



计算机科学

COMPUTER SCIENCE

基于对比学习的多关系属性图聚类方法

谢卓, 康乐, 周丽娟, 张志鸿

引用本文

谢卓, 康乐, 周丽娟, 张志鸿. 基于对比学习的多关系属性图聚类方法[J]. 计算机科学, 2023, 50(11): 62-70.

XIE Zhuo, KANG Le, ZHOU Lijuan, ZHANG Zhihong. Clustering Method Based on Contrastive Learning for Multi-relation Attribute Graph [J]. Computer Science, 2023, 50(11): 62-70.

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)

面向兴趣点推荐系统的自然噪声过滤算法

Natural Noise Filtering Algorithm for Point-of-Interest Recommender Systems

计算机科学, 2023, 50(11): 132-142. <https://doi.org/10.11896/jsjcx.230400045>

基于节点聚类复杂度的图聚类方法

Graph Clustering Algorithm Based on Node Clustering Complexity

计算机科学, 2023, 50(11): 77-87. <https://doi.org/10.11896/jsjcx.230600003>

融合无监督SimCSE的短文本聚类研究

Study on Short Text Clustering with Unsupervised SimCSE

计算机科学, 2023, 50(11): 71-76. <https://doi.org/10.11896/jsjcx.220900214>

基于谱聚类的边缘服务器放置算法

Edge Server Placement Algorithm Based on Spectral Clustering

计算机科学, 2023, 50(10): 248-257. <https://doi.org/10.11896/jsjcx.220900211>

基于图对比学习的多模态交通流量协同预测方法

Co-Forecasting for Multi-modal Traffic Flow Based on Graph Contrastive Learning

计算机科学, 2023, 50(10): 135-145. <https://doi.org/10.11896/jsjcx.230700127>

基于对比学习的多关系属性图聚类方法

谢卓¹ 康乐² 周丽娟¹ 张志鸿¹

¹ 郑州大学计算机与人工智能学院 郑州 450000

² 清华大学计算机科学与技术系 北京 100000

(zhuoxiez@163.com)

摘要 现实世界包含复杂的图数据,其节点之间通常包含多种关系,这种图被称为多关系属性图。图聚类是挖掘图数据相似信息的技术之一,然而现有的图聚类的方法大多只适用于单关系图。即使有的方法考虑到了多关系图,也往往是将图表示学习与聚类看作两个单独的过程。受 Deep Graph Infomax(DGI)算法的启发,文中设计了一种基于对比学习的多关系属性图的聚类方法(CCLMAG),用于解决上述问题:1)通过引入社区级互信息机制,弥补了 DGI 算法无法融合簇信息的缺点;2)引入嵌入融合模块来聚合不同关系上的节点嵌入;3)引入聚类优化模块将图表示学习与聚类两个过程联系起来,使得学习到的节点表示更适合聚类任务。在 3 个公开数据集和 1 个构建的期货数据集上的大量实验表明,所提方法优于目前最先进的基线方法,且具有实际应用价值。

关键词: 聚类;多关系属性图;图对比学习;图表示学习;无监督学习

中图法分类号 TP391

Clustering Method Based on Contrastive Learning for Multi-relation Attribute Graph

XIE Zhuo¹, KANG Le², ZHOU Lijuan¹ and ZHANG Zhihong¹

¹ School of Computer and Artificial Intelligence, Zhengzhou University, Zhengzhou 450000, China

² Department of Computer Science and Technology, Tsinghua University, Beijing 100000, China

Abstract In the real world, there are many complex graph data which includes multiple relations between nodes, namely multi-relation attribute graph. Graph clustering is one of the approaches for mining similar information from graph data. However, most existing graph clustering methods assume that only single type of relation exists between nodes. Even for those that considering the multi-relation of a graph, they use only node attributes for training, or regard graph representation learning and clustering as two completely independent processes. Recently, Deep Graph Infomax(DGI) has shown promising results on many downstream tasks. But there are two major limitations for DGI. Firstly, DGI does not fully explore the various relations among nodes. Secondly, DGI does not jointly optimize the graph representation learning and clustering tasks, resulting in suboptimal clustering results. To address the above-mentioned problems, this paper proposes a novel framework, called clustering method based on contrastive learning for multi-relation attribute graph(CCLMAG), for learning the node embedding suitable for clustering in a unsupervised way. To be more specific, 1) The community-level mutual information mechanism is applied to solve the problem of ignoring cluster information by DGI; 2) the Embedding Fusion Module is augmented to aggregate the embedding of nodes in different relationships; 3) the clustering optimization module is added to link the graph representation learning and clustering so that the learned node representation is more suitable for the clustering task, thus enhancing the interpretability of the clustering results. Extensive experimental results on three multi-relation attribute graph datasets and a real-world futures dataset demonstrate the superiority of CCLMAG compared with the state-of-the-art methods.

Keywords Clustering, Multi-relation attribute graph, Graph contrastive learning, Graph representation learning, Unsupervised learning

到稿日期:2022-09-17 返修日期:2022-12-05

基金项目:国家自然科学基金(62006211);河南省重大公益专项(201300210300)

This work was supported by the National Natural Science Foundation of China(62006211) and Major Public Welfare Projects in Henan Province(201300210300).

通信作者:周丽娟(ieljzhou@zzu.edu.cn)

1 引言

现实世界中存在着广泛的图结构数据,如社交网络、引文网络、蛋白质相互作用网络、期货交易网络等^[1]。近年来,对图数据的分析和挖掘正逐渐成为学术界和工业界一个积极研究的课题。在各种图数据挖掘技术中,图聚类被认为是图数据挖掘的重要任务之一,其目标是在图数据中挖掘出若干个互不相交的簇,以使得同一簇中的节点更有可能具有相似的特征。

传统的聚类方法^[2-4]由于只考虑了数据之间的浅层关系,忽略了图结构信息,因此无法有效地处理图数据的聚类问题。随着深度学习技术的迅速发展,利用深度学习强大的表示能力进行深度聚类的思想引起了广泛的关注。现有的图聚类方法大多只关注单个图,它们通常假设节点之间只存在一种类型的关系^[5-7]。然而,真实世界中的图数据往往要复杂得多,需要使用多种视图以更好地表示真实图数据,其中每一个视图都代表着节点之间的一种关系^[8]。例如,在期货交易场景中,以账户为图的节点,两个账户之间可能存在共用多种设备的情况,如共用 Android 设备、共用 iOS 设备、共用 Windows 设备和共用 Linux 设备等,每种共用设备都可以被看作两个节点之间的一种关系。虽然不同类型的关系可以独立形成不同的图,但是这些图之间并不是完全孤立的。此外,图中的节点可能会包含各种属性信息,这在许多实际应用中起着非常重要的作用。例如,在期货交易网络中,同属一个串通组的账户往往具有相似的交易行为,如果我们能够将账户的交易行为作为节点的属性,那么无疑可以提升聚类结果的准确性。这种节点具有属性特征且节点之间存在多种类型关系的复杂图被称为多关系属性图,它可以从更全面的角度对真实世界中的场景进行建模,在图数据挖掘中具有较大的研究价值。

近年来,虽然学术界已经进行一些关于多关系图聚类的研究,但是仍有一些问题需要进一步考虑。以往的一些方法只关注图的拓扑结构而忽略了节点的属性^[9];还有一些方法虽然考虑了节点的属性,却需要节点标签参与训练^[10-11],然而节点标签的获取通常是非常耗时或昂贵的;此外,许多工作将图表示学习和图聚类两个过程分步处理^[10,12],导致图聚类的结果完全依赖于图表示学习。

综上所述,有以下两个问题亟待解决。1)如何在无监督的前提下有效融合多种关系中的信息。单一关系的信息不足以获得良好的结果,不适当的信息融合也会影响模型的性能。2)如何联合优化图表示学习和聚类任务,换句话说,如何在图表示学习中更好地学习与聚类有关的信息。图表示学习和聚类通常是两个相互独立的任务,图表示学习的目的是重建原图,因此所学习到的节点嵌入并不一定包含足够的节点聚类所需要的信息。因此,我们需要学习到更适合节点聚类任务的节点嵌入。

针对上述问题,本文提出了一种无监督多关系属性图聚类的方法。该方法将 DGI^[6]算法扩展到多关系属性图中。为了在学习到的节点嵌入中保留簇信息,本文引入了社区级互信息机制;为了能够更好融合不同关系中的节点嵌入,引入基于注意力机制的嵌入融合模块,通过注意力使不同关系上的

节点表示信息进行交互,提高在多关系属性图上的表示能力;为了使学习到的节点嵌入更适合聚类任务以及增强聚类结果的可解释性,引入了一个聚类优化模块,使表示学习与聚类以端到端的方式同步学习。本文方法能够在统一的框架中训练和优化模型,从而使得不同的模块可以相互受益。最终将本文模型与一些最先进的基线任务在几个公开数据集和一个构建的期货数据集上进行对比,实验结果验证了本文模型的优越性。

2 相关工作

本章将介绍两个相关主题的一些最新发展,即图聚类和图对比学习。这些现有的方法都已取得良好的效果,对相关领域的发展起到了重要推动作用,对本文的工作也有着很强的借鉴意义。

2.1 图聚类

虽然传统的聚类方法如 K-Means^[2], Spectral^[3]已经取得良好的聚类结果,但是将它们应用复杂的图结构数据上时表现并不好。得益于强大的节点表示能力,近年来基于深度学习技术的聚类方法得到了广泛的应用。为了解决节点嵌入时忽略网络全局拓扑结构的高阶信息的问题, CANE^[13]提出在对抗学习框架下同时学习节点表示及识别社区,最终社区划分结果得到显著提升,然而该方法却忽略了节点的属性。Kang 等^[14]提出了一种融合变分图注意自编码器的深度聚类模型,可以有效聚合节点的属性和图的拓扑结构,同时能够优化聚类标签分配和学习适合于聚类的表示特征。Yang 等^[15]提出了一种基于节点相似性和图嵌入的复杂图聚类算法,首先通过图嵌入方法将节点映射为低维向量表示,然后计算节点相似性构建偏好图,进而获得初始质心,最后使用 K-Means 对低维向量进行聚类,以获得最终的结果。然而该方法也无法直接应用于多关系图,但是这两种方法都无法解决多关系属性图聚类的问题。DGI^[6]通过最大化节点嵌入和图级表示之间的互信息,提取出了有效的节点嵌入,但其同样无法直接应用于多关系属性图。DMGI^[12]作为一种无监督算法,弥补了 DGI 中不适用于多关系图的不足,应用共识正则化技术和注意力系数来整合最终的嵌入结果,但是模型训练时图表示学习和节点聚类被完全割裂。本文引入聚类优化模块,使学习到的节点嵌入更加适合于聚类方法,大大改善了聚类的结果。

2.2 图对比学习

虽然基于图的深度学习技术已经取得相当大的成就,但是,之前的大多数工作都是基于(半)监督学习,这导致模型具有标签依赖性强、泛化性及鲁棒性较差等缺点。近年来,作为一种自监督方法,对比学习引起了众多专家学者的兴趣,其在不使用标签的情况下依旧展现出了巨大的潜力。得益于强大的无监督学习的能力,对比学习已经在计算机视觉^[16]和自然语言处理^[17]等领域取得重大的成就。基于对比学习的图数据挖掘技术通过对比正负样本对来学习节点表示,其目标是将原始的图数据映射到一个对比表示空间中,在该空间中最大化正样本之间的相似度的同时最小化负样本之间的相似度。与传统的只关注正负样本的对比学习不同, Heco^[18]采用

交叉视图对比机制,从网络中提取两个视图分别学习节点嵌入,从而同时捕获局部和高阶结构,然后使用交叉视图对比学习和一种视图屏蔽机制,从两个视图中提取正样本和负样本,这使得两个视图能够协同监督彼此,并最终学习到足够好的节点嵌入,但是该方法也将图表示学习和聚类两个过程完全割裂。为了解决对比学习在随机负采样时会引入假阳性样本的问题,Zhao等^[19]提出了一种无偏图对比学习的方法,使用节点的类信息来避免引入假阳性样本,同时也能够将图嵌入和图聚类两个过程进行联合优化,但是该方法是专为单关系属性图设计的,无法较好地融合多关系图中的信息。因此,本文提出了一个嵌入融合模块,能够有效聚合多关系图中的信息。

3 预备知识

定义 1(多关系属性图) 一个多关系属性图可以表示为 $\mathcal{G} = \{\mathcal{G}^1, \mathcal{G}^2, \dots, \mathcal{G}^{|\mathcal{R}|}\} = \{V, E, \mathcal{R}\}$, 其中 $\mathcal{E} = \{e_1, e_2, \dots, e_N\}$, $\mathcal{E} = \{E_1, E_2, \dots, E_{|\mathcal{R}|} | E_r \subseteq V \times V, r \in \mathcal{R}\}$ 描述了节点之间的 $|\mathcal{R}|$ 种关系, $\mathbf{X} \in \mathbb{R}^{N \times d}$ 是 N 个节点的属性所组成的属性矩阵,对于关系 $r \in \mathcal{R}$, 有 $\mathcal{G}^r = \{V, E_r, \mathcal{R}\}$ 。同时,这 $|\mathcal{R}|$ 个边集也可以表示为邻接矩阵的集合 $\mathcal{A} = \{\mathcal{A}^{(1)}, \mathcal{A}^{(2)}, \dots, \mathcal{A}^{(|\mathcal{R}|)}\}$, 其中 $\mathcal{A}^{(r)} \in \{0, 1\}^{N \times N}$ 是图 \mathcal{G}^r 的邻接矩阵。注意,此处需满足 $|\mathcal{R}| > 1$ 。

定义 2(Deep Graph Infomax(DGI)) 文献[6]为学习节点表示提出了一种基于最大化互信息原理的自监督(无监督)学习方法,被称为 DGI。更具体地说,DGI 的目的是为每个节点 v_i 学习一个低维向量表示 $\mathbf{h}_i \in \mathbb{R}^d$, 使用函数 $Readout(\{h_1, h_2, \dots, h_N\})$ 得到图级表示 s 。最后,DGI 引入了一个判别器 \mathcal{D} 来区分正样本的嵌入 h_n 和负样本的嵌入 \tilde{h}_n , 可以使得 h_n 与图级表示 s 更加接近,而 \tilde{h}_n 与图级表示 s 更加远离,损失函数的表达式如下:

$$\mathcal{L} = \frac{1}{N+M} \left(\sum_{i=1}^N \mathbb{E}_{(X, A)} [\log \mathcal{D}(h_i, s)] + \sum_{j=1}^M \mathbb{E}_{(X, \tilde{A})} [\log(1 - \mathcal{D}(\tilde{h}_j, s))] \right) \quad (1)$$

尽管 DGI 很有效,但是它是为单关系属性图(Single Attributed Graph)设计的,无法直接应用于多关系图上。将 DGI 扩展到多关系图,只能通过将 DGI 单独应用于每个关系形成的单个图上,再计算每个图上的嵌入的平均值,并将其作为最终的节点表示。但是这种做法实际上假设不同关系是等价的,然而在多关系属性图中,各种关系对节点的影响往往是不同的。其次,DGI 只是一种通用的节点嵌入算法,并没有考虑到具体的下游任务。

4 研究方法

4.1 模型总览

如图 1 所示,本文提出了一种新的无监督多关系属性图聚类(Clustering Method Based on Contrastive Learning for Multi-relation Attribute Graph, CCLMAG)方法。该方法包含 3 个模块,分别是信息提取模块、嵌入融合模块和聚类优化模块。

1)信息提取模块。该模块被应用于多关系图的特定关系

上,分别提取出图级表示和社区级表示,计算节点与图级表示和社区级表示的互信息并使得互信息最大化,以此对节点的信息进行编码得到节点的嵌入,再将得到的特定关系的节点嵌入应用于嵌入融合模块。

2)嵌入融合模块。该模块的目的是将信息提取模块得到的不同关系的节点嵌入通过注意力机制进行融合,增强不同关系嵌入之间的交互,突出关键信息,最终聚合形成多关系图上的嵌入。

3)聚类优化模块。该模块利用对比学习的方法,通过最大化前 K 个最近邻节点的相似性以及最小化其他节点的相似性,来提高节点嵌入的质量,优化聚类结果,增强聚类结果的可解释性。

4.2 信息提取模块

受 DGI^[6]工作的启发,对于每一个关系 $r \in \mathcal{R}$, 我们引入一个特定关系类型的编码器 $\mathcal{E}_1^{(r)}$ 来生成特定关系的节点嵌入矩阵 $\mathbf{H}^{(r)}$ 。使用的编码器 $\mathcal{E}_1^{(r)}$ 是一个单层的 GCN, 表达式如下:

$$\mathbf{H}^{(r)} = \mathcal{E}_1^{(r)}(X, A^{(r)} | \mathbf{W}^{(r)}) = \sigma(\hat{D}_r^{-\frac{1}{2}} \hat{A}^{(r)} \hat{D}_r^{-\frac{1}{2}} X \mathbf{W}^{(r)}) \quad (2)$$

其中, $\hat{A}^{(r)} = A^{(r)} + I$, $\hat{D}_i = \sum_j A_{ij}$; $\mathbf{W}^{(r)} \in \mathbb{R}^{f \times d}$ 是可训练的权重矩阵; f 表示输入向量的维度; d 表示嵌入空间的维度; σ 表示非线性激活函数,如 RELU;使用向量 $\mathbf{s}^{(r)}$ 来表征图 $G^{(r)}$ 。我们使用均值函数来计算 $\mathbf{s}^{(r)}$, 其表达式如下:

$$\mathbf{s}^{(r)} = \sigma \left(\frac{1}{n} \sum_{i=1}^n h_i^{(r)} \right) \quad (3)$$

其中, $h_i^{(r)}$ 表示嵌入矩阵 $\mathbf{H}^{(r)}$ 的第 i 行,即节点 v_i 在当前关系下经编码器 $\mathcal{E}_1^{(r)}$ 后的嵌入结果; σ 表示非线性激活函数。

逐行打乱原始的属性矩阵,可以得到新的被扰乱的属性矩阵,即 $\mathcal{X} \rightarrow \tilde{\mathcal{X}}$, 将其作为负样本。再次使用前文所述的编码器 $\mathcal{E}_1^{(r)}$, 就能得到负样本的属性嵌入矩阵 $\tilde{\mathbf{H}}$ 。我们使用一个评价器 \mathcal{D} 来帮助最大化节点表示与图级表示之间的互信息, $\mathcal{D}(h_i^{(r)}, s^{(r)})$ 表示概率。我们的目标是使得原图的节点表示与图级表示计算得到的概率最大,新生成图的节点表示与图级表示计算得到的概率最小。其表达式如下:

$$\mathcal{L}_s^{(r)} = \frac{1}{N} \left(\sum_{i=1}^N \log \mathcal{D}(h_i^{(r)}, \mathbf{s}^{(r)}) + \sum_{j=1}^N \log(1 - \mathcal{D}(\tilde{h}_j^{(r)}, \mathbf{s}^{(r)})) \right) \quad (4)$$

其中, $\tilde{h}_j^{(r)}$ 表示矩阵 $\tilde{\mathbf{H}}$ 的第 j 行; $\mathcal{D}: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ 是一个评价器,用来计算两个向量的接近度。本文使用的是一个双线性评分函数,表达式如下:

$$\mathcal{D}(h_i^{(r)}, \mathbf{s}^{(r)}) = \sigma(h_i^{(r)\top} \mathbf{M}^{(r)} \mathbf{s}^{(r)}) \quad (5)$$

其中, σ 表示非线性激活函数, $\mathbf{M}^{(r)} \in \mathbb{R}^{d \times d}$ 是一个可训练矩阵。

考虑到节点与其自身所在的簇的互信息之间的影响,本文使用一个 $\mathcal{E}_2^{(r)}$ 编码器来获得节点 i 所在的簇在特定关系上的表示 $c_i^{(r)}$ 。所使用的 $\mathcal{E}_2^{(r)}$ 编码器是一个 GAT 编码器。不直接使用 $\mathcal{E}_1^{(r)}$ 编码器得到 $c_i^{(r)}$ 的原因是, $\mathbf{s}^{(r)}$ 向量本身就是 h_i 的一个线性组合,而 $\mathbf{s}^{(r)}$ 在节点与图级的互信息计算中已经使用过。类似于 $\mathcal{E}_1^{(r)}$ 编码器, $\mathcal{E}_2^{(r)}$ 的损失函数如下:

$$\mathcal{L}_c^{(r)} = \frac{1}{N} \left(\sum_{i=1}^N \log \mathcal{D}(h_i^{(r)}, c_i^{(r)}) + \sum_{j=1}^N \log(1 - \mathcal{D}(h_j^{(r)}, c_i^{(r)})) \right) \quad (6)$$

综合考虑各个关系上的损失函数,可以得到总的嵌入损失,表达式如下:

$$\mathcal{L}_s = \sum_{r \in \mathcal{R}} \mathcal{L}_s^{(r)} \quad (7)$$

$$\mathcal{L}_c = \sum_{r \in \mathcal{R}} \mathcal{L}_c^{(r)} \quad (8)$$

图1所示的模型共包含3个模块:信息提取模块、嵌入

融合模块和聚类优化模块。其中, $\mathbf{H}^{(r)}$ 是经过信息提取模块之后的第 r 个关系上的图的嵌入表示; α^r 是第 r 个关系的权重; H 是经过嵌入融合模块之后聚合的整个多关系属性图的嵌入表示; $Corrupt$ 表示随机打乱节点顺序,是一种数据增强的手段; $\tilde{\mathbf{H}}^{(r)}$ 是第 r 个关系上的图在数据增强之后的嵌入表示。

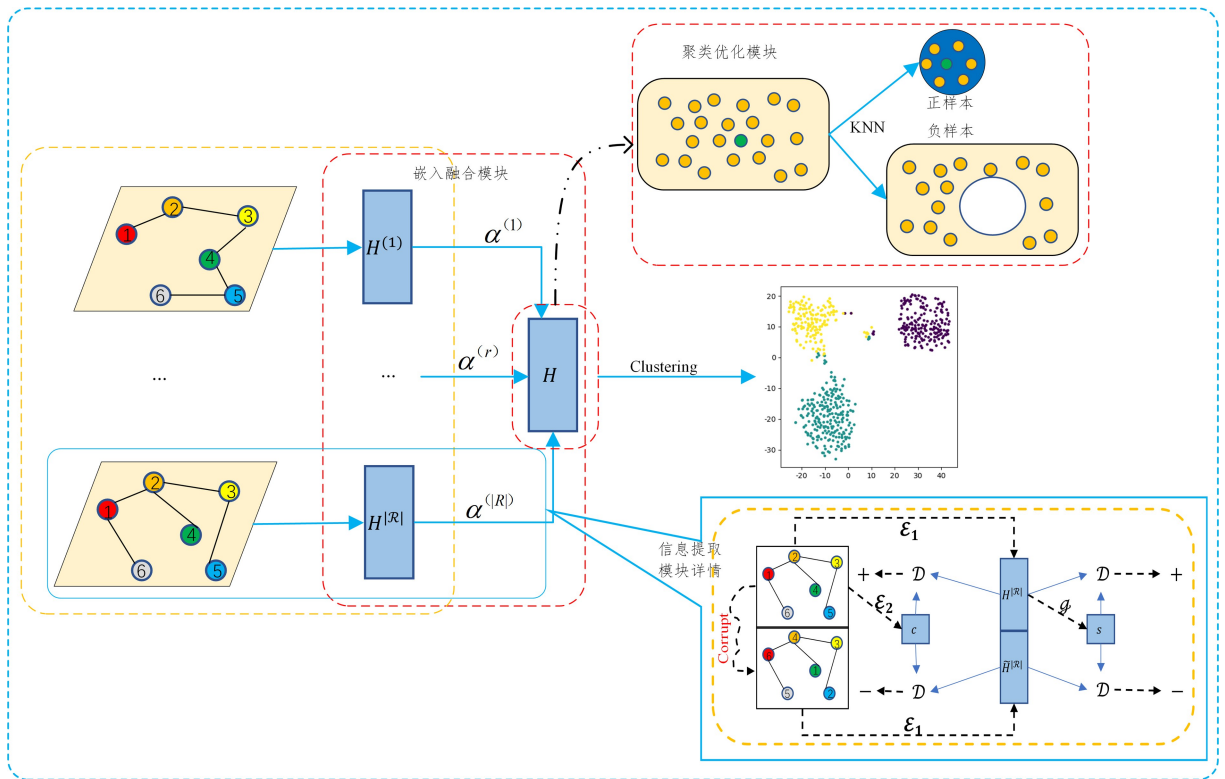


图1 CCLMAG 模型总览(电子版为彩图)

Fig.1 General framework of CCLMAG

4.3 嵌入融合模块

在多关系图中,针对某一个下游任务,节点之间的各种关系对该任务的影响是不同的^[8],因此本文提出了一个基于注意力机制^[20]的嵌入融合模块,用于聚合来自不同关系的节点嵌入。通过引入注意力机制来融合各个关系上的节点嵌入,使模型能够学习节点在不同关系中的重要性,通过不同关系嵌入之间的交互来反映其重要性,在最终节点嵌入中突出关键信息,提高节点嵌入的表示能力。节点嵌入的计算式如下:

$$h_i = \sum_{r \in \mathcal{R}} \alpha_i^{(r)} h_i^{(r)} \quad (9)$$

其中, $\alpha_i^{(r)}$ 表示关系 r 中的节点 v_i 对生成最终嵌入 h_i 的重要性。 $\alpha_i^{(r)}$ 的定义如下:

$$\alpha_i^{(r)} = \frac{\exp(\mathbf{q}^{(r)} \cdot \mathbf{h}_i^{(r)})}{\sum_{r' \in \mathcal{R}} \exp(\mathbf{q}^{(r')} \cdot \mathbf{h}_i^{(r')})} \quad (10)$$

其中, $\mathbf{q}^{(r)}$ 表示关系 $r \in \mathcal{R}$ 的特征向量。

通过这种方式,能够自适应地计算不同关系的权重,与均值池化的方式相比,各个关系类型上的信息能够获得良好的交互,以学习到更好的节点嵌入。

4.4 聚类优化模块

为了将图表示学习与节点聚类联系起来,使得学习到的节点嵌入更适合于聚类任务,本文结合对比学习的思想,提出了一个聚类优化模块,以增强图聚类的性能。

图聚类任务的目标是将相似的节点划分为同一个簇,

基于这一点,能够使用对比学习来使得相似节点在嵌入空间中更接近,不相似的节点在嵌入空间中相互远离。问题的关键是如何在嵌入空间中较好地区分与给定节点相似的节点和不相似的节点,换句话说,即如何定义对比学习中的正/负样本。在先前的工作中,大多都是通过数据增强来构造正负样本,但是不恰当的数据增强很可能会影响模型的性能^[21]。因此,为了避免不恰当的数据增强带来的影响,我们使用节点嵌入向量的距离来描述节点的相似性,将相似的节点作为正样本,将其他节点作为负样本。受文献^[5]的启发,使用 K-Nearest Neighbor(KNN)算法来获得给定节点的最近邻的 K 个邻居,并将这 K 个节点作为正样本,将其他的节点作为负样本。

如图1中的聚类优化模块所示,在嵌入空间中共有 N 个嵌入节点,对于一个给定的节点嵌入 h_i (图1中的绿色节点) 和一个给定的邻居数量 K , 可以通过 KNN 算法找到给定节点的 K 个近邻并将其作为正样本集合 $N_{P(i)}$, 将其余的 $N-K-1$ 个节点作为负样本集合 $N_{N(i)}$ 。由此就可以将给定节点与每一个正样本都组成正样本对,与每一个负样本也能组成负样本对。则单个节点 i 的聚类优化损失函数就可以写为:

$$\ell_i = -\log \frac{\sum_{h_j \in N_{P(i)}} \exp(\delta(h_i, h_j))}{\sum_{h_p \in N_{N(i)}} \exp(\delta(h_i, h_p))} \quad (11)$$

其中, $\delta(\cdot)$ 是一个余弦相似度计算函数, 用于计算两个向量的相似性(两个向量越相似, 该函数的值就越大)。再考虑到所有节点, 能够得到式(12)所示的总的邻域对比损失函数。

$$\mathcal{L}_{\text{copt}} = \frac{1}{N} \sum_{i=1}^N L_i \quad (12)$$

4.5 模型训练

合并式(7)、式(8)和式(11), 可以得到最终需要优化的损失函数, 表达式如下:

$$\mathcal{L}_{\text{total}} = \lambda_1 \mathcal{L}_g + \lambda_2 \mathcal{L}_c + \lambda_3 \mathcal{L}_{\text{copt}} + \lambda_4 \|\Theta\|^2 \quad (13)$$

其中, Θ 是所要优化的所有参数的集合, 采用 L2 正则化的方式; $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ 是几个权重参数, 用于保持不同部分的比重。通过优化上述等式, 学习到的节点嵌入保留了聚类任务所需要的信息, 提升了聚类任务的性能, 增强了聚类结果的可解释性。在网络训练结束之后, 再对最终学习到的节点嵌入执行聚类算法, 如 k -means^[2] 或 DBSCAN^[4], 以得到最终的聚类结果。

本文算法的描述如算法 1 所示。

算法 1 基于对比学习的多关系属性图聚类算法

输入: 属性矩阵 $\mathcal{G} \in \mathbb{R}^{N \times d}$, 邻接矩阵 $\mathcal{A} = \{\mathcal{A}^{(1)}, \mathcal{A}^{(2)}, \dots, \mathcal{A}^{(K)}\}$, 类簇数目 K

输出: 节点簇的集合

1. FOR $r \in \mathcal{R}$ DO:
2. 根据式(2)得到特定关系的嵌入向量 $\mathbf{h}^{(r)}$;
3. 根据式(3)和编码器 $\mathbf{E}_2^{(r)}$ 分别得到图级表示 $\mathbf{s}^{(r)}$ 和社区级表示 $\mathbf{c}_i^{(r)}$;
4. 根据式(7)和式(8)对特定关系的向量 $\mathbf{h}^{(r)}$ 进行进一步优化。
5. 根据式(9), 得到融合嵌入向量 \mathbf{h} ;
6. 根据式(12)对融合嵌入向量 \mathbf{h} 进行进一步优化, 使得 \mathbf{h} 更适合于聚类任务;
7. 使用 K-Means 算法对融合嵌入向量 \mathbf{h} 进行聚类, 获得最终的聚类划分的节点簇的集合。

5 实验

5.1 数据集

为了验证模型的有效性和鲁棒性, 我们在 3 个图数据集和 1 个真实场景的数据集上分别进行了实验。这些数据集的统计信息如表 1 所列。

表 1 数据集信息统计
Table 1 Datasets statistics

数据集	节点数量	特征维度	关系	边的数量	任务
DBLP	4 057 (authors)	334	APA	6 772 278	将作者按照研究领域分为 4 类
			APCPA	5 000 495	
			APTPA	11 113	
IMDB	4 780 (movies)	1 232	MAM	98 010	将电影按照类型分为 3 类
			MDM	21 018	
ACM	3 025 (papers)	1 830	PAP	29 281	将文章按照类型分为 3 类
			PSP	2 210 761	
			AWA	10 914	
FG	8 028 (accounts)	44	ALA	6 071	挖掘出隐藏的串通交易组的账户
			AIA	1 893	
			ANA	52 430	

在这些数据集中, DBLP^[1], IMDB^[2], ACM^[3] 是 3 个被广泛使用的公开数据集, FG 是来自郑州商品交易所的真实场景数据集, 本文使用这几个数据集构建了定义 1 中的多关系

属性图。数据集的详细描述如下:

1) DBLP。该数据集共包含 14 328 篇论文(P)、4 057 位作者(A)、20 个会议(C)和 8 789 期(T)。所生成的多关系属性图以论文作者(A)为顶点, 共包含 3 种类型的关系, 分别是 APA, APCPA 和 APTPA。节点的属性特征是由 bag-of-words 所表示的关键词。该数据集上的任务是将作者表示的节点按照研究领域分为 4 类: 数据库、数据挖掘、机器学习和信息检索。

2) IMDB。该数据集共包含 4 780 部电影(M)、5 841 位演员(A)和 2 269 位导演(D)。所生成的多关系属性图以电影为节点, 共包含两种类型的关系, 分别是 MAM 和 MDM。节点的属性特征是由 bag-of-words 所表示的电影情节概要。该数据集上的任务是将电影表示节点按照电影类型分为 3 类: 动作片、喜剧片和戏剧片。

3) ACM。该数据集共包含 3 025 篇论文(P)、5 835 为作者(A)和 56 个主题(S)。所生成的多关系属性图以论文(P)为节点, 共包含两种类型的关系, 分别是 PAP 和 PSP。节点的属性特征是由 bag-of-words 所表示的关键词。该数据集上的任务是将论文表示的节点按照文章类型分为 3 类: 数据库、无线通信和数据挖掘。

4) FG。该数据集是本文构建的期货数据集, 截取自郑州商品交易所在 2020 年 9 月 1 日至 2020 年 9 月 15 日(共计 11 个交易日)内的某期货合约上有过交易的账户(A)的交易数据和终端信息数据。该数据集以账户为图的顶点, 共包含 4 种类型的关系, 即共用 Windows 类型设备、共用 Linux 类型设备、共用 iOS 类型设备和共用 Android 类型设备, 分别记为 AWA, ALA, AIA, ANA。将账户在合约上的交易情况作为节点属性。在评价实验结果时所使用的节点标签来自交易所风控专家的人工标注。该数据集上的任务是挖掘出串通交易组的账户。

5.2 评价指标

本文使用了几个被广泛采用的评价指标: Accuracy (ACC), Normalized Mutual Information(NMI), Average Rand Index(ARI) and macro F1-score(F1)。对于所有度量指标, 其值越大就意味着模型的性能越好。

5.3 基线方法

为了更好地评估本文方法的性能, 本文选择了 10 个最先进的基线方法进行比较。这些基线任务大致上可以划分为以下 5 类。

1) 直接聚类: k -means^[2] 和 Spectral^[3]。这两种方法是直接对数据进行聚类, 其中 k -means 只使用了节点的属性而没有考虑到拓扑结构, Spectral 只使用了图的拓扑结构而没有考虑节点属性。

2) 无属性图: DeepWalk^[22], Node2Vec^[23] 和 LINE^[24]。DeepWalk 和 Node2Vec 都是基于 random-walk 和 skip-gram 的节点嵌入模型。LINE 通过设计的一阶相似度和二阶相似度来获得节点嵌入。

3) 属性图: GCN^[25] 和 DGI^[6]。GCN 根据节点局部结构来学习节点的嵌入。DGI 通过最大化节点嵌入和图级表示向量之间的互信息来获得良好的节点嵌入。

4)多关系属性图:DMGI^[12]和HDMI^[26]。DMGI使用一致性正则化方法来聚合不同关系上的节点嵌入,并使用注意力机制得到最终的节点嵌入。HDMI通过高阶互信息来设计包含外部和内部互信息的联合监督信号,以学习节点嵌入,并且还使用注意力机制来聚合不同关系上的节点嵌入。

5)联合优化:SDCN^[27]。SDCN通过结合自编码器和GCN的表示,将结构信息集成到深度聚类中,实现了图表示学习与聚类任务的联合优化。

其中,对于单关系图的嵌入算法,我们在每个关系形成的图中单独执行算法,再将各个关系上所得到的结果进行均值操作来得到最终的嵌入结果。

5.4 实验设置

为了公平比较,在公开数据集上进行实验时,我们将所有实验中的节点嵌入的维度都设置为64。对于其余的基线

任务的超参数,我们保留了其相应论文或实验代码中所使用的参数设置。对于本文模型,使用Adam优化算法^[28]来对模型进行优化,使用网格搜索技术来调整超参数,按照6:2:2的比例随机划分训练集、验证集和测试集,并在验证集上表现最好的超参数。采用提前停止机制并将最大等待次数设置为100,以防止模型过拟合。 λ_1 设置为2, λ_2 设置为1, λ_3 设置为0.01, λ_4 设置为0.0001。多头注意力的头数设置为8,最近邻的K值设置为6,学习率设置为0.001。

5.5 实验结果

3个公开数据集中不包含噪音且所要划分的簇的数量是已知的,因此采用K-Means算法进行聚类。由于K-Means算法的性能会受到初始簇中心的影响,因此本文重复K-Means过程10次,并取10次结果的平均值作为最终结果,如表2所列。

表2 模型在公开数据集的测试集上的性能

Table 2 Model performance on test sets DBLP,IMDB and ACM

(单位:%)

方法	DBLP				IMDB				ACM			
	Acc	F1	NMI	ARI	Acc	F1	NMI	ARI	Acc	F1	NMI	ARI
K-means	24.82	17.79	11.63	2.49	30.48	20.11	1.45	3.05	30.09	29.83	38.43	38.11
Spectral	24.17	9.70	0.18	0.00	33.86	18.80	0.10	0.01	33.49	16.78	0.13	0.00
LINE	24.83	24.46	54.76	59.88	32.75	28.70	0.10	-0.90	33.69	33.66	39.40	34.33
Node2Vec	28.44	30.33	19.33	12.86	34.34	31.55	0.22	-0.50	33.17	32.41	0.39	0.29
DeepWalk	25.44	22.67	1.76	1.33	33.97	32.28	1.45	2.15	32.42	31.40	41.61	35.10
DGI	25.17	22.09	60.38	63.27	32.20	26.93	0.31	-0.23	34.53	36.06	48.97	50.15
GCN	23.17	21.95	62.34	64.21	20.58	18.39	5.45	4.40	39.73	40.82	51.40	53.01
DMGI	27.24	27.85	70.34	68.61	35.83	33.88	5.69	4.49	32.96	33.79	67.18	73.05
HDMI	24.58	24.40	68.70	73.54	30.31	28.05	9.87	11.72	39.21	40.20	64.86	66.98
SDCN	35.85	36.54	67.09	60.69	35.46	36.90	9.92	10.08	56.02	60.85	65.17	67.06
CCLMAG	32.21	31.84	70.88	74.41	40.56	35.43	11.28	15.89	64.36	64.88	70.81	76.86

注:每一列的最优结果都已加粗表示。

通过与其他基线任务的比较,我们有以下发现:

1)LINE,Node2Vec和DeepWalk的表现相比其他模型都较差,这是因为它们没有利用节点的属性信息。与这几个模型相比,本文模型的结果都远远高于它们,表明了本文模型能够有效结合图的拓扑结构信息和节点属性信息,具有优越的性能。

2)K-means只利用节点的属性信息,Spectral只利用了图的结构信息,两者都只进行了简单的聚类过程,导致其在这3个数据集上的性能较差。与这种浅层模型相比,本文模型取得了较好的结果,说明了本文的深度模型具有优越的节点表示学习能力。

3)DGI和GCN的表现相比DMGI和HDMI都较差,这是因为DGI和GCN无法良好聚合多关系上的信息。与DMGI相比,本文模型在ACC,F1,NMI,ARI这4个指标上分别提高了3%~30%;与HDMI相比,本文模型在以上4个指标上分别提高了2%~15%。上述结果表明,本文模型引入的嵌入融合模块能够自动学习聚合不同关系上的信息,获得更加良好的效果。

4)SDCN的性能相比其他基线模型较优,这是因为它利用了节点嵌入和聚类的联合优化策略。与之相比,本文模型在NMI和ARI指标上都要优于SDCN,分别高出至少1%和5%。这说明联合训练策略可以有效获得具有有用聚类信息的节点表示。在DBLP和IMDB数据集上,CCLMAG的ACC或F1值略低于SDCN,这可能是由于SDCN经过了多层卷积,学习到的结果在准确率的计算上略有优势。

5)所有模型在IMDB数据集上的表现比在ACM数据集和DBLP数据集都要差得多。这可能是由于IMDB数据集中的标签存在噪音:原始的IMDB数据集中每部电影都具有多种标签,而在实验中只选择了其中的一个作为它的类标签。

5.6 参数敏感性

本节研究了模型对参数的敏感性,并展示了模型关于不同参数在ACM数据集上的表现结果(以NMI评价指标为例),如图2所示。

1)嵌入向量的维度对结果的影响。我们测试了嵌入向量的维度对结果的影响,如图2(a)所示。可以看到,随着嵌入维度的增大,模型性能在总体上呈现一种先提升后下降的

¹⁾ <https://dblp.uni-trier.de/>

²⁾ <https://www.imdb.com/>

³⁾ <http://dl.acm.org/>

趋势,并且当嵌入维度为 64 时,模型性能可以达到最高点。出现这样结果的原因是,模型需要一个合适的维度来编码数据,维度过小则无法表达全部信息;而维度过大,嵌入结果中可能会引入过多的冗余信息,影响模型的性能。

2)注意力头数对结果的影响。我们测试了嵌入融合模块中多头注意力机制的注意力头数对结果的影响,如图 2(b)

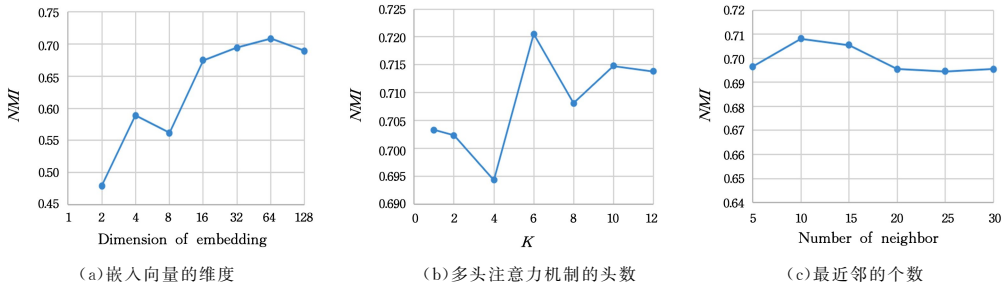


图 2 几个超参数对模型性能的影响

Fig. 2 Influence of some super parameters on model performance

3)邻域选择范围对结果的影响。我们测试了聚类优化模块中选择节点邻域时超参数 K 对结果的影响,如图 2(c)所示。可以看到,当邻域范围 K 改变时,模型性能基本上没有变化(波动非常轻微),这说明邻域范围 K 对模型性能的影响基本可以忽略不计。出现该结果的原因是每一类别的节点数量较多,对于一个目标节点来说,周围与它相似的节点基本上

所示。可以看到,当注意力头数改变时,实验结果也在一定范围内波动,当值小于 6 时,实验结果明显偏低,这表明当注意力头数较少时,模型无法较好地融合各个关系上的信息;而随着注意力头数的增多,实验结果呈现一种下降的趋势,这表明过多的注意力头数可能会引入噪声数据,导致模型的性能变差。

与它都属于同一类别,因此不管采集多少个邻居,实验结果都不会有太大的波动。

5.7 消融实验

本节主要验证了相比 DGI 文章新添加部分的有效性,分别验证了嵌入融合模块、社区级互信息和聚类优化模块 3 个方面对模型性能的影响。实验结果如表 3 所列。

表 3 消融实验的结果

Table 3 Results of ablation experiment

(单位:%)

方法	DBLP				IMDB				ACM			
	Acc	F1	NMI	ARI	Acc	F1	NMI	ARI	Acc	F1	NMI	ARI
CCLMAG1	18.20	17.56	68.45	72.92	36.50	34.95	8.76	11.30	62.66	63.94	69.90	76.09
CCLMAG2	27.23	25.77	68.01	70.98	33.08	29.97	4.27	4.82	57.69	57.90	69.15	75.34
CCLMAG3	31.86	30.28	68.70	72.06	33.82	30.31	3.29	3.67	31.59	34.14	69.55	75.56
CCLMAG	32.21	31.84	70.88	74.41	40.56	36.43	11.28	15.89	64.36	64.88	70.81	76.86

可以看到,模型在去除任一模块后,在所有数据集的所有评价指标上性能都有所下降,说明这几个模块对模型的性能都具有积极影响。

1)嵌入融合模块对性能的影响(CCLMAG1)。为了评价融合模块对性能的影响,我们将融合模块更改为均值池化操作。结果如表 3 所列,可以看到在所有数据集上嵌入融合模块比均值池化操作的性能要好,这说明融合模块对模型的性能提升有着积极的影响。

2)社区级互信息对性能的影响(CCLMAG2)。为了评价社区级互信息对性能的影响,我们将社区级互信息从模型中去除,然后比较结果。结果如表 3 所列,可以看到在所有数据集的所有评价指标上,将模型中的社区级互信息去除之后,性能都会变差,并且同样可以发现,两者的性能差距与融合模块相比明显较大,这也能够说明社区级互信息对性能的影响比融合模块对性能的影响更大。

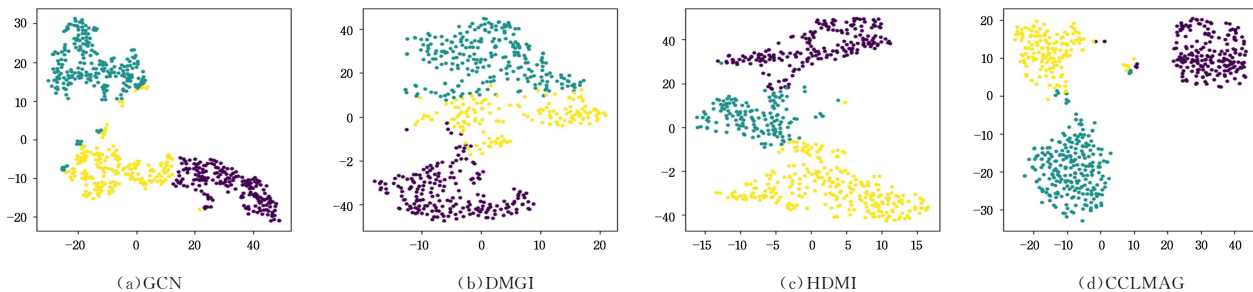
3)聚类优化模块对性能的影响(CCLMAG3)。为了评价聚类优化模块对性能的影响,我们将聚类优化模块从模型中去除,然后比较结果。结果如表 3 所列,可以看到在所有数据集的所有评价指标上,当把聚类优化模块去除之后,性能都会

变差,这说明聚类优化模块的存在可以提升聚类任务的性能。并且从 5.8 节中也可以看到,经过聚类优化模块之后的节点嵌入,形成了明显的可分离的簇,增强了聚类结果的可解释性。

5.8 可视化分析

为了更直观地比较聚类结果,本文将学习到的节点嵌入投影到一个二维空间中,进行可视化分析任务。具体地,本文使用 t-SNE^[29] 工具来可视化 ACM 数据集中的测试集的文章节点的嵌入,按照算法得到的聚类结果对节点进行着色。

如图 3 所示,我们选取了 ACM 数据集中所划分的结果进行展示,并将其作为一种应用于单一关系图的数据。GCN 无法良好处理多关系图数据,导致不同类的节点混合纠缠在一起,而 DMGI 和 HDMI 的可视化结果则要比 GCN 略好,同时不同类别的颜色边界也较为清晰,但是不同类别的节点却没有明显分离的情况。相比这 3 个模型的结果,所提模型不仅能够对不同类别的节点划分出清晰的边界,而且能使得同类节点紧密地聚合在一起,不同类别的节点明显分离,能够直观划分出不同类别的节点所表示的节点簇。



注:每个点表示一篇论文,同种颜色表示论文属于同一类型。

图3 ACM数据集嵌入结果可视化

Fig.3 Visualization of ACM embedding

5.9 在私有数据集上的有效性

为了评价模型在实际应用场景中的有效性,本文在郑州商品交易所的期货交易场景中应用模型,以检测出期货交易中隐藏的串通账户组。串通账户指,一些用户为了谋取私利而串通交易、恶意炒作,这种行为无疑危害着市场的公平性。串通账户可能具有相似的交易行为,或共用相同的设备。由于金融服务行业的特殊性,串通账户的确定应该尽可能谨慎,以避免对用户造成不必要的打扰,即要保证结果的准确率尽可能高,因此结合文献[30]的实验思路,在这里本文使用 Acc 作为评价指标。

由于该数据集中含有大量噪音(有的用户不属于任一串通组),本文事先根据已有的交易规则对数据进行了过滤,但是依旧存在一些无法依靠规则过滤的噪音数据。因此,本文使用 DBSCAN 算法^[4]替代 K-Means 算法对节点嵌入进行聚类,从而得到最终结果。同样,SDCN 算法也无法解决这种含有噪音的聚类任务。

综上所述,本节只选取了 8 个基线任务进行比较,实验结果如图 4 所示。嵌入空间的维度被设置为 32,其他参数设置与前文在公开数据集上的实验一致。可以看到,与其他基线任务相比,本文方法在 Acc 指标上具有明显的优势,相比其他基线任务普遍能够提升 5%~20%。这表明本文方法在真实场景的数据集上也同样是有效的。

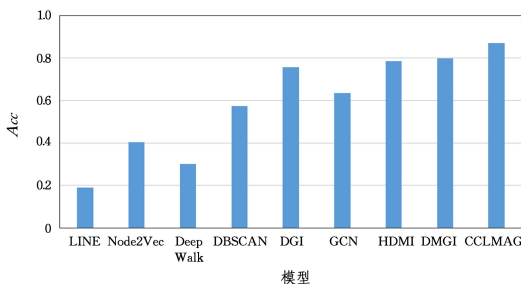


图4 真实数据集上的实验结果

Fig.4 Experimental results on real-world FG dataset

结束语 针对多关系属性图的聚类问题,本文提出了一种结合对比学习思想的多关系属性图聚类算法。引入社区级互信息机制,使得学习到的节点嵌入能够保留节点的簇信息,增加嵌入融合模块能够更好地融合不同关系上的信息,聚类优化模块能够将图表示学习和聚类任务联系起来,增强了聚类结果的准确性和可解释性。在公开数据集

和构建的期货数据集上与众多基线任务进行对比,结果表明本文模型具有良好的性能。

本文对节点表示与聚类两个任务的联合优化不够彻底,下一步会考虑将聚类任务与节点表示两者完全整合,以进一步提升模型的表现能力。

参考文献

- [1] SU X, XUE S, LIU F Z, et al. A Comprehensive Survey on Community Detection With Deep Learning[J]. arXiv:2105.12584, 2021.
- [2] HARTIGAN J A, WONG M. Algorithm AS 136: A k-means clustering algorithm[J]. Journal of the Royal Statistical Society (Series C), 1979, 28(1): 100-108.
- [3] ANDREW N, MICHAEL J, YAIR W. On spectral clustering: Analysis and an algorithm[C]// Advances in Neural Information Processing Systems 14. 2001: 849-856.
- [4] MARTIN E, HANS-PETER K, SANDER J, et al. A density-based algorithm for discovering clusters in large spatial databases with noise[C]// Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. 1996: 226-231.
- [5] WANG T, YANG G Y, WU J H, et al. Dual Contrastive Attributed Graph Clustering Network[J]. arXiv:2206.07897, 2022.
- [6] PETAR V, WILLIAM F, WILLIAM L, et al. Deep Graph Informax[J]. arXiv:1809.10341, 2018.
- [7] BU Z, WANG Y Y, MA L N, et al. Attributed Graph Clustering Approach Based on Dynamic Cluster Formation Game[J]. Chinese Journal of Computers, 2021, 44(9): 1824-1840.
- [8] FAN S H, WANG X, SHI C, et al. One2Multi Graph Autoencoder for Multi-view Graph Clustering[C]// Proceedings of The Web Conference 2020. 2020: 3070-3076.
- [9] SHI Y, HAN F Q, HE X W, et al. mvn2vec: Preservation and collaboration in multi-view network embedding[J]. arXiv:1801.06597, 2018.
- [10] WANG X, JI H Y, SHI C, et al. Heterogeneous Graph Attention Network[C]// The World Wide Web Conference. 2019: 2022-2032.
- [11] MICHAEL S, THOMAS N K, PETER B, et al. Modeling relational data with graph convolutional networks[C]// European Semantic Web Conference 2018. 2018: 593-607.

- [12] PARK C Y, KIM D, HAN J W, et al. Unsupervised Attributed Multiplex Network Embedding[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2020;5371-5378.
- [13] WANG J, CAO J, LI W, et al. CANE: community-aware network embedding via adversarial training[C]// Knowledge and Information Systems. 2021;411-438.
- [14] KANG Y, KOU Y Q, XIE S Y, et al. Deep Clustering Model Based on Fusion Variational Graph Attention Self-encoder [J]. Computer Science. 2021, 48(S2):81-87, 116.
- [15] YANG X H, WANG L, YE L, et al. Complex Network Community Detection Algorithm Based on Node Similarity and Network Embedding [J]. Computer Science. 2022, 49(3):121-128.
- [16] PAN S Y, QIN Y M, LI T Y, et al. Momentum contrast for unsupervised visual representation learning[C]// CVF Conference on Computer Vision and Pattern Recognition Workshops. 2022: 2646-2655.
- [17] LU S H, CHEN L, LU G Y, et al. Weakly-Supervised Contrastive Learning Framework for Few-Shot Sentiment Classification Tasks [J]. Journal of Computer Research and Development, 2022, 59(9):2003-2014.
- [18] WANG X, LIU N, HAN H, et al. Self-supervised Heterogeneous Graph Neural Network with Co-contrastive Learning [C]// Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. 2021;1726-1736.
- [19] ZHAO H, YANG X, WANG Z R, et al. Graph Debaised Contrastive Learning with Joint Representation Clustering[C]// International Joint Conferences on Artificial Intelligence Organization. 2021;3434-3440.
- [20] YOU Q Z, JIN H L, WANG Z W, et al. Image captioning with semantic attention[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition. 2016;4651-4659.
- [21] YANG Y H, WU L, HONG R C, et al. Enhanced Graph Learning for Collaborative Filtering via Mutual Information Maximization[C]// Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2021;71-80.
- [22] BRYAN P, RAMI A, STEVEN S, et al. Deepwalk: Online learning of social representations [C]// Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2014;701-710.
- [23] ADITYA G, JURE L. node2vec: Scalable feature learning for networks[C]// Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016;855-864.
- [24] 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. 2015;1067-1077.
- [25] THOMAS K M W. Semi-Supervised Classification with Graph Convolutional Networks[J]. arXiv:1609.02907, 2016.
- [26] JING B Y, PARK C Y, TONG H H, et al. HDMI: High-order Deep Multiplex Infomax[C]// Proceedings of the Web Conference 2021. 2021;2414-2424.
- [27] BO D Y, WANG X, SHI C, et al. Structural deep clustering network[C]// Proceedings of The Web Conference 2020. 2020: 1400-1410.
- [28] DIEDERIK P K, JIMMY B. Adam: A method for stochastic optimization[J]. arXiv:1412.6980, 2014.
- [29] LAURENS V D M, GEOFFREY H. Visualizing data using t-SNE[C]// Journal of Machine Learning Research. 2008;2579-2605.
- [30] WANG H B, ZHOU C, WU J, et al. Deep Structure Learning for Fraud Detection[C]// 2018 IEEE International Conference on Data Mining. 2018;567-576.



XIE Zhuo, born in 1997, postgraduate, is a member of China Computer Federation. His main research interest is deep learning on graphs.



ZHOU Lijuan, born in 1987, Ph.D, associate professor, is a member of China Computer Federation. Her main research interests include cross modal semantic understanding and computer vision.

(责任编辑:喻藜)