Heterogeneity-Aware Knowledge Sharing for Graph Federated Learning

Wentao Yu
joint work with Sheng Wan, Shuo Chen, Bo Han, Chen Gong

TLDR: We address heterogeneity by aligning semantic knowledge and structural knowledge across clients.

paper: https://arxiv.org/abs/2601.21589
code: https://github.com/blgpb/FedSSA

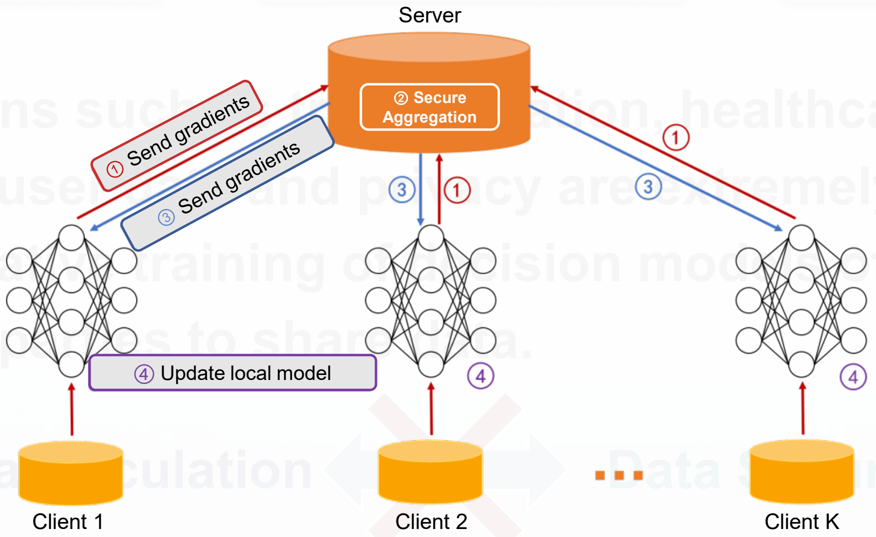
Contents

Motivation

Federated Learning

Goal : Data remains strictly local on each client, fully respecting ensuring **data security** and **user privacy** while enabling synergistic modeling across decentralized clients.

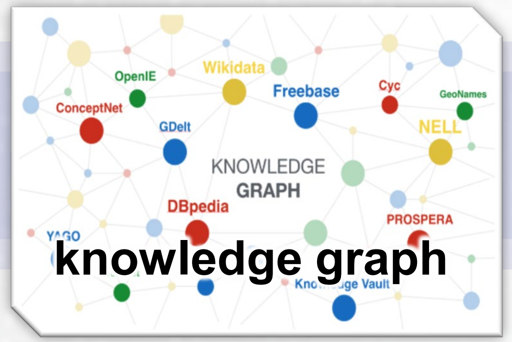
Insight : Data Stays Local, Models Move: Usable yet Invisible

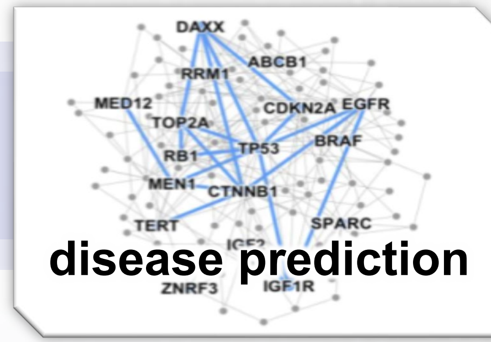


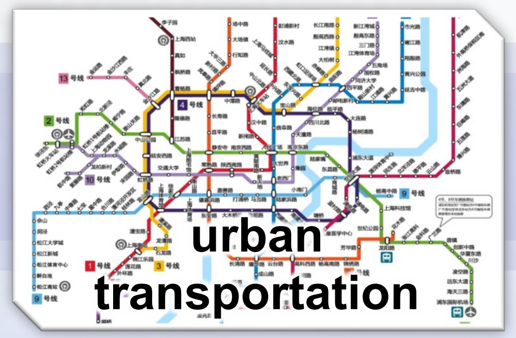
Motivation

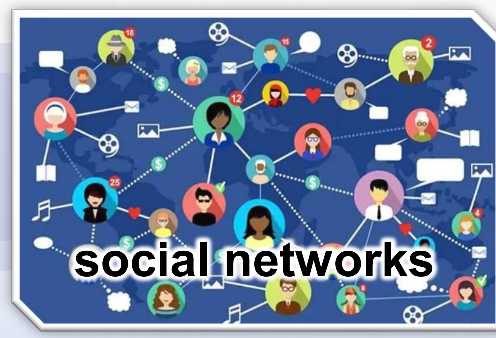
Graph Data

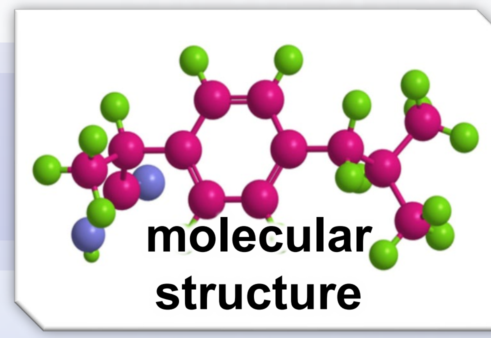
Most existing federated learning methods focus on unstructured data, while research on structured graph data remains very limited.





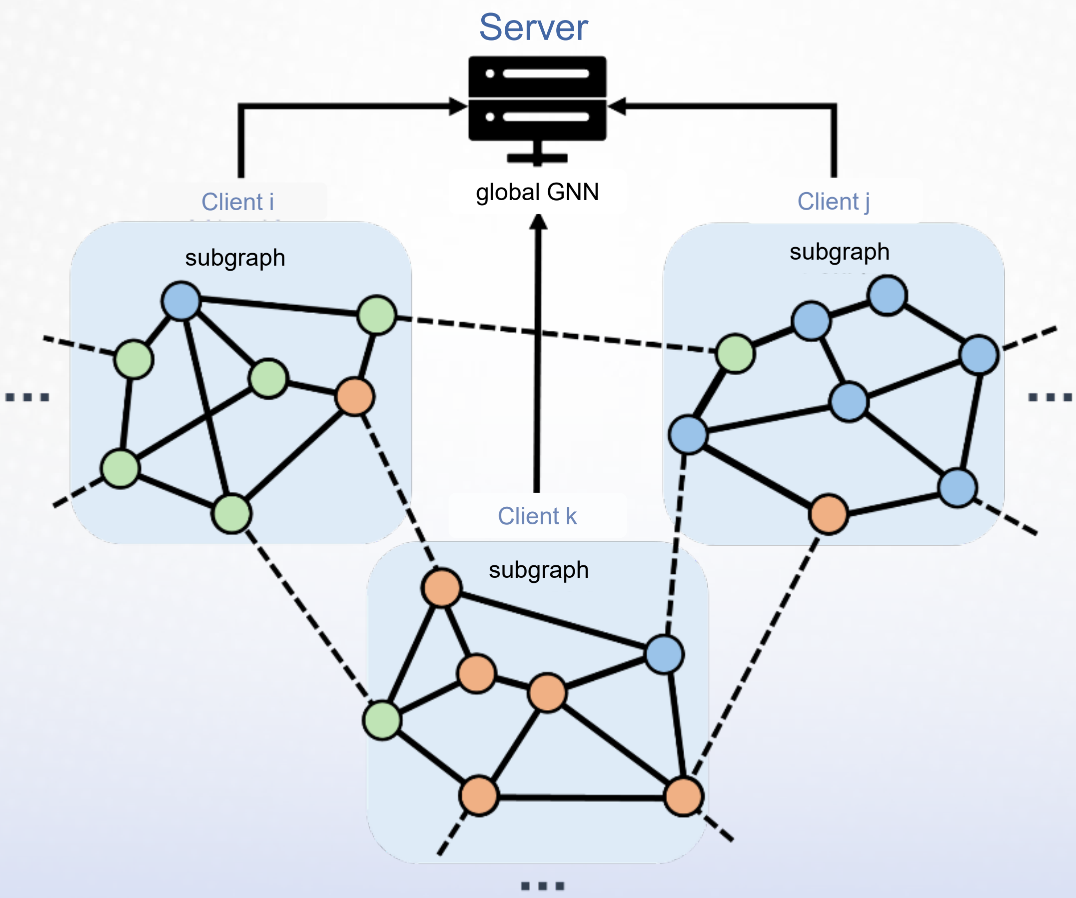






Motivation

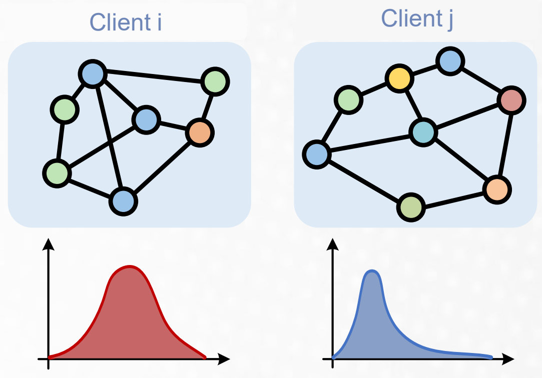
Challenges in Graph Federated Learning: Heterogeneity Caused by Inconsistent Graph Data Distributions Across Clients



Motivation

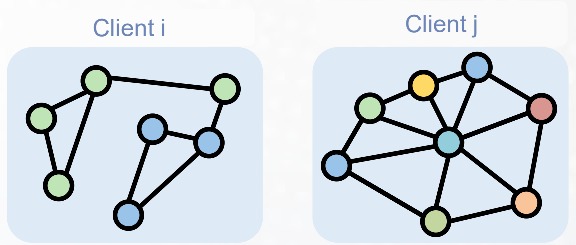
■ Node Feature Heterogeneity

Inconsistent Node Feature Distributions
Across Graphs from Different Clients



■ Structural Heterogeneity

Inconsistent Node Connection Patterns
Across Graphs from Different Clients



How can we explicitly address node feature heterogeneity and structural heterogeneity in graph federated learning?

Contents

南京理工大学
NANJING UNIVERSITY OF SCIENCE & TECHNOLOGY

Methodology

南京理工大学
NANJING UNIVERSITY OF SCIENCE & TECHNOLOGY

Research Question: How to address node feature heterogeneity and structural heterogeneity?

Problems:

Node Feature Heterogeneity

Structural Heterogeneity

Solutions:

semantic knowledge sharing

We infer class-wise node distributions, so that we can cluster clients based on inferred distributions and construct cluster-level representative distributions. We then **minimize the divergence** between local and cluster-level distributions to facilitate semantic knowledge sharing.

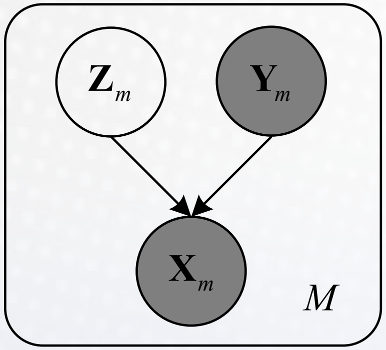
structural knowledge sharing

We propose a spectral energy measure to characterize structural information, so that we can cluster clients based on spectral energy and build cluster-level spectral GNNs. We then **align** the spectral characteristics of local spectral GNNs with those of cluster-level spectral GNNs to enable structural knowledge sharing.

For Node Feature Heterogeneity

■ **Key Innovations : inferring class-wise node distributions**

① **On the client side, we use latent variables to characterize the graph data distribution of each class-wise node feature.**



The graphical model of our proposed variational model, where $\mathbf{Z}_m$ is a latent variable, $\mathbf{X}_m$ and $\mathbf{Y}_m$ are observed variables.

log marginal likelihood:

$$\log p(\mathbf{X}_m, \mathbf{Y}_m) \geq \mathbb{E}_{\mathbf{Z}_m \sim q(\mathbf{Z}_m|\mathbf{X}_m,\mathbf{Y}_m)} \log \frac{p(\mathbf{X}_m|\mathbf{Z}_m,\mathbf{Y}_m)p(\mathbf{Z}_m)p(\mathbf{Y}_m)}{q(\mathbf{Z}_m|\mathbf{X}_m,\mathbf{Y}_m)}$$

ELBO:

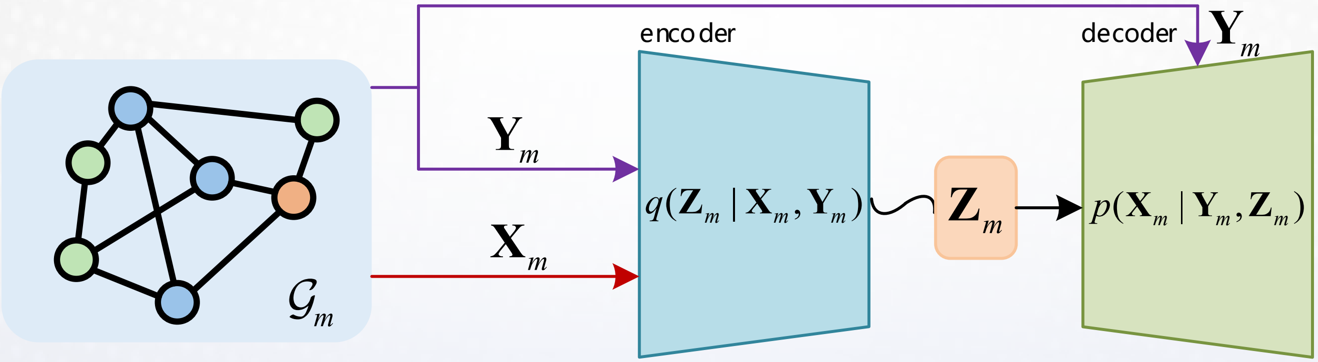
$$\mathcal{L}_{\mathrm{ELBO}}(\mathbf{X}_m, \mathbf{Y}_m) = \mathbb{E}_{\mathbf{Z}_m \sim q(\mathbf{Z}_m|\mathbf{X}_m,\mathbf{Y}_m)} \log p(\mathbf{X}_m|\mathbf{Z}_m,\mathbf{Y}_m) + \log p(\mathbf{Y}_m) - D_{\mathrm{KL}}(q(\mathbf{Z}_m|\mathbf{X}_m,\mathbf{Y}_m)||p(\mathbf{Z}_m))$$

For Node Feature Heterogeneity

- **Key Innovations : inferring class-wise node distributions**

  ① **On the client side, we use latent variables to characterize the graph data distribution of each class-wise node feature.**

the architecture of VGAE:



the latent distribution $q(\mathbf{Z}_m^c)$ of nodes with class label c on the m-th client:

$$\mathcal{I}_m^c \triangleq \{i \in \{1,2,\cdots,n_m\} \mid (\mathbf{Y}_m)_{[i,1]} = c\},$$

$$q(\mathbf{Z}_m^c) \triangleq q\left((\mathbf{Z}_m)_{[\mathcal{I}_m^c,:]} \,\middle|\, (\mathbf{X}_m)_{[\mathcal{I}_m^c,:]}, (\mathbf{Y}_m)_{[\mathcal{I}_m^c,:]}\right)$$

南京理工大学 NANJING UNIVERSITY OF SCIENCE & TECHNOLOGY

For Node Feature Heterogeneity

- **Key Innovations : clustering via inferred distributions**

② **Cluster based on node feature distribution uploaded by the client.**

client upload:

$$q(\mathbf{Z}_m^c) = \mathcal{N}(\boldsymbol{\mu}_m^c, \boldsymbol{\Sigma}_m^c) \text{ for c} = 1,2,\cdots,C$$

reparameterization sampling:

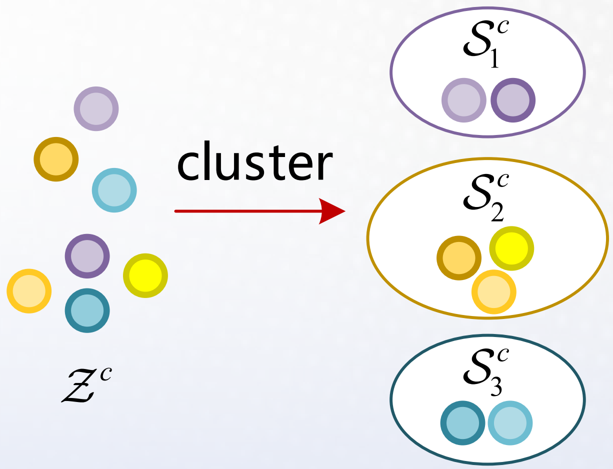
$$\tilde{\mathbf{Z}}_m^c = \boldsymbol{\mu}_m^c + (\boldsymbol{\Sigma}_m^c)^{\frac{1}{2}}\epsilon_m^c, \epsilon_m^c \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$$

the collection of clients with class c:

$$\mathcal{Z}^c = \{\tilde{\mathbf{Z}}_m^c | m \in \mathcal{M}\}$$

cluster into K_{node} clusters:

$$\text{Cluster}: \mathcal{Z}^c \rightarrow \{\mathcal{S}_1^c, \mathcal{S}_2^c, \cdots, \mathcal{S}_{K_{\text{node}}}^c\}$$



南京理工大学 NANJING UNIVERSITY OF SCIENCE & TECHNOLOGY

For Node Feature Heterogeneity

- **Key Innovations : cluster-based node feature distribution alignment**

③ **Approximate the node feature distributions of all nodes within a cluster into a single Gaussian distribution.**

mixed Gaussian distributions in cluster $\mathcal{S}_i^c$:

$$p(\mathbf{S}_i^c) = \sum_{m \in \mathcal{S}_i^c} w_m^c \cdot \mathcal{N}(\boldsymbol{\mu}_m^c, \boldsymbol{\Sigma}_m^c),$$

To reduce the computational complexity of mixed Gaussian distributions, we approximate them as a representative Gaussian distribution using moment matching:

$$q(\mathbf{S}_i^c) = \mathcal{N}(\boldsymbol{\mu}_i^c, \boldsymbol{\Sigma}_i^c)$$

the first moment:

$$\boldsymbol{\mu}_i^c = \sum_{m \in \mathcal{S}_i^c} w_m^c \boldsymbol{\mu}_m^c,$$

the second moment:

$$\Sigma_i^c = \left(\sum_{m \in \mathcal{S}_i^c} \omega_m^c (\boldsymbol{\Sigma}_m^c + \boldsymbol{\mu}_m^c (\boldsymbol{\mu}_m^c)^\top) \right) - \boldsymbol{\mu}_i^c (\boldsymbol{\mu}_i^c)^\top$$

minimize KL divergence to facilitate semantic knowledge sharing:

$$\mathcal{L}_{\text{node}} = \sum_{c=1}^{C} \sum_{i=1}^{K_{\text{node}}} \sum_{m \in \mathcal{S}_i^c} D_{\text{KL}} \left(q(Z_m^c) \parallel q(\mathcal{S}_i^c) \right)$$

南京理工大学 NANJING UNIVERSITY OF SCIENCE & TECHNOLOGY

For Structural Heterogeneity

■ Key Innovations : construct spectral energy measure

① **Construct the spectral energy measure and characterize graph structural information on each client.**

K-order polynomial graph filters:

$$\mathbf{P}_m = \sum_{k=0}^{K} \underbrace{w_m^k}_{\text{coefficients}} \overbrace{\mathbf{H}_m^k}^{\text{polynomial}}$$

The Laplacian matrix describes the structural information of graph and is reflected in the energy distribution along the spectral bands.

where $\mathbf{H}_m^k = (\mathbf{L}_m)^k \mathbf{X}_m \in \mathbb{R}^{n_m \times d}$

Since $\mathbf{L}_m = \mathbf{U}_m \Lambda_m \mathbf{U}_m^{\mathrm{T}}$, $\widetilde{\mathbf{X}}_m = \mathbf{U}_m^{\mathrm{T}} \mathbf{X}_m$

Therefore, $\mathbf{H}_m^k = \mathbf{U}_m \Lambda_m^k \widehat{\mathbf{X}}_m$. Its physical meaning represents the signal of the k-th spectral band. $w_m^k \in \mathbb{R}$ represents learnable coefficients of the k-th spectral band.

compute the average response vector for each band:

$$\mathbf{F}_m^k = w_m^k \mathbf{H}_m^k \in \mathbb{R}^{n_m \times d},$$

$$\mathbf{E}_m^k = \frac{1}{n_m} \sum_{i=1}^{n_m} \left(\mathbf{F}_m^k\right)_{[i,:]} \in \mathbb{R}^d$$

define the spectral energy measure:

$$\mathbf{S}_m = \left[\mathbf{E}_m^0, \mathbf{E}_m^1, \cdots, \mathbf{E}_m^K\right] \in \mathbb{R}^{d \times (K+1)}$$

This measure essentially acts as a **spectral fingerprint**, which reflects the energy distribution of structural topologies as learned by spectral GNN.

Methodology

For Structural Heterogeneity

- **Key Innovations : clustering via spectral energies**

② **Clustering clients based on the subspace distance spanned by the spectral energy measure.**

QR decomposition:

$$\mathbf{S}_m = \mathbf{Q}_m \mathbf{R}_m$$

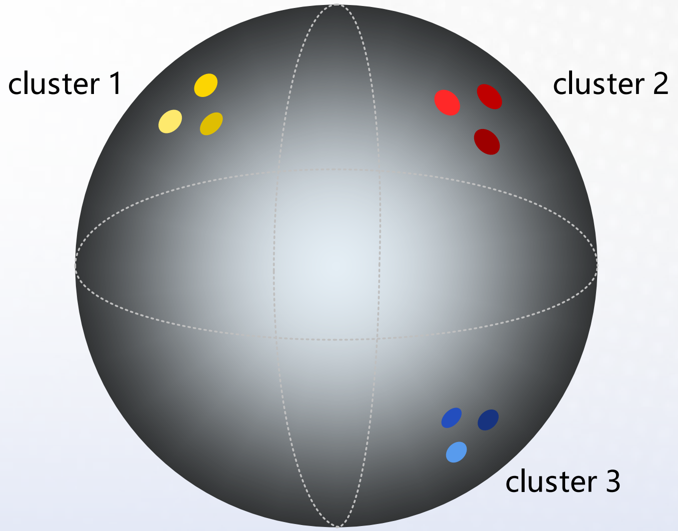
$$\mathbf{Q}_{\mathrm{m}}^{\top} \mathbf{Q}_{\mathrm{m}} = \mathbf{I}_{\mathrm{K+1}}$$

Chordal distance between two spectral energy measures:

$$d_{\mathrm{Chordal}}(\mathbf{Q}_m, \mathbf{Q}_n) = \left(K + 1 - \|\mathbf{Q}_m^{\top} \mathbf{Q}_n\|_{\mathrm{F}}^2\right)^{\frac{1}{2}}$$

cluster into K_{struct} clusters:

$$\text{Cluster:} \{\mathbf{S}_m\}_{m=1}^{M} \rightarrow \{\mathcal{T}_1^c, \mathcal{T}_2^c, \cdots, \mathcal{T}_{K_{\mathrm{struct}}}^c\}$$



For Structural Heterogeneity

- **Key Innovations : structural knowledge alignment**

③ **Align the spectral characteristics of local spectral GNNs with those of cluster-level spectral GNNs.**

compute the mean of learnable coefficients:

$$\overline{w}_j^k \leftarrow \frac{1}{|\mathcal{T}_j|} \sum_{m \in \mathcal{T}_j} w_m^k$$

minimize the discrepancy between the coefficients of local spectral GNNs and those of cluster-level spectral GNNs:

$$\mathcal{L}_{\text{align}} = \sum_{k=0}^{K} \sum_{j=1}^{K_{\text{struct}}} \sum_{m \in \mathcal{T}_j} \left| w_m^k - \overline{w}_j^k \right|$$

add a regularization term:

$$\mathcal{L}_{\text{reg}} = \sum\nolimits_{m=1}^{M} \sum\nolimits_{k=0}^{K} \left(\lambda_1 \left| w_m^k \right| + \frac{\lambda_2}{2} \left(w_m^k \right)^2 \right)$$

$$\mathcal{L}_{\text{struct}} = \mathcal{L}_{\text{align}} + \mathcal{L}_{\text{reg}}$$

Contents

Experimental Results

南京理工大学 NANJING UNIVERSITY OF SCIENCE & TECHNOLOGY

■ **Main Results**: node classification performance on **homophilic** graph datasets

	Cora			CiteSeer			PubMed			-
Methods	5 Clients	10 Clients	20 Clients	5 Clients	10 Clients	20 Clients	5 Clients	10 Clients	20 Clients	-
Local	81.30±0.21	79.94±0.24	80.30±0.25	69.02±0.05	67.82±0.13	65.98±0.17	84.04±0.18	82.81±0.39	82.65±0.03	-
FedAvg (McMahan et al., 2017)	74.45±5.64	69.19±0.67	69.50±3.58	71.06±0.60	63.61±3.59	64.68±1.83	79.40±0.11	82.71±0.29	80.97±0.26	-
FedProx (Li et al., 2020)	72.03±4.56	60.18±7.04	48.22±6.18	71.73±1.11	63.33±3.25	64.85±1.35	79.45±0.25	82.55±0.24	80.50±0.25	-
FedPer (Arivazhagan et al., 2019)	81.68±0.40	79.35±0.04	78.01±0.32	70.41±0.32	70.53±0.28	66.64±0.27	85.80±0.21	84.20±0.28	84.72±0.31	-
GCFL (Xie et al., 2021)	81.47±0.65	78.66±0.27	79.21±0.70	70.34±0.57	69.01±0.12	66.33±0.05	85.14±0.33	84.18±0.19	83.94±0.36	-
FedGNN (Wu et al., 2021)	81.51±0.68	70.12±0.99	70.10±3.52	69.06±0.92	55.52±3.17	52.23±6.00	79.52±0.23	83.25±0.45	81.61±0.59	-
FedSage+(Zhang et al., 2021)	72.97±5.94	69.05±1.59	57.97±12.60	70.74±0.69	65.63±3.10	65.46±0.74	79.57±0.24	82.62±0.31	80.82±0.25	-
FED-PUB (Baek et al., 2023)	83.70±0.19	81.54±0.12	81.75±0.56	72.68±0.44	72.35±0.53	67.62±0.12	86.79±0.09	86.28±0.18	85.53±0.30	-
FedGTA (Li et al., 2023)	80.06±0.63	80.59±0.38	79.01±0.31	70.12±0.10	71.57±0.34	69.94±0.14	87.75±0.01	86.80±0.01	87.12±0.05	-
AdaFGL (Li et al., 2024)	82.01±0.51	80.09±0.08	79.74±0.05	71.44±0.27	72.34±0.09	70.95±0.45	86.91±0.28	86.97±0.10	86.59±0.21	-
FedTAD (Zhu et al., 2024)	80.31±0.26	80.87±0.11	80.07±0.15	70.34±0.37	69.43±0.75	68.09±0.69	84.00±0.13	84.61±0.17	84.33±0.18	-
FedIIH (Yu et al., 2025a)	84.11±0.17	81.85±0.09	83.01±0.15	72.86±0.25	76.50±0.06	73.36±0.41	87.80±0.18	87.65±0.18	87.19±0.25	-
FedSSA (Ours)	**84.67±0.05**	**82.32±0.04**	**84.13±0.09**	**73.06±0.08**	**77.65±0.10**	**74.09±0.09**	**88.11±0.07**	**87.78±0.13**	**87.37±0.14**	-

	Amazon-Computer			Amazon-Photo			ogbn-arxiv			Avg.
Methods	5 Clients	10 Clients	20 Clients	5 Clients	10 Clients	20 Clients	5 Clients	10 Clients	20 Clients	All
Local	89.22±0.13	88.91±0.17	89.52±0.20	91.67±0.09	91.80±0.02	90.47±0.15	66.76±0.07	64.92±0.09	65.06±0.05	79.57
FedAvg (McMahan et al., 2017)	84.88±1.96	79.54±0.23	74.79±0.24	89.89±0.83	83.15±3.71	81.35±1.04	65.54±0.07	64.44±0.10	63.24±0.13	74.58
FedProx (Li et al., 2020)	85.25±1.27	83.81±1.09	73.05±1.30	90.38±0.48	80.92±4.64	82.32±0.29	65.21±0.20	64.37±0.18	63.03±0.04	72.84
FedPer (Arivazhagan et al., 2019)	89.67±0.34	89.73±0.04	87.86±0.43	91.44±0.37	91.76±0.23	90.59±0.06	66.87±0.05	64.99±0.18	64.66±0.11	79.94
GCFL (Xie et al., 2021)	89.07±0.91	90.03±0.16	89.08±0.25	91.99±0.29	92.06±0.25	90.79±0.17	66.80±0.12	65.09±0.08	65.08±0.04	79.90
FedGNN (Wu et al., 2021)	88.08±0.15	88.18±0.41	83.16±0.13	90.25±0.70	87.12±2.01	81.00±4.48	65.47±0.22	64.21±0.32	63.80±0.05	75.23
FedSage+(Zhang et al., 2021)	85.04±0.61	80.50±1.13	70.42±0.85	90.77±0.44	76.81±8.24	80.58±1.15	65.69±0.09	64.52±0.14	63.31±0.20	73.47
FED-PUB (Baek et al., 2023)	90.74±0.05	90.55±0.13	90.12±0.09	93.29±0.19	92.73±0.18	91.92±0.12	67.77±0.09	66.58±0.08	66.64±0.12	81.59
FedGTA (Li et al., 2023)	86.69±0.18	86.66±0.23	85.01±0.87	93.33±0.12	93.50±0.21	92.61±0.15	60.32±0.04	60.22±0.09	58.74±0.14	79.45
AdaFGL (Li et al., 2024)	80.20±0.05	83.62±0.26	84.53±0.23	86.69±0.19	89.85±0.83	88.11±0.05	52.73±0.19	51.77±0.36	50.94±0.08	76.97
FedTAD (Zhu et al., 2024)	82.20±1.20	85.50±0.33	83.91±1.54	92.29±0.39	90.59±0.09	89.18±0.84	65.35±0.14	64.06±0.25	64.45±0.13	78.87
FedIIH (Yu et al., 2025a)	90.74±0.13	90.86±0.23	90.44±0.05	93.42±0.02	94.22±0.08	93.55±0.09	70.30±0.06	69.34±0.02	68.65±0.04	83.10
FedSSA (Ours)	**91.09±0.07**	**91.30±0.13**	**90.62±0.09**	**93.86±0.15**	**94.62±0.14**	**93.76±0.07**	**70.86±0.13**	**69.47±0.08**	**68.77±0.13**	**83.53**

achieve the best performance among all methods

Experimental Results

■ **Main Results:** node classification performance on heterophilic graph datasets

Methods	Roman-empire 5 Clients	Roman-empire 10 Clients	Roman-empire 20 Clients	Amazon-ratings 5 Clients	Amazon-ratings 10 Clients	Amazon-ratings 20 Clients	Minesweeper 5 Clients	Minesweeper 10 Clients	Minesweeper 20 Clients	-
Local	33.65±0.13	28.42±0.26	23.89±0.32	45.03±0.31	**45.89±0.19**	46.02±0.02	71.35±0.17	69.96±0.16	69.31±0.09	-
FedAvg (McMahan et al., 2017)	38.93±0.32	35.43±0.32	32.00±0.39	41.26±0.53	41.66±0.14	42.20±0.21	72.60±0.08	69.68±0.09	71.36±0.16	-
FedProx (Li et al., 2020)	27.95±0.59	26.43±1.41	23.12±0.49	36.92±0.02	36.86±0.14	36.96±0.05	71.91±0.27	70.66±0.20	71.50±0.37	-
FedPer (Arivazhagan et al., 2019)	20.75±1.75	15.51±1.13	15.45±2.76	36.62±0.30	32.34±1.01	36.96±0.03	58.73±10.45	65.35±7.02	53.80±11.40	-
GCFL (Xie et al., 2021)	40.65±0.14	40.51±0.24	37.85±0.25	36.92±0.05	36.86±0.14	36.96±0.02	72.04±0.13	71.88±0.12	69.20±0.18	-
FedGNN (Wu et al., 2021)	30.26±0.11	29.09±0.13	26.60±0.02	36.80±0.06	36.72±0.00	36.45±0.09	72.15±0.13	71.08±0.07	71.71±0.27	-
FedSage+(Zhang et al., 2021)	57.26±0.00	49.07±0.00	38.36±0.00	36.82±0.00	36.71±0.00	37.03±0.02	77.74±0.00	72.80±0.00	69.70±0.00	-
FED-PUB (Baek et al., 2023)	40.80±0.26	36.77±0.30	32.67±0.39	44.41±0.41	44.85±0.17	45.39±0.50	72.18±0.02	71.69±0.71	71.41±0.87	-
FedGTA (Li et al., 2023)	61.56±0.27	60.94±0.19	59.65±0.28	41.22±0.66	39.40±0.44	39.24±0.12	73.54±1.56	72.65±1.21	69.63±4.54	-
AdaFGL (Li et al., 2024)	67.64±0.18	64.55±0.09	62.42±0.26	41.70±0.06	42.30±0.00	42.59±0.14	73.24±1.13	70.79±2.14	71.26±1.31	-
FedTAD (Zhu et al., 2024)	45.26±0.19	44.71±0.38	42.04±0.13	43.59±0.33	43.35±0.29	44.50±0.26	72.39±0.43	71.99±0.13	72.74±0.03	-
FedIIH (Yu et al., 2025a)	68.32±0.05	66.44±0.28	64.61±0.13	44.26±0.24	44.24±0.10	45.19±0.04	74.29±0.02	73.23±0.04	72.81±0.02	-
FedSSA (Ours)	**68.67±0.10**	**66.81±0.09**	**65.14±0.15**	**45.18±0.14**	45.11±0.15	**46.13±0.05**	**82.26±0.14**	**82.16±0.08**	**82.60±0.11**	-

Methods	Tolokers 5 Clients	Tolokers 10 Clients	Tolokers 20 Clients	Questions 5 Clients	Questions 10 Clients	Questions 20 Clients	Avg. 5 Clients	Avg. 10 Clients	Avg. 20 Clients	Avg. All
Local	67.81±0.17	70.04±0.23	62.34±0.67	66.73±0.57	57.96±0.10	60.00±0.21	56.91	54.45	52.31	54.56
FedAvg (McMahan et al., 2017)	60.74±0.31	54.73±0.50	56.36±0.39	65.68±0.23	58.91±0.22	60.33±0.15	55.84	52.08	52.45	53.46
FedProx (Li et al., 2020)	42.90±0.24	41.15±0.22	40.42±0.62	47.36±0.38	45.46±0.34	46.83±0.11	45.41	44.11	43.77	44.43
FedPer (Arivazhagan et al., 2019)	46.61±9.88	54.97±13.23	44.82±11.61	58.38±9.39	59.40±9.71	62.32±1.56	44.22	45.51	42.67	44.13
GCFL (Xie et al., 2021)	64.39±1.17	59.90±0.85	58.82±0.70	60.51±1.18	59.85±0.16	60.31±0.48	54.90	53.80	52.63	53.78
FedGNN (Wu et al., 2021)	43.10±0.27	41.57±0.07	40.70±0.74	47.55±0.02	45.65±0.12	47.39±0.13	45.97	44.82	44.57	45.12
FedSage+(Zhang et al., 2021)	75.06±0.00	71.31±0.00	69.73±0.00	64.95±0.00	65.06±0.00	59.33±0.00	62.37	58.99	54.83	58.73
FED-PUB (Baek et al., 2023)	70.88±0.58	72.46±0.68	65.26±0.59	67.71±3.99	59.64±0.52	62.48±2.92	59.20	57.08	55.44	57.24
FedGTA (Li et al., 2023)	60.83±0.45	55.18±1.20	57.89±1.61	65.56±1.91	58.29±1.57	61.70±0.35	60.54	57.29	57.62	58.49
AdaFGL (Li et al., 2024)	59.26±2.18	54.78±2.12	56.61±2.93	64.23±2.09	58.82±1.14	62.84±0.49	61.21	58.25	59.14	59.54
FedTAD (Zhu et al., 2024)	60.91±0.25	53.39±1.73	56.47±1.58	68.89±1.20	58.44±1.06	61.51±1.56	58.21	54.38	55.45	56.01
FedIIH (Yu et al., 2025a)	71.09±0.26	71.32±0.09	70.30±0.10	68.32±0.03	67.99±0.09	65.40±0.07	65.26	64.64	63.66	64.52
FedSSA (Ours)	**75.82±0.05**	**73.96±0.10**	**72.29±0.11**	**69.51±0.15**	**68.69±0.18**	**65.74±0.07**	**68.29**	**67.35**	**66.10**	**67.34**

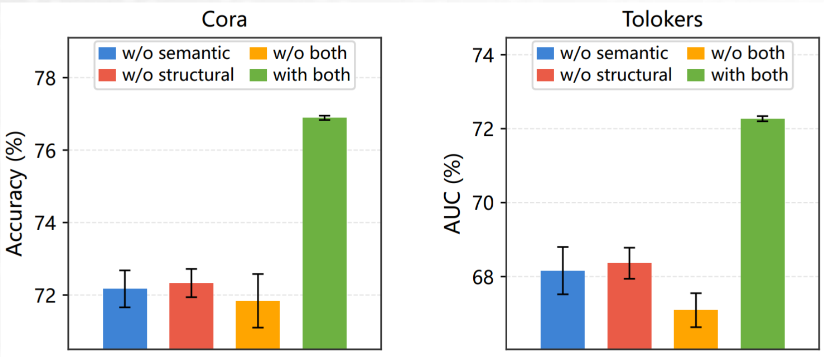
Compared with SOTA, the average performance improvement is 2.82%.

achieve the best average performance among all methods

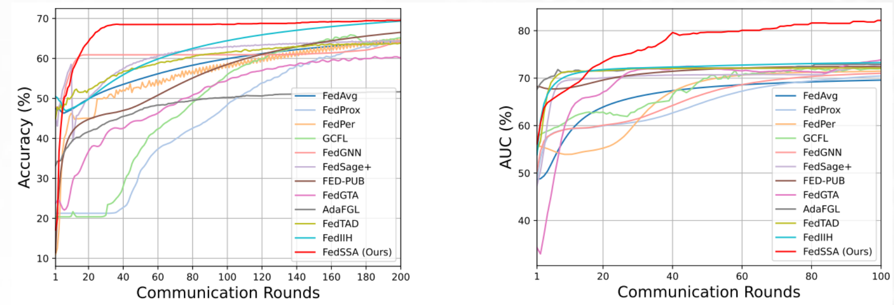
Experimental Results

南京理工大学
NANJING UNIVERSITY OF SCIENCE & TECHNOLOGY

ablation studies:

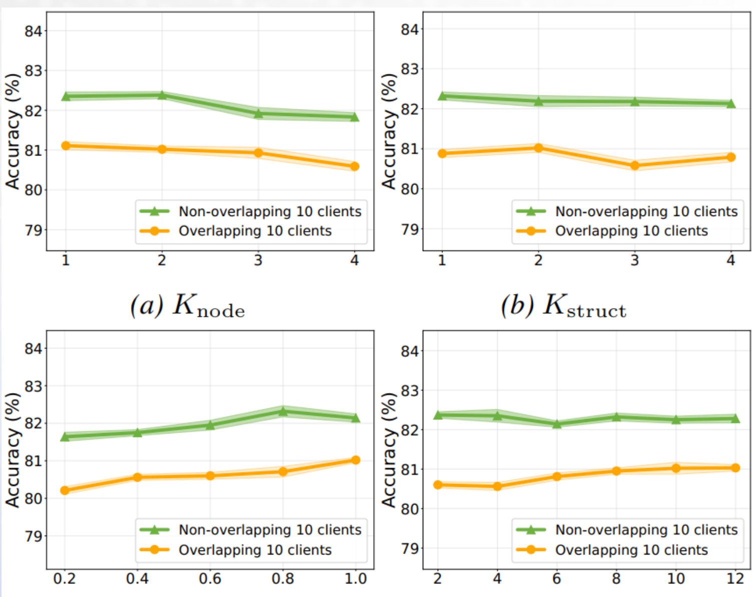


convergence curves:



(a) ogbn-arxiv (b) Minesweeper

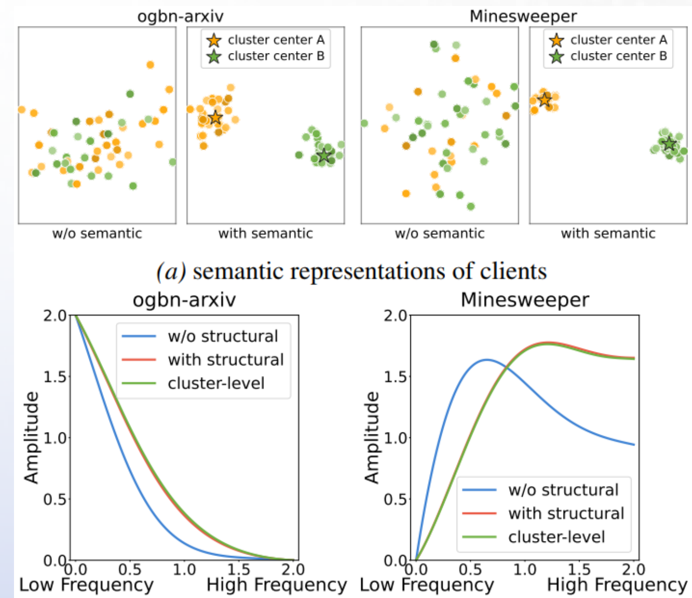
sensitivity analysis:



(a) K_{node} (b) K_{struct}

(c) λ_1 (d) λ_2

case studies:



(a) semantic representations of clients

(b) spectral properties of local models

Contents

Conclusion

We propose a graph **Fed**erated learning method via **S**emantic and **S**tructural **A**lignment (FedSSA) to address heterogeneity in Graph Federated Learning (GFL).

- On one hand, we **minimize the divergence** between local distributions and cluster-level distributions to mitigate the heterogeneity in node features.
- On the other hand, we **align the spectral characteristics** of local spectral GNNs with those of cluster-level spectral GNNs to address the heterogeneity in structural topologies.

Our proposed FedSSA achieves strong performance on eleven datasets and outperforms the second-best method by a large margin of **2.82%** in terms of classification accuracy.

Thank you!

Wentao Yu

joint work with Sheng Wan, Shuo Chen, Bo Han, Chen Gong

paper: https://arxiv.org/abs/2601.21589

code: https://github.com/blgpb/FedSSA