

基于折叠路径聚合的属性网络节点嵌入方法

白明昌

(甘肃农业大学 信息科学技术学院, 兰州 730070)

摘要: 属性网络嵌入是图分析领域具有挑战性的任务之一,旨在从网络的拓扑结构和节点特征中学习节点的低维向量表示,同时最大限度地保持其结构和固有特性,然而现有方法多数仅研究网络间的基本关系,未考虑节点邻居的相对重要性。基于此,提出一种基于折叠路径聚合的属性网络节点嵌入方法,有效挖掘属性网络中的复合关系并充分度量节点邻居的重要程度。基于拓扑结构捕获节点的直接邻居并构建结构-属性二部图,挖掘“节点-属性-节点”折叠路径中所蕴含的复合关系,捕获节点的间接邻居。设计语义路径内部聚合策略,通过卷积神经网络聚合器聚合间接邻居表示和直接邻居表示,同时融合节点属性以捕获不同语义之间细粒度的特征交互,并根据语义路径间聚合策略整合2种细粒度嵌入表示,得到最终的节点嵌入。在Flickr、ArXiv和Pubmed这3个真实数据集上的实验结果表明,该方法的节点分类性能优于先进的属性网络嵌入方法,且与Node2Vec方法相比,Macro-F1和Micro-F1值分别高出0.067~0.234。

关键词: 折叠路径;注意力机制;卷积神经网络;属性图;节点嵌入

开放科学(资源服务)标志码(OSID):



源代码链接: <https://github.com/baimingch/NEFA>

中文引用格式:白明昌.基于折叠路径聚合的属性网络节点嵌入方法[J].计算机工程,2023,49(7):76-84.

英文引用格式:BAI M C. Node embedding method based on folded path aggregation on attributed network [J]. Computer Engineering, 2023, 49(7): 76-84.

Node Embedding Method Based on Folded Path Aggregation on Attributed Network

BAI Mingchang

(College of Information Science and Technology, Gansu Agricultural University, Lanzhou 730070, China)

[Abstract] Attributed network embedding is a challenging task in the field of graph analysis. It aims to learn the low-dimensional vector representation of nodes from the network topology and node attributes of the network while maintaining its structure and inherent characteristics to the greatest extent possible. However, most of the existing methods often only consider the basic relationship between networks and neglect the relative importance of node neighbors. To address this shortcoming, this paper proposes an attributed network node embedding method based on folding path aggregation, which aims to effectively mine the composite relationships in the attributed network and fully measure the importance of node neighbors. First, the direct neighbors of nodes are captured based on the topology, and the structure-attribute bipartite graph is constructed. Furthermore, the compound relationship present in the node-attribute-node folding path is mined to capture the indirect neighbors of the node. Second, an internal aggregation strategy for semantic paths is designed. The indirect- and direct-neighbor representations are aggregated using a convolution neural network aggregator, and the node attributes are fused to capture the fine-grained feature interaction between different semantics. The final embedding can be achieved by integrating two fine-grained embedding representations through the aggregation between semantic paths. Experimental results on three real-world datasets, Flickr, ArXiv and Pubmed, show that the node classification performance of the proposed method is better than that of the advanced attribute network embedding method, and the values of Macro-F1 and Micro-F1 are 0.067-0.234 higher than Node2Vec, respectively.

[Key words] folded path; attention mechanism; convolutional neural network; attributed graph; node embedding

DOI: 10.19678/j.issn.1000-3428.0064318

基金项目:教育部产学研合作协同育人新工科建设项目(201801224003);甘肃省高等教育教学成果培育项目。

作者简介:白明昌(1975—),男,副教授、博士,主研方向为人工智能、数据挖掘。

收稿日期:2022-03-28 修回日期:2022-08-30 E-mail:1164731416@qq.com

0 概述

图(网络)常被用来建模复杂对象间的关系,其中节点表示复杂系统,而边用来刻画节点间的关联。随着现实世界信息的多元化,节点往往附着有描述其特征信息,即节点属性。例如,在学术引文网络中,节点代表论文,节点间的引用关系为其间的边。此外,节点(论文)附有标题、关键词、所属的期刊、会议或者研究主题等描述其特定特征的属性信息。在社交网络中,除社交用户(节点)及其之间密切的交流(边)外,用户的性别、年龄、居住城市等个人资料作为刻画节点特征的属性信息存在。不同于普通网络(即仅具有拓扑连接信息的网络),属性网络具有结构复杂、信息丰富等特点^[1],其与结构互补的丰富属性信息可显著地提升网络分析任务的性能^[2-4]。此外,有关社会科学的一些相关研究还表明,网络的形成不仅受节点间复杂交互关系的影响,还受到节点属性的影响,共享属性的节点往往具有更内聚的连接^[5-7]。因而面向属性网络的研究,如何利用这些信息来分析网络的统计特性、潜在特征和交互规律以及揭示潜在的模式受到越来越多研究人员的关注^[8-9]。

在众多网络分析任务中,网络嵌入作为一种可以最大限度地保持网络的结构和固有特性的有效工具而被广泛研究^[10-13],其旨在融合网络拓扑结构和节点属性来学习低维节点表示,学习到的节点表示可以用于各种下游任务,如链路预测^[14]、节点分类^[15]和聚类^[16]等。早期的网络嵌入方法多面向拓拓扑网络^[17-19](也称为简单网络,即网络仅有节点和边),经典的方法包括:DeepWalk^[17]基于随机游走生成了节点序列,随后输入到Word2Vec^[18]模型中来获得节点表示,针对大规模网络,TANG等^[19]提出了LINE算法,通过定义一阶相似度和二阶相似度得到节点表示,文献[20]对DeepWalk进行改进,提出一种Node2Vec方法,其不仅改进了随机游走策略,而且利用深度优先搜索和广度优先搜索对深层次的网络结构进行探索。但是,上述代表性方法仅基于网络拓扑结构信息获得节点表示,忽略了网络中蕴含的节点属性,这使得学习到的节点表示不佳。为了充分利用节点的属性信息,研究人员提出许多面向属性网络的嵌入模型,其核心思想是在保持固有拓扑结构的同时引入属性对于节点的补充信息,进而学习到融合结构和属性节点更精确的低维潜在表示。代表性的工作包括:ANECP^[21]将拓扑结构、节点属性以及社区结构结合在一起,使用深度神经网络捕获拓扑结构和属性信息中潜在的高度非线性关系,社区结构的引入从微观角度解决了数据稀疏的问题;MAN^[22]从多个角度利用节点属性和节点之间的潜在邻近依赖性提出了多邻近注意力网络。然而,现有的属性网络嵌入方法只考虑“节点-节点”和“节点-属性”这样的基本关系,即节点之间的交互和节点拥有某些属性,而忽略了“节点-属性-节点”的复合关系。复合关系揭示了属性网络中“节点拥有共同

属性”的潜在联系并以此捕获属性网络中的内在高阶关系。显然,复合关系比基本关系传递了更多的信息^[23-24]。

针对以上问题,本文提出一种基于折叠路径聚合的属性网络节点嵌入方法NEFA。通过构造结构-属性交互二部图捕获直接邻居和间接邻居,并设计语义路径内部聚合,在充分考虑节点属性的前提下,分别聚合直接邻居表示和间接邻居表示以获得节点在不同语义的表示,最终基于语义路径间聚合策略使用水平卷积和垂直卷积,融合直接邻居和间接邻居的表示得到最终嵌入。

1 相关定义

1.1 符号说明与问题定义

设属性网络为 $G=(V, E, F)$,其中,节点集表示为 $V=\{v_1, v_2, \dots, v_n\}$,边集表示为 $E \subseteq V \times V$,设图中节点所包含的属性表示为集合 $F=\{f_1, f_2, \dots, f_m\}$ 。描述网络中拓扑结构关系的邻接矩阵为 $A_{n \times n}$,若 $(v_i, v_j) \in E$,则 $A_{ij}=1$,否则 $A_{ij}=0$ 。节点-属性关系矩阵由 $Q_{n \times m}$ 表示。

定义1(属性网络嵌入) 给定 $G=(V, E, F)$,属性网络嵌入旨在把图 G 中的节点 $v_i \in V$ 映射到一个低维空间中,即学习得到一个映射 $\varphi: V \rightarrow \mathbb{R}^d$,其中 $d \ll n$, φ 保留并融合了网络结构和节点的属性信息,学习到的节点表示可以作为后续下游任务的输入。

定义2(属性网络中的复合关系) 给定 $G=(V, E, F)$ 、节点集 $V=\{v_1, v_2, \dots, v_n\}$ 、属性集合 $F=\{f_1, f_2, \dots, f_m\}$ 和节点-属性关系矩阵 Q ,若三元组 (v_i, f_k, v_j) 满足 $Q_{ik}=1 \wedge Q_{jk}=1$,则满足此约束的关系为复合关系。

本文的主要符号及含义如表1所示。

表1 符号及含义

Table 1 Notations and meanings

符号	含义
$G=(V, E, F)$	图 G 是由节点集合 V 、边集 E 和属性集 F 组成
$SAG=(V \cup F, ESAG)$	SAG表示结构-属性二部图, $V \cup F$ 代表节点集,ESAG代表边集
$h_{v_i}, h_{v_i}^s, \tilde{h}_{v_i}$	节点 v_i 的初始特征、投影特征向量、最终嵌入
A, q	语义路径内部、语义路径间聚合注意力向量
$\mathcal{N}_{v_i}^S, \mathcal{N}_{v_i}^F$	节点 v_i 基于拓扑结构、结构-属性二部图的邻居
$h_{v_i}^{SE}, h_{v_i}^{AE}$	节点 v_i 基于拓扑结构、结构-属性二部图的向量表示
$h_{v_i}^S, h_{v_i}^A$	节点 v_i 通过多头注意力扩展后的拓扑结构上、结构-属性二部图上的向量表示
$h_{P(v_i, v_j)}$	折叠路径 $P(v_i, v_j)$ 的向量表示
o_v, o_h	垂直卷积和水平卷积的结果

1.2 基于折叠路径聚合的属性网络节点嵌入方法基本框架

图1所示为基于折叠路径聚合的属性网络节点嵌入方法的总体框架。该框架以属性网络作为

输入,首先通过语义路径内部聚合得到节点在不同语义下的嵌入,然后利用语义路径间聚合得到节点的最终嵌入。图1中的圆形代表节点,条形代表节点携带的属性。以 v_3 为例,携带属性 f_2, f_4 ,那么节点 v_3 和属性 f_2, f_4 可以构建连边,由此构造结构-属性二部图。在拓扑结构上,节点 v_3 的直接邻居为 $\mathcal{N}_3^s = \{v_1, v_2, v_4\}$;在结构-属性二部图上,通过

折叠路径可以得到节点 v_3 的间接邻居 $\mathcal{N}_3^f = \{v_1, v_2, v_3, v_4, v_5\}$ 。通过设计语义路径内部聚合策略,首先根据语义级注意力机制将2种不同语义的表示进行拼接,之后执行水平卷积和垂直卷积操作得到不同维度细粒度的信息,最后通过神经网络融合不同维度细粒度的信息得到节点 v_3 的最终表示。

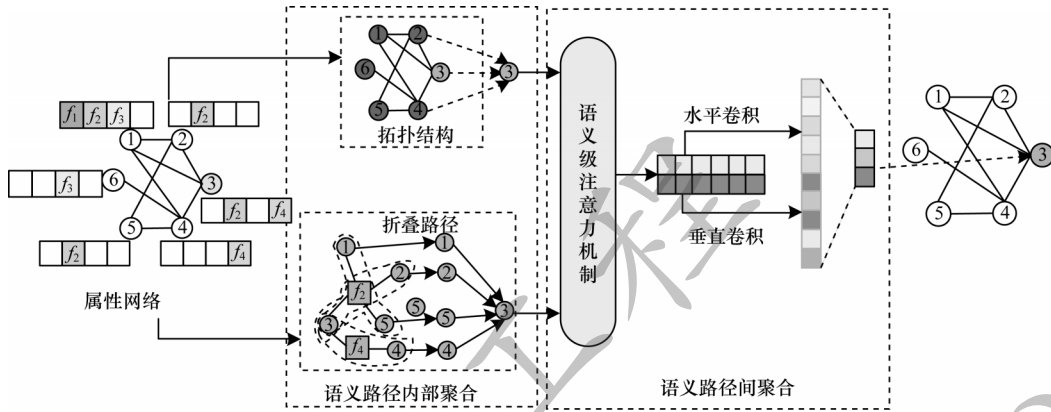


图1 基于折叠路径聚合的属性网络节点嵌入方法的总体框架

Fig.1 Overall framework of node embedding method based on folded path aggregation on attributed network

2 NEFA 方法

2.1 结构-属性二部图

在属性网络中,节点与其携带的属性之间可以很自然地建模为结构-属性二部图。通过结构-属性二部图可以很容易捕获一些基本关系,如某个节点拥有某些属性。此外,节点属性间的复杂交互信息更有助于节点的表示。属性(即节点-属性-节点)可作为桥梁来连接不同的节点,其可帮助算法更清晰地捕获节点与属性之间的复合关系。显然,这种复合关系包含了节点之间更多的信息。本文将“节点-属性-节点”定义为折叠路径。值得注意的是,在“节点-属性-节点”的复合关系中,起始节点与尾结点之间可能不是拓扑邻居。为了和拓扑结构中的直接邻居进行区分,本文将其称为通过折叠路径的间接邻居。具体定义如下:

定义3(结构-属性二部图) 设 $SAG=(V \cup F, E_{SAG})$,其中 $E_{SAG} \subseteq V \times F$ 。矩阵 $Q_{n \times m}$ 是该结构-属性二部图的邻接矩阵。对于 $\forall v_i \in V, \forall f_j \in F$,若节点 v_i 携带属性 f_j ,则 $Q_{ij}=1$,否则 $Q_{ij}=0$ 。

2.2 语义路径内部聚合

与拓扑网络中的节点表示不同,属性网络中的节点表示不仅包含节点在网络中的复杂拓扑邻域信息,而且具有本身所蕴含的属性信息。节点的局部邻域结构与节点所包含的属性信息存在显著关系。本节从拓扑结构和结构-属性二部图2种不同的语义路径出发,获得节点的直接邻居与间接邻居,以更精确地捕获属性网络中节点的局部邻域结构与节点属性之间的关系,最后通过聚合不同邻居得到不同语

义下的节点表示。由于基于结构-属性二部图的间接邻居聚合相对复杂,为了清晰起见,下文均以间接邻居为例进行说明。

为使节点嵌入更加准确,本文增强节点 v_i 初始特征表示,具体如下:

$$h'_{v_i} = W_1 h_{v_i} \quad (1)$$

其中: h_{v_i} 和 h'_{v_i} 分别表示节点 v_i 初始特征向量和变换后的向量表示; W_1 为可学习的变换权重矩阵。节点的初始特征可通过 W_1 进行变换得到比初始特征相对好的节点表示,同时转换节点维度以适应后期的学习。

与拓扑结构相比,结构-属性二部图的特殊之处在于结构-属性二部图中的边只存在于节点集和属性集之间,而节点集或属性集节点之间没有直接边连接,即在结构-属性二部图中,仅存在节点和属性之间的连接,同类节点间并没有连接。但值得注意的是,仅仅考虑节点和属性之间的直接连接不足以有效挖掘结构-属性二部图中所蕴含的丰富信息,而节点间的隐式关系(节点-属性-节点)有助于属性网络中的节点表示学习任务建模属性网络所蕴含的丰富且复杂的语义信息。本文将这种节点间的隐式关系定义为折叠路径。

在结构-属性二部图中,给定一条折叠路径 P ,语义路径内部聚合通过对 P 进行编码,学习嵌入在目标节点、基于折叠路径的间接邻居及其局部邻域中的结构和语义信息。设 $P(v_i, v_j)$ 是连接目标节点 v_i 和其基于折叠路径的间接邻居 $v_j \in \mathcal{N}_{v_i}^f = \{v_j | \forall f_k \in F, (v_i, f_k) \in E_{SAG} \wedge (v_j, f_k) \in E_{SAG}\}$ 的一条

折叠路径。为清晰起见,设 $M^{P(v_i, v_j)} = \{f_k | (v_i, f_k) \in E_{SAG} \wedge (v_j, f_k) \in E_{SAG}\}$ 为连接节点 v_i 和 v_j 之间的属性形成的集合。语义路径内部聚合使用特殊的路径编码器将沿折叠路径的所有节点特征转换为单个向量,这不仅充分利用了拥有相同属性的节点信息,而且也捕获到了中间属性节点这一桥梁的信息。具体地:

$$\mathbf{h}_{P(v_i, v_j)} = f_\theta(P(v_i, v_j)) = f_\theta\left(\mathbf{h}'_{v_i}, \mathbf{h}'_{v_j}, \left\{\mathbf{h}'_{f_k}, \forall f_k \in M^{P(v_i, v_j)}\right\}\right) \quad (2)$$

其中: $\mathbf{h}_{P(v_i, v_j)} \in \mathbb{R}^d$ 是折叠路径的向量表示。

本文给出了2个候选编码器:

1) 平均编码器。该函数将沿折叠路径的节点向量进行元素间乘积后再平均。

$$\mathbf{h}_{P(v_i, v_j)} = \text{mean}\left(\mathbf{h}'_{v_i}, \mathbf{h}'_{v_j}, \left\{\mathbf{h}'_{f_k}, \forall f_k \in M^{P(v_i, v_j)}\right\}\right) \quad (3)$$

2) 线性编码器。该函数是对平均编码器的扩展,它在线性编码器后面附加了线性变换:

$$\mathbf{h}_{P(v_i, v_j)} = \mathbf{W}_P \cdot \text{mean}\left(\mathbf{h}'_{v_i}, \mathbf{h}'_{v_j}, \left\{\mathbf{h}'_{f_k}, \forall f_k \in M^{P(v_i, v_j)}\right\}\right) \quad (4)$$

将折叠路径编码为向量后,本文采用图注意力层对与目标节点 v_i 相关的折叠路径加权求和,核心思想是不同的折叠路径将在不同程度上对目标节点的表示有所贡献。因此,可以通过学习与目标节点 v_i 有关的每条折叠路径的重要性权重 α_{ij} , 对折叠路径进行加权求和得到目标节点 v_i 在结构-属性二部图语义下的表示 $\mathbf{h}_{v_i}^{\text{AE}}$:

$$\mathbf{h}_{v_i}^{\text{AE}} = \sigma\left(\sum_{v_j \in \mathcal{N}_{v_i}^F} \alpha_{ij} \cdot \mathbf{h}_{P(v_i, v_j)}\right) \quad (5)$$

其中: σ 是 LeakyReLU 函数。

重要性权重 α_{ij} 的计算方式为:

$$\alpha_{ij} = \frac{\exp\left(\sigma\left(\mathbf{a}^T \cdot \left[\mathbf{h}'_{v_i} \parallel \mathbf{h}_{P(v_i, v_j)}\right]\right)\right)}{\sum_{v_j \in \mathcal{N}_{v_i}^F} \exp\left(\sigma\left(\mathbf{a}^T \cdot \left[\mathbf{h}'_{v_i} \parallel \mathbf{h}_{P(v_i, v_j)}\right]\right)\right)} \quad (6)$$

其中: $\mathcal{N}_{v_i}^F$ 是节点 v_i 在结构-属性二部图上基于折叠路径的间接邻居的集合; \mathbf{a} 是注意力向量; \parallel 表示向量拼接运算。值得注意的是,权重系数 α_{ij} 是不对称的,即 $\alpha_{ij} \neq \alpha_{ji}$, 这源于节点 v_i 和 v_j 具有不同的基于折叠路径的间接邻居,因此其彼此之间的贡献程度可能不同。

此外,可以将注意力机制扩展到多头,这有助于学习过程并使训练过程更加稳定,即执行 K 个独立的注意机制,然后将其输出进行拼接。

由上述内容可得出:

$$\mathbf{h}_{v_i}^A = \parallel_{k=1}^K \left(\sum_{v_j \in \mathcal{N}_{v_i}^F} [\alpha_{ij}]_k \cdot \mathbf{h}_{P(v_i, v_j)} \right) \quad (7)$$

至此,得到节点 v_i 在结构-属性二部图上的表示 $\mathbf{h}_{v_i}^A$ 。类似地,在拓扑结构上可以直接聚合其直接邻居,得到节点 v_i 在拓扑结构上的向量表示 $\mathbf{h}_{v_i}^S$ 。

2.3 语义路径间聚合

基于上述语义路径内部聚合,分别得到节点 v_i 在结构-属性二部图上的向量表示 $\mathbf{h}_{v_i}^A$ 以及在拓扑结构上的向量表示 $\mathbf{h}_{v_i}^S$, 这2种节点表示分别揭示了不同的语义信息。下一步需要融合这些不同语义的表示并更新节点信息。现有的图卷积模型通常利用基于池化函数来更新节点信息,忽略了不同语义信息之间至关重要的特征交互。本节基于CNN的信息更新方法来捕捉不同语义之间多粒度的特征交互。

受文献[24]的启发,本文认为拓扑结构和属性在属性图中并不是同等重要的。因此,本文采用语义级注意网络来更新特定语义下的权重。通过赋予拓扑结构和结构-属性二部图2种不同的语义路径的重要性权重来指导算法关注特定语义的信息。面向结构-属性二部图,对所有节点在该图上的向量表示进行转换取平均,进而使用 Softmax 来得到在结构-属性二部图这一语义下的重要程度:

$$e_A = \mathbf{q}^T \cdot \frac{1}{|V|} \sum_{v_i \in V} \tanh(\mathbf{W}_2 \cdot \mathbf{h}_{v_i}^A + \mathbf{b}) \quad (8)$$

其中: \mathbf{q}^T 是注意力向量; \mathbf{W}_2 是权重矩阵; \mathbf{b} 是偏置向量。同理,可以得到拓扑结构语义的重要程度 e_S 。归一化的注意力系数 β_s 可通过式(9)得到:

$$\beta_s = \frac{\exp(e_s)}{\exp(e_s) + \exp(e_A)} \quad (9)$$

同理可得 β_A 。如前所述,现有的基于图神经网络的模型基于池化来粗略地融合和更新信息,但没有考虑到不同语义信息之间的特征交互。鉴于这一问题,本文在拓扑结构和结构-属性二部图语义下得到节点表示连接成矩阵,即 $\mathbf{M}_{v_i} = [\beta_S \cdot \mathbf{h}_{v_i}^S] \parallel [\beta_A \cdot \mathbf{h}_{v_i}^A]$, 并基于卷积神经网络来捕获特征交互。

为了捕获不同语义信息间丰富的细粒度特征交互,本文设计2种不同的卷积核:第1种是 2×1 的垂直卷积核,对矩阵 \mathbf{M}_{v_i} 进行垂直卷积操作;第2种是 $2 \times d$ 的水平卷积核,对矩阵 \mathbf{M}_{v_i} 进行水平卷积操作。

具体来说,为了从嵌入维度的角度提取显著特征交互,本文使用垂直卷积来提取特征交互。垂直卷积覆盖本文涉及到的2种语义信息,并沿着嵌入维度方向滑动,表示为 $V' \in \mathbb{R}^{2 \times 1}$, $1 \leq t \leq m$ 。每个滤波器 V' 通过从左到右滑动与矩阵 \mathbf{M}_{v_i} 的每列相互作用,产生卷积结果 \mathbf{c}' , 如式(10)所示:

$$\mathbf{c}' = [c'_1, c'_2, \dots, c'_t, \dots, c'_d]$$

$$c'_t = \text{ReLU}\left(\mathbf{M}_{v_i}[:, t] \odot V'\right) \quad (10)$$

其中: \odot 表示卷积操作。通过用 m 个卷积核,对矩阵 \mathbf{M}_{v_i} 执行垂直卷积可以有效提取 m 个不同子空间中的特征交互。最后, m 个垂直卷积的结果被连接成一个向量 \mathbf{o}_i , 如式(11)所示:

$$\mathbf{o}_i = \mathbf{c}'^1 \parallel \mathbf{c}'^2 \parallel \dots \parallel \mathbf{c}'^m \quad (11)$$

此外,为了捕获不同语义之间的交互,通过水平卷积在矩阵 M_{v_i} 上沿着不同语义方向滑动,并与所有水平维度进行交互。具体来说,使用 n 个水平卷积核 $H^t \in \mathbb{R}^{2 \times d}, 1 \leq t \leq n$ 。类似地, n 个水平卷积核的结果被连接成向量 o_h , 如式(12)所示:

$$\begin{aligned} \tilde{c}' &= \text{ReLU}(M_{v_i} \odot H^t) \\ o_h &= \tilde{c}'^1 \parallel \tilde{c}'^2 \parallel \dots \parallel \tilde{c}'^n \end{aligned} \quad (12)$$

与垂直卷积不同,本文只有结构和属性2种不同的语义信息,每个水平卷积每次得到一个卷积结果 \tilde{c}' 。

垂直卷积和水平卷积捕获的特征交互信号通过全连接层以提取高级全局交互特征,如式(13)所示:

$$\tilde{h}_{v_i} = \text{ReLU}(W_f[o_h \parallel o_s] + b_f) \quad (13)$$

其中: W_f 和 b_f 是权重矩阵和偏置。

本文增加了一个全连接层以提高 NEFA 算法的性能。以节点分类任务为例,可采用交叉熵来进行模型的优化训练:

$$L = - \sum_{v_i \in V_L} Y_{v_i} \ln(C \cdot \tilde{h}_{v_i}) \quad (14)$$

其中: V_L 是拥有标签的节点集合; Y_{v_i} 为节点 v_i 的标签; \tilde{h}_{v_i} 为节点 v_i 最终表示; C 是分类器的参数。最后通过反向传播对模型进行优化,学习节点的向量表示。算法伪代码见算法1。

算法1 NEFA 前向传播

输入 属性网络 $G=(V, E, F)$, 节点初始特征 $h_{v_i}, \forall v_i \in V$, 多头注意力数 K

输出 最终节点表示 \tilde{h}_{v_i}

1. 构造结构-属性交互二部图 SAG
2. 特征转换: $h_{v_i}' \leftarrow W_1 \cdot h_{v_i}$
3. for $k=1, 2, \dots, K$ do:
4. for $v_i \in V$ do:
5. 分别找到节点 v_i 的直接邻居和间接邻居集合 $N_{v_i}^s, N_{v_i}^f$
6. 根据式(2)计算包含节点 v_i 的折叠路径的表示 $h_{p(v_i, v_j)}$
7. 根据式(6)、式(7)分别计算节点 v_i 在2种语义下的表示 $h_{v_i}^s$ 和 $h_{v_i}^f$
8. end for
9. end for
10. 根据式(8)、式(9)分别计算2种语义的重要程度得到矩阵 M_{v_i}
11. 根据式(11)~式(13)得到节点 v_i 的表示 \tilde{h}_{v_i}
12. 计算交叉熵损失: $L = - \sum_{v_i \in V_L} Y_{v_i} \ln(C \cdot \tilde{h}_{v_i})$
13. 反向传播和更新参数;
14. 返回: 最终节点表示 \tilde{h}_{v_i}

2.4 结构-属性二部图的合理性分析

属性网络中的表示学习方法可根据对于属性图的处理分为2类:1)将属性视为节点的辅助信息处理,这种方法大多基于深度学习的方法对节点和属

性统一建模,例如 DANE^[25] 基于随机游走来构建具有网络高阶相似性的特征矩阵,并基于深度编码器融合该特征矩阵与属性矩阵到统一框架中,学习其对应的表示向量;2)将属性视为另一类节点来处理,例如 Gat2Vec^[26] 构建了可表示节点与属性间关系的二分网络,并在原网络和此二分网络中进行随机游走得到2种序列,将这2种序列混合作为训练数据,并输入到 skip-gram 模型中,即可得到同时反映拓扑结构和节点属性的节点表示,这类方法不仅将属性作为节点的辅助信息考虑,而且也可整合节点和属性间交互信息。

属性网络嵌入主要的挑战是如何在网络表示空间中集成和平衡拓扑结构和属性信息。将节点和属性建模为二部图,首先可以将属性这种异质信息与节点联系起来,捕获节点和属性之间的交互。其次可以与拓扑结构互相补充共同促进节点表示的质量。同时,有助于建模拓扑结构和节点属性的相对重要性,以平衡两者在嵌入过程中的重要程度。相比之下,统一建模可能无法捕获节点和属性之间的交互,同时对平衡两者之间的信息有一定的限制。因此,本文通过拓扑结构捕获直接邻居,通过结构-属性二部图捕获融合属性的间接邻居。为了获得不同语义下的节点表示,采用注意力机制自适应地分配高权重给重要的邻居节点。此外,为了建模基于结构和基于属性的表示之间细粒度的联系,通过卷积操作,在不同特征维度对网络结构和节点属性的信息进行融合,获得最终节点的嵌入。

3 实验结果与分析

本节通过实验验证 NEFA 方法的有效性。基于本文方法特性,设置了4个验证实验,旨在解答以下4个问题:1)NEFA 在节点分类任务中的表现;2)NEFA 对参数的敏感程度;3)NEFA 的主要组成部分对其有何影响;4)NEFA 在不同语义下的性能。

3.1 实验数据集

在3个真实数据集 Flickr、ArXiv 和 Pubmed 上进行大量的实验来评估本文算法的有效性。表2给出了真实数据集的统计信息。

表2 数据集统计信息

Table 2 Data set statistics

数据集	节点数/个	边数/条	属性数/个	类别数/个
Flickr	7 575	239 738	12 047	9
ArXiv	856	2 660	30	19
Pubmed	19 717	44 338	500	3

由表2可知:

1)Flickr 是来自图片和视频托管网站的数据集,由7 575个用户(节点)和239 738条社交关系(边)组

成,节点具有12 047个描述其特定特征的属性信息。该数据集可根据用户喜欢的图片类型划分为9个类别。

2)ArXiv是一个引用网络,包含了856个节点和2 660条边,其中,节点代表论文,边代表引用关系。30个属性表示特定关键字在论文摘要中出现的频率。根据论文所属的研究领域共被划分为19个类别。

3)Pubmed是来自Pubmed数据库的关于糖尿病的引文网络,包含了3个类别共19 717篇论文以及44 338条边。从每篇论文的标题和摘要中通过删除停止词和低频率词后,生成500个单词作为属性,每篇论文的属性可以用一个500维的二进制向量表示。

3.2 对比方法及参数设置

为了评估本文所提方法的性能,分别与以下2类基准算法进行比较:第1类是基于网络拓扑结构的表示学习算法,即DeepWalk和Node2Vec;第2类是综合了网络拓扑结构和节点属性信息的表示学习算法,即ANECF和MAN。

DeepWalk^[17]:该方法通过对随机游走序列进行建模来学习图中节点的嵌入表示。

Node2Vec^[20]:该方法通过随机游走得到节点序列,随后采用skip-gram模型学习节点表示。

ANECF^[21]:该方法将拓扑结构、节点属性和社区信息的一阶和高阶接近性纳入嵌入过程,并建立三者与节点表示之间的一致和互补关系。

MAN^[22]:该方法从低阶近似和高阶近似的角度出发,利用节点属性和节点之间的潜在邻近依赖性得到节点表示。

为分析各组件对于NEFA的贡献,设置了以下变体方法:

NEFA w/o attribute:仅考虑基于折叠路径的间接邻居,不考虑路径中间的节点属性。

NEFA linear:使用线性路径编码器聚合折叠路径。

NEFA w/o attention:将语义路径内部聚合和语义路径间聚合时的重要性权重设置为相同。

NEFA-pooling:用常见的平均池化替代卷积神经网络。

NEFA w/o h-CNN:在卷积过程中去除水平卷积。

NEFA w/o v-CNN:在卷积过程中去除垂直卷积。

NEFA-S:仅利用直接邻居获得的节点表示。

NEFA-A:仅利用间接邻居获得的节点表示。

对于DeepWalk和Node2Vec,设置步长为80,窗口大小设置为10,步数为10。其他的基准方法参数设置均遵循原始论文,最后将节点表示的维数设置为64。

3.3 结果分析

3.3.1 节点分类任务

为了比较不同方法在节点分类任务上的性能,本文随机选择30%的节点作为训练集,其余节点作为测试集来训练SVM分类器。表3给出了不同方法重复10次的平均Macro-F1、Micro-F1,其中最优值以粗体突出显示。

表3 节点分类的性能对比

Table 3 Performance comparison of node classification

方法	Flickr		ArXiv		Pubmed	
	Macro-F1	Micro-F1	Macro-F1	Micro-F1	Macro-F1	Micro-F1
DeepWalk	0.615 5	0.638 1	0.466 4	0.539 7	0.787 0	0.773 3
Node2Vec	0.635 7	0.685 7	0.510 3	0.561 5	0.795 5	0.809 4
ANECF	0.735 1	0.798 3	0.592 8	0.648 2	0.823 7	0.828 3
MAN	0.821 7	0.842 6	0.671 2	0.731 1	0.856 5	0.859 1
NEFA	0.853 6	0.862 1	0.744 0	0.792 3	0.873 2	0.876 6

从表3可以看出,NEFA在3个数据集上的性能最佳。与只考虑网络拓扑结构的嵌入方法相比,Node2Vec比DeepWalk具有更好的性能,但由于本文方法融合了属性信息,这使得NEFA在3个数据集上的Macro-F1和Micro-F1值比Node2Vec高出0.067到0.234不等。这表明节点属性在网络嵌入学习中的重要性,属性信息可提升网络嵌入学习方法的性能。此外,本文方法NEFA在所有数据集上的性能均优于ANECF和MAN,说明捕获节点的重要性的“节点-属性-节点”折叠路径中的语义信息对网络嵌入是有益的,这进一步说明不同的节点在进行嵌入表示时具有不同的重要程度。此外,通过卷积神经网络捕获不同语义之间的细粒度交互对于网络嵌入也是至关重要的。

3.3.2 参数敏感性分析

本节从3个方面研究NEFA在数据集Flickr、ArXiv和Pubmed上对重要参数的敏感性以及分析不同参数对本文方法性能的影响:1)从节点向量嵌入维度的角度出发研究其对于NEFA的影响;2)通过分析语义路径间聚合注意力向量维度来观察其对于NEFA的影响;3)挖掘多头注意力机制数量的变化对于NEFA的性能影响。

1)节点向量嵌入维度

为验证嵌入向量的维度对本文方法的影响,通过在{16, 32, 64, 128, 256}的范围内改变 d ,在3个数据集上的性能比较如图2所示。从图2可以看出:模型的性能在ArXiv数据集上维度取64时性能表现最佳,然后维度的增加使得Micro-F1值降低,模型的性能不断下降;在Flickr和Pubmed数据集上,当最终嵌

入维度为128时,模型性能最好,后续同样随着维度的增加,模型性能在下降。原因在于NEFA需要一个合适的维度来编码节点信息,更大或者更小的维度可能会导致信息的不充分学习或者带来冗余,使得节点嵌入的效果不佳。

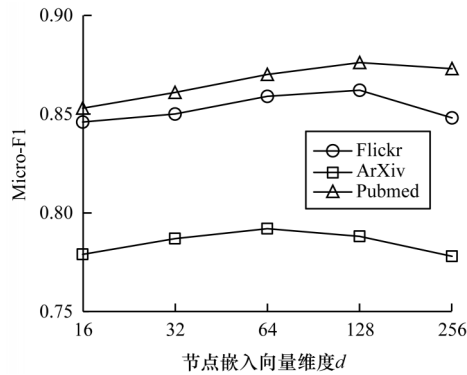


图2 节点嵌入向量维度在3个数据集上的影响
Fig.2 Effect of node embedding vector dimension on three data sets

2) 语义路径间聚合注意力向量维度

语义路径间聚合的性能受与其有关的注意向量 q 的维度影响。因此,本文探讨了注意向量 q 在不同维度下的实验结果,如图3所示。从图3可以看出,模型的性能随着注意向量 q 维度的增加而增长,当 q 的维度取128时模型性能达到最佳。之后,随着维度的增加模型性能开始下降,这可能是因为随着维度的增大,学习到的特征越来越稀疏,造成过度拟合。

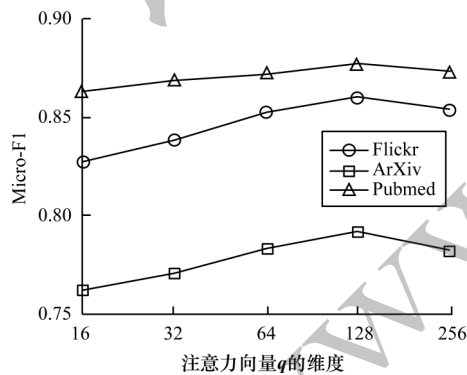


图3 注意力向量 q 在不同维度下的实验结果
Fig.3 Experimental results of attention vector q in different dimensions

3) 多头注意力机制数量

为测试多头注意力机制的效果,在3个数据集上通过设置不同的 K 值以确定多头注意力机制的数量对模型性能的影响。特别地,当 $K=1$ 时,模型退化为单头注意力机制。实验结果如图4所示, K 值的增加使得模型在3个数据集上的性能都得到了提升,NEFA算法在 K 取8时性能表现最佳。

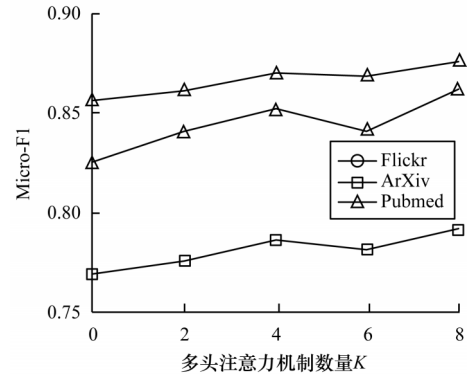


图4 不同 K 值对多头注意力数量的影响
Fig.4 Effect of different K values on multiple attention number

3.3.3 消融研究

为验证模型各部分的有效性,本文进一步对不同的NEFA变体进行了实验,结果如表4所示。其中,NEFA w/o attention仅考虑基于折叠路径的间接邻居,不考虑路径中间的节点属性,NEFA_{linear}为使用线性路径编码器,NEFA w/o attribute将语义路径内部聚合和语义路径间聚合时的重要性权重设置为相同。

表4 消融研究性能对比

Table 4 Performance comparison of ablation study

变体方法	Flickr		ArXiv		Pubmed	
	Macro-F1	Micro-F1	Macro-F1	Micro-F1	Macro-F1	Micro-F1
NEFA w/o attention	0.8367	0.8469	0.7702	0.7739	0.8513	0.8572
NEFA w/o attribute	0.8459	0.8532	0.7374	0.7848	0.8626	0.8650
NEFA _{linear}	0.8486	0.8569	0.7411	0.7861	0.8719	0.8724
NEFA _{pooling}	0.8474	0.8572	0.7404	0.7853	0.8701	0.8688
NEFA w/o h-CNN	0.8446	0.8541	0.7390	0.7832	0.8673	0.8665
NEFA w/o v-CNN	0.8432	0.8536	0.7381	0.7827	0.8695	0.8671
NEFA _{avg}	0.8536	0.8621	0.7440	0.7923	0.8732	0.8766

比较NEFA_{avg}和NEFA w/o attribute可以看到,聚合折叠路径中间属性相比直接聚合基于折叠路径的间接邻居能够提高模型性能,这证明了折叠路径中属性的有效性;NEFA_{avg}和NEFA_{linear}的结果表明,平均编码器确实有助于小幅提高NEFA的性能;从NEFA_{avg}和NEFA w/o attention的对比可以看出,在利用图神经网络的思想聚合节点邻居的过程中,考虑每个邻居节点的不同的重要性是很有必要的,这有助于提高模型的性能。

3.3.4 案例研究

为研究节点在不同语义表示的能力,本节通过案例来进一步阐述注意力机制在节点表示中的重要性,进而探索在不同语义下节点表示的性能。图5和图6给出了在ArXiv数据集上针对待表示节点邻居的注意力权重赋值以及在不同语义下节点的性能。图7给出了在NEFA算法中节点属性对于算法

成本的影响。

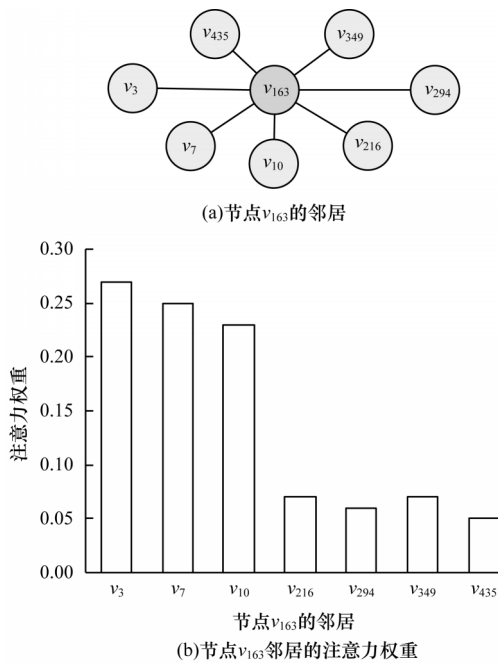


图 5 节点 v_{163} 的邻居及其对应的注意力权重
Fig.5 Neighbors of node v_{163} and their corresponding attention weights

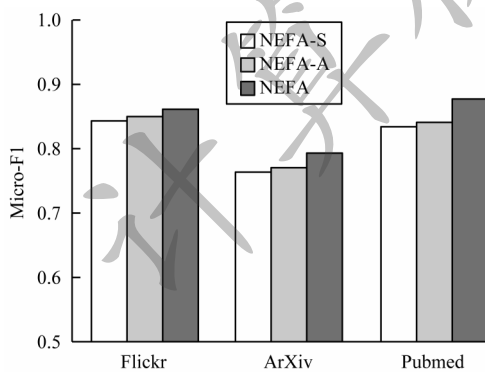


图 6 不同语义下节点表示的性能
Fig.6 Performance of node representation under different semantics

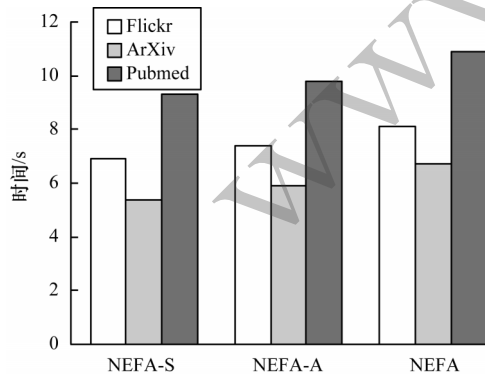


图 7 不同邻居聚合下的时间效率
Fig.7 Time efficiency under different neighbor aggregation

在 ArXiv 数据集中选择论文 ID 为 163 的节点 v_{163} , 在模型训练结束后, 再提取模型学习到的注意

力权重。图 5 给出了可视化结果, 由图 5 可知, 不同邻居对节点 v_{163} 的表示贡献不同, 这可能源于算法可捕获不同语义之间细粒度的特征交互, 通过特征交互和拓扑信息可精准度量邻居节点对于待表示节点的贡献。此外, 通过对 ArXiv 数据集的详细分析可知, 节点 v_3, v_7, v_{10} 与节点 v_{163} 具有相同的类别, 这进一步表明本文的注意力机制可以对目标节点有意义的一些邻居分配较高的权重, 从而提高节点表示的质量。

此外, 本节给出了 NEFA 在不同语义下的节点表示的性能。具体来说, 在本文中有 2 种语义关系, 分别为来自直接邻居和来自间接邻居的语义关系。为了验证不同语义下的节点表示能力, 这里设置 2 个变体: NEFA-S 表示只利用直接邻居获得的节点表示; NEFA-A 表示只利用间接邻居获得的节点表示。图 6 展示了在 3 个数据集上 2 种变体与本文方法的性能对比结果。从图 6 可以看出, NEFA 性能最好, NEFA-S 和 NEFA-A 相对较差, 这表明了同时融合 2 种不同的语义关系能学习到更好的节点表示。

此外, 图 7 分析了不同邻居聚合方法下节点属性对 NEFA 方法运算成本的影响。从图 7 可以看出, 节点属性的引入确实会致使时间成本的耗费增加, 但值得注意的是, 虽然时间成本上耗费有所增加, 但考虑属性可使得获得的节点表示更佳。

4 结束语

本文设计一种在属性网络中基于折叠路径聚合的属性网络节点嵌入方法。为充分利用属性网络的信息, 构建结构-属性二部图并与原始拓扑结构相结合, 挖掘基于拓扑结构的直接邻居和折叠路径的间接邻居, 不同语义的信息往往可较好地刻画出节点表示, 利用语义路径内部聚合得到节点在 2 种不同语义下的嵌入。最后根据语义路径间聚合得到不同语义的重要性权重并通过卷积神经网络生成节点的最终嵌入。在真实数据集上的实验结果表明, 本文方法具有较好的性能。由于本文方法未充分探索网络结构和节点属性的内在联系, 下一步将研究网络结构和节点属性间的深层次内在联系以减少运算成本。此外, 折叠路径仅考虑节点拥有共同属性, 并未建模共同属性的数量, 这也是未来需要研究的方向。

参考文献

[1] XU C, FENG S, ZUO Y. Relation-aware dynamic attributed graph attention network for stocks recommendation [J]. Pattern Recognition, 2022, 121: 108119.
[2] QIU R H, YIN H Z, HUANG Z, et al. GAG: global attributed graph neural network for streaming session-based recommendation[C]//Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, USA: ACM Press, 2020:

- 669-678.
- [3] LI X, WU Y, ESTER M, et al. SCHAIN-IRAM: an efficient and effective semi-supervised clustering algorithm for attributed heterogeneous information networks[J]. IEEE Transactions on Knowledge and Data Engineering, 2022, 34(4):1980-1992.
- [4] LI X, WU Y, ESTER M, et al. Semi-supervised clustering in attributed heterogeneous information networks [C]// Proceedings of the 26th International Conference on World Wide Web. New York, USA: ACM Press, 2017: 1621-1629.
- [5] HUANG Z, ZHONG X, WANG Q, et al. Detecting community in attributed networks by dynamically exploring node attributes and topological structure[J]. Knowledge-Based Systems, 2020, 196: 105760.
- [6] ZHUO S, BU J, ZHANG Z, et al. Cross multi-type objects clustering in attributed heterogeneous information network [J]. Knowledge-Based Systems, 2020, 194: 105458.
- [7] ISLAM M S, ALI M E, KANG Y B, et al. Keyword aware influential community search in large attributed graphs[J]. Information Systems, 2022, 104: 101914.
- [8] XIE X, SONG M, LIU C, et al. Effective influential community search on attributed graph[J]. Neurocomputing, 2021, 444: 111-125.
- [9] WANG X F, LI J H, YANG L, et al. Weakly-supervised learning for community detection based on graph convolution in attributed networks[J]. International Journal of Machine Learning and Cybernetics, 2021, 12(12): 3529-3539.
- [10] 刘知远, 孙茂松, 林衍凯, 等. 知识表示学习研究进展[J]. 计算机研究与发展, 2016, 53(2): 247-261.
- LIU Z Y, SUN M S, LIN Y K, et al. Knowledge representation learning: a review[J]. Journal of Computer Research and Development, 2016, 53(2): 247-261. (in Chinese)
- [11] 冶忠林, 赵海兴, 张科, 等. 基于邻节点和关系模型优化的网络表示学习[J]. 计算机研究与发展, 2019, 56(12): 2562-2577.
- YE Z L, ZHAO H X, ZHANG K, et al. Network representation learning using the optimizations of neighboring vertices and relation model[J]. Journal of Computer Research and Development, 2019, 56(12): 2562-2577. (in Chinese)
- [12] PAN Y, ZOU J, QIU J. Joint network embedding of network structure and node attributes via deep autoencoder [J]. Neurocomputing, 2022, 468: 198-210.
- [13] MENDONÇA M R F, BARRETO A M S, ZIVIANI A. Approximating network centrality measures using node embedding and machine learning[J]. IEEE Transactions on Network Science and Engineering, 2021, 8(1): 220-230.
- [14] GUPTA C, JAIN Y, DE A, et al. Integrating transductive and inductive embeddings improves link prediction accuracy[C]//Proceedings of the 30th ACM International Conference on Information & Knowledge Management. New York, USA: ACM Press, 2021: 3043-3047.
- [15] PARK H, NEVILLE J. Exploiting interaction links for node classification with deep graph neural networks [C]// Proceedings of the 28th International Joint Conference on Artificial Intelligence. New York, USA: ACM Press, 2019: 3223-3230.
- [16] GONG M G, CHEN C, XIE Y, et al. Community preserving network embedding based on memetic algorithm [C]// Proceedings of IEEE Transactions on Emerging Topics in Computational Intelligence. Washington D. C., USA: IEEE Press, 2018: 108-118.
- [17] PEROZZI B, AL-RFOU R, SKIENA S. DeepWalk: online learning of social representations[C]//Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, USA: ACM Press, 2014: 701-710.
- [18] RONG X. Word2Vec parameter learning explained [EB/OL]. [2023-02-20]. <https://arxiv.org/abs/1411.2738>.
- [19] TANG J, QU M, WANG M Z, et al. LINE: large-scale information network embedding [C]//Proceedings of the 24th International Conference on World Wide Web. New York, USA: ACM Press, 2015: 1067-1077.
- [20] GROVER A, LESKOVEC J. Node2Vec: scalable feature learning for networks [C]//Proceedings of International Conference on Knowledge Discovery & Data Mining. Washington D. C., USA: IEEE Press, 2016: 855-864.
- [21] HUANG T, ZHOU L H, WANG L Z, et al. Attributed network embedding with community preservation [C]// Proceedings of the 7th IEEE International Conference on Data Science and Advanced Analytics. Washington D. C., USA: IEEE Press, 2020: 334-343.
- [22] LIU X X, WANG K, WANG C D, et al. Attention-based multi-proximity preserved attributed network embedding [C]//Proceedings of International Joint Conference on Neural Networks. Washington D. C., USA: IEEE Press, 2021: 1-8.
- [23] 陈亦琦, 钱铁云, 李万理, 等. 基于复合关系图卷积的属性网络嵌入方法[J]. 计算机研究与发展, 2020, 57(8): 1674-1682.
- CHEN Y Q, QIAN T Y, LI W L, et al. Exploiting composite relation graph convolution for attributed network embedding [J]. Journal of Computer Research and Development, 2020, 57(8): 1674-1682. (in Chinese)
- [24] ZHANG D K, YIN J, ZHU X Q, et al. Attributed network embedding via subspace discovery [J]. Data Mining and Knowledge Discovery, 2019, 33(6): 1953-1980.
- [25] GAO H C, HUANG H. Deep attributed network embedding [C]//Proceedings of the 27th International Joint Conference on Artificial Intelligence. New York, USA: ACM Press, 2018: 3364-3370.
- [26] SHEIKH N, KEFATO Z, MONTRESOR A. Gat2Vec: representation learning for attributed graphs [J]. Computing, 2019, 101(3): 187-209.