

文章编号: 1671-9352(2024)03-0001-13 DOI: 10.6040/j.issn.1671-9352.7.2023.787

基于元路径属性融合的异质网络表示学习

王静红^{1,2,3}, 吴芝冰¹, 黄鹏¹, 杨家腾¹, 李笔^{4*}

(1. 河北师范大学计算机与网络空间安全学院, 河北 石家庄 050024; 2. 河北省网络与信息安全重点实验室(河北师范大学), 河北 石家庄 050024; 3. 供应链大数据分析与安全河北省工程研究中心(河北师范大学), 河北 石家庄 050024; 4. 河北师范大学商学院, 河北 石家庄 050024)

摘要: 针对信息网络的表示学习进行研究, 提出了一种基于元路径信息融合的异质图神经网络(metapath attribute fusion graph neural network, MAFGNN), 通过在异质网络中引入元路径之前将目标节点的邻居信息包括元路径信息融入到节点中, 实现目标节点和邻居信息的融合。该方法首先将不同类型的节点属性特征进行维度转换便于后续的融合操作, 通过计算目标节点和邻居节点权重值完成目标节点信息的融合操作。然后根据特定元路径对目标节点进行融合, 最后在不同元路径间实现不同语义信息的融合操作。在多个异质信息数据集上进行实验表明, MAFGNN 模型在处理异质网络节点嵌入方面相比于最先进的基准实验有最好的性能和更加准确的预测结果。

关键词: 元路径; 异质信息网络; 异质图嵌入; 信息融合; 注意力机制

中图分类号: TP181 **文献标志码:** A

引用格式: 王静红, 吴芝冰, 黄鹏, 等. 基于元路径属性融合的异质网络表示学习[J]. 山东大学学报(理学版), 2023, 59(3): 1-13.

Heterogeneous network representation learning based on metapath attribute fusion

WANG Jinghong^{1,2,3}, WU Zhibing¹, HUANG Peng¹, YANG Jiateng¹, LI Bi^{4*}

(1. School of Computer and Cyberspace Security, Hebei Normal University, Shijiazhuang 050024, Hebei, China; 2. Hebei Key Laboratory of Network and Information Security (Hebei Normal University), Shijiazhuang 050024, Hebei, China; 3. Hebei Provincial Engineering Research Center for Supply Chain Big Data Analytics and Security (Hebei Normal University), Shijiazhuang 050024, Hebei, China; 4. School of Business, Hebei Normal University, Shijiazhuang 050024, Hebei, China)

Abstract: Focusing on the research on representation learning of information networks, a metapath attribute fusion graph neural network (MAFGNN) based on metapath information fusion is proposed, which is to integrate the neighbor information of the target node, including the metapath information, into the node before introducing the metapath in the heterogeneous network to achieve the fusion of target node and neighbor information. This method first converts the attribute features of different types of nodes into dimensions to facilitate subsequent fusion operations. The fusion operation of target node information is completed by calculating the weight values of target nodes and neighbor nodes. Then target nodes are fused according to specific metapaths, and finally different semantic information is fused between different metapaths. Experiments on multiple heterogeneous information datasets show that the MAFGNN model has the best performance and more accurate prediction results than the most advanced benchmark experiments in dealing with heterogeneous network node embedding.

Key words: metapath; heterogeneous information network; heterogeneous graph embedding; information fusion; attention mechanism

收稿日期: 2023-04-29; 网络出版时间: 2023-11-29 13:14:13

网络出版地址: <https://link.cnki.net/urlid/37.1389.N.20231128.1346.002>

基金项目: 河北省自然科学基金资助项目(F2021205014); 河北省高等学校科学技术研究项目(ZD2022139); 中央引导地方科技发展资金项目(226Z1808G)

第一作者: 王静红(1967—), 女, 教授, 硕士生导师, 博士, 研究方向为人工智能、数据挖掘、模式识别等. E-mail: wangjinghong@126.com

* 通信作者: 李笔(1966—), 男, 副教授, 硕士, 研究方向为机器学习、复杂网络、智能信息系统等. E-mail: libilb@263.net

0 引言

随着互联网数据信息的逐渐增加,现实世界中的数据大多都是用图网络的结构形式来表示,如引文网络^[1]和社交网络。随着信息数据的复杂化和多类型化,根据信息网络模型中数据是否属于同类型,可将其划分为同质信息网络和异质信息网络^[2]。在同质信息网络中,节点代表相同的实体。异质信息网络^[3]是包含两种或者以上的节点或者边类型,每种节点或者边都包含本类型的属性特征。由于异质信息网络的复杂性,传统的图表示学习方法可能无法充分捕捉到节点之间的多样性和关联关系。怎样将异质信息网络进行学习表示是一项巨大的挑战。异质信息网络表示学习是图表示学习的一个具体应用场景。图表示学习(也称为图嵌入)^[4]是将图节点或边节点进行低维向量表示,是社区发现和推荐任务^[5]的上游工作。图表示学习是从最开始受到 word2vec^[6]和 DeepWalk^[7]节点序列嵌入的启发到后来的改进工作,如 metapath2vec^[8]设计的元路径^[9]在图神经网络中起到了保留节点间语义信息的功能的嵌入工作。

与同质网络嵌入不同的是,异质信息网络包含了节点与边之间的信息异质性,用传统的随机游走等方式通常不能较好地体现这种异质性,所以异质网络表示学习大多采用了元路径的方式。基于元路径引导的随机游走生成的节点序列不仅包含结构信息,还包含语义信息。BHIN2vec^[10]通过相对训练比率来关注具有高损失值的任务,通过随机游走来生成具有丰富语义的异质节点,并使用一种扩展的 skip-gram^[6]技术来平衡各种类型的关系,将异质网络嵌入视为多个基于关系的任务。HAN^[11]在节点级和语义级上学习不同元路径的重要性表示,但是缺少了在元路径内目标节点在信息网络中的重要性。上述的表示学习方式都没有考虑到网络结构中元路径属性对节点的影响作用,忽略了元路径上节点属性特征的信息融合。元路径聚合图神经网络(metapath aggregated graph neural network, MAGNN)^[12]在元路径层面,引入了注意力机制来进行节点嵌入。随着网络动态性的发展,动态异构网络嵌入(dynamic heterogeneous network embedding, DyHNE)^[13]解决了异质网络表示学习方法中扩展性差、无法处理动态网络的问题。它是一种基于矩阵扰动理论的增量更新方法,在考虑网络中异质性和演化过程的同时学习节点嵌入,为保证算法的高效性同时保留了一阶和二阶近似的元路径。DyHATR^[14]则是使用不同层级的注意力机制来获取不同时间的节点信息。

本文提出了一种全新的基于元路径的异质信息网络,采用基于元路径上节点属性特征融合的方式来解决异质信息的融合问题。具体地说,该模型首先在输入上对目标节点和邻居节点进行特定的特征类型转换,从而便于后续融合操作,然后将目标节点和邻居节点进行属性融合,得到的输出作为下一层网络的输入;其次在指定元路径上采用目标节点和邻居节点的节点级注意力学习在该元路径下的节点嵌入;最后设计双层多层感知机(multi-layer perceptron, MLP),根据不同的元路径采用语义级注意力学习最后的节点嵌入。

总而言之,这篇文章的创新点如下:

(1)设计了一种新的解决目标节点和邻居节点特征融合方式,将目标节点和邻居节点特征维度投影到相同潜在空间,然后将同维度特征进行融合,可以让模型对处理复杂的异质信息网络有着很好的效果,扩展了异质信息网络的类型。

(2)提供了一种新的异质信息网络模型,解决异质网络嵌入的问题,在不同元路径考虑到网络中重要的语义信息,采用在元路径上的节点嵌入和在不同元路径上的语义级嵌入对属性特征的融合问题有很好的帮助效果。

(3)最后在与多个传统的网络嵌入模型和异质图神经网络模型进行实验比较,使用多个数据集在该模型进行表示学习,在异质网络上节点属性特征元路径上语义信息提取结果相比于之前的模型都有很好的效果。

1 相关工作

网络表示学习又可以称为图表示学习、图嵌入,可进行节点分类和聚类等任务,并且其又分为同质网络嵌入和异质网络嵌入两方面。

1.1 同质网络嵌入

面对现实世界中图结构数据的复杂性和多元性,网络嵌入是将图结构数据映射到低维向量空间,在保证较好地保留图信息的前提下,为下游工作如用户偏好推荐、社区划分和优化算法等提供更高效的方式。一般的图数据的表示分为节点层面的表示和全图层面的表示,但节点作为图数据的基本组成元素,在节点层面的表示上能够更好地挖掘图数据的有效或潜在的信息。同时为解决图数据特征在拓扑结构上的稀疏性和高计算复杂性,理解节点间的语义关系,掌握不同的语义路径,图表示学习为研究图结构数据被提出了大量的算法理论。例如,受到了词嵌入^[15]的启发,DeepWalk 算法通过设计出随机游走的方式,在网络结构中采样到节点序列,然后得到节点嵌入。LINE^[16]算法对大型复杂图进行嵌入,与 DeepWalk 不同的是可以在有向图上进行学习表示,提出当两个节点的邻居节点具有相似性,那么这两个节点在表示上也具有一定的相似性。Node2vec 是通过改变采样策略,使用深度优先搜索和广度优先搜索的方式进行采样。struct2vec^[17]在结构上考虑两个节点的相似性。近年来图神经网络(graph neural network, GNN)^[18]的有效推广,图卷积网络(graph convolutional network, GCN)^[19]引入卷子算子来进行嵌入。图注意力网络(graph attention network, GAT)^[20]又在 GNN 的基础上加入注意力机制^[21],通过计算邻居节点上的权重来学习节点级别的嵌入,然后又推广到了多头注意力机制来提高嵌入能力。但现实世界中大多数图信息网络都是异质性的,考虑到图信息网络的异质性,又提出了一些新的方法。

1.2 异质信息网络嵌入

由于现实世界中数据信息的异质性和复杂性,一般对于信息网络采用异质网络的形式表示它的网络结构。与同质信息网络不同的是,异质信息网络表示学习则是将网络中复杂的多类型节点信息和节点间链接信息做低维向量表示,需要考虑不同类型节点对于邻居节点的重要程度以及节点间连接关系的语义信息。比如 Dong^[8]对异质网络的研究引入了元路径,采用元路径获取节点序列的方式输入到 skip-gram 模型得到异质网络的节点嵌入。然而该模型只使用了一条元路径,可能会忽略其他元路径上有用的语义信息。HIN2Vec^[22]通过利用节点之间不同类型的连接关系来学习异质信息网络中的潜在向量。ESim^[23]通过用户指定的元路径来解决异质信息网络的表示学习,实现了采用多条元路径的方式来解决节点嵌入。HMSG^[24]则是在多条元路径中生成基于元路径的子图,对不同类型节点信息聚合操作得到子图聚合向量,然后加入注意力机制给每个子图赋予不同的权重,最后得到异质网络的节点嵌入。HAHE^[25]提出了在元路径和元路径实例上使用层次注意力机制来捕获元路径的语义信息和路径实例上的个性化偏好,在不同元路径传达不同的语义问题上得到了解决。随着异质图神经网络(heterogeneous graph neural network, HGNN)^[26]的推广, HAN 在节点级别上和语义级别上实现了对异质网络信息的表示学习,克服了异质信息网络属性的异质性融合问题。MAGNN 也在元路径层面引入了注意力机制来进行节点嵌入。

2 基本定义

本节给出本文所有相关内容的重要定义。

定义 1 异质信息网络。与同质信息网络不同的是,异质信息网络是具有多种节点类型和多种边类型的网络。定义异质信息网络为 $G=(V,E)$,其中 V 代表节点集, E 代表边集, V 中的节点和 E 中的边都有多种类型。

由于网络异质性,那么还存在着节点类型映射函数 $\Phi:V\rightarrow A$, A 表示节点类型集合;边类型映射函数 $\psi:E\rightarrow R$,其中 R 表示边类型集合, $|A|+|R|>2$ 。对于异质信息网络中所有的节点 V 和边 E 存在类型映射关系, $v\in V,\Phi(v)\in A$ 表示每个节点属于对应的节点类型, $e\in E,\psi(e)\in R$ 表示每条边属于对应的边类型。例如图 1(a)和(b)所示的异质信息网络描述了 3 种类型的节点(作者、论文、期刊)和定义它们之间关系的边类型(发表/被发表、录用/被录用)的异质信息网络。图 1(c)和(d)为图 1(b)的两条元路径和对应的实例。图 1(e)和(f)为图(b)中 p_2 对应的基于 PAP 元路径的邻居和子图。

定义 2 元路径。图结构中元路径表示为 $\pi:A_1\stackrel{R_1}{\rightarrow}A_2\stackrel{R_2}{\rightarrow}\dots\stackrel{R_l}{\rightarrow}A_{l+1}$ 可以简写为 (A_1,A_2,\dots,A_{l+1}) ,其中边类型 $R=R_1\circ R_2\circ\dots\circ R_l$,其中“ \circ ”表示连接关系上的复合运算, L 表示元路径 p 的长度。

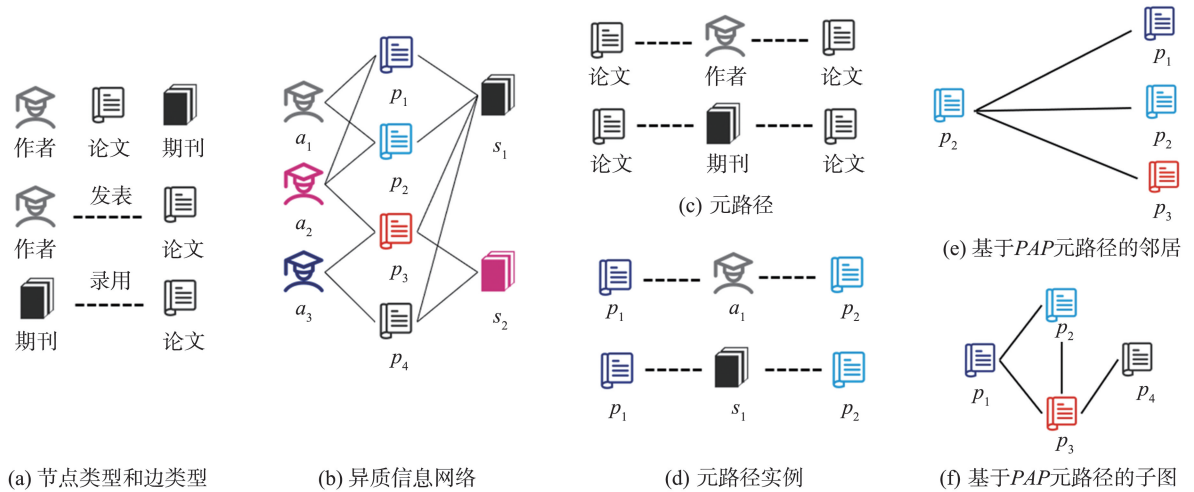


图 1 异质信息网络和元路径

Fig. 1 Heterogeneous information networks and metapaths

元路径分为对称元路径和非对称元路径,若元路径 p 与反向元路径 p^{-1} 相同,则该元路径为对称元路径。如果 $\forall i=1,2,\dots,n, A_i=\Phi(v_i), R_i=\psi(v_i, v_{i+1})$ 路径 p 经过节点 v_1, v_2, \dots, v_n 则表示路径 p 是在元路径 π 上的一个实例,简单来说就是节点和边具体化的元路径 p 称为元路径的实例。一条元路径包含着丰富节点信息和边信息,对于处理图信息网络表示学习有很好的帮助效果。本文将元路径中的起始节点统一为目标节点。

例如在图 1(c)和(d)中, $p_1-a_1-p_2$ 是基于“论文→作者→论文”(PAP)元路径下的元路径实例,同样 $p_1-s_1-p_2$ 是基于“论文→期刊→论文”(PJP)下的元路径实例。

定义 3 基于元路径的邻居。在给定目标节点 V 和元路径 P ,基于元路径 P 得到包含类型与目标节点相同的节点 N_V^P 被定义为目标节点 V 基于元路径 P 的邻居节点集合。

例如图 1(d)和(e)中,基于元路径 PAP,关于节点 p_2 的邻居节点有 p_1, p_2, p_3 包含目标节点本身。

定义 4 基于元路径的子图。基于元路径 P 在异质信息网络 G 上构成的只有一种类型的节点集的同质图 G^P 。在该同质图中,邻居节点是根据元路径上对称位置的节点集合,具有同种类型的信息和属性。

定义 5 异质网络的表示学习。给定的异质信息网络 $G=(V, E)$,通过学习目标函数 $f: V \rightarrow Z, Z \in \mathbf{R}^{|V| \times d} (d \ll |V|)$ 将目标节点进行 d 维上学习表示,能够解决邻接矩阵稀疏性问题和捕获丰富的结构、语义和属性信息。

表 1 为文章中使用的符号。

表 1 本文使用各种符号汇总

Table 1 This article uses various symbols to summarize

符号	解释
G	$G=(V, E)$ 异质图信息网络
V, G	异质网络中的节点集、边集
P	元路径
G^P	基于元路径 P 的子图
h, h'	初始节点特征向量、转换后节点特征向量
p	基于元路径 P 的路径实例
W	特定节点类型转换矩阵
$e_{i,j}^P$	基于元路径 P 的节点对 (i, j) 重要性系数
h_{node}^P	元路径 P 的节点级注意力向量
Z_v	语义层级节点嵌入向量

3 模型介绍

本文提出了一种解决异质信息网络节点嵌入的 MAFGNN。MAFGNN 模型结构如图 2 所示。该模型首先在输入上对目标节点和邻居节点进行特定的特征类型转换便于后续融合操作。然后进行目标节点和邻居节点的属性融合并将结果输出作为下一层网络的输入。其次在指定元路径上采用目标节点和邻居节点的节点级注意力学习在该元路径下的节点嵌入。最后通过双层 MLP 根据不同的元路径采用语义级注意力学习最后的节点嵌入。在模型中经过损失函数和神经网络反向传播机制来学习设置的所有权重系数。

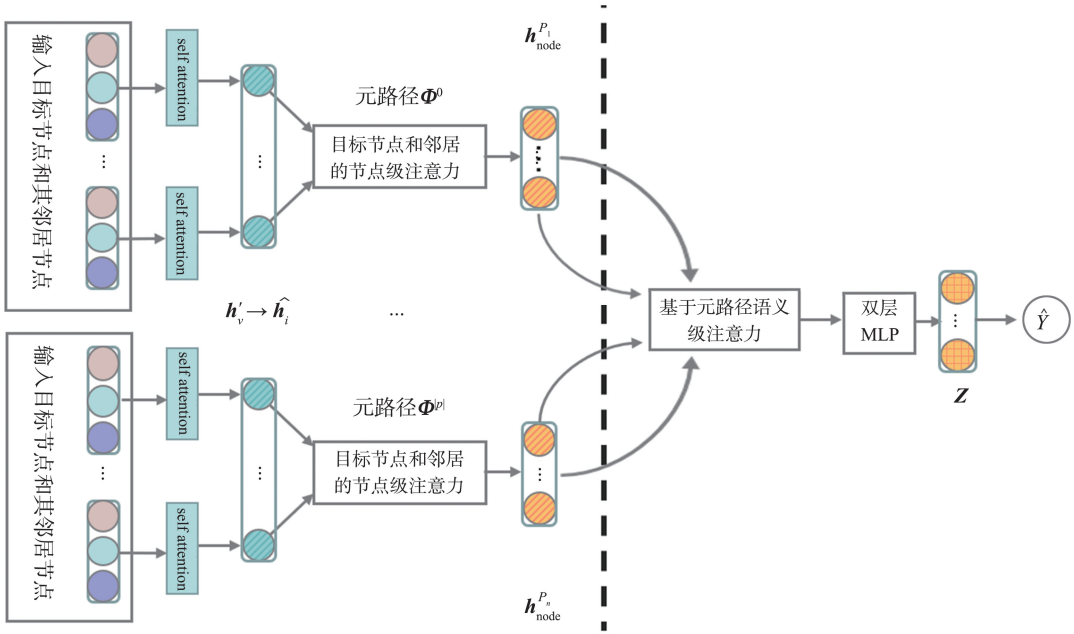


图2 MAFGNN 模型示意图
Fig.2 Schematic diagram of MAFGNN model

3.1 节点特征转换

在处理与节点属性相关的异质信息网络时,面对不同类型的节点特征需要将其转换为相同维度的特征向量,从而实现不同类型节点之间的属性融合操作,获取更多的语义信息。例如在引文网络中的论文节点和作者节点,它们分别具有不同维度的特征向量,在不同的特征空间中表示。为了便于后续不同类型节点间特征向量的操作,因此首先需要将节点特征投射到相同的潜向量空间。

在 MAFGNN 模型中输入的是节点特征信息和异质信息网络邻接矩阵信息。节点特征向量可以表示为 $\mathbf{h}=(\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_N), \mathbf{h}_i \in \mathbf{R}^{N \times M}$ 。其中不同类型节点有着不同的特征维度数 M 和节点个数 N 。通过设计了节点特征维度转换层,可以将不同维度的特征向量经过特定类型的线性变化投影到相同维度的潜向量空间,对于节点类型为 $a \in A$ 的节点 $v \in V_A$,有

$$\mathbf{h}'_v = \mathbf{W}_A \mathbf{h}_v^A, \quad (1)$$

式中: \mathbf{h}'_v 是经过线性投影得到的节点特征向量; \mathbf{W}_A 是特定线性转换的参数权重矩阵; \mathbf{h}_v^A 是节点的原始特征向量。输出结果为 $\mathbf{h}'_v=(\mathbf{h}'_1, \mathbf{h}'_2, \dots, \mathbf{h}'_N)$ 。

不同类型的节点经过 MAFGNN 中的节点特征转换层之后可以转换到相同维度的潜向量空间,具有相同维度的节点特征向量可以在后续节点属性聚合上进行更方便的操作。

3.2 目标节点和邻居节点的属性融合

通过节点特征转换可以得到不同类型节点的相同维度节点向量,对于目标节点和异质信息网络中的邻居节点属性融合得到新的节点表示。属性融合过程可以获取到异质信息网络中潜在的语义信息和结构信息。在这里只选取目标节点的一阶近邻来作为邻居节点,其中目标节点 v 和邻居节点集 $u \in U$,需要使用一种注意机制来计算目标节点和邻居节点的注意力系数从而来进行属性融合过程。定义 $(\mathbf{h}'_v, \mathbf{h}'_u)$ 为目标节点和邻居节点的节点对,定义 $e_{v,u}$ 为目标节点和一阶近邻的邻居节点之间的权重系数,有

$$\mathbf{e}_{v,u} = \mathbf{a}(\mathbf{W}\mathbf{h}'_v, \mathbf{W}\mathbf{h}'_u), \quad (2)$$

式中: \mathbf{a} 为可学习的线性权重向量参数; \mathbf{W} 表示为可学习的权重矩阵参数,用于学习特征的线性变换; \mathbf{h}'_v 和 \mathbf{h}'_u 表示经过节点特征转换后的目标节点和邻居节点特征向量。对 $\mathbf{e}_{v,u}$ 使用 softmax 归一化函数对所有权重系数进行归一化计算,有

$$a_{v,u} = \text{softmax}(\mathbf{e}_{v,u}) = \frac{\exp(\sigma(\mathbf{a}^T(\mathbf{W}\mathbf{h}'_v \parallel \mathbf{W}\mathbf{h}'_u)))}{\sum_{k \in U} \exp(\sigma(\mathbf{a}^T(\mathbf{W}\mathbf{h}'_v \parallel \mathbf{W}\mathbf{h}'_k)))}, \quad (3)$$

式中: $a_{v,u}$ 为归一化后的权重系数; σ 表示激活函数, 在所做的实验中此处激活函数设计为 ELU 非线性函数, 参数设置为 0.2 (负输入斜率 $\alpha=0.2$), 用于前馈神经网络的传播; \parallel 表示向量拼接操作; \mathbf{a}^T 为元路径下注意力系数, 可通过前馈传播学习得到。式(3)中目标节点的邻居信息包含本身节点, 意味着在对目标节点和邻居节点的属性融合包含本身因素的影响, 并且也体现出可以更好地获取目标节点在异质信息网络中的结构信息和语义信息。

最后经过相应的权重系数对目标节点特征向量进行聚合得到如下输出: $\hat{\mathbf{h}} = a_{v,u} \mathbf{h}'_v = a_{v,u} (\mathbf{h}'_1, \mathbf{h}'_2, \dots, \mathbf{h}'_N)$ 。

3.3 元路径上的属性融合

根据异质信息网络的拓扑结构给定一组元路径 $\{P_1, P_2, \dots, P_n\}$ 来提取语义信息 (n 为指定元路径数量), 基于元路径 P 进行随机游走得到的同类型节点集作为该元路径下的同质网络 G^P , 意味着目标节点通过元路径 P 与其邻居相连, 邻居节点都是同种类型的节点, 最后给出同质网络中节点和其邻居的邻接矩阵 M^P 。元路径就是连接两个实体的一条特定的路径, 对称元路径就是第一个节点和最后一个节点类型相同的路径。具体来说, 基于元路径得到同质网络的方式是, 指定一组元路径来提取异质网络中不同的语义信息, 在该条元路径下通过间接连接的节点即为同质网络中的邻居。如图 3(a)–(d) 异质信息网络中, 由于节点的异质性, 可以通过节点类型来区分元路径类型, 当元路径类型为 PAP 时, 对应元路径实例节点类型只能为“论文类型、作者类型、论文类型”的节点连接得到的路径实例。然后将路径实例两端类型相同的节点作为该目标节点的邻居节点从而构成类型相同的同质图网络。

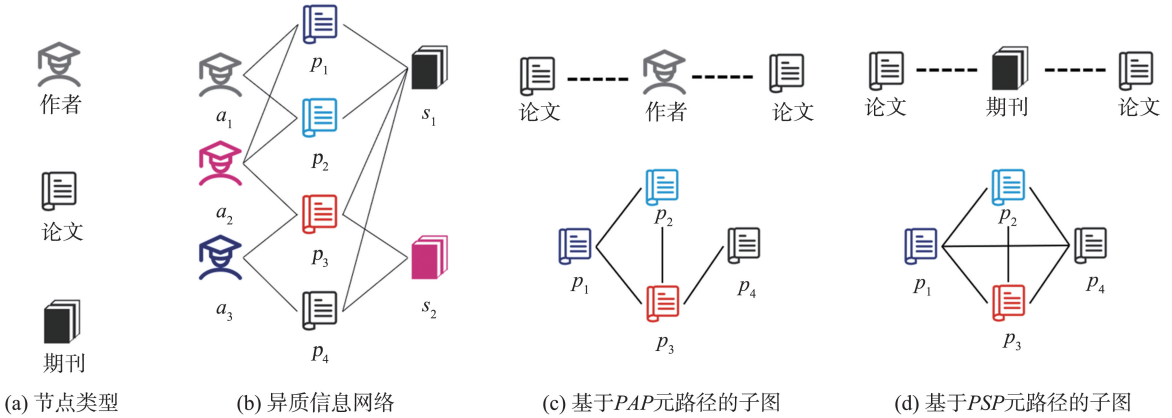


图3 基于不同的元路径得到的网络结构

Fig.3 Network structures based on different metapaths

基于上述中的目标节点在异质网络拓扑结构中的位置和同质图 G^P , 使用基于元路径的节点级注意力。在给定元路径 P_i 的情况下, 对基于元路径 P_i 目标节点和邻居节点的节点对 $(\hat{\mathbf{h}}_i, \hat{\mathbf{h}}_j)$ 设计了邻居节点 $\hat{\mathbf{h}}_j$ 对目标节点 $\hat{\mathbf{h}}_i$ 的重要程度。图 4(a) 和 (b) 分别为基于 PAP 和 PSP 不同元路径目标节点的邻居节点, 其中基于元路径邻居节点对目标节点的重要性可以表示为

$$\Phi_{i,j}^P = A_{\text{node}}(\hat{\mathbf{h}}_i, \hat{\mathbf{h}}_j; P), \quad (4)$$

式中: $\Phi_{i,j}^P$ 表示节点 $\hat{\mathbf{h}}_j$ 对节点 $\hat{\mathbf{h}}_i$ 的重要性; A_{node} 表示执行节点级注意力的深度神经网络; 邻居节点 $\hat{\mathbf{h}}_j$ 可以包含 $\hat{\mathbf{h}}_i$ 节点; $(\hat{\mathbf{h}}_i, \hat{\mathbf{h}}_j)$ 为 3.2 节属性融合后的节点向量。在相同元路径下, 所有节点共享 A_{node} 。获得了基于元路径的目标节点和邻居节点的重要性系数之后, 目标节点基于元路径的节点嵌入可以表示为

$$\mathbf{h}_{\text{node}}^P = \sum_{j \in J} \text{softmax}(\Phi_{i,j}^P) \hat{\mathbf{h}}_j, \quad (5)$$

$$\text{softmax}(\Phi_{i,j}^P) = \frac{\exp(\sigma(\Phi_{i,j}^P))}{\sum_{K \in J} \exp(\sigma(\Phi_{i,K}^P))} = \frac{\exp(\sigma(A_{\text{node}}(\hat{\mathbf{h}}_i, \hat{\mathbf{h}}_j; P))}{\sum_{K \in J} \exp(\sigma(A_{\text{node}}(\hat{\mathbf{h}}_i, \hat{\mathbf{h}}_K; P)))}, \quad (6)$$

式中 J 为目标节点的邻居节点。在该公式中重要性系数都取决于节点的特征向量。经过式(6)计算可以得到基于一组元路径 $\{P_1, P_2, \dots, P_n\}$ 的节点级嵌入 $\{\mathbf{h}_{\text{node}}^{P_1}, \mathbf{h}_{\text{node}}^{P_2}, \dots, \mathbf{h}_{\text{node}}^{P_n}\}$ 。这里的节点级嵌入使用自注意力方式, 将注意力头数设置为 K , 当注意力头数 $K > 1$ 时, 有如下嵌入方式

$$\mathbf{h}_{\text{node}}^p = \prod_{K=1}^K \sum_{j \in J} \text{softmax}(\Phi_{i,j}^p) \mathbf{h}_j. \quad (7)$$

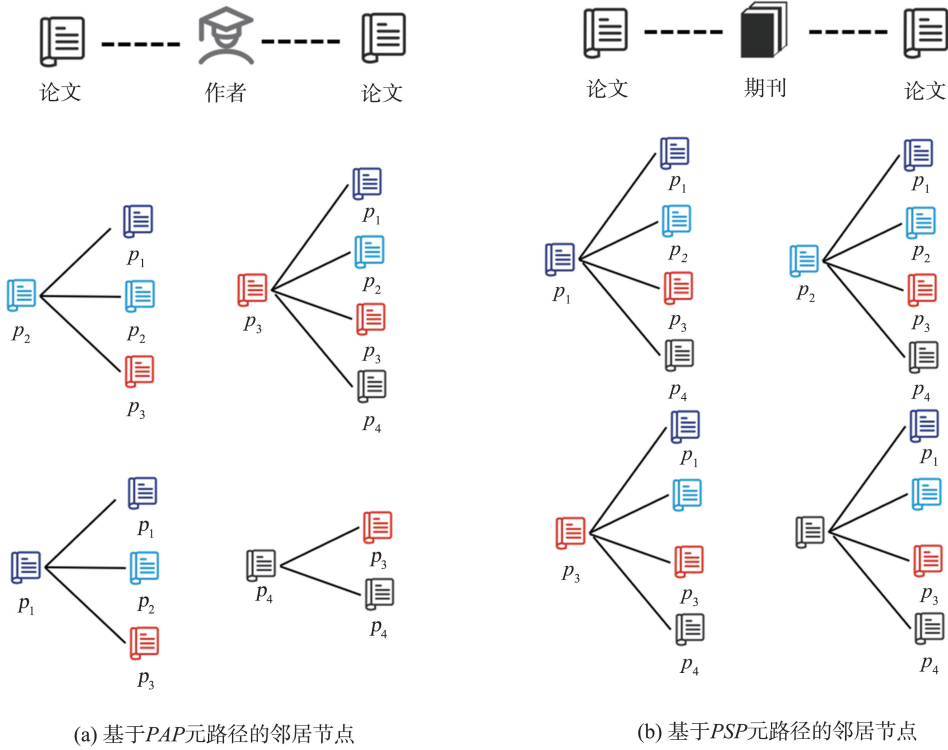


图4 基于不同元路径目标节点的邻居节点

Fig.4 Neighbor nodes based on different metapath target nodes

最后根据不同的元路径去了解它们之间的重要性,设计了双层非线性变换来进行不同语义的嵌入。设计双层非线性变换的目的是提高网络模型的复杂度,使模型的收敛速度更加的稳定。在每一个元路径上的所有目标节点的语义级重要性做均值处理,在指定元路径下目标节点对不同元路径的语义信息不存在影响因素。首先第1层使用 tanh 函数,第2层使用 LeakyReLU 函数,增设不同的权重矩阵 \mathbf{W}_1 、 \mathbf{W}_2 和同一个偏移量 b ,每个元路径的重要性表达可以表示为

$$\psi_p = \frac{1}{|V|} \sum_{v \in V} \mathbf{q}^T \text{LeakyReLU}(\mathbf{W}_1(\tanh(\mathbf{W}_2 \mathbf{h}_{\text{node}}^p) + \mathbf{b})), \quad (8)$$

式中: ψ_p 为不同元路径语义级重要性系数; $|V|$ 为元路径上的目标节点数量; \mathbf{q}^T 为不同元路径的语义级向量,可通过前馈神经网络学习得到。最后在不同元路径上进行归一化处理,经过融合操作可以得到最后目标节点的嵌入向量

$$\mathbf{Z}_v = \frac{\exp(\psi_p)}{\sum_{p \in P} \exp(\psi_p)} \mathbf{h}_{\text{node}}^p, \quad (9)$$

式中: \mathbf{Z}_v 为目标节点的模型输出的嵌入向量; $\mathbf{h}_{\text{node}}^p$ 为目标节点在融合邻居节点特征的节点嵌入。

3.4 损失函数和训练方式

本文模型最终输出为节点嵌入向量并用于节点分类任务,通过使用半监督的方式来学习模型中的参数,设置输出预测向量和真实向量之间的交叉熵为

$$L = - \sum_{v \in V} \hat{Y}_v \ln(\mathbf{C}^T \mathbf{Z}_v), \quad (10)$$

式中: \hat{Y}_v 为节点的真实标签; \mathbf{C} 为分类器参数。MAFGNN 模型通过网络模型的反向传播来优化参数,同时使用早停机制来进行的训练。MAFGNN 模型的算法过程如表2所示。假设元路径条数为 n , 节点个数为 $|V|$ 。第1次循环遍历所有目标节点,计算目标节点的节点级嵌入向量。第2次循环遍历所有元路径,计算目标节点的语义级嵌入向量。因此,表2中算法1的时间复杂度为 $O(n|V|)$ 。

表2 MAFGNN 算法过程
Table 2 MAFGNN algorithm process

算法1 基于元路径与属性融合的异质网络表示学习

输入: 异质信息网络 $G = \{V, E\}$, 节点特征 $\mathbf{h} = (\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_N)$, 元路径集合 $\{P_1, P_2, \dots, P_n\}$, 节点类型 $A = \{A_1, A_2, \dots, A_{|A|}\}$;
输出: 节点最终表示 Z 。

- 1 for v in V do //对所有目标节点进行操作。
- 2 通过公式(1)将目标节点和目标节点的邻居节点进行特征维度转换得到转换后的向量表示 $\mathbf{h}'_v, \mathbf{h}'_u$;
- 3 通过公式(2)和(3)计算转换后的目标节点和邻居节点之间的权重系数并得到输出结果: $\mathbf{h} = a_{v,u} \mathbf{h}'_v = a_{v,u} (\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_N)$;
- 4 end for //终止循环。
- 5 for $\{P_1, P_2, \dots, P_n\}$ in P do
- 6 for v in V do
- 7 通过公式(4)计算基于元路径 P_i 的邻居节点重要性 $\Phi_{i,j}^P$;
- 8 通过公式(5)目标节点基于元路径的节点嵌入 \mathbf{h}_{node}^P ;
- 9 if 注意力头数 $K > 1$ do
- 10 公式(7)计算嵌入方式
- 11 end for //终止循环
- 12 end for //终止循环
- 13 通过公式(8)得到每一条元路径的重要性 ψ_p ;
- 14 通过公式(9)得到最终的目标节点表示向量 Z_v ;
- 15 return $Z_v \in V$ //返回目标节点向量。

4 实验

4.1 数据集

本文实验采用网络上公开的 DBLP(Digital Bibliography & Library Project)、IMDB(Internet Movie Database)和 ACM(Association for Computing Machinery)3种异质网络数据资源,经过相应的处理方式得到可处理的异质网络结构,表3是所使用的数据集的详细描述。

表3 实验所需数据集
Table 3 Experimental dataset required

数据集	节点类型	节点数量	边关系	边关系数量	元路径
DBLP	A	4 057			
	P	14 328	A—P	19 654	APA
	T	7 723	P—T	85 810	APCPA
	V	20	P—V	14 328	APTPA
IMDB	M	4 278			
	D	2 081	M—D	4 278	MAM
	A	5 257	M—A	12 828	MDM
ACM	A	5 912			
	P	3 025	P—A	9 936	PAP
	S	57	P—S	3 025	PSP

DBLP 是计算机科学书目提供有关主要计算机学期刊和会议的开放书目信息。实验选取 DBLP 数据集的 1 个子集,通过数据预处理后得到的节点包含 4 057 位作者(author, A)、14 328 篇论文(paper, P)、7 723 个学术术语(term, T)、20 个出版场所(venue, V)的异质信息数据。其中作者分为数据库、数据挖掘、机器学习和信息检索 4 个领域,按照每位作者发表论文的出版场所,将作者的领域贴上标签。每个作者特征都是由关键词的词袋特征组成。最后选取 3 类元路径 $\{APA, APCPA, APTPA\}$, 随机选取 2 000 个作者节点作为训练集,500 个节点作为验证集,500 个节点作为训练集进行实验。

IMDB 数据集是一个互联网电影数据库,包含关于电影演员、电影、电视节目、电影影评、导演等信息。

实验使用网上搜集的 IMDB 的一个子集,通过预处理之后得到包含 4 278 部电影(movie, M)、2 081 位导演(director, D)和 5 257 位演员(actor, A)。其中电影数据分为 3 类节点(动作片、喜剧片、戏剧片),电影特征由词袋特征组成。最后实验选取两类元路径 $\{MAM, MDM\}$,随机选取 2 000 条电影节点作为训练集,500 个节点作为验证集,500 个节点作为测试集来进行实验。

ACM 数据集是关于计算机学术文献数据集,本实验通过提取 ACM 数据集的一个子集数据,经过预处理得到一个包含 5 912 位作者、3 025 篇论文、57 个主题(subject, S)的 ACM 子集,其中论文具有 3 种标签:数据库、无线通信、数据挖掘。使用的元路径为 $\{PAP, PSP\}$ 。从中随机选取 1 000 篇论文作为训练集,500 篇论文作为验证集,1 525 篇论文作为测试集。

4.2 基准实验

本文提出的 MAFGNN 模型,通过与一些传统得同质网络模型和异质网络模型的基准实验进行了对比,包括同(异质)网络嵌入模型和同(异质)图神经网络模型来验证 MAFGNN 模型的优越性。为了验证 MAFGNN 模型的有效性,选取如下模型进行对比。

(1) DeepWalk:基于随机游走和 word2vec 两种算法的图嵌入算法,该算法能挖掘网络结构中的隐藏信息,可以得到图中节点的潜在向量表示。但该方法只能用于同质图,因此可以忽略节点的异质性来对异质图进行实验。

(2) metapath2vec:采用元路径的随机游走方式,解决了节点的异质性,生成的序列使用 skip-gram 得到异质网络节点嵌入。该对比实验测试了 metapath2vec 所有元路径并报告实验中具有最佳性能的元路径。

(3) GCN:一个传统的同质图神经网络,该模型在图的傅里叶域进行卷积运算。经过测试所有基于元路径的同质图并报告来自最佳元路径的结果。

(4) GAT:一个传统的同质图神经网络,该模型通过给邻居节点分配权重来聚合图中节点的邻居信息得到节点嵌入。同样测试所有基于元路径的同质图并报告来自最佳元路径的结果。

(5) HERec:一种异质网络推荐算法,通过使用元路径的方式学习同质图节点的表示,然后通过融合方式得到最后每个节点的输出。采用异质信息网络数据中的所有元路径进行实验并最终报告最佳性能。

(6) HAN:一种采用节点层次注意力和语义层次注意力的半监督异质图神经网络,该模型在基于不同的元路径得到同类型节点的同质图来学习目标节点嵌入。

4.3 参数设置

MAFGNN 模型中设置了随机矩阵转换参数,来学习不同类型节点特征维度不同情况下的转换矩阵参数。通过实验对比不同矩阵输出维度对模型准确率的影响,使该网络模型可以更好地适应复杂的信息网络。在实验参数设置中,将学习率设置为 0.005,优化模型使用 Adam 优化器,权值衰减设置为 0.001,语义层注意向量维度设置为 128,注意力头数设置为 8。基于异质图神经网络的 GCN、GAT、HAN 模型上将 dropout 设置为 0.5,使用相同划分的训练集、验证集和测试集以确保公平性。设定训练图神经网络最大轮数 epoch 参数设置为 250,并以 100 个 patience 作为耐心值来训练。对于传统的同质图神经网络,忽视异质网络中不同类型的节点进行实验。在基于随机游走的 DeepWalk、metapath2vec 模型上设置窗口大小为 5,随机游走步长为 100,负采样数设置为 5,然后将算法的嵌入维度设置为 64。

实验采用 Python 语言编写,所用到的数据集全部通过预处理之后得到符合自己网络模型的数据格式。实验设备为 Legion Y7000 2020H;处理器-CPU:Inter (R) Core (TM) i7-10750H CPU @ 2.60GHz;内存:SODIMM 16G;GPU:NVIDIA GeForce GTX 1650 4G;操作系统:windows 10 家庭中文版。

4.4 节点分类实验

节点分类在图嵌入模型中被广泛使用,使用 Macro-F1 分数和 Micro-F1 分数作为节点嵌入的评价指标,其数值越接近 100 则表示模型训练的精确率越高。采用 $k=5$ 的 KNN 分类器来进行节点分类,模型被训练好之后经过图神经网络的前馈传播便可得到目标节点的嵌入向量。由于图结构数据的方差可能非常高,所以将所有实验进行 20 次实验取平均值进行报告。表 4 为各个模型的 Macro-F1 分数和 Micro-F1 分数的平均值。

表4 节点分类实验结果
Table 4 Node Classification Experiment Result

Dataset	Metrics	DeepWalk	metapath2vec	GCN	GAT	HERec	HAN	MAFGNN
DBLP	Macro-F1	84.81	91.89	92.38	91.73	92.34	93.80	95.81
	Micro-F1	86.26	92.80	93.09	92.55	93.27	93.99	96.00
IMDB	Macro-F1	50.35	45.15	51.81	52.99	47.64	55.54	56.76
	Micro-F1	54.33	48.81	54.61	56.97	50.99	57.60	59.20
ACM	Macro-F1	57.38	53.99	61.89	66.39	73.92	77.32	84.83
	Micro-F1	57.57	54.31	61.51	71.57	73.84	78.23	85.77

在表4的测试结果中,将模型在某个数据集上的最好结果进行了加粗表示。观察表4可知,在DBLP数据上多数模型对大型复杂网络都有较好的性能,而本研究内容中模型在几种基准实验中是最优说明该模型可以处理较为复杂的异质信息网络。面对IMDB和ACM这样的小规模数据,部分模型并不具备较好的学习效果,而MAFGNN模型在这两个数据集上训练的准确率在60附近和80附近,与基准实验相比也具有最好性能,说明该模型在对小规模异质信息网络也具有较好的性能。对表4中3个数据集,本文的MAFGNN模型都具有更好的表现,显示出在基于特定的元路径下,对目标节点属性融合的方法,对异质信息表示学习有更好地性能。MAFGNN与HAN对比说明将目标节点邻居信息引入训练过程要比直接在元路径上进行节点级嵌入更加有效。MAFGNN与HERec相比说明采用特定元路径的方式固然重要,但在异质信息网络中融合目标节点和邻居节点信息有着更加重要的作用。MAFGNN与传统的同质图神经网络模型GCN和GAT模型相比说明注意力机制对基于不同原路径的邻居的权重学习要优于对所有邻居的权重学习。

4.5 节点聚类实验

使用Kmeans算法在DBLP、IMDB和ACM数据集上进行了节点聚类评估。与节点分类任务相似,将算法得到的节点嵌入到的目标节点预测类型作为KMeans的输入,将Kmeans的集群数 K 设置为4,即节点实际类别数量,然后通过计算标准化互信息(NMI)和调兰德系数(ARI)来评估各模型聚类结果质量。其中NMI系数和ARI系数都表示越接近1节点聚合性能越高。为了消除算法聚合性能受KMeans质心的影响,将本实验重复10次取均值,表5为各个模型的节点聚类实验结果。

表5 节点聚类实验结果
Table 5 Node clustering experiment results

Dataset	Metrics	DeepWalk	metapath2vec	GCN	GAT	HERec	HAN	MAFGNN
DBLP	NMI	76.53	74.30	75.01	71.50	76.73	81.98	88.21
	ARI	81.35	78.50	80.49	77.26	80.98	87.37	92.41
IMDB	NMI	1.45	1.20	5.45	8.45	5.45	20.94	25.78
	ARI	2.14	1.70	4.40	7.46	4.40	23.70	29.52
ACM	NMI	41.61	21.22	40.44	56.26	40.70	59.17	68.61
	ARI	35.10	21.00	29.59	53.69	37.13	59.48	71.39

在表5中,模型在某个数据集上的最好结果在表内采用了加粗表示。从表中的3个数据集和7个模型的对比实验中可以看出,本文的MAFGNN模型始终比基准实验要好,由此可知大多数图神经网络在对节点嵌入上有着更优的性能。其中基于图神经网络的模型(GCN、GAT、HERec、HAN和MAFGNN)要比基于随机游走的方法有更好的性能,基于异质图神经网络的模型(HERec、HAN和MAFGNN)要比传统的图神经网络(GAN和GAT)性能更加优秀。最后,MAFGNN模型在3个数据集上的NMI和ARI两种参数上都要比这些基准实验性能更好,说明了MAFGNN模型可以更好地学习异质信息网络中的语义信息和节点嵌入。

4.6 可视化实验

为了进行更加直观的比较,继续进行节点嵌入向量的可视化实验,对比MAFGNN模型在其他基准实验中的优越性。在可视化实验的参数设置中按照对比实验设置,使用2-SNE模型进行降维,然后在DBLP数据集进行测试。经过对比GAT、HAN和MAFGNN模型的可视化结果明显可以看出,MAFGNN的边界更为明显,表示MAFGNN模型对于节点嵌入有更好的学习效果。

通过观察图5(a)和(b)可以看出,基于元路径对目标节点的聚合过程对节点聚类有更好的结果。在图5(b)和(c)的对比实验结果可以分析出,通过在特定元路径下对目标节点的邻居节点进行属性聚合会使模

型在节点聚类上有更好的性能。GAT 作为传统节点嵌入模型通过聚合邻居信息得到节点间权重系数,不能有效的处理 DBLP 异质信息数据,也不能学习到 DBLP 数据集中不同类型图节点间的语义信息,对节点嵌入还有一定的匮乏性。HAN 通过使用元路径得到 DBLP 数据中作者类型节点的同质图,并没有获取到原异质图中的邻居信息,图 5(b) 也显示 HAN 模型只是对 DBLP 中作者节点进行了粗略的划分。相反,图 5(c) MAFGNN 模型在对作者节点的划分效果更加明显,表明在元路径下聚合邻居信息会有更好的嵌入效果。

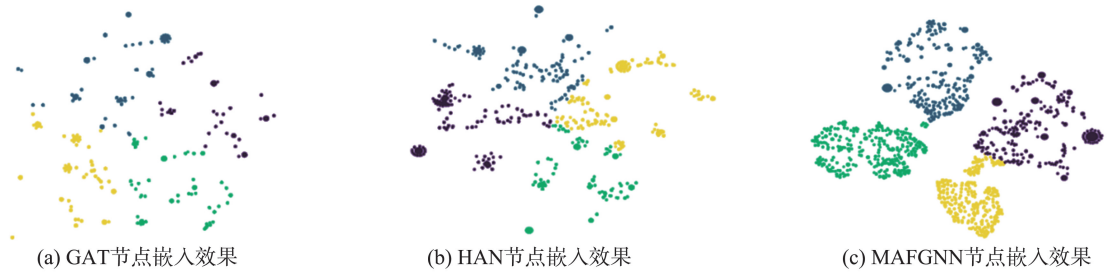


图 5 节点嵌入实验结果
Fig.5 Node embedding experiment results

4.7 参数分析实验

在 MAFGNN 模型参数分析实验中,本文分别研究了训练轮数 epoch、注意力头数 heads 和语义级注意力向量参数对 MAFGNN 模型的影响。

(1) 首先研究训练 epoch 参数对本模型的影响。在保证其他参数与数据集一致的情况下,通过增大 epoch 来观察 HAN 模型和 MAFGNN 模型的损失 Loss 和训练集的 F1 值,横坐标设置为训练轮数 epoch,纵坐标分别为模型损失值 Loss 和评价指标 F1 值,结果如图 6 和 7 所示。从图 6 可以看出,随着 epoch 增加, HAN 模型和 MAFGNN 模型在训练过程中的 Loss 会逐渐减少并趋于稳定;在 epoch 大于 10 之后,MAFGNN 的 Loss 几乎都比 HAN 模型更小,所以 MAFGNN 比 HAN 性能更好,收敛速度更快。从图 7(a) 和 (b) 可以看出,随着 epoch 的增加,MAFGNN 模型的 F1 大于 HAN 模型,并且 MAFGNN 模型的收敛速度更快。

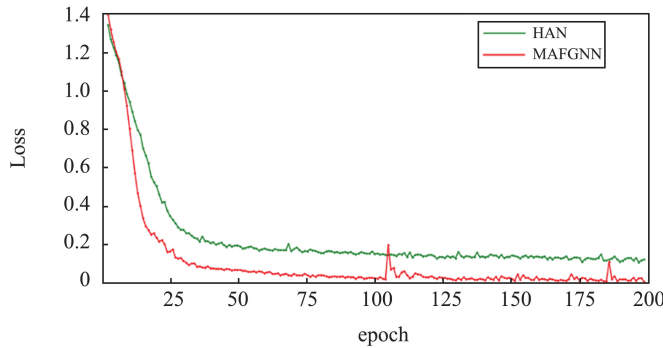
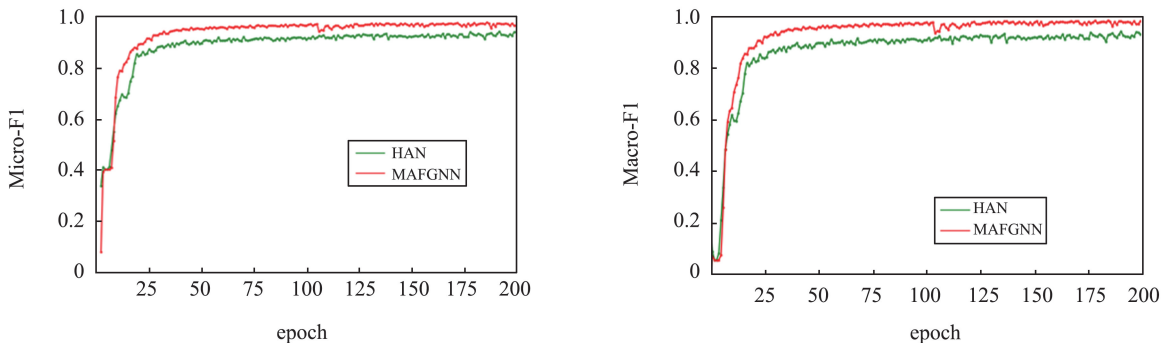


图 6 DBLP 数据集训练轮数和 Loss
Fig.6 Number of training rounds and Loss for DBLP dataset



(a) 分类器的准确性 (b) 分类器的召回率

图 7 DBLP 数据集训练轮数和 Micro-F1
Fig.7 DBLP dataset training rounds and Micro-F1

(2) 在训练头数 heads 参数实验中,通过改变多头注意力头数来观察对 MAFGNN 模型的聚类效果的影响,结果如图 8 所示。随着注意头数的增加,MAFGNN 模型在性能上有所增加。当注意力头数为 1 时,多头注意力被删除,使得目标节点过于集中自身的注意力,让模型缺乏了对邻居节点信息的聚合,效果不太明显。随着注意力头数的增加,MAFGNN 模型性能变化明显并且更为稳定。最后通过研究,当使用注意力头数为 10 或者以上,NMI 系数并没有明显的提高同时还会导致准确率降低,当注意力头数大于 12 以上时,对模型的训练在多数情况下出现内存不足的情况,根据 Bolya 等^[27]对多头注意力的研究,注意力头数超过一定范围会出现准确率崩溃和内存不足的情况,所以在该研究中对注意力头数太大的情况不予考虑。

(3) 在语义级注意力向量嵌入实验的设计中,增加了多层非线性层来更加全面地学习语义级嵌入。不同于以往的非线性层的设计,该实验增设了 3 层非线性层,第 1 层和第 2 层之间通过 tanh 激活函数来连接,第 2,3 层之间通过 LeakyReLU 激活函数来连接。使用不同的激活函数可以提高模型的平滑程度,同样解决了模型中梯度消失和爆炸等计算量大的问题,同时增设 3 层非线性层增加了本模型的复杂程度,也减缓了参数收敛速度,使得模型更加的稳定。在图 9 中,通过观察增加语义级向量嵌入维度对 MAFGNN 模型的准确率的影响,可以发现 MAFGNN 模型的准确率随着语义层向量维度的增加而增加,当语义级注意力向量维度设置为 32 时表现最好,在之后 MAFGNN 的性能开始降低,随着语义层向量维度增加可能会导致模型出现过拟合现象从而失去了较好的性能。

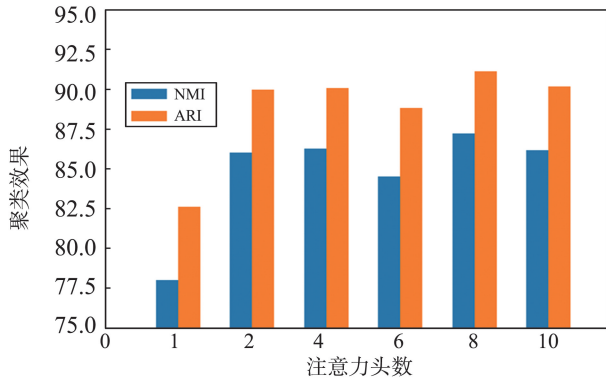


图 8 注意力头数实验

Fig.8 Attention headcount experiment

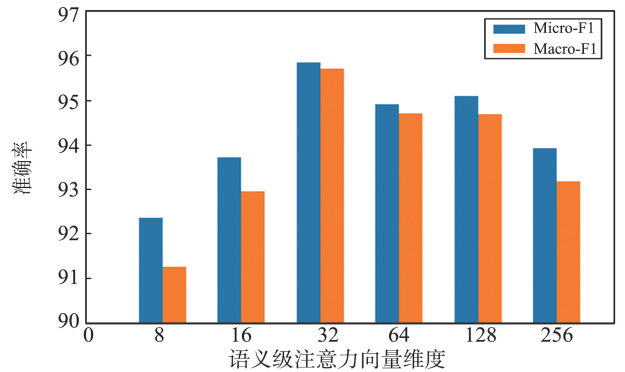


图 9 语义级注意力向量维度实验图

Fig.9 Semantic level attention vector dimension experiment

5 总结

本文提出的 MAFGNN 异质网络嵌入算法,设计了基于元路径对目标节点进行属性融合的方式来获取不同元路径间的语义信息,并采用自注意力机制来计算目标节点和邻居间的权重系数,挖掘节点之间的关系信息。最后在 3 个公开的数据集上进行实验最后得到了较好的效果。在未来的工作中,可以借鉴本文的研究工作的思路,尝试引入新的算法来融合元路径和目标节点的属性,挖掘节点之间的关系信息,还可以用该模型解决下游任务如节点分类和社区划分等。

参考文献:

- [1] ATWOOD J, TOWSLEY D. Diffusion-convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2016, 29:2001-2009.
- [2] SHI Chuan, LI Yitong, ZHANG Jiawei, et al. A survey of heterogeneous information network analysis[J]. IEEE Transactions on Knowledge and Data Engineering, 2016, 29(1):17-37.
- [3] SUN Yizhou, HAN Jiawei. Mining heterogeneous information networks: a structural analysis approach[J]. Association for Computing Machinery, 2013, 14(2):20-28.
- [4] CUI Peng, WANG Xiao, PEI Jian, et al. A survey on network embedding[J]. IEEE Transactions on Knowledge and Data Engineering, 2018, 31(5):833-852.
- [5] CAO B, LIU N N, YANG Q. Transfer learning for collective link prediction in multiple heterogenous domains[C]//WROBEL S. Proceedings of the 27th International Conference on Machine Learning. Haifa, Israel: DAUMÉ III H, 2010:159-166.
- [6] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space[EB/OL]. (2013-

- 09-07)[2023-09-23]. <https://arxiv.org/abs/1301.3781>.
- [7] PEROZZI B, AL-RFOU R, SKIENA S. Deepwalk; online learning of social representations[C]//MACSKASSY S. Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York; Association for Computing Machinery, 2014:701-710.
- [8] DONG Y, CHAWLA N V, SWAMI A. Metapath2vec; scalable representation learning for heterogeneous networks[C]//MATWIN S. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Halifax, Canada; Association for Computing Machinery, 2017:135-144.
- [9] SUN Yizhou, HAN Jiawei, YAN Xifeng, et al. Pathsim; meta path-based top-k similarity search in heterogeneous information networks[J]. Proceedings of the VLDB Endowment, 2011, 4(11):992-1003.
- [10] LEE S, PARK C, YU H. BHIN2vec; balancing the type of relation in heterogeneous information network[C]//ZHU Wenwu. Proceedings of the 28th ACM International Conference on Information and Knowledge Management. New York; Association for Computing Machinery, 2019:619-628.
- [11] WANG X, JI H, SHI C, et al. Heterogeneous graph attention network[C]//LING L. The World Wide Web Conference. New York; Association for Computing Machinery, 2019:2022-2032.
- [12] FU Xinyu, ZHANG Jiani, MENG Ziqiao, et al. MAGNN; metapath aggregated graph neural network for heterogeneous graph embedding[C]//HUANG Y N. Proceedings of The Web Conference 2020. Taipei; Association for Computing Machinery, 2020:2331-2341.
- [13] WANG Xiao, LU Yuanfu, SHI Chuan, et al. Dynamic heterogeneous information network embedding with meta-path based proximity[J]. IEEE Transactions on Knowledge and Data Engineering, 2020, 34(3):1117-1132.
- [14] XUE Hansheng, YANG Luwei, JIANG Wen, et al. Modeling dynamic heterogeneous network for link prediction using hierarchical attention with temporal RNN [C]//BIE T D. Machine Learning and Knowledge Discovery in Databases. Ghent, Belgium; Springer, 2020:282-298.
- [15] MIKOLOV T, SUTSKEVER I, CHEN K, et al. Distributed representations of words and phrases and their compositionality [J]. Advances in Neural Information Processing Systems, 2013, 26:3111-3119.
- [16] TANG Jian, QU Meng, WANG Mingzhe, et al. LINE; large-scale information network embedding[C]//GANGEMI A. Proceedings of the 24th International Conference on World Wide Web. Florence, Italy; Association for Computing Machinery, 2015:1067-1077.
- [17] RIBEIRO L F R, SAVERESE P H P, FIGUEIREDO D R. Struc2vec; learning node representations from structural identity [C]//MATWIN S. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Halifax, Canada; Association for Computing Machinery, 2017:385-394.
- [18] WU Zonghan, PAN Shirui, CHEN Fengwen, et al. A comprehensive survey on graph neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(1):4-24.
- [19] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[EB/OL]. (2017-02-22)[2023-09-23]. <https://arxiv.org/abs/1609.02907>.
- [20] VELIČKOVIĆ P, CUCURULL G, CASANOVA A, et al. Graph attention networks[EB/OL]. (2018-02-04)[2023-09-23]. <https://arxiv.org/abs/1710.10903>.
- [21] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30:6000-6010.
- [22] FU T, LEE W C, LEI Z. HIN2Vec; explore meta-paths in heterogeneous information networks for representation learning [C]//LIM E P. Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. Singapore; Association for Computing Machinery, 2017:1797-1806.
- [23] SHANG Jingbo, QU Meng, LIU Jialu, et al. Meta-path guided embedding for similarity search in large-scale heterogeneous information networks[EB/OL]. (2016-10-31)[2023-09-23]. <https://arxiv.org/abs/1610.09769>.
- [24] GUAN Mengya, CAI Xinjun, SHANG Jiaying, et al. HMSG; heterogeneous graph neural network based on metapath sub-graph learning[EB/OL]. (2021-09-07)[2023-09-23]. <https://arxiv.org/abs/2109.02868>.
- [25] ZHOU Sheng, BU Jiajun, WANG Xin, et al. HAHE; hierarchical attentive heterogeneous information network embedding [EB/OL]. (2019-05-14)[2023-09-23]. <https://arxiv.org/abs/1902.01475>.
- [26] ZHANG Chuxu, SONG Dongjin, HUANG Chao, et al. Heterogeneous graph neural network[C]//TEREDESAL A. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Anchorage, USA; Association for Computing Machinery, 2019:793-803.
- [27] BOLYA D, FU C Y, DAI X L, et al. Hydra attention; efficient attention with many heads[EB/OL]. (2023-02-12)[2023-09-23]. <https://arxiv.org/abs/2209.07484>.