# Learning to Collaborate: Personalized Federated GNNs

**Mohit Kataria**[1]* Mohit.Kataria@scai.iitd.ac.in
**Nikita Malik**[2]* bsz218185@iitd.ac.in
**Satyam Govind Shukla**[4]* shuklasatyam774@gmail.com
**Shreyash Bhilwade**[3] ee3210178@iitd.ac.in
**Sandeep Kumar**[2,1,3] ksandeep@ee.iitd.ac.in

[1] Yardi School of Artificial Intelligence, Indian Institute of Technology Delhi
[2] Bharti School of Telecommunication Technology and Management, Indian Institute of Technology Delhi
[3] Department of Electrical Engineering, Indian Institute of Technology Delhi
[4] Department of Computer Science, National Institute of Technology Karnataka

## Abstract

Federated graph learning enables organizations with distributed graph data to collaboratively train GNNs without sharing raw features. However, in node classification and related tasks, local graphs are typically non-IID in both structure and attributes, making a single global model suboptimal. We propose pFEDGNN, a personalized federated graph learning framework that *privately infers the global graph structure* across clients and then *derives a client-level collaboration graph* via coarsening principles. Concretely, we estimate a global Laplacian in a privacy-preserving manner and obtain the collaboration graph by lifting with a partition matrix ($L_C = P\, L_G\, P^\top$), where edge weights encode collaboration strength among clients. This structure guides graph-aware parameter aggregation, allowing clients with similar data distributions to share more while preserving personalization. Experiments on diverse graph benchmarks show that pFEDGNN significantly improves node classification performance over strong FL/pFL baselines; notably, our method *learns the collaboration graph in one shot*, reducing both communication and computation compared to iterative approaches.

## 1 Introduction

Graph-structured data arise in diverse domains such as social networks [1], molecular chemistry [2], and biomedical knowledge graphs [3], where effective analysis is vital for applications like drug discovery, disease diagnosis, and social behavior modeling. Graph Neural Networks (GNNs) have become state-of-the-art for these tasks [4], learning node, edge, and graph level representations that surpass traditional methods. In many practical scenarios, graph data are inherently distributed across multiple organizations or data silos [5]. For example, hospitals hold patient records [6], pharmaceutical labs maintain molecular datasets [7], and financial institutions control transaction networks [8]. In such settings, privacy and regulatory constraints prohibit direct sharing, yet most graph-based semi-supervised methods assume centralized access to raw data [9], raising privacy concerns. A natural solution is Federated Learning (FL) with GNNs [10], where multiple clients train local models on limited labeled data, and a server aggregates them into a global model in a privacy-preserving manner.

FL [11] enables collaborative training without exposing raw data: clients perform local updates and only share model parameters or gradients, aggregated by a central server. This paradigm, pioneered by [12, 13], allows multiple parties to jointly train models while preserving data locality. While FL mitigates direct privacy risks, it faces unique challenges on graph data [14]. First, the global

*Equal contribution.

, Learning to Collaborate: Personalized Federated GNNs (Extended Abstract). Presented at the Fourth Learning on Graphs Conference (LoG 2025), Hybrid Event, December 10–12, 2025.

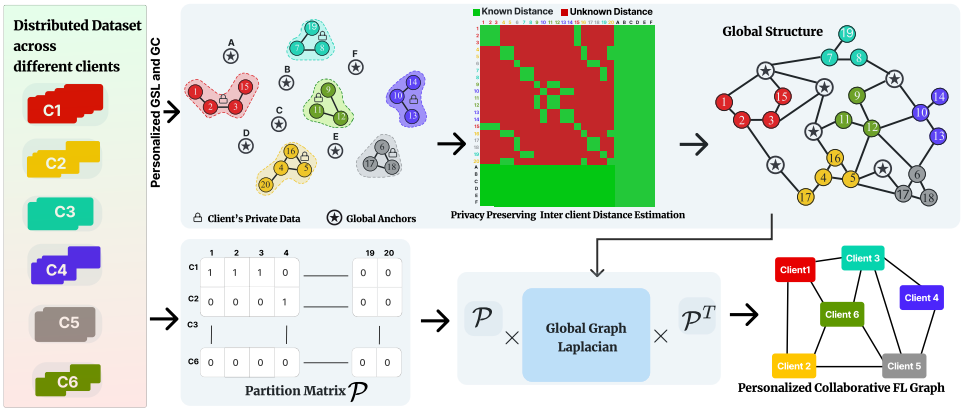



**Figure 1:** Pipeline of *pFedGNN*. Client datasets remain private, yielding an incomplete distance matrix where inter-client entries are unknown. From this partial information, we infer the global graph structure via the Laplacian $L$. In parallel, we learn the partition matrix $P$, which maps data points to clients. Using both, the client-level collaboration graph is obtained as $PLP^\top$, which then guides personalized federated learning.

structure of the graph essential for many GNN-based tasks is unavailable, as edges between entities residing in different clients cannot be directly observed. This hampers the ability to capture cross-client dependencies. Second, graph data are inherently heterogeneous: clients differ in distributions, structural patterns, and label spaces, causing divergent local updates and degrading the performance of a single global model. To address these issues, Personalized FL (pFL) has emerged as a promising paradigm. Early methods [15–17] learn multiple global or cluster-specific models for groups of similar clients, while recent approaches [18–21] emphasize client-specific personalization, where the global model serves primarily for knowledge transfer. Thus, FL's objective has shifted from producing one global model to enabling many high-performing, client-specific models while retaining collaboration benefits.

A key open problem is determining which clients should collaborate and to what extent. Relationships among participants can be naturally modeled as a graph, where nodes denote clients and edges capture potential collaboration (e.g., similarity or trust). However, balancing individual utility with collaborative gains remains difficult [22], since heterogeneity is unknown under privacy constraints. Existing methods lack mechanisms to adaptively decide collaboration intensity, limiting their flexibility in handling diverse heterogeneity or malicious clients. We address this gap with a three-stage framework (see Figure 1) for personalized federated graph learning: (i) infer a privacy-preserving estimate of the global graph structure, (ii) derive a client-level collaboration graph via coarsening principles, and (iii) perform personalized federated learning guided by this collaboration graph. In the collaboration graph, each node represents a client's personalized GNN model, and edge weights quantify collaboration intensity based on model similarity. This serves as a proxy for data similarity, enabling adaptive, fine-grained collaboration.

## 2 Background and Related Work

**Graphs and GNNs.** A graph is denoted $\mathcal{G} = (\mathcal{V}, \mathcal{E}, X)$ with node set $\mathcal{V}$, edge set $\mathcal{E}$, and feature matrix $X \in \mathbb{R}^{|\mathcal{V}| \times d}$. Its structure is commonly represented by the adjacency matrix $\mathbf{A}$, the degree matrix $\mathbf{D}$, and the graph Laplacian $L = \mathbf{D} - \mathbf{A}$. GNNs [23–26] leverage this structure to learn node- and graph-level representations. A two-layer GCN [27] computes $f_\theta(\mathbf{X}, \mathbf{A}) = \mathrm{softmax}(\mathbf{A}\, \sigma(\mathbf{A}\mathbf{X}\mathbf{W}_1)\mathbf{W}_2)$, with parameters $\theta = \{\mathbf{W}_1, \mathbf{W}_2\}$. GNNs are widely used for semi-supervised node classification, link prediction, and graph-level learning across domains.

**Graph Coarsening (GC).** GC compresses a graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, X)$ with $N$ nodes and features $X \in \mathbb{R}^{N \times d}$ into a smaller graph $\mathcal{G}_c(\widetilde{\mathcal{V}}, \widetilde{\mathcal{E}}, \widetilde{X})$ with $n \ll N$ supernodes and features $\widetilde{X} \in \mathbb{R}^{n \times d}$. This is

done via a partition matrix $\mathcal{P} \in \mathbb{R}^{n \times N}$ that merges structurally or feature-similar nodes, producing $\widetilde{X} = \mathcal{P}X$ while approximately preserving key structural and spectral properties [28–31].

**FL and Personalized FL.** FL enables $M$ clients $C_1, \ldots, C_M$ to collaboratively train models without sharing raw data, exchanging updates with a central server instead [32–35]. While preserving privacy, standard FL targets a single global model, which underperforms on heterogeneous (non-IID) data. pFL mitigates this by adapting models to individual clients while retaining cross-client knowledge transfer. Approaches include regularization-based methods (Ditto [36], pFedMe [37]), clustering (CFL [38]), adaptive aggregation (FedAMP [39], FedRep [40]), hypernetworks (pFedHN [41]), and collaboration graphs (pFedGraph [42]). Yet, most rely on iterative similarity estimation or communication-heavy updates, motivating our one-shot, coarsening-inspired strategy for efficient collaboration graph inference.

## 3 Methodology

Our framework consists of three key stages: (i) inferring a privacy-preserving estimate of the global graph structure, (ii) deriving a client-level collaboration graph via coarsening, and (iii) performing personalized federated graph machine learning guided by this collaboration graph.

**Inference of Global Structure.** A central challenge in federated graph learning is that the global structure required for effective GNN training is inaccessible due to privacy constraints: raw data and cross-client edges cannot be shared. We address this using the anchor-based distance reconstruction framework from PPDA [43], which estimates global structure from incomplete, privacy-preserving distance measurements. Let $\mathcal{C} = \{C_1, \ldots, C_M\}$ be the set of clients, each holding private data $D_m = \{x_1^{(m)}, \ldots, x_{n_m}^{(m)}\} \subset \mathbb{R}^{d_h}$. The server provides $K$ public anchors $\mathcal{A} = \{a_1, \ldots, a_K\}$. Each client computes squared distances between its local data and the anchors, transmitting only these values:

$$D_\Omega(i,j) = \begin{cases} \|x_i^{(m)} - a_j\|_2^2, & (i,j) \in \Omega, \\ \text{unobserved}, & \text{otherwise}, \end{cases}$$

where $\Omega \subseteq [N] \times [K]$ is the set of observed distance entries (N is the total number of data points across all clients). Then structured matrix completion is used to recover the full distance matrix $\widehat{D}$ from $D_\Omega$. The global graph Laplacian is then estimated from an affinity matrix $\widehat{W}$ derived from $\widehat{D}$ as $L_G = \mathrm{Diag}(\widehat{W}\mathbf{1}) - \widehat{W}$. Thus, $L_G$ provides a faithful approximation of the global structure while ensuring rigorous privacy guarantees.

**Lemma 1 (Privacy-Preserving Global Structure, adapted from [43])** *Let $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^{d_h}$ be the union of all client data and $\mathcal{A}$ the set of $K$ anchors. If $K < d_h$, the original features $\{x_i\}$ cannot be uniquely reconstructed from $\{\|x_i - a_j\|_2^2\}_{i,j}$. Hence, $L_G$ preserves global structure without exposing private data.*

**Inference of the Collaboration Graph.** Given the global Laplacian $L_G$, we infer collaboration relationships among clients using graph coarsening. Let $P \in \{0,1\}^{M \times N}$ denote the partition matrix mapping each of the $N$ data points to one of the $M$ clients. The client-level (coarsened) Laplacian is $L_C = P L_G P^\top$ [28, 29, 31].

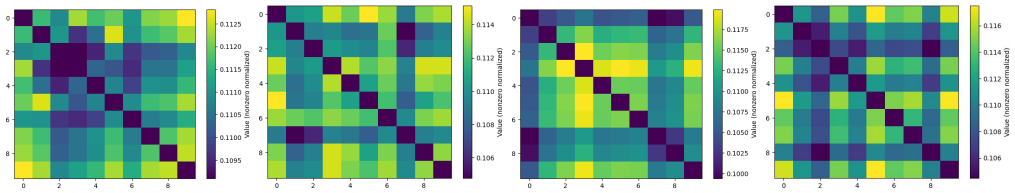
**Remark 1 (Collaboration Graph via Coarsening)** *With $L_G$ as the global Laplacian and $P$ as the client partition matrix, the coarsened Laplacian $L_C = P L_G P^\top$ induces a collaboration graph $\mathcal{G}_C$ over clients. Each entry $(L_C)_{ij}$ encodes the collaboration strength between $C_i$ and $C_j$: larger weights indicate stronger structural similarity, while smaller weights reflect dissimilarity.*

This formulation yields the collaboration graph in *one shot*, supported by coarsening theory, and avoids the heavy cost of iterative similarity estimation or repeated refinement common in existing pFL methods [42].

**Personalized Federated GNN.** With the collaboration graph $\mathcal{G}_C$ in place, we move beyond a single global model toward personalization. Purely global models ignore heterogeneity, while purely local models forgo collaboration. Personalized FL balances these extremes: each client maintains a model

**Table 1:** Node classification accuracy of baseline methods vs. *pFedGNN* on benchmark datasets.

| Method | Cora | CiteSeer | PubMed | Squirrel | RE | AR | CS |
|---|---|---|---|---|---|---|---|
| FedPub | 81.47 | 70.74 | 86.47 | 44.25 | 42.03 | 45.00 | 89.40 |
| FedAvg | 80.85 | 69.24 | 85.55 | 39.98 | 45.04 | 45.04 | 89.53 |
| FedGTA | 81.26 | 69.15 | 85.44 | 38.22 | 44.66 | 45.06 | 89.46 |
| FedDC | 83.83 | 75.17 | 81.81 | 35.62 | 27.00 | 38.81 | 90.25 |
| MOON | 82.18 | 71.09 | 85.64 | 40.44 | 44.99 | 44.54 | 89.91 |
| Scaffold | 75.58 | 75.26 | 70.14 | 35.34 | 20.21 | 36.81 | 83.10 |
| **pFedGNN** | **89.18** | **79.61** | **87.17** | **46.36** | **45.04** | **52.21** | **92.47** |



**(a)** Citeseer **(b)** Cora **(c)** Chameleon **(d)** Squirrel

**Figure 2:** Visualization of the inferred collaboration graphs. Each heatmap shows the interaction strength between pairs of clients, with stronger values indicating higher similarity and greater potential for collaboration. Results are shown here for 10 clients.

that integrates its own knowledge with that of structurally similar peers. Formally, let $\theta_i \in \mathbb{R}^p$ denote the parameters of client $C_i$. We define the update rule as a graph-regularized message-passing step:

$$\theta_i^{(t+1)} = \alpha\, \theta_i^{(t)} + (1-\alpha) \sum_{j \in \mathcal{N}(i)} w_{ij}\, \theta_j^{(t)},$$

where $\alpha \in [0,1]$ controls self-reliance vs. collaboration, $\mathcal{N}(i)$ is the neighborhood of client $i$ in $\mathcal{G}_C$, and $w_{ij} = \frac{(L_C)_{ij}}{\sum_{j'}(L_C)_{ij'}}$ are normalized edge weights. This update mirrors message passing in GNNs: each client retains a self-weighted local model while aggregating information from neighbors in $\mathcal{G}_C$. Personalization thus emerges naturally from the inferred collaboration structure, enabling tailored models that respect both heterogeneity and privacy.

## 4 Experiments

We evaluate *pFedGNN* on benchmark graph datasets against strong baselines. This section outlines the datasets, baselines, and implementation details, followed by results and analysis.

**Experimental Setup.** We evaluate our approach on 7 widely used benchmark datasets spanning three categories. (a) *Citation Networks:* Cora, Citeseer, Pubmed [44]; (b) *Co-author and Wiki-page Networks:* Coauthor-CS, [45], and Squirrel [46]; (c) *General Networks:* Roman-Empire (Article Syntax), Amazon-Ratings (Social) [47]. Dataset statistics are in Table 2. More details about training setting and baselines are included in Appendix A.2.

**Results.** Table 1 shows that, while preserving data privacy, our approach successfully infers both the global structure and client-level collaboration strengths, which are then leveraged in the pFedGNN framework. This leads to consistently strong node classification performance compared to baselines. Unlike existing methods that struggle with graph heterogeneity, pFedGNN achieves substantial accuracy gains across all datasets.

**Visualization of Collaboration Graph.** Figure 4 shows collaboration graphs for Cora, Citeseer, Chameleon and Squirrel. Heatmaps illustrate client–client interaction strengths, with darker cells denoting stronger collaboration. The inferred graphs adaptively capture heterogeneous client relationships, crucial for pFL as they guide collaboration by underlying data similarity rather than a single global model.

## 5 Conclusion

In federated graph-based machine learning, graphs within the same domain often exhibit non-IID properties. To tackle this challenge, we employ a personalized federated graph learning approach. A critical issue in personalized federated graph learning is finding the right balance between individual utilities and collaborative benefits. We address this with pFedGNN, where the collaboration graph guides each client to collaborate more with similar and beneficial peers, thereby enhancing local graph data homogeneity. Experiments across various graph networks show that pFedGNN outperforms baseline methods in GNN node classification, demonstrating its effectiveness in adaptive learning.

## References


[1] Fenxiao Chen, Yun-Cheng Wang, Bin Wang, and C-C Jay Kuo. Graph representation learning: a survey. *APSIPA Transactions on Signal and Information Processing*, 9:e15, 2020. **(Cited at p. 1.)**

[2] Xiaomin Fang, Lihang Liu, Jieqiong Lei, Donglong He, Shanzhuo Zhang, Jingbo Zhou, Fan Wang, Hua Wu, and Haifeng Wang. Geometry-enhanced molecular representation learning for property prediction. *Nature Machine Intelligence*, 4(2):127–134, 2022. **(Cited at p. 1.)**

[3] Michelle M Li, Kexin Huang, and Marinka Zitnik. Graph representation learning in biomedicine and healthcare. *Nature biomedical engineering*, 6(12):1353–1369, 2022. **(Cited at p. 1.)**

[4] Gabriele Corso, Hannes Stark, Stefanie Jegelka, Tommi Jaakkola, and Regina Barzilay. Graph neural networks. *Nature Reviews Methods Primers*, 4(1):17, 2024. **(Cited at p. 1.)**

[5] Huanding Zhang, Tao Shen, Fei Wu, Mingyang Yin, Hongxia Yang, and Chao Wu. Federated graph learning–a position paper. *arXiv preprint arXiv:2105.11099*, 2021. **(Cited at p. 1.)**

[6] Georgios A Kaissis, Marcus R Makowski, Daniel Rückert, and Rickmer F Braren. Secure, privacy-preserving and federated machine learning in medical imaging. *Nature Machine Intelligence*, 2(6):305–311, 2020. **(Cited at p. 1.)**

[7] Martijn Oldenhof, Gergely Ács, Balázs Pejó, Ansgar Schuffenhauer, Nicholas Holway, Noé Sturm, Arne Dieckmann, Oliver Fortmeier, Eric Boniface, Clément Mayer, et al. Industry-scale orchestrated federated learning for drug discovery. In *Proceedings of the aaai conference on artificial intelligence*, volume 37, pages 15576–15584, 2023. **(Cited at p. 1.)**

[8] R Vijay Anand, G Magesh, I Alagiri, Madala Guru Brahmam, Balamurugan Balusamy, Chithirai Pon Selvan, Haya Mesfer Alshahrani, Masresha Getahun, and Ben Othman Soufiene. Design of an improved model using federated learning and lstm autoencoders for secure and transparent blockchain network transactions. *Scientific Reports*, 15(1):1615, 2025. **(Cited at p. 1.)**

[9] Zixing Song, Xiangli Yang, Zenglin Xu, and Irwin King. Graph-based semi-supervised learning: A comprehensive review. *IEEE Transactions on Neural Networks and Learning Systems*, 34 (11):8174–8194, 2022. **(Cited at p. 1.)**

[10] Han Xie, Jing Ma, Li Xiong, and Carl Yang. Federated graph classification over non-iid graphs. *Advances in neural information processing systems*, 34:18839–18852, 2021. **(Cited at p. 1.)**

[11] Ke Zhang, Carl Yang, Xiaoxiao Li, Lichao Sun, and Siu Ming Yiu. Subgraph federated learning with missing neighbor generation. *Advances in neural information processing systems*, 34: 6671–6682, 2021. **(Cited at p. 1.)**

[12] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR, 2017. **(Cited at p. 1.)**

[13] Jakub Konečnỳ, H Brendan McMahan, Felix X Yu, Peter Richtárik, Ananda Theertha Suresh, and Dave Bacon. Federated learning: Strategies for improving communication efficiency. *arXiv preprint arXiv:1610.05492*, 2016. **(Cited at p. 1.)**

[14] Tian Li, Anit Kumar Sahu, Ameet Talwalkar, and Virginia Smith. Federated learning: Challenges, methods, and future directions. *IEEE signal processing magazine*, 37(3):50–60, 2020. **(Cited at p. 1.)**

[15] Avishek Ghosh, Jichan Chung, Dong Yin, and Kannan Ramchandran. An efficient framework for clustered federated learning. *Advances in neural information processing systems*, 33: 19586–19597, 2020. **(Cited at p. 2.)**

[16] Yishay Mansour, Mehryar Mohri, Jae Ro, and Ananda Theertha Suresh. Three approaches for personalization with applications to federated learning. *arXiv preprint arXiv:2002.10619*, 2020. **(Not cited.)**

[17] Nikita Kotelevskii, Maxime Vono, Alain Durmus, and Eric Moulines. Fedpop: A bayesian approach for personalised federated learning. *Advances in Neural Information Processing Systems*, 35:8687–8701, 2022. **(Cited at p. 2.)**

[18] Chenning Li, Xiao Zeng, Mi Zhang, and Zhichao Cao. Pyramidfl: A fine-grained client selection framework for efficient federated learning. In *Proceedings of the 28th annual international conference on mobile computing and networking*, pages 158–171, 2022. **(Cited at p. 2.)**

[19] Yongheng Deng, Feng Lyu, Ju Ren, Huaqing Wu, Yuezhi Zhou, Yaoxue Zhang, and Xuemin Shen. Auction: Automated and quality-aware client selection framework for efficient federated learning. *IEEE Transactions on Parallel and Distributed Systems*, 33(8):1996–2009, 2021. **(Not cited.)**

[20] Alysa Ziying Tan, Han Yu, Lizhen Cui, and Qiang Yang. Towards personalized federated learning. *IEEE transactions on neural networks and learning systems*, 34(12):9587–9603, 2022. **(Not cited.)**

[21] Youssef Allouah, Abdellah El Mrini, Rachid Guerraoui, Nirupam Gupta, and Rafael Pinot. Fine-tuning personalization in federated learning to mitigate adversarial clients. *Advances in Neural Information Processing Systems*, 37:100816–100844, 2024. **(Cited at p. 2.)**

[22] Yutao Huang, Lingyang Chu, Zirui Zhou, Lanjun Wang, Jiangchuan Liu, Jian Pei, and Yong Zhang. Personalized cross-silo federated learning on non-iid data. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pages 7865–7873, 2021. **(Cited at pp. 2 and 9.)**

[23] Liang Yao, Chengsheng Mao, and Yuan Luo. Graph convolutional networks for text classification. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 7370–7377, 2019. **(Cited at p. 2.)**

[24] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks, 2018. URL https://arxiv.org/abs/1710.10903. **(Not cited.)**

[25] Yuzhong Peng, Yanmei Lin, Xiao-Yuan Jing, Hao Zhang, Yiran Huang, and Guang Sheng Luo. Enhanced graph isomorphism network for molecular admet properties prediction. *Ieee Access*, 8:168344–168360, 2020. **(Not cited.)**

[26] Will Hamilton, Zhitao Ying, and Jure Leskovec. Inductive representation learning on large graphs. *Advances in neural information processing systems*, 30, 2017. **(Cited at p. 2.)**

[27] Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks, 2017. URL https://arxiv.org/abs/1609.02907. **(Cited at p. 2.)**

[28] Andreas Loukas. Graph reduction with spectral and cut guarantees. *Journal of Machine Learning Research*, 20(116):1–42, 2019. **(Cited at p. 3.)**

[29] Mohit Kataria, Aditi Khandelwal, Rocktim Das, Sandeep Kumar, and Jayadeva Jayadeva. Linear complexity framework for feature-aware graph coarsening via hashing. In *NeurIPS 2023 Workshop: New Frontiers in Graph Learning*, 2023. **(Cited at p. 3.)**

[30] Manoj Kumar, Anurag Sharma, Shashwat Saxena, and Sandeep Kumar. Featured graph coarsening with similarity guarantees. In *International Conference on Machine Learning*, pages 17953–17975. PMLR, 2023. **(Not cited.)**

[31] Mohit Kataria, Sandeep Kumar, et al. Ugc: Universal graph coarsening. *Advances in Neural Information Processing Systems*, 37:63057–63081, 2024. **(Cited at p. 3.)**

[32] H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agüera y Arcas. Communication-efficient learning of deep networks from decentralized data, 2023. URL https://arxiv.org/abs/1602.05629. **(Cited at pp. 3 and 8.)**

[33] Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. Federated optimization in heterogeneous networks, 2020. URL https://arxiv.org/abs/1812.06127. **(Cited at p. 8.)**

[34] Durmus Alp Emre Acar, Yue Zhao, Ramon Matas Navarro, Matthew Mattina, Paul N. Whatmough, and Venkatesh Saligrama. Federated learning based on dynamic regularization, 2021. URL https://arxiv.org/abs/2111.04263. **(Not cited.)**

[35] Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, Sashank J. Reddi, Sebastian U. Stich, and Ananda Theertha Suresh. Scaffold: Stochastic controlled averaging for federated learning, 2021. URL https://arxiv.org/abs/1910.06378. **(Cited at p. 3.)**

[36] Tian Li, Shengyuan Hu, Ahmad Beirami, and Virginia Smith. Ditto: Fair and robust federated learning through personalization, 2021. URL https://arxiv.org/abs/2012.04221. **(Cited at pp. 3 and 9.)**

[37] Canh T. Dinh, Nguyen H. Tran, and Tuan Dung Nguyen. Personalized federated learning with moreau envelopes, 2022. URL https://arxiv.org/abs/2006.08848. **(Cited at p. 3.)**

[38] Felix Sattler, Klaus-Robert Müller, and Wojciech Samek. Clustered federated learning: Model-agnostic distributed multi-task optimization under privacy constraints, 2019. URL https://arxiv.org/abs/1910.01991. **(Cited at pp. 3 and 9.)**

[39] Yutao Huang, Lingyang Chu, Zirui Zhou, Lanjun Wang, Jiangchuan Liu, Jian Pei, and Yong Zhang. Personalized cross-silo federated learning on non-iid data, 2021. URL https://arxiv.org/abs/2007.03797. **(Cited at p. 3.)**

[40] Manoj Ghuhan Arivazhagan, Vinay Aggarwal, Aaditya Kumar Singh, and Sunav Choudhary. Federated learning with personalization layers, 2019. URL https://arxiv.org/abs/1912.00818. **(Cited at p. 3.)**

[41] Aviv Shamsian, Aviv Navon, Ethan Fetaya, and Gal Chechik. Personalized federated learning using hypernetworks, 2021. URL https://arxiv.org/abs/2103.04628. **(Cited at p. 3.)**

[42] Rui Ye, Zhenyang Ni, Fangzhao Wu, Siheng Chen, and Yanfeng Wang. Personalized federated learning with inferred collaboration graphs. In *International Conference on Machine Learning*, pages 39801–39817. PMLR, 2023. **(Cited at p. 3.)**

[43] Nikita Malik, Nimesh Agrawal, and Sandeep Kumar. Ppda: A privacy preserving framework for distributed graph learning. In *International Conference on Neural Information Processing*, pages 308–331. Springer, 2024. **(Cited at p. 3.)**

[44] Deepak Bhaskar Acharya and Huaming Zhang. Feature selection and extraction for graph neural networks, 2019. URL https://arxiv.org/abs/1910.10682. **(Cited at p. 4.)**

[45] Oleksandr Shchur, Maximilian Mumme, Aleksandar Bojchevski, and Stephan Günnemann. Pitfalls of graph neural network evaluation. *ArXiv preprint*, 2018. **(Cited at p. 4.)**

[46] Hongbin Pei, Bingzhe Wei, Kevin Chen-Chuan Chang, Yu Lei, and Bo Yang. Geom-gcn: Geometric graph convolutional networks, 2020. URL https://arxiv.org/abs/2002.05287. **(Cited at p. 4.)**

[47] Oleg Platonov, Denis Kuznedelev, Michael Diskin, Artem Babenko, and Liudmila Prokhorenkova. A critical look at the evaluation of gnns under heterophily: Are we really making progress? *arXiv preprint arXiv:2302.11640*, 2023. **(Cited at p. 4.)**

[48] Jianyu Wang, Qinghua Liu, Hao Liang, Gauri Joshi, and H Vincent Poor. Tackling the objective inconsistency problem in heterogeneous federated optimization. *Advances in neural information processing systems*, 33:7611–7623, 2020. **(Cited at p. 8.)**

[49] Xunkai Li, Yinlin Zhu, Boyang Pang, Guochen Yan, Yeyu Yan, Zening Li, Zhengyu Wu, Wentao Zhang, Rong-Hua Li, and Guoren Wang. Openfgl: A comprehensive benchmark for federated graph learning, 2025. URL https://arxiv.org/abs/2408.16288. **(Cited at p. 8.)**

[50] Liam Collins, Hamed Hassani, Aryan Mokhtari, and Sanjay Shakkottai. Exploiting shared representations for personalized federated learning, 2023. URL https://arxiv.org/abs/2102.07078. **(Cited at p. 9.)**

[51] Hong-You Chen and Wei-Lun Chao. On bridging generic and personalized federated learning for image classification, 2022. URL https://arxiv.org/abs/2107.00778. **(Cited at p. 9.)**

[52] Othmane Marfoq, Giovanni Neglia, Laetitia Kameni, and Richard Vidal. Personalized federated learning through local memorization, 2022. URL https://arxiv.org/abs/2111.09360. **(Cited at p. 9.)**

[53] Vassilis Kalofolias. How to learn a graph from smooth signals, 2016. URL https://arxiv.org/abs/1601.02513. **(Cited at p. 9.)**

# A Appendix

## A.1 Dataset Statistics.

The data sets are summarized below in Table 2.

**Table 2:** Summary of graph datasets.

| Data | Nodes | Edges | Features | Class |
|---|---|---|---|---|
| Cora | 2,708 | 5,429 | 1,433 | 7 |
| CiteSeer (Cite.) | 3,327 | 4,732 | 3,703 | 6 |
| PubMed (Pub.) | 19,717 | 44,338 | 500 | 3 |
| Co-author CS (CS) | 18,333 | 81,894 | 6,805 | 15 |
| Chameleon (Cham.) | 2,277 | 36,101 | 2,325 | 5 |
| Squirrel (Sqi.) | 5,201 | 216,933 | 2,089 | 5 |
| Co-author Physics (Phy.) | 34,493 | 247,962 | 8,415 | 5 |
| Roman-empire (RE) | 22,662 | 32,927 | 300 | 18 |
| Amazon-ratings (AR) | 24,492 | 93,050 | 300 | 5 |

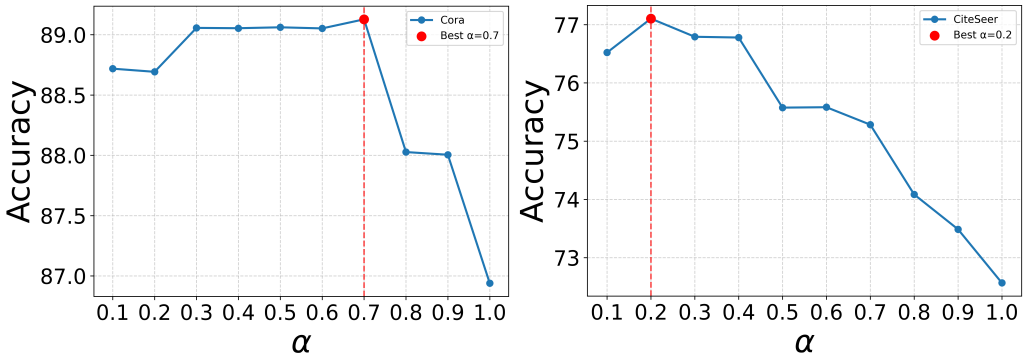
## A.2 Training Setting and Baselines

*Training Settings.* Experiments use 50 communication rounds, as personalized FL converges quickly, with each client running 200 iterations [48]. The local model is a 2-layer GNN with hidden dimensions [120, 84]. We train with SGD (lr=0.01). All runs are on a server with 72-core Intel(R) Xeon(R) Platinum 8360Y CPU, 1 TB RAM, and an NVIDIA RTX A6000 GPU (48 GB), under Ubuntu 20.04.1.

*Baselines.* For node classification comparisons (Table 1), we adopt baselines from the OpenFGL benchmark [49].

## A.3 Abalation Study

**Effect of Varying Personalization ($\alpha$).** The parameter $\alpha$ controls the trade-off between local learning and collaboration: lower $\alpha$ values encourage stronger inter-client cooperation, while higher values emphasize local client models. The results in Figure 3 show that moderate $\alpha$ values (e.g., $\alpha = 0.7$ for Cora, $\alpha = 0.2$ for CiteSeer) achieve the best balance between personalization and collaboration, confirming the adaptability of *pFedGNN* to varying heterogeneity levels. Excessive reliance on either extreme—purely local or fully global—reduces generalization, whereas a balanced trade-off enables models to adapt effectively to client-specific and shared knowledge simultaneously.




**(a)** Cora **(b)** Citeseer

**Figure 3:** Effect of the varying ($\alpha$) parameter on test accuracy for (a) Cora and (b) CiteSeer.

## A.4 Additional Experiments.

We conducted additional experiments, where we considered 11 other baseline methods, including local model training without collaboration (referred to as *Local*), as well as $FedAvg$ [32] and $FedProx$ [33], along with their fine-tuned (FT) versions. Additionally, we evaluate 6 representative

**Table 3:** This table presents a comparison of node classification performance across various existing methods on graph datasets. pFedGNN- demonstrates significant improvements on 3 out of the 5 datasets.

| Method | Citeseer | Cora | CS | DBLP | Physics | Pubmed |
|---|---|---|---|---|---|---|
| Local | 68.63 | 57.94 | 96.78 | 87.63 | 96.20 | 46.20 |
| FedAvg | 54.84 | 38.54 | 92.14 | 66.02 | 97.07 | 66.13 |
| FedAvgFT | 70.30 | 44.84 | 97.34 | 88.99 | 96.94 | 65.83 |
| FedProx | 55.00 | 38.39 | 92.00 | 62.15 | 97.13 | 64.07 |
| FedProxFT | 70.22 | 45.43 | 97.38 | 88.71 | 96.91 | 65.73 |
| FedAmp | 72.74 | 45.40 | 97.23 | 90.51 | 96.86 | 57.37 |
| Ditto | 64.30 | 46.10 | 96.95 | 88.22 | 96.85 | 67.47 |
| FedRep | 67.35 | 11.46 | 95.13 | 88.52 | 96.84 | 58.70 |
| Knn-per | 67.66 | 48.54 | 97.14 | 88.81 | 97.10 | 68.60 |
| FedRodPer | 71.46 | 17.13 | 95.38 | 87.60 | 96.74 | 58.20 |
| CFL | 71.14 | 36.59 | 98.90 | 89.06 | 96.94 | 65.83 |
| pFedGNN- | 92.93 | 92.84 | 95.96 | 92.04 | 96.60 | 85.42 |

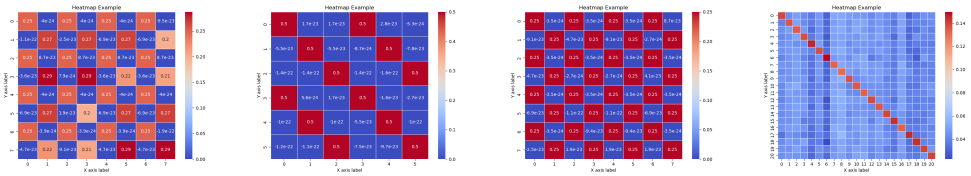
**Table 4:** This table illustrates the node classification accuracy (for pFedGNN-) when two different similarity measures are used to generate the collaboration graph.

| Method | Citeseer | Cora | DBLP | PubMed | Physics |
|---|---|---|---|---|---|
| cosine | 92.93 | 92.84 | 92.04 | 85.42 | 96.60 |
| inner product | 91.71 | 89.14 | 90.96 | 86.41 | 95.21 |

personalized federated learning (FL) methods. Among these, $CFL$ [38] employs a clustering-based approach, while $Ditto$ [36] is based on regularized optimization. $FedAMP$ [22] modifies the aggregation process, and $FedRep$ [50] focuses on sharing a backbone representation. $FedRoD$ [51] utilizes a multi-branch architecture, and $KNN-Per$ [52] employs feature memorization. The results for these are shown in Table 3, where we compared a non-private variant of pFedGNN called 'pFedGNN-' where we are learning the global structure of the graph between the clients without using PPDA and simply by using smooth signals [53].

Existing methods are inadequate in handling graph datasets effectively. pFedGNN- demonstrates substantial improvements in node classification accuracy across all these datasets. pFedGNN- showcases good performance across different datasets and against 11 baslines. For instance, on the Citeseer dataset, pFedGNN- improves accuracy from 20.19% (compared to best existing method) to 38.09% (compared to worst-performing method). Similarly, for Cora, the accuracy increases from 35.20% to 81.38%, for DBLP from 2.98% to 29.89%, and for PubMed from 16.82% to 39.22%.

While pFedGNN is not the top-performing method on the CS and Physics datasets, it remains competitive, with performance comparable to the best existing approaches. These results highlight pFedGNN's significant advantages in improving classification accuracy on challenging datasets, although its performance on certain datasets may still be matched by other methods.



**(a)** DBLP **(b)** Citeseer **(c)** CS **(d)** Cora

**Figure 4:** This figure illustrates the collaboration graphs, where we considered varying numbers of clients across different datasets: 8 for DBLP, 6 for Citeseer, 8 for CS, and 21 for Cora. Each cell in the heatmap represents the interaction strength between different clients.

Table 3 contains results for pFedGraph when we apply cosine similarity to optimize the collaboration graph. We compare two types of model similarity metrics, including cosine-based and inner-product-based. Results are reported in Table 4. From the table, we see that the proposed cosine-based optimization tends to perform the best.

### A.5 Visualization of Collaboration Graph

Here we visualize the collaboration graph of DBLP, Citeseer, CS Cora. Figure 4 illustrates the collaboration graph for these datasets in the form of heat-map. We have considered the various number of clients ranging from 6 to 20. Cosine is the measure of similarity while constructing the collaboration graph.