FEDERATED GRAPH LEARNING

Nhu Ngoc Hoang
nhu.ngoc.hoang@estudiantat.upc.edu

INTRODUCTION

Graph neural networks have emerged as a powerful paradigm for learning from graph-structured data, achieving state-of-the-art results across many domains, but typically require centralized access to large graph datasets, which is becoming increasingly infeasible due to privacy regulations and the sheer volume of data distributed across multiple organizations. Federated graph learning addresses this challenging by combining graph neural networks with federated learning, a powerful paradigm that enables collaborative training of machine learning models without sharing raw data, upholding data privacy and ownership.

PROBLEM FORMULATION

FEDERATED LEARNING

Federated learning [MMR+17] is a machine learning paradigm where multiple clients collaboratively train a model without exchanging raw data. There are M clients $\mathcal{C} = \{c_1, c_2, \ldots, c_M\}$ where each client c_k owns a private dataset $\mathcal{P}_k$. The goal of federated learning is to optimize some global objective function. Federated learning involves multiple rounds of updates orchestrated by a central server.

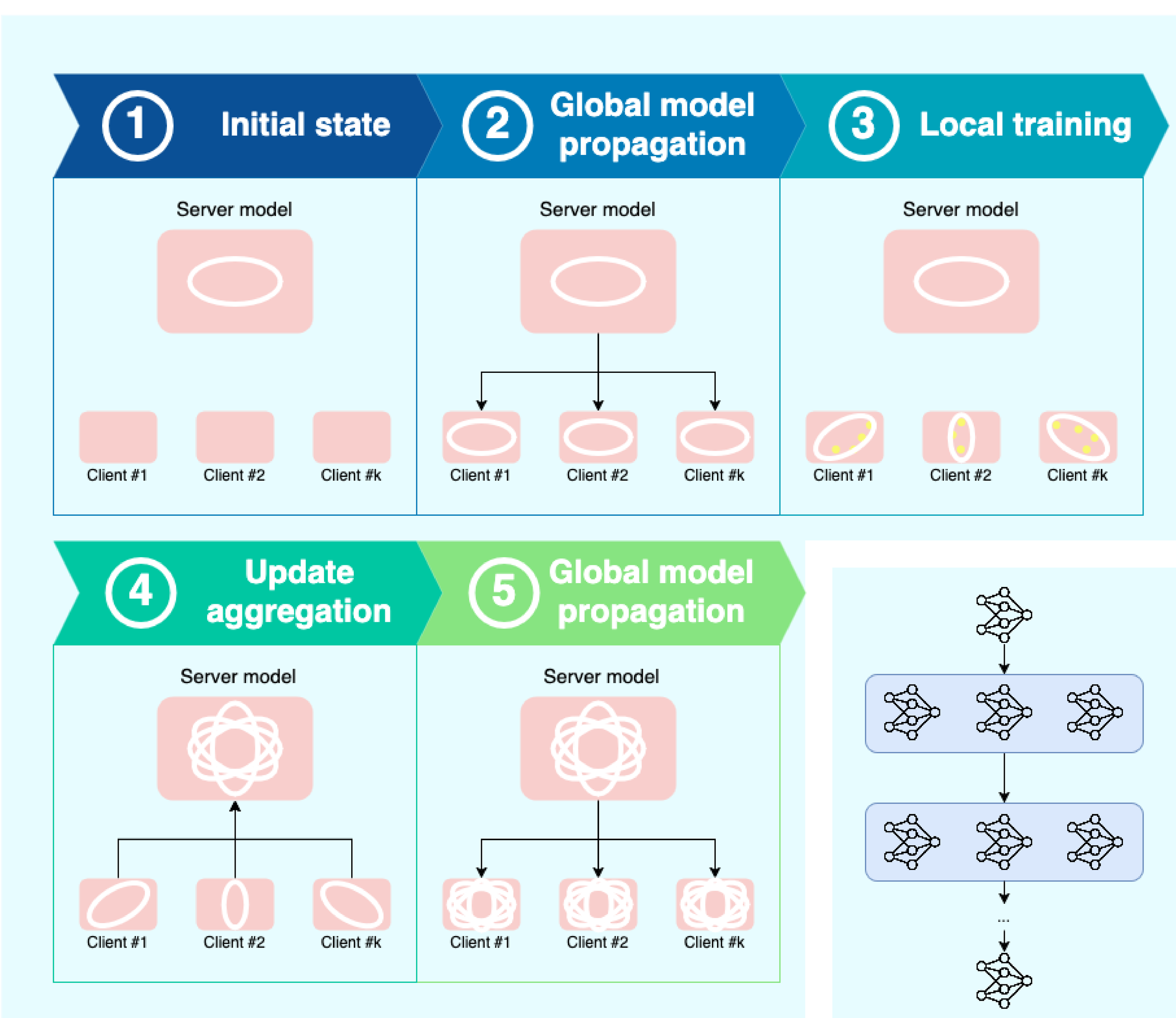

Figure 1. Federated learning protocol & graph neural networks simplified architecture.

GRAPH NEURAL NETWORKS

Graph neural networks have naturally emerged as powerful models for handling graph data [LXD+25]. Graph neural networks are a class of deep learning model designed to encode both structural and feature-based information in graphs via neighborhood propagation. Their ability to encode such rich contextual and topological information leads to high performance in tasks on the graph level (e.g. graph property prediction), node level (e.g. node classification), or edge level (e.g. missing edge prediction).

FEDERATED GRAPH LEARNING

Federated graph learning facilitates the training of graph neural networks across multiple data-owning clients without sharing raw graphs, ensuring decentralization and privacy of data.

TAXONOMY

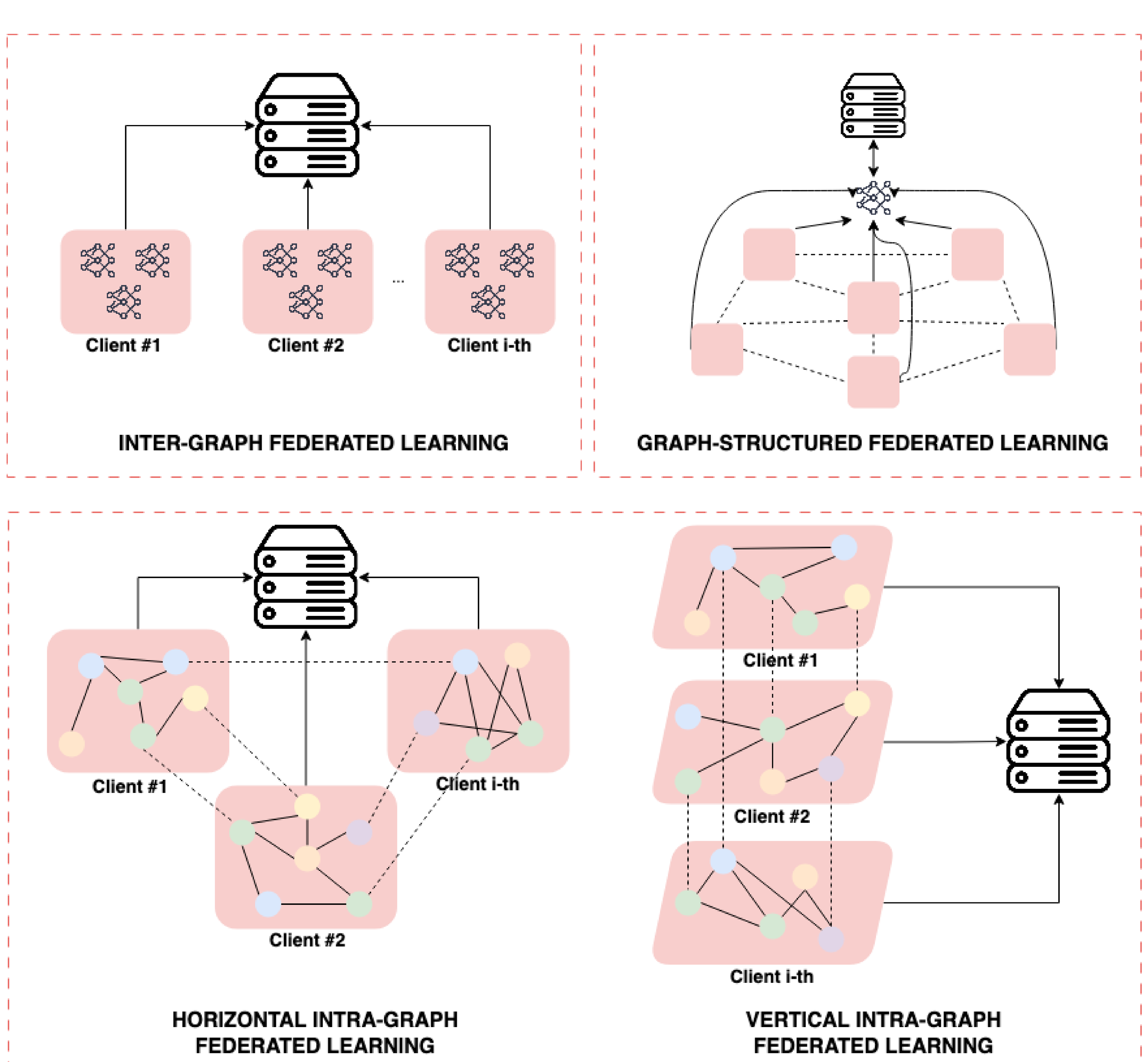

Figure 2. Federated graph learning taxonomy based on data partitioning.

CORE CHALLENGES & APPROACHES

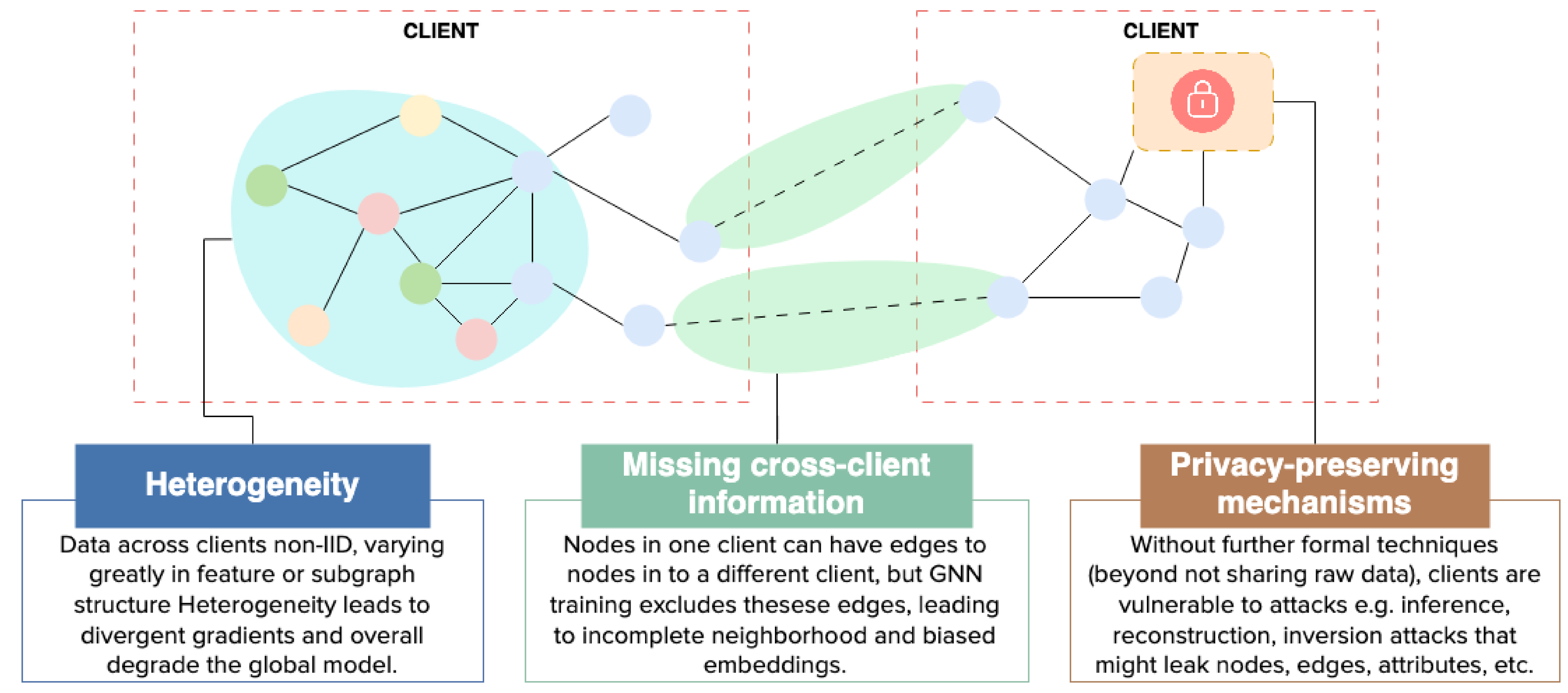

Figure 3. Three core challenges of federated graph learning.

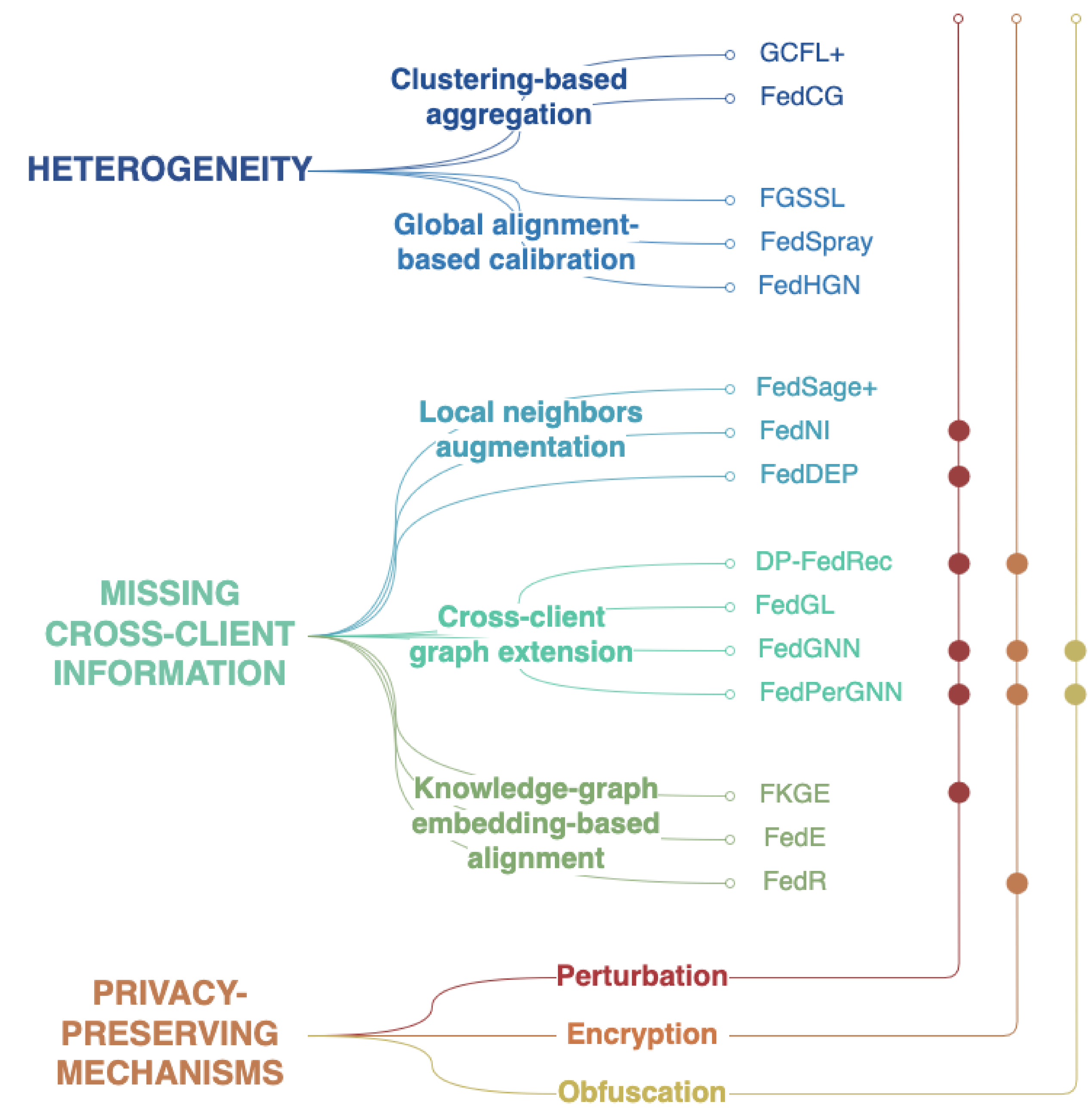

Figure 4. Summary of approaches by challenge addressed.

APPLICATIONS

Biomedical & healthcare
- Disease prediction
- Disease classification
- Molecular property prediction

Computer vision
- Image classification
- Video-based trajectory prediction

Recommendation systems
- Use bipartite user-item graphs to improve recommendation

Finance
- Fraud detection

OPEN CHALLENGES & FUTURE DIRECTIONS

- **Scalability**: Real-world graphs can have millions of nodes, while existing literature usually rely on simulating with graphs of smaller scale. For large-scale graphs, increased communication overhead and failure-prone devices might introduce significant bottlenecks.
- **Communication efficiency**: The learning process involves exchanging large messages (node embeddings, gradient matrices, etc.). Reducing communication is critical, especially for large-scale graphs.
- **Decentralized protocols**: Most federated graph learning schemes assume a trusted coordinating server. Fully decentralized algorithms (e.g. based on consensus or graph-based aggregation) remain largely unexplored.
- **Heterogeneity**: Graph heterogeneity remains a deep challenge. While clustering or calibrating can alleviate some issues, a unified theory is lacking.
- **Security and robustness**: Adversarial attacks on federated graph learning remain largely unexplored. While privacy-preserving mechanisms have emerged, further robustness-enhancing techniques need future research.

SELECTED REFERENCES

[KMA+21] Peter Kairouz, H. Brendan McMahan, Brendan Avent, Aurélien Bellet, and Mehdi Bennis et al, *Advances and open problems in federated learning*, Foundations and Trends® in Machine Learning (2021), no. 1–2, 1–210.

[KW17] Thomas N. Kipf and Max Welling, *Semi-supervised classification with graph convolutional networks*, 5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings, OpenReview.net, 2017.

[LXD+25] Rui Liu, Pengwei Xing, Zichao Deng, Anran Li, Cuntai Guan, and Han Yu, *Federated graph neural networks: Overview, techniques, and challenges*, IEEE Transactions on Neural Networks and Learning Systems (2025), no. 3, 4279–4295.

[MMR+17] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas, *Communication-Efficient Learning of Deep Networks from Decentralized Data*, Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 20–22 Apr 2017.

[WWL+22] Chuhan Wu, Fangzhao Wu, Lingjuan Lyu, Tao Qi, Yongfeng Huang, and Xing Xie, *A federated graph neural network framework for privacy-preserving personalization*, Nature Communications **13** (2022), no. 1.

[ZSW+21] Huanding Zhang, Tao Shen, Fei Wu, Mingyang Yin, Hongxia Yang, and Chao Wu, *Federated graph learning - A position paper*, CoRR (2021).

[ZWW+22] Kai Zhang, Yu Wang, Hongyi Wang, Lifu Huang, Carl Yang, Xun Chen, and Lichao Sun, *Efficient federated learning on knowledge graphs via privacy-preserving relation embedding aggregation*, Findings of the Association for Computational Linguistics: EMNLP 2022 (Abu Dhabi, United Arab Emirates) (Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang, eds.), Association for Computational Linguistics, December 2022, pp. 613–621.

[ZYL+21] 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 (M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan, eds.), vol. 34, Curran Associates, Inc., 2021, pp. 6671–6682.