SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

February 2026, Vol. 69, Iss. 2, 122102:1–122102:11
https://doi.org/10.1007/s11432-024-4467-3

Multi-topology contrastive graph representation learning

Yu XIE[1], Jie JIA[1], Chao WEN[1], Deyu LI[1] & Ming LI[2*]

[1]*Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan 030006, China*
[2]*Zhejiang Key Laboratory of Intelligent Education Technology and Application, Zhejiang Normal University, Jinhua 321004, China*

Received 8 May 2024/Revised 7 March 2025/Accepted 4 June 2025/Published online 19 September 2025

Abstract Self-supervised graph representation learning has received significant attention by virtue of tackling the label scarcity issue in graph data. However, prior methods underutilize graph structures with multiple forms and subgraph structures at different scales, thus failing to deeply explore the diversity and complexity of graph data. In this paper, we present a novel multi-topology contrastive graph representation learning (MCGRL) framework, which aims to improve the effectiveness of node representation learning by capturing multi-granularity information in different topologies. Specifically, we generate multiple topologies from different viewpoints and then contrast the learned multi-granularity node representations in different topologies to preserve the rich multi-topology interactions and complementary information. Drawing upon an in-depth scrutiny of the classical Intersection over Union, we propose a subgraph-level similarity constraint (SIoU) to explore the semantic consistency among multiple topologies and dynamically characterize different-granularity subgraph information. Empirical experiments on real-world datasets demonstrate the effectiveness of our proposed method compared with current state-of-the-art methods.

Keywords graph representation learning, self-supervised contrastive learning, graph neural networks, multiple topologies, semantic consistency

Citation Xie Y, Jia J, Wen C, et al. Multi-topology contrastive graph representation learning. Sci China Inf Sci, 2026, 69(2): 122102, https://doi.org/10.1007/s11432-024-4467-3

1 Introduction

Recently, graph neural networks (GNNs) [1–3] have been widely adopted in efficiently modeling graph-structured data and play a paramount role in practical application scenarios, such as finding interpersonal relations in a social network [4], calculating drug similarities in drug discovery [5], and predicting congestion in a transport network [6,7]. Most existing GNNs adopt a supervised or semi-supervised learning paradigm that leverages labeled information within graph data to complete specific downstream tasks [8]. However, labeled graph data are not only expensive and labor-intensive to collect in the real world, but also require high-quality annotation. For example, manual annotations of user relationships in social networks and protein-protein interactions in biological networks require substantial labor and time [9,10]. As a result, supervised or semi-supervised approaches exhibit significant limitations in practical applications.

To alleviate these problems, self-supervised graph representation learning, particularly graph contrastive learning, has attracted widespread attention because of its remarkable performance on a spectrum of graph tasks [11]. Deep graph infomax (DGI) [12] is a pioneering approach in graph contrastive learning that relies on maximizing the mutual information [13] between node representations and corresponding graph-level representations. GRACE [14] focuses on contrasting views at the node level and promotes consistency between node representations in two augmented views. Motivated by BYOL (bootstrap your own latent) [15], bootstrapped graph representation learning (BGRL) [16] introduces a self-supervised graph representation learning approach that eliminates the necessity for negative pairs. Based on DGI and multi-view graph representation learning (MVGRL) [17], Graph group discrimination (GGD) [18] presents a novel learning paradigm that can directly discriminate between different groups of node samples instead of maximizing the mutual information between similar instances, as in the above methods.

Despite significant advances in graph-contrastive learning, the following drawbacks exist. During the data collection process, acquisition devices, environmental conditions and so on may introduce noise, deformations, and

* Corresponding author (email: mingli@zjnu.edu.cn)

other impacts into the obtained data, leading to partial information loss. Although numerous contrastive learning methods employ data augmentation to mitigate information loss and enhance model generalization capabilities, augmented graphs do not guarantee semantic consistency with the original graphs. Meanwhile, the lack of consideration of associations between graph topologies from different perspectives hampers the mining of potential information within multiple topologies, resulting in an inadequate characterization of graph data, making it challenging to deeply explore the diversity and complexity present in graph data.

To address these issues, we introduce a novel multi-topology contrastive graph representation learning (MCGRL) method. The main idea is to contrast the multi-granularity information that exists in diverse topologies to achieve efficient representation learning for unlabeled graph data. First, we generate different types of topologies representing different views of the graph. By synthesizing insights from diverse perspectives, it not only compensates for deficiencies within each viewpoint but also comprehensively captures the diversity of the graph structure. Then, a global node representation learning module consisting of an online encoder and a target encoder is developed for each view to capture the global structure information in multiple topologies. To further explore the semantic consistency among multiple topologies, a subgraph-level similarity constraint (SIoU) is proposed that leverages pseudo-labels assigned through a classical clustering method to dynamically regulate intra-class compactness and inter-class separation. By adjusting the structural relationships among nodes, SIoU facilitates the adaptive formation and continuous optimization of subgraph structures, thereby enhancing the quality of learned representations. Subsequently, the refined subgraph connectivity is fed into a shared graph neural network that learns local node representations by aggregating neighborhood information within the adjusted subgraph structure and ensures that fine-grained local dependencies are effectively preserved. Finally, we construct a multi-topology contrastive graph representation learning loss to learn the final node representations with semantic coherence from diverse views. Contrastive loss enforces semantic coherence across multiple views, ensuring that nodes with similar semantics remain close while maintaining sufficient discrimination between distinct clusters. MCGRL effectively captures both local and global structural information across multiple topologies and yields discriminative node representations.

In summary, we make three contributions.

- We propose a multi-topology contrastive graph representation learning method that aims to effectively extract global and local information from multiple views.
- A subgraph-level similarity constraint named SIoU is designed to systematically explore the intrinsic interaction information among diverse topologies and extract the semantic consistency preserved node representations.
- We demonstrate the superiority of MCGRL by theoretical analysis. Furthermore, extensive experiments conducted on a series of benchmark datasets show that our method achieves superior or comparable results to state-of-the-art self-supervised graph representation learning methods.

The remainder of this paper is organized as follows. In Section 2, we review related work. In Section 3, the technical details of our proposed framework are introduced, and we provide a theoretical analysis of the expressive power of MCGRL in Appendix B. Section 4 presents the experimental results, and Section 5 concludes the paper.

2 Related work

2.1 Multi-topology learning on graphs

In the real world, graph structures with varying topologies exhibit distinct relationships among nodes. Generating node representations within a multi-topology learning framework allows us to capture more comprehensive graph information, which is crucial for various graph-related machine-learning tasks. Recently, multi-dimensional graph convolutional networks (mGCN) [19] delve into the varying interactions among nodes across dimensions and the intrinsic connections for the same node in different dimensions. Adaptive multi-channel graph convolutional networks (AM-GCN) [20] note that the capacity of graph neural networks to fuse topological structures and node features falls short of an optimal or satisfactory level and propose an adaptive fusion mechanism to improve the model's ability to learn effective graph representations and boost performance on node-level tasks. GPS [21] automatically generates multi-scale positive views using graph-pooling modules, capturing hierarchical information at different granularities. By incorporating the generated views into a joint contrastive learning framework, GPS improves the model's ability to generalize and learn more robust graph representations. Multi-GCN [22] empirically demonstrates that incorporating multi-view information into the learning process can achieve promising performance, demonstrating the advantage of using multiple perspectives of the graph over a single view. Additionally, MAGCN [23] theoretically further elucidates why multi-view methods surpass single-view methods. Although the aforementioned methods have undergone rigorous empirical verification or theoretical analysis in multi-topology

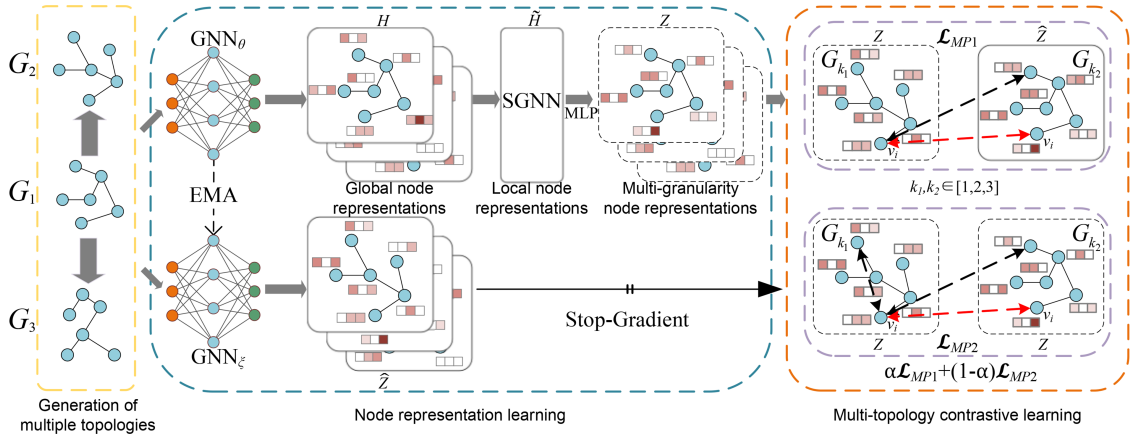

Figure 1 (Color online) Overview of our proposed MCGRL. The graph-level GNNs (GGNN) consisting of GNN_θ and GNN_ξ are designed to obtain the global node representations from multiple topologies. Then, the local node representations are learned by the subgraph-level GNNs (SGNN) module, as illustrated in Figure 2. Finally, the entire network undergoes joint optimization via the overall multi-topology contrastive graph representation learning loss. Here, the number of topologies is set to 3 in this paper for brevity, MLP denotes the projection layer, red and black arrows represent positive and negative sample pairs, respectively.

learning, most of these approaches are anchored in semi-supervised settings. Under current circumstances, there remains a dearth of consideration on how to introduce the multi-topology learning paradigm into the domain of self-supervised graph representation learning.

2.2 Self-supervised graph contrastive learning

As a self-supervised deep learning paradigm for graphs, graph contrastive learning aims to maximize the similarity between positive samples while increasing the dissimilarity among negative samples in graphs. Through self-supervised graph contrastive learning, we can find a reasonable latent space for all nodes in a graph, which is similar to subspace clustering [24–27]. Early methods, exemplified by DGI [12], extend the principles of Deep InfoMax [28] to the graph domain by maximizing the mutual information between local node features and global graph features, whereas InfoGraph [29] further refines this approach to maximize the mutual information of representations between the graph-level and different substructure-levels. MVGRL [17] maximizes the mutual information between cross-view representations of nodes and graphs, further promoting graph representation learning. Most recently, BGRL [16] has eliminated the need for negative samples by using an exponential moving average (EMA) [15]. Multi-level graph contrastive prototypical clustering (MLG-CPC) [30] introduces an end-to-end clustering framework to alleviate the semantic bias, thereby enabling more effective representation learning. Multi-scale subgraph contrastive learning (MSSGCL) [31] generates global and local views at different scales using subgraph sampling, and constructs multiple contrastive relationships based on semantic associations. However, the aforementioned methods exhibit shortcomings in characterizing the interrelationships and semantic consistency between different views. Based on the above foundations, we explore in depth the interaction information among topologies and the correlation between different topologies from the perspective that the same nodes should have the same semantic characteristics under different topologies.

3 Methodology

3.1 Notations and problem definition

Suppose an undirected graph $G = \{V, E, A, X\}$, where $V = \{v_1, \ldots, v_N\}$ represents the set of nodes in which N is the number of nodes, $E \subseteq V \times V$ represents the set of edges where $e_{ij} = 1$ indicates an edge exists between nodes v_i and v_j, otherwise, $e_{ij} = 0$. A implies the adjacency matrix and X denotes the node feature matrix. Multi-topology graphs are defined as $\mathcal{G} = \{G_1, G_2, \ldots, G_K\}$ where $G_k = \{V, E^k, A^k, X^k\}$ and K is the number of topologies. In this paper, we attempt to learn multi-topology and multi-granularity information-preserving node representations Z for downstream tasks, such as node classification. The notations used in this paper are illustrated in Appendix A.

3.2 Overall framework

As shown in Figure 1, our proposed MCGRL method consists mainly of three components: generation of multiple topologies, node representation learning, and multi-topology contrastive learning. To characterize the rich topological structural information in a graph, we first generate multiple adjacency matrices. Subsequently, for each topology, we design GGNN with an EMA and SGNN to learn the global and local node representations, respectively. To explore the interaction information among multiple topologies, we devise a novel similarity constraint in SGNN to dynamically characterize semantic consistency. Finally, we introduce the multi-topology contrastive graph representation learning loss for learning informative and discriminative node representations.

3.3 Generation of multiple topologies

In this subsection, we explore distinct methodologies for generating adjacency matrices from different perspectives, such as cosine similarity calculation of node attributes, graph diffusion, and edge perturbation. For computational efficiency, we adopt a strategy in which a node index is randomly chosen from the adjacency matrix, functioning as the splitting point for cropping the original graph into a standardized subgraph size for batch training. Concurrently, attribute masking is implemented to improve model generalization.

3.3.1 *Cosine similarity of node attributes*

By calculating the cosine similarity between node attributes, we derive a similarity matrix $S \in \mathbb{R}^{N\times N}$:

$$S = \cos(X, X), \quad s_{ij} = \frac{X_i^{\mathrm{T}} \cdot X_j}{||X_i|| \cdot ||X_j||}, \tag{1}$$

where s_{ij} denotes the similarity score between nodes v_i and v_j. To generate the topology, the edges for each node are established by selecting the top m nodes with the highest similarity, where X_i denotes the feature vector of node v_i.

3.3.2 *Graph diffusion*

Graph diffusion is adopted to morph the original graph structure and generate a new topology without compromising global information and inherent semantics. The general graph diffusion process is defined as [32]:

$$\mathcal{T} = \sum_{i=0}^{\infty} \mu_i M^i, \tag{2}$$

where μ is the weighting coefficient controlling the local and global signal distributions subject to $\sum_{i=0}^{\infty} \mu_i = 1$, and $M \in \mathbb{R}^{N\times N}$ represents a generalized transition matrix derived from the adjacency matrix A. Here, we employ personalized pagerank (PPR) as an illustrative example of graph diffusion, articulated as follows:

$$\mathcal{T}_{\mathrm{PPR}}(A) = \lambda(I - (1-\lambda)D^{-1/2}AD^{-1/2})^{-1}, \tag{3}$$

where λ serves as a parameter governing the teleport probability in a random walk, and D and I represent the degree and identity matrices of A, respectively.

We refrain from providing an exhaustive discussion of other topics such as edge perturbation, subgraph sampling, and attribute masking. Interested readers are encouraged to explore the relevant literature and its references for a detailed understanding of topology generation techniques [33–35].

3.4 Node representation learning

To obtain multi-granularity node representations, we first extract the global node representations from GGNN with an exponential moving average for each topology. The node representation matrix learned by GNN_θ is defined as $H^k = \mathrm{GNN}_\theta(A^k, X^k)$ and that learned by GNN_ξ is $\widehat{Z}^k = \mathrm{GNN}_\xi(A^k, X^k)$, where k represents the k-th topology and $k \in [1, \ldots, K]$. GNN_θ and GNN_ξ comprise a GNN layer followed by a linear layer, and the parameters of GNN_ξ are updated via EMA.

To investigate the interconnectivity and semantic consistency among multiple topologies, we present an SGNN module aimed at extracting the local information of different topologies, as depicted in Figure 2. The specific procedure unfolds as follows. We first assign pseudo-labels to each node via a clustering approach to the

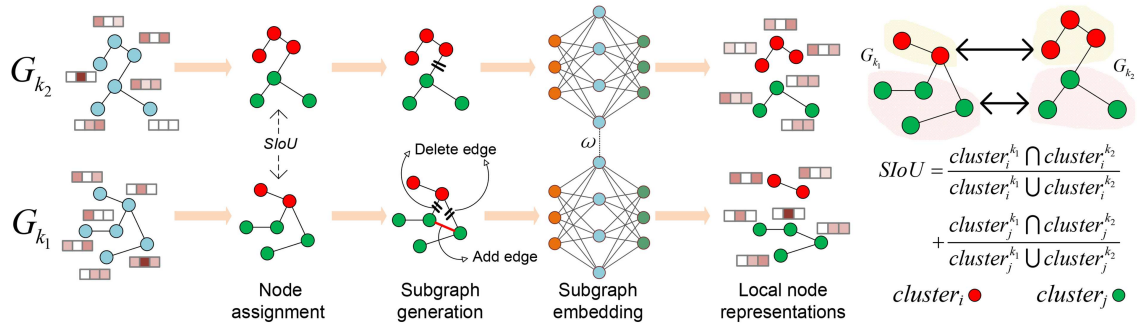

Figure 2 (Color online) Architecture of SGNN. Pseudo-labels are assigned to nodes via the clustering algorithm and constrained by SIoU, which drives the construction of a new adjacency matrix that will be input to a shared GNN for learning the fine-grained node representations. The diagram on the right provides a detailed illustration of the computational procedure for SIoU.

node representations H learned from GNN_θ, which can partition the nodes into $|C|$ distinct clusters, denoted as $C = \{\mathrm{cluster}_1, \mathrm{cluster}_2, \ldots, \mathrm{cluster}_{|C|}\}$, where each cluster represents a set of nodes with similar feature-based characteristics. Next, we compute the pairwise similarity between nodes as $s_{ij} = \mathrm{sim}(h_i, h_j)$ and use it as a criterion for performing edge modifications. Specifically, for nodes belonging to the same pseudo-label cluster, the existing edges are preserved and additional edges are introduced between node pairs with high similarity to enhance the connectivity of the corresponding subgraph. This step ensures that nodes within the same cluster remain well-connected, thereby strengthening the intra-cluster relationships.

$$\mathcal{A}^s_{ij} = \begin{cases} 1, & A_{ij} = 1 \text{ and } \mathrm{cluster}(v_i) = \mathrm{cluster}(v_j), \\ 1, & A_{ij} = 0 \text{ and } \mathrm{cluster}(v_i) = \mathrm{cluster}(v_j) \\ & \text{and } j = \mathrm{argmax}_{j \neq i} s_{ij}, \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

For adjacent nodes v_i and v_j, where $A_{ij} = 1$ but with different pseudo-labels, that is, $\mathrm{cluster}(v_i) \neq \mathrm{cluster}(v_j)$, we apply an edge-pruning operation. In this step, edges are removed if their similarity $s_{ij} \leqslant \epsilon$, because such weak connections may introduce noise. After completing these edge operations, we obtain the final refined adjacency matrix $\mathcal{A}$, which better reflects the intrinsic cluster structures, while maintaining the essential connectivity properties. We can then update the node representations for the k-th topology using a shared graph neural network as

$$\tilde{H}^k = \mathrm{GNN}_\omega(\mathcal{A}^k, H^k). \tag{5}$$

To ensure semantic consistency during the clustering process for different topologies, we introduce a novel SIoU metric based on the intersection and union operations of their clustering results, which is inspired by the classical Intersection over Union in object detection [36]. For the i-th cluster in the k_1-th topology and the j-th cluster in the k_2-th topology, their SIoU score is defined as

$$\mathrm{SIoU}(\mathrm{cluster}_i^{k_1}, \mathrm{cluster}_j^{k_2}) = \frac{\mathrm{cluster}_i^{k_1} \cap \mathrm{cluster}_j^{k_2}}{\mathrm{cluster}_i^{k_1} \cup \mathrm{cluster}_j^{k_2}}. \tag{6}$$

For any cluster, its corresponding cluster in another topology is defined as the cluster that has the highest SIoU score with it. Then, SIoU is explored to constrain the semantic consistency between different topologies as follows:

$$\mathcal{L}_{\mathrm{SIoU}}(Z(\mathrm{cluster}_i^{k_1}), Z(\mathrm{cluster}_j^{k_2})) = -\log \frac{\mathrm{dist}(Z(\mathrm{cluster}_i^{k_1}), Z(\mathrm{cluster}_j^{k_2}))}{\sum_{c=1}^{|C|} \mathrm{dist}(Z(\mathrm{cluster}_i^{k_1}), Z(\mathrm{cluster}_c^{k_2}))}, \tag{7}$$

$$\mathcal{L}_{\mathrm{SIoU}} = \frac{1}{K(K-1)|C|} \sum_{k_1=1, k_2 \neq k_1}^{K} \sum_{i=1}^{|C|} \mathcal{L}_{\mathrm{SIoU}}(Z(\mathrm{cluster}_i^{k_1}), Z(\mathrm{cluster}_j^{k_2})), \tag{8}$$

where $\mathrm{cluster}_j^{k_2}$ is assumed to be the corresponding cluster of $\mathrm{cluster}_i^{k_1}$ in the k_2-th topology ($k_2 = [1, \ldots, K]$ and $k_2 \neq k_1$) for brevity, and the subgraph representation $Z(\cdot) \in R^{d \times 1}$ for a cluster is calculated by

$$Z(\mathrm{cluster}_i^k) = \mathrm{readout}(H^k(\mathrm{cluster}_i^k)), \tag{9}$$

where $H^k(\text{cluster}_i^k)$ denotes the node representation matrix of the i-th cluster in the k-th topology. d refers to the dimension of representations, and readout$(\cdot)$ is a pooling function, such as the average pooling [37].

Finally, we obtain the multi-granularity node representations by weighting the global and local node representations for the k-th topology and projecting them through an MLP layer:

$$Z^k = \text{MLP}(\eta H^k + (1-\eta)\tilde{H}^k), \tag{10}$$

where η is a hyperparameter ranging from 0 to 1.

3.5 Multi-topology contrastive learning

To preserve the complex associations and differences in multiple topologies, we conduct multi-topology contrastive learning, whose loss includes $\mathcal{L}_{\text{MP1}}$ and $\mathcal{L}_{\text{MP2}}$.

$\mathcal{L}_{\text{MP1}}$ contrasts the multi-granularity node representations from one topology with the global node representations obtained by GNN_ξ from other topologies.

$$\mathcal{L}_{\text{MP1}}^{k_1k_2}(v_i) = -\log \frac{\exp(\text{sim}(z_{v_i}^{k_1}, \widehat{z}_{v_i}^{k_2}))}{\sum_{j=1}^{N}\exp(\text{sim}(z_{v_i}^{k_1}, \widehat{z}_{v_j}^{k_2}))}, \tag{11}$$

where $z_{v_i}^{k_1} \in Z^{k_1}$, $\widehat{z}_{v_i}^{k_2}, \widehat{z}_{v_j}^{k_2} \in \widehat{Z}^{k_2}$. Because our study involves multiple topologies for graph contrastive learning, the final formula for $\mathcal{L}_{\text{MP1}}$ is formulated as follows:

$$\mathcal{L}_{\text{MP1}} = \frac{1}{K(K-1)N}\sum_{k_1=1,k_2\neq k_1}^{K}\sum_{i=1}^{N}\mathcal{L}_{\text{MP1}}^{k_1k_2}(v_i). \tag{12}$$

$\mathcal{L}_{\text{MP2}}$ contrasts the multi-granularity node representations within the same topology or between different topologies. We start with intra-topological contrastive learning, which designates other nodes within a topology as negative samples and introduces the same nodes from different topologies as positive samples.

$$\begin{aligned}\mathcal{L}_{\text{intra}}^{k_1k_2}(v_i) &= -\log\frac{\exp(\text{sim}(z_{v_i}^{k_1}, z_{v_i}^{k_2}))}{\exp(\text{sim}(z_{v_i}^{k_1}, z_{v_i}^{k_2})) + \phi},\\ \phi &= \sum_{j=1,j\neq i}^{N}\exp(\text{sim}(z_{v_i}^{k_1}, z_{v_j}^{k_1})),\end{aligned} \tag{13}$$

where ϕ represents the similarity summation of all negative pairs computed for node v_i within the same topology. Then, similar to $\mathcal{L}_{\text{MP1}}$, inter-topological contrastive learning is defined as

$$\mathcal{L}_{\text{inter}}^{k_1k_2}(v_i) = -\log\frac{\exp(\text{sim}(z_{v_i}^{k_1}, z_{v_i}^{k_2}))}{\sum_{j=1}^{N}\exp(\text{sim}(z_{v_i}^{k_1}, z_{v_j}^{k_2}))}. \tag{14}$$

By integrating inter- and intra-topological contrastive learning, we can calculate the $\mathcal{L}_{\text{MP2}}$ loss:

$$\mathcal{L}_{\text{MP2}} = \frac{1}{K(K-1)N}\sum_{k_1=1,k_2\neq k_1}^{K}\sum_{i=1}^{N}(\mathcal{L}_{\text{intra}}^{k_1k_2}(v_i) + \mathcal{L}_{\text{inter}}^{k_1k_2}(v_i)). \tag{15}$$

Finally, the overall objective function for our multi-topology contrastive graph-representation learning can be expressed as follows:

$$\mathcal{L} = \alpha\mathcal{L}_{\text{MP1}} + (1-\alpha)\mathcal{L}_{\text{MP2}} + \beta\mathcal{L}_{\text{SIoU}}, \tag{16}$$

where α and β are hyperparameters for balancing the multi-topology contrastive loss and semantic consistency loss. Appendix B provides a rigorous mathematical analysis of the expressive power of MCGRL via information theory and examines the factors contributing to its superior performance over traditional graph contrastive representation learning methods.

Table 1 Statistics of the datasets.

Dataset	#Nodes	#Edges	#Content words	#Features	#Classes	#Label rate	#Testing nodes
Cora	2708	5429	3880564	1433	7	0.052	1000
CiteSeer	3327	4732	12274336	3703	6	0.036	1000
PubMed	19717	44338	9858500	500	3	0.003	1000

4 Experiments

This section describes the experimental setups. We then compare the proposed method with ten advanced models in semi-supervised and self-supervised learning settings to evaluate its performance. Subsequently, we perform a parameter sensitivity analysis and an ablation study. Finally, the learned representations are visualized to validate the performance of our method intuitively, as provided in Appendix C.

4.1 Datasets

Three widely used benchmark datasets are adopted: Cora, CiteSeer, and PubMed citation networks [38], where nodes represent publications and edges correspond to citation relationships. The statistical results for these datasets are presented in Table 1.

• Cora. Cora contains a number of machine learning publications, in which each publication cites at least one other paper, or is cited by another publication. This citation network includes seven classes corresponding to case based, genetic algorithms, neural networks, probabilistic methods, reinforcement learning, rule learning, and theory. We select about 5.2% of the nodes in this dataset for training.

• CiteSeer. CiteSeer contains scientific publications grouped into six classes: agents, artificial intelligence, database, information retrieval, machine language, and human-computer interaction. Each publication is represented by a 0/1-valued vector. We select about 3.6% of the nodes in this dataset for training.

• PubMed. PubMed contains diabetes-related scientific publications in three classes. Each publication is represented by a term frequency-inverse document frequency vector. We select about 0.3% of the nodes in this dataset for training.

4.2 Experiment setup

In our framework, we set the dimensions of node representations to 512. During the node allocation process, the K-Means algorithm is employed to cluster the global node representations, with the number of clusters set to the actual number of classes. Finally, we aggregate the multi-granularity node representations obtained from different topologies and train a linear classifier in a semi-supervised setting. For the Cora, CiteSeer, and PubMed datasets, we randomly select 20 nodes from each class to train the linear classifier, while using 1000 nodes to test the results. During the evaluation phase, each experiment is repeated 20 times and the model's performance is evaluated based on the average classification accuracy. For all experiments, we train our framework using the Adam Optimizer [39] with an initial learning rate of 3e−4.

Running environment. The MCGRL framework is implemented on the PyTorch platform with 3 NVIDIA GeForce RTX 3090 and conducted on a Linux server with a 24-core Intel CPU, 125.7 GB RAM and 72 GB GPU memory. The operating system is Ubuntu 18.04.6 LTS.

4.3 Comparison with state-of-the-art baselines

We compare MCGRL with ten state-of-the-art graph representation learning methods, including semi-supervised methods such as GCN [1] and MAGCN [23], as well as self-supervised methods including DGI [12], GRACE [14], GMI [40], MVGRL [17], BGRL [16], CCA-SSG [41], GGD [18], and GraphMAE2 [42].

The results, presented in Table 2, demonstrate that our framework achieves either superior or comparable performance across the three node classification benchmarks. Specifically, on CiteSeer and PubMed, MCGRL yields accuracies of 75.2% and 82.5%, respectively, representing improvements of at least 1.7% and 1.1% over the best-performing existing methods. The performance of our MCGRL on Cora is similar to that of the state-of-the-art methods. Overall, these results consistently prove the effectiveness of our MCGRL in node classification, which can be attributed to the fact that our method comprehensively considers the rich interactions among multiple topologies and the abundant complementary information between topologies, thereby enabling the learning of more discriminative representations.

Table 2 Node classification performance on three benchmarks.

Method	Cora	CiteSeer	PubMed
GCN	81.5	70.3	79.0
MAGCN	84.5 ± 0.2	73.5 ± 0.3	80.6 ± 0.2
DGI	82.3 ± 0.6	71.8 ± 0.7	76.8 ± 0.6
GRACE	81.9 ± 0.4	71.2 ± 0.5	80.6 ± 0.4
GMI	82.7 ± 0.2	73.0 ± 0.3	80.1 ± 0.2
MVGRL	83.5 ± 0.4	73.3 ± 0.5	80.1 ± 0.7
BGRL	82.7 ± 0.6	71.4 ± 0.8	79.6 ± 0.5
CCA-SSG	84.0 ± 0.4	73.1 ± 0.3	81.1 ± 0.4
GGD	83.9 ± 0.4	73.0 ± 0.6	81.3 ± 0.8
GraphMAE2	84.5 ± 0.6	73.4 ± 0.3	81.4 ± 0.5
MCGRL	84.6 ± 0.1	75.2 ± 0.2	82.5 ± 0.1

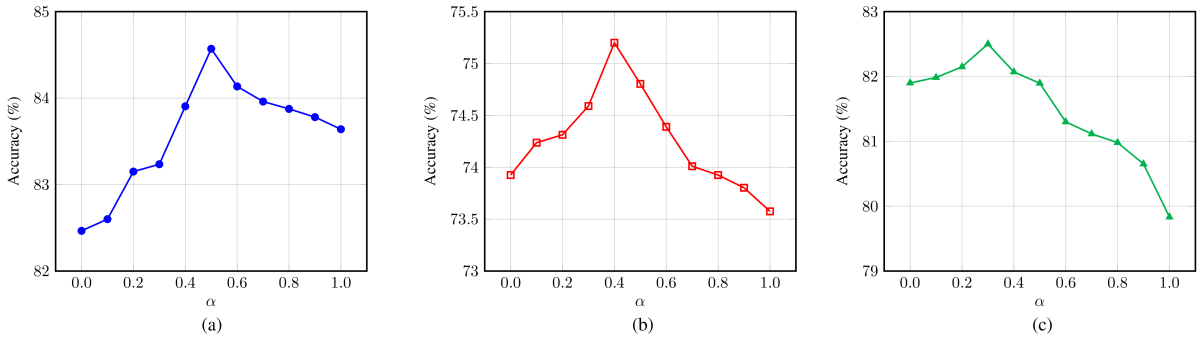

Figure 3 (Color online) Classification accuracy of MCGRL with different α values on three datasets. (a) Cora; (b) CiteSeer; (c) PubMed.

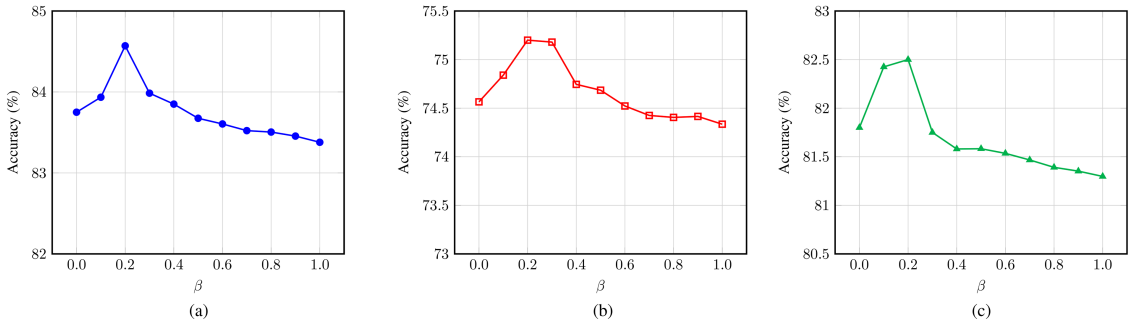

Figure 4 (Color online) Classification accuracy of MCGRL with different β values on three datasets. (a) Cora; (b) CiteSeer; (c) PubMed.

4.4 Hyperparameter sensitivity analysis

In this subsection, we delve deeply into the impact of the balancing factors α and β on the performance of node classification tasks utilizing multi-topology contrastive losses. The experimental results, as illustrated in Figures 3 and 4, delineate the trends in accuracy for Cora, CiteSeer, and PubMed datasets as α and β vary. Our observations suggest that appropriately optimizing the ratio between $\mathcal{L}_{\mathrm{MP1}}$ and $\mathcal{L}_{\mathrm{MP2}}$ losses plays a crucial role in significantly enhancing the model's performance. Moreover, controlling β within certain bounds significantly enhances the overall accuracy of the node classification, thus underscoring the importance of maintaining semantic consistency across multi-topology structures.

We further conduct a careful analysis of the hyperparameter sensitivity for η, aiming to gain a more comprehensive understanding of its critical role in exploring multi-granularity information for multi-topology contrastive learning. The results in Figure 5 clearly demonstrate that as η increases, the model performance gradually improves, reaching an optimal accuracy at $\eta = 0.8$. This confirms that balancing the weights between the global and local node representations contributes to the effective modeling of multi-granularity information, thereby enhancing the expressive power of our MCGRL. However, our research reveals a significant finding that the introduction of

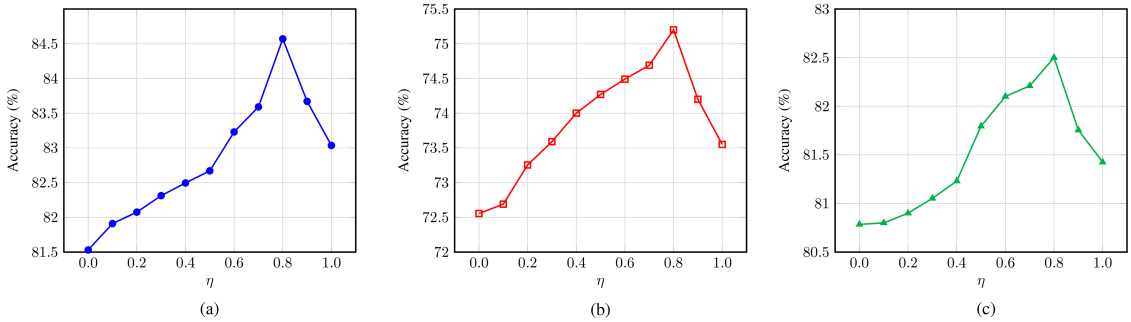

Figure 5 (Color online) Classification accuracy of MCGRL with different η values on three datasets. (a) Cora; (b) CiteSeer; (c) PubMed.

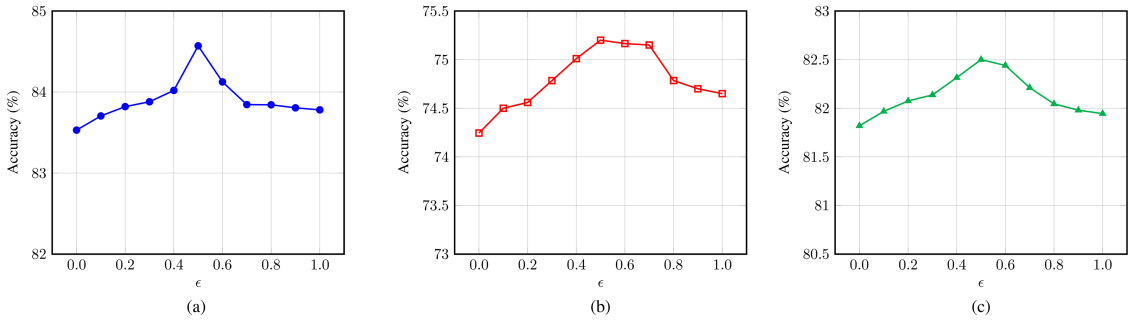

Figure 6 (Color online) Classification accuracy of MCGRL with different ϵ values on three datasets. (a) Cora; (b) CiteSeer; (c) PubMed.

Table 3 Ablation study (%) on three datasets.

Method	Cora	CiteSeer	PubMed
MCGRL	84.6	75.2	82.5
MCGRL without SGNN and $\mathcal{L}_{\mathrm{SIoU}}$	83.5	74.1	81.4
MCGRL without $\mathcal{L}_{\mathrm{SIoU}}$	83.8	74.6	81.8
MCGRL without G_1	83.2	74.2	80.3
MCGRL without G_2	84.0	72.6	81.6
MCGRL without G_3	81.1	73.8	80.5

excessive global information adversely affects the quality of node representations. This observation emphasizes the importance of avoiding excessive attention to global information during contrastive multi-topology learning. When η is too small, the model performs worse, validating our hypothesis that moderately utilizing local information in multiple topologies can improve model performance. Overall, by balancing the relationships between global and local information for multi-granularity node representation learning, the generalization performance of the model on node classification tasks can be effectively enhanced.

In Figure 6, we observe that as the threshold for deleting edges ϵ increases, the classification accuracy initially exhibits an ascending trend, followed by a decline. This trend can be attributed to the fact that increasing the threshold progressively reduces the number of edges between nodes from different categories in the original topology, leading to the removal of unnecessary edges. As the threshold further increases, some crucial edges are removed, causing a decrease in the classification accuracy.

4.5 Ablation study

In this subsection, we conduct an ablation study to determine the effectiveness of each component by designing five variants of MCGRL: without SGNN and $\mathcal{L}_{\mathrm{SIoU}}$, without $\mathcal{L}_{\mathrm{SIoU}}$, without the graph after edge perturbation for the original topology (G_1), without the k-nearest neighboring graph (G_2), and without the diffusion graph (G_3). The ablation results in Table 3 demonstrate that the removal of any component leads to a decline in model performance. For instance, the results for CiteSeer reveal that the deletion of $\mathcal{L}_{\mathrm{SIoU}}$ decreases performance to some extent, emphasizing the role of semantic consistency constraints across multiple topologies. When the model eliminates any topology from different perspectives, a significant performance drop is observed, with classification

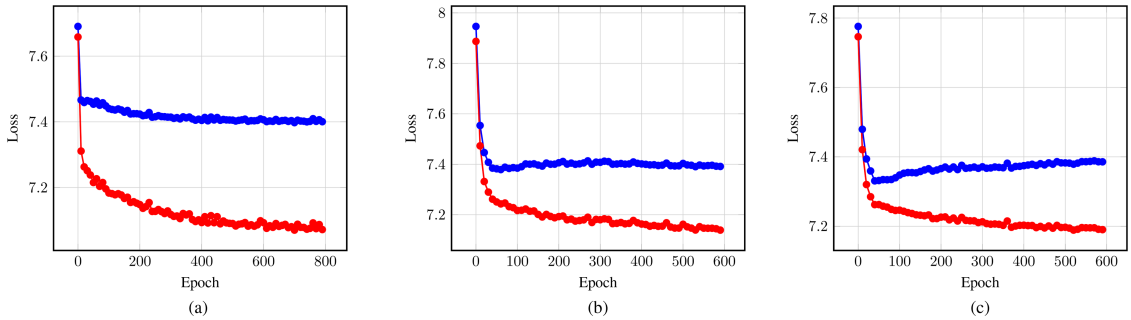

Figure 7 (Color online) Loss curves of our MCGRL with and without $\mathcal{L}_{\mathrm{SIoU}}$ loss on three datasets, illustrated by the blue and red curves, respectively. (a) Cora; (b) CiteSeer; (c) PubMed.

accuracy decreasing by approximately -1.0%, -2.6%, and -1.0%, respectively. This underscores the pivotal role of multiple topologies in graph contrastive learning, and substantiates the effectiveness of multi-topology contrastive learning in integrating diverse and information-rich topologies to yield discriminative node representations.

Furthermore, we plot the loss curves with and without $\mathcal{L}_{\mathrm{SIoU}}$ for the three datasets in Figure 7. Compared to the MCGRL without $\mathcal{L}_{\mathrm{SIoU}}$ on the lower side, the upper curve exhibits earlier convergence. We attribute this phenomenon to two primary factors. First, $\mathcal{L}_{\mathrm{SIoU}}$ reduces the distance between positive samples by maximizing the similarity between the representations of same-category subgraphs under different topological structures. Second, $\mathcal{L}_{\mathrm{SIoU}}$ further increases the distance between negative samples by maximizing the dissimilarity of different-category subgraphs under the same topological structure. In summary, $\mathcal{L}_{\mathrm{SIoU}}$ loss pulls nodes of the same class closer and pushes nodes of different classes farther apart, facilitating model convergence.

5 Conclusion

In this paper, we devise an innovative multi-topology contrastive graph representation learning framework that extracts global and local information from diverse topologies to learn multi-granularity node representations. Moreover, the similarity constraint SIoU is designed to systematically explore the intrinsic interaction information among different topologies. Extensive experiments on real-world datasets demonstrate the superior performance of MCGRL in node classification.

However, several key issues remain that warrant further investigation. First, the reliance on fixed topology generation rules may limit its effectiveness in dynamic or complex graph structures. Future work can explore the incorporation of dynamic graph neural networks or generative adversarial networks to design a more adaptive dynamic topology generation strategy. Additionally, our study primarily focuses on single-domain graph data without considering multi-domain or cross-modal graph data. Extending our MCGRL to cross-domain or cross-modal tasks and using more advanced GNNs as backbones to address more complex scenarios are desirable for future work.

Acknowledgements This work was supported by "Pioneer" and "Leading Goose" R&D Program of Zhejiang (Grant No. 2024C03262), National Natural Science Foundation of China (Grant Nos. 62472270, 62106131, 62473241, U21A20473, 62172370), Research Project Supported by Shanxi Scholarship Council of China (Grant No. 2022-007), and Fund Program for the Scientific Activities of Selected Returned Overseas Professionals in Shanxi Province (Grant No. 20220002).

Supporting information Appendixes A–C. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks. In: Proceedings of the International Conference on Learning Representations, 2017
2 Chen J X, Lin G Q, Chen J X, et al. Towards efficient allocation of graph convolutional networks on hybrid computation-in-memory architecture. Sci China Inf Sci, 2021, 64: 160409
3 Xia R T, Zhang C X, Zhang Y, et al. A novel graph oversampling framework for node classification in class-imbalanced graphs. Sci China Inf Sci, 2024, 67: 162101
4 Qiu J, Tang J, Ma H, et al. DeepInf: social influence prediction with deep learning. In: Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2018. 2110–2119
5 Wang Y, Min Y, Chen X, et al. Multi-view graph contrastive representation learning for drug-drug interaction prediction. In: Proceedings of the Web Conference, 2021. 2921–2933
6 Li F, Feng J, Yan H, et al. Dynamic graph convolutional recurrent network for traffic prediction: benchmark and solution. ACM Trans Knowl Discov Data, 2023, 17: 1–21

7 Li Z, Li L, Peng Y, et al. A two-stream graph convolutional neural network for dynamic traffic flow forecasting. In: Proceedings of the International Conference on Tools with Artificial Intelligence, 2020. 355–362
8 Jin T S, Dai H Q, Cao L J, et al. Deepwalk-aware graph convolutional networks. Sci China Inf Sci, 2022, 65: 152104
9 Wang G, Liu X, Wang K, et al. Deep-learning-enabled protein-protein interaction analysis for prediction of SARS-CoV-2 infectivity and variant evolution. Nat Med, 2023, 29: 2007–2018
10 Wei X, Liu Y, Sun J, et al. Dual subgraph-based graph neural network for friendship prediction in location-based social networks. ACM Trans Knowl Discov Data, 2023, 17: 1–28
11 Fu S, Peng Q, He Y, et al. Unsupervised multiplex graph diffusion networks with multi-level canonical correlation analysis for multiplex graph representation learning. Sci China Inf Sci, 2025, 68: 132102
12 Veličković P, Fedus W, Hamilton W L, et al. Deep graph infomax. In: Proceedings of the International Conference on Learning Representations, 2019
13 Linsker R. Self-organization in a perceptual network. Computer, 1988, 21: 105–117
14 Zhu Y, Xu Y, Yu F, et al. Deep graph contrastive representation learning. 2020. ArXiv:2006.04131
15 Grill J B, Strub F, Altché F, et al. Bootstrap your own latent a new approach to self-supervised learning. In: Proceedings of the Advances in Neural Information Processing Systems, 2020. 21271–21284
16 Thakoor S, Tallec C, Azar M G, et al. Bootstrapped representation learning on graphs. In: Proceedings of the ICLR Workshop on Geometrical and Topological Representation Learning, 2021
17 Hassani K, Khasahmadi A H. Contrastive multi-view representation learning on graphs. In: Proceedings of the International Conference on Machine Learning, 2020. 4116–4126
18 Zheng Y, Pan S, Lee V, et al. Rethinking and scaling up graph contrastive learning: an extremely efficient approach with group discrimination. In: Proceedings of the Advances in Neural Information Processing Systems, 2022. 10809–10820
19 Ma Y, Wang S, Aggarwal C C, et al. Multi-dimensional graph convolutional networks. In: Proceedings of the IEEE International Conference on Data Mining, 2019. 657–665
20 Wang X, Zhu M, Bo D, et al. AM-GCN: adaptive multi-channel graph convolutional networks. In: Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2020. 1243–1253
21 Ju W, Gu Y Y, Mao Z Y, et al. GPS: graph contrastive learning via multi-scale augmented views from adversarial pooling. Sci China Inf Sci, 2025, 68: 112101
22 Khan M R, Blumenstock J E. Multi-GCN: graph convolutional networks for multi-view networks, with applications to global poverty. In: Proceedings of the 33rd AAAI Conference on Artificial Intelligence, 2019. 606–613
23 Yao K, Liang J, Liang J, et al. Multi-view graph convolutional networks with attention mechanism. Artif Intell, 2022, 307: 103708
24 Peng X, Feng J, Lu J, et al. Cascade subspace clustering. In: Proceedings of the Conference on Artificial Intelligence, 2017. 2478–2484
25 Peng X, Feng J, Zhou J T, et al. Deep subspace clustering. IEEE Trans Neural Netw Learn Syst, 2020, 31: 5509–5521
26 Chen J, Mao H, Woo W L, et al. One-step adaptive graph learning for incomplete multiview subspace clustering. IEEE Trans Knowl Data Eng, 2025, 37: 2771–2783
27 Peng X, Lu J, Yi Z, et al. Automatic subspace learning via principal coefficients embedding. IEEE Trans Cybern, 2016, 47: 3583–3596
28 Hjelm R D, Fedorov A, Lavoie-Marchildon S, et al. Learning deep representations by mutual information estimation and maximization. In: Proceedings of the International Conference on Learning Representations, 2019
29 Sun F Y, Hoffmann J, Verma V, et al. InfoGraph: unsupervised and semi-supervised graph-level representation learning via mutual information maximization. In: Proceedings of the International Conference on Learning Representations, 2019
30 Zhang Y, Yuan Y, Wang Q. Multi-level graph contrastive prototypical clustering. In: Proceedings of the International Joint Conference on Artificial Intelligence, 2023. 4611–4619
31 Liu Y, Zhao Y, Wang X, et al. Multi-scale subgraph contrastive learning. In: Proceedings of the International Joint Conference on Artificial Intelligence, 2023. 2215–2223
32 Jin M, Zheng Y, Li Y F, et al. Multi-scale contrastive siamese networks for self-supervised graph representation learning. In: Proceedings of the International Joint Conference on Artificial Intelligence, 2021. 1477–1483
33 Wu L, Lin H, Tan C, et al. Self-supervised learning on graphs: contrastive, generative, or predictive. IEEE Trans Knowl Data Eng, 2023, 35: 4216–4235
34 Jiao Y, Xiong Y, Zhang J, et al. Sub-graph contrast for scalable self-supervised graph representation learning. In: Proceedings of the IEEE International Conference on Data Mining, 2020. 222–231
35 Chen J, Fang H, Saad Y. Fast approximate kNN graph construction for high dimensional data via recursive Lanczos bisection. J Mach Learn Res, 2009, 10: 1989–2012
36 Yu J, Jiang Y, Wang Z, et al. UnitBox: an advanced object detection network. In: Proceedings of the ACM International Conference on Multimedia, 2016. 516–520
37 Corso G, Cavalleri L, Beaini D, et al. Principal neighbourhood aggregation for graph nets. In: Proceedings of the Advances in Neural Information Processing Systems, 2020. 13260–13271
38 Sen P, Namata G, Bilgic M, et al. Collective classification in network data. AI Mag, 2008, 29: 93–106
39 Kingma D P, Jimmy Ba J. Adam: a method for stochastic optimization. In: Proceedings of the International Conference on Learning Representations, 2015
40 Peng Z, Huang W, Luo M, et al. Graph representation learning via graphical mutual information maximization. In: Proceedings of the Web Conference, 2020. 259–270
41 Zhang H, Wu Q, Yan J, et al. From canonical correlation analysis to self-supervised graph neural networks. In: Proceedings of the Advances in Neural Information Processing Systems, 2021. 76–89
42 Hou Z, He Y, Cen Y, et al. GraphMAE2: a decoding-enhanced masked self-supervised graph learner. In: Proceedings of the Web Conference, 2023. 737–746