 GCN 对拓扑结构信息的获取能力。

FMI^[30]:节点与节点之间存在直接相连的低阶结构和特殊连接模式形成的高阶结构,该模型引入节点间高阶结构关系信息学习节点表示。

SSC-HCL^[22]:将各层 GCN 生成的特征表示作为对比视图,同时利用少量的标记信息和大量的无标记信息生成自监督信息,并构建半监督对比损失。

4.1.3 参数设置

所有基线模型都先按照原始论文中建议的超参数进行初始化,并对部分模型进行调整。对于 MGCN-LPA,隐藏层维度 $nhid1$ 和 $nhid2$ 分别在 {512,768,1024} 和 {8,16,32,64} 中搜索,Cora 和 Citeseer 数据集的属性近邻个数 k 在 {20,30,40,50,60} 中搜索, Pubmed 的 k 在 {1000,2000,3000,4000,5000} 中搜索,损失函数平衡系数 β 在 {4,6,8,10,12} 中搜索,常用超参数学习率 lr 以及 $dropout$ 比率为固定值 0.05 和 0.5,训练过程中使用 Adam 优化器更新权重矩阵。最终, MGCN-LPA 超参数的最终设置如表 3 所列。

表 3 MGCN-LPA 参数设置

Table 3 Parameter settings of MGCN-LPA

数据集	$nhid1$	$nhid2$	k	β	lr	$dropout$
Cora	1024	32	40	10	0.05	0.5
Citeseer	1024	32	40	10	0.05	0.5
Pubmed	512	16	3000	6	0.05	0.5

4.2 实验结果

本节通过节点分类任务评估 MGCN-LPA 和基线模型的实验性能,并记录各模型的准确度值。实验结果如表 4 所列,其中粗体表示最优结果,下划线表示次优结果。

表 4 节点分类实验结果

Table 4 Results of node classification

		(%)		
类别	模型名称	Cora	Citeseer	Pubmed
ML	LR	56.8	55.8	73.3
	SVM	30.1	38.5	66.2
	NBM	52.2	51.7	70.4
MLP	MLP ^[29]	57.8	54.7	73.3
	Graph-MLP ^[29]	79.5	73.1	79.7
	GCN ^[27]	81.5	70.3	79.0
	SGC ^[17]	81.0	71.9	78.9
GNN	GAT ^[9]	83.0	72.5	79.0
	GAT-DSM ^[31]	81.8	70.6	78.7
	DPDNN ^[11]	<u>84.3</u>	73.3	<u>80.2</u>
	FMI ^[30]	79.5	70.7	78.4
	SSC-HCL ^[22]	83.7	<u>73.5</u>	80.1
GNN+LPA	GCN-LPA	83.0	68.8	78.1
	MGCN-LPA	93.6	86.6	80.7

从实验结果可以得出以下结论:

1) 在 3 个数据集上, MGCN-LPA 的实验表现均优于当前较为先进的基线模型,特别是 Cora 和 Citeseer 数据集上的分类结果相比表现最优的基线模型提升了 9.3% 和 12%。结果表明, MGCN-LPA 改善了节点间特征和标签信息的传播方式,提高了模型的预测性能。

2) 相比仅使用属性信息作为输入的机器学习方法和 MLP,同时使用属性和拓扑信息作为输入的模型预测效果更好,例如引入拓扑对比约束的 Graph-MLP 相比仅使用交叉熵损失的 MLP 提升显著。

3) 由于一阶邻域信息难以保留节点间复杂的关系,使得采用局部邻域聚合的模型难以充分获取潜在的特征信息,而采用扩大节点拓扑接收域方式的 DPDNN 以及融合属性空间和拓扑空间信息 MGCN-LPA,能够在信息传递过程中保留更丰富的特征信息,因此在不同数据集上的泛化性更强。

4) GCN-LPA 在 LPA 和 GCN 均使用相同节点标签的情况下,其结果相比原始的 GCN 表现不佳,而采用多通道方式融合拓扑和属性空间信息的 MGCN-LPA 始终优于 GCN 和 GCN-LPA 的实验结果,模型的表征能力更强。

5) 对比 Cora 和 Citeseer 数据集, Pubmed 的节点数和边数更多,而 MGCN-LPA 训练过程中使用的标记信息较少,同类节点间标签和特征信息传递的路径更长,使得基于 LPA 的路径权重优化效果减弱,因此限制了模型的性能。

4.3 消融实验

为了验证 LPA、拓扑空间信息和属性空间信息对 MGCN-LPA 实验性能的影响,在 3 个数据集上进行两组消融实验。其中,不使用 LPA 的 MGCN-LPA 变体为 MGCN,仅使用拓扑空间信息的变体为 TGCN-LPA(与 GCN-LPA 相同),仅使用属性空间信息的变体为 AGCN-LPA。为保持实验的公平性,各模型参数均按照表 3 进行设置。消融实验结果如图 3 和图 4 所示。

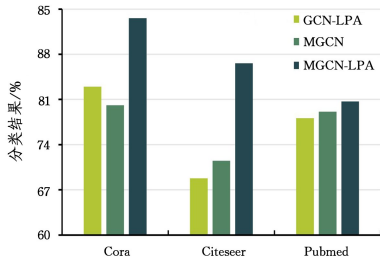


图3 消融实验(LPA和多通道)

Fig. 3 Ablation study(LPA and multi-channel)

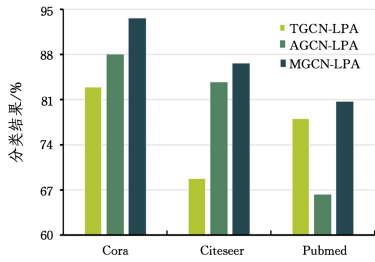


图4 消融实验(拓扑和属性)

Fig. 4 Ablation study(topology and attribute)

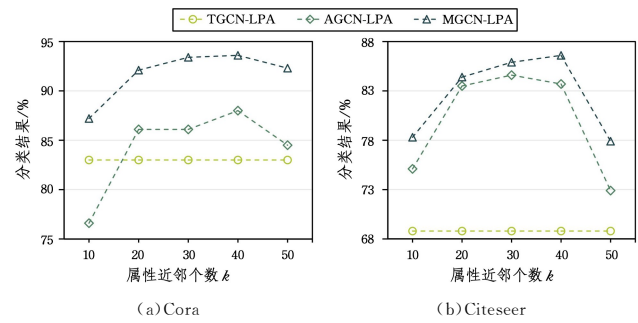
第一组消融实验探究了LPA和多通道策略对模型性能的影响。观察图3中的实验结果,GCN-LPA在Cora数据集上的实验表现优于MGCN,在Citeseer和Pubmed数据集上的表现弱于MGCN,而同时使用LPA和多通道策略的MGCN-LPA在3个数据集上均获得了最佳的实验表现,证明了在GCN-LPA的基础上引入多通道策略,不仅能改善节点间特征和标签信息的传播,还能使模型关注更多不同空间的特征信息,提升模型的预测性能。

第二组消融实验探究了拓扑空间和属性空间信息对模型性能的影响。观察图4中的实验结果,TGCN-LPA在Pubmed数据集上的实验表现优于AGCN-LPA,在Cora和

Citeseer数据集上的表现弱于AGCN-LPA,而同时使用拓扑空间和属性空间信息的MGCN-LPA在3个数据集上均获得了最佳的实验表现,证明了多通道策略能够使模型提取和保留更丰富的原始图信息,增强模型的代表能力。

4.4 参数实验

相比学习率、隐藏层维度和损失函数平衡系数等常见参数,属性近邻个数 k 是影响MGCN-LPA性能的重要参数。为了评估不同 k 值对模型性能的影响,在Cora和Citeseer数据集上进行参数实验(属性和拓扑空间变体作为参照),实验结果如图5所示。从结果看,使用AGCN-LPA和MGCN-LPA均呈现先上升再下降的趋势,使用过大或过小的 k 值实验表现均有限,这是因为 k 值过小时属性关系图无法保留充足的属性空间信息,而 k 值过大时属性关系图会生成大量噪声边从而导致模型性能下降。

图5 参数 k 的实验结果Fig. 5 Experimental results of parameter k

4.5 模型收敛实验

本节在Cora和Citeseer数据集上进行模型收敛实验,分析MGCN-LPA及其变体在不同训练次数下训练损失值、训练集分类结果和测试集分类结果的变化趋势。收敛实验结果如图6和图7所示。

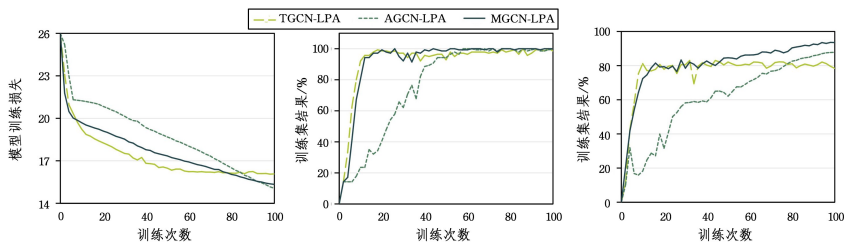


图6 模型收敛实验(Cora)

Fig. 6 Convergence analysis of models(Cora)

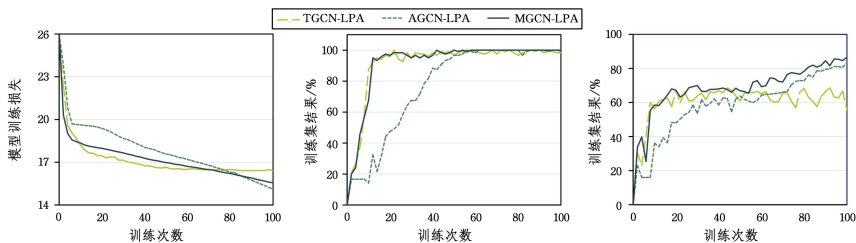


图7 模型收敛实验(Citeseer)

Fig. 7 Convergence analysis of models(Citeseer)

观察图6(a)和图7(a)的实验结果,拓扑空间变体TGCN-LPA的训练损失值最先趋于平稳,属性空间变体

AGCN-LPA的训练损失值先快速下降再保持一定斜率线性下降,同时使用拓扑和属性空间信息的MGCN-LPA在训练

初期损失值下降最快,但随着训练次数的进一步增加,MGCN-LPA的训练损失值受AGCN-LPA的影响增大,因此也呈现固定斜率下降的趋势。观察图6(b)和图7(b)的实验结果可知,进行10次训练后TGCN-LPA和MGCN-LPA分类结果即可快速收敛,而AGCN-LPA在50次训练后才逐渐收敛。观察图6(c)和图7(c)的实验结果可知,MGCN-LPA和TGCN-LPA在初始阶段的测试集结果显著优于AGCN-LPA并快速上升,但随着训练次数的增加,TGCN-LPA的表现趋于稳定,而MGCN-LPA测试结果随着AGCN-LPA进一步增长。

综合分析图6和图7,引入多通道策略的MGCN-LPA收敛表现同时受拓扑空间和属性空间通道的影响。从训练损失值看,MGCN-LPA的训练损失大多介于TGCN-LPA和AGCN-LPA之间。从测试集的预测结果看,MGCN-LPA的测试性能会先随着拓扑变体快速提升,然后随着属性变体的提高进一步增长。整体上看,多通道策略能够充分发挥拓扑空间或属性空间的优势信息,相比使用单一特征信息的变体,MGCN-LPA的表征能力更强,预测性能更好。

4.6 可视化

节点表示中蕴含了节点属性和拓扑结构信息,可视化后能直观反映原始图的某些特征以及模型对原始图的表征能力。首先使用t-SNE^[32]将各模型在Cora数据集上的节点表示降至2维,然后根据节点标签将二维平面上的数据点标记为7种不同的颜色,实验结果如图8所示。

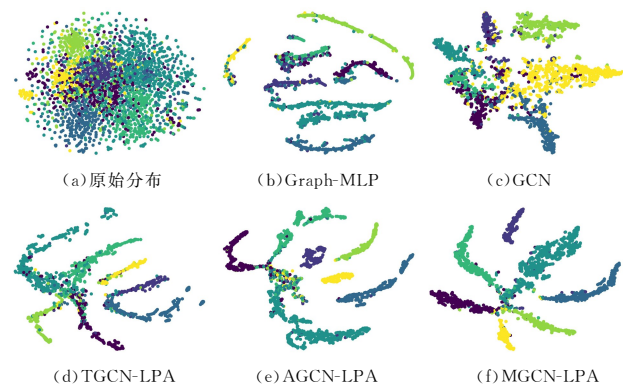


图8 节点表示可视化(电子版为彩图)

Fig. 8 Visualization of node representations

观察图8可知,所有模型均能从任意分布的原始数据中提取特征信息形成社区结构,但是各模型形成的社区数量、节点的类内相似性和类间界限差异较大。其中,GCN的类内节点分布松散、类间差异不够明显,而Graph-MLP和MGCN-LPA及其变体的类内节点分布更加接近。但是,Graph-MLP将7类节点分成了12个社区,而MGCN-LPA能够将7类节点分为7个社区,同时相比拓扑空间和属性空间的变体,MGCN-LPA多色节点重合区域更小。可视化实验直观地反映了MGCN-LPA保留同一社区节点相似特征的能力,证明了LPA和多通道策略的引入增强了模型对原始图信息的表征能力,使同类节点表示在向量空间中更加接近。

4.7 训练时间

为了比较模型的训练时间,记录Cora和Citeseer数据集上迭代100次后单次迭代的平均时间(包括前向传播、损失

函数计算、反向传播过程),实验结果如图9所示。相比仅使用交叉熵损失进行优化的GCN,GCN-LPA在训练时增加了调整邻域关系的LPA损失,使得训练时间略微增加,而MGCN-LPA相比GCN-LPA需要同时优化拓扑和属性空间的LPA损失,因此训练时间大约为GCN-LPA的2倍。

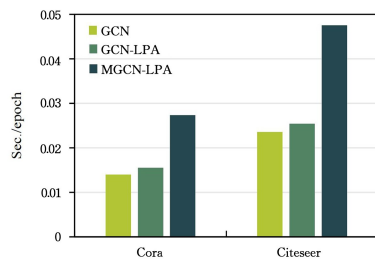


图9 训练时间对比

Fig. 9 Training time comparison

4.8 实验分析

Wang等将传统社区发现算法LPA与GCN相结合,构建了GCN-LPA模型。从实验结果看,GCN-LPA在标签数量有限的情况下,对模型性能的提升并不明显,并且对不同数据集的泛化性同样受限。这是因为GCN-LPA中LPA通过拓扑结构向邻居节点传播标签特征,忽略了标签信息与节点属性之间的关联。然而,现实世界中图数据的标签与属性、拓扑之间的关系更加复杂,仅对某一特征空间的信息进行增强,对表征能力的提升是有限的。因此,本文使用LPA同时在拓扑空间和属性空间增大同类节点间路径的权重,即同时关注标签与属性、拓扑之间的依赖关系,从而提升模型在不同数据集上的泛化性。

结束语 本文提出了一种融合标签传播算法的多通道图卷积模型MGCN-LPA,通过增大同类节点间路径的权重,来改善节点间特征和标签信息的传播,以增强模型对原始图信息的表征能力。在3个基准数据集的节点分类任务中,MGCN-LPA的实验表现均优于当前较为先进的基线模型,证明了引入多通道策略和LPA优化的有效性。此外,消融实验结果表明,融合拓扑和属性信息的MGCN-LPA能够更加充分地提取和保留原始图中蕴含的特征,相比仅使用拓扑或属性信息的方法表征能力和泛化性更强。在当前工作中,通过修改信息传递和聚合方式提升模型性能,忽略了图数据中可能存在的噪声。因此,在后续工作中将对图数据的节点属性、拓扑结构以及标签信息进行增强,提高模型的鲁棒性和泛化性。

参考文献

- [1] MCLAREN C D, BRUNER M W. Citation network analysis[J]. International Review of Sport and Exercise Psychology, 2022, 15(1): 179-198.
- [2] GHAREHCHOPOGH F S. An improved Harris Hawks optimization algorithm with multi-strategy for community detection in social network[J]. Journal of Bionic Engineering, 2023, 20(3): 1175-1197.
- [3] ZOU M, GAN Z, CAO R, et al. Similarity-navigated graph neural networks for node classification[J]. Information Sciences, 2023, 633: 41-69.

- [4] NASIRI E, BERAHMAND K, LI Y. Robust graph regularization nonnegative matrix factorization for link prediction in attributed networks [J]. *Multimedia Tools and Applications*, 2023, 82(3):3745-3768.
- [5] XU M. Understanding graph embedding methods and their applications[J]. *SIAM Review*, 2021, 63(4):825-853.
- [6] BHATTI U A, TANG H, WU G, et al. Deep learning with graph convolutional networks: An overview and latest applications in computational intelligence[J/OL]. *International Journal of Intelligent Systems*, 2023; 1-28. <https://dl.acm.org/doi/abs/10.1155/2023/8342104>.
- [7] LI C H, ZHU X F. Dual-channel graph random convolutional networks for semi-supervised node classification[J]. *Journal of Chinese Computer Systems*, 2023, 44(8):1656-1664.
- [8] GUO M H, XU T X, LIU J J, et al. Attention mechanisms in computer vision: A survey [J]. *Computational Visual Media*, 2022, 8(3):331-368.
- [9] VELICKOVIC P, CUCURULL G, CASANOVA A, et al. Graph attention networks[EB/OL]. [2023-12-12]. <https://arxiv.org/abs/1710.10903v3>.
- [10] LI B, JING B, TONG H, et al. Graph communal contrastive learning[C]//*Proceedings of the 22nd International Conference on World Wide Web*. New York: ACM, 2022; 1203-1213.
- [11] XUE L, NONG L P, ZHANG W H, et al. An improved graph convolution network semi-supervised node classification [J]. *Computer Applications and Software*, 2021, 38(10):153-158, 163.
- [12] RAGHAVAN U N, ALBERT R, KUMARA S. Near linear time algorithm to detect community structures in large-scale networks[J]. *Physical Review E*, 2007, 76(3):036106.
- [13] WANG H, LESKOVEC J. Unifying Graph Convolutional Neural Networks and Label Propagation [EB/OL]. [2023-12-12]. <https://arxiv.linfen3.top/abs/2002.06755v1>.
- [14] CUNNINGHAM P, DELANY S J. k-Nearest neighbor classifiers-A Tutorial[J]. *ACM Computing Surveys (CSUR)*, 2021, 54(6):1-25.
- [15] LIU J Q, TU W X, ZHU E. Survey on graph convolutional neural network[J]. *Computer Engineering & Science*, 2023, 45(8):1472-1481.
- [16] ZHANG L Y, SUN H H, SHI B B. Review of node classification methods based on graph convolutional neural networks[J/OL]. *Computer Science*, 2023; 1-19. <http://kns.cnki.net/kcms/detail/50.1075.TP.20230925.1655.162.html>.
- [17] WU F, SOUZA A, ZHANG T, et al. Simplifying graph convolutional networks[C]//*Proceedings of the 36th International Conference on Machine Learning*. Cambridge: PMLR, 2019; 6861-6871.
- [18] TAILOR S A, OPOLKA F, LIO P, et al. Do we need anisotropic graph neural networks? [C]//*Proceedings of the 2022 International Conference on Learning Representations*. Open-Review.net, 2022; 1-19.
- [19] BRODY S, ALON U, YAHAV E. How attentive are graph attention networks? [C]//*Proceedings of the 2022 International Conference on Learning Representations*. OpenReview.net, 2022; 1-26.
- [20] GUO X C, ZHANG W Y, XIA Z X. Two-way data augmentation graph convolutional networks [J]. *Computer Engineering and Design*, 2023, 44(8):2345-2351.
- [21] RANI V, NABI S T, KUMAR M, et al. Self-supervised Learning: A Succinct Review [J]. *Archives of Computational Methods in Engineering*, 2023, 30(4):2761-2775.
- [22] LI Y Q, WANG J, WANG F, et al. Semi-Supervised Node Classification Algorithm Based on Hierarchical Contrastive Learning [J]. *Pattern Recognition and Artificial Intelligence*, 2023, 36(8):712-720.
- [23] ZHONG L, YANG J, CHEN Z, et al. Contrastive graph convolutional networks with generative adjacency matrix [J]. *IEEE Transactions on Signal Processing*, 2023, 71:772-785.
- [24] HUANG J, DU L, CHEN X, et al. Robust mid-pass filtering graph convolutional networks[C]//*Proceedings of the 2023 International World Wide Web Conference*. New York: ACM, 2023:328-338.
- [25] NIE S M, DU Y H, LU T L, et al. Survey on local community detection algorithms[J]. *Software Guide*, 2020, 19(5):271-275.
- [26] WANG X, ZHU M, BO D, et al. Am-gen: Adaptive multi-channel graph convolutional networks[C]//*Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: ACM, 2020; 1243-1253.
- [27] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[EB/OL]. [2023-12-19]. <https://arxiv.org/abs/1609.02907v4>.
- [28] YUAN L N, LI X, WANG X D, et al. Graph embedding models: a survey[J]. *Journal of Frontiers of Computer Science and Technology*, 2022, 16(1):59-87.
- [29] HU Y, YOU H, WANG Z, et al. Graph-mlp: node classification without message passing in graph[EB/OL]. [2023-12-19]. <https://arxiv.org/abs/2106.04051v1>.
- [30] ZHENG W P, GE H L, LIU M L, et al. Node classification algorithm fusing 2-connected motif-structure information[J/OL]. *Journal of Computer Applications*, 2023; 1-10. <http://kns.cnki.net/kcms/detail/51.1307.tp.20230816.1507.006.html>.
- [31] ZENG J X, WANG P H, DING Y D, et al. Graph neural network based node embedding enhancement model for node classification[J]. *Journal of Zhe Jiang University(Engineering Science)*, 2023, 57(2):219-225.
- [32] CAI T T, MA R. Theoretical foundations of t-SNE for visualizing high-dimensional clustered data[J]. *The Journal of Machine Learning Research*, 2022, 23(1):13581-13634.



YUAN Lining, born in 1995, doctoral student, is a member of CCF (No. H5844M). His main research interests include machine learning and graph neural network.



FENG Wengang, born in 1982, Ph.D. professor, Ph.D supervisor. His main research interests include pattern recognition and public security intelligence.