

跨视图一致性表示的多视图属性图聚类

陈晓惠,荆雪纯,曹付元*

(山西大学 计算机与信息技术学院,山西 太原 030006)

摘要:针对基于图卷积神经网络的多视图聚类算法中视图嵌入表示存在的两个问题:跨视图特征表示一致性不足、跨视图簇划分一致性不足,本文提出了一种跨视图一致性表示的多视图属性图聚类算法(Multi-view Attribute Graph Clustering Algorithm with Cross-view Consistent Representations, CCRAGC)。该算法通过约束视图间的节点相似度矩阵逼近单位矩阵来加强视图嵌入表示之间的特征级一致性;同时,把视图嵌入表示映射到聚类级子空间,使子空间中的软标签矩阵尽可能相似,以此来加强视图间簇划分的一致性。研究表明,CCRAGC对于计算机协会(Association for Computing Machinery, ACM)、dblp 计算机科学书目(dblp computer science bibliography, DBLP)、互联网电影资料库(Internet Movie Database, IMDB)三个广泛使用的数据集都有效,且准确率(Accuracy, Acc)相对于性能最优的基准算法分别提高了1.21%、0.37%、5.74%。

关键词:聚类;特征一致性;簇级一致性;属性图

中图分类号:TP311 **文献标志码:**A **文章编号:**0253-2395(2025)05-0921-12

Cross-view Consistent Representation for Multi-view Attribute Graph Clustering

CHEN Xiaohui, JING Xuechun, Cao Fuyuan*

(School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China)

Abstract: Aiming at the two problems of view embedding representations in GCN-based multi-view clustering algorithms: insufficient cross-view feature consistency and insufficient cross-view clustering consistency, this paper proposes a multi-view attribute graph clustering algorithm (CCRAGC) with cross-view consistent representations. The algorithm strengthens the feature-level consistency between view-embedded representations by calculating the node similarity matrix between views, and then constraining the node similarity matrix to approximate the unit matrix; at the same time, it maps the view-embedded representations to the clustering-level subspaces, so that the soft-label matrices in the subspaces are as similar as possible to a way of enhancing the learning of view-embedded representations and correlation of clustering tasks. The results of the study show that CCRAGC is effective for three widely used datasets, namely ACM, DBLP, and IMDB, and Acc improves by 1.21%, 0.37%, and 5.74%, respectively, with respect to the benchmark algorithm with the best performance.

Key words: clustering; feature consistency; cluster consistency; attribute graph

收稿日期:2024-01-07; **接受日期:**2024-02-27

基金项目:国家自然科学基金(62376145);山西省科技创新人才团队(202204051002016);山西省基础研究计划(202203021212416)

作者简介:陈晓惠(1998-),女,山西晋城人,硕士研究生,研究方向为多视图聚类。E-mail:13393465764@163.com

* **通信作者:**曹付元(CAO Fuyuan), E-mail:cfy@sxu.edu.cn

引文格式:陈晓惠,荆雪纯,曹付元.跨视图一致性表示的多视图属性图聚类[J].山西大学学报(自然科学版),2025,48(5):921-932. DOI:10.13451/j.sxu.ns.2024030.

0 引言

多视图聚类是一种利用多个特征来对数据进行聚类的机器学习技术。它通过整合不同来源的信息,更全面地理解数据的本质,从而产生更准确、更鲁棒的聚类结果。自2010年以来,深度神经网络由于其良好的非线性映射能力和在不同场景下的灵活性,已经广泛应用于多视图聚类任务。深度多视图聚类算法可以分为三类:基于深度嵌入聚类(Deep Embedded Clustering, DEC)的多视图聚类^[1-5]、基于子空间的多视图聚类^[6-9]和基于图卷积神经网络(Graph Convolutional Networks, GCN)的多视图聚类^[10-16]。基于DEC的多视图聚类算法和基于子空间的多视图聚类算法在欧式空间中能够取得良好的性能。然而,在处理图数据这种非欧式数据时,算法的性能表现并不尽如人意。传统的多视图数据每个视图只有一个特征表示,然而对于图数据,每个视图的数据类型比传统数据更复杂,不仅涉及节点的属性特征,还包括表示节点连接关系的拓扑结构。由于图数据中的拓扑结构是非欧式的,其复杂性远超传统的数据类型,这无疑增加了处理的难度,传统的特征提取方法难以从中提取出有意义的特征表示。图卷积神经网络^[17]在处理图数据时展现出卓越的性能,它通过定义图卷积操作,将每个节点的特征与其邻居节点的特征进行聚合,形成新的节点表示,该表示能够有效地捕捉节点之间的复杂关系。因此,基于GCN的多视图聚类算法已成为处理图数据的有效方法,它在处理属性图聚类任务时表现出色。属性图聚类任务会考虑节点的属性特征,将其与图的拓扑结构信息相结合,以共同决定节点的聚类归属。现有的算法中,直接使用GCN作为编码器生成视图的嵌入表示是一种常见的做法,如多视图属性图卷积聚类网络(Multi-view Attribute Graph Convolution Networks for Clustering, MAGCN)^[11]、一对多图自编码器多视图图聚类(One2Multi Graph Autoencoder for Multi-view Graph Clustering, O2MAC)^[12]、基于一致性约束的多重图嵌入多视角聚类(Consistent Multiple Graph Embedding for Multi-view Clustering,

CMGEC)^[13]等方法都采用了这种方式。除此之外,还有一些其他的结合方式。如结构深度聚类网络(Structural Deep Clustering Network, SDCN)^[14]用自编码器学习属性特征矩阵,并将自编码器每层学到的表示传递给相应的GCN层,从而将属性特征集成到GCN编码器中。深度融合聚类网络(Deep Fusion Clustering Network, DFCN)^[15]设计了一个动态信息融合模块,对从自编码器和图自编码器中提取的属性特征和拓扑结构信息进行精细处理。基于马尔可夫扩散的图神经网络(Graph Neural Networks on Markov Diffusion, Markov-GNN)^[16]从原始的静态结构中生成一系列拓扑结构,增强了GCN的表达能力。但现有的算法仍存在以下问题:(1)多视图图编码器采用独立的GCN编码器对每个视图进行处理,这种做法简单直观,但独立的GCN编码器无法充分挖掘各个视图嵌入表示的内在联系,视图嵌入表示之间缺乏必要的协同和一致性约束,这会导致视图嵌入表示之间的特征一致性不足。(2)现有的多视图属性图聚类算法通常采用端到端的框架,将聚类和特征表示学习集成在一个模型中,然后对模型进行训练。在训练过程中,先对视图嵌入表示进行预训练学习,然后加入聚类模块,对视图嵌入表示进行微调。然而,这种做法导致在预训练时视图嵌入表示的学习和聚类任务割裂,提取的视图嵌入表示在聚类级的一致性不足。

针对以上问题,本文提出了跨视图一致性表示的多视图属性图聚类算法(Multi-view Attribute Graph Clustering Algorithm with Cross-view Consistent Representations, CCRAGC)。在CCRAGC算法中,首先,通过多视图图编码器获得多个视图的嵌入表示;其次,利用特征级一致性模块加强视图嵌入表示之间的特征一致性;再次,利用簇级一致性模块使视图嵌入表示的学习过程与聚类任务的联系更加紧密;继续,通过注意力机制融合多个视图的嵌入表示,并利用融合表示重构图的拓扑结构;最后,加入自监督聚类模块进一步指导整个模型的学习。本文的主要贡献总结如下:

(1)针对跨视图特征表示一致性不足的问题

题,设计了特征级一致性模块,通过约束视图之间的节点相似度矩阵逼近单位矩阵来促进视图嵌入表示之间的特征级一致性。

(2)针对跨视图簇划分一致性不足的问题,设计了簇级一致性模块,通过把视图嵌入表示映射到聚类级子空间,让子空间中的软标签矩阵尽可能的相似来加强视图的嵌入表示之间的簇级一致性。

(3)在三个公开的数据集上与经典的算法比较,验证了所提出算法的有效性。

1 相关工作

深度多视图聚类算法可以分为三类:基于DEC的多视图聚类、基于子空间的多视图聚类和基于GCN的多视图聚类。

基于DEC的多视图聚类算法的特点是采用了合并自编码器和聚类层的方式,用深度自编码器学习低维嵌入,然后利用KL散度(Kullback-Leibler Divergence, KL Divergence)^[1]实现聚类。KL散度是一种用于度量两个概率分布之间差异的方法,在聚类层中,通过定义软标签分布和目标分布,并利用KL散度来拉近这两个分布之间的距离,可以优化聚类质心,改善聚类分配。作为首个基于DEC的多视图聚类算法,DEC^[1]由两部分组成,第一部分预训练一个自编码器模型,第二部分去除自编码器模型中的解码器,并加入聚类层,使用KL散度进行训练。基于局部结构保持的改进深度嵌入聚类(Improved Deep Embedded Clustering with Local Structure Preservation, IDEC)^[2]则提出保持DEC中的解码器不变,增加特征表示的重构损失项,然后联合原有的聚类损失项作为总目标函数进行训练,该方法证明了局部结构在深度聚类中的重要性。除了利用深度自编码器,深度对抗多视图聚类网络(Deep Adversarial Multi-view Clustering Network, DAMC)^[3]还引入了对抗性学习。对于每个嵌入表示,除了生成重构样本外,还利用鉴别器识别生成的样本,从而更好地揭示多视图数据的内在结构和非线性特征。与以上方法相似,基于协作训练的深度嵌入多视图聚类(Deep Embedded Multi-view Clustering with Collaborative Training, DEM-

VC)^[4]也是先利用深度自编码器对每个视图特征进行单独学习,但在聚类层采用多视图协同训练的方法,让每个视图轮流成为参考视图,指导整个网络的学习,该方法能更好地利用多视图的一致信息和互补信息。基于多层次特征学习的对比多视图聚类(Multi-level Feature Learning for Contrastive Multi-view Clustering, MFLVC)^[5]提出了一个多层次特征学习框架,该方法把深度自编码器和对比学习结合,利用自编码器重建低级特征,利用对比学习实现高级特征和语义特征的一致性,能减少视图私有信息对聚类的不利影响。

基于子空间的多视图聚类算法假设不同簇的样本来自多个子空间,通常先估计每对样本的亲密度,形成亲和矩阵,然后对亲和矩阵进行谱聚类。深度多视图稀疏子空间聚类(Deep Multi-view Sparse Subspace Clustering, DMVSSC)^[6]利用基于典型相关分析(Canonical Correlation Analysis, CCA)^[7]的自表达模块,把从多视图数据中提取的特征进行集成,能加强多视图的特征交互,更好地探索多视图的互补信息。与DMVSSC相比,多视图深度子空间聚类(Multi-view Deep Subspace Clustering, MVD-SCN)^[8]分别学习特定于视图的自表示矩阵和多视图通用的自表示矩阵,使多视图的特征提取阶段不再独立存在,能更好地探索多视图之间的关系。

与基于DEC的多视图聚类算法和基于子空间的多视图聚类算法相比,基于GCN的多视图聚类算法在处理图数据时具有显著的优势和更高的性能。MAGCN^[11]在生成视图的嵌入表示时为不同节点分配了可学习的权重,在生成聚类表示同时考虑了多视图的几何关系一致性和概率分布一致性。但是,MAGCN是一个专为处理由单拓扑结构和多属性特征矩阵组成的图数据设计的聚类模型,而O2MAC^[12]则可以处理由多拓扑结构和单属性特征矩阵组成的图数据,它的聚类损失和重构损失在形式上与基于DEC的多视图聚类算法相似,不同之处在于它采用了GCN作为编码器。与以上两种方法相比,CMGEC^[13]的适用范围更广泛一些,它可以处理由多拓扑结构和多属性特征矩阵组成的

图数据。CMGEC一方面利用互信息最大化模块来维持不同视图中相同节点的相似性,另一方面利用图融合网络学习不同视图的共识图,该模型能编码更多的互补信息。

2 方法

CCRAGC模型框架如图1所示。首先,给定一个输入 $G = \{A_1, A_2, \dots, A_K, X\}$, 其中 K 表示视图的个数, X 表示属性特征矩阵, A_1, A_2, \dots, A_K 表示 K 个邻接矩阵, 先通过多视图图编码器学习每个视图的嵌入表示 Z_1, Z_2, \dots, Z_K 。然后, 计算两两视图之间的节点相似度矩阵, 如第 k 个视图和第 k' 个视图之间的节点相似度矩阵为 $S_{(k,k')} \in \mathbb{R}^{N \times N}$, N 为节点的数量, 接着计算视图之间的特征级一致性损失, 并利用该损失项引导视图嵌入表示 Z_k 的学习, 加强跨视图特征一致性。同时, 把各视图的嵌入表示 Z_k 映射到聚类级子空间, 让各视图的软标签矩阵 Y_k 趋于一致, 从而加强嵌入表示 Z_k 与聚类任务的紧密性, 提高视图嵌入表示之间的簇级一致性。接下来, 利用注意力机制融合各视图的嵌入表示并对输入的拓扑结构进行重构。最后, 添加自监督模块, 通过最小化总目标函数, 更新整个框架。下文详细介绍模型的每个模块。

2.1 多视图图编码器

受图自编码器的启发, 多视图图编码器由多个GCN编码器组成, 每个编码器对应一个视图。每个编码器可以利用拓扑结构的信息, 聚合相邻节点的特征, 从而更好地捕捉节点之间的关系, 学到更具代表性的视图嵌入表示。GCN编码器本质是一个神经网络, 通过层与层之间的传播, 属性特征矩阵 X 和第 k 个视图的邻接矩阵 A_k 最终被编码为 Z_k , 其第 l 层和第 $l+1$ 层之间的传播公式为:

$$Z_k^{(l+1)} = \sigma(\tilde{D}_k^{-\frac{1}{2}} \tilde{A}_k \tilde{D}_k^{-\frac{1}{2}} Z_k^{(l)} W_k^{(l)}), \quad (1)$$

其中 $Z_k^{(l+1)}$ 为第 k 个视图在第 $l+1$ 层的嵌入表示, σ 为神经网络的非线性激活函数, $\tilde{A}_k = A + I$, A 为邻接矩阵, I 为单位矩阵, \tilde{D}_k 为 \tilde{A}_k 的度矩阵, $W_k^{(l)}$ 为神经网络第 l 层的权重矩阵。设置 $Z_k^{(0)} = X$, 即第一层嵌入的输入为属性特征矩阵 X 。通过多个GCN编码器, 可以得到各个视图的嵌入表示, 本文采用的GCN网络层数为2, 选用的激活函数为Relu。

2.2 特征级一致性

通过多视图图编码器获得了各个视图的嵌入表示, 但是由于各视图的GCN编码器是相互独立的, 视图嵌入表示在跨视图特征一致性学习存在不足。为了解决这个问题, 本文设计了一个跨视图损失函数, 该损失函数鼓励相同节

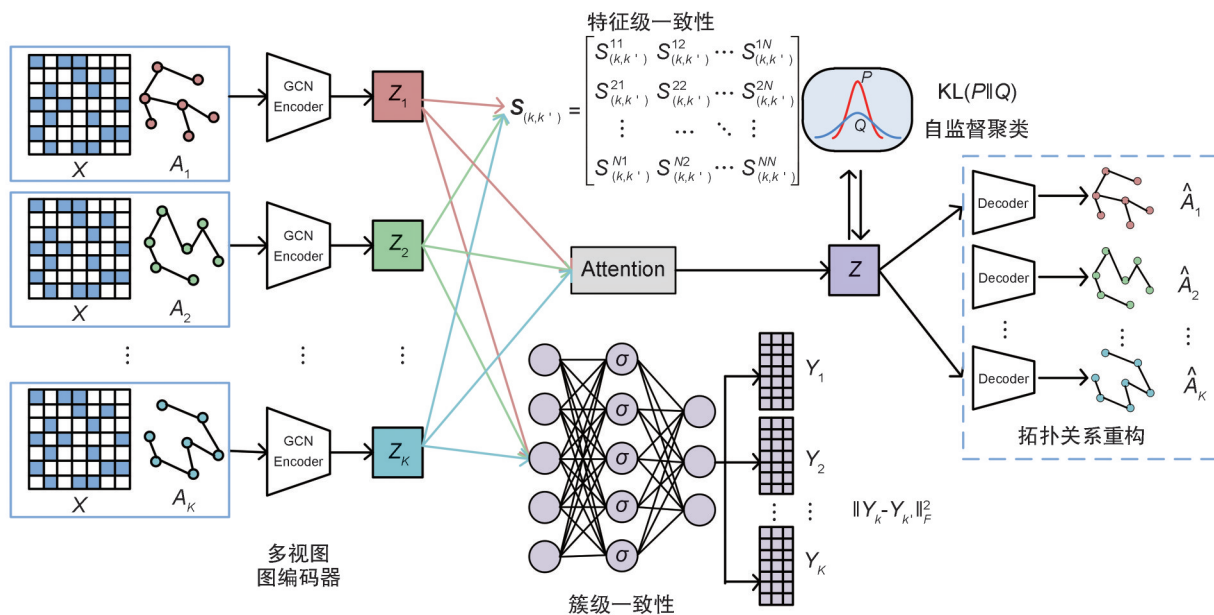


图1 CCRAGC模型框架图

Fig. 1 Overall structure of CCRAGC model

点在不同视图中的特征向量尽可能相似,同时不同节点的特征向量尽可能不同。具体来说,先计算视图之间的节点相似度矩阵 $S_{(k,k')} \in \mathbb{R}^{N \times N}$, $S_{(k,k')}$ 的元素 $S_{(k,k')}^{ij}$ 是 Z_k^i 和 $Z_{k'}^j$ 之间的余弦相似度, $S_{(k,k')}^{ij}$ 的计算公式为:

$$S_{(k,k')}^{ij} = \frac{(Z_k^i)(Z_{k'}^j)^T}{\|Z_k^i\| \|Z_{k'}^j\|}, \forall i, j \in [1, N], \quad (2)$$

其中 N 为节点的数量, Z_k^i 为第 k 个视图中第 i 个节点的特征向量, $Z_{k'}^j$ 为第 k' 个视图中第 j 个节点的特征向量, $S_{(k,k')}^{ij}$ 表示第 k 个视图中第 i 个节点的特征向量与第 k' 个视图中第 j 个节点的特征向量之间的余弦相似度。那么,鼓励 $S_{(k,k')}$ 的对角线元素等于 1, 也就是鼓励第 k 个视图和第 k' 个视图中相同节点的特征向量尽可能相似; 鼓励 $S_{(k,k')}$ 的非对角线元素等于 0, 也就是鼓励第 k 个视图和第 k' 个视图中不同节点的特征向量尽可能不同。由此, 得到第 k 个视图与第 k' 个视图的特征一致损失函数:

$$L_{(k,k')} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left((S_{(k,k')} - I)^{ij} \right)^2, \quad (3)$$

其中 I 表示单位矩阵, $S_{(k,k')}$ 表示第 k 个视图的嵌入表示和第 k' 个视图的嵌入表示之间的相似度矩阵。再将其扩展到所有视图:

$$L_{\text{feature}} = \sum_{k=1}^K \sum_{k' \neq k} L_{(k,k')} \quad (4)$$

通过最小化式(4), 促进了不同视图之间的信息交流, 提高了跨视图特征一致性表示学习。

2.3 簇级一致性

该函数将 Z_k 映射到子空间, 即 $Y_k = \text{clusterMLP}(Z_k)$, 其中 $\text{clusterMLP}(\cdot)$ 由两层非线性的多层感知机 (Multilayer Perceptron, MLP) 堆叠, Y_k 是一个 $N \times C$ 的矩阵, N 表示节点数量, C 表示簇的个数。 Y_k 可以看作第 k 个视图中节点的软标签矩阵, Y_k^{ij} 表示第 i 个节点属于第 j 个簇的概率。各视图的 Y_k 应该是相似的, 本文选用差异相似度函数度量 K 个视图软标签矩阵的相似度, 并最小化簇级一致损失函数:

$$L_{\text{cluster}} = \sum_{k \neq k'}^K \|Y_k - Y_{k'}\|_F^2, \quad (5)$$

其中 Y_k 表示第 k 个视图的软标签矩阵, $Y_{k'}$ 表示第 k' 个视图的软标签矩阵, 在该损失项的监督下, 视图的特征表示学习与聚类任务可以更加

紧密地联系起来, 视图嵌入表示的跨视图簇级一致性也得以提高。

2.4 拓扑关系重构

由于不同的视图可能包含不同的语义信息, 因此需要将它们进行有效的融合, 以获得更全面的节点表示。注意力融合可以自动地调整不同视图在节点表示中的贡献, 其核心思想是学习每个视图的权重, 本文计算第 k 个视图权重的公式如下:

$$e_k = \frac{1}{N} \sum_{i=1}^N \mathbf{h}^T \cdot \tanh(\mathbf{G} \cdot Z_k^i + \mathbf{b}), \quad (6)$$

其中 N 为节点数量, \mathbf{h} 表示一个注意力向量, \mathbf{G} 表示一个权重矩阵, Z_k^i 表示第 k 个视图中第 i 个节点的特征向量, \mathbf{b} 是一个偏差向量。然后将 e_k 归一化后得到最终的权重:

$$x_k = \text{Softmax}(e_k) = \frac{\exp(e_k)}{\sum_{k=1}^K \exp(e_k)}. \quad (7)$$

将计算出的不同视图的权重进行自适应融合, 以获得多视图的融合表示 Z :

$$Z = \sum_{k=1}^K x_k \cdot Z_k \quad (8)$$

因为多视图的融合表示 Z 利用了每个视图的拓扑结构, 因此 Z 应该能够重构原来的拓扑矩阵, 恢复多视图的结构信息。本文采用内积作为多视图解码器来重构图的结构。具体来说, 第 k 个视图的拓扑结构被重构为:

$$\hat{A}_k = \text{Sigmoid}(Z \cdot N_k Z^T), k = (1, \dots, K), \quad (9)$$

其中 \hat{A}_k 表示第 k 个视图的重构邻接矩阵, N_k 表示第 k 个视图的可训练权重矩阵。得到每个视图的重构邻接矩阵后, 令其重构误差最小化, 即最小化损失函数:

$$L_r = \sum_{k=1}^K \text{loss}(A_k, \hat{A}_k). \quad (10)$$

2.5 自监督聚类

遵循经典的端到端框架中的做法, 本文采用自监督聚类模块来指导 Z 的学习。使用 t 分布作为核度量 Z_i 和聚类质心的相似性。 Z 的软标签分布 Q 为:

$$Q_{ij} = \frac{(1 + \|Z_i - \mu_j\|^2)^{-1}}{\sum_j (1 + \|Z_i - \mu_j\|^2)^{-1}}, \quad (11)$$

其中 μ_j 为聚类质心, Z_i 表示第 i 个节点的特征向量, Q_{ij} 表示将第 i 个节点分配给第 j 个簇的概率。同时, 用目标分布 P 强化软标签分布 Q 的预测结果, P 的计算方法为:

$$P_{ij} = \frac{Q_{ij}^2 / f_i}{\sum_j Q_{ij}^2 / f_j}, \quad (12)$$

其中 $f_j = \sum_i q_{ij}$, 以归一化每个质心的损失贡献, 以防止数量过大的簇扭曲隐藏的特征空间。 P_{ij} 的值介于 0 到 1, 它可以利用 KL 散度拉近与软标签分布 Q 的距离, 从当前高置信度分配中优化聚类质心, KL 散度损失为:

$$L_c = \text{KL}(P \| Q) = \sum_i \sum_j P_{ij} \log \frac{P_{ij}}{Q_{ij}}. \quad (13)$$

每 T 次迭代后更新目标分布 P 。最后, 第 i 个节点的聚类标签可以通过 Q_{ij} 获得:

$$s_i = \arg \max_j Q_{ij}. \quad (14)$$

2.6 总目标函数

通过结合式(4), 式(5), 式(10)和式(13), 所提出的 CCRAGC 的总目标函数为:

$$L = L_r + \alpha L_{\text{feature}} + \beta L_{\text{cluster}} + L_c, \quad (15)$$

其中 α 和 β 为超参数。

CCRAGC 的具体算法过程如下。

Input: 邻接矩阵 A_1, \dots, A_k , 属性特征矩阵 X , 迭代次数 $\text{Iter1}, \text{Iter2}$, 目标分布更新间隔 T , 簇数量 C , 超参数 α, β 。

Output: 最终的聚类结果。

- (1) for iterator = 1 to Iter1 do
 \ \ 对视图嵌入表示进行预训练学习
- (2) 通过式(1)生成视图的嵌入表示 Z_k ;
- (3) 通过式(2)计算视图之间的节点相似度矩阵;
- (4) 通过式(3)、式(4)计算特征一致损失函数 L_{feature}
- (5) $Y_k = \text{clusterMLP}(Z_k)$;
 \ \ 计算每个视图的软标签矩阵
- (6) 通过式(5)计算簇级一致性损失函数 L_{cluster}
- (7) 通过式(6)、式(7)计算每个视图的权重;
- (8) 通过式(8)计算经过注意力融合生成

的多视图融合表示 Z ;

(9) 通过式(9)计算重构的图的连接关系;

(10) 通过式(10)重构损失 L_r ;

(11) 通过最小化 $L_r + \alpha L_{\text{feature}} + \beta L_{\text{cluster}}$ 更新编码器、解码器和 clusterMLP 中的可训练权重矩阵

(12) end for

(13) 基于 Z 计算初始聚类中心 μ ;

(14) for iterator = 1 to Iter2 do

\ \ 加入聚类模块微调

(15) if iterator % $T == 0$ then

(16) 通过式(11)计算软标签分布 Q ;

(17) 通过式(12)计算目标分布 P ;

(18) end if

(19) 通过式(13)计算聚类损失;

(20) 通过最小化式(15)更新整个框架;

(21) end for

(22) 从软标签分布 Q 中获得最终的聚类结果。

3 实验

在本节中, 在计算机协会 (Association for Computing Machinery, ACM)、dblp 计算机科学书目 (dblp computer science bibliography, DBLP) 和互联网电影资料库 (Internet Movie Database, IMDb) 三个多视图属性图数据集上通过对比数种经典算法以验证所提出算法的有效性。同时, 通过可视化任务、消融实验和参数敏感性分析等实验进一步体现算法的优越性。

3.1 数据集

实验数据集 ACM 是一个包含两个邻接矩阵和一个属性特征矩阵的论文网络数据集。两个邻接矩阵分别是基于论文是否由同一作者撰写 (co-paper) 和论文是否包含相同主题 (co-subject) 构成的拓扑关系图。属性特征矩阵表示了 3 025 篇论文的特征向量, 每篇论文的特征向量维度为 1 870, 特征值由论文内容的关键词决定。论文根据其研究领域可分为 3 类。

实验数据集 DBLP 是一个包含三个邻接矩阵和一个属性特征矩阵的作者合作关系数据集。三个邻接矩阵分别是基于作者是否共同撰写论

文(co-author)、作者是否在同一次会议上发表论文(co-conference)和作者是否以相同术语发表论文(co-term)构成的拓扑关系图。属性特征矩阵表示了4 057位作者的特征向量,每位作者的特征向量维度为334,特征值由描述作者的关键词决定。作者根据其研究内容可分为4类。

实验数据集IMDB是一个包含两个邻接矩阵和一个属性特征矩阵的电影网络数据集。两个邻接矩阵分别是基于电影是否由同一演员扮演(co-actor)和电影是否由同一导演执导(co-director)构成的拓扑关系图。属性特征矩阵表示了4 780部电影的特征向量,每部电影的特征向量维度为1 232,特征值由电影故事情节的关键词决定。电影根据其类型可分为3类。

三个数据集的详细介绍如表1所示。

3.2 基准算法

本文将CCRAGC算法与数种经典算法进行比较,这些方法包括两种单视图图聚类方法(大规模信息网络嵌入(Large-scale Information Network Embedding, LINE)^[18]、变分图自编码器(Variational Graph Autoencoders, GAE)^[19]),七种多视图图嵌入方法(可扩展多元网络嵌入(Scalable Multiplex Network Embedding, MNE)^[20]、基于原理的多层网络嵌入(Principled Multilayer Network Embedding, PMNE)系列^[21]、基于低秩稀疏分解的鲁棒多视图谱聚类(Robust Multi-view Spectral Clustering via Low-rank and Sparse Decomposition, RMSC)^[22]、参数加权多视图聚类(Parameter-weighted Multiview Clustering, PwMC)^[23]、自加权多视图聚类(Self-weighted Multiview Clustering, SwMC)^[23]),三种多视图属性图聚类方法(一对多图自编码器(One2Multi Graph Autoencoder, O2MA)^[12]、

O2MAC^[12]、CMGEC^[13])。

3.3 实验设置

本文的模型是基于tensorflow深度学习框架实现的,硬件环境是使用Ubuntu系统,AMD EPYC 7352 CPU,内存为40 GB。

设计CCRAGC模型结构时,多视图图编码器由两层GCN组成,两层GCN的维度分别设置为32、16;簇级一致性模块,clusterMLP(\cdot)隐藏层的维度为16,输出维度为类别数;注意力融合模块,注意力向量 h 的维度为16;自监督聚类模块,目标分布 P 的更新间隔 $T=5$ 。训练CCRAGC模型时,需要进行两轮迭代。第一轮迭代通过最小化 $L_r + \alpha L_{\text{feature}} + \beta L_{\text{cluster}}$ 更新编码器、解码器和clusterMLP中的可训练权重矩阵,在所有数据集上都进行200次迭代。对于学习率的设置,ACM和IMDB设置为0.001,DBLP设置为0.000 5。第二轮迭代通过最小化式(15)更新整个框架,在所有数据集上也进行200次迭代,从而实现对整个模型的微调,学习率与第一轮迭代的设置相同。此外,根据参数敏感性分析结果,超参数 $\alpha=0.1$, $\beta=0.1$ 。

3.4 聚类结果与分析

聚类结果使用准确率(Accuracy, Acc)^[24]、F1分数(F1 Score, F1)^[25]、标准化互信息(Normalized Mutual Information, NMI)^[26]和调整兰德指数(Adjusted Rand Index, ARI)^[24]作为评价指标。Acc衡量了聚类结果中被正确分配的样本数的比例,是最直观的评估指标。F1是一种综合考虑了精度和召回率的指标,值越高表示聚类性能越好。NMI衡量了聚类结果和真实类别之间的相似度,是一种基于信息论的指标。ARI通过比较聚类结果与真实类别之间的相似度来评估聚类性能,同时考虑了随机分

表1 实验数据集

Table 1 Experimental datasets

数据集	ACM	DBLP	IMDB
节点数	3 025	4 057	4 780
特征维度	1 870	334	1 232
邻接矩阵及其边数	co-paper: 29 281 co-subject: 2 210 761	co-author: 11 113 co-conference: 5 000 495 co-term: 6 776 335	co-actor: 98 010 co-director: 21 018
类别数	3	4	3

配的影响,是一种稳定、可靠的评估指标。所提出的CCRAGC算法与数种经典算法的性能对比见表2至表4,其中加粗的数值表示聚类性能最好,‘-’表示内存溢出。由于聚类算法的结果依赖于随机初始化,表2至表4展示在相同实验条件下算法随机运行10次的平均值。对于基准算法,直接展示其原始论文中列出的结果。

表2 在ACM数据集上的聚类结果

Table 2 Clustering results on ACM dataset

算法	Acc	F1	NMI	ARI
LINE	0.647 9	0.659 4	0.394 1	0.343 3
GAE	0.821 6	0.822 5	0.491 4	0.544 4
MNE	0.637 0	0.647 9	0.299 9	0.248 6
PMNE(n)	0.693 6	0.695 5	0.464 8	0.430 2
PMNE(r)	0.649 2	0.661 8	0.406 3	0.345 3
PMNE(c)	0.699 8	0.700 3	0.477 5	0.443 1
RMSC	0.631 5	0.574 6	0.397 3	0.331 2
PwMC	0.416 2	0.378 3	0.033 2	0.039 5
SwMC	0.383 1	0.470 9	0.083 8	0.018 7
O2MA	0.888 0	0.889 4	0.651 5	0.698 7
O2MAC	0.904 2	0.905 3	0.692 3	0.739 4
CMGEC	0.908 9	0.907 2	0.690 9	0.723 2
CCRAGC	0.921 0	0.921 8	0.735 3	0.781 5

表3 在DBLP数据集上的聚类结果

Table 3 Clustering results on DBLP dataset

算法	Acc	F1	NMI	ARI
LINE	0.868 9	0.854 6	0.667 6	0.698 8
GAE	0.885 9	0.874 3	0.692 5	0.741 0
MNE	—	—	—	—
PMNE(n)	0.792 5	0.796 6	0.591 4	0.526 5
PMNE(r)	0.383 5	0.368 8	0.087 2	0.068 9
PMNE(c)	—	—	—	—
RMSC	0.899 4	0.824 8	0.711 1	0.764 7
PwMC	0.325 3	0.280 8	0.019 0	0.015 9
SwMC	0.653 8	0.560 2	0.376 0	0.380 0
O2MA	0.904 0	0.897 6	0.725 7	0.770 5
O2MAC	0.907 4	0.901 3	0.728 7	0.778 0
CMGEC	0.910 3	0.904 2	0.723 7	0.785 9
CCRAGC	0.914 0	0.905 3	0.752 0	0.799 1

注:“—”表示该方法在处理此数据集时内存溢出。

从表2—表4可以观察到:在三个数据集上,CCRAGC在Acc、F1、NMI和ARI四个指标上基本都达到了最优,以上结果验证了CCRAGC算法的优越性。除此之外,还可以得到以下观察结果:

(1)相较于单视图图聚类方法(LINE、GAE),CCRAGC利用多个邻接矩阵,从不同的节点关系中提取更丰富的特征,提高了最终表示在聚类任务上的鲁棒性。

表4 在IMDB数据集上的聚类结果

Table 4 Clustering results on IMDB dataset

算法	Acc	F1	NMI	ARI
LINE	0.426 8	0.287 0	0.003 1	-0.009 0
GAE	0.429 8	0.406 2	0.040 2	0.047 3
MNE	0.395 8	0.331 6	0.001 7	0.000 8
PMNE(n)	0.495 8	0.390 6	0.035 9	0.036 6
PMNE(r)	0.469 7	0.318 3	0.001 4	0.011 5
PMNE(c)	0.471 9	0.388 2	0.028 5	0.028 4
RMSC	0.270 2	0.377 5	0.005 4	0.001 8
PwMC	0.245 3	0.316 4	0.002 3	0.001 7
SwMC	0.267 1	0.371 4	0.005 6	0.000 4
O2MA	0.469 7	0.422 9	0.052 4	0.075 3
O2MAC	0.450 2	0.415 9	0.042 1	0.056 4
CMGEC	0.484 4	0.510 1	0.051 4	0.046 9
CCRAGC	0.553 2	0.444 2	0.061 0	0.103 1

(2)相较于多视图图嵌入方法(MNE、PMNE系列、RMSC、PwMC、SwMC),CCRAGC利用GCN作为编码器,可以同时结构信息和属性信息进行学习,这有助于提高模型对节点表示学习的质量。

(3)相较于基于GCN的多视图属性图聚类方法(O2MA、O2MAC、CMGEC),CCRAGC利用特征级一致性模块来加强视图之间的信息交互,对于不同视图的相同节点,使其特征向量趋于一致;对于不同节点,最大化其特征向量的差异性,从而保留更丰富的节点特征,提高模型的鲁棒性。同时,CCRAGC还利用簇级一致性模块来将视图特征表示学习部分和聚类部分更紧密地结合在一起,从而使得视图的嵌入表示更加关注与聚类相关的特征,提高其在聚类任务上的准确性和有效性。此外,相对于CMGEC模型,CCRAGC利用注意力机制给视图分配可学习的权重,让多视图的嵌入表示在融合的过程中可以更根据聚类任务需要更灵活地分配对不同视图的关注度。

(4)对于ACM数据集,与次高值相比,CCRAGC在Acc、F1、NMI、ARI四个评价指标上分别提高了1.21%、1.46%、4.3%、4.21%。对于DBLP数据集,与次高值相比,CCRAGC在

Acc、F1、NMI、ARI 四个评价指标上分别提高了 0.37%、0.11%、2.33%、1.32%。对于 IMDB 数据集,与次高值相比,CCRAGC 在 Acc、NMI、ARI 三个评价指标上分别提高了 5.74%、0.86%、2.78%,在 F1 指标上没有取得最优。综合来看,在三个数据集上的聚类结果表明 CCRAGC 可以提升聚类性能。

3.5 可视化任务

为了更直观地验证 CCRAGC 算法的有效性,本节对多视图融合后的最终表示 Z 进行可视化处理。实验利用 t 分布随机邻域嵌入 (t -Distributed Stochastic Neighbor Embedding, t -SNE)^[27] 算法将 Z 投影到二维空间,以便观察簇的分布。考虑到 IMDB 数据集在评估指标上的值较低,该数据集的可视化结果难以观察到清晰的聚类效果,而 DBLP 数据集所需的计算时间相对较长。因此,为了更有效地展示和分析聚类结果,本文选择 ACM 数据集进行可视化分析。图 2 展示了 CCRAGC 算法在 ACM 数据集上执行聚类任务时的可视化结果,不同颜色代表不同的簇。可以观察到,随着迭代的进行,簇的结构逐渐变得更加清晰。

3.6 消融实验

本节通过消融实验进一步证明特征级一致性损失 $L_{feature}$ 和簇级一致性损失 $L_{cluster}$ 的有效性,实验结果如图 3 所示。(a)到(c)分别为在 ACM、DBLP、IMDB 三个数据集上的聚类结果。图中有四个不同颜色的条形图,对应图例中的四个简称,分别代表的含义是:CCRAGC-o 表示没有 $L_{feature}$ 和 $L_{cluster}$ 约束的 CCRAGC; CCRAGC-c 表示没有 $L_{feature}$ 约束的 CCRAGC; CCRAGC-f 表示没有 $L_{cluster}$ 约束的 CCRAGC; CCRAGC 表示同时使用 $L_{feature}$ 和 $L_{cluster}$ 的模型。

从图 3 可以观察到:CCRAGC-o 的聚类性能最低,在 CCRAGC-o 的基础上,添加特征级一致性损失 $L_{feature}$ 或者簇级一致性损失 $L_{cluster}$ 都可以提升 CCRAGC-o 的聚类性能,特征级一致性损失 $L_{feature}$ 对聚类性能的提升更显著。同时添加特征级一致性损失 $L_{feature}$ 和簇级一致性损失 $L_{cluster}$ 时,模型的聚类性能最优。实验结果验证了这两个约束的有效性。

3.7 参数敏感性分析

由于 DBLP 所需计算时间较长,为了节省时间和计算资源,本节仅在 ACM 和 IMDB 数据

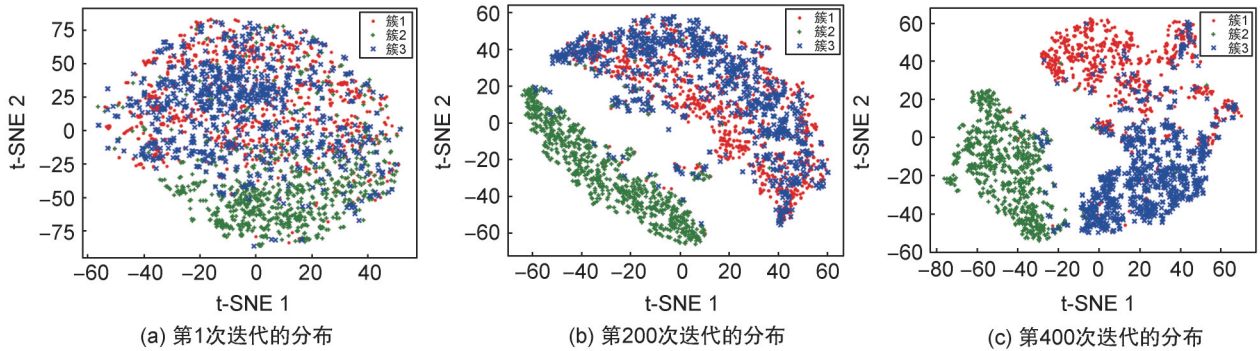


图2 CCRAGC算法在ACM数据集上不同迭代次数的可视化结果

Fig. 2 Visualization of different iterations of the CCRAGC algorithm on the ACM dataset

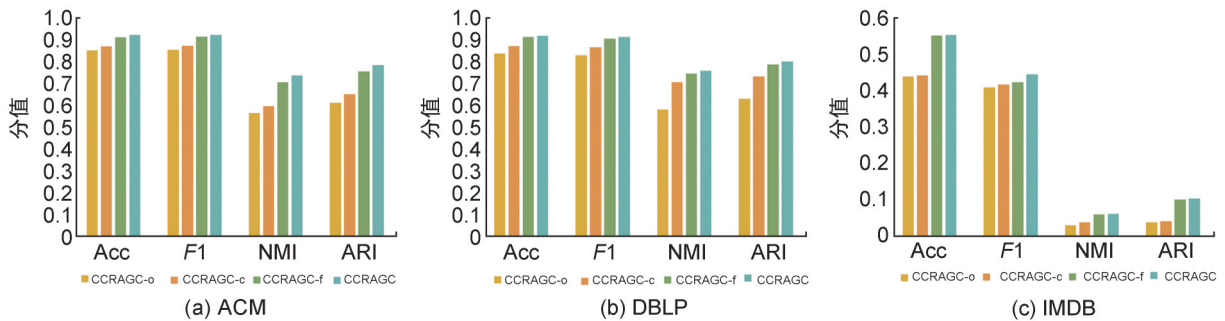


图3 CCRAGC及其变体在三个数据集上的聚类性能比较

Fig. 3 Comparison of clustering performances of CCRAGC and its variants on three datasets

集上分析超参数 α 和 β 对聚类性能 (Acc、F1) 的影响, α 反映了特征级一致性损失的重要性, β 反映了簇级一致性损失的重要性, 实验结果如图 4 所示。在实验中, 对 α 和 β 的调整范围为 $[0.01, 0.05, 0.1, 0.5, 1]$, 使用控制变量法来分析其影响。

实验结果显示, 在 ACM 数据集上, 当 α 和 β 在 $0.01 \sim 1$ 之间变化时, 相应的评价指标 Acc、F1 基本保持不变, 这说明 α 和 β 在合理范围的波动时, 基本不会对聚类性能造成影响; 在 IMDB 数据集上, 当 α 在 $0.05 \sim 0.1$ 、 β 在 $0.1 \sim 1$ 之间变化时, 模型聚类性能最稳定。当 $\alpha = 0.1, \beta = 0.1$ 时, 聚类性能达到最优。总体来说, CCRAGC 在大多数参数组合下可以获得不错的聚类性能, 这说明算法 CCRAGC 对特征级一致性损失的权重 α 和簇级一致性损失的权重 β 的敏感性较低。

4 结论

基于 GCN 的多视图聚类算法利用 GCN 作为编码器, 能很好地融合图拓扑结构信息和属性信息。但是, 每个视图的 GCN 编码器是相互独立的, 这会导致视图嵌入表示的跨视图特征一致性不足; 此外, 虽然视图嵌入表示学习和聚类任务是集成在一个框架中学习的, 但训练过程还是分为特征表示学习和加入自监督模块的微调两个阶段, 这会造成学到的视图嵌入表示不能很好地符合聚类任务的要求, 即跨视图簇级一致性不足。本文提出了一种跨视图一致性表示的多视图属性图聚类算法 CCRAGC, 它通过计算视图之间的相似度矩阵并对其添加约束来指导视图嵌入表示的学习过程。此外, 为了不让视图嵌入表示的学习过程和聚类任务更紧密, 将学到的视图嵌入表示映射到子空间并对其软标签矩阵进行对齐。本文在 ACM、

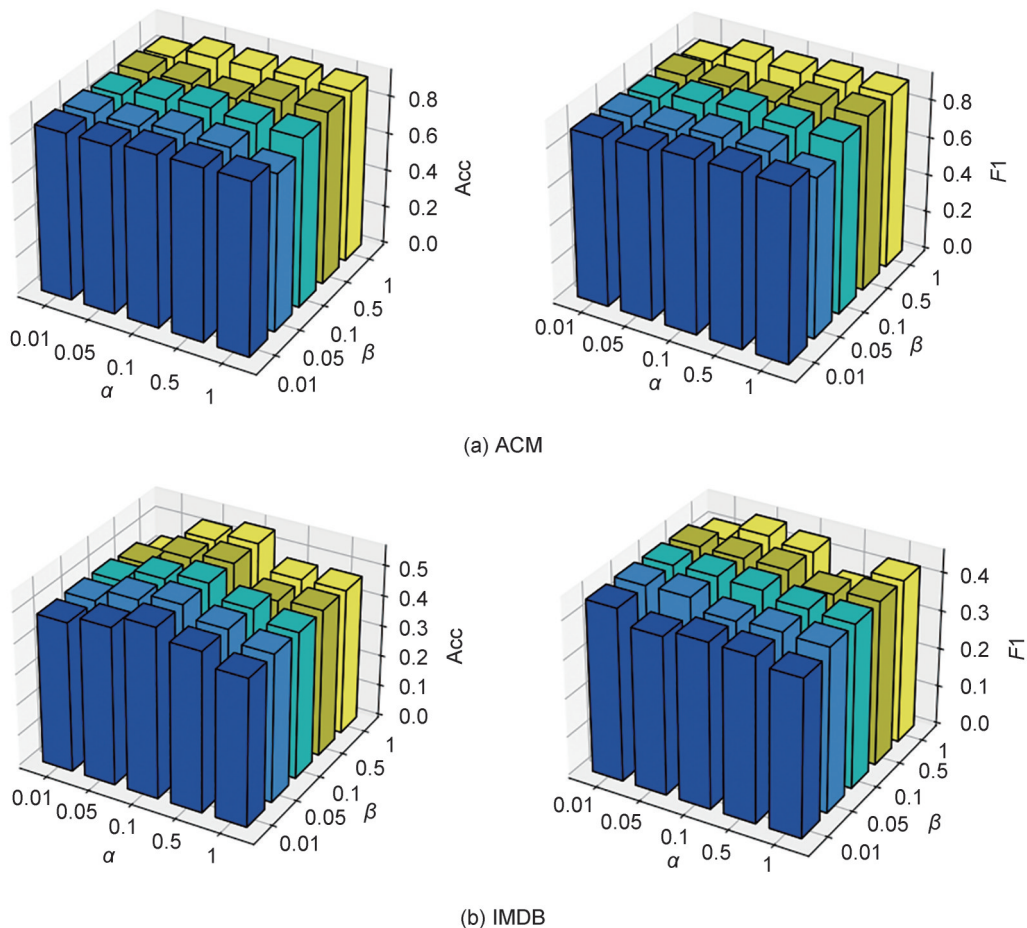


图4 超参数 α 和 β 在 ACM、IMDB 数据集上的敏感性分析

Fig. 4 Sensitivity analysis of hyperparameters α and β on ACM and IMDB datasets

DBLP 和 IMDB 三个广泛使用的数据集上对提出的算法 CCRAGC 进行了实验,证明了 CCRAGC 算法的有效性。尽管如此,所提出的算法还存在一些可以改进的地方,目前算法主要适用于无向图,且要求节点的连接关系是静态的。然而,现实世界中的图数据可能具有其他特性,例如动态性、有向性等。因此,将所提出的算法扩展到其他类型的图,进一步提高聚类的准确性和鲁棒性,是值得深入研究的工作。

参考文献:

- [1] XIE J Y, GIRSHICK R, FARHADI A. Unsupervised Deep Embedding for Clustering Analysis[C]//Proceedings of the 33rd International Conference on International Conference on Machine Learning-Volume 48. New York: ACM, 2016: 478–487. DOI: 10.5555/3045390.3045442.
- [2] GUO X F, GAO L, LIU X W, *et al.* Improved Deep Embedded Clustering with Local Structure Preservation[C]//Proceedings of the 26th International Joint Conference on Artificial Intelligence. New York: ACM, 2017: 1753–1759. DOI: 10.5555/3172077.3172131.
- [3] LI Z Y, WANG Q Q, TAO Z Q, *et al.* Deep Adversarial Multi-view Clustering Network[C]//Proceedings of the 28th International Joint Conference on Artificial Intelligence. New York: ACM, 2019: 2952–2958. DOI: 10.5555/3367243.3367449.
- [4] XU J, REN Y Z, LI G F, *et al.* Deep Embedded Multi-view Clustering with Collaborative Training[J]. *Inform Sciences*, 2021, **573**: 279–290. DOI: 10.1016/j.ins.2020.12.073.
- [5] XU J, TANG H Y, REN Y Z, *et al.* Multi-level Feature Learning for Contrastive Multi-view Clustering[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2022: 16030–16039. DOI: 10.1109/CVPR52688.2022.01558.
- [6] TANG X L, TANG X, WANG W L, *et al.* Deep Multi-view Sparse Subspace Clustering[C]//Proceedings of the 2018 VII International Conference on Network, Communication and Computing. New York: ACM, 2018: 115–119. DOI: 10.1145/3301326.3301391.
- [7] ANDREW G, ARORA R, BILMES J, *et al.* Deep Canonical Correlation Analysis[J]. *30th Int Conf Mach Learn ICML 2013, 2013(PART 3)*: 2284–2292.
- [8] ZHU P F, HUI B Y, ZHANG C Q, *et al.* Multi-view Deep Subspace Clustering Networks[J]. *IEEE Trans Cybern*, 2024,**54**(7):4280–4293. DOI: 10.1109/TCYB.2024.3372309.
- [9] WANG J, WU B, REN Z W, *et al.* Multi-scale Deep Multi-view Subspace Clustering with Self-weighting Fusion and Structure Preserving[J]. *Expert Syst Appl*, 2023, **213**: 119031. DOI: 10.1016/j.eswa.2022.119031.
- [10] LIU J, CAO F Y, JING X C, *et al.* Deep Multi-view Graph Clustering Network with Weighting Mechanism and Collaborative Training[J]. *Expert Syst Appl*, 2024, **236**: 121298. DOI: 10.1016/j.eswa.2023.121298.
- [11] CHENG J F, WANG Q Q, TAO Z Q, *et al.* Multi-view Attribute Graph Convolution Networks for Clustering[C]//Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence. New York: ACM, 2021: 2973–2979. DOI: 10.5555/3491440.3491851.
- [12] FAN S H, WANG X, SHI C, *et al.* One2Multi Graph Autoencoder for Multi-view Graph Clustering[C]//Proceedings of The Web Conference 2020. New York: ACM, 2020: 3070–3076. DOI: 10.1145/3366423.3380079.
- [13] WANG Y M, CHANG D X, FU Z Q, *et al.* Consistent Multiple Graph Embedding for Multi-view Clustering[J]. *IEEE Trans Multimed*, 1008, **25**: 1008–1018. DOI: 10.1109/TMM.2021.3136098.
- [14] BO D Y, WANG X, SHI C, *et al.* Structural Deep Clustering Network[C]//Proceedings of The Web Conference 2020. New York: ACM, 2020: 1400–1410. DOI: 10.1145/3366423.3380214.
- [15] TU W X, ZHOU S H, LIU X W, *et al.* Deep Fusion Clustering Network[J]. *Proc AAAI Conf Artif Intell*, 2021, **35**(11): 9978–9987. DOI: 10.1609/aaai.v35i11.17198.
- [16] RAHMAN M K, AGRAWAL A, AZAD A. MarkovGNN: Graph Neural Networks on Markov Diffusion[C]//Proceedings of the Web Conference 2022. New York: ACM, 2022: 1019–1029. DOI: 10.1145/3487553.3524713.
- [17] KIPF T N, WELLING M. Semi-supervised Classification with Graph Convolutional Networks[EB/OL]. (2016-09-09) [2025-01-07]. <https://doi.org/10.48550/arXiv.1609.02907>.
- [18] 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: ACM, 2015: 1067–1077. DOI: 10.1145/2736277.2741093.
- [19] KIPF TN, WELLING M. Variational Graph Autoencoders[EB/OL]. (2016-11-21)[2025-01-07]. <https://doi.org/10.48550/arXiv.1611.07308>.
- [20] ZHANG H M, QIU L W, YI L L, *et al.* Scalable Multiplex Network Embedding[C]//Proceedings of the 27th International Joint Conference on Artificial Intelligence. New York: ACM, 2018: 3082–3088. DOI: 10.5555/3304889.3305089.
- [21] LIU W Y, CHEN P Y, YEUNG S, *et al.* Principled Multi-

- layer Network Embedding[C]//2017 IEEE International Conference on Data Mining Workshops (ICDMW). New York: IEEE, 2017: 134–141. DOI: 10.1109/ICDMW.2017.23.
- [22] XIA R K, PAN Y, DU L, *et al.* Robust Multi-view Spectral Clustering via Low-rank and Sparse Decomposition[C]// Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence. New York: ACM, 2014: 2149–2155. DOI: 10.5555/2892753.2892850.
- [23] NIE F P, LI J, LI X L. Self-weighted Multiview Clustering with Multiple Graphs[C]//Proceedings of the 26th International Joint Conference on Artificial Intelligence. New York: ACM, 2017: 2564–2570. DOI: 10.5555/3172077.3172245.
- [24] LIANG J Y, BAI L, DANG C Y, *et al.* The k -means-type Algorithms Versus Imbalanced Data Distributions[J]. *IEEE Trans Fuzzy Syst*, 2012, **20**(4): 728–745. DOI: 10.1109/TFUZZ.2011.2182354.
- [25] BOUYER A, ROGHANI H. LSMD: A Fast and Robust Local Community Detection Starting from Low Degree Nodes in Social Networks[J]. *Future Gener Comput Syst*, 2020, **113**: 41–57. DOI: 10.1016/j.future.2020.07.011.
- [26] STREHL A, GHOSH J. Cluster Ensembles-A Knowledge Reuse Framework for Combining Multiple Partitions[J]. *J Mach Learn Res*, 2002, **3**: 583–617. DOI: 10.1162/153244303321897735.
- [27] VAN DER MAATEN L, HINTON G. Visualizing Data Using t-SNE[J]. *J Mach Learn Res*, 2008, **9**: 2579–2605.