S^2FGL: Spatial Spectral Federated Graph Learning

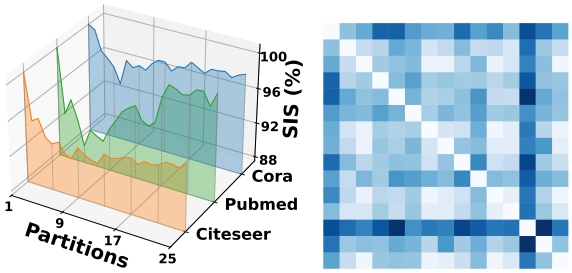
Zihan Tan [*1] **Suyuan Huang** [*1] **Guancheng Wan** [1] **Wenke Huang** [1] **He Li** [1] **Mang Ye** [1]

Abstract

Federated Graph Learning (FGL) combines the privacy-preserving capabilities of federated learning (FL) with the strong graph modeling capability of Graph Neural Networks (GNNs). Current research addresses subgraph-FL only from the structural perspective, neglecting the propagation of graph signals on spatial and spectral domains of the structure. From a spatial perspective, subgraph-FL introduces edge disconnections between clients, leading to disruptions in label signals and a degradation in the class knowledge of the global GNN. From a spectral perspective, spectral heterogeneity causes inconsistencies in signal frequencies across subgraphs, which makes local GNNs overfit the local signal propagation schemes. As a result, spectral client drifts occur, undermining global generalizability. To tackle the challenges, we propose a global knowledge repository to mitigate label signal disruption and a frequency alignment to address spectral client drifts. The combination of **S**patial and **S**pectral strategies forms our framework S^2FGL. Extensive experiments on multiple datasets demonstrate the superiority of S^2FGL. The code is available at https://github.com/Wonder7racer/S2FGL.git

1. Introduction

Graph Neural Networks (GNNs) have demonstrated remarkable efficacy in modeling graph-structured data (Wan et al., 2025a; Fang et al., 2025), thereby finding applications across various domains, such as social networks (Fan et al., 2020; Zhang et al., 2022b), epidemiology (Liu et al., 2024), and fraud detection (Wang et al., 2019; Tang et al., 2022). However, in real-world scenarios, graph data is often generated at the edge devices rather than in centralized systems (Zhang et al., 2021a). To address this, Federated Graph Learning (FGL) has emerged (Fu et al., 2022; Liu & Yu, 2022; Tan et al., 2025; 2024; Huang et al., 2022; Wan et al., 2024b; 2025b), leveraging the data privacy-preserving capabilities of Federated Learning (FL) (Huang et al., 2024; 2023b;c; 2022) to enable the efficient distributed training of GNNs (Huang et al., 2024). A significant use case of FGL is subgraph-FL, where each participant possesses a subgraph of the same overarching graph data.



(a) (b)

Figure 1. (a) According to existing FGL mainstream research, the range of 5 to 20 is the closest to real-world application scenarios. This range is encountering the toughest **label signal disruption** challenges. (b) The heat map of Kullback-Leibler divergence of eigenvalue distributions across clients. Inconsistency in graph signal frequency of subgraphs leads to **spectral client drifts**.

Although numerous FGL methods have attempted to provide solutions based purely on structure, including identifying structurally similar collaborators (Baek et al., 2023; Xie et al., 2021; Li et al., 2024), enhancing structural knowledge exchange (Tan et al., 2023; Huang et al., 2023a; Tan et al., 2025), and retrieving generic information under structural shifts (Wan et al., 2024a; Tan et al., 2024). These approaches overlooked the propagation of graph signals within the structure. Signal propagation can be analyzed from two perspectives: the spatial and the spectral domain. Specifically, the spatial domain governs the explicit transmission of signals among linked nodes, while the spectral domain characterizes signal diffusion across varying frequency spectra.

From the spatial perspective, due to edge loss, we hypothesize that nodes in subgraph-FL lose label signals from

[*]Equal contribution [1]National Engineering Research Center for Multimedia Software, School of Computer Science, Wuhan University, Wuhan, China. Correspondence to: Mang Ye <yemang@whu.edu.cn>.

Proceedings of the 42^{nd} International Conference on Machine Learning, Vancouver, Canada. PMLR 267, 2025.

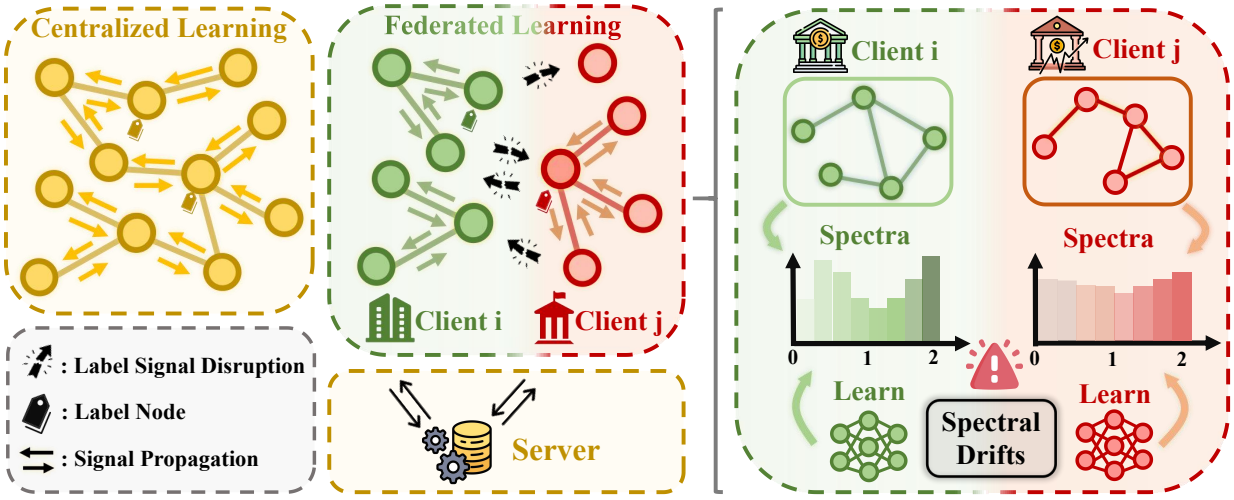


Figure 2. Illustration of the problems. (a) From the spatial perspective, nodes in subgraph-FL lose label signals from originally nearby labeled nodes due to edge loss, namely **label signal disruption**. Correspondingly, GNNs suffer from poor class knowledge, leading to a deteriorated global GNN. (b) From the spectral perspective, spectral heterogeneity induces inconsistencies in signal frequencies across subgraphs, leading to **spectral client drifts** in the signal propagation schemes of GNNs and degraded global generalizability.

originally nearby labeled nodes. This degradation hampers the ability of GNNs to learn comprehensive semantic knowledge, resulting in poor global performance and reduced generalizability. Correspondingly, we define this phenomenon as **Label Signal Disruption** (LSD), which naturally exists in subgraph-FL. For verifying, inspired by graph active learning research (Han et al., 2023), we investigate how the Structure Inertia Score (SIS) varies in subgraph-FL. Specifically, SIS evaluates the influence and significance of labels on graphs. In Fig. 1(a), we empirically demonstrate that the SIS initially decreases and then increases as the number of partitions increases. Correspondingly, existing methods fail to address LSD and suffer from poor class knowledge. Based on our analysis, we pose the question: **I)** *How can we address label signal disruption in subgraph-FL?*

From the spectral perspective, inconsistencies in signal frequencies across subgraphs of clients induce spectral client drifts in the signal transmission schemes of GNNs, thereby undermining the collaboration. To verify this phenomenon, we examine graph spectra across clients and demonstrate the heterogeneity in Fig. 1(b). It reveals inconsistent eigenvalue distribution across clients. As a result, GNNs learn distinct signal propagation schemes of subgraphs and optimize in different spectral directions, leading to **spectral client drifts** and degraded generalizability. Based on the above analysis, we pose the question: **II)** *How can we alleviate spectral client drifts under spectral heterogeneity?*

To address the issue of label signal disruption in Question **I)**, we propose Node Label Information Reinforcement (NLIR). Specifically, we leverage structurally representative nodes to construct a prototype-based global repository of class knowledge. During training, NLIR calculates the similarity distribution between all representative prototypes with node features, which provides multidimensional class localization of nodes. Furthermore, our strategy injects class knowledge into the local GNN, effectively mitigating the LSD issue.

Considering the spectral client drift posed by spectral heterogeneity in **II)**, we propose Frequency-aware Graph Modeling Alignment (FGMA). Our method utilizes the similarity relationship of the nodes to reconstruct a spectrum that incorporates GNNs adjacency awareness. FGMA then projects the high-frequency and low-frequency components of the features onto this spectrum. Subsequently, by aligning the local projections with the global one, we encourage the GNNs to learn a globally generic frequency processing scheme, thereby mitigating spectral client drift.

In conclusion, our key contributions are:

- We identify and empirically reveal the issue of label signal disruption in subgraph-FL. In addition, we reveal the spectral heterogeneity in subgraph scenarios.
- We design our framework S^2FGL including strategy Node Label Information Reinforcement and Frequency-aware Graph Modeling Alignment, effectively addressing the label signal disruption and the challenge of spectral client drifts in subgraph-FL.
- We conduct extensive experiments on various datasets, validating the superiority of our proposed S^2FGL.

2. Related Work

Federated Graph Learning. Federated graph learning leverages the powerful graph modeling capabilities of GNNs

along with the privacy-preserving attributes of federated learning, thus gaining increasing attention these days (He et al., 2021a; Fu et al., 2022; Liu & Yu, 2022; Wan et al., 2025b). Current FGL research can generally be categorized into two types: intra-graph FGL and inter-graph FGL. Intra-graph FGL research primarily focuses on subgraph-FL scenarios, where each client participates in the collaboration with a part of the whole graph (Zhang et al., 2021b). Correspondingly, the training targets include missing link prediction (Chen et al., 2021; Baek et al., 2023), node classification (Huang et al., 2023a; Li et al., 2024; Wan et al., 2024a; Zhu et al., 2024), and so on. On the other hand, clients in inter-graph FGL own independent local graph data, such as multiple graphs from different domains (Tan et al., 2023; Xie et al., 2021). In this paper, we focus on subgraph-FL scenarios of intra-graph FGL. We are the first to reveal and address the label signal disruption and spectral heterogeneity challenge among subgraphs, while existing methods inevitably fail due to the lack of targeted solutions.

Federated Learning. Federated learning (Huang et al., 2023c; 2024; Yang et al., 2023; Wan et al., 2024a) has gained increasing attention in recent years as it addresses the issue of data silos while ensuring data privacy. Several research directions have emerged from FL, including robustness (Xu et al., 2022; Hong et al., 2023; Zhu et al., 2023; Fang & Ye, 2022), fairness (Chen et al., 2024; Ezzeldin et al., 2023; Ray Chaudhury et al., 2022), and asynchronous federated learning (Xu et al., 2023; Zhang et al., 2023d). Generally, FL can be categorized into two main types by their optimization objective: traditional FL (tFL) and personalized FL (Hu et al., 2024; Shang et al., 2022; Lv et al., 2024; Smith et al., 2017). Research of tFL aims at aggregating a highly generalizable global model (McMahan et al., 2017; Li et al., 2020; Acar et al., 2021; Zhang et al., 2022a). For instance, FedNTD(Lee et al., 2022) preserves the global perspective on local data for the not-true classes, FEDGEN (Zhu et al., 2021) ensembles user information in a data-free manner to regulate local training, SCAFFOLD (Karimireddy et al., 2020) uses variance reduction for the client drift phenomenon. Instead, strategies of pFL aim to customize models that perform optimally for each client (Wu et al., 2023; Zhou & Konukoglu, 2023; Li et al., 2021; Zhang et al., 2023b). Specifically, FedALA (Zhang et al., 2023c) proposed adaptive masks to achieve personalized aggregation, DBE (Zhang et al., 2023a) stores domain biases for eliminating them, FedRoD (Chen & Chao, 2022) leverages two heads for global and personalized tasks.

Graph Spectrum Being related closely to graph connectivity, signal propagation, and structure, graph spectra have proven essential in performing various tasks on graph-structured data. For instance, it plays an essential role in anatomy detection, (Gao et al., 2023; Tang et al., 2022), graph condensation (Kreuzer et al., 2021; Liu et al., 2023), and graph contrastive learning (Bo et al., 2023a; Liu et al., 2022). Additionally, spectral GNNs (Wu et al., 2020) based on spectral filters are showing powerful ability in modeling graph data and attracting more attention. Specifically, existing research either (He et al., 2021b; Defferrard et al., 2016; He et al., 2022; Wang & Zhang, 2023) leverages various orthogonal polynomials to approximate arbitrary filters, or utilizes neural networks to parameterize the filters (Liao et al., 2019; Bo et al., 2023b). Although the potential of graph spectrum has been explored in various scenarios and tasks, the spectral domain in generalizable subgraph-FL has remained unexplored. Consequently, current methods suffer from optimization diverging on spectra and are trapped in suboptimal learning. Instead, our approach remarkably mitigates the challenge by targeted alignment on spectra.

Graph Signal Propagation: *Graph signal propagation describes how node signals diffuse on graph structures. In the* ***spatial*** *domain, propagation occurs through explicit signal passing along edges. In the* ***spectral*** *domain, propagation is characterized by how signals distribute across different frequency components.* ***Label Signal Disruption:*** *As subgraphs experience edge loss, nodes lose critical label signals containing class knowledge from their formerly adjacent labeled neighbors. Consequently, it limits the ability of GNNs to capture class distinctions accurately, leading to a degraded global model.* ***Spectral Client Drifts:*** *Inconsistencies in signal frequencies on graph spectra across subgraphs lead to diverging signal propagation schemes, causing spectral drift and degrading the generalizability of the global model.*

3. Problem Statement

Notation. Let the graph data be represented as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of nodes with $|\mathcal{V}| = N$ vertices, and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ denotes the set of edges connecting these nodes. The adjacency matrix is represented by $\mathbf{A} \in \mathbb{R}^{N \times N}$, where $\mathbf{A}_{uv} = 1$ indicates the presence of an edge $e_{uv} \in \mathcal{E}$, and $\mathbf{A}_{uv} = 0$ otherwise. The Laplacian matrix is given by $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where $\mathbf{D}$ is the degree matrix. The unitary matrix $\mathbf{U}$ is composed of the eigenvectors of $\mathbf{L}$. To distinguish between local and global properties, we introduce the following additional notation: the symbol i represents local properties or entities, whereas g denotes global properties or entities. Additional notation includes $\mathbf{h}$, which denotes the feature vector matrix of the graph $\mathcal{G}$, and $\mathbf{h}'$, representing the reconstructed feature matrix that emphasizes local structure. The cosine similarity matrix of $\mathbf{h}'$ is denoted by $\mathbf{S}'$ and the projections of the feature matrix onto low- and high-frequency eigenvectors are represented by $\mathbf{p}_l$ and $\mathbf{p}_h$, respectively.

Definition 3.1. Personalized PageRank (PPR): *The PPR*

matrix quantifies the influence each node has on every other node within the graph and is defined as:

$$P = \alpha \left(I - (1-\alpha)D^{-1}A\right)^{-1}. \tag{1}$$

Here, $\alpha \in (0,1)$ is the damping factor, typically set to 0.85, representing the probability of continuing the random walk. $\mathbf{I}$ is the identity matrix of size $N \times N$, $\mathbf{A}$ is the adjacency matrix of the graph, and $\mathbf{D}$ is the degree matrix with D_{ii} denoting the degree of node i.

Definition 3.2. Structure Inertia Score (SIS): *The SIS quantifies the cumulative influence of the training nodes on the entire graph and is defined as:*

$$SIS(P,t) = \sum_{i=1}^{n} \max_{j} \left(P_{i,j} \cdot t_j\right). \tag{2}$$

Here, P is the PPR matrix, and $t \in \{0,1\}^n$ is a binary vector indicating the training nodes, where $t_j = 1$ if node j is part of the training set, and $t_j = 0$ otherwise. The SIS aggregates the maximum personalized PageRank values from each node to any labeled node, effectively measuring the strongest influence each node in the graph receives from the training set. A higher SIS indicates greater structural inertia, suggesting that the labels have a significant influence over the network overall structure.

4. Methodology

4.1. Motivation

Signal propagation over graph structures fundamentally underpins the signal transmission paradigm in GNNs. Therefore, rather than focusing solely on challenges arising from static graph structures in FGL, as prior works have done, it is crucial to consider the dynamics of signal propagation in FGL. Specifically, the spatial domain governs the explicit transmission of signals between connected nodes, while the spectral domain captures signal diffusion across different frequency components. Accordingly, we empirically validate the presence of two major challenges: **label signal disruption** and **spectral client drifts**, from both the spatial and spectral perspectives.

Motivation of NLIR. Graph data is fragmented across multiple clients in FGL, which inevitably disrupts these label signals, especially as the number of partitions increases. Label signal disruption severs key pathways for propagating label information, resulting in incomplete and biased local feature representations that ultimately degrade GNNs ability semantically. For validation, we empirically investigate how the Structure Inertia Score (SIS) varies in subgraph-FL and experimentally demonstrate the relationship between the decrease in SIS and the client scale. Specifically, the SIS first decreases and then increases as the number of client partitions increases, highlighting the significant role of LSD in the observed performance decline. We aim to mitigate the impact of label signal disruption in FGL by preserving valuable label knowledge across fragmented data. Correspondingly, Node Label Information Reinforcement (NLIR) is introduced. By selecting nodes with both structural representativeness and rich label information for global repository and injecting it during local training, NLIR reduces the information loss inherent in subgraph partitioning. Specifically, throughout the training process, NLIR assesses the similarity distributions between node features and all representative prototypes, thereby enabling multidimensional localization of node classes. It effectively reconstructs label signal pathways within each client, enhancing local feature modeling by restoring context around labeled nodes.

Motivation of FGMA. We reveal **spectral client drifts** in the subgraph scenario, where GNNs across different clients capture high and low-frequency information inconsistently due to local adaptations. This inconsistency results from frequency-based signal propagation variations, which amplify biases during collaboration and compromise the generalizability of the global model. Spectral client drifts manifest when models optimized for distinct local frequency characteristics are combined, leading to a global model that is unable to generalize well across diverse frequency profiles. To address spectral client drifts, we propose the Frequency-aware Graph Modeling Alignment. Specifically, FGMA constructs the local graph spectra of the GNN adjacency awareness after calculating node similarity matrix. By projecting features separately onto high- and low-frequency components of the reconstructed spectra and aligh the local and global projections, FGMA promotes the generalizability of the global GNN graph signal propagation paradigm, thereby reducing frequency-based discrepancies duringcollaboration.

4.2. Node Label Information Reinforcement

First of all, we introduce the Structure-Aware Label Centrality (SALC) metric. The SALC, denoted as Λ_u^{SALC}, is defined as the combination of the Label Influence Centrality Λ_u^l and the Structural Prominence Score Λ_u^s:

$$\Lambda_u^{\text{SALC}} = \Lambda_u^s + \Lambda_u^l, \tag{3}$$

where Λ_u^s assesses the structural representativeness of node u, while Λ_u^l quantifies the influence propagation of labels.

$$\Lambda_u^s = \max\left(\tilde{P}_{u,v} \cdot \tau_v\right), \quad \Lambda_u^l = \sum_{v \in \mathcal{V}_L} \tilde{P}_{v,u}^{(L)}, \tag{4}$$

where $\tilde{P}_{u,v}$ is the (u,v)-th element of the standard Personalized PageRank (PPR) matrix $\tilde{\mathbf{P}}$, and τ_v represents the prior importance score of node v, typically initialized to 1 for all

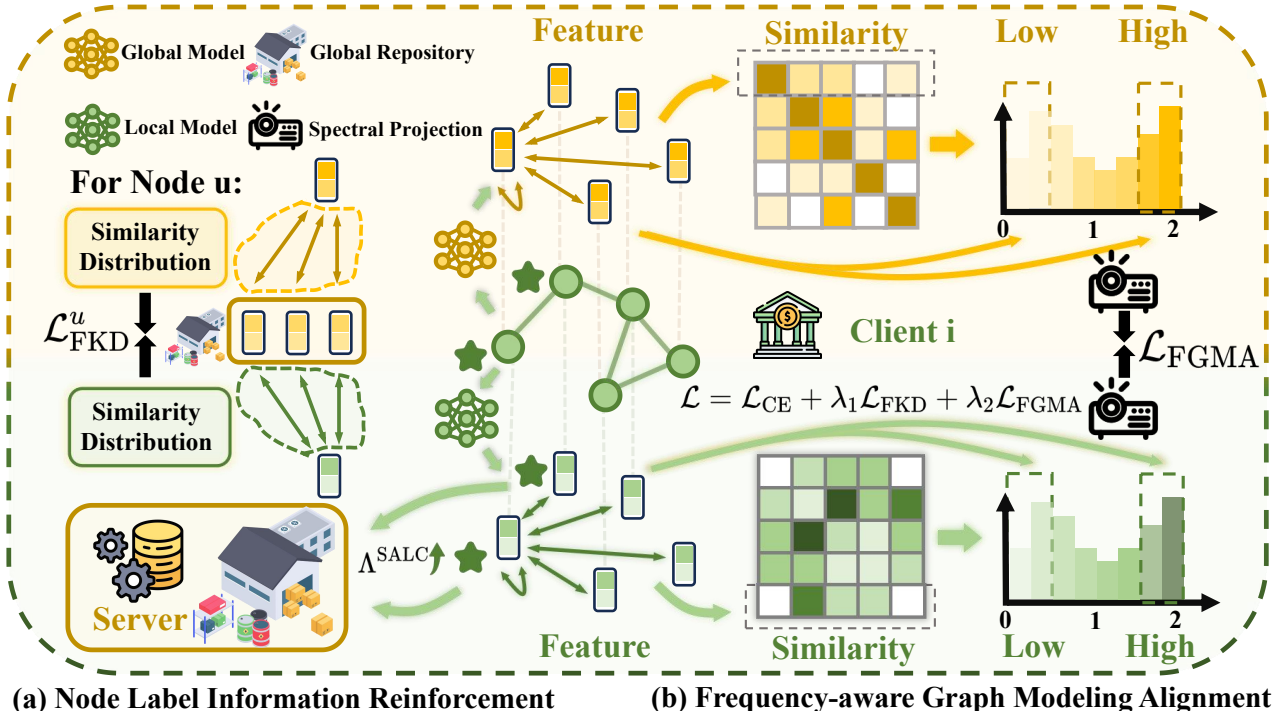


Figure 3. Framework illustration. (a) Node Label Information Reinforcement (NLIR) leverages a prototype-based structurally representative global repository of class knowledge. It provides **multidimensional class localization** of nodes through similarity distribution and allows L_{FKD} to **inject** the class knowledge during training. (b) Frequency-aware Graph Modeling Alignment (FGMA) **aligns** local high and low **spectral adjacency awareness** with the global GNN for a generic signal propagation scheme, mitigating spectral client drifts.

nodes. The Structural Prominence Score captures the maximum influence exerted by any node on node u, weighted by its importance.

As clarified, the PPR matrix used for computing Label Influence Centrality is denoted as $\tilde{\mathbf{P}}^{(L)}$, and defined as:

$$\tilde{\mathbf{P}}^{(L)} = \alpha \left(\mathbf{I} - (1-\alpha)\mathbf{D}^{-1}\mathbf{A}'\right)^{-1}.$$

Here, $\mathcal{V}_L$ denotes the set of labeled nodes, and $\tilde{P}^{(L)}_{v,u}$ represents the influence of node v on node u as captured by $\tilde{\mathbf{P}}^{(L)}$. To accurately capture the influence of labeled nodes, the inclusion of self-loops in $\mathbf{A}'$ ensures that each labeled node's own label contributes to its Λ^l_u. Compared to the intuitive approach of directly selecting labeled nodes, the SALC metric Λ^{SALC}_u considers both structural representativeness of the nodes and diffusion of label signals, thus avoiding biases caused by isolated labeled nodes. It is also capable of selecting unlabeled nodes that still possess rich label signals and structural advantages. This improves knowledge quality, enriches the repository, and mitigates the LSD problem. After computing the SALC scores for all nodes, we rank the nodes based on their Λ^{SALC}_u values and select the top K nodes, where the default value of K is $1/3$ of the total number of nodes. Subsequently, for each class c, the local prototype $\mathbf{H}^i_c$ at each client is computed as the mean feature vector of the selected nodes belonging to class c:

$$\mathbf{H}^i_c = \frac{1}{|\mathcal{V}^i_c|} \sum_{u \in \mathcal{V}^i_c} \mathbf{h}^i_u, \tag{5}$$

where $\mathcal{V}^i_c$ represents the set of nodes categorized as class c on client i, and $\mathbf{h}^i_u$ is the feature vector of node u on client i. Once the local prototypes are computed, clients upload their prototypes to the server along with the node count. For each class c, the server aggregates the prototypes from α percent of the clients by weighting each local prototype according to its sample size. Four global anchor prototypes are constructed for each class. Each global prototype $\mathbf{H}^{g,k}_c$ is computed as:

$$\mathbf{H}^{g,k}_c = \frac{1}{\sum_{i \in \mathcal{N}^k_c} |\mathcal{V}^i_c|} \sum_{i \in \mathcal{N}^k_c} |\mathcal{V}^i_c| \mathbf{H}^i_c, \quad \mathcal{H} = \begin{bmatrix} \mathbf{H}^{g,1}_1 \\ \mathbf{H}^{g,2}_1 \\ \vdots \\ \mathbf{H}^{g,4}_C \end{bmatrix}, \tag{6}$$

where $\mathbf{H}^{g,k}_c$ represents the k-th global prototype for class c, $\mathcal{N}^k_c$ is the set of clients randomly selected to contribute to the k-th global prototype for class c, and C is the number of classes. The global repository $\mathcal{H}$ contains all the global prototypes and will be broadcast back to clients. Once the

global knowledge repository is constructed, it is distributed to the local clients along with the model parameters. The global features $\mathbf{h}_u^g$ used in the following loss formulation refer to the frozen inference features extracted locally using the distributed global model. To guide local training, we introduce a federated knowledge distillation loss function designed to align local feature representations with the global prototypes:

$$\mathcal{L}_{\text{FKD}} = \frac{1}{|\mathcal{V}_i|} \sum_{u \in \mathcal{V}_i} \text{KL}\left(\sigma\left(\varphi(\mathbf{h}_u^i, \mathcal{H})\right), \sigma\left(\varphi(\mathbf{h}_u^g, \mathcal{H})\right)\right), \tag{7}$$

where $\text{KL}(\cdot,\cdot)$ represents the Kullback-Leibler divergence, and $\sigma(\cdot)$ is the softmax function applied to similarity scores computed by the similarity function $\varphi(\mathbf{h}, \mathcal{H})$, which returns a vector of cosine similarities between the feature $\mathbf{h}$ and all global prototypes.

4.3. Frequency-aware Graph Modeling Alignment

Reconstruct. As defined in Sec. 3, the feature vector matrix of graph $\mathcal{G}$ is denoted as $\mathbf{H}$. In order to emphasize the local structure and more accurately capture meaningful similarities between nodes, we use the original feature matrix $\mathbf{H}$. For each node u, we identify its k_{sim} most similar neighbors based on the cosine similarity of their feature vectors $\mathbf{h}_u$. We then construct a *sparse self-similarity matrix* $\mathbf{S}'$ as:

$$\mathbf{S'}_{u,v} = \begin{cases} \frac{\mathbf{h}_u \cdot \mathbf{h}_v}{\|\mathbf{h}_u\|_2 \|\mathbf{h}_v\|_2} & v \text{ among the top } k_{\text{sim}} \\ 0 & \text{otherwise} \end{cases}. \tag{8}$$

To characterize the topological structure between nodes, we calculate the graph Laplacian matrix $\mathbf{L}'$ based on this sparse similarity matrix $\mathbf{S}'$ as:

$$\mathbf{L}' = \mathbf{D}' - \mathbf{S}', \tag{9}$$

where $\mathbf{D}'$ is the diagonal degree matrix of $\mathbf{S}'$.

Projection. We perform eigendecomposition on the Laplacians $\mathbf{L'}^i$ and $\mathbf{L'}^g$. For $\mathbf{L'}^i$, let $\{\mathbf{u}_m^{\text{low},i}\}_{m=1}^{k_{\text{eig}}}$ be the eigenvectors corresponding to the smallest eigenvalues (low-frequency), and $\{\mathbf{u}_m^{\text{high},i}\}_{m=1}^{k_{\text{eig}}}$ for the largest eigenvalues (high-frequency). Similarly, for $\mathbf{L'}^g$, we obtain $\{\mathbf{u}_m^{\text{low},g}\}_{m=1}^{k_{\text{eig}}}$ and $\{\mathbf{u}_m^{\text{high},g}\}_{m=1}^{k_{\text{eig}}}$. The feature matrix $\mathbf{H}$ is then projected onto each of these eigenvectors. For instance:

$$\mathbf{Z}_m^{\text{low}} = (\mathbf{u}_m^{\text{low}} \mathbf{u}_m^{\text{low}T})\mathbf{H}, \quad \mathbf{Z}_m^{\text{high}} = (\mathbf{u}_m^{\text{high}} \mathbf{u}_m^{\text{high}T})\mathbf{H}. \tag{10}$$

Applying these projections for each $m \in \{1, \ldots, k_{\text{eig}}\}$, we obtain several sets of projected feature matrices used in the loss computation. Specifically, local features $\mathbf{H}^i$ are projected onto low/high-frequency eigenvectors of the local graph, yielding $\mathbf{Z}_m^{i,\text{low}}$ and $\mathbf{Z}_m^{i,\text{high}}$, respectively. Similarly, frozen global inference features $\mathbf{H}^g$ are projected onto corresponding eigenvectors from $\mathbf{L'}^g$, yielding $\mathbf{Z}_m^{g,\text{low}}$ and $\mathbf{Z}_m^{g,\text{high}}$.

Loss $\mathcal{L}_{\text{FGMA}}$ is then defined as the sum of MSE over all eigenvector-projected pairs:

$$\mathcal{L}_{\text{FGMA}} = \sum_{m=1}^{k_{\text{eig}}} \left(\text{MSE}(\mathbf{Z}_m^{i,\text{low}}, \mathbf{Z}_m^{g,\text{low}}) + \text{MSE}(\mathbf{Z}_m^{i,\text{high}}, \mathbf{Z}_m^{g,\text{high}})\right). \tag{11}$$

This loss addresses spectral heterogeneity by aligning client and global signal characteristics in both spectral domains.

Overall Objective. Combining strategies Node Label Information Reinforcementand Frequency-aware Graph Modeling Alignment, our framework S^2FGLreinforces label knowledge during local modeling and mitigates spectral heterogeneity. The final objective is:

$$\mathcal{L} = \mathcal{L}_{\text{CE}} + \lambda_1 \mathcal{L}_{\text{FKD}} + \lambda_2 \mathcal{L}_{\text{FGMA}}, \tag{12}$$

where $\mathcal{L}_{\text{CE}}$ denotes the standard cross-entropy loss for node classification, while λ_1 and λ_2 are balancing coefficients for the proposed methods.

5. Experiments

5.1. Experimental Setup

5.1.1. Datasets

- **Cora** (McCallum et al., 2000) dataset consists of 2708 scientific publications classified into one of seven classes. There are 5429 edges in the network of citations. 1433 distinct words make up the dictionary.
- **Citeseer** (Giles et al., 1998) dataset consists of 3312 scientific publications classified into one of six classes and 4732 edges. The dictionary contains 3703 unique words.
- **Pubmed** (Sen et al., 2008) dataset consists of 19717 scientific papers on diabetes that have been categorized into one of three categories in the PubMed database. The citation network has 44338 edges. A word vector from a dictionary with 500 unique terms that is TF/IDF weighted is used to describe each publication in the dataset.
- **Texas** and **Wisconsin** datasets are subsets of the WebKB dataset (Craven et al., 1998). The WebKB dataset was introduced in 1998, comprising web pages from the computer science departments of various universities, including the University of Texas and the University of Wisconsin. The dataset is commonly used for tasks such as webpage classification and link prediction, serving as a benchmark for evaluating machine learning models in graph-based learning scenarios.
- **Minesweeper** (Baranovskiy et al., 2023) dataset is a synthetic graph dataset inspired by the Minesweeper game. In this dataset, the graph is structured as a regular 100x100 grid, where each node represents a cell connected to its neighboring nodes, except for edge nodes, which have

Table 1. Comparison with the state-of-the-art methods on homophilic and heterophilic graph datasets. We report node classification accuracies and each value difference from FedAvg with the best result in bold.

Methods	Cora	CiteSeer	PubMed	Texas	Wisconsin	Minesweeper
FedAvg [ASTAT17]	81.9 ± 0.7	74.3 ± 0.4	87.3 ± 0.3	72.8 ± 0.1	77.6 ± 0.2	79.6 ± 0.1
FedProx [arXiv18]	82.1 ± 0.5 $_{\uparrow 0.2}$	74.4 ± 0.3 $_{\uparrow 0.1}$	87.9 ± 0.4 $_{\uparrow 0.6}$	73.5 ± 3.7 $_{\uparrow 0.7}$	77.3 ± 3.4 $_{\downarrow 0.3}$	79.7 ± 0.1 $_{\uparrow 0.1}$
FedNova [NeurIPS20]	81.6 ± 1.2 $_{\downarrow 0.3}$	74.4 ± 0.4 $_{\uparrow 0.1}$	88.2 ± 0.5 $_{\uparrow 0.9}$	73.0 ± 4.4 $_{\uparrow 0.2}$	77.4 ± 4.2 $_{\downarrow 0.2}$	79.9 ± 0.4 $_{\uparrow 0.3}$
FedFa [ICLR23]	82.7 ± 0.5 $_{\uparrow 0.8}$	74.9 ± 0.6 $_{\uparrow 0.6}$	87.8 ± 0.5 $_{\uparrow 0.5}$	73.9 ± 3.6 $_{\uparrow 1.1}$	78.1 ± 4.6 $_{\uparrow 0.5}$	80.1 ± 0.3 $_{\uparrow 0.5}$
FedSage+ [NeurIPS19]	82.3 ± 0.7 $_{\uparrow 0.4}$	75.2 ± 0.3 $_{\uparrow 0.9}$	88.2 ± 0.7 $_{\uparrow 0.9}$	73.7 ± 4.0 $_{\uparrow 0.9}$	$\mathbf{79.0 \pm 3.3}$ $_{\uparrow 1.4}$	79.9 ± 0.2 $_{\uparrow 0.3}$
FedStar [AAAI23]	82.6 ± 0.5 $_{\uparrow 0.7}$	74.5 ± 0.3 $_{\uparrow 0.2}$	88.1 ± 0.6 $_{\uparrow 0.8}$	74.3 ± 2.7 $_{\uparrow 1.5}$	78.3 ± 4.7 $_{\uparrow 0.7}$	79.8 ± 0.1 $_{\uparrow 0.2}$
FedPub [ICML23]	82.3 ± 0.8 $_{\uparrow 0.4}$	74.8 ± 0.7 $_{\uparrow 0.5}$	88.0 ± 0.4 $_{\uparrow 0.7}$	73.4 ± 3.5 $_{\uparrow 0.6}$	77.8 ± 3.1 $_{\uparrow 0.2}$	79.9 ± 0.2 $_{\uparrow 0.3}$
FGSSL [IJCAI23]	82.6 ± 0.4 $_{\uparrow 0.7}$	74.9 ± 0.2 $_{\uparrow 0.6}$	87.6 ± 0.7 $_{\uparrow 0.3}$	73.6 ± 4.6 $_{\uparrow 0.8}$	77.8 ± 3.8 $_{\uparrow 0.2}$	79.9 ± 0.2 $_{\uparrow 0.3}$
FedGTA [VLDB24]	82.4 ± 0.8 $_{\uparrow 0.5}$	75.1 ± 0.5 $_{\uparrow 0.8}$	87.7 ± 0.9 $_{\uparrow 0.4}$	72.6 ± 4.2 $_{\downarrow 0.2}$	77.8 ± 4.1 $_{\uparrow 0.2}$	80.2 ± 0.3 $_{\uparrow 0.6}$
FGGP [AAAI24]	82.5 ± 0.4 $_{\uparrow 0.6}$	74.7 ± 0.5 $_{\uparrow 0.4}$	87.5 ± 0.4 $_{\uparrow 0.2}$	73.6 ± 2.8 $_{\uparrow 0.8}$	78.2 ± 3.4 $_{\uparrow 0.6}$	80.4 ± 0.3 $_{\uparrow 0.8}$
S^2FGL (ours)	$\mathbf{83.4 \pm 0.5}$ $_{\uparrow 1.5}$	$\mathbf{76.0 \pm 0.3}$ $_{\uparrow 1.7}$	$\mathbf{88.6 \pm 0.1}$ $_{\uparrow 1.3}$	$\mathbf{74.7 \pm 0.1}$ $_{\uparrow 1.9}$	78.4 ± 0.1 $_{\uparrow 0.8}$	$\mathbf{80.5 \pm 0.3}$ $_{\uparrow 0.9}$

fewer neighbors. The primary task is to predict which nodes contain mines. This dataset is commonly used to evaluate the performance of GNNs under heterophily.

5.1.2. Baselines

We compare S^2FGL with several state-of-the-art approaches, including traditional federated learning methods such as **FedAvg** (McMahan et al., 2017), **FedProx** (Li et al., 2020), and **FedNova** (Wang et al., 2020); personalized federated learning algorithms like **FedFa** (Zhou & Konukoglu, 2023); federated graph learning approaches including **FGSSL** (Huang et al., 2023a) and **FGGP** (Wan et al., 2024a); as well as personalized federated graph learning methods such as **FedSage+** (Zhang et al., 2021b), **FedStar** (Tan et al., 2023), **FedPub** (Baek et al., 2023), and **FedGTA** (Li et al., 2024). This comprehensive set of baseline methods spans various federated and graph learning paradigms, allowing us to evaluate the performance and advantages of our proposed S^2FGL across diverse scenarios.

5.1.3. Implement Details

Following prevalent methodologies in FGL research, we employ the Louvain community detection algorithm to partition the graph into subgraphs assigned to different clients. For each dataset, we divide the nodes into training, validation, and testing sets with ratios of 60%, 20% and 20% respectively. Additionally, we simulate various collaborative scenarios by configuring the number of clients to 10 for Cora, Citeseer, Pubmed, and Minesweeper datasets, and 3 for Texas and Wisconsin datasets. The primary evaluation metric is the node classification accuracy on the clients' test sets. We conduct each experiment five times and report the average accuracy from the last five communication epochs as the final performance. We conduct experiments with the ACM-GCN (Luan et al., 2022), which achieves a strong ability on both homophilic and heterophilic graph datasets.

5.2. Experiment Results

In this section, we comprehensively evaluate the proposed S^2FGL by addressing the following questions:

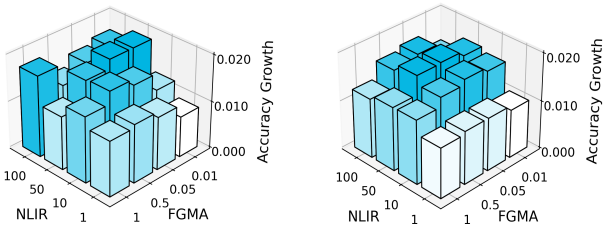
- **Q1: How does S^2FGL perform compared to existing methods in subgraph-FL?**
- **Q2: What is the impact of each component of S^2FGL on its performance?**
- **Q3: Does S^2FGL exhibit good stability?**
- **Q4: Does the effectiveness of NLIR help alleviate the impact of SIS decline in subgraph-FL?**

Q1:How does S^2FGL perform compared to existing methods in subgraph-FL?

We present the results of node classification tasks across various FGL scenarios using multiple graph datasets, and we summarize the final average test accuracy in Tab. 1. Our proposed method, S^2FGL, demonstrates superior performance by outperforming all other baseline approaches in five out of the six datasets. This consistent superiority highlights the effectiveness of S^2FGL in handling diverse graph structures and data distributions inherent to different datasets.

Q2:What is the impact of each component of S^2FGL on its performance?

To evaluate the individual contributions of the proposed **NLIR** and **FGMA** strategies within the **S^2FGL** framework, we conducted ablation experiments on the **Cora** and **Citeseer** datasets. In this study, we systematically removed each



(a) Cora (b) Citeseer

Figure 4. Analysis of the performance growth between S^2FGL and FedAvg under varying scaling factors of NLIR and FGMA

component to assess its impact on model performance. The results of the ablation study are presented in Tab. 2 which demonstrate that both **NLIR** and **FGMA** independently contribute to the overall performance.

NLIR	FGMA	Dataset	
		Cora	Citeseer
✗	✗	81.9	74.3
✓	✗	83.2	75.6
✗	✓	82.6	75.0
✓	✓	**83.4**	**76.0**

Table 2. Ablation study of key components (NLIR, FGMA) of S^2FGL on Cora and Citeseer datasets.

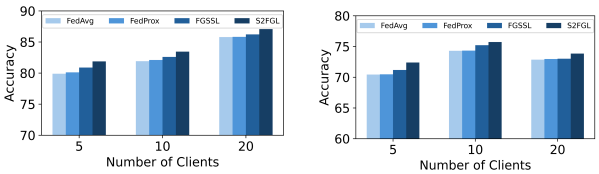
Q3: Does S^2FGL exhibit good stability?

We evaluated the stability and adaptability of our approach on the Cora and Citeseer datasets through separate hyperparameter tuning and client partitioning experiments. Specifically, we varied the scaling factors of NLIR and FGMA to investigate how different parameter settings influence the model performance and robustness. Additionally, we conducted experiments with varying client scales

Varying Scaling Factors of NLIR and FGMA. For the NLIR method, we test scaling factors set to 100, 50, 10, and 1. For the FGMA method, the settings were 0.01, 0.05, 0.5, and 1. The results, shown in Fig. 4, indicate that our method maintains consistent performance across these varying hyperparameter configurations.

Varying Client Scales. We assessed performance with different numbers of client partitions: 5, 10 and 20 clients. We compared S^2FGL with other methods, including **FedAvg** , **FedProx** , and **FGSSL**. The outcomes, presented in Fig. 5, demonstrate that S^2FGL consistently delivers reliable results regardless of the number of partitions.

Overall, S^2FGL exhibits strong stability and adaptability across varying hyperparameter settings and client partition configurations on both the Cora and Citeseer datasets. These findings confirm that S^2FGL not only sustains its effectiveness under diverse conditions but also adapts seamlessly to varying client scales, demonstrating its suitability for real-world subgraph-FL scenarios.



(a) Cora (b) Citeseer

Figure 5. Analysis of performance under different client numbers.

Q4: Does the effectiveness of NLIR help alleviate the impact of SIS decline in subgraph-FL? Through experiments, we observe that as the number of clients increases, the Subgraph Information Similarity (SIS) consistently decreases, indicating greater structural divergence across local graphs. Notably, the performance gains brought by NLIR become more pronounced under these conditions. This trend, illustrated in Fig. 6, highlights the targeted effectiveness of NLIR in mitigating the adverse effects of reduced SIS in subgraph-FL settings.

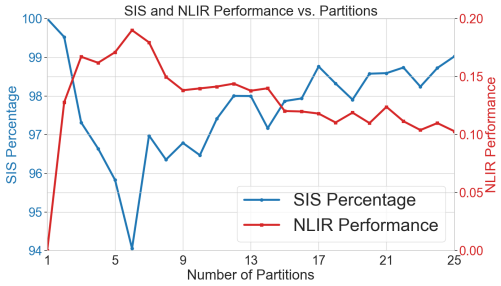


Figure 6. Performance improvement of global GNN using the NLIR method in relation to the variation of the Structure Inertia Score (SIS) across different numbers of client partitions.

6. Conclusion

In this paper, we identify two critical challenges in subgraph-FL from the perspective of graph signal propagation: label signal disruption and spectral client drifts. To overcome these obstacles, we propose two key strategies: NLIR and FGMA. NLIR constructs a global repository of structurally representative nodes with rich label influence to deal with the LSD issue, while FGMA promotes a generic signal propagation paradigm, mitigating spectral client drift. By integrating these strategies, S^2FGL effectively addresses both the spatial and spectral challenges in FGL. Extensive experiments on various datasets validate that S^2FGL significantly boosts global generalizability.

Acknowledgement

This work is supported by the National Key Research and Development Program of China (2024YFC3308400), and National Natural Science Foundation of China under Grant (62361166629, 62176188, 62225113, 623B2080), the Wuhan University Undergraduate Innovation Research Fund Project. The supercomputing system at the Supercomputing Center of Wuhan University supported the numerical calculations in this paper.

Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none of which we feel must be specifically highlighted here.

References

Acar, D. A. E., Zhao, Y., Matas, R., Mattina, M., Whatmough, P., and Saligrama, V. Federated learning based on dynamic regularization. In *ICLR*, 2021.

Baek, J., Jeong, W., Jin, J., Yoon, J., and Hwang, S. J. Personalized subgraph federated learning. In *ICML*, pp. 1396–1415, 2023.

Baranovskiy, D., Oseledets, I., and Babenko, A. A critical look at the evaluation of gnns under heterophily: Are we really making progress? *arXiv preprint arXiv:2302.11640*, 2023.

Bo, D., Fang, Y., Liu, Y., and Shi, C. Graph contrastive learning with stable and scalable spectral encoding. *NeurIPS*, 36:45516–45532, 2023a.

Bo, D., Shi, C., Wang, L., and Liao, R. Specformer: Spectral graph neural networks meet transformers. *arXiv preprint arXiv:2303.01028*, 2023b.

Chen, H.-Y. and Chao, W.-L. On bridging generic and personalized federated learning for image classification. In *ICLR*, 2022.

Chen, M., Zhang, W., Yuan, Z., Jia, Y., and Chen, H. Fede: Embedding knowledge graphs in federated setting. In *IJCKG*, pp. 80–88, 2021.

Chen, Y., Huang, W., and Ye, M. Fair federated learning under domain skew with local consistency and domain diversity. In *CVPR*, pp. 12077–12086, 2024.

Craven, M., DiPasquo, D., Freitag, D., McCallum, A., Mitchell, T., Nigam, K., and Slattery, S. Learning to extract symbolic knowledge from the world wide web. *AAAI/IAAI*, pp. 2, 1998.

Defferrard, M., Bresson, X., and Vandergheynst, P. Convolutional neural networks on graphs with fast localized spectral filtering. In *NeurIPS*, 2016.

Ezzeldin, Y. H., Yan, S., He, C., Ferrara, E., and Avestimehr, A. S. Fairfed: Enabling group fairness in federated learning. In *AAAI*, 2023.

Fan, W., Ma, Y., Li, Q., Wang, J., Cai, G., Tang, J., and Yin, D. A graph neural network framework for social recommendations. *TKDE*, 2020.

Fang, X. and Ye, M. Robust federated learning with noisy and heterogeneous clients. In *CVPR*, pp. 10072–10081, 2022.

Fang, X., Easwaran, A., Genest, B., and Suganthan, P. N. Adaptive hierarchical graph cut for multi-granularity out-of-distribution detection. *IEEE TAI*, 2025.

Fu, X., Zhang, B., Dong, Y., Chen, C., and Li, J. Federated graph machine learning: A survey of concepts, techniques, and applications. *arXiv preprint arXiv:2207.11812*, 2022.

Gao, Y., Wang, X., He, X., Liu, Z., Feng, H., and Zhang, Y. Addressing heterophily in graph anomaly detection: A perspective of graph spectrum. In *Proceedings of the ACM Web Conference 2023*, pp. 1528–1538, 2023.

Giles, C. L., Bollacker, K. D., and Lawrence, S. Citeseer: An automatic citation indexing system. In *Proceedings of the third ACM conference on Digital libraries*, pp. 89–98, 1998.

Han, H., Liu, X., Ma, L., Torkamani, M., Liu, H., Tang, J., and Yamada, M. Structural fairness-aware active learning for graph neural networks. In *ICLR*, 2023.

He, C., Balasubramanian, K., Ceyani, E., Yang, C., Xie, H., Sun, L., He, L., Yang, L., Yu, P. S., Rong, Y., et al. Fedgraphnn: A federated learning system and benchmark for graph neural networks. In *ICLR*, 2021a.

He, M., Wei, Z., Xu, H., et al. Bernnet: Learning arbitrary graph spectral filters via bernstein approximation. In *NeurIPS*, pp. 14239–14251, 2021b.

He, M., Wei, Z., and Wen, J.-R. Convolutional neural networks on graphs with chebyshev approximation, revisited. In *NeurIPS*, pp. 7264–7276, 2022.

Hong, J., Wang, H., Wang, Z., and Zhou, J. Federated robustness propagation: sharing adversarial robustness in heterogeneous federated learning. In *AAAI*, pp. 7893–7901, 2023.

Hu, M., Yue, Z., Xie, X., Chen, C., Huang, Y., Wei, X., Lian, X., Liu, Y., and Chen, M. Is aggregation the only choice? federated learning via layer-wise model recombination. In *SIGKDD*, pp. 1096–1107, 2024.

Huang, W., Ye, M., and Du, B. Learn from others and be yourself in heterogeneous federated learning. In *CVPR*, pp. 10143–10153, 2022.

Huang, W., Wan, G., Ye, M., and Du, B. Federated graph semantic and structural learning. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, pp. 3830–3838, 2023a.

Huang, W., Ye, M., Shi, Z., and Du, B. Generalizable heterogeneous federated cross-correlation and instance similarity learning. *TPAMI*, pp. 712–728, 2023b.

Huang, W., Ye, M., Shi, Z., Li, H., and Du, B. Rethinking federated learning with domain shift: A prototype view. In *CVPR*, pp. 16312–16322. IEEE, 2023c.

Huang, W., Ye, M., Shi, Z., Wan, G., Li, H., Du, B., and Yang, Q. A federated learning for generalization, robustness, fairness: A survey and benchmark. *TPAMI*, 2024.

Karimireddy, S. P., Kale, S., Mohri, M., Reddi, S. J., Stich, S. U., and Suresh, A. T. Scaffold: Stochastic controlled averaging for on-device federated learning. In *ICML*, pp. 5132–5143, 2020.

Kreuzer, D., Beaini, D., Hamilton, W., Létourneau, V., and Tossou, P. Rethinking graph transformers with spectral attention. In *NeruIPS*, pp. 21618–21629, 2021.

Lee, G., Jeong, M., Shin, Y., Bae, S., and Yun, S.-Y. Preservation of the global knowledge by not-true distillation in federated learning. In Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., and Oh, A. (eds.), *NeurIPS*, pp. 38461–38474, 2022.

Li, T., Sahu, A. K., Zaheer, M., Sanjabi, M., Talwalkar, A., and Smith, V. Federated optimization in heterogeneous networks. *MLSys*, 2:429–450, 2020.

Li, X., Wu, Z., Zhang, W., Zhu, Y., Li, R.-H., and Wang, G. Fedgta: Topology-aware averaging for federated graph learning. *arXiv preprint arXiv:2401.11755*, 2024.

Li, X.-C., Zhan, D.-C., Shao, Y., Li, B., and Song, S. Fedphp: Federated personalization with inherited private models. In *ECML*, pp. 587–602, 2021.

Liao, R., Zhao, Z., Urtasun, R., and Zemel, R. S. Lanczosnet: Multi-scale deep graph convolutional networks. In *ICLR*, 2019.

Liu, N., Wang, X., Bo, D., Shi, C., and Pei, J. Revisiting graph contrastive learning from the perspective of graph spectrum. In *NeurIPS*, pp. 2972–2983, 2022.

Liu, R. and Yu, H. Federated graph neural networks: Overview, techniques and challenges. *arXiv preprint arXiv:2202.07256*, 2022.

Liu, Y., Bo, D., and Shi, C. Graph condensation via eigenbasis matching. *arXiv preprint arXiv:2310.09202*, 2023.

Liu, Z., Wan, G., Prakash, B. A., Lau, M. S., and Jin, W. A review of graph neural networks in epidemic modeling. *arXiv preprint arXiv:2403.19852*, 2024.

Luan, S., Hua, C., Lu, Q., Zhu, J., Zhao, M., Zhang, S., Chang, X.-W., and Precup, D. Revisiting heterophily for graph neural networks. In *NeurIPS22*, pp. 1362–1375, 2022.

Lv, F., Shang, X., Zhou, Y., Zhang, Y., Li, M., and Lu, Y. Personalized federated learning on heterogeneous and long-tailed data via expert collaborative learning. *arXiv preprint arXiv:2408.02019*, 2024.

McCallum, A. K., Nigam, K., Rennie, J., and Seymore, K. Automating the construction of internet portals with machine learning. *Information Retrieval*, 3(2):127–163, 2000.

McMahan, B., Moore, E., Ramage, D., Hampson, S., and y Arcas, B. A. Communication-efficient learning of deep networks from decentralized data. In *AISTATS*, pp. 1273–1282, 2017.

Ray Chaudhury, B., Li, L., Kang, M., Li, B., and Mehta, R. Fairness in federated learning via core-stability. In *NeurIPS*, pp. 5738–5750, 2022.

Sen, P., Namata, G., Bilgic, M., Getoor, L., Galligher, B., and Eliassi-Rad, T. Collective classification in network data. *AI magazine*, 29(3):93–93, 2008.

Shang, X., Lu, Y., Huang, G., and Wang, H. Federated learning on heterogeneous and long-tailed data via classifier re-training with federated features. In *IJCAI*, 2022.

Smith, V., Chiang, C.-K., Sanjabi, M., and Talwalkar, A. S. Federated multi-task learning. In *NeurIPS*, 2017.

Tan, Y., Liu, Y., Long, G., Jiang, J., Lu, Q., and Zhang, C. Federated learning on non-iid graphs via structural knowledge sharing. In *AAAI*, pp. 9953–9961, 2023.

Tan, Z., Wan, G., Huang, W., and Ye, M. Fedssp: Federated graph learning with spectral knowledge and personalized preference. In *NeurIPS*, 2024.

Tan, Z., Wan, G., Huang, W., Li, H., Zhang, G., Yang, C., and Ye, M. Fedspa: Generalizable federated graph learning under homophily heterogeneity. In *CVPR*, 2025.

Tang, J., Li, J., Gao, Z., and Li, J. Rethinking graph neural networks for anomaly detection. In *ICML*, pp. 21076–21089, 2022.

Wan, G., Huang, W., and Ye, M. Federated graph learning under domain shift with generalizable prototypes. In *AAAI*, pp. 15429–15437, 2024a.

Wan, G., Tian, Y., Huang, W., Chawla, N. V., and Ye, M. S3gcl: Spectral, swift, spatial graph contrastive learning. In *ICML*, pp. 49973–49990, 2024b.

Wan, G., Huang, Z., Zhao, W., Luo, X., Sun, Y., and Wang, W. Rethink graphode generalization within coupled dynamical system. In *ICML*, 2025a.

Wan, G., Shi, Z., Huang, W., Zhang, G., Tao, D., and Ye, M. Energy-based backdoor defense against federated graph learning. In *ICLR*, 2025b.

Wang, D., Lin, J., Cui, P., Jia, Q., Wang, Z., Fang, Y., Yu, Q., Zhou, J., Yang, S., and Qi, Y. A semi-supervised graph attentive network for financial fraud detection. In *ICDM*, pp. 598–607, 2019.

Wang, J., Liu, Q., Liang, H., Joshi, G., and Poor, H. V. Tackling the objective inconsistency problem in heterogeneous federated optimization. In *NeurIPS*, pp. 7611–7623, 2020.

Wang, X. and Zhang, M. How powerful are spectral graph neural networks. In *ICML*, pp. 23341–23362, 2023.

Wu, X., Liu, X., Niu, J., Zhu, G., and Tang, S. Bold but cautious: Unlocking the potential of personalized federated learning through cautiously aggressive collaboration. In *ICCV*, pp. 19375–19384, 2023.

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., and Philip, S. Y. A comprehensive survey on graph neural networks. *TNNLS*, pp. 4–24, 2020.

Xie, H., Ma, J., Xiong, L., and Yang, C. Federated graph classification over non-iid graphs. In *NeruIPS*, pp. 18839–18852, 2021.

Xu, C., Qu, Y., Xiang, Y., and Gao, L. Asynchronous federated learning on heterogeneous devices: A survey. *Computer Science Review*, 50:100595, 2023.

Xu, J., Chen, Z., Quek, T. Q., and Chong, K. F. E. Fedcorr: Multi-stage federated learning for label noise correction. In *CVPR*, pp. 10184–10193, 2022.

Yang, X., Huang, W., and Ye, M. Dynamic personalized federated learning with adaptive differential privacy. *Advances in Neural Information Processing Systems*, 36: 72181–72192, 2023.

Zhang, H., Shen, T., Wu, F., Yin, M., Yang, H., and Wu, C. Federated graph learning – a position paper. In *arXiv preprint arXiv:2105.11099*, 2021a.

Zhang, J., Li, Z., Li, B., Xu, J., Wu, S., Ding, S., and Wu, C. Federated learning with label distribution skew via logits calibration. In *ICML*, pp. 26311–26329, 2022a.

Zhang, J., Hua, Y., Cao, J., Wang, H., Song, T., XUE, Z., Ma, R., and Guan, H. Eliminating domain bias for federated learning in representation space. In *NeurIPS*, pp. 14204–14227, 2023a.

Zhang, J., Hua, Y., Wang, H., Song, T., Xue, Z., Ma, R., Cao, J., and Guan, H. Gpfl: Simultaneously learning global and personalized feature information for personalized federated learning. In *CVPR*, pp. 5041–5051, 2023b.

Zhang, J., Hua, Y., Wang, H., Song, T., Xue, Z., Ma, R., and Guan, H. Fedala: Adaptive local aggregation for personalized federated learning. In *AAAI*, pp. 11237–11244, 2023c.

Zhang, K., Yang, C., Li, X., Sun, L., and Yiu, S. M. Subgraph federated learning with missing neighbor generation. In *NeurIPS*, pp. 6671–6682, 2021b.

Zhang, T., Gao, L., Lee, S., Zhang, M., and Avestimehr, S. Timelyfl: Heterogeneity-aware asynchronous federated learning with adaptive partial training. In *CVPR*, pp. 5064–5073, 2023d.

Zhang, Y., Gao, S., Pei, J., and Huang, H. Improving social network embedding via new second-order continuous graph neural networks. In *KDD*, pp. 2515–2523, 2022b.

Zhou, T. and Konukoglu, E. Fedfa: Federated feature augmentation. In *ICLR*, 2023.

Zhu, B., Wang, L., Pang, Q., Wang, S., Jiao, J., Song, D., and Jordan, M. I. Byzantine-robust federated learning with optimal statistical rates. In *AISTATS*, pp. 3151–3178, 2023.

Zhu, Y., Li, X., Wu, Z., Wu, D., Hu, M., and Li, R.-H. Fedtad: Topology-aware data-free knowledge distillation for subgraph federated learning. *arXiv preprint arXiv:2404.14061*, 2024.

Zhu, Z., Hong, J., and Zhou, J. Data-free knowledge distillation for heterogeneous federated learning. In *ICML*, pp. 12878–12889, 2021.