

# UNSUPERVISED FEDERATED GRAPH MATCHING WITH GRAPHLET FEATURE EXTRACTION AND SEPARATE TRUST REGION

**Anonymous authors**

Paper under double-blind review

## ABSTRACT

Graph matching in the setting of federated learning is still an open problem. This paper proposes an unsupervised federated graph matching algorithm, UFGM, for inferring matched node pairs on different graphs across clients while maintaining privacy requirement, by leveraging graphlet theory and trust region optimization. First, the nodes' graphlet features are captured to generate pseudo matched node pairs on different graphs across clients as pseudo training data for tackling the dilemma of unsupervised graph matching in federated setting and leveraging the strength of supervised graph matching. An approximate graphlet enumeration method is proposed to sample a small number of graphlets and capture nodes' graphlet features. Theoretical analysis is conducted to demonstrate that the approximate method is able to maintain the quality of graphlet estimation while reducing its expensive cost. Second, we propose a separate trust region algorithm for pseudo supervised federated graph matching while maintaining the privacy constraints. In order to avoid expensive cost of the second-order Hessian computation in the trust region algorithm, we propose two weak quasi-Newton conditions to construct a positive definite scalar matrix as the Hessian approximation with only first-order gradients. We theoretically derive the error introduced by the separate trust region due to the Hessian approximation and conduct the convergence analysis of the approximation method.

## 1 INTRODUCTION

Federated graph learning (FGL) is a promising paradigm that enables collaborative training of shared machine learning models over large-scale distributed graph data, while preserving privacy of local data (Zheng et al., 2020; Chen et al., 2021; Zhang et al., 2021a). Only recently, researchers have started to attempt to study the FGL problems (Suzumura et al., 2019; Mei et al., 2019; Zhou et al., 2020b; Jiang et al., 2020; Wang et al., 2020a; Chen et al., 2021; Ke & Honorio, 2021; Wu et al., 2021; Wang et al., 2021a; He et al., 2021b;c). Most of them concentrate on node classification (Zhang et al., 2021b; Wang et al., 2022a; Chen et al., 2022a; Baek et al., 2022; Xie et al., 2023; Zhang et al., 2023; Li et al., 2023), graph classification (Xie et al., 2021; He et al., 2021a; Tan et al., 2022; Wang et al., 2022b), network embedding (Ni et al., 2021; Zhang et al., 2022; Hu et al., 2023; Zhu et al., 2023), and link prediction (Chen et al., 2022c; Baek et al., 2022). Graph matching (i.e., network alignment) is one of the most important research topics in the graph domain, which aims to match the same entities (i.e., nodes) across two or more graphs (Zhang & Yu, 2015; Zhang et al., 2015; Liu et al., 2016; 2017; Malmi et al., 2017; Vijayan & Milenkovic, 2018; Nassar et al., 2018; Zhou et al., 2018b; Chu et al., 2019; Wang et al., 2019b). It has been widely applied to many real-world applications ranging from protein network matching in bioinformatics (Kelley et al., 2003; Singh et al., 2008), user account linking in different social networks (Shu et al., 2016; Mu et al., 2016; Zhong et al., 2018; Li et al., 2018; Zhou et al., 2018a; Feng et al., 2019; Li et al., 2019a), and knowledge translation in multilingual knowledge bases (Xu et al., 2019b; Zhu et al., 2019), to geometric keypoint matching in computer vision (Fey et al., 2020).

While the existing techniques have achieved remarkable performance in the above graph learning domains, there is still a paucity of techniques of effective federated graph matching (FGM), which is much more difficult to study. Directly sharing and inferring matched node pairs on different graphs

across clients and local graphs over multiple clients gives rise to a serious privacy leakage concern and thus limits the applicability of graph matching in the centralized setting, such as user account linking in social networks and financial crime detection on transaction networks (Suzumura et al., 2019; Wang et al., 2019a; Zhang et al., 2021a; NSF; IBM), where the social network data and the bank customer and transfer data contain many sensitive information, advocating the invention of novel FGM techniques. In this work, we aim to answer the following questions: (1) How to train effective FGM models on distributed clients with maintaining high matching performance? (2) How to make FGM models with strong privacy protection for cross-client information exchange?

Research activities on centralized graph matching can be classified into two groups: supervised graph matching (Man et al., 2016; Zhou et al., 2018a; Yasar & Çatalyürek, 2018; Li et al., 2019b;a; Chu et al., 2019; Fey et al., 2020) and unsupervised graph matching (Zhou et al., 2018b; Heimann et al., 2018; Zhong et al., 2018; Li et al., 2018; Huynh et al., 2020b). The former utilizes a set of pre-matched node pairs between pairwise graphs belonging to the same entities as training data to learn an effective graph matching model by minimizing the distances (or maximizing the similarities) between the pre-matched node pairs. The latter fails to employ the strength of training data and thus often leads to sub-optimal solutions. However, supervised graph matching using the pre-matched node pairs as the training data is improper for the FGM scenarios due to privacy risks of direct cross-client information exchange when the graphs to be matched are distributed over different clients.

This motivates us to capture nodes’ graphlet features to generate pseudo matched node pairs on different graphs across clients as the pseudo training data for leveraging the strength of supervised graph matching. A graphlet is a small graph of size up to  $k$  nodes of a larger graph, such as triangle, wedge, or  $k$ -clique, which describes the local topology of a larger graph. A node’s local topology can be measured by a graphlet feature vector, where each component denotes the frequency of one type of graphlets. Thus, a graphlet feature vector is one of node structure representation (Shervashidze et al., 2009; Kondor et al., 2009; Soufiani & Airoldi, 2012; Jin et al., 2018; Tu et al., 2019). It is highly possible that the nodes in different graphs with the small distances regarding their graphlet features correspond to the same entities. Thus, they can be treated as the pseudo matched node pairs for pseudo supervised FGM.

However, graphlet enumeration one by one on large graphs is impossible due to expensive cost. We propose to leverage Monte Carlo Markov Chain (MCMC) technique for sampling a small number of graphlets. The number of graphlet samples is much smaller than that of all graphlets in the graphs, which dramatically improves the efficiency of graphlet enumeration. Theoretical analysis is conducted to demonstrate that the estimated graphlet count based on the MCMC sampling strategy is close to the actual count of all graphlets, which implies that the graphlet samples and all graphlets share similar distributions.

In order to maintain the privacy requirement of federated learning, we first encrypt local raw graph data on each client with a key shared by all clients (not accessed by the server). The encrypted graph data from all clients are accessed by only the server (not by other clients) for matching the graphs with each other. Note that stochastic gradient descent (SGD) optimization widely used in deep learning fails to work on the clients in the FGM, since each client can access only its own local graph data and thus cannot update local loss based on the pseudo matched node pairs. We propose a separate trust region algorithm for pseudo supervised FGM while maintaining the privacy constraints. Specifically, we separate model optimization from model evaluation in the trust region algorithm: (1) the server aggregates the local model parameter  $M_b^s$  on each client  $s$  into a global model parameter  $M_b$  at global iteration  $b$ , runs and evaluates  $M_b$  on the all pseudo training data  $\tilde{D}^{st}$  and the encrypted graph data, and computes the individual loss  $\mathcal{L}^s(M_b)$ , the gradient  $\nabla\mathcal{L}^s(M_b)$ , and the Hessian  $\nabla^2\mathcal{L}^s(M_b)$  for each client  $s$ ; (2) client  $s$  receives its individual  $\mathcal{L}^s(M_b)$ ,  $\nabla\mathcal{L}^s(M_b)$ , and  $\nabla^2\mathcal{L}^s(M_b)$  from the server and optimizes  $M_{b+1}^s$ .

Unfortunately, the second-order Hessian computation  $\nabla^2\mathcal{L}^s(M_b)$  in the separate trust region algorithm is time-consuming over large graphs. We propose to explore quasi-Newton conditions to construct a positive definite scalar matrix  $\alpha_b\mathbf{I}$ , where  $\alpha_b \geq 0$  is a scalar and  $\mathbf{I}$  is an identity matrix. Client  $s$  uses only first-order gradients  $\nabla\mathcal{L}^s(M_b)$  to compute the Hessian approximation, i.e.,  $z^T\nabla^2\mathcal{L}^s(M_b)z \approx \alpha_b z^T z$ . We theoretically derive the error by the separate trust region due to the Hessian approximation and conduct the convergence analysis of the approximation method.

To our best knowledge, this work is the first to offer an unsupervised federated graph matching solution for inferring matched node pairs on different graphs across clients while maintaining the privacy requirement of federated learning, by leveraging the graphlet theory and trust region optimization. Our UFGM method exhibits three compelling advantages: (1) The combination of the unsupervised FGM and the encryption of local raw graph data is able to provide strong privacy protection for sensitive local data; (2) The graphlet feature extraction can leverage the strength of supervised graph matching with the pseudo training data for improving the matching quality; and (3) The separate trust region for pseudo supervised FGM is helpful to enhance the efficiency while maintaining the privacy constraints.

Empirical evaluation on real datasets shows the superior performance of our UFGM model against several state-of-the-art centralized graph matching, federated domain adaption, and FGL methods.

## 2 BACKGROUND

### 2.1 SUPERVISED GRAPH MATCHING

Given a set of  $S$  graphs  $G = \{G^1, \dots, G^S\}$ . Each graph is denoted as  $G^s = (V^s, E^s)$  ( $1 \leq s \leq S$ ), where  $V^s = \{v_1^s, v_2^s, \dots\}$  is the set of nodes and  $E^s = \{(v_i^s, v_j^s) : 1 \leq i, j \leq |V^s|, i \neq j\}$  is the set of edges. Each  $G^s$  has a binary adjacency matrix  $\mathbf{A}^s$ , where each entry  $\mathbf{A}_{ij}^s = 1$  if there exists an edge  $(v_i^s, v_j^s) \in E^s$ ; otherwise  $\mathbf{A}_{ij}^s = 0$ .  $\mathbf{A}_{i:}^s$  specifies the  $i^{\text{th}}$  row vector of  $\mathbf{A}^s$  and is used to denote the representation of a node  $v_i^s$ .

The entire training data consist of a set of training data between pairwise graphs, i.e.,  $D = \{D^{12}, \dots, D^{1S}, \dots, D^{(S-1)S}\}$ . Each  $D^{st}$  ( $1 \leq s < t \leq S$ ) specifies a set of pre-matched node pairs  $D^{st} = \{(v_i^s, v_j^t) | v_i^s \leftrightarrow v_j^t, v_i^s \in V^s, v_j^t \in V^t\}$ , where  $v_i^s \leftrightarrow v_j^t$  represents that two nodes  $v_i^s$  and  $v_j^t$  are the equivalent ones in two graphs  $G^s$  and  $G^t$  and are treated as the same entity. The objective of supervised graph matching is to utilize  $D^{st}$  as the training data to identify the one-to-one matchings between nodes  $v_i^s$  and  $v_j^t$  in the test data.

Based on structure, attribute, or embedding features, existing efforts often aim to learn a matching function  $M$  to map the node pairs  $(v_i^s, v_j^t) \in D^{st}$  with different features across two graphs into common space, i.e., minimize the distances (or maximize the similarities) between source nodes  $M(v_i^s)$  and target ones  $M(v_j^t)$  (Man et al., 2016; Zhou et al., 2018a; Yasar & Çatalyürek, 2018; Li et al., 2019b;a). The node pairs  $(v_i^s, v_j^t) \in D^{st}$  with the smallest distances in the test data are selected as the matching results. This work follows these existing efforts to design the loss function.

$$\mathcal{L} = \sum_{s=1}^S \sum_{t=s+1}^S \mathbb{E}_{(v_i^s, v_j^t) \in D^{st}} \|M(v_i^s) - M(v_j^t)\|_2^2 \quad (1)$$

Graph convolutional networks (GCNs) have demonstrated their superior learning performance in network embedding tasks (Kipf & Welling, 2017). In this paper, if there are no specific descriptions, we utilize the GCNs to learn the embedding representation with the same dimensions of each node  $v_i^s$  in each graph  $G^s$ , based on its original structure features  $\mathbf{A}_{i:}^s$ . The embedding representation of  $v_i^s$  is denoted by  $\mathbf{v}_i^s$ . Thus, the objective of supervised graph matching is reformulated as follows.

$$\mathcal{L} = \sum_{s=1}^S \sum_{t=s+1}^S \mathbb{E}_{(v_i^s, v_j^t) \in D^{st}} \|M(\mathbf{v}_i^s) - M(\mathbf{v}_j^t)\|_2^2 \quad (2)$$

### 2.2 FEDERATED GRAPH MATCHING

In this paper, without loss of generality, we assume that each client contains only one local graph in the federated setting, but it is straightforward to extend to the case of multiple local graphs owned by each client. Given  $S$  clients with a set of  $S$  graphs  $\mathcal{G} = \{G^1, \dots, G^S\}$  and their local training data  $D = \{D^{12}, \dots, D^{1S}, \dots, D^{(S-1)S}\}$ , and a server, federated graph matching (FGM) aims to learn a global graph matching model  $M$  on the server by optimizing the problem below.

$$\begin{aligned} \min_{M \in \mathbb{R}^d} \mathcal{L}(M) &= \sum_{s=1}^S \mathcal{L}^s(M) = \sum_{s=1}^S \sum_{t=s+1}^S \frac{N^{st}}{N} L^{st}(M) \\ \text{where } L^{st}(M) &= \frac{1}{N^{st}} \sum_{(v_i^s, v_j^t) \in D^{st}} l_{ij}^{st}(M) \end{aligned} \quad (3)$$

where  $l_{ij}^{st}(M) = \|M(\mathbf{v}_i^s) - M(\mathbf{v}_j^t)\|_2^2$  denotes the loss function of the prediction on the pre-matched node pair  $(v_i^s, v_j^t) \in D^{st}$  made with  $M$ .  $\mathcal{L}^s(M)$  and  $\mathcal{L}(M)$  are the local loss function on client  $s$  and the global one respectively.  $N^{st} = |D^{st}|$  denotes the size of local training dataset  $D^{st}$ .  $N$  is the size of total training data  $D$ , i.e.,  $N = N^{12} + \dots + N^{1S} + \dots + N^{(S-1)S}$ . A local graph matching model  $M^s$  is optimized based on the local loss  $\mathcal{L}^s(M)$ . In the FGM,  $M$  is iteratively updated with the aggregation of all  $M^1, \dots, M^S$  on  $S$  clients in each round, i.e.,  $M = \sum_{s=1}^S \sum_{t=s+1}^S \frac{N^{st}}{N} M^s$ .

Observed from Eq.(3), when calculating the local loss  $\mathcal{L}^s(M)$  on client  $s$  for optimizing the local model  $M^s$ , we need to access the pre-matched node pairs  $\{v_i^s, v_j^t\} \in D^{st}$  and the graph  $G^t$  on client  $t$ . This operation obviously violates the privacy requirement of federated learning. Thus, it is difficult to utilize the pre-matched node pairs for supervised FGM.

### 3 MONTE CARLO MARKOV CHAIN FOR GRAPHLET FEATURE EXTRACTION

As discussed in the last section, the supervised graph matching usually achieves better performance than the unsupervised one. In addition, supervised FGM may lead to serious privacy concerns. In this work, we explore to capture nodes' graphlet features to generate pseudo matched node pairs on different graphs across clients as the pseudo training data for leveraging the strength of supervised graph matching while keeping the local graph data safe.

In order to prohibit other clients and server from accessing local raw graphs and embedding representations on any client  $s$  for maintaining the privacy requirement of FGM, we first utilize an efficient matrix generation method (Randall, 1993) to produce a random nonsingular matrix  $\mathbf{K}$  as a key. Each client employs  $\mathbf{K}$  to encrypt its network embedding  $\hat{\mathbf{v}}_i^s = \mathbf{v}_i^s \mathbf{K}$  from the original one  $\mathbf{v}_i^s$  and uses its inverse  $\mathbf{K}^{-1}$  to decrypt from  $\hat{\mathbf{v}}_i^s$  to  $\mathbf{v}_i^s = \hat{\mathbf{v}}_i^s \mathbf{K}^{-1}$ . The encrypted  $\hat{\mathbf{v}}_i^s$  from all clients will be uploaded to the server for graph matching. It is important that  $\mathbf{K}$  is kept secret between senders and recipients. In our setting,  $\mathbf{K}$  is shared by all clients, but not accessed by the server.

The first step of graphlet feature extraction is to enumerate all graphlets in a graph  $G = (V, E)$ . Concretely, let  $G_k$  be the set of all  $C$  connected induced  $k$ -subgraphs (with  $k$  nodes) in  $G$ . Let  $\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_R$  be all  $R$  types of non-isomorphic  $k$ -graphlets (with  $k$  nodes) for which we would like to count. We denote a  $k$ -subgraph  $g \in G_k$  that is isomorphic to a  $k$ -graphlet  $\mathcal{G}_r$  ( $1 \leq r \leq R$ ) as  $g \sim \mathcal{G}_r$ . The number of  $k$ -graphlets of type  $r$  in  $G$  is equal to

$$n_{kr}(G) = \sum_{g \in G_k} \mathbb{I}(g \sim \mathcal{G}_r) \quad (4)$$

where  $\mathbb{I}(\cdot)$  is an indicator function.

However, graphlet enumeration one by one on large-scale graphs is impossible due to expensive cost. We propose a MCMC sampling technique for which one can calculate the stationary distribution  $p$  on the  $k$ -subgraphs in  $G_k$ . We only sample a small number of  $k$ -subgraphs  $g_{k1}, \dots, g_{kO}$  in  $G$ , where the size  $O \ll C$ . Then we use Horvitz-Thompson inverse probability weighting to estimate the graphlet counts as follows.

$$\tilde{n}_{kr}(G) = \frac{1}{O} \sum_{o=1}^O \frac{\mathbb{I}(g_{ko} \sim \mathcal{G}_r)}{p(g_{ko})} \quad (5)$$

Next, we describe how to expand from 1-subgraphs to  $k$  subgraphs in the graphlet enumeration. For any  $(k-1)$ -subgraph  $g_{k-1}$ , we expand it to a  $k$ -subgraph by adding a node from its neighborhood  $\mathcal{N}_v(g_{k-1})$  at random in terms of a certain probability distribution, where  $\mathcal{N}_v(g_{k-1})$  is the set of all nodes adjacent to a certain node in  $g_{k-1}$  but not including all nodes in  $g_{k-1}$ .

This expansion operation can explore any subgraph in  $G_k$ . It iteratively builds a  $k$ -subgraph  $g_k$  from a starting node. First, suppose that a starting node  $v_1$  is sampled from the distribution  $q$ , which can be computed from local information. We assume that  $q(v) = \frac{f(\text{deg}(v))}{F}$ , where  $f(x)$  is a certain function (usually a polynomial) and  $F$  is a user-defined normalizing factor. Thus, a 1-subgraph  $g_1 = \{v_1\}$  is generated. Second, it samples an edge  $(v_1, v_2)$  uniformly in  $\mathcal{N}_e(g_1)$ , where  $\mathcal{N}_e(g_1)$  is the set of all edges that connect a node in  $g_1$  and a node outside of  $g_1$ . Thus, a node  $v_2$  is then attached to  $g_1$ , forming a 2-subgraph  $g_2 = g_1 \cup v_2 \cup (v_1, v_2)$ . Similarly, at each iteration, it samples an edge  $(v_i, v_{j+1})$  ( $1 \leq i \leq j$ ) from  $\mathcal{N}_e(g_j)$  uniformly at random and attach the node  $v_{j+1}$  to the

subgraph  $g_j$ , forming a  $j + 1$ -subgraph  $g_{j+1} = g_j \cup v_{j+1} \cup (v_i, v_{j+1})$ . After  $k - 1$  iterations, we obtain a  $k$ -subgraph  $g_k$ . Once  $g_k$  has been sampled we need to classify it into a graphlet type, i.e.,  $g_k \sim \mathcal{G}_r$ . The method repeats the above process  $O$  times until  $O$   $k$ -graphlets  $g_{k1}, g_{k1}, \dots, g_{kO}$  are produced.

We conduct the theoretical analysis to evaluate the permanence of our graphlet enumeration based on the MCMC sampling, in terms of the difference between the estimated and actual graphlet counts.

In the estimation  $\tilde{n}_{kr}(G)$  in Eq.(5), a key problem is to calculate  $p(g_{ko})$ . The probability  $p(g_k)$  of getting a  $k$ -subgraph  $g_k$  via subgraph expansion from a  $(k - 1)$ -subgraph  $g_{k-1}$  is given by the sum  $p(g_k) = \sum_{g_{k-1} \subset g_k} \mathbb{P}(g_k | g_{k-1}) p(g_{k-1})$ , where the sum is taken over all connected  $(k - 1)$ -subgraphs  $g_{k-1} \subset g_k$ , and  $\mathbb{P}(g_k | g_{k-1})$  is the probability of getting from  $g_{k-1}$  to  $g_k$  in the expansion process.

$$p(g_k) = \sum_{g_{k-1} \subset g_k} p(g_{k-1}) \frac{\deg_{g_{k-1}}(V_{g_k} - V_{g_{k-1}})}{|\mathcal{N}_e(g_{k-1})|} = \sum_{g_{k-1} \subset g_k} p(g_{k-1}) \frac{|E_{g_k}| - |E_{g_{k-1}}|}{\sum_{v \in V_{g_{k-1}}} \deg(v) - 2|E_{g_{k-1}}|} \quad (6)$$

where for a subgraph  $g_k \subseteq G$ ,  $V_{g_k}$  the set of its nodes and  $E_{g_k}$  is the set of its edges.  $\deg_{g_{k-1}}(V)$  specifies the number of nodes in  $g_{k-1}$  that are connected to a node set  $V$ .  $\deg(v)$  denotes the number of associated edges of a node  $v$ .

In order to calculate  $p(g_k)$ , we need to consider all possible orderings of nodes in  $g_k$ . Assume that the original node ordering of  $g_k$  via the subgraph expansion is  $x_k = \{v_1, v_2, \dots, v_k\}$ . Let  $\mathcal{S}(g_k) = [v_1, v_2, \dots, v_k]$  be the set of all possible node sequences of  $x_k$ . Notice that an induced subgraph  $h_l(x_k) = \{v_1, v_2, \dots, v_l, x_k, G\}$  of graph  $G$  with the first  $l$  nodes  $\{v_1, v_2, \dots, v_l\}$  in  $x_k$  must be a connected subgraph for any  $l$  ( $1 \leq l \leq k$ ). Thus, we have

$$\mathcal{S}(g_k) = \{[v_1, \dots, v_k] | \{v_1, \dots, v_k\} = V_{g_k}, g_k | \{v_1, \dots, v_l\} \text{ is connected}\} \quad (7)$$

The following theorems give an explicit solution of the probability  $p(g_k)$  of getting a  $k$ -subgraph  $g_k$  via subgraph expansion and the variance of the estimation  $\tilde{n}_{kr}(G)$  of graphlet counts.

**Theorem 1.** *Let  $x_k = \{v_1, v_2, \dots, v_k\}$  be the original node ordering of  $g_k$  via the subgraph expansion,  $\mathcal{S}(g_k) = [v_1, v_2, \dots, v_k]$  be the set of all possible node sequences of  $x_k$ ,  $x_k[i]$  be the  $i^{\text{th}}$  node in  $x_k$ ,  $F$  be a user-defined normalizing factor in the subgraph expansion, and  $h_l(x_k) = \{v_1, v_2, \dots, v_l, x_k, G\}$  be an induced subgraph of graph  $G$  with the first  $l$  nodes  $\{v_1, v_2, \dots, v_l\}$  in  $x_k$ , then the probability of getting a  $k$ -subgraph  $g_k$  via the subgraph expansion is*

$$p(g_k) = \sum_{x_k \in \mathcal{S}(g_k)} \frac{f(\deg(x_k[1]))}{F} \prod_{l=1}^{k-1} \frac{|E_{h_{l+1}(x_k)}| - |E_{h_l(x_k)}|}{\sum_{i=1}^l \deg(x_k[i]) - 2|E_{h_l(x_k)}|} \quad (8)$$

**Theorem 2.** *Let  $\tilde{n}_{kr}(G) = \frac{1}{O} \sum_{o=1}^O \frac{\mathbb{I}(g_{ko} \sim \mathcal{G}_r)}{p(g_{ko})}$  be the estimation of graphlet counts,  $d_1, \dots, d_k$  be the  $k$  highest degrees of nodes in  $G$ , and denote  $D = \prod_{l=2}^{k-1} (d_1 + \dots + d_k)$ . If  $q$  for sampling the starting node is the stationary distribution of the node random walk, then the upper bound of the variance  $\text{Var}(\tilde{n}_{kr}(G))$  is*

$$\text{Var}(\tilde{n}_{kr}(G)) \leq \frac{1}{O} n_{kr}(G) \frac{2|E_G|}{|\mathcal{S}(\mathcal{G}_r)|} D \quad (9)$$

Please refer to Appendix A.2 for detailed proof of Theorems 1 and 2.

It is observed that the variance  $\text{Var}(\tilde{n}_{kr}(G))$  is small when the distribution of  $p(g_k)$  is close to uniform distribution. A larger  $p(g_k)$  results in a smaller variance of the estimator. Thus, the variation can be reduced by an appropriate choice of  $q$  for sampling the starting node, say a smaller normalizing factor  $F$ . In this case, the estimated graphlet count  $\tilde{n}_{kr}(G)$  is close to the actual count  $n_{kr}(G)$ , which implies that the graphlet samples and all graphlets share similar distributions.

We capture the graphlet features of a node by computing the frequency of each type of graphlet with size up to  $k$  that is associated with this node. For the node pairs between pairwise graphs, we compute the cosine similarity scores based on the graphlet features on all  $R$  types of graphlet. The top- $K$  node pairs with the largest similarities between pairwise graphs  $G^s$  and  $G^t$  are treated as the pseudo matched node pairs and added to the pseudo training data  $\tilde{D}^{st}$ .

#### 4 SEPARATE TRUST REGION FOR UNSUPERVISED FEDERATED GRAPH MATCHING

In this work, according to the graphlet-based pseudo training data  $\tilde{D}^{st}$  and the encrypted network embedding  $\hat{v}_i^s$ , we propose a separate trust region algorithm for pseudo supervised FGM while maintaining the privacy constraints. Specifically, we separate model optimization from model evaluation in the trust region algorithm: (1) the server aggregates the local model parameter  $M_b^s$  on each client  $s$  into a global model parameter  $M_b$  at global iteration  $b$ , runs and evaluates  $M_b$  on all the pseudo training data  $\tilde{D}^{st}$  and the encrypted network embeddings  $\hat{v}_i^s$ , and computes the individual loss  $\mathcal{L}^s(M_b)$ , the gradient  $\nabla \mathcal{L}^s(M_b)$ , and the Hessian  $\nabla^2 \mathcal{L}^s(M_b)$  for each client  $s$ ; (2) client  $s$  receives its individual  $\mathcal{L}^s(M_b)$ ,  $\nabla \mathcal{L}^s(M_b)$ , and  $\nabla^2 \mathcal{L}^s(M_b)$  from the server and optimizes  $M_{b+1}^s$ .

$$\begin{aligned} \text{Server : Compute } M_b &= \sum_{s=1}^S \sum_{t=s+1}^S \frac{N^{st}}{N} M_b^s, \quad L^{st}(M_b) = \frac{1}{N^{st}} \sum_{(v_i^s, v_j^t) \in \tilde{D}^{st}} \|M_b(\hat{v}_i^s) - M_b(\hat{v}_j^t)\|_2^2, \\ \mathcal{L}^s(M_b) &= \sum_{t=s+1}^S \frac{N^{st}}{N} L^{st}(M_b), \quad \nabla \mathcal{L}^s(M_b), \quad \text{and } \nabla^2 \mathcal{L}^s(M_b) \end{aligned} \quad (10)$$

$$\begin{aligned} \text{Client } s : \text{ Optimize } z^* &= \arg \min u_b(z) = \mathcal{L}^s(M_b) + (\nabla \mathcal{L}^s(M_b))^T z + \frac{1}{2} z^T \nabla^2 \mathcal{L}^s(M_b) z, \quad \text{s.t. } \|z\| \leq \Delta^s \\ \text{Update } M_{b+1}^s &= M_b^s + z^* \end{aligned} \quad (11)$$

where  $\Delta^s > 0$  is the trust-region radius.  $z^*$  is the trust-region step. The individual loss  $\mathcal{L}^s(M_b)$  aims to minimize the sum of distance between nodes on client  $s$  and nodes on other clients in the pseudo training data  $\tilde{D}^{st}$ . The node pairs with the smallest distance between pairwise encrypted network embeddings are selected as the matching results.

A key challenge in the separate trust region algorithm is to compute the second-order Hessian computation  $\nabla^2 \mathcal{L}^s(M_b)$ . It is time-consuming over large-scale graph data. We propose to explore quasi-Newton conditions to construct a positive definite scalar matrix  $\alpha_b \mathbf{I}$ , where  $\alpha_b \geq 0$  is a scalar and  $\mathbf{I}$  is an identify matrix, as the Hessian approximation with only first-order gradients, i.e.,  $z^T \nabla^2 \mathcal{L}^s(M_b) z \approx \alpha_b z^T z$ .

Concretely, the quasi-Newton condition is given as follows.

$$\nabla^2 \mathcal{L}^s(M_b) z_b = y_b \quad (12)$$

where  $z_b = M_{b+1} - M_b$  and  $y_b = \nabla \mathcal{L}^s(M_{b+1}) - \nabla \mathcal{L}^s(M_b)$ . The condition is derived from the following quadratic model.

$$u_{b+1}(z) = \mathcal{L}^s(M_{b+1}) + (\nabla \mathcal{L}^s(M_{b+1}))^T z + \frac{1}{2} z^T \nabla^2 \mathcal{L}^s(M_{b+1}) z \quad (13)$$

The quadratic model is an approximation of the objective function at iteration  $b+1$  and satisfies the following three interpolation conditions:

$$(1) u_{b+1}(0) = \mathcal{L}^s(M_{b+1}), \quad (2) \nabla u_{b+1}(0) = \nabla \mathcal{L}^s(M_{b+1}), \quad (3) \nabla u_{b+1}(-z_b) = \nabla \mathcal{L}^s(M_b) \quad (14)$$

It is difficult to satisfy the quasi-Newton equation in Eq.(12) with a nonsingular scalar matrix (Farid et al., 2010). A recent study introduced a weak condition form by projecting the quasi-Newton equation in Eq.(12) in the direction  $z_b$  (J. E. Dennis & Wolkowicz, 1993).

$$z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b = z_b^T y_b \quad (15)$$

The choice of  $z_b$  may influence the quality of the curvature information provided by the weak quasi-Newton condition. Another weak condition is directly derived from an interpolation emphasizing more on function values rather than from the projection of the quasi-Newton condition (xiang Yuan, 1991).

$$u_{b+1}(-z_b) = \mathcal{L}^s(M_b) \quad (16)$$

By combining sub-conditions (1) and (2) in Eq.(14) and replacing (3) with Eq.(16), we can get another weak quasi-Newton condition.

$$z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b = 2 \left( \mathcal{L}^s(M_b) - \mathcal{L}^s(M_{b+1}) + z_b^T \nabla \mathcal{L}^s(M_{b+1}) \right) \quad (17)$$

By integrating two types of weak quasi-Newton conditions together, we have a generalized weak quasi-Newton condition.

$$\begin{aligned} z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b &= (1 - \omega) z_b^T y_b + \omega \left[ 2(\mathcal{L}^s(M_b) - \mathcal{L}^s(M_{b+1})) + 2z_b^T \nabla \mathcal{L}^s(M_{b+1}) \right] \\ &= z_b^T y_b + \omega \left[ 2(\mathcal{L}^s(M_b) - \mathcal{L}^s(M_{b+1})) + (\nabla \mathcal{L}^s(M_b) + \nabla \mathcal{L}^s(M_{b+1}))^T z_b \right] \end{aligned} \quad (18)$$

where  $\omega \geq 0$  is the weight. If  $\nabla^2 \mathcal{L}^s(M_{b+1})$  is set to be a scalar matrix  $\alpha_{b+1}^*(\omega) \mathbf{I}$ , then we have

$$\alpha_{b+1}(\omega) = \frac{z_b^T y_b + \omega \left[ 2(\mathcal{L}^s(M_b) - \mathcal{L}^s(M_{b+1})) + (\nabla \mathcal{L}^s(M_b) + \nabla \mathcal{L}^s(M_{b+1}))^T z_b \right]}{z_b^T z_b} \quad (19)$$

The following theorems derive the error introduced by the separate trust region due to the Hessian approximation and conduct the convergence analysis of the approximation method.

**Theorem 3.** Let  $d$  be the dimension of the flattened  $M_{b+1}$ ,  $\otimes$  be an appropriate tensor product,  $\mathcal{A}_{b+1} \in \mathbb{R}^{d \times d \times d}$  and  $\mathcal{B}_{b+1} \in \mathbb{R}^{d \times d \times d \times d}$  are the tensors of  $\mathcal{L}^s(M_{b+1})$  at iteration  $b + 1$  satisfying

$$\mathcal{A}_{b+1} \otimes z_b^3 = \sum_{i,j,k=1}^d \frac{\partial^3 \mathcal{L}^s(M_{b+1})}{\partial M^i \partial M^j \partial M^k} z_b^i z_b^j z_b^k \quad (20)$$

and

$$\mathcal{B}_{b+1} \otimes z_b^4 = \sum_{i,j,k,l=1}^d \frac{\partial^4 \mathcal{L}^s(M_{b+1})}{\partial M^i \partial M^j \partial M^k \partial M^l} z_b^i z_b^j z_b^k z_b^l. \quad (21)$$

Suppose that  $\mathcal{L}^s(M_{b+1})$  is sufficiently smooth, if  $\|z_b\|$  is small enough, then we have

$$z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b - \alpha_{b+1}(\omega) z_b^T z_b = \left( \frac{1}{2} - \frac{\omega}{6} \right) \mathcal{A}_{b+1} \otimes z_b^3 - \left( \frac{1}{6} - \frac{\omega}{12} \right) \mathcal{B}_{b+1} \otimes z_b^4 + \mathcal{O}(\|z_b\|^5) \quad (22)$$

**Theorem 4.** Suppose  $\|\nabla \mathcal{L}^s(M_b)\| \neq 0$ , the solution  $z_b$  of the separate trust region optimization  $\arg \min u_b(z) = \mathcal{L}^s(M_b) + (\nabla \mathcal{L}^s(M_b))^T z + \frac{1}{2} z^T \nabla^2 \mathcal{L}^s(M_b) z$ , s.t.  $\|z\| \leq \Delta^s$  in Eq.(11) satisfies

$$u_b(0) - u_b(z_b) \geq \frac{1}{2} \|\nabla \mathcal{L}^s(M_b)\| \min \left\{ \Delta^s, \frac{\|\nabla \mathcal{L}^s(M_b)\|}{\alpha_b} \right\} \quad (23)$$

Please refer to Appendix A.2 for detailed proof of Theorems 3 and 4.

Finally, the separate trust region based on two weak quasi-Newton conditions is given below.

$$z^* = \arg \min u_b(z) \approx \mathcal{L}^s(M_b) + (\nabla \mathcal{L}^s(M_b))^T z + \frac{1}{2} \alpha_b(\omega) z^T z, \text{ s.t. } \|z\| \leq \Delta^s \quad (24)$$

## 5 EXPERIMENTAL EVALUATION

In this section, we have evaluated the performance of our UFGM model and other comparison methods for federated graph matching over several representative federated graph datasets to date. We show that UFGM with graphlet feature extraction and separate trust region is able to achieve higher matching accuracy and faster convergence in federated settings against several state-of-the-art centralized graph matching, ~~federated graph learning~~ and federated domain adaption methods.

**Datasets.** We focus on three representative graph learning benchmark datasets: social networks (SNS) (Zhang et al., 2015), protein-protein interaction networks (PPI) (Zitnik & Leskovec, 2017), and DBLP coauthor graphs (DBLP) (DBL). Without loss of generality, we assume that each client contains only one local graph in the federated setting. For the supervised learning methods, the training data ratio over the above three datasets is all fixed to 20%. We train the models on the training set and test them on the test set for three datasets. The detailed descriptions of the federated datasets are presented in Appendix A.5.

**Baselines.** To our best knowledge, this work is the first to offer an unsupervised federated graph matching solution for inferring matched node pairs on different graphs across clients while maintaining the privacy requirement of federated learning, by leveraging the graphlet theory and trust region optimization. Thus, we choose three types of baselines that are most close to the task of federated graph matching: centralized graph matching, ~~federated graph learning~~ and federated domain adaption. We compare the UFGM model with six state-of-the-art centralized graph matching models: **NextAlign** (Zhang et al., 2021c), **NetTrans** (Zhang et al., 2020), **CPUGA** (Pei et al.,

Table 1: Final performance on SNS

Type	Algorithm	<i>Hits</i> @1	<i>Hits</i> @5	<i>Hits</i> @10	<i>Hits</i> @50	<i>Loss</i>
Centralized Graph Matching	NextAlign	<b>0.430</b>	<b>0.512</b>	<b>0.571</b>	<b>0.635</b>	2.149
	NetTrans	0.379	0.439	0.447	0.496	1.611
	CPUGA	0.230	0.238	0.252	0.297	2.551
	ASAR-GM	0.199	0.229	0.252	0.337	1.410
	SIGMA	0.220	0.232	0.253	0.262	1.330
	SeedGNN	0.319	0.340	0.342	0.388	2.919
Federated Domain Adaption	DualAdapt	0.001	0.002	0.002	0.002	2.049
	EFDA	0.001	0.001	0.002	0.002	3.427
	WSDA	0.003	0.005	0.007	0.011	5.129
	FedKA	0.001	0.001	0.010	0.013	3.715
	UFGM	0.371	0.440	0.411	0.459	<b>0.501</b>

Table 2: Final performance on PPI

Type	Algorithm	<i>Hits</i> @1	<i>Hits</i> @5	<i>Hits</i> @10	<i>Hits</i> @50	<i>Loss</i>
Centralized Graph Matching	NextAlign	<b>0.951</b>	<b>0.962</b>	<b>0.972</b>	<b>0.979</b>	2.115
	NetTrans	0.921	0.932	0.958	0.960	1.571
	CPUGA	0.248	0.392	0.433	0.563	2.598
	ASAR-GM	0.299	0.394	0.453	0.668	1.699
	SIGMA	0.499	0.560	0.633	0.782	1.652
	SeedGNN	0.884	0.943	0.959	0.960	3.039
Federated Domain Adaption	DualAdapt	0.006	0.006	0.007	0.011	2.106
	EFDA	0.007	0.011	0.014	0.029	3.249
	WSDA	0.009	0.011	0.013	0.016	2.746
	FKA	0.005	0.006	0.006	0.008	2.227
	UFGM	0.771	0.880	0.902	0.930	<b>0.659</b>

2022), **ASAR-GM** (Ren et al., 2022), **SeedGNN** (Yu et al., 2022), and **SIGMA** (Li et al., 2022), ~~six representative federated graph learning architectures: **FedGraphNN** (He et al., 2021a), **FKGE** (Peng et al., 2021), **SpreadGNN** (He et al., 2022), **SFL** (Chen et al., 2022b), **FederatedScope-GNN** (Wang et al., 2022b), and **FedStar** (Tan et al., 2022);~~ and four recent federated domain adaption methods: **DualAdapt** (Peng et al., 2020), **EFDA** (Kang et al., 2022), **WSDA** (Jiang & Koyejo, 2023), and **FedKA** (Sun et al., 2022). The detailed descriptions of the baselines are presented in Appendix A.5.

**Evaluation metrics.** By following the same settings in two representative graph matching models (Yasar & Çatalyürek, 2018; Fey et al., 2020), We employ a popular measure, *Hits*@*K*, to evaluate and compare our UFGM model to previous lines of work, where *Hits*@*K* measures the proportion of correctly matched nodes ranked in the top-*K* list. A larger *Hits*@*K* value indicates a better graph matching result. We use final *Hits*@*K* to evaluate the quality of the federated federated learning algorithms. In addition, we plot the measure curves regarding *Hits*@*K* and Loss Function Values (*Loss*) with increasing rounds to verify the convergence of different federated learning methods: (Karimireddy et al., 2020; Mitra et al., 2021; Liu et al., 2020; Reddi et al., 2021; Karimireddy et al., 2021; Wang et al., 2021b). A smaller Loss score shows a better federated learning result.

**Final *Hits*@*K* and *Loss* on SNS and PPI.** Tables 1 and 2 show the quality of six centralized graph matching, ~~six federated graph learning~~, and four federated domain adaption algorithms over SNS and PPI respectively. We have observed that our UFGM federated graph matching solution outperforms all the competitors of ~~federated graph learning and~~ federated domain adaption in most experiments. UFGM achieves the highest *Hits*@*K* values (> 0.771 over SNS and > 0.371 on PPI respectively) and the lowest *Loss* values (= 0.659 over SNS and = 0.501 on PPI respectively), which are better than other ~~four~~ baseline methods in all tests. In addition, the *Hits*@*K* scores achieved by UFGM is close or much better than the centralized graph matching method. Compared with the best centralized graph matching method, NextAlign, the *Hits*@1, *Hits*@5, *Hits*@10, and



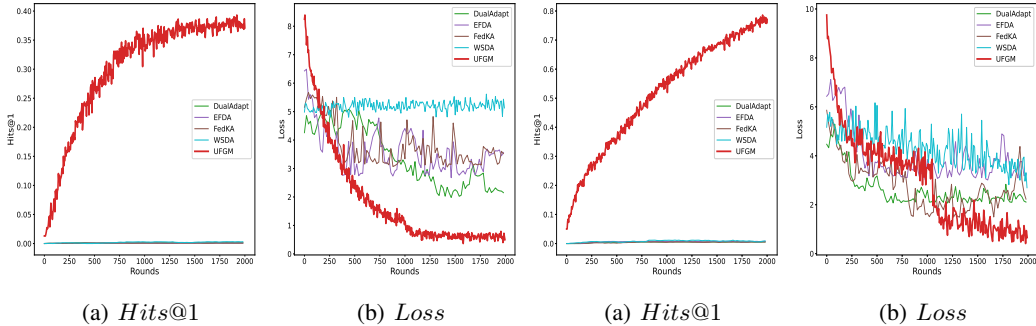


Figure 1: Convergence on SNS

Figure 2: Convergence on PPI

*Hits@50* scores by UFGM are only 15.3% lower respectively. A reasonable explanation is that the combination of graphlet feature extraction, separate trust region, and pseudo supervised learning is able to achieve higher matching accuracy and faster convergence in federated settings. In addition, the promising performance of UFGM over both datasets implies that UFGM has great potential as a general federated graph matching solution over federated datasets, which is desirable in practice.

***Hits@K* Convergence on SNS and PPI.** Figures 1 and 2 exhibit the *Hits@K* curves of **five** federated learning models for graph matching over SNS and PPI respectively. It is obvious that the performance curves by federated learning algorithms initially keep increasing with training rounds and remains relatively stable when the curves are beyond convergence points, i.e., turning points from a sharp *Hits@K* increase to a flat curve. This phenomenon indicates that most federated learning algorithms are able to converge to the invariant solutions after enough training rounds. However, among **six federated graph learning and four** federated domain adaption approaches, our UFGM method can significantly speedup the convergence on two datasets in most experiments, showing the superior performance of UFGM in federated settings. Compared to the learning results by other federated learning models, based on training rounds at convergence points, UFGM, on average, achieves 31.8% and 35.4% convergence improvement on two datasets respectively.

***Loss* Convergence on SNS and PPI.** Figures 1 and 2 also present the *Loss* curves achieved by **five** federated learning models on two datasets respectively. We have observed that the reverse trends, in comparison with the *Hits@K* curves. In most experiments, our UFGM is able to achieve the fastest convergence, especially, UFGM can converge around 1,000 training rounds and then always keep stable on two datasets. A reasonable explanation is that UFGM fully utilizes the proposed graphlet feature extraction techniques to generate the pseudo training data and employ the strength of supervised graph matching for accelerating the training convergence.

## 6 CONCLUSIONS

In this work, we have proposed an unsupervised federated graph matching algorithm. First, an approximate graphlet enumeration method is proposed to capture nodes’ graphlet features to generate pseudo matched node pairs as pseudo training data. Second, a separate trust region algorithm is proposed for pseudo supervised federated graph matching while maintaining the privacy constraints. Finally, empirical evaluation on real datasets demonstrates the superior performance of our UFGM.

## 7 REPRODUCIBILITY STATEMENT

We include the citations and URLs of all datasets used in this work and all codes of third-party baselines in Sections 5 and A.5. Since the datasets used are all public datasets and our methodologies, the experiment environment, the datasets, the training strategies, the baselines, the implementation details, and the hyperparameter settings are explicitly described in Section 3, 4, 5, and A.5, our codes and experiments can be easily reproduced on top of a GPU server. We promise to release our open-source codes on GitHub and maintain a project website with detailed documentation for long-term access by other researchers and end-users after the paper is accepted.

## REFERENCES

- <http://dblp.uni-trier.de/xml/>.
- <https://research.ibm.com/blog/privacy-preserving-federated-learning-finance>.
- <https://new.nsf.gov/news/us-uk-launch-innovation-prize-challenges-privacy>.
- Jinheon Baek, Wonyong Jeong, Jiongdoo Jin, Jaehong Yoon, and Sung Ju Hwang. Personalized subgraph federated learning. *CoRR*, abs/2206.10206, 2022. doi: 10.48550/arXiv.2206.10206. URL <https://doi.org/10.48550/arXiv.2206.10206>.
- Yunsheng Bai, Hao Ding, Ken Gu, Yizhou Sun, and Wei Wang. Learning-based efficient graph similarity computation via multi-scale convolutional set matching. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pp. 3219–3226, 2020.
- Hizir Can Bayram and Islem Rekik. A federated multigraph integration approach for connective brain template learning. In Tanveer F. Syeda-Mahmood, Xiang Li, Anant Madabhushi, Hayit Greenspan, Quanzheng Li, Richard M. Leahy, Bin Dong, and Hongzhi Wang (eds.), *Multimodal Learning for Clinical Decision Support - 11th International Workshop, ML-CDS 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, October 1, 2021, Proceedings*, volume 13050 of *Lecture Notes in Computer Science*, pp. 36–47. Springer, 2021. doi: 10.1007/978-3-030-89847-2\_4. URL [https://doi.org/10.1007/978-3-030-89847-2\\_4](https://doi.org/10.1007/978-3-030-89847-2_4).
- Debora Caldarola, Massimiliano Mancini, Fabio Galasso, Marco Ciccone, Emanuele Rodolà, and Barbara Caputo. Cluster-driven graph federated learning over multiple domains. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2021, virtual, June 19-25, 2021*, pp. 2749–2758. Computer Vision Foundation / IEEE, 2021. doi: 10.1109/CVPRW53098.2021.00309. URL [https://openaccess.thecvf.com/content/CVPR2021W/LLID/html/Caldarola\\_Cluster-Driven\\_Graph\\_Federated\\_Learning\\_Over\\_Multiple\\_Domains\\_CVPRW\\_2021\\_paper.html](https://openaccess.thecvf.com/content/CVPR2021W/LLID/html/Caldarola_Cluster-Driven_Graph_Federated_Learning_Over_Multiple_Domains_CVPRW_2021_paper.html).
- Nicholas Carlini, Florian Tramèr, Eric Wallace, Matthew Jagielski, Ariel Herbert-Voss, Katherine Lee, Adam Roberts, Tom B. Brown, Dawn Song, Úlfar Erlingsson, Alina Oprea, and Colin Raffel. Extracting training data from large language models. In Michael Bailey and Rachel Greenstadt (eds.), *30th USENIX Security Symposium, USENIX Security 2021, August 11-13, 2021*, pp. 2633–2650. USENIX Association, 2021. URL <https://www.usenix.org/conference/usenixsecurity21/presentation/carlini-extracting>.
- Soumen Chakrabarti, Harkanwar Singh, Shubham Lohiya, Prachi Jain, and Mausam. Joint completion and alignment of multilingual knowledge graphs. In Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang (eds.), *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, pp. 11922–11938. Association for Computational Linguistics, 2022. URL <https://aclanthology.org/2022.emnlp-main.817>.
- Chaochao Chen, Jun Zhou, Longfei Zheng, Huiwen Wu, Lingjuan Lyu, Jia Wu, Bingzhe Wu, Ziqi Liu, Li Wang, and Xiaolin Zheng. Vertically federated graph neural network for privacy-preserving node classification. In Luc De Raedt (ed.), *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022*, pp. 1959–1965. ijcai.org, 2022a. doi: 10.24963/ijcai.2022/272. URL <https://doi.org/10.24963/ijcai.2022/272>.
- Chuan Chen, Weibo Hu, Ziyue Xu, and Zibin Zheng. Fedgl: Federated graph learning framework with global self-supervision. *CoRR*, abs/2105.03170, 2021.
- Fengwen Chen, Guodong Long, Zonghan Wu, Tianyi Zhou, and Jing Jiang. Personalized federated learning with a graph. In Luc De Raedt (ed.), *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022*, pp. 2575–2582. ijcai.org, 2022b. doi: 10.24963/ijcai.2022/357. URL <https://doi.org/10.24963/ijcai.2022/357>.

- Hongxu Chen, Hongzhi Yin, Xiangguo Sun, Tong Chen, Bogdan Gabrys, and Katarzyna Musial. Multi-level graph convolutional networks for cross-platform anchor link prediction. In *KDD '20: The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, CA, USA, August 23-27, 2020*, pp. 1503–1511, 2020a.
- Lu Chen, Yanbin Zhao, Boer Lyu, Lesheng Jin, Zhi Chen, Su Zhu, and Kai Yu. Neural graph matching networks for chinese short text matching. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pp. 6152–6158, 2020b.
- Mingyang Chen, Wen Zhang, Zhen Yao, Xiangnan Chen, Mengxiao Ding, Fei Huang, and Hua-jun Chen. Meta-learning based knowledge extrapolation for knowledge graphs in the federated setting. In Luc De Raedt (ed.), *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022*, pp. 1966–1972. ijcai.org, 2022c. doi: 10.24963/ijcai.2022/273. URL <https://doi.org/10.24963/ijcai.2022/273>.
- Xiyuan Chen, Mark Heimann, Fatemeh Vahedian, and Danai Koutra. Cone-align: Consistent network alignment with proximity-preserving node embedding. In *CIKM '20: The 29th ACM International Conference on Information and Knowledge Management, Virtual Event, Ireland, October 19-23, 2020*, pp. 1985–1988, 2020c.
- Xiaokai Chu, Xinxin Fan, Di Yao, Zhihua Zhu, Jianhui Huang, and Jingping Bi. Cross-network embedding for multi-network alignment. In *The World Wide Web Conference, WWW 2019, San Francisco, CA, USA, May 13-17, 2019*, pp. 273–284, 2019.
- Xingbo Du, Junchi Yan, and Hongyuan Zha. Joint link prediction and network alignment via cross-graph embedding. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*, pp. 2251–2257, 2019.
- Zhou Fan, Cheng Mao, Yihong Wu, and Jiaming Xu. Spectral graph matching and regularized quadratic relaxations: Algorithm and theory. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*, pp. 2985–2995, 2020.
- Mahboubeh Farid, Wah June Leong, and Malik Abu Hassan. A new two-step gradient-type method for large-scale unconstrained optimization. *Comput. Math. Appl.*, 59(10):3301–3307, 2010.
- Jie Feng, Mingyang Zhang, Huandong Wang, Zeyu Yang, Chao Zhang, Yong Li, and Depeng Jin. Dplink: User identity linkage via deep neural network from heterogeneous mobility data. In *The World Wide Web Conference, WWW 2019, San Francisco, CA, USA, May 13-17, 2019*, pp. 459–469, 2019.
- Matthias Fey, Jan Eric Lenssen, Christopher Morris, Jonathan Masci, and Nils M. Kriege. Deep graph matching consensus. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*, 2020.
- Quankai Gao, Fudong Wang, Nan Xue, Jin-Gang Yu, and Gui-Song Xia. Deep graph matching under quadratic constraint. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*, pp. 5069–5078. Computer Vision Foundation / IEEE, 2021. doi: 10.1109/CVPR46437.2021.00503. URL [https://openaccess.thecvf.com/content/CVPR2021/html/Gao\\_Deep\\_Graph\\_Matching\\_Under\\_Quadratic\\_Constraint\\_CVPR\\_2021\\_paper.html](https://openaccess.thecvf.com/content/CVPR2021/html/Gao_Deep_Graph_Matching_Under_Quadratic_Constraint_CVPR_2021_paper.html).
- Lingbing Guo, Qiang Zhang, Zequn Sun, Mingyang Chen, Wei Hu, and Hua-jun Chen. Understanding and improving knowledge graph embedding for entity alignment. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvári, Gang Niu, and Sivan Sabato (eds.), *International Conference on Machine Learning, ICML 2022, 17-23 July 2022, Baltimore, Maryland, USA*, volume 162 of *Proceedings of Machine Learning Research*, pp. 8145–8156. PMLR, 2022. URL <https://proceedings.mlr.press/v162/guo22i.html>.
- Pietro Hiram Guzzi and Tijana Milenkovic. Survey of local and global biological network alignment: the need to reconcile the two sides of the same coin. *Briefings in Bioinformatics*, 19(3):472–481, 2018.

- Stefan Haller, Lorenz Feineis, Lisa Hutschenreiter, Florian Bernard, Carsten Rother, Dagmar Kainmüller, Paul Swoboda, and Bogdan Savchynskyy. A comparative study of graph matching algorithms in computer vision. In Shai Avidan, Gabriel J. Brostow, Moustapha Cissé, Giovanni Maria Farinella, and Tal Hassner (eds.), *Computer Vision - ECCV 2022 - 17th European Conference, Tel Aviv, Israel, October 23-27, 2022, Proceedings, Part XXIII*, volume 13683 of *Lecture Notes in Computer Science*, pp. 636–653. Springer, 2022. doi: 10.1007/978-3-031-20050-2\\_37. URL [https://doi.org/10.1007/978-3-031-20050-2\\_37](https://doi.org/10.1007/978-3-031-20050-2_37).
- Chaoyang He, Keshav Balasubramanian, Emir Ceyani, Yu Rong, Peilin Zhao, Junzhou Huang, Murali Annavaram, and Salman Avestimehr. Fedgraphnn: A federated learning system and benchmark for graph neural networks. *CoRR*, abs/2104.07145, 2021a. URL <https://arxiv.org/abs/2104.07145>.
- Chaoyang He, Keshav Balasubramanian, Emir Ceyani, Yu Rong, Peilin Zhao, Junzhou Huang, Murali Annavaram, and Salman Avestimehr. Fedgraphnn: A federated learning system and benchmark for graph neural networks. *CoRR*, abs/2104.07145, 2021b.
- Chaoyang He, Emir Ceyani, Keshav Balasubramanian, Murali Annavaram, and Salman Avestimehr. Spreadgnn: Serverless multi-task federated learning for graph neural networks. *CoRR*, abs/2106.02743, 2021c.
- Chaoyang He, Emir Ceyani, Keshav Balasubramanian, Murali Annavaram, and Salman Avestimehr. Spreadgnn: Decentralized multi-task federated learning for graph neural networks on molecular data. In *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*, pp. 6865–6873. AAAI Press, 2022. URL <https://ojs.aaai.org/index.php/AAAI/article/view/20643>.
- Jiawei He, Zehao Huang, Naiyan Wang, and Zhaoxiang Zhang. Learnable graph matching: Incorporating graph partitioning with deep feature learning for multiple object tracking. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*, pp. 5299–5309. Computer Vision Foundation / IEEE, 2021d. doi: 10.1109/CVPR46437.2021.00526. URL [https://openaccess.thecvf.com/content/CVPR2021/html/He\\_Learnable\\_Graph\\_Matching\\_Incorporating\\_Graph\\_Partitioning\\_With\\_Deep\\_Feature\\_Learning\\_CVPR\\_2021\\_paper.html](https://openaccess.thecvf.com/content/CVPR2021/html/He_Learnable_Graph_Matching_Incorporating_Graph_Partitioning_With_Deep_Feature_Learning_CVPR_2021_paper.html).
- Mark Heimann, Haoming Shen, Tara Safavi, and Danai Koutra. REGAL: representation learning-based graph alignment. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management, CIKM 2018, Torino, Italy, October 22-26, 2018*, pp. 117–126, 2018.
- Yuke Hu, Wei Liang, Ruofan Wu, Kai Xiao, Weiqiang Wang, Xiaochen Li, Jinfei Liu, and Zhan Qin. Quantifying and defending against privacy threats on federated knowledge graph embedding. In Ying Ding, Jie Tang, Juan F. Sequeda, Lora Aroyo, Carlos Castillo, and Geert-Jan Houben (eds.), *Proceedings of the ACM Web Conference 2023, WWW 2023, Austin, TX, USA, 30 April 2023 - 4 May 2023*, pp. 2306–2317. ACM, 2023. doi: 10.1145/3543507.3583450. URL <https://doi.org/10.1145/3543507.3583450>.
- Lisa Hutschenreiter, Stefan Haller, Lorenz Feineis, Carsten Rother, Dagmar Kainmüller, and Bogdan Savchynskyy. Fusion moves for graph matching. In *2021 IEEE/CVF International Conference on Computer Vision, ICCV 2021, Montreal, QC, Canada, October 10-17, 2021*, pp. 6250–6259. IEEE, 2021. doi: 10.1109/ICCV48922.2021.00621. URL <https://doi.org/10.1109/ICCV48922.2021.00621>.
- Thanh Trung Huynh, Nguyen Thanh Toan, Van Vinh Tong, Thanh Dat Hoang, Duong Chi Thang, Nguyen Quoc Viet Hung, and Abdul Sattar. A comparative study on network alignment techniques. *Expert Syst. Appl.*, 140, 2020a.
- Thanh Trung Huynh, Van Vinh Tong, Thanh Tam Nguyen, Hongzhi Yin, Matthias Weidlich, and Nguyen Quoc Viet Hung. Adaptive network alignment with unsupervised and multi-order convolutional networks. In *36th IEEE International Conference on Data Engineering, ICDE 2020, Dallas, TX, USA, April 20-24, 2020*, pp. 85–96, 2020b.

- Jr. J. E. Dennis and H. Wolkowicz. Sizing and least-change secant methods. *SIAM Journal on Numerical Analysis*, 30(5):1291–1314, 1993.
- Enyi Jiang and Oluwasanmi O Koyejo. Weakly-supervised domain adaptation in federated learning. In *ICLR 2023 Submission*, 2023.
- Meng Jiang, Taeho Jung, Ryan Karl, and Tong Zhao. Federated dynamic GNN with secure aggregation. *CoRR*, abs/2009.07351, 2020.
- Zheheng Jiang, Hossein Rahmani, Plamen P. Angelov, Sue Black, and Bryan M. Williams. Graph-context attention networks for size-varied deep graph matching. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022, New Orleans, LA, USA, June 18-24, 2022*, pp. 2333–2342. IEEE, 2022. doi: 10.1109/CVPR52688.2022.00238. URL <https://doi.org/10.1109/CVPR52688.2022.00238>.
- Di Jin, Luzhi Wang, Yizhen Zheng, Xiang Li, Fei Jiang, Wei Lin, and Shirui Pan. CGMN: A contrastive graph matching network for self-supervised graph similarity learning. In Luc De Raedt (ed.), *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022*, pp. 2101–2107. ijcai.org, 2022. doi: 10.24963/ijcai.2022/292. URL <https://doi.org/10.24963/ijcai.2022/292>.
- Jiashun Jin, Zheng Tracy Ke, and Shengming Luo. Network global testing by counting graphlets. In Jennifer G. Dy and Andreas Krause (eds.), *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018*, volume 80 of *Proceedings of Machine Learning Research*, pp. 2338–2346. PMLR, 2018.
- Hua Kang, Zhiyang Li, and Qian Zhang. Communicational and computational efficient federated domain adaptation. *IEEE Trans. Parallel Distributed Syst.*, 33(12):3678–3689, 2022. doi: 10.1109/TPDS.2022.3167457. URL <https://doi.org/10.1109/TPDS.2022.3167457>.
- Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, Sashank J. Reddi, Sebastian U. Stich, and Ananda Theertha Suresh. SCAFFOLD: stochastic controlled averaging for federated learning. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*, volume 119 of *Proceedings of Machine Learning Research*, pp. 5132–5143. PMLR, 2020.
- Sai Praneeth Karimireddy, Martin Jaggi, Satyen Kale, Mehryar Mohri, Sashank J. Reddi, Sebastian U. Stich, and Ananda Theertha Suresh. Mime: Mimicking centralized stochastic algorithms in federated learning. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*, 2021.
- Chuyang Ke and Jean Honorio. Federated myopic community detection with one-shot communication. *CoRR*, abs/2106.07255, 2021.
- Brian P. Kelley, Roded Sharan, Richard M. Karp, Taylor Sittler, David E. Root, Brent R. Stockwell, and Trey Ideker. Conserved pathways within bacteria and yeast as revealed by global protein network alignment. *Proceedings of the National Academy of Sciences*, 100(20):11394–11399, 2003.
- Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*, 2017.
- Risi Kondor, Nino Shervashidze, and Karsten M. Borgwardt. The graphlet spectrum. In Andrea Pohoreckýj Danyluk, Léon Bottou, and Michael L. Littman (eds.), *Proceedings of the 26th Annual International Conference on Machine Learning, ICML 2009, Montreal, Quebec, Canada, June 14-18, 2009*, volume 382 of *ACM International Conference Proceeding Series*, pp. 529–536. ACM, 2009.
- Xiangnan Kong, Jiawei Zhang, and Philip S. Yu. Inferring anchor links across multiple heterogeneous social networks. In *22nd ACM International Conference on Information and Knowledge Management, CIKM’13, San Francisco, CA, USA, October 27 - November 1, 2013*, pp. 179–188, 2013.

- Anusha Lalitha, Osman Cihan Kilinc, Tara Javidi, and Farinaz Koushanfar. Peer-to-peer federated learning on graphs. *CoRR*, abs/1901.11173, 2019. URL <http://arxiv.org/abs/1901.11173>.
- Chaozhuo Li, Senzhang Wang, Philip S. Yu, Lei Zheng, Xiaoming Zhang, Zhoujun Li, and Yanbo Liang. Distribution distance minimization for unsupervised user identity linkage. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management, CIKM 2018, Torino, Italy, October 22-26, 2018*, pp. 447–456, 2018.
- Chaozhuo Li, Senzhang Wang, Hao Wang, Yanbo Liang, Philip S. Yu, Zhoujun Li, and Wei Wang. Partially shared adversarial learning for semi-supervised multi-platform user identity linkage. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management, CIKM 2019, Beijing, China, November 3-7, 2019*, pp. 249–258, 2019a.
- Chaozhuo Li, Senzhang Wang, Yukun Wang, Philip S. Yu, Yanbo Liang, Yun Liu, and Zhoujun Li. Adversarial learning for weakly-supervised social network alignment. In *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, Honolulu, Hawaii, USA, January 27 - February 1, 2019*, pp. 996–1003, 2019b.
- Wuyang Li, Xinyu Liu, and Yixuan Yuan. SIGMA: semantic-complete graph matching for domain adaptive object detection. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022, New Orleans, LA, USA, June 18-24, 2022*, pp. 5281–5290. IEEE, 2022. doi: 10.1109/CVPR52688.2022.00522. URL <https://doi.org/10.1109/CVPR52688.2022.00522>.
- Xunkai Li, Wentao Zhang, Rong-Hua Li, Yulin Zhao, Yinlin Zhu, and Guoren Wang. A new paradigm for federated structure non-iid subgraph learning. In *ICLR 2023 Submission*, 2023.
- Chang Liu, Shaofeng Zhang, Xiaokang Yang, and Junchi Yan. Self-supervised learning of visual graph matching. In Shai Avidan, Gabriel J. Brostow, Moustapha Cissé, Giovanni Maria Farinella, and Tal Hassner (eds.), *Computer Vision - ECCV 2022 - 17th European Conference, Tel Aviv, Israel, October 23-27, 2022, Proceedings, Part XXIII*, volume 13683 of *Lecture Notes in Computer Science*, pp. 370–388. Springer, 2022a. doi: 10.1007/978-3-031-20050-2\_22. URL [https://doi.org/10.1007/978-3-031-20050-2\\_22](https://doi.org/10.1007/978-3-031-20050-2_22).
- Li Liu, William K. Cheung, Xin Li, and Lejian Liao. Aligning users across social networks using network embedding. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9-15 July 2016*, pp. 1774–1780, 2016.
- Linfeng Liu, Michael C. Hughes, Soha Hassoun, and Liping Liu. Stochastic iterative graph matching. In Marina Meila and Tong Zhang (eds.), *Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18-24 July 2021, Virtual Event*, volume 139 of *Proceedings of Machine Learning Research*, pp. 6815–6825. PMLR, 2021. URL <http://proceedings.mlr.press/v139/liu21i.html>.
- Siyuan Liu, Shuhui Wang, Feida Zhu, Jinbo Zhang, and Ramayya Krishnan. HYDRA: large-scale social identity linkage via heterogeneous behavior modeling. In *International Conference on Management of Data, SIGMOD 2014, Snowbird, UT, USA, June 22-27, 2014*, pp. 51–62, 2014.
- Wei Liu, Li Chen, Yunfei Chen, and Wenyi Zhang. Accelerating federated learning via momentum gradient descent. *IEEE Trans. Parallel Distributed Syst.*, 31(8):1754–1766, 2020.
- Weijie Liu, Hui Qian, Chao Zhang, Jiahao Xie, Zebang Shen, and Nenggan Zheng. From one to all: Learning to match heterogeneous and partially overlapped graphs. In *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*, pp. 4109–4119. AAAI Press, 2022b. URL <https://ojs.aaai.org/index.php/AAAI/article/view/20329>.

- Xiao Liu, Haoyun Hong, Xinghao Wang, Zeyi Chen, Evgeny Kharlamov, Yuxiao Dong, and Jie Tang. Selfkg: Self-supervised entity alignment in knowledge graphs. In Frédérique Laforest, Raphaël Troncy, Elena Simperl, Deepak Agarwal, Aristides Gionis, Ivan Herman, and Lionel Médini (eds.), *WWW '22: The ACM Web Conference 2022, Virtual Event, Lyon, France, April 25 - 29, 2022*, pp. 860–870. ACM, 2022c. doi: 10.1145/3485447.3511945. URL <https://doi.org/10.1145/3485447.3511945>.
- Yangwei Liu, Hu Ding, Danyang Chen, and Jinhui Xu. Novel geometric approach for global alignment of PPI networks. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, AAAI 2017, February 4-9, 2017, San Francisco, California, USA.*, pp. 31–37, 2017.
- Gengyu Lyu, Yanan Wu, and Songhe Feng. Deep graph matching for partial label learning. In Luc De Raedt (ed.), *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022*, pp. 3306–3312. ijcai.org, 2022. doi: 10.24963/ijcai.2022/459. URL <https://doi.org/10.24963/ijcai.2022/459>.
- Eric Malmi, Aristides Gionis, and Evimaria Terzi. Active network alignment: A matching-based approach. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, CIKM 2017, Singapore, November 06 - 10, 2017*, pp. 1687–1696, 2017.
- Tong Man, Huawei Shen, Shenghua Liu, Xiaolong Jin, and Xueqi Cheng. Predict anchor links across social networks via an embedding approach. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9-15 July 2016*, pp. 1823–1829, 2016.
- Guangxu Mei, Ziyu Guo, Shijun Liu, and Li Pan. SGNN: A graph neural network based federated learning approach by hiding structure. In *2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, December 9-12, 2019*, pp. 2560–2568, 2019.
- Chuzheng Meng, Sirisha Rambhatla, and Yan Liu. Cross-node federated graph neural network for spatio-temporal data modeling. In Feida Zhu, Beng Chin Ooi, and Chunyan Miao (eds.), *KDD '21: The 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, Singapore, August 14-18, 2021*, pp. 1202–1211. ACM, 2021. doi: 10.1145/3447548.3467371. URL <https://doi.org/10.1145/3447548.3467371>.
- Lin Meng, Yuxiang Ren, Jiawei Zhang, Fanghua Ye, and Philip S. Yu. Deep heterogeneous social network alignment. In *2019 IEEE First International Conference on Cognitive Machine Intelligence (CogMI), Los Angeles, CA, USA, December 12-14, 2019*, pp. 43–52, 2019.
- Aritra Mitra, Rayana Jaafar, George Pappas, and Hamed Hassani. Linear convergence in federated learning: Tackling client heterogeneity and sparse gradients. In *The 35th Conference on Neural Information Processing Systems, (NeurIPS'21)*, Online, December 6-14 2021.
- Xin Mu, Feida Zhu, Ee-Peng Lim, Jing Xiao, Jianzong Wang, and Zhi-Hua Zhou. User identity linkage by latent user space modelling. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016*, pp. 1775–1784, 2016.
- Huda Nassar, Nate Veldt, Shahin Mohammadi, Ananth Grama, and David F. Gleich. Low rank spectral network alignment. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web, WWW 2018, Lyon, France, April 23-27, 2018*, pp. 619–628, 2018.
- Xiang Ni, Xiaolong Xu, Lingjuan Lyu, Changhua Meng, and Weiqiang Wang. A vertical federated learning framework for graph convolutional network. *CoRR*, abs/2106.11593, 2021. URL <https://arxiv.org/abs/2106.11593>.
- NSF. U.s. and u.k. launch innovation prize challenges in privacy-enhancing technologies to tackle financial crime and public health emergencies. <https://new.nsf.gov/news/us-uk-launch-innovation-prize-challenges-privacy,2022>.
- Qiyang Pan and Yifei Zhu. Fedwalk: Communication efficient federated unsupervised node embedding with differential privacy. In Aidong Zhang and Huzefa Rangwala (eds.), *KDD '22: The 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 14 - 18, 2022*, pp. 1317–1326. ACM, 2022.

- Qiyang Pan, Yifei Zhu, and Lingyang Chu. Lumos: Heterogeneity-aware federated graph learning over decentralized devices. In *39th IEEE International Conference on Data Engineering, ICDE 2023, Anaheim, CA, USA, April 3-7, 2023*, pp. 1914–1926. IEEE, 2023.
- Jin-Duk Park, Cong Tran, Won-Yong Shin, and Xin Cao. Gradalign+: Empowering gradual network alignment using attribute augmentation. In Mohammad Al Hasan and Li Xiong (eds.), *Proceedings of the 31st ACM International Conference on Information & Knowledge Management, Atlanta, GA, USA, October 17-21, 2022*, pp. 4374–4378. ACM, 2022. doi: 10.1145/3511808.3557605. URL <https://doi.org/10.1145/3511808.3557605>.
- Shichao Pei, Lu Yu, Guoxian Yu, and Xiangliang Zhang. Graph alignment with noisy supervision. In Frédérique Laforest, Raphaël Troncy, Elena Simperl, Deepak Agarwal, Aristides Gionis, Ivan Herman, and Lionel Médini (eds.), *WWW '22: The ACM Web Conference 2022, Virtual Event, Lyon, France, April 25 - 29, 2022*, pp. 1104–1114. ACM, 2022. doi: 10.1145/3485447.3512089. URL <https://doi.org/10.1145/3485447.3512089>.
- Hao Peng, Haoran Li, Yangqiu Song, Vincent W. Zheng, and Jianxin Li. Differentially private federated knowledge graphs embedding. In Gianluca Demartini, Guido Zuccon, J. Shane Culpepper, Zi Huang, and Hanghang Tong (eds.), *CIKM '21: The 30th ACM International Conference on Information and Knowledge Management, Virtual Event, Queensland, Australia, November 1 - 5, 2021*, pp. 1416–1425. ACM, 2021. doi: 10.1145/3459637.3482252. URL <https://doi.org/10.1145/3459637.3482252>.
- Xingchao Peng, Zijun Huang, Yizhe Zhu, and Kate Saenko. Federated adversarial domain adaptation. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net, 2020. URL <https://openreview.net/forum?id=HJezF3VYPB>.
- Yuanyuan Qiao, Yuewei Wu, Fan Duo, Wenhui Lin, and Jie Yang. Siamese neural networks for user identity linkage through web browsing. *IEEE Trans. Neural Networks Learn. Syst.*, 31(8): 2741–2751, 2020.
- Kyle Kai Qin, Flora D. Salim, Yongli Ren, Wei Shao, Mark Heimann, and Danai Koutra. G-CREWE: graph compression with embedding for network alignment. In *CIKM '20: The 29th ACM International Conference on Information and Knowledge Management, Virtual Event, Ireland, October 19-23, 2020*, pp. 1255–1264, 2020.
- Jingwei Qu, Haibin Ling, Chenrui Zhang, Xiaoqing Lyu, and Zhi Tang. Adaptive edge attention for graph matching with outliers. In Zhi-Hua Zhou (ed.), *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI 2021, Virtual Event / Montreal, Canada, 19-27 August 2021*, pp. 966–972. ijcai.org, 2021. doi: 10.24963/ijcai.2021/134. URL <https://doi.org/10.24963/ijcai.2021/134>.
- Liang Qu, Ningzhi Tang, Ruiqi Zheng, Quoc Viet Hung Nguyen, Zi Huang, Yuhui Shi, and Hongzhi Yin. Semi-decentralized federated ego graph learning for recommendation. In Ying Ding, Jie Tang, Juan F. Sequeda, Lora Aroyo, Carlos Castillo, and Geert-Jan Houben (eds.), *Proceedings of the ACM Web Conference 2023, WWW 2023, Austin, TX, USA, 30 April 2023 - 4 May 2023*, pp. 339–348. ACM, 2023. doi: 10.1145/3543507.3583337. URL <https://doi.org/10.1145/3543507.3583337>.
- Miklós Z. Rácz and Anirudh Sridhar. Correlated stochastic block models: Exact graph matching with applications to recovering communities. In Marc’Aurelio Ranzato, Alina Beygelzimer, Yann N. Dauphin, Percy Liang, and Jennifer Wortman Vaughan (eds.), *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual*, pp. 22259–22273, 2021. URL <https://proceedings.neurips.cc/paper/2021/hash/baf4f1a5938b8d520b328c13b51ccf11-Abstract.html>.
- Dana Randall. Efficient generation of random nonsingular matrices. *Random Struct. Algorithms*, 4(1):111–118, 1993.



- Sashank J. Reddi, Zachary Charles, Manzil Zaheer, Zachary Garrett, Keith Rush, Jakub Konečný, Sanjiv Kumar, and Hugh Brendan McMahan. Adaptive federated optimization. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021, 2021*.
- Fuxin Ren, Zhongbao Zhang, Jiawei Zhang, Sen Su, Li Sun, Guozhen Zhu, and Congying Guo. BANANA: when behavior analysis meets social network alignment. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020*, pp. 1438–1444, 2020.
- Qibing Ren, Qingquan Bao, Runzhong Wang, and Junchi Yan. Appearance and structure aware robust deep visual graph matching: Attack, defense and beyond. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022, New Orleans, LA, USA, June 18-24, 2022*, pp. 15242–15251. IEEE, 2022. doi: 10.1109/CVPR52688.2022.01483. URL <https://doi.org/10.1109/CVPR52688.2022.01483>.
- Yuxiang Ren, Charu C. Aggarwal, and Jiawei Zhang. Meta diagram based active social networks alignment. In *35th IEEE International Conference on Data Engineering, ICDE 2019, Macao, China, April 8-11, 2019*, pp. 1690–1693, 2019.
- Elsa Rizk and Ali H. Sayed. A graph federated architecture with privacy preserving learning. In *22nd IEEE International Workshop on Signal Processing Advances in Wireless Communications, SPAWC 2021, Lucca, Italy, September 27-30, 2021*, pp. 131–135. IEEE, 2021. doi: 10.1109/SPAWC51858.2021.9593148. URL <https://doi.org/10.1109/SPAWC51858.2021.9593148>.
- Donald Shenaj, Eros Fanì, Marco Toldo, Debora Caldarola, Antonio Tavera, Umberto Michieli, Marco Ciccone, Pietro Zanuttigh, and Barbara Caputo. Learning across domains and devices: Style-driven source-free domain adaptation in clustered federated learning. In *IEEE/CVF Winter Conference on Applications of Computer Vision, WACV 2023, Waikoloa, HI, USA, January 2-7, 2023*, pp. 444–454. IEEE, 2023.
- Nino Shervashidze, S. V. N. Vishwanathan, Tobias Petri, Kurt Mehlhorn, and Karsten M. Borgwardt. Efficient graphlet kernels for large graph comparison. In David A. Van Dyk and Max Welling (eds.), *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics, AISTATS 2009, Clearwater Beach, Florida, USA, April 16-18, 2009*, volume 5 of *JMLR Proceedings*, pp. 488–495. JMLR.org, 2009.
- Kai Shu, Suhang Wang, Jiliang Tang, Reza Zafarani, and Huan Liu. User identity linkage across online social networks: A review. *SIGKDD Explorations*, 18(2):5–17, 2016.
- Rohit Singh, Jinbo Xu, and Bonnie Berger. Global alignment of multiple protein interaction networks with application to functional orthology detection. *Proceedings of the National Academy of Sciences*, 105(35):12763–12768, 2008. ISSN 0027-8424.
- Hossein Azari Soufiani and Edoardo M. Airoldi. Graphlet decomposition of a weighted network. In Neil D. Lawrence and Mark A. Girolami (eds.), *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics, AISTATS 2012, La Palma, Canary Islands, Spain, April 21-23, 2012*, volume 22 of *JMLR Proceedings*, pp. 54–63. JMLR.org, 2012.
- Yuwei Sun, Ng S. T. Chong, and Hideya Ochiai. Multi-source domain adaptation based on federated knowledge alignment. *CoRR*, abs/2203.11635, 2022. doi: 10.48550/arXiv.2203.11635. URL <https://doi.org/10.48550/arXiv.2203.11635>.
- Zequn Sun, Chengming Wang, Wei Hu, Muhao Chen, Jian Dai, Wei Zhang, and Yuzhong Qu. Knowledge graph alignment network with gated multi-hop neighborhood aggregation. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pp. 222–229, 2020.
- Toyotaro Suzumura, Yi Zhou, Nathalie Barcardo, Guangnan Ye, Keith Houck, Ryo Kawahara, Ali Anwar, Lucia Larise Stavarache, Daniel Klyashorny, Heiko Ludwig, and Kumar Bhaskaran. Towards federated graph learning for collaborative financial crimes detection. *CoRR*, abs/1909.12946, 2019.

- Hao-Ru Tan, Chuang Wang, Si-Tong Wu, Tie-Qiang Wang, Xu-Yao Zhang, and Cheng-Lin Liu. Proxy graph matching with proximal matching networks. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, pp. 9808–9815. AAAI Press, 2021. URL <https://ojs.aaai.org/index.php/AAAI/article/view/17179>.
- Yue Tan, Yixin Liu, Guodong Long, Jing Jiang, Qinghua Lu, and Chengqi Zhang. Federated learning on non-iid graphs via structural knowledge sharing. *CoRR*, abs/2211.13009, 2022. doi: 10.48550/arXiv.2211.13009. URL <https://doi.org/10.48550/arXiv.2211.13009>.
- Zhihua Tian, Yuan Ding, Rui Zhang, Jian Liu, and Kui Ren. Towards federated learning of deep graph neural networks. In *ICLR 2023 Submission*, 2023.
- Kun Tu, Jian Li, Don Towsley, Dave Braines, and Liam D. Turner. gl2vec: learning feature representation using graphlets for directed networks. In Francesca Spezzano, Wei Chen, and Xiaokui Xiao (eds.), *ASONAM '19: International Conference on Advances in Social Networks Analysis and Mining, Vancouver, British Columbia, Canada, 27-30 August, 2019*, pp. 216–221. ACM, 2019.
- Vipin Vijayan and Tijana Milenkovic. Multiple network alignment via multimagna++. *IEEE/ACM Trans. Comput. Biology Bioinform.*, 15(5):1669–1682, 2018.
- Vipin Vijayan, Shawn Gu, Eric T. Krebs, Lei Meng, and Tijana Milenkovic. Pairwise versus multiple global network alignment. *IEEE Access*, 8:41961–41974, 2020.
- Binghui Wang, Ang Li, Hai Li, and Yiran Chen. Graphfl: A federated learning framework for semi-supervised node classification on graphs. *CoRR*, abs/2012.04187, 2020a.
- Binghui Wang, Ang Li, Meng Pang, Hai Li, and Yiran Chen. Graphfl: A federated learning framework for semi-supervised node classification on graphs. In Xingquan Zhu, Sanjay Ranka, My T. Thai, Takashi Washio, and Xindong Wu (eds.), *IEEE International Conference on Data Mining, ICDM 2022, Orlando, FL, USA, November 28 - Dec. 1, 2022*, pp. 498–507. IEEE, 2022a. doi: 10.1109/ICDM54844.2022.00060. URL <https://doi.org/10.1109/ICDM54844.2022.00060>.
- Chenxu Wang, Yang Wang, Zhiyuan Zhao, Dong Qin, Xiapu Luo, and Tao Qin. Credible seed identification for large-scale structural network alignment. *Data Min. Knowl. Discov.*, 34(6): 1744–1776, 2020b.
- Chunnan Wang, Bozhou Chen, Geng Li, and Hongzhi Wang. FL-AGCNS: federated learning framework for automatic graph convolutional network search. *CoRR*, abs/2104.04141, 2021a.
- Daixin Wang, Yuan Qi, Jianbin Lin, Peng Cui, Quanhui Jia, Zhen Wang, Yanming Fang, Quan Yu, Jun Zhou, and Shuang Yang. A semi-supervised graph attentive network for financial fraud detection. In Jianyong Wang, Kyuseok Shim, and Xindong Wu (eds.), *2019 IEEE International Conference on Data Mining, ICDM 2019, Beijing, China, November 8-11, 2019*, pp. 598–607. IEEE, 2019a.
- Fudong Wang, Nan Xue, Jin-Gang Yu, and Gui-Song Xia. Zero-assignment constraint for graph matching with outliers. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020*, pp. 3030–3039, 2020c.
- Jianyu Wang, Zheng Xu, Zachary Garrett, Zachary Charles, Luyang Liu, and Gauri Joshi. Local adaptivity in federated learning: Convergence and consistency. In *The International Workshop on Federated Learning for User Privacy and Data Confidentiality in Conjunction with ICML 2021, (FL-ICML'21)*, Online, December 6-14 2021b.
- Runzhong Wang, Junchi Yan, and Xiaokang Yang. Graduated assignment for joint multi-graph matching and clustering with application to unsupervised graph matching network learning. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020d.

- Runzhong Wang, Junchi Yan, and Xiaokang Yang. Unsupervised learning of graph matching with mixture of modes via discrepancy minimization. *IEEE Trans. Pattern Anal. Mach. Intell.*, 45(8): 10500–10518, 2023.
- Shaokai Wang, Xutao Li, Yunming Ye, Shanshan Feng, Raymond Y. K. Lau, Xiaohui Huang, and Xiaolin Du. Anchor link prediction across attributed networks via network embedding. *Entropy*, 21(3):254, 2019b.
- Tao Wang, He Liu, Yidong Li, Yi Jin, Xiaohui Hou, and Haibin Ling. Learning combinatorial solver for graph matching. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020*, pp. 7565–7574, 2020e.
- Tianzhe Wang, Zetian Jiang, and Junchi Yan. Multiple graph matching and clustering via decayed pairwise matching composition. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pp. 1660–1667, 2020f.
- Yaqing Wang, Chunyan Feng, Ling Chen, Hongzhi Yin, Caili Guo, and Yunfei Chu. User identity linkage across social networks via linked heterogeneous network embedding. *World Wide Web*, 22(6):2611–2632, 2019c.
- Zhen Wang, Weirui Kuang, Yuexiang Xie, Liuyi Yao, Yaliang Li, Bolin Ding, and Jingren Zhou. Federatedscope-gnn: Towards a unified, comprehensive and efficient package for federated graph learning. In Aidong Zhang and Huzefa Rangwala (eds.), *KDD '22: The 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 14 - 18, 2022*, pp. 4110–4120. ACM, 2022b. doi: 10.1145/3534678.3539112. URL <https://doi.org/10.1145/3534678.3539112>.
- Deborah A. Weighill, Marouen Ben Guebila, Camila M. Lopes-Ramos, Kimberly Glass, John Quackenbush, John Platig, and Rebekka Burkholz. Gene regulatory network inference as relaxed graph matching. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, pp. 10263–10272. AAAI Press, 2021. URL <https://ojs.aaai.org/index.php/AAAI/article/view/17230>.
- Chuhan Wu, Fangzhao Wu, Yang Cao, Yongfeng Huang, and Xing Xie. Fedgmn: Federated graph neural network for privacy-preserving recommendation. In *The International Workshop on Federated Learning for User Privacy and Data Confidentiality in Conjunction with ICML 2021, (FL-ICML'21)*, Online, December 6-14 2021.
- Yuting Wu, Xiao Liu, Yansong Feng, Zheng Wang, and Dongyan Zhao. Neighborhood matching network for entity alignment. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pp. 6477–6487, 2020.
- Ya xiang Yuan. A modified bfgs algorithm for unconstrained optimization. *IMA Journal of Numerical Analysis*, 11(3):325–332, 1991.
- Han Xie, Jing Ma, Li Xiong, and Carl Yang. Federated graph classification over non-iid graphs. In Marc’Aurelio Ranzato, Alina Beygelzimer, Yann N. Dauphin, Percy Liang, and Jennifer Wortman Vaughan (eds.), *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual*, pp. 18839–18852, 2021. URL <https://proceedings.neurips.cc/paper/2021/hash/9c6947bd95ae487c81d4e19d3ed8cd6f-Abstract.html>.
- Han Xie, Li Xiong, and Carl Yang. Federated node classification over graphs with latent link-type heterogeneity. In Ying Ding, Jie Tang, Juan F. Sequeda, Lora Aroyo, Carlos Castillo, and Geert-Jan Houben (eds.), *Proceedings of the ACM Web Conference 2023, WWW 2023, Austin, TX, USA, 30 April 2023 - 4 May 2023*, pp. 556–566. ACM, 2023. doi: 10.1145/3543507.3583471. URL <https://doi.org/10.1145/3543507.3583471>.

- Wei Xie, Xin Mu, Roy Ka-Wei Lee, Feida Zhu, and Ee-Peng Lim. Unsupervised user identity linkage via factoid embedding. In *IEEE International Conference on Data Mining, ICDM 2018, Singapore, November 17-20, 2018*, pp. 1338–1343, 2018.
- Kexuan Xin, Zequn Sun, Wen Hua, Wei Hu, Jianfeng Qu, and Xiaofang Zhou. Large-scale entity alignment via knowledge graph merging, partitioning and embedding. In Mohammad Al Hasan and Li Xiong (eds.), *Proceedings of the 31st ACM International Conference on Information & Knowledge Management, Atlanta, GA, USA, October 17-21, 2022*, pp. 2240–2249. ACM, 2022. doi: 10.1145/3511808.3557374. URL <https://doi.org/10.1145/3511808.3557374>.
- Hongteng Xu, Dixin Luo, Hongyuan Zha, and Lawrence Carin. Gromov-wasserstein learning for graph matching and node embedding. In *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, pp. 6932–6941, 2019a.
- Kun Xu, Liwei Wang, Mo Yu, Yansong Feng, Yan Song, Zhiguo Wang, and Dong Yu. Cross-lingual knowledge graph alignment via graph matching neural network. In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, pp. 3156–3161, 2019b.
- Junchi Yan, Shuang Yang, and Edwin R. Hancock. Learning for graph matching and related combinatorial optimization problems. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020*, pp. 4988–4996, 2020.
- Xu Yang, Zhi-Yong Liu, and Hong Qiaoxu. Incorporating discrete constraints into random walk-based graph matching. *IEEE Trans. Syst. Man Cybern. Syst.*, 50(4):1406–1416, 2020.
- Chun-Han Yao, Boqing Gong, Hang Qi, Yin Cui, Yukun Zhu, and Ming-Hsuan Yang. Federated multi-target domain adaptation. In *IEEE/CVF Winter Conference on Applications of Computer Vision, WACV 2022, Waikoloa, HI, USA, January 3-8, 2022*, pp. 1081–1090. IEEE, 2022.
- Abdurrahman Yasar and Ümit V. Çatalyürek. An iterative global structure-assisted labeled network aligner. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2018, London, UK, August 19-23, 2018*, pp. 2614–2623, 2018.
- Zhenzhang Ye, Tarun Yenamandra, Florian Bernard, and Daniel Cremers. Joint deep multi-graph matching and 3d geometry learning from inhomogeneous 2d image collections. In *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*, pp. 3125–3133. AAAI Press, 2022. URL <https://ojs.aaai.org/index.php/AAAI/article/view/20220>.
- Liren Yu, Jiaming Xu, and Xiaojun Lin. Seedgnn: Graph neural networks for supervised seeded graph matching. *CoRR*, abs/2205.13679, 2022. doi: 10.48550/arXiv.2205.13679. URL <https://doi.org/10.48550/arXiv.2205.13679>.
- Tianshu Yu, Runzhong Wang, Junchi Yan, and Baoxin Li. Learning deep graph matching with channel-independent embedding and hungarian attention. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*, 2020.
- Tianshu Yu, Runzhong Wang, Junchi Yan, and Baoxin Li. Deep latent graph matching. In Marina Meila and Tong Zhang (eds.), *Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18-24 July 2021, Virtual Event*, volume 139 of *Proceedings of Machine Learning Research*, pp. 12187–12197. PMLR, 2021. URL <http://proceedings.mlr.press/v139/yu21d.html>.
- Huanding Zhang, Tao Shen, Fei Wu, Mingyang Yin, Hongxia Yang, and Chao Wu. Federated graph learning - A position paper. *CoRR*, abs/2105.11099, 2021a.
- Jiawei Zhang and Philip S. Yu. Integrated anchor and social link predictions across social networks. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015*, pp. 2125–2132, 2015.

- Jiawei Zhang, Xiangnan Kong, and Philip S. Yu. Predicting social links for new users across aligned heterogeneous social networks. In *2013 IEEE 13th International Conference on Data Mining, ICDM 2013, Dallas, TX, USA, December 7-10, 2013*, pp. 1289–1294, 2013.
- 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. In Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang (eds.), *Findings of the Association for Computational Linguistics: EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, pp. 613–621. Association for Computational Linguistics, 2022. URL <https://aclanthology.org/2022.findings-emnlp.43>.
- Ke Zhang, Carl Yang, Xiaoxiao Li, Lichao Sun, and Siu-Ming Yiu. Subgraph federated learning with missing neighbor generation. In Marc’Aurelio Ranzato, Alina Beygelzimer, Yann N. Dauphin, Percy Liang, and Jennifer Wortman Vaughan (eds.), *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual*, pp. 6671–6682, 2021b. URL <https://proceedings.neurips.cc/paper/2021/hash/34adeb8e3242824038aa65460a47c29e-Abstract.html>.
- Si Zhang and Hanghang Tong. FINAL: fast attributed network alignment. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2016, San Francisco, CA, USA, August 13-17, 2016*, pp. 1345–1354, 2016.
- Si Zhang and Hanghang Tong. Network alignment: Recent advances and future directions. In *CIKM ’20: The 29th ACM International Conference on Information and Knowledge Management, Virtual Event, Ireland, October 19-23, 2020*, pp. 3521–3522, 2020.
- Si Zhang, Hanghang Tong, Ross Maciejewski, and Tina Eliassi-Rad. Multilevel network alignment. In *The World Wide Web Conference, WWW 2019, San Francisco, CA, USA, May 13-17, 2019*, pp. 2344–2354, 2019.
- Si Zhang, Hanghang Tong, Yinglong Xia, Liang Xiong, and Jiejun Xu. Nettrans: Neural cross-network transformation. In Rajesh Gupta, Yan Liu, Jiliang Tang, and B. Aditya Prakash (eds.), *KDD ’20: The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, CA, USA, August 23-27, 2020*, pp. 986–996. ACM, 2020. doi: 10.1145/3394486.3403141. URL <https://doi.org/10.1145/3394486.3403141>.
- Si Zhang, Hanghang Tong, Long Jin, Yinglong Xia, and Yunsong Guo. Balancing consistency and disparity in network alignment. In Feida Zhu, Beng Chin Ooi, and Chunyan Miao (eds.), *KDD ’21: The 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, Singapore, August 14-18, 2021*, pp. 2212–2222. ACM, 2021c. doi: 10.1145/3447548.3467331. URL <https://doi.org/10.1145/3447548.3467331>.
- Xinwei Zhang, Mingyi Hong, and Jie Chen. GLASU: A communication-efficient algorithm for federated learning with vertically distributed graph data. *CoRR*, abs/2303.09531, 2023. doi: 10.48550/arXiv.2303.09531. URL <https://doi.org/10.48550/arXiv.2303.09531>.
- Yutao Zhang, Jie Tang, Zhilin Yang, Jian Pei, and Philip S. Yu. COSNET: connecting heterogeneous social networks with local and global consistency. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australia, August 10-13, 2015*, pp. 1485–1494, 2015.
- Kaixuan Zhao, Shikui Tu, and Lei Xu. IA-GM: A deep bidirectional learning method for graph matching. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, pp. 3474–3482. AAAI Press, 2021. URL <https://ojs.aaai.org/index.php/AAAI/article/view/16461>.
- Da Zheng, Minjie Wang, Quan Gan, Zheng Zhang, and George Karypis. Learning graph neural networks with deep graph library. In Amal El Fallah Seghrouchni, Gita Sukthankar, Tie-Yan Liu, and Maarten van Steen (eds.), *Companion of The 2020 Web Conference 2020, Taipei, Taiwan, April 20-24, 2020*, pp. 305–306. ACM / IW3C2, 2020.

- Vincent W. Zheng, Mo Sha, Yuchen Li, Hongxia Yang, Yuan Fang, Zhenjie Zhang, Kian-Lee Tan, and Kevin Chen-Chuan Chang. Heterogeneous embedding propagation for large-scale e-commerce user alignment. In *IEEE International Conference on Data Mining, ICDM 2018, Singapore, November 17-20, 2018*, pp. 1434–1439, 2018.
- Zexuan Zhong, Yong Cao, Mu Guo, and Zaiqing Nie. Colink: An unsupervised framework for user identity linkage. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018*, pp. 5714–5721, 2018.
- Fan Zhou, Lei Liu, Kunpeng Zhang, Goce Trajcevski, Jin Wu, and Ting Zhong. Deeplink: A deep learning approach for user identity linkage. In *2018 IEEE Conference on Computer Communications, INFOCOM 2018, Honolulu, HI, USA, April 16-19, 2018*, pp. 1313–1321, 2018a.
- Fan Zhou, Zijing Wen, Goce Trajcevski, Kunpeng Zhang, Ting Zhong, and Fang Liu. Disentangled network alignment with matching explainability. In *2019 IEEE Conference on Computer Communications, INFOCOM 2019, Paris, France, April 29 - May 2, 2019*, pp. 1360–1368, 2019.
- Fan Zhou, Chengtai Cao, Goce Trajcevski, Kunpeng Zhang, Ting Zhong, and Ji Geng. Fast network alignment via graph meta-learning. In *39th IEEE Conference on Computer Communications, INFOCOM 2020, Toronto, ON, Canada, July 6-9, 2020*, pp. 686–695, 2020a.
- Jingya Zhou and Jianxi Fan. Translink: User identity linkage across heterogeneous social networks via translating embeddings. In *2019 IEEE Conference on Computer Communications, INFOCOM 2019, Paris, France, April 29 - May 2, 2019*, pp. 2116–2124, 2019.
- Jun Zhou, Chaochao Chen, Longfei Zheng, Huiwen Wu, Jia Wu, Xiaolin Zheng, Bingzhe Wu, Ziqi Liu, and Li Wang. Vertically federated graph neural network for privacy-preserving node classification. *CoRR*, abs/2005.11903, 2020b.
- Tao Zhou, Ee-Peng Lim, Roy Ka-Wei Lee, Feida Zhu, and Jiuxin Cao. Retrofitting embeddings for unsupervised user identity linkage. In *Advances in Knowledge Discovery and Data Mining - 24th Pacific-Asia Conference, PAKDD 2020, Singapore, May 11-14, 2020, Proceedings, Part I*, pp. 385–397, 2020c.
- Xiaoping Zhou, Xun Liang, Xiaoyong Du, and Jichao Zhao. Structure based user identification across social networks. *IEEE Trans. Knowl. Data Eng.*, 30(6):1178–1191, 2018b.
- Jing Zhu, Danai Koutra, and Mark Heimann. CAPER: coarsen, align, project, refine - A general multilevel framework for network alignment. In Mohammad Al Hasan and Li Xiong (eds.), *Proceedings of the 31st ACM International Conference on Information & Knowledge Management, Atlanta, GA, USA, October 17-21, 2022*, pp. 4747–4751. ACM, 2022a. doi: 10.1145/3511808.3557563. URL <https://doi.org/10.1145/3511808.3557563>.
- Qiannan Zhu, Xiaofei Zhou, Jia Wu, Jianlong Tan, and Li Guo. Neighborhood-aware attentional representation for multilingual knowledge graphs. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*, pp. 1943–1949, 2019.
- Renbo Zhu, Xukun Luo, Meng Ma, and Ping Wang. Adaptive graph convolutional network for knowledge graph entity alignment. In Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang (eds.), *Findings of the Association for Computational Linguistics: EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, pp. 6011–6021. Association for Computational Linguistics, 2022b. URL <https://aclanthology.org/2022.findings-emnlp.444>.
- Xiangrong Zhu, Guangyao Li, and Wei Hu. Heterogeneous federated knowledge graph embedding learning and unlearning. In Ying Ding, Jie Tang, Juan F. Sequeda, Lora Aroyo, Carlos Castillo, and Geert-Jan Houben (eds.), *Proceedings of the ACM Web Conference 2023, WWW 2023, Austin, TX, USA, 30 April 2023 - 4 May 2023*, pp. 2444–2454. ACM, 2023. doi: 10.1145/3543507.3583305. URL <https://doi.org/10.1145/3543507.3583305>.
- Marinka Zitnik and Jure Leskovec. Predicting multicellular function through multi-layer tissue networks. *Bioinform.*, 33(14):i190–i198, 2017.

## A APPENDIX

### A.1 RELATED WORK

**Centralized Graph Matching.** Graph matching, also well known as network alignment, which aims to identify the same entities (i.e., nodes) across multiple graphs, has been a heated topic in recent years (Chu et al., 2019; Xu et al., 2019a; Wang et al., 2020f; Chen et al., 2020a;c; Zhang & Tong, 2016; Mu et al., 2016; Heimann et al., 2018; Li et al., 2019b; Fey et al., 2020; Qin et al., 2020; Feng et al., 2019; Ren et al., 2020). Research activities can be classified into three broad categories. (1) Topological structure-based techniques, which rely on only the structural information of nodes to match two or multiple graphs, including CrossMNA (Chu et al., 2019), MOANA (Zhang et al., 2019), GWL (Xu et al., 2019a), DPMC (Wang et al., 2020f), MGCN (Chen et al., 2020a), GraphSim (Bai et al., 2020), ZAC (Wang et al., 2020c), GRAMPA (Fan et al., 2020), CONE-Align (Chen et al., 2020c), DeepMatching (Wang et al., 2020b), Exact Graph Matching (Rácz & Sridhar, 2021), qc-DGM (Gao et al., 2021), OTTER (Weighill et al., 2021), IA-GM (Zhao et al., 2021), GMTracker (He et al., 2021d), Proxy Graph Matching (Tan et al., 2021), Fusion Moves for Graph Matching (Hutschenreiter et al., 2021), D-GAP (Lyu et al., 2022), CPUGA (Pei et al., 2022), CAPER (Zhu et al., 2022a); (2) Structure and/or attribute-based approaches, which utilize highly discriminative structure and attribute features for ensuring the matching effectiveness, such as FINAL (Zhang & Tong, 2016), ULink (Mu et al., 2016), gsaNA (Yasar & Çatalyürek, 2018), REGAL (Heimann et al., 2018), SNNA (Li et al., 2019b), CENALP (Du et al., 2019), GAlign (Huynh et al., 2020b), Deep Graph Matching Consensus (Fey et al., 2020), CIE (Yu et al., 2020), RE (Zhou et al., 2020c), Meta-NA (Zhou et al., 2020a), G-CREWE (Qin et al., 2020), GA-MGM (Wang et al., 2020d), EAGM (Qu et al., 2021), DLGM (Yu et al., 2021), SIGMA (Liu et al., 2021), CGMN (Jin et al., 2022), FOTA (Liu et al., 2022b), SCGM (Liu et al., 2022a), and Grad-Align+ (Park et al., 2022); (3) Heterogeneous methods employ heterogeneous structural, content, spatial, and temporal features to further improve the matching performance, including SCAN-PS (Zhang et al., 2013), MNA (Kong et al., 2013), HYDRA (Liu et al., 2014), COSNET (Zhang et al., 2015), Factoid Embedding (Xie et al., 2018), HEP (Zheng et al., 2018), LHNE (Wang et al., 2019c), ActiveIter (Ren et al., 2019), NAME (Zhou et al., 2019), TransLink (Zhou & Fan, 2019), DPLink (Feng et al., 2019), DETA (Meng et al., 2019). BANANA (Ren et al., 2020), SAUIL (Qiao et al., 2020), GCAN (Jiang et al., 2022), and Deep Multi-Graph Matching (Ye et al., 2022); Several papers review key achievements of graph matching across online information networks including state-of-the-art algorithms, evaluation metrics, representative datasets, and empirical analysis (Shu et al., 2016; Guzzi & Milenkovic, 2018; Huynh et al., 2020a; Vijayan et al., 2020; Yan et al., 2020; Zhang & Tong, 2020; Haller et al., 2022). It has been widely applied to many real-world applications, including protein network alignment in bioinformatics (Liu et al., 2017; Vijayan et al., 2020), user account linking in multiple social networks (Shu et al. (2016); Mu et al. (2016); Feng et al. (2019)), object matching in computer vision (Fey et al., 2020; Wang et al., 2020c;e; Yang et al., 2020), knowledge translation in multilingual knowledge bases (Xu et al., 2019b; Zhu et al., 2019; Sun et al., 2020; Wu et al., 2020; Zhu et al., 2022b; Chakrabarti et al., 2022; Liu et al., 2022c; Guo et al., 2022; Zhu et al., 2022b; Xin et al., 2022) and text matching (Chen et al., 2020b).

**Federated Graph Learning.** With the increasing privacy awareness, commercial competition, and regulation restrictions, real-world graph data is often generated locally and remains distributed graphs of multiple data silos among a large number of clients (Zheng et al., 2020; Chen et al., 2021; Zhang et al., 2021a). Federated graph learning (FGL) is a promising paradigm that enables collaborative training of shared machine learning models over large-scale distributed graph data, while preserving privacy of local data. Based on how graph data can be distributed across clients, existing FGL techniques on machine unlearning can be broadly classified into three categories below. (1) Graph-level FGL: each client possesses a set of graphs and all clients collaborate to train a shared model to predict graph properties, including (Xie et al., 2021; He et al., 2021a; Zhang et al., 2022; Chen et al., 2022c; Tan et al., 2022; Hu et al., 2023; Qu et al., 2023). Typical graph-level FGL task is graph classification/regression, which have been applied multiple domains, such as molecular property prediction (Xie et al., 2021; He et al., 2022) and brain network analysis (Bayram & Rekik, 2021); (2) Subgraph-level FL: each client contains a subgraph of a global graph, a part of node features, and a part of FGL model (Zhang et al., 2021b; Ni et al., 2021; Wang et al., 2022a; Chen et al., 2022a; Baek et al., 2022; Xie et al., 2023; Zhang et al., 2023; Li et al., 2023; Zhu et al., 2023; Tian et al., 2023). The clients aim to collaboratively train a global model with the partial features

and subgraphs to predict node properties. Typical graph-level FGL task is node classification and link prediction; (3) Node-level FGL: the clients are connected by a graph and thus each of them is treated as a node (Lalitha et al., 2019; Meng et al., 2021; Caldarola et al., 2021; Rizk & Sayed, 2021). Namely, the clients, rather than the data, are graph-structured. For example, each client performs learning with its own data and they exchange data through the communication graph (Lalitha et al., 2019; Meng et al., 2021). The server maintains the graph structure and uses a GNN to aggregate information (either models or data) collected from the clients (Caldarola et al., 2021; Rizk & Sayed, 2021).

A recent work studied the problem of federated knowledge graphs embedding with a byproduct of knowledge graph alignment (Peng et al., 2021). It exploits adversarial generation between pairs of knowledge graphs to translate identical entities and relations of different domains into near embedding spaces. To our best knowledge, this work is the first to have the potential to tackle the problem of general federated graph matching. However, it is a supervised learning method with aligned entities and relations as training data. In addition, it is possible that neural models may memorize inputs and reconstruct inputs from corresponding outputs (Carlini et al., 2021). The method exchanges the embeddings of entities and relations between clients and server. Adversarial samples and gradients are interchanged among the clients. Although a host client cannot access the embeddings of the other’s, the exchange of translational mapping matrices (i.e., the gradients in the generators of the other clients) makes it possible for the host client to reconstruct the former’s embeddings with the inverse of translational mapping matrices. The combination of the above two properties dramatically limits the applicability of the method in real scenarios. This work is the first to offer an unsupervised federated graph matching solution for inferring matched node pairs on different graphs across clients while maintaining the privacy requirement of federated learning, by leveraging the graphlet theory and trust region optimization.

## A.2 PROOF OF THEOREMS

**Theorem 1.** *Let  $x_k = \{v_1, v_2, \dots, v_k\}$  be the original node ordering of  $g_k$  via the subgraph expansion,  $\mathcal{S}(g_k) = [v_1, v_2, \dots, v_k]$  be the set of all possible node sequences of  $x_k$ ,  $x_k[i]$  be the  $i^{\text{th}}$  node in  $x_k$ ,  $F$  be a user-defined normalizing factor in the subgraph expansion, and  $h_l(x_k) = \{v_1, v_2, \dots, v_l, x_k, G\}$  be an induced subgraph of graph  $G$  with the first  $l$  nodes  $\{v_1, v_2, \dots, v_l\}$  in  $x_k$ , then the probability of getting a  $k$ -subgraph  $g_k$  via the subgraph expansion is*

$$p(g_k) = \sum_{x_k \in \mathcal{S}(g_k)} \frac{f(\deg(x_k[1]))}{F} \prod_{l=1}^{k-1} \frac{|E_{h_{l+1}(x_k)}| - |E_{h_l(x_k)}|}{\sum_{i=1}^l \deg(x_k[i]) - 2|E_{h_l(x_k)}|} \quad (25)$$

*Proof.* We can consider a subgraph expansion process as a way of sampling a sequence  $x_k = \{v_1, v_2, \dots, v_k\}$ , ordered from the first node sampled to the last one, that is then used to generate a  $k$ -subgraph  $g_k$ . Denote the set of such sequences as  $V_G^k$ . Let  $h_l = \{v_1, v_2, \dots, v_l\}$  is a  $l$ -subgraph of graph  $G$  obtained by the subgraph expansion process on step  $l$ . The probability of sampling node  $v_{l+1}$  on the step  $l+1$  to produce a  $(l+1)$ -subgraph  $h_{l+1} = \{v_1, v_2, \dots, v_l, v_{l+1}\}$  is equal to

$$\mathbb{P}(v_{l+1} | h_l) = \frac{\deg_{h_l}(v_{l+1})}{|\mathcal{N}_e(h_l)|} = \frac{|E_{h_{l+1}}| - |E_{h_l}|}{\sum_{i=1}^l \deg(v_i) - 2|E_{h_l}|} \quad (26)$$

where  $\mathcal{N}_e(h_l)$  is the set of all edges that connect a node in  $h_l$  and a node outside of  $h_l$ .  $\deg_{h_l}(v_{l+1})$  specifies the number of nodes in  $h_l$  that are connected to the node  $v_{l+1}$ .

Thus, the probability  $\tilde{p}(x_k)$  of sampling a sequence  $x_k = \{v_1, v_2, \dots, v_k\} \in \mathcal{S}(g_k)$  in the subgraph expansion process is equal to

$$\tilde{p}(x_k) = q(v_1) \prod_{l=1}^{k-1} \mathbb{P}(v_{l+1} | h_l) = \frac{f(\deg(v_1))}{F} \prod_{l=1}^{k-1} \frac{|E_{h_{l+1}}| - |E_{h_l}|}{\sum_{i=1}^l \deg(v_i) - 2|E_{h_l}|} \quad (27)$$

Notice that



$$p(g_k) = \sum_{x_k \in \mathcal{S}(g_k)} \tilde{p}(x_k) \quad (28)$$

Since

$$\mathcal{S}(g_k) = \{[v_1, \dots, v_k] | \{v_1, \dots, v_k\} = V_{g_k}, g_k | \{v_1, \dots, v_l\} \text{ is connected}\} \quad (29)$$

then we have

$$p(g_k) = \sum_{x_k \in \mathcal{S}(g_k)} \frac{f(\deg(x_k[1]))}{F} \prod_{l=1}^{k-1} \frac{|E_{h_{l+1}(x_k)}| - |E_{h_l(x_k)}|}{\sum_{i=1}^l \deg(x_k[i]) - 2|E_{h_l(x_k)}|} \quad (30)$$

where  $x_k = \{v_1, v_2, \dots, v_k\}$  be the original node ordering of  $g_k$  via the subgraph expansion process.  $x_k[i]$  be the  $i^{\text{th}}$  node in  $x_k$ .  $h_l(x_k) = \{v_1, v_2, \dots, v_l, x_k, G\}$  be an induced subgraph of graph  $G$  with the first  $l$  nodes  $\{v_1, v_2, \dots, v_l\}$  in  $x_k$

Therefore, the proof is concluded.

**Theorem 2.** Let  $\tilde{n}_{kr}(G) = \frac{1}{O} \sum_{o=1}^O \frac{\mathbb{I}(g_{ko} \sim \mathcal{G}_r)}{p(g_{ko})}$  be the estimation of graphlet counts,  $d_1, \dots, d_k$  be the  $k$  highest degrees of nodes in  $G$ , and denote  $D = \prod_{l=2}^{k-1} (d_1 + \dots + d_k)$ . If  $q$  for sampling the starting node is the stationary distribution of the node random walk, then the upper bound of the variance  $\text{Var}(\tilde{n}_{kr}(G))$  is

$$\text{Var}(\tilde{n}_{kr}(G)) \leq \frac{1}{O} n_{kr}(G) \frac{2|E_G|}{|\mathcal{S}(\mathcal{G}_r)|} D \quad (31)$$

*Proof.* Consider sampling the starting node  $v_1$  independently and from an arbitrary distribution  $q$  when we have access to all the nodes. Sampling nodes independently implies that the subgraph expansion process will result in independent  $k$ -subgraph samples. Thus, the variance of the graphlet count estimator can be decomposed into the variance of the individual  $k$ -subgraph samples. The variance of the estimator  $\tilde{n}_{kr}(G)$  is then

$$\text{Var}(\tilde{n}_{kr}(G)) = \frac{1}{O} \text{Var} \left( \frac{\mathbb{I}(g_{kO} \sim \mathcal{G}_r)}{p(g_{kO})} \right) = \frac{1}{O} \left( \sum_{g_k \in \mathcal{G}_k} \frac{\mathbb{I}(g_k \sim \mathcal{G}_r)}{p(g_k)} - n_{kr}(G)^2 \right) \quad (32)$$

It is observed that the variance  $\text{Var}(\tilde{n}_{kr}(G))$  is small when the distribution of  $p(g_k)$  is close to uniform distribution. A larger  $p(g_k)$  results in a smaller variance of the estimator. Thus, the variation can be reduced by an appropriate choice of  $q$  for sampling the starting node, say a smaller normalizing factor  $F$ . In this case, the estimated graphlet count  $\tilde{n}_{kr}(G)$  is close to the actual count  $n_{kr}(G)$ , which implies that the graphlet samples and all graphlets share similar distributions.

Let

$$\phi_o = \frac{\mathbb{I}(g_{ko} \sim \mathcal{G}_r)}{p(g_{ko})} \quad (33)$$

The variance can be rewritten as follows.

$$\text{Var}(\tilde{n}_{kr}(G)) = \frac{1}{O} \text{Var}(\phi_o) \quad (34)$$

Notice that  $n_{kr}(G) = \mathbb{E}\phi_o$ , and  $\tilde{n}_{kr}(G) = \frac{1}{O} \sum_{o=1}^O \phi_o$  for the estimator.

We can bound the variance in Eq.(32) by the second moment, which is bounded by,

$$\mathbb{E}\phi_o^2 \leq \mathbb{E}\phi_o \max \phi_o = n_{kr}(G) \max \phi_o \quad (35)$$

By seeking to control the the maximum of  $\phi_o$ , we have

$$\max_{g_k} \frac{1}{p(g_k)} \leq \max_{x_k} \frac{1}{|\mathcal{S}(g_k)|\tilde{p}(x_k)} \leq \max \frac{\prod_{l=1}^{k-1} (d_1 + \dots + d_l)}{|\mathcal{S}(\mathcal{G}_r)|q(d_1)} \quad (36)$$

and we obtain

$$\max_x \frac{|\mathcal{N}_v(x)|}{\mathcal{S}(\mathcal{G}_r)|\tilde{p}(x)} \leq \max \frac{\prod_{l=1}^{k-1} (d_1 + \dots + d_l)}{|\mathcal{S}(\mathcal{G}_r)|q(d_1)} \quad (37)$$

Thus, we can construct a bound on  $\text{Var}(\phi_o)$  and  $\text{Var}(\tilde{n}_{kr}(G))$ .

Therefore, the proof is concluded.

**Theorem 3.** Let  $d$  be the dimension of the flattened  $M_{b+1}$ ,  $\otimes$  be an appropriate tensor product,  $\mathcal{A}_{b+1} \in \mathbb{R}^{d \times d \times d}$  and  $\mathcal{B}_{b+1} \in \mathbb{R}^{d \times d \times d \times d}$  are the tensors of  $\mathcal{L}^s(M_{b+1})$  at iteration  $b+1$  satisfying

$$\mathcal{A}_{b+1} \otimes z_b^3 = \sum_{i,j,k=1}^d \frac{\partial^3 \mathcal{L}^s(M_{b+1})}{\partial M^i \partial M^j \partial M^k} z_b^i z_b^j z_b^k \quad (38)$$

and

$$\mathcal{B}_{b+1} \otimes z_b^4 = \sum_{i,j,k,l=1}^d \frac{\partial^4 \mathcal{L}^s(M_{b+1})}{\partial M^i \partial M^j \partial M^k \partial M^l} z_b^i z_b^j z_b^k z_b^l. \quad (39)$$

Suppose that  $\mathcal{L}^s(M_{b+1})$  is sufficiently smooth, if  $\|z_b\|$  is small enough, then we have

$$z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b - \alpha_{b+1}(\omega) z_b^T z_b = \left(\frac{1}{2} - \frac{\omega}{6}\right) \mathcal{A}_{b+1} \otimes z_b^3 - \left(\frac{1}{6} - \frac{\omega}{12}\right) \mathcal{B}_{b+1} \otimes z_b^4 + \mathcal{O}\left(\|z_b\|^5\right) \quad (40)$$

*Proof.* By utilizing the Taylor expansion, we obtain

$$\begin{aligned} \mathcal{L}^s(M_b) &= \mathcal{L}^s(M_{b+1}) - (\nabla \mathcal{L}^s(M_{b+1}))^T z_b + \frac{1}{2} z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b - \\ &\quad \frac{1}{6} \mathcal{A}_{b+1} \otimes z_b^3 + \frac{1}{24} \mathcal{B}_{b+1} \otimes z_b^4 + \mathcal{O}\left(\|z_b\|^5\right) \end{aligned} \quad (41)$$

and

$$\begin{aligned} (\nabla \mathcal{L}^s(M_b))^T z_b &= (\nabla \mathcal{L}^s(M_{b+1}))^T z_b - z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b + \\ &\quad \frac{1}{2} \mathcal{A}_{b+1} \otimes z_b^3 - \frac{1}{6} \mathcal{B}_{b+1} \otimes z_b^4 + \mathcal{O}\left(\|z_b\|^5\right) \end{aligned} \quad (42)$$

In addition, we have

$$\alpha_{b+1}(\omega) = \frac{z_b^T y_b + \omega \left[ 2(\mathcal{L}^s(M_b) - \mathcal{L}^s(M_{b+1})) + (\nabla \mathcal{L}^s(M_b) + \nabla \mathcal{L}^s(M_{b+1}))^T z_b \right]}{z_b^T z_b} \quad (43)$$

By combining Eqs.(41), (42), and (43), we get

$$\begin{aligned}
& z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b - \alpha_{b+1}(\omega) z_b^T z_b \\
&= z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b - z_b^T y_b - \\
& \quad \omega \left[ 2(\mathcal{L}^s(M_b) - \mathcal{L}^s(M_{b+1})) + (\nabla \mathcal{L}^s(M_b) + \nabla \mathcal{L}^s(M_{b+1}))^T z_b \right] \\
&= \left( \frac{1}{2} - \frac{\omega}{6} \right) \mathcal{A}_{b+1} \otimes z_b^3 - \left( \frac{1}{6} - \frac{\omega}{12} \right) \mathcal{B}_{b+1} \otimes z_b^4 + \mathcal{O}(\|z_b\|^5)
\end{aligned} \tag{44}$$

Therefore, the proof is concluded.

Sensitivity analysis of weight  $\omega$ . Based on Eq.(43), if  $\omega = 0$ , we have

$$\alpha_{b+1}(0) = \frac{z_b^T y_b}{z_b^T z_b} \tag{45}$$

Then, it derives the following equation based on Eq.(40).

$$z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b - \alpha_{b+1}(\omega) z_b^T z_b = \frac{1}{2} \mathcal{A}_{b+1} \otimes z_b^3 - \frac{1}{6} \mathcal{B}_{b+1} \otimes z_b^4 + \mathcal{O}(\|z_b\|^5) \tag{46}$$

According to Eq.(46) and Theorem 3, it is reasonable to believe that if the weight parameter  $\omega$  is chosen such that

$$\left| \frac{1}{2} - \frac{\omega}{6} \right| < \frac{1}{2} \tag{47}$$

and

$$\left| \frac{1}{6} - \frac{\omega}{12} \right| < \frac{1}{6} \tag{48}$$

i.e.,  $0 < \omega < 4$ , then  $\alpha_{b+1}(\omega) z_b^T z_b$  may capture the second order curvature  $z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b$  with a high precision.

Now, let us further compare several possible choices of  $\omega$  and the corresponding formulas for  $\alpha_{b+1}(\omega)$ .

(1) If  $\omega = 1$ , then

$$\alpha_{b+1}(1) = \frac{z_b^T y_b + 2(\mathcal{L}^s(M_b) - \mathcal{L}^s(M_{b+1})) + (\nabla \mathcal{L}^s(M_b) + \nabla \mathcal{L}^s(M_{b+1}))^T z_b}{z_b^T z_b} \tag{49}$$

The resulting matrix  $\alpha_{b+1}(1)\mathbb{I}$  satisfies the weak quasi-Newton equation in Eq.(17). Based on Eq.(40), we have

$$z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b - \alpha_{b+1}(1) z_b^T z_b = \frac{1}{3} \mathcal{A}_{b+1} \otimes z_b^3 - \frac{1}{12} \mathcal{B}_{b+1} \otimes z_b^4 + \mathcal{O}(\|z_b\|^5) \tag{50}$$

(2) If  $\omega = 2$ , then

$$\alpha_{b+1}(2) = \frac{z_b^T y_b + 4(\mathcal{L}^s(M_b) - \mathcal{L}^s(M_{b+1})) + 2(\nabla \mathcal{L}^s(M_b) + \nabla \mathcal{L}^s(M_{b+1}))^T z_b}{z_b^T z_b} \tag{51}$$

The following equation is derived from Eq.(40).

$$z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b - \alpha_{b+1}(2) z_b^T z_b = \frac{1}{6} \mathcal{A}_{b+1} \otimes z_b^3 + \mathcal{O}(\|z_b\|^5) \quad (52)$$

(3) If  $\omega = 3$ , then

$$\alpha_{b+1}(3) = \frac{z_b^T y_b + 6(\mathcal{L}^s(M_b) - \mathcal{L}^s(M_{b+1})) + 3(\nabla \mathcal{L}^s(M_b) + \nabla \mathcal{L}^s(M_{b+1}))^T z_b}{z_b^T z_b} \quad (53)$$

According to Eq.(40), we obtain

$$z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b - \alpha_{b+1}(3) z_b^T z_b = \frac{1}{12} \mathcal{B}_{b+1} \otimes z_b^4 + \mathcal{O}(\|z_b\|^5) \quad (54)$$

**Theorem 4.** Suppose  $\|\nabla \mathcal{L}^s(M_b)\| \neq 0$ , the solution  $z_b$  of the separate trust region optimization  $\arg \min u_b(z) = \mathcal{L}^s(M_b) + (\nabla \mathcal{L}^s(M_b))^T z + \frac{1}{2} z^T \nabla^2 \mathcal{L}^s(M_b) z$ , s.t.  $\|z\| \leq \Delta^s$  in Eq.(11) satisfies

$$u_b(0) - u_b(z_b) \geq \frac{1}{2} \|\nabla \mathcal{L}^s(M_b)\| \min \left\{ \Delta^s, \frac{\|\nabla \mathcal{L}^s(M_b)\|}{\alpha_b} \right\} \quad (55)$$

*Proof.* We have the separate trust region optimization based on two weak quasi-Newton conditions as follows.

$$z^* = \arg \min u_b(z) \approx \mathcal{L}^s(M_b) + (\nabla \mathcal{L}^s(M_b))^T z + \frac{1}{2} \alpha_b z^T z, \text{ s.t. } \|z\| \leq \Delta^s \quad (56)$$

Since  $\|\nabla \mathcal{L}^s(M_b)\| \neq 0$ , the solution of the separate trust region optimization based on two weak quasi-Newton conditions in Eq.(56) can be solved as follows.

(1) if  $\|\nabla \mathcal{L}^s(M_b)\| \leq \alpha_b \Delta^s$ ,  $z_b = -\frac{1}{\alpha_b} \nabla \mathcal{L}^s(M_b)$ ;

(2) if  $\|\nabla \mathcal{L}^s(M_b)\| > \alpha_b \Delta^s$ , the optimal solution  $s_k$  will be on the boundary of the separate trust region, i.e.,  $z_b$  is the solution of the following problem.

$$z^* = \arg \min u_b(z) \approx \mathcal{L}^s(M_b) + (\nabla \mathcal{L}^s(M_b))^T z + \frac{1}{2} \alpha_b z^T z, \text{ s.t. } \|z\| = \Delta^s \quad (57)$$

From Eq.(57), we have the solution  $z_b = -\frac{\Delta^s}{\|\nabla \mathcal{L}^s(M_b)\|} \nabla \mathcal{L}^s(M_b)$ .

Thus, the general solution of the separate trust region optimization based on two weak quasi-Newton conditions in Eq.(56) can be rewritten as follows.

$$z_b = -\frac{1}{\tilde{\alpha}_b} \nabla \mathcal{L}^s(M_b), \quad \text{where } \tilde{\alpha}_b = \max \left\{ \alpha_b, \frac{\|\nabla \mathcal{L}^s(M_b)\|}{\Delta^s} \right\} \quad (58)$$

If  $\|\nabla \mathcal{L}^s(M_b)\| \leq \alpha_b \Delta^s$ , then  $z_b = -\frac{1}{\alpha_b} \nabla \mathcal{L}^s(M_b)$ . Thus, we obtain

$$\begin{aligned} u_b(0) - u_b(z_b) &= -(\nabla \mathcal{L}^s(M_b))^T \left( -\frac{1}{\alpha_b} \nabla \mathcal{L}^s(M_b) \right) \\ &\quad - \frac{1}{2} \left( -\frac{1}{\alpha_b} \nabla \mathcal{L}^s(M_b) \right)^T \alpha_b I \left( -\frac{1}{\alpha_b} \nabla \mathcal{L}^s(M_b) \right) \\ &= \frac{\|\nabla \mathcal{L}^s(M_b)\|^2}{\alpha_b} - \frac{1}{2} \frac{\|\nabla \mathcal{L}^s(M_b)\|^2}{\alpha_b} \\ &= \frac{1}{2} \frac{\|\nabla \mathcal{L}^s(M_b)\|^2}{\alpha_b} \end{aligned} \quad (59)$$

If  $\|\nabla\mathcal{L}^s(M_b)\| > \alpha_b\Delta^s$ , then  $z_b = -\frac{\Delta^s}{\|\nabla\mathcal{L}^s(M_b)\|}\nabla\mathcal{L}^s(M_b)$ . Hence, we have

$$\begin{aligned}
 u_b(0) - u_b(z_b) &= -(\nabla\mathcal{L}^s(M_b))^T \left( -\frac{\Delta^s}{\|\nabla\mathcal{L}^s(M_b)\|}\nabla\mathcal{L}^s(M_b) \right) \\
 &\quad - \frac{1}{2} \left( -\frac{\Delta^s}{\|\nabla\mathcal{L}^s(M_b)\|}\nabla\mathcal{L}^s(M_b) \right)^T \alpha_b I \left( -\frac{\Delta^s}{\|\nabla\mathcal{L}^s(M_b)\|}\nabla\mathcal{L}^s(M_b) \right) \\
 &= \Delta^s \|\nabla\mathcal{L}^s(M_b)\| - \frac{1}{2}\alpha_b(\Delta^s)^2 \\
 &> \Delta^s \|\nabla\mathcal{L}^s(M_b)\| - \frac{1}{2}\Delta^s \|\nabla\mathcal{L}^s(M_b)\| \\
 &= \frac{1}{2}\Delta^s \|\nabla\mathcal{L}^s(M_b)\|
 \end{aligned} \tag{60}$$

By integrating Eqs.(59) and (60), we obtain Eq.(55).

Therefore, the proof is concluded.

### A.3 ADDITIONAL EXPERIMENTS

**Final Performance and Convergence on SNS, PPI, and DBLP.** Table 3 and Figures-3-6 exhibit the quality of six centralized graph matching, ~~six federated graph learning~~, and four federated domain adaption algorithms over SNS, PPI, and DBLP respectively, based on  $Hits@1$ ,  $Hits@5$ ,  $Hits@10$ ,  $Hits@50$ , and  $Loss$ . Similar trends are observed for the comparison of federated graph matching effectiveness and convergence in these figures: our UFGM method achieves the close or much better than the centralized graph matching method, regarding  $Hits@1$  ( $>0.37$ ),  $Hits@5$  ( $>0.43$ ),  $Hits@10$  ( $>0.41$ ), and  $Hits@50$  ( $>0.45$ ) on three datasets respectively. Our UFGM method achieves better performance than all the competitors of ~~federated graph learning and~~ federated domain adaption in most experiments. In addition, our UFGM method can significantly speedup the convergence on two datasets in most experiments, compared with all federated learning algorithms. Especially, UFGM can converge around 1,000 training rounds and then always keep stable on SNS. This demonstrates that UFGM fully utilizes the proposed graphlet feature extraction techniques to generate the pseudo training data and employ the strength of supervised graph matching for accelerating the training convergence. The above experiment results demonstrate that UFGM is effective as well as efficient for addressing the federated graph matching problem. This advantage is very important for large-scale federated graph matching. For example, innovators were asked to develop privacy-preserving federated learning solutions that help tackle the challenge of international money laundering across large-scale local transaction network owned by multiple banks (NSF, 2022). Federated graph matching (FGM) can be utilized to infer cross-graph edges over multiple clients (e.g., identify the same potential criminals transferring money between multiple organizations) and derive a latent global graph (i.e., a global financial transaction network) (Suzumura et al., 2019; Wang et al., 2019a; Zhang et al., 2021a).

Table 3: Final performance on DBLP

Type	Algorithm	$Hits@1$	$Hits@5$	$Hits@10$	$Hits@50$	$Loss$
Centralized Graph Matching	NextAlign	<b>0.572</b>	<b>0.609</b>	0.632	0.690	<b>0.222</b>
	NetTrans	0.529	0.592	0.616	0.632	1.881
	CPUGA	0.136	0.199	0.276	0.296	2.232
	ASAR-GM	0.172	0.237	0.260	0.271	2.052
	SIGMA	0.276	0.360	0.378	0.421	1.992
	SeedGNN	0.530	0.582	<b>0.637</b>	<b>0.702</b>	4.185
Federated Domain Adaption	DualAdapt	0.000	0.001	0.001	0.001	4.023
	EFDA	0.000	0.000	0.000	0.000	2.452
	WSDA	0.000	0.001	0.001	0.001	3.332
	FKA	0.001	0.001	0.002	0.002	4.601
	UFGM	0.453	0.552	0.591	0.659	0.332

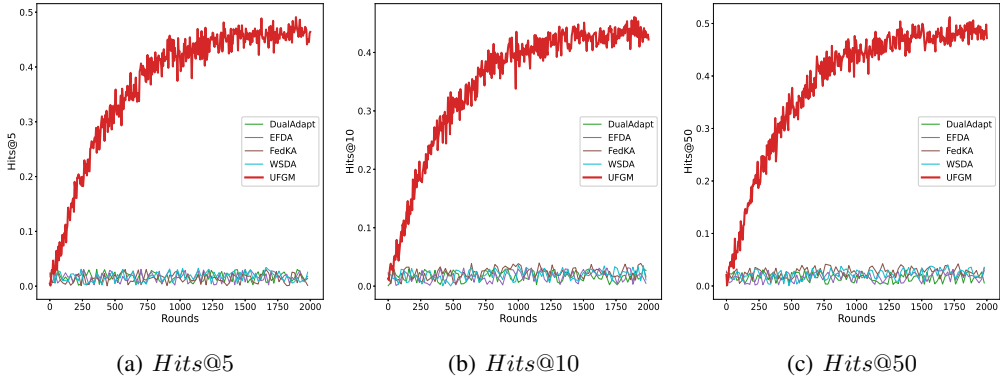


Figure 3: Convergence on SNS

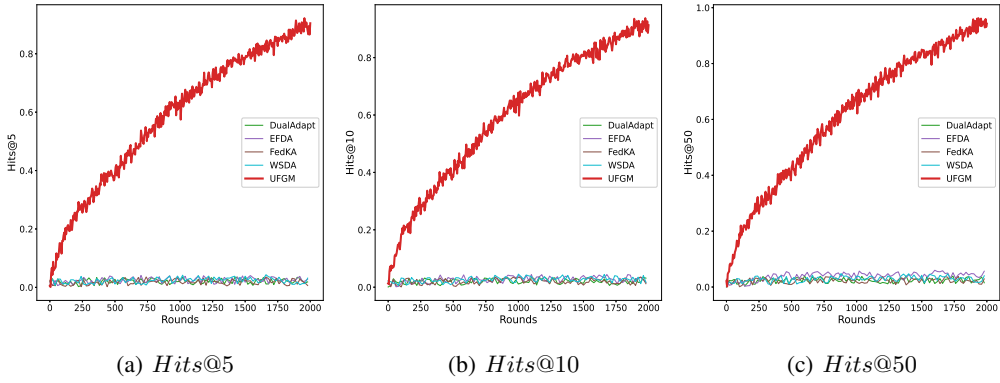


Figure 4: Convergence on PPI

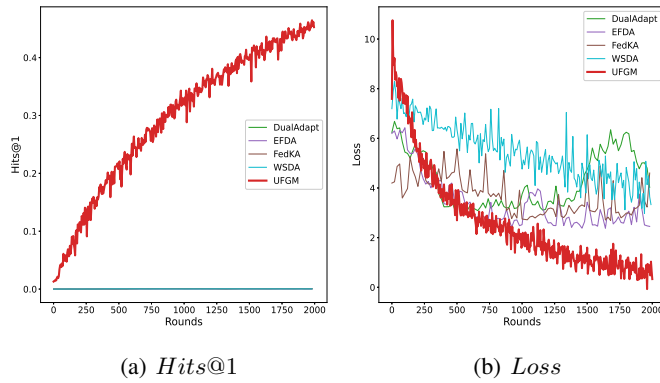


Figure 5: Convergence on DBLP

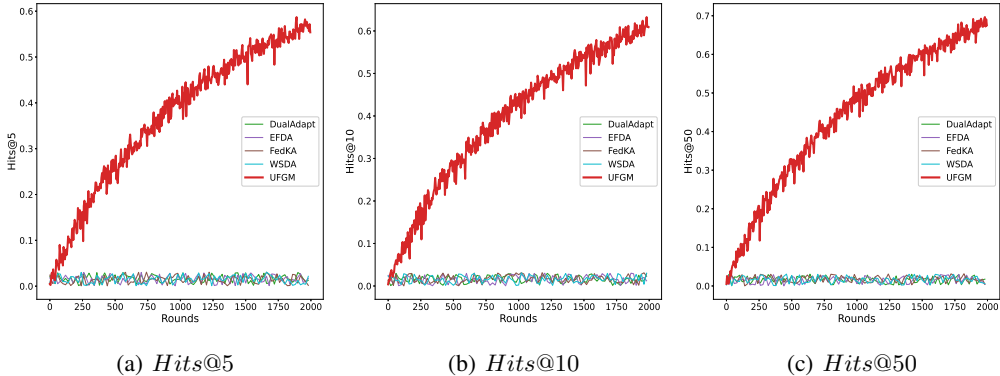


Figure 6: Convergence on DBLP

Table 4: Final performance of centralized learning on SNS

Algorithm	Dataset	$Hits@1$	$Hits@5$	$Hits@10$	$Hits@50$	$Loss$
UFGM	SNS	0.371	0.440	0.411	0.459	0.501
	PPI	0.771	0.880	0.902	0.930	0.659
	DBLP	0.453	0.552	0.591	0.659	0.332
UFGM-C	SNS	0.387	0.417	0.478	0.486	0.427
	PPI	0.786	0.911	0.922	0.932	0.495
	DBLP	0.471	0.563	0.635	0.718	0.182

**Final Performance of Centralized Learning.** We evaluate two versions of UFGM to show the strength of our UFGM method for federated graph matching. UFGM is the federated version with graph data encryption, graphlet feature extraction, model evaluation on the server, model optimization with the trust region on the clients, and Hessian approximation. UFGM-C is the centralized version with raw graph data uploaded to the server, graphlet feature extraction, model evaluation and model optimization with the standard stochastic gradient descent on the server. The experiment results in Table 4 exhibit that the performance of the centralized version, UFGM-C, is close to our federated version, UFGM over all three datasets. This further validates that our UFGM algorithm can achieve superior performance for the federated graph matching.

Table 5: Final performance of UFGM on large-scale datasets

Dataset	$Hits@1$	$Hits@5$	$Hits@10$	$Hits@50$	$Loss$
DBLP 100K	0.536	0.659	0.671	0.735	1.115
DBLP 200K	0.405	0.496	0.559	0.619	1.720

**Final Performance on Large-scale Datasets.** In order to evaluate the scalability of our UFGM algorithm on large-scale datasets, we select and split the original DBLP dataset into 20 graphs by publication year, ranging from 2002-2022, such that each graph has around 100,000 and 200,000 authors as nodes and coauthor relationships as edges respectively. Thus, most authors occur in all 20 graphs but different graphs contain few emeritus and new authors. The experiment results in Table 5 demonstrate that our UFGM method scales well on two large-scale datasets.

Table 6: Final performance with new baselines on DBLP

Type	Algorithm	Hits@1	Hits@5	Hits@10	Hits@50	Loss
Unsupervised Centralized Graph Matching	GANN-GM	0.034	0.058	0.082	0.126	4.125
	REGAL	0.349	0.425	0.472	0.551	N/A
Unsupervised Federated Domain Adaption	LADD	0.002	0.003	0.004	0.011	4.120
	FMTDA	0.008	0.011	0.016	0.029	1.597
	UFGM	<b>0.453</b>	<b>0.552</b>	<b>0.591</b>	<b>0.659</b>	<b>0.332</b>

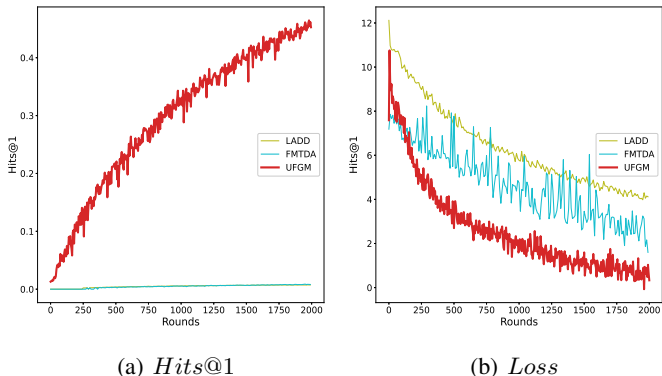


Figure 7: Convergence on DBLP

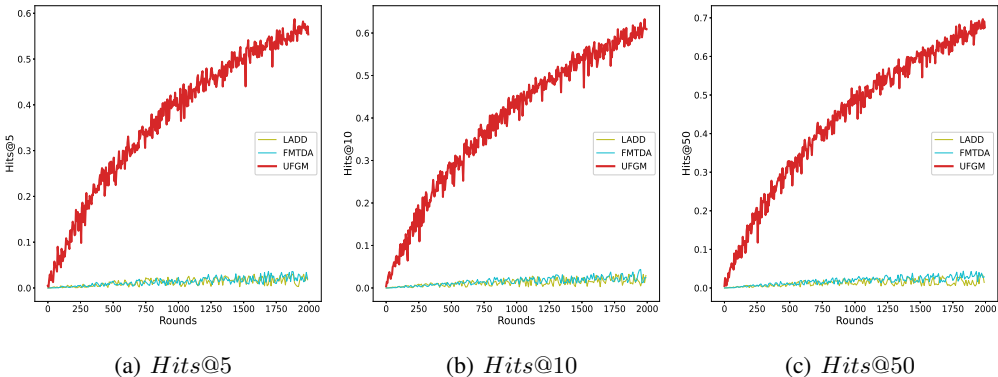


Figure 8: Convergence on DBLP

**Final Performance and Convergence with New Baselines on DBLP.** Table 6 and Figures 7 and 8 exhibit the quality of our UFGM method with two unsupervised centralized graph matching approaches of **GANN-GM** (Wang et al., 2023) and **REGAL** (Heimann et al., 2018) and two unsupervised federated domain adaption algorithms of **LADD** (Shenaj et al., 2023) and **FMTDA** (Yao et al., 2022), and two unsupervised federated graph learning methods of **FedWalk** (Pan & Zhu, 2022) and **Lumos** (Pan et al., 2023). Similar trends are observed for the comparison among these unsupervised federated learning methods: our UFGM method outperforms these baselines in all experiments, in terms of both final performance and convergence. Notice that REGAL is a matrix factorization-based graph alignment method and thus there are no loss functions in it.



Table 7: Final performance of quasi-Newton approximation on three datasets

Algorithm	Dataset	Hits@1	Hits@5	Hits@10	Hits@50	Loss	Runing Time (m)
UFGM	SNS	0.371	0.440	0.411	0.459	0.501	168
	PPI	0.771	0.880	0.902	0.930	0.659	153
	DBLP	0.453	0.552	0.591	0.659	0.332	732
UFGM-E	SNS	0.380	0.441	0.472	0.497	0.453	399
	PPI	0.786	0.907	0.937	0.947	0.633	367
	DBLP	0.487	0.606	0.657	0.687	0.228	1,556

**Final Performance of quasi-Newton Approximation.** We evaluate two versions of UFGM to show the strength of the quasi-Newton approximation for improving the efficiency while maintaining the quality federated graph matching. UFGM is the approximate version with the quasi-Newton approximation. UFGM-E is the exact version with the exact Hessian computation. The experiment results in Table 7 exhibit that the approximate version UFGM achieves slightly lower performance than the exact version UFGM-E but has much smaller running time. This demonstrates that the quasi-Newton approximation method is able to dramatically improve the efficiency while maintaining the utility constraints.

#### A.4 PARAMETER SENSITIVITY

In this section, we conduct more experiments to validate the sensitivity of various parameters in our UFGM method for the federated graph matching task.

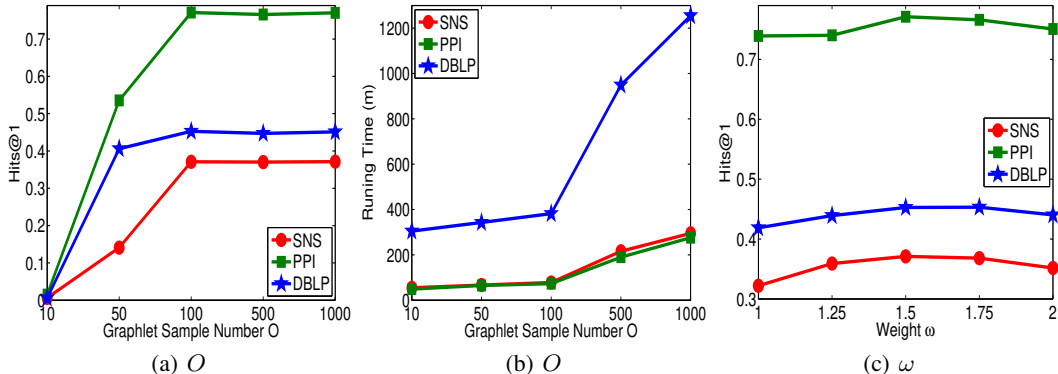


Figure 9: Final Hits@1 with varying parameters on three datasets

**Impact of graphlet sample numbers.** Figure 9 (a) measures the performance effect of sampled graphlet numbers in the Monte Carlo Markov Chain sampling for graphlet enumeration and estimation by varying  $O$  from 10 to 1,000. We have witnessed the performance curves by UFGM initially increase quickly and then become stable when  $O$  continuously increases. Initially, a large  $O$  can help utilize the strength of effective graphlet feature extraction for generating the pseudo training data for tackling the dilemma of unsupervised graph matching in federated setting and employing the strength of supervised graph matching. Later on, when  $O$  continues to increase and goes beyond some thresholds, the performance curves become stable. A rational guess is that after the enough graphlet features have been already extracted at a certain threshold and considered in the FGM training, our UFGM model is able to generate a good graph matching result. When  $O$  continuously increases, this does not affect the performance of graph matching any more. Figure 9 (b) reports the corresponding running time of our UFGM model by varying sampled graphlet number  $O$  from 10 to 1,000. We make the observation on the quality and efficiency over three datasets: both the performance scores and the running time keep increasing when the sampled graphlet number is increasing. A rational guess is that a larger sampled graphlet number exchanges better performance with more sampling and processing time.

**Impact of weight  $\omega$  between two types of weak quasi-Newton conditions.** Figures 9 (c) shows the influence of weight of two types of weak quasi-Newton conditions in our UFGM model by varying it from 1 to 2. It is observed that the performance initially raises when the  $\omega$  increases. Intuitively, a large  $\omega$  can help the algorithm well balance two types of weak quasi-Newton conditions and thus help improve the quality of separate trust region and graph matching. Later on, the performance curves decrease quickly when the  $\omega$  continuously increases. A reasonable explanation is that a too large  $\omega$  may ruin the first type of weak quasi-Newton condition and miss the optimal solution in the search process. Thus, it is important to determine the optimal  $\omega$  for separate trust region.

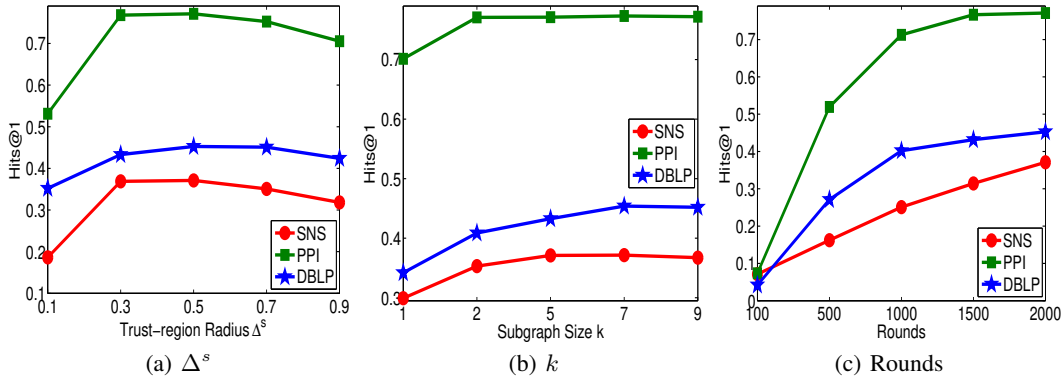


Figure 10: Final *Hits@1* with varying parameters on three datasets

**Influence of trust-region radius.** Figure 10 (a) demonstrates the influence of trust-region radius in the separate trust region in our UFGM model by varying it from 0.1 to 0.9. We have observed that the performance initially raises when the trust-region radius increases. Intuitively, a trust-region radius can help the algorithm quickly find the optimal solution and thus help improve the quality of federated graph matching. Later on, the performance curves decrease quickly when the trust-region radius continuously increases. A reasonable explanation is that a too large trust-region radius may miss the optimal solution with large step size in the search process. Thus, it is important to determine the optimal trust-region radius for the federated graph matching.

**Sensitivity of subgraph size.** Figure 10 (b) shows the influence of  $k$ -graphlets with  $k$  nodes in the graphlet feature extraction in our UFGM model by varying it from 1 to 9. We make the observation: on the quality over three datasets: the performance curves keep increasing when the maximum subgraph size for the graphlet counting increases and then become stable when  $k$  continuously increases. A rational guess is that a larger subgraph size initially makes UFGM capture more graphlet features and be more resilient to the unavailability of the training data. Later on, when  $k$  continues to increase and goes beyond some thresholds, the performance curves become stable. A reasonable explanation is that after the enough graphlet features have been already extracted at a certain threshold and considered in the FGM training, our UFGM model is able to generate a good graph matching result. When  $k$  continuously increases, this does not affect the performance of graph matching any more.

**Impact of training round.** Figure 10 (c) exhibits the sensitivity of training rounds of our UFGM model by varying them from 100 and 2,000. As we can see, the performance curves continuously increase with increasing training rounds. This is consistent with the fact that more training rounds makes the graph matching models be resilient to the federated setting. It is observed that our UFGM converges very fast on three datasets. From rounds 1,500 to 2,000, the *Hits@1* scores oscillate within the range of 7.8% on three datasets.

Table 8: Final performance of quasi-Newton approximation on three datasets

Pseudo Training Data	Dataset	Hits@1	Hits@5	Hits@10	Hits@50	Loss
20%	SNS	0.077	0.116	0.176	0.292	0.557
	PPI	0.372	0.440	0.497	0.627	0.669
	DBLP	0.226	0.291	0.309	0.336	0.392
40%	SNS	0.157	0.198	0.306	0.335	0.582
	PPI	0.519	0.588	0.702	0.796	0.691
	DBLP	0.312	0.378	0.397	0.442	0.449
60%	SNS	0.302	0.332	0.347	0.407	0.512
	PPI	0.628	0.776	0.825	0.917	0.686
	DBLP	0.381	0.397	0.458	0.559	0.358
80%	SNS	0.362	0.407	0.416	0.438	0.531
	PPI	0.752	0.802	0.857	0.927	0.689
	DBLP	0.406	0.497	0.533	0.610	0.349
100%	SNS	0.371	0.440	0.411	0.459	0.501
	PPI	0.771	0.880	0.902	0.930	0.659
	DBLP	0.453	0.552	0.591	0.659	0.332

**Influence of pseudo training data.** Table 8 tests the influence of the pseudo training data for the performance of graph matching by varying the ratio of the pseudo training data from 20% to 100%. The ratio 100% corresponds to the number of the pseudo matched node pairs used in our current experiments. The numbers are 3,041 on SNS, 1,264 over PPI, and 2,817 on DBLP respectively. As we can see, the performance scores continuously increase with increasing pseudo training data. This is consistent with the fact that more training data makes the graph matching models achieve better performance.

#### A.5 EXPERIMENTAL DETAILS

**Environment.** The experiments were conducted on a compute server running on Red Hat Enterprise Linux 7.2 with 2 CPUs of Intel Xeon E5-2650 v4 (at 2.66 GHz) and 8 GPUs of NVIDIA GeForce GTX 2080 Ti (with 11GB of GDDR6 on a 352-bit memory bus and memory bandwidth in the neighborhood of 620GB/s), 256GB of RAM, and 1TB of HDD. Overall, the experiments took about 2 days in a shared resource setting. We expect that a consumer-grade single-GPU machine (e.g., with a 2080 Ti GPU) could complete the full set of experiments in around 3-4 days, if its full resources were dedicated. The codes were implemented in Python 3.7.3 and PyTorch 1.0.14. We also employ Numpy 1.16.4 and Scipy 1.3.0 in the implementation. Since the datasets used are all public datasets and our methodologies and the hyperparameter settings are explicitly described in Section 3, 4, 5, and A.5, our codes and experiments can be easily reproduced on top of a GPU server. We promise to release our open-source codes on GitHub and maintain a project website with detailed documentation for long-term access by other researchers and end-users after the paper is accepted.

Table 9: Statistics of the datasets

Dataset	#Clients/#Graphs	#Avg. Nodes	#Nodes	#Avg. Edges	#Edges
SNS	3	14,331	14,262 ~ 14,573	51,358	48,105 ~ 53,381
PPI	50	1,767	1,767	32,320	31,179 ~ 32,358
DBLP	20	10,038	9,984 ~ 10,168	56,314	54,891 ~ 60,058

**Datasets.** We study federated graph matching tasks on three representative graph matching benchmark datasets: social networks (SNS) <sup>1</sup>, protein-protein interaction networks (PPI) <sup>2</sup>, and DBLP coauthor graphs (DBLP) <sup>3</sup>. The above three graph datasets are all public datasets, which allow re-

<sup>1</sup><https://www.aminer.cn/cosnet>

<sup>2</sup><http://snap.stanford.edu/ohmnet/>

<sup>3</sup><http://dblp.uni-trier.de/xml/>

searchers to use for non-commercial research and educational purposes. Among three datasets used in the experiment, social networks (SNS), protein-protein interaction networks (PPI), and DBLP coauthor graphs (DBLP) contain 3, 50, 20 different graphs respectively. These three datasets are widely used in training/evaluating the graph matching. The SNS dataset from (Zhang et al., 2015) has 3 different graphs of Flickr, Last.fm, and MySpace. The PPI dataset from (Zitnik & Leskovec, 2017) has 50 different graphs, each representing a tissue with proteins as nodes. As for the DBLP dataset, we select and split the original DBLP dataset into 20 graphs by publication year, ranging from 2002-2022. Thus, most authors occur in all 20 graphs but different graphs contain few emeritus and new authors.

**Training.** For each of the above three datasets, we use one client to maintain only one local graph in the federated setting. We randomly assign the graphs in the three datasets to 3, 50, 20 clients respectively in the experiments. We choose all of these graphs and clients to participate in the training of the models of federated graph matching. For the supervised learning methods, the training data ratio over the above three datasets is all fixed to 20%. We train the models on the training set and test them on the test set for three datasets. In addition, we run each experiment for 3 trials for obtaining more stable results.

**Baselines.** We compare three types of baselines that are most close to the task of federated graph matching: centralized graph matching, ~~federated graph learning~~ and federated domain adaption. (1) **Centralized graph matching baselines.** We compare the UFGM model with six state-of-the-art models. **NextAlign** is a semi-supervised network alignment method that achieves a good trade-off between alignment consistency and alignment disparity (Zhang et al., 2021c). **NetTrans** is an end-to-end supervised graph matching model that learns a composition of nonlinear operations to transform one network to another in a hierarchical manner (Zhang et al., 2020). **CPUGA** is a robust supervised graph alignment model designed with non-sampling learning to distinguish noise from benign data in the given labeled data (Pei et al., 2022). **ASAR-GM** is a robust visual graph matching approach that enlarges the disparity among appearance-similar key-points in graph, orthogonal to de facto adversarial training (Ren et al., 2022). **SeedGNN** is a supervised approach that can learn from a training set how to match unseen graphs with only a few seeds (Yu et al., 2022). **SIGMA** is a semantic-complete graph matching framework that completes mismatched semantics and reformulates the adaptation with graph matching (Li et al., 2022). (2) ~~**Federated graph learning baselines.** We evaluate the UFGM model with six representative federated graph learning architectures. **FedGraphNN** is an open research federated learning system and a benchmark to facilitate GNN-based FL research (He et al., 2021a). **FKGE** is a decentralized scalable learning framework that learns knowledge graph embedding in an asynchronous and peer-to-peer manner while being privacy-preserving (Peng et al., 2021). **SpreadGNN** is a multi-task federated training framework capable of operating in the presence of partial labels and the absence of a central server for GNNs over molecular graphs (He et al., 2022). **SFL** is a structured federated learning framework to learn both the global and personalized models simultaneously using client-wise relation graphs and clients’ private data (Chen et al., 2022b). **FederatedScope-GNN** is an easy-to-use FGL package that provides a unified view for modularizing and expressing FGL algorithms (Wang et al., 2022b). **FedStar** is an FGL framework that extracts and shares the common underlying structure information for inter-graph federated learning tasks (Tan et al., 2022).~~ (2) **Federated domain adaption baselines.** We compare the model performance with four recent federated domain adaption methods. **DualAdapt** aims to align the representations learned among the different nodes with the data distribution of the target node (Peng et al., 2020). **EFDA** extends domain adaptation with the constraints of federated learning to train a model for the target domain and preserve the data privacy of all the source and target domains (Kang et al., 2022). **WSDA** leverages auxiliary information to reduce the risk of federated domain adaption on the target client during local training (Jiang & Koyejo, 2023). **FedKA** aligns features from different clients and those of the target task (Sun et al., 2022).

**Implementation.** For six state-of-the-art centralized graph matching models of NextAlign <sup>4</sup>, NetTrans <sup>5</sup>, CPUGA <sup>6</sup>, ASAR-GM <sup>7</sup>, SeedGNN <sup>8</sup>, and SIGMA <sup>9</sup>, we used the open-source implementation and default parameter settings by the original authors for the experiments. All hyperparameters are standard values from reference codes or prior works. ~~For six representative federated graph learning architectures of FedGraphNN <sup>10</sup>, FKGE <sup>11</sup>, SpreadGNN <sup>12</sup>, SFL <sup>13</sup>, FederatedScope-GNN <sup>14</sup>, and FedStar <sup>15</sup>, we also use the default parameters in the authors' implementation.~~ For four recent federated domain adaption methods of DualAdapt <sup>16</sup>, EFDA <sup>17</sup>, WSDA <sup>18</sup>, and FedKA <sup>19</sup>, we utilized the same model architecture as the official implementation provided by the authors and used the same datasets to validate the performance of these federated graph matching models in all experiments. All models were trained for 2,000 rounds, with a batch size of 500, and a learning rate of 0.05. The above open-source codes from the GitHub are licensed under the MIT License, which only requires preservation of copyright and license notices and includes the permissions of commercial use, modification, distribution, and private use.

For our UFGM model, we performed hyperparameter selection by performing a parameter sweep on sampled graphlet numbers  $O \in \{1, 5, 10, 15, 20\}$ , weight of two types of weak quasi-Newton conditions  $\omega \in \{1, 1.25, 1.5, 1.75, 2\}$ , trust-region radius  $\Delta^s \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ , subgraph size for graphlet feature extraction  $k \in \{1, 2, 5, 7, 9\}$ , training round  $\in \{100, 500, 1,000, 1,500, 2,000\}$ , and learning rate  $\in \{0.001, 0.005, 0.01, 0.05, 0.1, 0.5\}$ . We select the best parameters over 50 epochs of training and evaluate the model at test time. **In our current implementation, we first utilize an efficient matrix generation method (Randall, 1993) to produce a random nonsingular matrix  $\mathbf{K}$  and then orthogonalize it to preserve the distances between the embedding vectors.**

### Hyperparameter settings.

Unless otherwise explicitly stated, we used the following default parameter settings in the experiments.

Table 10: Hyperparameter settings

Parameter	Value
Training data ratio for supervised learning methods	20%
Sampled graphlet numbers $O$	10
Weight of two types of weak quasi-Newton conditions $\omega$	1.5
Trust-region radius $\Delta^s$	0.5
Subgraph size $k$ for graphlet feature extraction	5
Training round	2,000
Batch size for training the model	500
Learning rate	0.05

<sup>4</sup><https://github.com/sizhang92/NextAlign-KDD21>

<sup>5</sup><https://github.com/sizhang92/NetTrans-KDD20>

<sup>6</sup><https://github.com/scpei/CPUGA>

<sup>7</sup><https://github.com/Thinklab-SJTU/ThinkMatch>

<sup>8</sup><https://openreview.net/forum?id=iYvbPx8GTta>

<sup>9</sup><https://github.com/CityU-AIM-Group/SIGMA>

<sup>16</sup><https://drive.google.com/file/d/1OekTpqB6qLfjIE2XUjQPm3F110KDMFc0/view?usp=sharing>

<sup>17</sup><https://github.com/yuetan031/fedstar>

<sup>18</sup>[https://openreview.net/forum?id=\\_1gu0EX0mM3](https://openreview.net/forum?id=_1gu0EX0mM3)

<sup>19</sup><https://github.com/yuweisunn/federated-knowledge-alignment>