Federated Graph-level Clustering Network with Attribute Inference

Renda Han[1], **Junlong Wu**[1], **Wenxuan Tu**[1,2], **Jingxin Liu**[3], **Haotian Wang**[1], **Jieren Cheng**[1,2]

[1]School of Computer Science and Technology, Hainan University, 570000, Haikou, China
[2] Hainan Blockchain Technology Engineering Research Center, 570000, Haikou, China
[3] School of Cyberspace Security, Hainan University, 570000, Haikou, China
{hanrenda, ttlong_346, twx, ljx_0417, 20010101, 992730}@hainanu.edu.cn

Abstract

With the rise of vertical segmentation in real-world data, federated graph-level clustering has gained significant attention in recent years. However, the inherent missing attributes in graph datasets held by certain clients lead to suboptimal local parameter updates and misaligned global parameter consensus. This results in knowledge shifts during negotiation to ultimately impair overall clustering performance. This issue remains largely underexplored in the current advanced research. To bridge this gap, we propose a novel deep learning network called **Fed**erated Graph-level Clustering Network with **A**ttribute **I**nference (FedAI), which utilizes high-confidence prior knowledge from each domain and multi-party collaborative optimization to achieve efficient reasoning of unknown features. Specifically, on the client, high-confidence graph samples are projected into a latent space. We then extract and upload irreversible path digest information and attribute-oriented inference signals from them. On the server, we first identify affinity relationships hierarchically via the improved graph kernel method. We then infer the features of clients lacking node attributes through a prior structure-guide recovery operator, facilitating inter-client knowledge transfer for better clustering. Experimental results on 15 cross-dataset and cross-domain non-IID graph datasets demonstrate that FedAI consistently outperforms existing methods.

Code — https://github.com/H00001/FedAI

Introduction

In recent years, with the explosive growth of real-world graph-structured data (Liang et al. 2024a, 2025; Liu et al. 2025c,d; Liang et al. 2024b,c; Liu et al. 2024b) that encode rich and complex relationships, Federated Graph Learning (FGL) (Xie et al. 2021; Li et al. 2024; Xia and Zhang 2024), as an emerging distributed machine learning paradigm, has shown broad application potential in social network analysis (Ambekar et al. 2024), smart healthcare (Gurumurthy, Pal, and Sharma 2025), financial risk control (Liu et al. 2024a). Its core value lies in effectively solving the non-IID data challenge resulting from vertical partitioning of data via collaborative training of the Graph Neural Network (GNN) model, while meeting the rigid requirements of privacy protection.

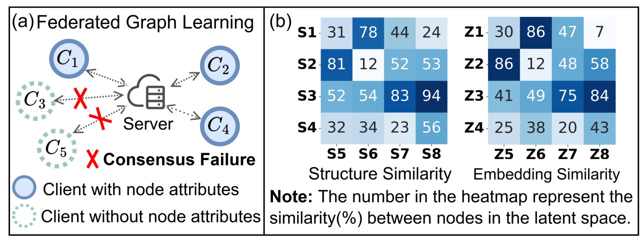

Figure 1: (a) A key issue of existing FGC methods: clients without node attributes struggle to reach consensus, leading to knowledge drift in the server. (b) We can observe that the similarity between structure and attributes is evaluated under a shared latent space, where higher structural similarity implies a greater embedding similarity.

Currently, in real-world scenarios, labeled data is often scarce or entirely unavailable. As a result, unsupervised Federated Graph Clustering (FGC) methods attracted significant research interest in recent years as a promising subfield within FGL. Among these, FedNCN (Liu et al. 2025b) and FedGCN (Liu et al. 2025a) represent the first and most recent attempts at addressing node- and graph-level federated clustering tasks, respectively. FedNCN mends the destroyed links by clustering prior knowledge without sharing private graph data. Notably, FedGCN generates structure-oriented graph embeddings and extracts prototypes on each client. The server uses an improved Gaussian mixture distribution to separate multi-source structural signals for local model learning. This approach not only significantly improves clustering quality but also effectively protects data privacy.

However, these advanced methods fundamentally rely on the assumption of fully observable node attributes, overlooking the issue that this assumption often fails in the real world due to privacy constraints, observation sparsity, or data corruption. Moreover, some methods (Tan et al. 2023; Liu et al. 2025a) typically treat structural encodings as node attributes. Since structural information is already inherently available, such graph structure embeddings do not effectively enrich the information beneficial for clustering. Es-

Corresponding author: Jieren Cheng
Renda Han and Junlong Wu contributed equally.

sentially, the semantic information of attributes is hard to recover. For example, node attributes in datasets such as MUTAG (Cai, Zhang, and Fan 2025) and BZR_MD (Cai et al. 2024a), which are commonly used in the domains of small molecules, are completely missing and unobserved. For local learning, the graph as a non-Euclidean structure is inherently fragile (Tu et al. 2024). Missing node attributes hinder the extraction of reliable node representations, disrupt the relational identification, and ultimately lead to a decline in local clustering performance. In the Federated Learning (FL) (Song et al. 2025; Liao et al. 2024, 2025, 2023) framework, this issue is amplified. Since some clients hold graphs with completely missing node attributes, i.e., Client without Node Attribute (CNA), the learning process is severely disrupted. As shown in Fig. 1(a), the learned local model parameters are biased or inconsistent between CNAs and clients with node attributes. During the aggregation process, these flawed model parameters cause the knowledge shifts across clients, which further leads to the server consensus failure, ultimately degrading clustering performance.

To address this issue, an intuitive solution in FL is to leverage the known attributes to infer the missing attributes across clients. However, since CNAs only have graph structure, it is essential to explore the latent relationships between attributes and structures in clients with complete attributes and transfer this paradigm (i.e., use structure to infer attributes) to CNAs. Previous works (Xie et al. 2021; Chen et al. 2022) argue that structures and attributes have implicit relationships within the Hilbert space. As shown in Fig. 1 (b), we observe that higher structural similarity implies a greater embedding similarity in a shared latent space. Based on the above observation, we propose an elegant FGL framework for both cross-dataset and cross-domain graph-level attribute inference called **Fed**erated Graph-Level Clustering Network with **A**ttribute **I**nference (FedAI). The key idea of FedAI is to employ high-confidence prior knowledge to infer the unknown attribute in non-IID graphs. Specifically, in the client, we first sample graph instances with reliable representations to extract hard-to-recover digest information by the Hierarchical Random Walk (HRW) Graph kernel we proposed for FL, and upload them to the server. In the server, we propose a Kernel-aware Alignment and Inference Strategy (KAIS), which utilizes the HRW Graph Kernel (GK) to quantify the affinity on inter-client hierarchically. Then, we introduce an attribute-oriented inference operator to infer cluster-friendly attributes and perform graph-level clustering jointly. It is worth noting that our proposed HRW GK can measure the potential affinity relationship between clients without uploading the original sensitive graph data in the FL environment. In summary, our contributions are as follows:

- **New Research Task** To our knowledge, this is the first attempt to jointly address multi-client clustering and attribute inference within a unified FL framework.
- **New FGL Method** A novel federated graph-level clustering network called FedAI is proposed, which aims to utilize high-confidence prior knowledge to cross-domain inference cluster-friendly attributes and solve the issue of knowledge caused by missing node attributes.
- **Better Experimental Results** Extensive experiments demonstrate that FedAI consistently outperforms existing SOTA federated graph-level learning methods.

Related Work

Federated Graph-level Learning

Federated Graph-Level Learning aims to collaboratively train models across distributed parties to learn graph-level representations while preserving data privacy. Early methods such as GCFL and GCFL+ (Xie et al. 2021), first leverage gradient-driven subgroup discovery to mitigate heterogeneity across clients. Then, FedStar (Tan et al. 2023) decouples structure and feature representations, which mitigates feature misalignment. Moreover, FedGCN (Liu et al. 2025a) collects topology-driven features from non-IID graphs across clients to generate global consensus representations. More recently, FedPKA (Wu et al. 2025) uses a community-driven dynamic aggregation method to customize personalized model parameters for each client. FCLG (Li, Agrawal, and Ramnath 2024) jointly enhances local graph representation learning and alleviates inter-client distributional discrepancies in federated settings. As for anomaly detection (Wang et al. 2025b, 2023, 2025a) in FGL, FGAD (Cai et al. 2024c) introduces an anomaly generator to perturb normal graphs and trains an anomaly detector to distinguish between normal and anomalous graphs. Then, LG-FGAD (Cai et al. 2024b) based on FGAD trains a discriminator by maximizing/minimizing mutual information from a multi-level perspective. AGDiff (Cai et al. 2025) uses subtle perturbations in graph representations, generating pseudo-anomalous graphs that closely resemble normal ones. However, these methods fail to consider the issue of node attribute missing in the datasets held by clients, which is a common challenge in real-world scenarios.

Attribute-missing Graph Clustering

Recent advances in GNNs have proven effective in learning from complex graph data (Liu et al. 2024c; Yang et al. 2024; Jin et al. 2022). However, most methods assume complete node attributes and largely ignore the challenge of missing information. To this end, a number of GNN-based methods have explored data imputation techniques, such as multi-modal networks (Liu et al. 2023), and Gaussian mixture models (Liu et al. 2025a). While effective in reconstructing missing features, these techniques are typically not designed for clustering tasks, and thus fail to ensure that the recovered attributes contribute to cluster-discriminative representations. To bridge this gap, recent studies, for instance, AMGC (Tu et al. 2024), iteratively refines imputation and clustering using current cluster assignments as guidance. Although this bidirectional interaction improves both stages, its performance heavily depends on the quality of the initial clusters and involves non-trivial optimization procedures. To overcome this issue, FPGC (Xie et al. 2025) further constructs individualized models for each node using a Squeeze-and-Excitation mechanism, capturing local inter-attribute dependencies. However, FPGC suffers from scalability issues due to the explosion of learnable model parame-

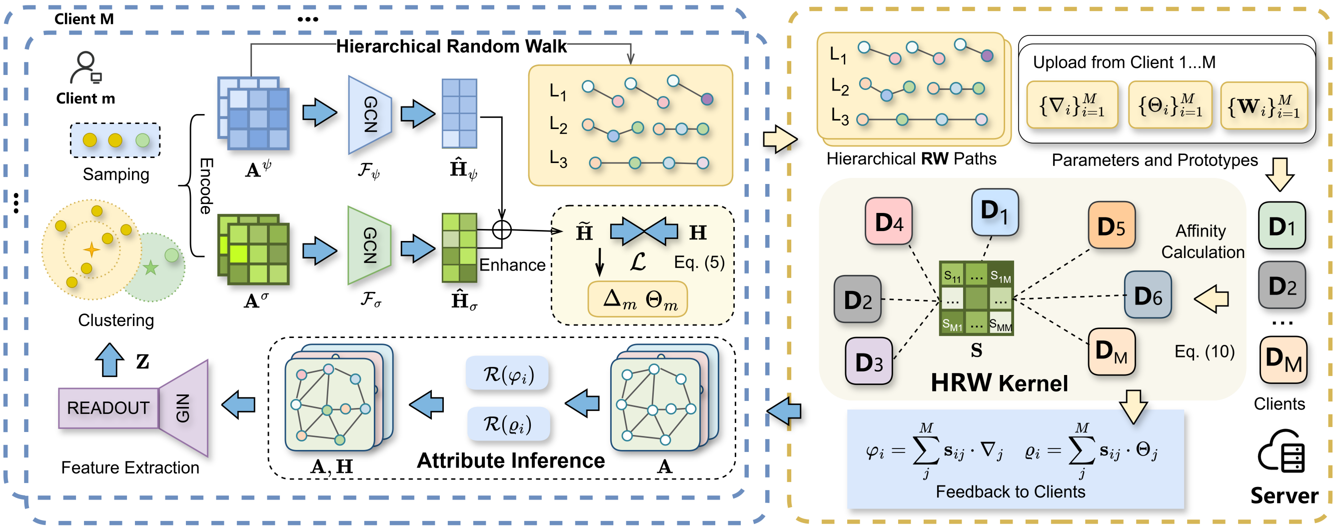

Figure 2: The network architecture of FedAI. Each client extracts hierarchical RW paths from local graphs, generating a structural digest signal. In parallel, a dual network is utilized to extract an attribute-oriented gradient signal. Upon receipt the both signals, the server leverages the digest signal to compute structural affinities across clients and subsequently constructs an inference operator. This operator is then used to guide the restoration of missing attributes on the CNAs.

ters. Moreover, MVCG (Li et al. 2025) integrates contrastive and generative objectives, which encourage more robust and discriminative embeddings. More recently, WAGE (Tu et al. 2025) enhances feature reconstruction without introducing excessive model complexity, achieving high-quality attribute recovery. However, these methods follow a centralized paradigm to solve the issue of missing node attributes and neglect the inherent privacy challenges arising in federated learning environments.

Methodology

In this section, we provide a detailed introduction to FedAI. The core idea is to leverage prior credible knowledge to infer missing node attributes, while alternately optimizing the quality of attribute inference and consensus clustering. The overall methodology is organized into four main modules: local model learning, hierarchical random walk Graph Kernel for FL, kernel-aware alignment and inference, and global parameter aggregation. An architectural overview of the proposed FedAI framework is presented in Fig. 2, highlighting the core components and their interactions. The pseudocode for FedAI is presented in Algorithm 1.

Notations

Suppose that our federated framework consists of M decentralized clients. Each client holds a private graph dataset $\mathcal{G} = \{G_i\}_{i=1}^{I}$, with I individual graphs and K clusters. Any graph G is represented as a tuple $(\mathbf{A}, \mathbf{X})$, where $\mathbf{A} \in \mathbb{R}^{n\times n}$ denotes the normalized adjacency matrix with self-loop, and $\mathbf{X} \in \mathbb{R}^{n\times d}$ is the node attribute matrix, with each row encoding a d-dimensional feature vector of a node. Note that on some clients, the original $\mathbf{X}$ is absent due to missing node attributes in the graph dataset, referred to as Clients without Node Attributes (CNA).

Local Model Learning

To enable high-quality representation learning across clients with attributed graph data and to reliably identify prior samples with high confidence, we employ the Graph Isomorphism Network (GIN) as the encoder $\mathcal{F}$, enhancing structural awareness and the expressiveness of the learned embeddings. The formula is as follows:

$$\mathbf{H} = \mathcal{F}(\mathbf{X}, \mathbf{A}|\mathbf{W}), \tag{1}$$

where $\mathbf{W}$ denotes the learnable parameter matrix of GIN. In this way, we extract the node-level latent embeddings $\mathbf{H}$. Then, we utilize $\mathbf{H}$ to obtain the graph-level representation $\mathbf{Z}$, as: $\mathbf{Z} = \omega(\mathbf{H})$, where $\omega(\cdot)$ represents the READOUT function. Here, we leverage the reconstruction loss to guide the optimization process. Next, we apply the improved k-medoids algorithm to identify high-confidence representatives from $\mathbf{Z}$. It selects actual data points as cluster centers to ensure interpretability. Formally, given a dissimilarity function $d(\cdot,\cdot)$, the goal is to minimize the total intra-cluster dissimilarity, as:

$$\min_{\mathcal{M}\subset\{\mathbf{z}_i\},\,|\mathcal{M}|=K} \sum\nolimits_{j=1}^{K} \sum\nolimits_{\mathbf{h}_i\in\mathcal{C}_j} d(\mathbf{z}_i, \mathbf{m}_j), \tag{2}$$

where $\mathcal{M} = \{\mathbf{m}_k\}_{k=1}^{K}$ is the medoid set, and each cluster $\mathcal{C}_j$ contains a point $\mathbf{m}_j$ that is closest to all points in the cluster. Then, we select the nearest $\gamma\%$ as a set of representative samples that are both structurally meaningful and appropriate for extracting the knowledge digest.

Attribute-oriented Inference Operator To facilitate the inference of missing attributes by leveraging structural priors, we construct an implicit correspondence between the two modalities using an attribute-aware gradient signal, which contains the underlying structure-to-attribute alignment in a differentiable manner. Specifically, we define a

Algorithm 1: Training Procedure of FedAI

Require: Each graph dataset $\mathcal{G}$; Number of clusters K; Number of client M; Global train epoch N_{global};
Ensure: Cluster results R;

```
Initialize model parameters W for each client;
for i ← 1 to N_c do
    Learn graph representation by Eq. (1);
    Execute the clustering task by k-medoids;
    /* On clients with attributes */
    Select γ% representative samples by Eq. (2);
    Extract Hierarchical RW path φ(G) by Eqs. (6)-(7);
    Extract attribute-oriented inference operator Θ and
    ∇ by Eqs. (3)-(5);
end for
/* On the server */
Collect Θ, ∇ and D from each client;
Construct the affinity matrix S by HRW kernel;
Generate the inference operator φ and ϱ by Eq. (11);
Feedback φ and ϱ to each CNA;
for i ← 1 to M do
    /*On clients without attributes*/
    Leverage the φ_i and ϱ_i to recover the attribute;
end for
return R
```

dual structure-to-attribute projector $\mathcal{F}_\psi$ and $\mathcal{F}_\sigma$ to extract structure embeddings $\hat{\mathbf{H}}_\psi$ and $\hat{\mathbf{H}}_\sigma$, respectively, as:

$$\hat{\mathbf{H}}_\psi = \mathcal{F}_\psi(\mathbf{A}^\psi; \Theta_\psi|\nabla_\psi), \quad \hat{\mathbf{H}}_\sigma = \mathcal{F}_\sigma(\mathbf{A}^\sigma; \Theta_\sigma|\nabla_\sigma), \quad (3)$$

where Θ_σ and Θ_ϕ are the parameters, and ∇_σ and ∇_ϕ are gradient. $\mathcal{F}_\theta$ and $\mathcal{F}_\sigma$ utilize degree distribution and random walk to encode $\mathbf{A}$ as $\mathbf{A}^\phi$ and $\mathbf{A}^\sigma$ for projecting, respectively. This dual approach enables the capture of multi-view structural information, ensuring an end-to-end differentiable signal that effectively integrates higher-order structural features. Next, we enhance the two learned representations, as:

$$\widetilde{\mathbf{H}} = \hat{\mathbf{H}}_\psi \hat{\mathbf{H}}_\psi^\top \hat{\mathbf{H}}_\sigma. \quad (4)$$

To guide the learning process, we employ the reconstruction loss $\mathcal{L}$ between the inferred and ground-truth attributes:

$$\mathcal{L} = \frac{1}{n} \sum \|\widetilde{\mathbf{H}} - \mathbf{H}\|^2. \quad (5)$$

Subsequently, we collect the model gradient $\nabla = (\nabla_\psi; \nabla_\sigma)$ and model parameters $\Theta = (\Theta_\psi, \Theta_\sigma)$ as attribute-oriented inference signals from the dual networks and upload them to the server. The design has two advantages: it 1) protects the private information by transmitting only gradients and parameters. 2) does not rely on any raw structural data.

Hierarchical RW Graph Kernel for FL

In order to effectively extract the potential correlation between clients while protecting privacy, we propose a novel graph kernel called the Hierarchical Random Walk (HRW) Graph Kernel. HRW Graph Kernel requires only the path digest information of each graph provided by the client, eliminating the necessity to upload the complete graph structure. This design not only protects privacy but also reduces the computational burden on the server by avoiding full-graph similarity and feature-mapping operations. The entire process is compatible with federated learning settings, where local clients extract HRW paths and the server performs the final relation computing. Here, HRW paths are constructed in a location-based manner to more accurately identify affinities between clients in an FL environment, allowing the model to capture both fine-grained and higher-order topological information across clients.

Random Walk Path Representation Let $G = (V, E)$ be an undirected graph. $\iota(\cdot)$ is Laplace Position Encoding (Tan et al. 2023). For a fixed walk length l, we define $\mathcal{RW}_l(G)$ as the set of random walk paths of length l sampled from G, where each path representation is denoted $\pi = (v_1, v_2, \ldots, v_{l+1})$. We formulate each path $\pi \in \mathcal{RW}_l(G)$ by concatenating the structural features of its nodes:

$$\iota(\pi) = \Pi(\omega(v_1) \cdot \iota(v_1), \cdots, \omega(v_{l+1}) \cdot \iota(v_{l+1})), \quad (6)$$

where Π denotes the concatenation operator and $\omega(v_i) = \vartheta_d(v_i) + \vartheta_b(v_i) + \vartheta_c(v_i)$ to emphasize more structurally important nodes in the random walk path representation. Here, $\vartheta_d(v_i)$, $\vartheta_b(v_i)$, and $\vartheta_c(v_i)$ represent the degree-, betweenness-, and closeness-centrality of node v_i.

Path-level Kernel Function Given two walk paths $\pi \in \mathcal{RW}_l(G)$ and $\pi' \in \mathcal{RW}_l(G')$, the kernel function k_p at the path level is defined as the inner product of their attributes:

$$k_p(\pi, \pi') = \langle \iota(\pi), \iota(\pi') \rangle. \quad (7)$$

This linear kernel (Wang, Tu, and Cheng 2025) captures aligned semantic similarity along corresponding walk positions. Then the HRW Graph Kernel k_{hrw} can be defined as:

$$k_{\text{hrw}}(G, G') = \sum_{l=1}^{L} \left[\sum_{\pi \in \mathcal{RW}_l(G)} \sum_{\pi' \in \mathcal{RW}_l(G)} k_p(\pi, \pi') \right], \quad (8)$$

where L is the diameter of each graph.

Hierarchical Graph Feature Mapping To construct a graph-level feature $\varphi_l(G)$ for graph G, we first aggregate the representations of all sampled paths of the same length l. The hierarchical feature vector at the level is defined as:

$$\varphi_l(G) = \sum\nolimits_{\pi \in \mathcal{RW}_l(G)} \iota(\pi), \quad (9)$$

Next, we define the hierarchical mapping as:

$$\Phi(G) = \text{CONCAT}\left(\varphi_1(G), \varphi_2(G), \ldots, \varphi_L(G)\right), \quad (10)$$

where the $\text{CONCAT}(\cdot)$ denotes the concatenation operation.

Kernel-aware Alignment and Inference

After completing all above steps, we proceed with attribute inference. Here, the server collects guidance signals $\{\nabla_i\}_{i=1}^M$ and $\{\Theta_i\}_{i=1}^M$, client graph digest $\{\mathcal{D}_i\}_{i=1}^M$, and parameters $\{\mathcal{W}_i\}_{i=1}^M$ from each client. $\{\mathcal{D}_i\}_{i=1}^M$ includes all hierarchical RW paths $\Phi(G)$ of representative samples, and this collection process is completed on the client. Next, to enable cross-client alignment, we first employ the HRW

Models	SM^2(7)				SM-SY^2(8)				SM-BIO^2(9)			
	ACC	NMI	ARI	F1	ACC	NMI	ARI	F1	ACC	NMI	ARI	F1
Local	54.3±0.5	14.5±2.3	9.3±2.2	51.5±1.8	54.1±2.7	2.9±3.0	4.0±2.6	42.4±2.5	52.1±1.2	6.7±0.2	3.5±1.8	50.3±1.1
FedSage*	55.6±1.4	12.2±1.3	7.6±0.6	50.2±1.0	49.4±1.7	10.4±1.5	11.2±2.3	42.6±2.1	57.4±2.2	5.2±2.1	4.2±2.7	49.9±0.5
GCFL*	61.1±1.8	8.7±2.4	9.4±2.4	43.3±1.6	52.0±0.9	4.3±2.1	0.8±1.1	49.6±2.3	60.1±1.8	4.7±2.4	3.2±2.3	47.3±1.5
FedStar*	58.9±2.4	12.0±1.2	0.1±0.8	49.7±2.8	53.1±2.6	3.9±1.3	0.2±0.1	41.2±2.4	59.5±1.6	5.3±1.5	3.8±2.0	51.7±2.2
LG-FGAD†	65.8±0.8	18.8±1.9	3.4±1.1	62.9±0.6	60.8±1.9	10.1±3.3	7.0±3.4	59.3±2.2	59.6±1.5	9.0±1.4	7.8±1.7	56.0±2.6
FGAD†	66.4±2.4	20.2±2.6	4.3±3.2	63.8±2.6	65.5±1.5	15.2±2.3	11.4±2.0	63.8±1.9	63.5±1.0	14.7±1.1	2.1±2.0	60.7±1.5
AGDiff†	70.2±1.4	19.3±2.9	15.6±3.9	67.3±2.8	62.7±1.4	13.5±1.8	12.6±2.4	58.6±1.3	61.3±2.5	10.6±1.9	3.4±0.2	57.2±1.6
FedGCN	75.9±0.8	23.1±1.6	31.1±3.4	67.1±1.5	70.8±2.5	14.2±7.0	16.9±8.5	58.9±5.6	69.2±0.6	14.0±2.7	17.5±3.1	59.1±0.9
FedAI	**78.6±0.4**	**26.9±1.0**	**32.8±1.3**	**70.5±1.1**	**74.5±0.6**	**26.4±3.9**	**31.1±3.5**	**68.1±1.3**	**73.2±1.7**	**20.0±4.2**	**23.2±2.4**	**62.3±4.7**

Models	SM-BIO-SY^2(10)				SN-SY^{11}(2)				CV^{15}(3)			
Local	55.5±3.1	6.4±2.8	12.0±2.9	52.1±2.0	8.7±0.2	5.6±6.3	1.6±0.2	6.4±0.5	26.4±1.8	29.3±1.0	12.5±1.3	27.3±1.7
FedSage*	57.6±1.9	20.6±1.9	17.6±2.4	46.7±1.8	15.6±1.1	7.6±1.0	3.4±2.7	2.9±1.8	19.6±0.8	22.7±0.4	12.5±1.3	18.2±0.8
GCFL*	59.1±2.0	14.4±2.2	13.7±2.8	52.3±1.9	19.3±0.6	4.5±2.3	1.2±1.1	8.7±0.9	27.9±1.6	27.5±2.2	13.1±1.9	27.4±1.3
FedStar*	57.9±2.6	15.7±2.4	16.1±3.0	52.3±2.2	19.0±2.9	4.1±2.5	2.3±2.5	7.9±2.2	22.7±1.0	20.3±1.7	10.3±2.1	20.2±3.2
LG-FGAD†	58.4±0.5	7.6±0.4	6.4±0.8	54.6±0.7	19.1±1.4	6.7±0.6	3.3±1.0	7.4±1.1	27.4±1.6	31.4±4.0	9.6±3.2	24.7±3.3
FGAD†	62.2±1.3	14.6±2.6	3.0±2.9	56.7±0.8	16.4±0.5	7.4±0.5	2.7±0.3	8.3±0.7	26.0±1.1	31.9±1.2	7.3±1.1	25.2±1.1
AGDiff†	61.4±1.3	15.6±2.8	13.5±1.4	50.4±3.0	15.8±2.0	4.3±2.0	3.4± 0.2	9.6±2.8	23.6±1.4	27.5±1.3	8.7±0.9	22.8±1.4
FedGCN†	68.6±1.3	13.5±2.1	17.2±3.6	59.4±3.8	18.3±3.1	4.8±5.0	2.3±2.6	11.2±3.5	30.4±2.6	31.8±2.1	16.6±2.3	27.1±2.9
FedAI	**72.3±1.1**	**21.2±1.9**	**22.3±2.6**	**61.2±1.3**	**24.6±0.3**	**11.5±2.0**	**5.3±0.7**	**14.3±1.1**	**33.6±1.4**	**33.3±1.6**	**19.1±1.6**	**31.3±1.4**

Table 1: Performance comparison across different federated graph-level clustering tasks in horizontal format. * denotes supervised methods adapted for unsupervised learning. † denotes anomaly detection methods adapted for clustering.

graph kernel to compute the structural affinity matrix $\mathbf{S}$ between different clients as: $s_{ij} = k_{\text{hrw}}(\mathcal{D}_i, \mathcal{D}_j)$, where $s_{ij} \in \mathbf{S}$. Next, the server uses the affinity to construct a consensus attribute-oriented inference operator φ and ϱ for each cluster in the client that lacks attributes as:

$$\varphi_i = \sum_j^M \hat{s}_{ij} \cdot \nabla_j, \quad \varrho_i = \sum_j^M \hat{s}_{ij} \cdot \Theta_j, \tag{11}$$

where $\hat{s}_{ij}$ is the normalized affinity score. In this way, we send φ_i and ϱ_i back to the i-th CNA. It leverages both the original graph structure and its encoded counterpart to construct a recovery network $\mathcal{R}$. In the process of inference, these clients utilize consensus inference signals derived from these dual structural views to recover the missing node attributes in latent space.

Global Parameter Aggregation

To further improve the generalization capability of the proposed method, we incorporate the previously derived affinity matrix $\mathbf{S}$ to calculate the weight of clients during the aggregation process, as $\mathbf{W}_i = \sum_j^M \hat{s}_{ij} \cdot \mathbf{W}_j$. In this way, we exploit inter-client correlations to construct tailored model parameters that further guide the model training. In fact, alternative aggregation strategies can also be utilized in our methods.

Experiments

Experiment Setup

Benchmark Datasets To evaluate the effectiveness of FedAI, we construct non-IID federated learning scenarios using datasets across five domains from the publicly available TUDataset collection (Morris et al. 2020), which include 1) a same-domain setting involving only small molecule datasets (SM(7)); 2) a cross-domain setting combining small molecule and biological datasets (SM-BIO(9));

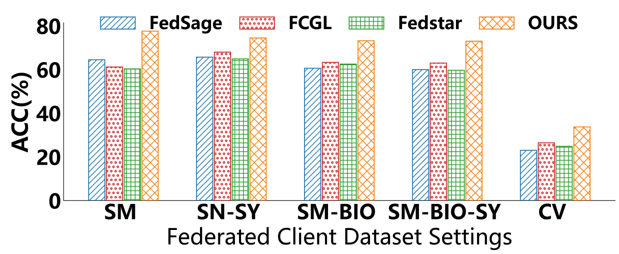

Figure 3: Performance comparison between three advanced federated graph-level classification methods with a limited number of labels and FedAI on five non-IID settings.

3) a more heterogeneous cross-domain setting comprising small molecule, biological, and synthetic datasets (SM-BIO-SY(10)); 4) a cross-domain setting involving social network and synthetic datasets (SN-SY(2)); 5) a same-domain setting with computer vision datasets (CV(3)); and 6) a cross-domain setting involving small molecule and synthetic datasets (SM-SY(8)).

Evaluation Metrics The clustering performance is evaluated by four widely used metrics: ACC, NMI, ARI, and F1 (Chen et al. 2025b,a,c).

Performance Comparison

To evaluate the effectiveness of FedAI, we consider three experimental settings: 1) Comparison with the SOTA FGC method, i.e., FedGCN, under a fully unsupervised scenario. 2) Adaptation and comparison with representative federated classification and anomaly detection methods, i.e., FedSage (Zhang et al. 2021), GCFL (Xie et al. 2021), FedStar (Tan et al. 2023), LG-FGAD (Cai et al. 2024b), FGAD (Cai et al. 2024c), and AGDiff (Cai et al. 2025) under a unified unsu-

Datasets	Domain	Classes	Graphs	A. Nodes	A. Edges	N. A
MUTAG		2	188	17.93	19.79	-
BZR_MD		2	306	21.30	225.06	-
PTC_MR		2	344	14.29	14.69	-
BZR	SM	2	405	35.75	38.36	+ (3)
COX2		2	467	41.22	43.45	+ (3)
DHFR		2	756	42.43	44.54	+ (3)
AIDS		2	2000	15.69	16.20	+ (4)
DD	BIO	2	1178	284.32	715.66	-
PROTEINS		2	1113	39.06	72.82	+ (1)
SYNTHETIC	SY	2	300	100.00	196.00	+ (1)
COLORS-3		11	10500	61.31	91.03	+ (4)
REDDIT-M	SN	11	11929	391.41	456.89	-
Letter-high		15	2250	4.67	4.50	* (2)
Letter-low	CV	15	2250	4.68	3.13	+ (2)
Letter-med		15	2250	4.67	3.21	+ (2)

Table 2: Description of benchmark datasets. "+(v)" indicates that the graphs have v attribute. "–" means the graphs lack node attributes, while "*" denotes that the node attributes are manually removed from the dataset. "N. A" denotes the dimension of the node attributes.

pervised objective for cross-paradigm evaluation. 3) Comparison with supervised baseline methods using partial labels shows that FedAI achieves competitive performance without any labels. The results of the unsupervised methods for the first two settings are shown in Table 1, the results of the supervised methods are shown in Fig. 3.

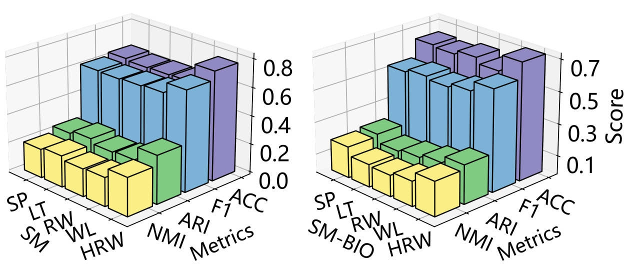

Figure 4: Graph Kernel ablation between HRW GK and other kernels: SP, LT, RW, and WL on two non-IID settings.

Discussion Based on the experimental results, we can draw the following conclusions: 1) Compared with existing advanced methods, FedAI achieves significant performance improvements, which fully verify that it can infer unknown attributes with high quality and effectively correct consensus bias through the collaboration of the two modules, thus significantly improving the overall clustering performance. 2) Compared to supervised methods with a small number of labels, FedAI still has certain performance advantages. These are mainly attributed to FedAI effectively inferring attribute information that is beneficial for clustering and enhances the negotiation quality of the server.

Scalability Evaluation

To further validate the scalability and robustness of FedAI, we conduct additional experiments on larger-scale datasets with an increased number of participating clients (16 clients). This non-IID setting includes MUTAG, BZR_MD, PTC_MR, BZR, COX2, DHFR, AIDS, DD, PROTEINS, SYNTHETIC, SW-620, OVCAR-8, PC-3, MOLT-4, SF-295, P388 datasets. The experimental results are shown in Table 3, which demonstrate that FedAI maintains stable convergence and consistently achieves superior clustering performance across various scales. This demonstrates that FedAI has a strong generalization ability to large-scale federated environments.

Models	ACC	NMI	ARI	F1
FedStar	62.6±3.1	13.9±2.5	16.3±1.5	47.6±2.4
FedGCN	65.3±2.3	20.8±1.7	21.4±1.3	52.7±2.0
Ours	**72.3±1.5**	**26.2±1.3**	**27.4±1.9**	**58.6±1.7**

Table 3: Performance comparison in terms of ACC, NMI, ARI, and F1 metrics under large-scale datasets and more clients.

Ablation Studies

Effect of KAIS To assess the effectiveness of KAIS, we compare its performance with KAIS both enabled and disabled. The experimental results presented in Table 1 ("Local" denotes that the KAIS mechanism is not enabled) lead to the following key observation: Enabling the KAIS mechanism results in a substantial performance improvement over the local method. This enhancement is primarily due to the fact that FedAI infers clustering-friendly attributes within the CNAs, effectively alleviating knowledge shift and improving overall clustering performance.

Settings	Models	ACC	NMI	ARI	F1
SM (7)	FedProx	73.5±2.0	25.5±2.3	9.3±2.2	67.5±1.8
	FedPer	76.2±1.9	24.7±2.8	14.3±3.0	53.7±2.1
	FedAvg	76.0±1.5	23.6±2.1	31.4±1.7	68.4±1.3
	Ours	**78.6±0.4**	**26.9±1.0**	**32.8±1.3**	**70.5±1.1**
SM-BIO (9)	FedProx	66.8±2.1	19.1±2.4	23.0±1.4	60.6±2.1
	FedPer	71.3±1.9	18.7±1.9	22.8±2.4	60.7±1.6
	FedAvg	69.8±1.5	17.5±1.8	21.4±2.0	61.4±1.7
	Ours	**73.2±1.7**	**20.0±4.2**	**23.2±2.4**	**62.3±4.7**
SM-BIO-SY (10)	FedProx	70.8±1.7	18.7±2.4	21.4±2.3	61.0±2.0
	FedPer	69.5±2.0	16.7±2.1	20.5±1.8	60.6±2.3
	FedAvg	69.8±2.4	17.2±2.0	19.7±1.5	60.2±1.6
	Ours	**72.3±1.1**	**21.2±1.9**	**22.3±2.6**	**61.2±1.3**

Table 4: Ablation study on three aggregation strategies in FedAI, evaluated by ACC, NMI, ARI, and F1 metrics.

Effect of Graph Kernels To further evaluate the performance of our proposed HRW, we select different graph kernels (e.g., SP (Shervashidze et al. 2011), LT (Johansson et al. 2014), WL (Borgwardt, Kriegel, and Hans-Peter 2005), RW (Kang, Tong, and Sun 2012)) to calculate the affinity between clients. The experimental results are illustrated in Fig. 4. The following conclusions can be summarized: HRW can achieve the best performance, which is mainly due to its strong multiple graph feature expression generated by ran-

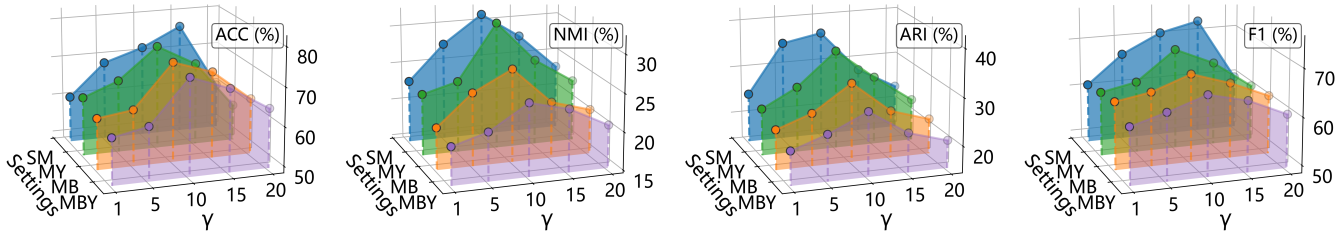

Figure 5: Hyperparameter γ sensitivity analysis results with variable $\gamma \in [1, 20]\%$, reporting ACC, NMI, ARI, and F1.

dom walk paths of different lengths. In addition, the Laplace Position structure encoding encourages the model to capture multi-scale features effectively.

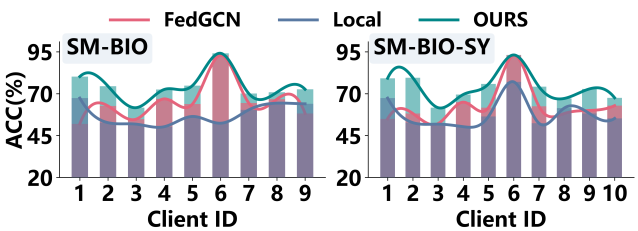

Figure 6: Client-wise ACC metric results under SM-BIO and SM-BIO-SY non-IID settings.

Parameter Aggregation Strategy Studies To verify the adaptability of FedAI to different global aggregation strategies for the model parameters. We evaluate the performance under various aggregation strategies (e.g., FedAVG (Li et al. 2019), FedProx (Li et al. 2020), FedPer (Arivazhagan et al. 2019)). The experimental results are shown in Table 4 and the following can be observed: 1) FedAI demonstrates substantial compatibility with various parameter aggregation strategies, and its consistently improved performance highlights the pivotal role of precise attribute inference in enhancing overall model performance. Even with the adoption of diverse strategies, the clustering performance of FedAI outperforms existing SOTA methods. 2) A personalized aggregation strategy (i.e., the one we used) can improve the clustering performance, however, suboptimal aggregation can disrupt inter-client consensus, diminishing overall clustering performance.

Influence of Hyperparameter γ γ is used to control the proportion of uploaded representational samples. We analyze its effect by varying its value across different settings. As illustrated in Fig. 5, the experimental results lead to the following key observations: Setting γ to 10% yields the best performance on most datasets. A smaller proportion fails to capture comprehensive client characteristics, while a larger proportion introduces extra noise due to decreased confidence in the uploaded samples.

Client-wise Performance Analysis

To validate that FedAI achieves comprehensive performance enhancement across all clients without compromising individual performance, we conduct client-wise comparative analyses of clustering performance among three methods, compared with FedGCN. The experimental results are shown in Fig. 6, and the following conclusions are obtained: FedAI can significantly improve the performance of all clients, whether the client lacks attributes or not, demonstrating its effectiveness in correcting global knowledge drift.

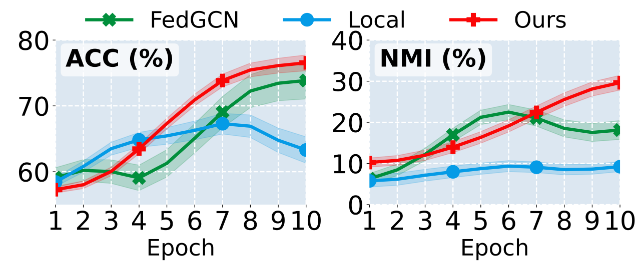

Figure 7: Convergence study under the SM non-IID setting, ACC and NMI employed as evaluation metrics.

Convergence Study

We present the performance trajectory of FedAI over 20 communication rounds to facilitate a comprehensive analysis of convergence speed and stability. As shown in Fig. 7, the experimental results demonstrate that, compared to existing SOTA methods, FedAI exhibits significantly faster convergence and greater training stability. This improvement is primarily attributed to the use of inferred attributes to achieve global consensus, effectively mitigating the issue of consensus failures caused by client heterogeneity.

Conclusion

In this paper, we investigate a new task, i.e., federated graph-level clustering with partial clients whose node attributes are missing. It is the first attempt to achieve cross-dataset/domain attribute inference under the circumstances of federated graph-level learning. To address this issue, we propose FedAI to alternatively optimize clustering and inference. In FedAI, the proposed KAIS can accurately infer unknown attributes to enhance the quality of global consensus for better clustering. Extensive experiments verify that FedAI consistently outperforms existing SOTA FGL methods. In the future, we will explore extending this method to the graph level, aiming to establish a more generalized FGL framework.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (NSFC) (Grant No. 62562026, 62506102), the Key Research and Development Program of Hainan Province (Grant No. ZDYF2024GXJS014, ZDYF2023GXJS163), and Collaborative Innovation Project of Hainan University (Grant No. XTCX2022XXB02).

References

Ambekar, S.; Yao, Y.; Li, R.; and Joe-Wong, C. 2024. FedGAT: A Privacy-Preserving Federated Approximation Algorithm for Graph Attention Networks. *arXiv preprint arXiv:2412.16144*.

Arivazhagan, M. G.; Aggarwal, V.; Singh, A. K.; and Choudhary, S. 2019. Federated Learning with Personalization Layers. *arXiv preprint arXiv:1912.00818*.

Borgwardt, K. M.; Kriegel; and Hans-Peter. 2005. Shortest-path Kernels on Graphs. In *Proceedings of the IEEE International Conference on Data Mining*, 74–81.

Cai, J.; Han, Y.; Guo, W.; and Fan, J. 2024a. Deep Graph-level Clustering using Pseudo-label-guided Mutual Information Maximization Network. *Neural Computing and Applications*, 36(16): 9551–9566.

Cai, J.; Zhang, Y.; and Fan, J. 2025. Self-Discriminative Modeling for Anomalous Graph Detection. In *Proceedings of the International Conference on Machine Learning*, 6387–6408.

Cai, J.; Zhang, Y.; Fan, J.; and Ng, S.-K. 2024b. Lg-fgad: An effective federated graph anomaly detection framework. In *Proceedings of the International Joint Conference on Artificial Intelligence*, 3760–3769.

Cai, J.; Zhang, Y.; Liu, F.; and Ng, S.-K. 2025. Leveraging Diffusion Model as Pseudo-Anomalous Graph Generator for Graph-Level Anomaly Detection. In *Proceedings of the International Conference on Machine Learning*, 3782–3794.

Cai, J.; Zhang, Y.; Lu, Z.; Guo, W.; and Ng, S.-k. 2024c. FGAD: Self-boosted Knowledge Distillation for An Effective Federated Graph Anomaly Detection Framework. *arXiv:2402.12761*.

Chen, M.-S.; Lai, P.-Y.; Liao, D.-Z.; Wang, C.-D.; and Lai, J.-H. 2025a. Graph Prompt Clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(7): 5794–5805.

Chen, M.-S.; Lai, P.-Y.; Liao, D.-Z.; Wang, C.-D.; and Lai, J.-H. 2025b. Homophily Induced Contrastive Attributed Graph Clustering. *IEEE Transactions on Circuits and Systems for Video Technology*, 35(10): 10213–10224.

Chen, M.-S.; Lin, J.-Q.; Wang, C.-D.; Huang, D.; and Lai, J.-H. 2025c. Contrastive Ensemble Clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 36(8): 14678–14690.

Chen, X.; Chen, S.; Yao, J.; Zheng, H.; Zhang, Y.; and Tsang, I. W. 2022. Learning on Attribute-Missing Graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(2): 740–757.

Gurumurthy, K.; Pal, H.; and Sharma, C. 2025. Federated Spectral Graph Transformers Meet Neural Ordinary Differential Equations for Non-IID Graphs. *arXiv preprint arXiv:2504.11808*.

Jin, D.; Wang, R.; Wang, T.; He, D.; Ding, W.; Huang, Y.; Wang, L.; and Pedrycz, W. 2022. Amer: A new Attribute-missing Network Embedding Approach. *IEEE Transactions on Cybernetics*, 53(7): 4306–4319.

Johansson, F.; Jethava, V.; Dubhashi, D.; and Bhattacharyya, C. 2014. Global Graph Kernels using Geometric Embeddings. In *Proceedings of the International Conference on Machine Learning*, 694–702.

Kang, U.; Tong, H.; and Sun, J. 2012. Fast Random Walk Graph Kernel. In *Proceedings of the SIAM International Conference on Data Mining*, 828–838.

Li, S.; Wang, C.; Xu, K.; Li, X.; He, G.; and Liu, X. 2025. Mutual-View Contrastive Generative Framework for Attribute-Missing Graph Clustering. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 1–5.

Li, T.; Sahu, A. K.; Zaheer, M.; Sanjabi, M.; Talwalkar, A.; and Smith, V. 2020. Federated Optimization in Heterogeneous Networks. In *Proceedings of the Machine Learning and Systems*, 429–450.

Li, X.; Agrawal, G.; and Ramnath, R. 2024. Federated Contrastive Learning of Graph-Level Representations. In *Proceedings of the IEEE International Conference on Big Data*, 747–752.

Li, X.; Huang, K.; Yang, W.; Wang, S.; and Zhang, Z. 2019. On the Convergence of Fedavg on Non-iid Data. *arXiv preprint arXiv:1907.02189*.

Li, X.; Wu, Z.; Zhang, W.; Zhu, Y.; Li, R.-H.; and Wang, G. 2024. Fedgta: Topology-aware Averaging for Federated Graph Learning. *arXiv preprint arXiv:2401.11755*.

Liang, K.; Meng, L.; Li, H.; Wang, J.; Lan, L.; Li, M.; and Liu, X. 2025. From Concrete to Abstract: Multi-View Clustering on Relational Knowledge. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(10): 9043–9060.

Liang, K.; Meng, L.; Liu, M.; Liu, Y.; Tu, W.; Wang, S.; Zhou, S.; Liu, X.; Sun, F.; and He, K. 2024a. A Survey of Knowledge Graph Reasoning on Graph Types: Static, Dynamic, and Multi-Modal. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12): 9456–9478.

Liang, K.; Meng, L.; Liu, Y.; Liu, M.; Wei, W.; Liu, S.; and et al. 2024b. Simple Yet Effective: Structure Guided Pre-trained Transformer for Multi-modal Knowledge Graph Reasoning. In *Proceedings of the ACM International Conference on Multimedia*, 1554–1563.

Liang, K.; Zhou, S.; Liu, M.; Liu, Y.; Tu, W.; Zhang, Y.; Fang, L.; Liu, Z.; and Liu, X. 2024c. Hawkes-Enhanced Spatial-Temporal Hypergraph Contrastive Learning Based on Criminal Correlations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 8733–8741.

Liao, T.; Fu, L.; Chen, J.; Wang, Z.; Zheng, Z.; and Chen, C. 2024. A Swiss Army Knife for Heterogeneous Federated

Learning: Flexible Coupling via Trace Norm. In *Proceedings of the Conference on Neural Information Processing Systems*, 139886–139911.

Liao, T.; Wu, Z.; Chen, C.; and Zheng, Z. 2023. Tensor Completion via Convolutional Sparse Coding with Small Samples-Based Training. *Pattern Recognition*, 141: 109624.

Liao, T.; Xu, Z.; Hu, Q.; Dai, H.-N.; Huang, H.; Zheng, Z.; and Chen, C. 2025. FedBRB: A Solution to the Small-to-Large Scenario in Device-Heterogeneity Federated Learning. *IEEE Transactions on Mobile Computing*, 1–14.

Liu, J.; Cheng, J.; Han, R.; Tu, W.; Wang, J.; and Peng, X. 2025a. Federated Graph-Level Clustering Network. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 18870–18878.

Liu, J.; Han, R.; Tu, W.; Wang, H.; Wu, J.; and Cheng, J. 2025b. Federated Node-Level Clustering Network with Cross-Subgraph Link Mending. In *Proceedings of the International Conference on Machine Learning*, 38540–38556.

Liu, M.; Liang, K.; Hu, D.; Yu, H.; Liu, Y.; Meng, L.; Tu, W.; Zhou, S.; and Liu, X. 2023. TMAC: Temporal Multi-Modal Graph Learning for Acoustic Event Classification. In *Proceedings of the ACM International Conference on Multimedia*, 3365–3374.

Liu, M.; Liang, K.; Wang, S.; Hu, X.; Zhou, S.; and Liu, X. 2025c. Deep Temporal Graph Clustering: A Comprehensive Benchmark and Datasets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(12): 11561–11578.

Liu, M.; Liang, K.; Yu, H.; Meng, L.; Wang, S.; Zhou, S.; and Liu, X. 2025d. Multiview Temporal Graph Clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 36(10): 18383–18396.

Liu, R.; Xing, P.; Deng, Z.; Li, A.; Guan, C.; and Yu, H. 2024a. Federated Graph Neural Networks: Overview, Techniques, and Challenges. *IEEE Transactions on Neural Networks and Learning Systems*, 36: 4279–4295.

Liu, S.; Liang, K.; Dong, Z.; Wang, S.; Yang, X.; Zhou, S.; Zhu, E.; and Liu, X. 2024b. Learn from View Correlation: An Anchor Enhancement Strategy for Multi-view Clustering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 26151–26161.

Liu, Z.; Yang, D.; Wang, S.; and Su, H. 2024c. Adaptive Multi-Channel Bayesian Graph Attention Network for IoT Transaction Security. *Digital Communications and Networks*, 10(3): 631–644.

Morris, C.; Kriege, N. M.; Bause, F.; Kersting, K.; Mutzel, P.; and Neumann, M. 2020. TUDataset: A Collection of Benchmark Datasets for Learning with Graphs. In *Proceedings of the ICML Workshop on Graph Representation Learning and Beyond*.

Shervashidze, N.; Schweitzer, P.; Van Leeuwen, E. J.; Mehlhorn, K.; and Borgwardt, K. M. 2011. Weisfeiler-Lehman Graph Kernels. *Journal of Machine Learning Research*, 12: 2539–2561.

Song, S.; Zheng, H.; Hu, Z.; Zheng, M.; Yang, L.; and Xu, A. 2025. GradPFL: Gradient-Driven Adaptive Clustering in Personalized Federated Learning. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 1–5.

Tan, Y.; Liu, Y.; Long, G.; Jiang, J.; Lu, Q.; and Zhang, C. 2023. Federated Learning on Non-IID Graphs via Structural Knowledge Sharing. In *Proceedings of the AAAI conference on Artificial Intelligence*, 9953–9961.

Tu, W.; Guan, R.; Zhou, S.; Ma, C.; Peng, X.; Cai, Z.; Liu, Z.; Cheng, J.; and Liu, X. 2024. Attribute-Missing Graph Clustering Network. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 15392–15401.

Tu, W.; Zhou, S.; Liu, X.; Cai, Z.; Zhao, Y.; Liu, Y.; and He, K. 2025. WAGE: Weight-Sharing Attribute-Missing Graph Autoencoder. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(7): 5760–5777.

Wang, J.; Tu, W.; and Cheng, J. 2025. Hierarchical Shortest-Path Graph Kernel Network. In *Preceedings of the Conference on Neural Information Processing Systems*.

Wang, X.; Dong, Y.; Jin, D.; Li, Y.; Wang, L.; and Dang, J. 2023. Augmenting Affective Dependency Graph via Iterative Incongruity Graph Learning for Sarcasm Detection. In *Proceedings of the AAAI conference on artificial intelligence*, 4702–4710.

Wang, X.; Sun, R.; Zhang, Y.; Feng, B.; He, D.; Wang, L.; and Jin, D. 2025a. Stealthy Yet Effective: Distribution-Preserving Backdoor Attacks on Graph Classification. In *Proceedings of the Conference on Neural Information Processing Systems*.

Wang, X.; Wang, Y.; He, D.; Yu, Z.; Li, Y.; Wang, L.; Dang, J.; and Jin, D. 2025b. Elevating Knowledge-Enhanced Entity and Relationship Understanding for Sarcasm Detection. *IEEE Transactions on Knowledge and Data Engineering*, 37: 3365–3371.

Wu, J.; Wang, H.; Liu, J.; Tu, W.; and et al. 2025. FedPKA: Federated Graph-Level Clustering Network with Personalized Knowledge Aggregation. In *Proceedings of the International Conference on Intelligent Computing*, 27–38.

Xia, Z.; and Zhang, X. 2024. Federated Graph Augmentation for Semisupervised Node Classification. *IEEE Transactions on Computational Social Systems*, 3232–3242.

Xie, H.; Ma, J.; Xiong, L.; and Yang, C. 2021. Federated Graph Classification over Non-IID Graphs. In *Proceedings of the Conference on Neural Information Processing Systems*, 18839–18852.

Xie, X.; Li, B.; Pan, E.; Guo, Z.; Kang, Z.; and Chen, W. 2025. One Node One Model: Featuring the Missing-Half for Graph Clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 21688–21696.

Yang, D.; Wang, Y.; Cai, Z.; and Li, Y. 2024. Spectrum Prediction via Graph Structure Learning. In *Prceedings of the IEEE Vehicular Technology*, 1–5.

Zhang, K.; Yang, C.; Li, X.; and et al. 2021. Subgraph Federated Learning with Missing Neighbor Generation. In *Proceedings of the Conference on Neural Information Processing Systems*, 6671–6682.