

Article

Not peer-reviewed version

Federated Graph Neural Networks for Heterogeneous Graphs with Data Privacy and Structural Consistency

[Haifeng Yang](#) ^{*} , Mengjie Wang , Linyan Dai , Yan Wu , [Junliang Du](#) ^{*}

Posted Date: 8 August 2025

doi: [10.20944/preprints202508.0540.v1](https://doi.org/10.20944/preprints202508.0540.v1)

Keywords: federated graph learning; structural alignment; consistency modeling; multi-source heterogeneous graphs



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Federated Graph Neural Networks for Heterogeneous Graphs with Data Privacy and Structural Consistency

Haifeng Yang ^{1,*}, Mengjie Wang ², Linyan Dai ³, Yan Wu ⁴ and Junliang Du ^{5,*}

¹ Northeastern University, Boston, USA

² New York University, New York, USA

³ University of California, Davis Sacramento, USA

⁴ University of Delaware, Newark, USA

⁵ Shanghai Jiao Tong University, Shanghai, China

* Correspondence: yang.haif@northeastern.edu (H.Y.); jldu@acm.org (J.D.)

Abstract

This paper addresses the problem of joint modeling for multi-source heterogeneous graph data in distributed environments by proposing a federated graph neural network classification framework driven by structural alignment and consistency regularization. The method preserves data locality by enabling each participant to learn node features and topological information through a local graph neural network encoder. A cross-source structural alignment module maps embeddings from different graphs into a shared representation space, mitigating semantic inconsistencies caused by structural differences. Additionally, a consistency regularization mechanism is introduced to enhance the robustness of node representations through multi-view perturbations, improving the model's generalization ability during training. At the global level, a federated averaging strategy is adopted to periodically aggregate local models, enabling collaborative optimization and enhancing the consistency and discriminative capacity of the global representation. To validate the effectiveness of the proposed approach, experiments are conducted in a multi-source heterogeneous graph environment using various node distribution strategies. The results show that the method outperforms existing federated graph learning models in terms of accuracy, clustering consistency, and structural expressiveness, achieving efficient multi-source graph classification while preserving data privacy.

Keywords: federated graph learning; structural alignment; consistency modeling; multi-source heterogeneous graphs

1. Introduction

With the continuous evolution of big data environments, the integration and mining of massive heterogeneous information have become central challenges in the construction of intelligent systems. Due to the diversity of data sources, multi-source data exhibit significant differences in structure, semantic granularity, and noise interference. These characteristics make it difficult for traditional centralized modeling approaches to achieve high-quality information fusion and knowledge discovery while ensuring both privacy and generalizability. In highly sensitive domains such as finance [1], healthcare [2], and large language models [3–6], achieving cross-platform and cross-institution collaborative learning that satisfies task demands and guarantees data security has emerged as a critical problem in intelligent mining systems [7].

Federated learning, as an emerging distributed training paradigm, enables collaborative model optimization without the need to share raw data. It effectively mitigates the conflict between data silos and privacy protection. However, most existing federated learning frameworks are designed for



structured or vectorized data and are not well-suited to modeling complex relational data. In real-world applications such as cloud networks, backend systems, and distributed systems, data are often naturally represented in the form of graphs [8–12]. The non-Euclidean relationships among nodes carry essential contextual and semantic dependencies. Therefore, integrating graph neural networks into federated learning frameworks to enable distributed relational modeling holds great promise for enhancing the understanding and classification of cross-domain graph data.

Graph neural networks have demonstrated strong expressive power in modeling high-order dependencies and structural semantics among nodes [13]. They have been widely applied in tasks such as community detection, recommendation systems, and risk analysis. Nevertheless, current GNN approaches typically assume that the entire graph is fully observable within a single platform, lacking the capacity to accommodate heterogeneous graph information from multiple sources. In distributed environments, graphs often suffer from missing structures, inconsistent edge weights, and incomplete node attributes, which negatively impact model generalization and robustness. Moreover, semantic shifts and distribution mismatches frequently exist across multi-source graphs, making it difficult for conventional training strategies to capture stable and transferable graph representations. This necessitates the development of GNN mechanisms that are adaptable to non-independent and identically distributed scenarios under a federated setting.

In this context, exploring collaborative mining strategies for multi-source heterogeneous graph data can alleviate the limitations of single-source modeling and enhance global representational capacity through cross-source complementarity. In real-world scenarios where data is limited and labels are scarce, the integration of multiple information sources offers a strong foundation for methods such as weak supervision, transfer learning, and semi-supervised learning [14]. By employing a federated learning mechanism to construct a graph learning framework involving multiple parties, it becomes possible to share structural knowledge and abstract patterns across sources while preserving local data privacy. Such a framework provides a practical path for collaborative classification and relational inference in multi-task environments. The combination of structure-aware modeling and collaborative optimization represents a promising learning paradigm for the next generation of distributed intelligent systems, advancing capabilities in interpretability, security, and adaptability.

In summary, addressing the challenges of structural heterogeneity, privacy sensitivity, and distribution shift in multi-source data fusion requires the design of a GNN-based classification framework with federated optimization capabilities, structural alignment mechanisms, and semantic collaborative representations. This research direction aligns with the data-driven modeling trends in complex scenarios and offers a solid foundation for improving the practicality and scalability of intelligent analysis systems in multi-party settings. By integrating key technologies such as graph modeling, federated collaboration, and heterogeneous alignment, the methodological and architectural contributions of this work hold significant theoretical and practical value for advancing intelligent mining applications in real-world multi-source environments [15].

2. Method

This network architecture illustrates a federated graph neural network classification framework designed for multi-source graph data, consisting of four main components: local encoding, structural alignment, consistency modeling, and global aggregation. Each participant locally performs GNN-based encoding and feature extraction, while a consistency module is introduced to enhance robustness across multiple views [16]. The structural alignment module maps local representations into a shared space. At the global level, federated aggregation and update mechanisms are applied to synchronize parameters across domains, ensuring collaborative model optimization under privacy-preserving constraints. The model architecture is shown in Figure 1.

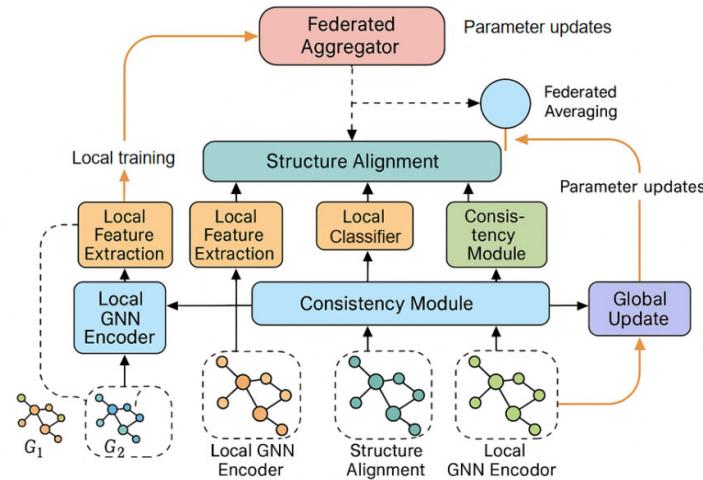


Figure 1. Federated Graph Classification with Alignment and Consistency Learning.

This study proposes a federated graph neural network classification mining framework for multi-source data environments to solve the graph representation learning problem under the coexistence of inconsistent cross-domain graph structure distribution and privacy protection [17]. The overall architecture is based on federated parameter synchronization and builds a collaborative optimization mechanism for local graph embedding, structural alignment, and global aggregation [18]. At each participant's local, the model receives its private graph data $G_i = (V_i, \mathcal{E}_i, X_i)$, where V_i is the node set, \mathcal{E}_i is the edge set, and $X_i \in R^{|V_i| \times d}$ is the node feature matrix. Preliminary feature propagation is achieved through graph convolution operations, and the node representation is updated as follows:

$$H_i^{(l+1)} = \sigma(\hat{A}_i H_i^{(l)} W^{(l)})$$

Where \hat{A}_i is the normalized adjacency matrix, $W^{(l)}$ is the trainable weight of the l th layer, $\sigma(\cdot)$ is the activation function, and the initial state is $H_i^{(0)} = X_i$. Considering the topological offset and semantic heterogeneity between multi-source graph structures, this study introduces a structural alignment mapping function $f_{align} : H_i \rightarrow \tilde{H}_i$ so that the embedding distributions between different sources are aligned in the common latent space. The alignment process is constrained by minimizing the alignment loss between the global center embedding \bar{H} and the embeddings of each participant:

$$L_{align} = \sum_i^N \|\tilde{H}_i - \bar{H}\|_2^2, \text{ where } \bar{H} = \frac{1}{N} \sum_{i=1}^N \tilde{H}_i.$$

During the local training process, each participant performs classification supervision optimization based on the set $V_i^{lab} \subset V_i$ of graph nodes of its label part, using cross-entropy loss as the basic objective function:

$$L_{sup}^{(i)} = - \sum_{v \in V_i^{lab}} \sum_{c=1}^C y_{v,c} \log(\hat{y}_{v,c})$$

$y_{v,c}$ is the true label of the node v in the category c , and $\hat{y}_{v,c}$ is the model prediction probability. After local training is completed, the model synchronizes global parameters through the federated averaging mechanism:

$$W^{(l)} \leftarrow \sum_{i=1}^N \frac{|V_i|}{\sum_j |V_j|} W_i^{(l)}$$

Ensure that the parameters of each layer can integrate multi-source knowledge while maintaining privacy. This optimization strategy iterates between local training, structural alignment, and global aggregation, gradually enhancing the model's structural expression and classification generalization capabilities for cross-domain graph data.

To improve the robustness and generalization of the global representation, the framework design introduces a consistency regularization term to constrain the consistency behavior of embedding under different perspectives. Let $H_i^{(a)}$ and $H_i^{(b)}$ be two embedding views generated by different perturbation strategies on the source i , then the consistency loss is defined as:

$$L_{cons} = \sum_{i=1}^N \|H_i^{(a)} - H_i^{(b)}\|_2^2$$

Finally, the model is solved end-to-end with a joint optimization objective function during the training phase, and its total loss function is:

$$L_{total} = \sum_{i=1}^N (L_{sup}^{(i)} + \lambda_1 L_{align}^{(i)} + \lambda_2 L_{cons}^{(i)})$$

Among them, λ_1 and λ_2 are weight coefficients, which are used to adjust the proportion of each loss in the overall optimization. By introducing structural alignment, consistency constraints, and federated average strategies, this method realizes collaborative classification modeling of multi-source graph representation while ensuring data privacy, providing a new paradigm support for distributed intelligent analysis under complex network data.

3. Performance Evaluation

3.1. Dataset

This study adopts the OGBN-ArXiv dataset as the primary basis for experimental evaluation. The dataset is part of the Open Graph Benchmark and contains a large-scale citation network composed of academic papers and their citation links across various research domains. Each node in the graph represents a paper, and edges represent citation relationships. Node features include textual content represented by word embeddings, while labels correspond to the paper's subject area, covering a total of 40 disciplines.

The dataset exhibits natural cross-domain structural differences and semantic heterogeneity. It can be partitioned into multiple subgraphs to simulate a multi-source data environment. For example, the graph can be divided based on publication years or subject labels. Each participant holds only a portion of the graph structure and specific class distributions, resulting in non-independent and identically distributed (Non-IID) data partitions. This setting aligns with the modeling requirements of federated graph learning.

By constructing a heterogeneous multi-participant graph environment based on the OGBN-ArXiv dataset and performing classification tasks, the proposed method can be effectively evaluated. This setup allows for assessing the model's robustness and generalization under scenarios involving missing structural information, distribution shifts, and cross-source collaboration. It also provides practical support for multi-source knowledge mining applications.

3.2. Experimental Results

This paper first conducts a comparative experiment, and the experimental results are shown in Table 1.

Table 1. Comparative experimental results.

Method	Accuracy (%)	Macro-F1 (%)	NMI (%)	Homogeneity (%)
Ours (FedGCL)	78.6	74.2	56.5	54.2
FedSage+ [19]	72.6	68.4	49.7	47.2
GCN-FL [20]	74.1	69.5	50.9	48.6
FedGraphNN [21]	75.3	70.8	52.4	50.1
HierFGL [22]	76.3	71.2	53.0	51.3

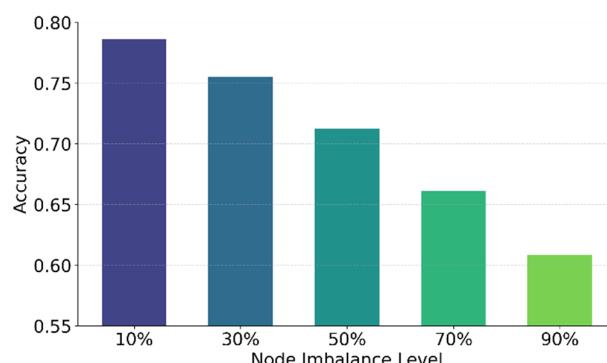
Experimental results show that the proposed federated graph neural network classification method (FedGCL) significantly outperforms existing mainstream models across multiple evaluation metrics. In particular, it demonstrates superior performance in Accuracy and macro-F1, two critical indicators. These results validate the effectiveness of the proposed structural alignment mechanism and consistency regularization strategy in modeling multi-source graph structures. The method effectively mitigates issues related to graph heterogeneity and feature distribution inconsistency across participants, thereby improving global classification performance.

From a horizontal comparison, early federated graph learning methods such as FedSage+ and GCN-FL enable basic distributed training. However, due to the lack of cross-source structural adaptation, they perform poorly in structural metrics like NMI and Homogeneity. This suggests their limited capacity in capturing global structural relationships. In contrast, models like HierFGL and FedGraphNN incorporate more complex synchronization mechanisms and feature-sharing strategies. While these approaches show improved performance, they still fall short in addressing the core challenge of multi-view consistency modeling in federated environments.

FedGCL's leading performance in NMI and Homogeneity further highlights its advantages in global structural alignment and semantic clustering. By introducing consistency constraints during local model training and applying structural alignment in the global phase, FedGCL effectively integrates high-order structural relations across multi-source graphs. This allows the model to capture more discriminative federated graph representations while maintaining privacy separation, which is critical for real-world tasks such as cross-domain node classification.

In summary, FedGCL achieves a unified design that combines structural awareness, local trainability, and global collaborative optimization. This reflects a methodological advancement in federated graph neural networks under heterogeneous multi-source scenarios. Its improved performance across key metrics not only demonstrates the model's capacity but also confirms the importance of unified structural spaces and consistent representation strategies in complex federated settings.

This paper also experiments on the impact of the size of the representation dimension on the uneven distribution of multi-source data nodes on the classification performance. The experimental results are shown in Figure 2.

**Figure 2.** Impact of Unbalanced Distribution of Multi-Source Data Nodes on Classification Performance.

Experimental results indicate that as the degree of node distribution imbalance increases, the model's classification accuracy shows a clear downward trend. This observation suggests that in scenarios where multi-source data have inconsistent class distributions, the global model's discriminative ability is significantly affected. When the data becomes highly skewed, the model struggles to learn adequate class boundaries, leading to degraded performance.

When the imbalance ranges from 10% to 30%, the model maintains relatively stable performance. This stability is attributed to the structural alignment mechanism, which helps mitigate the effects of local data bias. At this stage, there is still a degree of class overlap among participants. This allows the federated synchronization process to capture some global patterns, resulting in a more generalizable aggregated model.

However, when the imbalance exceeds 50%, classification accuracy declines sharply. This indicates that in the absence of sufficient class sharing, consistency regularization, and local structural information alone are insufficient for high-quality representation learning. The distribution shift weakens the convergence capacity of the structural alignment module in the embedding space and limits the model's ability to recognize global patterns after aggregation.

Further observations reveal that under severe imbalance, the model tends to overfit structural features, especially for classes with abundant data, while neglecting the learning of long-tail classes. This finding highlights the importance of designing client partition strategies that ensure diverse coverage. Such strategies are essential for improving the model's stability and robustness in federated graph settings with heterogeneous node distributions.

This paper also experiments on the impact of the overlap rate of node categories in multi-source graphs on the generalization ability of the model. The experimental results are shown in Figure 3.

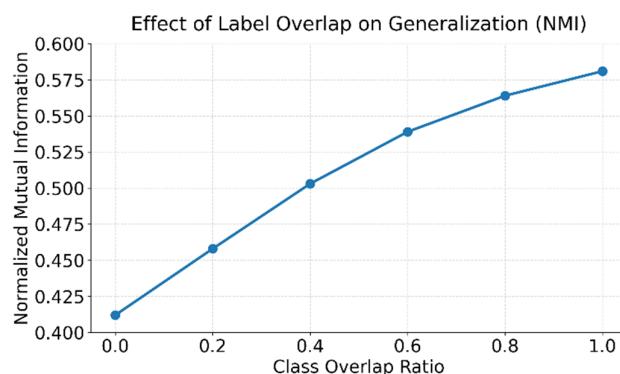


Figure 3. The impact of multi-source graph node category overlap rate on model generalization ability.

Experimental results show that as the overlap rate of node classes among multi-source graphs increases, the model's generalization ability improves consistently, as indicated by a gradual rise in the NMI metric. This suggests that in federated graph neural networks, the degree of class sharing among participants significantly affects the consistency of the global representation space. When more node labels are shared across participants, the model can learn more coherent semantic boundaries, which leads to more stable joint representations during aggregation.

At low levels of class overlap, the label spaces across clients have little intersection. As a result, the structural alignment module struggles to establish effective semantic mappings. This causes fragmentation in the embedding space and reduces the model's global clustering performance. In such non-shared label scenarios, the limited information flow during federated optimization hinders semantic consistency modeling and weakens the convergence of structural alignment.

As the overlap rate increases, semantic similarity across sources becomes stronger. This alleviates the non-independent and identically distributed (Non-IID) nature of federated learning. In this phase, the consistency constraint more effectively captures cross-source representation

differences and reaches convergence. Consequently, the model achieves better structural separation and semantic compactness in the global clustering space, resulting in clearer cluster boundaries.

When the overlap reaches a moderate to high level, the improvement in generalization begins to saturate. This indicates that moderate class sharing is crucial for enhancing federated graph modeling, rather than complete label space alignment. Introducing local sharing mechanisms while preserving data privacy is a key strategy for effective classification in heterogeneous multi-source graph environments.

4. Conclusions

This study focuses on the challenge of graph structure modeling in multi-source collaborative scenarios. It proposes a federated graph neural network classification framework that integrates structural alignment and consistency constraints. The framework aims to address the conflicts among data heterogeneity, privacy preservation, and expressive power in distributed environments. By jointly designing local graph neural embeddings, structural consistency mappings, and federated synchronization mechanisms, the method enables effective representation and classification of nodes across multiple sources without sharing raw data. This enhances the model's generalization and robustness in handling heterogeneous graph information.

In the proposed framework, a learnable structural alignment module is introduced to alleviate structural discrepancies among multi-source graphs. This allows different clients to express semantic features collaboratively in a shared space during the federated optimization process. At the same time, a consistency regularization mechanism improves the model's robustness against view perturbations and stabilizes the learned representations in the global embedding space. Experimental results demonstrate that the method achieves superior classification performance across multiple benchmark graph datasets. These findings highlight its effectiveness in modeling complex structural dependencies under privacy constraints.

This research provides a new modeling paradigm at the intersection of federated learning, graph neural networks, and multi-source heterogeneous data mining. It offers meaningful insights for application scenarios such as collaborative healthcare, financial risk control, and cross-domain recommendation, where both data privacy and modeling performance must be preserved. In particular, the proposed framework lays a theoretical and engineering foundation for decentralized structure understanding and decision-making in environments where data cannot be freely shared across institutions.

5. Future Research

Future research may extend the model's adaptability to dynamic graphs, attribute-heterogeneous graphs, and cross-modal graph structures. It may also explore personalized optimization strategies under federated learning to improve modeling efficiency in non-independent and identically distributed environments. Moreover, in large-scale scenarios with increasing node volumes and client numbers, reducing communication costs, improving aggregation efficiency, and enabling online incremental learning will be key directions for the practical deployment of federated graph neural networks.

References

1. T. Xu, X. Deng, X. Meng, H. Yang, and Y. Wu, "Clinical NLP with attention-based deep learning for multi-disease prediction," arXiv preprint arXiv:2507.01437, 2025.
2. X. Su, "Deep forecasting of stock prices via granularity-aware attention networks", 2024.
3. S. Lyu, Y. Deng, G. Liu, Z. Qi, and R. Wang, "Transferable modeling strategies for low-resource LLM tasks: A prompt and alignment-based," arXiv preprint arXiv:2507.00601, 2025.
4. Y. Wang, "Structured compression of large language models with sensitivity-aware pruning mechanisms," Journal of Computer Technology and Software, vol. 3, no. 9, 2024.

5. Q. Wu, "Task-aware structural reconfiguration for parameter-efficient fine-tuning of LLMs", 2024.
6. X. Quan, "Layer-wise structural mapping for efficient domain transfer in language model distillation," *Transactions on Computational and Scientific Methods*, vol. 4, no. 5, 2024.
7. X. Fu, B. Zhang, Y. Dong, et al., "Federated graph machine learning: A survey of concepts, techniques, and applications," *ACM SIGKDD Explorations Newsletter*, vol. 24, no. 2, pp. 32-47, 2022.
8. Y. Zou, N. Qi, Y. Deng, Z. Xue, M. Gong, and W. Zhang, "Autonomous resource management in microservice systems via reinforcement learning," *arXiv preprint arXiv:2507.12879*, 2025.
9. M. Gong, "Modeling microservice access patterns with multi-head attention and service semantics", 2025.
10. R. Meng, H. Wang, Y. Sun, Q. Wu, L. Lian, and R. Zhang, "Behavioral anomaly detection in distributed systems via federated contrastive learning," *arXiv preprint arXiv:2506.19246*, 2025.
11. Y. Ren, "Deep learning for root cause detection in distributed systems with structural encoding and multi-modal attention", 2024.
12. B. Fang and D. Gao, "Collaborative multi-agent reinforcement learning approach for elastic cloud resource scaling," *arXiv preprint arXiv:2507.00550*, 2025.
13. Deng, L.; Huang, Y.; Liu, X.; Liu, H.; Lu, Z. Graph2MDA: a multi-modal variational graph embedding model for predicting microbe-drug associations. *Bioinformatics* **2021**, *38*, 1118–1125, <https://doi.org/10.1093/bioinformatics/btab792>.
14. H. Cheng, H. Wen, X. Zhang, et al., "Contrastive continuity on augmentation stability rehearsal for continual self-supervised learning," *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 5707-5717, 2023.
15. J. Zhou and Y. Lei, "Multi-source heterogeneous data fusion algorithm based on federated learning," *Proceedings of the International Conference on Soft Computing in Data Science*, pp. 46-60, 2023.
16. Peng, S.; Zhang, X.; Zhou, L.; Wang, P. YOLO-CBD: Classroom Behavior Detection Method Based on Behavior Feature Extraction and Aggregation. *Sensors* **2025**, *25*, 3073, <https://doi.org/10.3390/s25103073>.
17. W. Zhu, "Fast adaptation pipeline for LLMs through structured gradient approximation," *Journal of Computer Technology and Software*, vol. 3, no. 6, 2024.
18. Z. Xu, K. Ma, Y. Liu, W. Sun, and Y. Liu, "Causal representation learning for robust anomaly detection in complex environments," 2025.
19. Z. Wang, W. Kuang, Y. Xie, et al., "FederatedScope-GNN: Towards a unified, comprehensive and efficient package for federated graph learning," *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 4110-4120, 2022.
20. Wu, C.; Wu, F.; Lyu, L.; Qi, T.; Huang, Y.; Xie, X. A federated graph neural network framework for privacy-preserving personalization. *Nat. Commun.* **2022**, *13*, 1–10, <https://doi.org/10.1038/s41467-022-30714-9>.
21. C. Wu, F. Wu, Y. Cao, et al., "FedGNN: Federated graph neural network for privacy-preserving recommendation," *arXiv preprint arXiv:2102.04925*, 2021.
22. R. Liu, P. Xing, Z. Deng, et al., "Federated graph neural networks: Overview, techniques, and challenges," *IEEE Transactions on Neural Networks and Learning Systems*, 2024.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.