Effective Federated Graph Matching

Yang Zhou [1] **Zijie Zhang** [1] **Zeru Zhang** [1] **Lingjuan Lyu** [2] **Wei-Shinn Ku** [1]

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., 2021a; 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., 2022b; 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).

Despite achieving 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. A real application is user account linking in different social networks (Shu et al., 2016; Li et al., 2019a). Since the user information contain many sensitive information, directly uploading the raw user data to the server for centralized graph matching (CGM) gives rise to a serious privacy risk (Zhang et al., 2021a). Another real scenario is financial crime detection on transaction networks with millions to billions of bank customers and transactions (Suzumura et al., 2019; Wang et al., 2019a). Data exchange among clients and server about sensitive bank customer and transfer data should be prohibited for privacy concerns. Recently, US and UK governments launched a privacy-enhancing technology challenge about federated learning for financial crime detection in July 2022 (NSF; IBM). The dataset contains SWIFT transfer data between bank accounts and individual bank account transaction data. Frequent interbank transactions between the same or affili-

[1]Auburn University, USA [2]Sony AI, Japan. Correspondence to: Yang Zhou <yangzhou@auburn.edu>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024. Copyright 2024 by the author(s).

ated entities may be potential money laundering activities. 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 CGM 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. If we choose to conduct both evaluation and optimization of the graph matching model at the server, then the computational capability of each client are unutilized. This will dramatically degrade the algorithm efficiency, especially large-scale graph matching is often time-consuming. 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 identify 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. In conclusion, the server computes the individual loss and the first-order gradients. The clients calculates the second-order Hessians and optimizes the local models. This design is helpful to make full use of the computational capability of each client to improve the FGM efficiency over large graphs.

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.

2. Background

2.1. Supervised Graph Matching

Given a set of S graphs $G = \{G^1, \cdots, G^S\}$. Each graph is denoted as $G^s = (V^s, E^s)$ $(1 \leq s \leq S)$, where $V^s = \{v_1^s, v_2^s, \cdots\}$ 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 ith 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., $D = \{D^{12}, \cdots, D^{1S}, \cdots, D^{(S-1)S}\}$. 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 goal 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 an 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 or largest similarities in the test data are selected as the matching results. This work follows these existing efforts to design the loss function.

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

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.

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

2.2. Federated Graph Matching

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

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

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., $N = N^{12} + \cdots + N^{1S} + \cdots + N^{(S-1)S}$. 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 $M^1, \cdot, M^S$ 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 $\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_R$ 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

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

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 $g_{k1}, \cdots, g_{kO}$ in G, where the size $O << C$. Then we use Horvitz-Thompson inverse probability weighting to estimate the graphlet counts as follows.

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

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 expend 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 $q(v) = \frac{f(\deg(v))}{F}$, 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 $g_{k1}, g_{k1}, \cdots, g_{kO}$ 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 $p(g_k) = \sum_{g_{k-1}} \mathbb{P}(g_k|g_{k-1})p(g_{k-1})$, 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.

$$\begin{aligned} p(g_k) &= \sum_{g_{k-1} \subset g_k} p(g_{k-1}) \frac{\deg_{g_{k-1}}\left(V_{g_k} - V_{g_{k-1}}\right)}{|\mathcal{N}_e(g_{k-1})|} \\ &= \sum_{g_{k-1} \subset g_k} p(g_{k-1}) \frac{|E_{g_k}| - \left|E_{g_{k-1}}\right|}{\sum_{v \in V_{g_{k-1}}} \deg(v) - 2\left|E_{g_{k-1}}\right|} \end{aligned} \tag{6}$$

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 $x_k = \{v_1, v_2, \cdots, v_k\}$. Let $\mathcal{S}(g_k) = [v_1, v_2, \cdots, v_k]$ be the set of all possible node sequences of x_k. Notice that an induced subgraph $h_l(x_k) = \{v_1, v_2, \cdots, v_l, x_k, G\}$ of graph G with the first l nodes $\{v_1, v_2, \cdots, v_l\}$ in x_k must be a connected subgraph for any l $(1 \leq l \leq k)$. Thus, we have

$$\begin{aligned} \mathcal{S}(g_k) &= \{[v_1, \ldots, v_k] | \{v_1, \ldots, v_k\} \\ &= V_{g_k}, g_k | \{v_1, \ldots, v_l\} \text{is connected}\} \end{aligned} \tag{7}$$

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

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

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, $d_1, \cdots, d_k$ be the k highest degrees of nodes in G, and denote $D = \prod_{l=2}^{k-1}(d_1 + \cdots + d_k)$.*

If q for sampling the starting node is the stationary distribution of the node random walk, then the upper bound of the variance $\mathrm{Var}(\tilde{n}_{kr}(G))$ *is*

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

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

It is observed that the variance $\mathrm{Var}(\tilde{n}_{kr}(G))$ 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 $\hat{\mathbf{v}}_i^s$, 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 $\hat{\mathbf{v}}_i^s$, 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.

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

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

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.

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

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.

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

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

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

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).

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

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).

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

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

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

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

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

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

$$\mathcal{A}_{b+1} \otimes z_b^3 = \sum_{i,j,k=1}^{d} \frac{\partial^3 \mathcal{L}^s(M_{b+1})}{\partial M^i \partial M^j \partial M^k} z_b^i z_b^j z_b^k \tag{18}$$

and

$$\mathcal{B}_{b+1} \otimes z_b^4 = \sum_{i,j,k,l=1}^{d} \frac{\partial^4 \mathcal{L}^s(M_{b+1})}{\partial M^i \partial M^j \partial M^k \partial M^l} z_b^i z_b^j z_b^k z_b^l. \tag{19}$$

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

$$\begin{aligned} z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b - \alpha_{b+1}(\omega) z_b^T z_b = \left(\frac{1}{2} - \frac{\omega}{6}\right) \mathcal{A}_{b+1} \otimes z_b^3 - \\ \left(\frac{1}{6} - \frac{\omega}{12}\right) \mathcal{B}_{b+1} \otimes z_b^4 + \mathcal{O}\left(\|z_b\|^5\right) \end{aligned} \tag{20}$$

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

$$u_b(0) - u_b(z_b) \geq \frac{1}{2} \|\nabla \mathcal{L}^s(M_b)\| \min\left\{\Delta^s, \frac{\|\nabla \mathcal{L}^s(M_b)\|}{\alpha_b}\right\} \tag{21}$$

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.

$$\begin{aligned} z^* = \operatorname{argmin} u_b(z) \approx \mathcal{L}^s(M_b) + (\nabla \mathcal{L}^s(M_b))^T z + \\ \frac{1}{2} \alpha_b(\omega) z^T z,\ \text{s.t.} \|z\| \leq \Delta^s \end{aligned} \tag{22}$$

5. Experimental Evaluation

In this section, we have evaluated the performance of our UFGM model and other comparison methods for federated graph matching over serval 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.,

Table 1: Final performance on SNS

Type	Algorithm	$Hits@1$	$Hits@5$	$Hits@10$	$Hits@50$	$Loss$
Centralized Graph Matching	NextAlign	**0.430**	**0.512**	**0.571**	**0.635**	**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
Federated Graph Learning	FedGraphNN	0.051	0.100	0.116	0.161	3.120
	FKGE	0.177	0.205	0.222	0.250	1.086
	SpreadGNN	0.078	0.146	0.175	0.192	**0.189**
	SFL	0.000	0.000	0.000	0.001	4.332
	FederatedScope-GNN	0.000	0.000	0.000	0.001	5.611
	FedStar	0.039	0.071	0.105	0.136	3.770
	UFGM	0.371	0.440	0.411	0.459	0.501

2020a), **CPUGA** (Pei et al., 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.

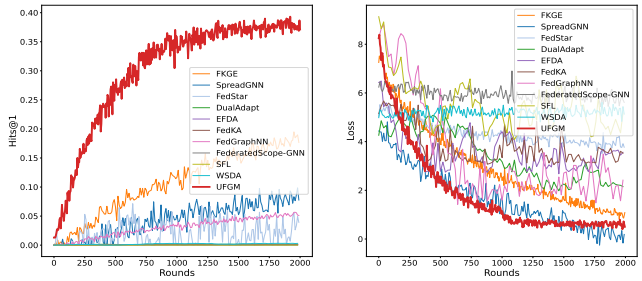
Evaluation metrics. By following the same settings in two representative graph matching models (Yasar & Çatalyürek, 2018; Fey et al., 2020), We employ a popular measure, $Hits@K$, to evaluate and compare our UFGM model to previous lines of work, where $Hits@K$ measures the proportion of correctly matched 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 ten 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 $Hits@50$ 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 eleven 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. How-

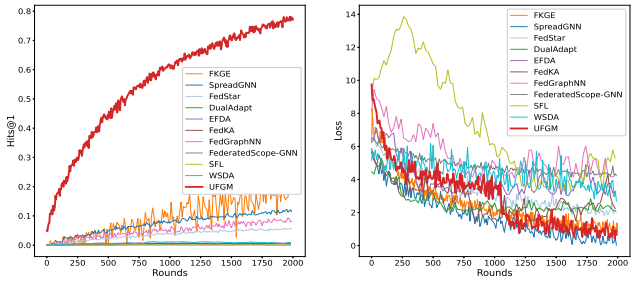
Table 2: Final performance on PPI

Type	Algorithm	$Hits@1$	$Hits@5$	$Hits@10$	$Hits@50$	$Loss$
Centralized Graph Matching	NextAlign	**0.951**	**0.962**	**0.972**	**0.979**	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
Federated Graph Learning	FedGraphNN	0.081	0.132	0.179	0.200	4.259
	FKGE	0.231	0.323	0.352	0.441	0.817
	SpreadGNN	0.115	0.179	0.213	0.236	**0.290**
	SFL	0.000	0.001	0.001	0.002	5.285
	FederatedScope-GNN	0.001	0.001	0.001	0.002	4.259
	FedStar	0.057	0.092	0.137	0.211	2.255
	UFGM	0.771	0.880	0.902	0.930	0.659



(a) $Hits@1$ (b) $Loss$

Figure 1: Convergence on SNS



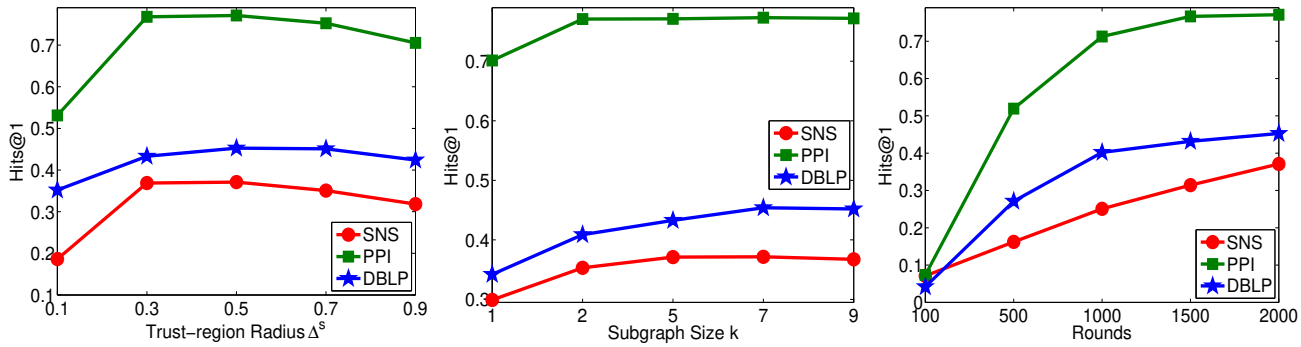
(a) $Hits@1$ (b) $Loss$

Figure 2: Convergence on PPI

ever, 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 eleven federated learning models on two datasets respectively. We have observed obvious 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.

Influence of trust-region radius. Figure 3 (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 learning. 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 learning.



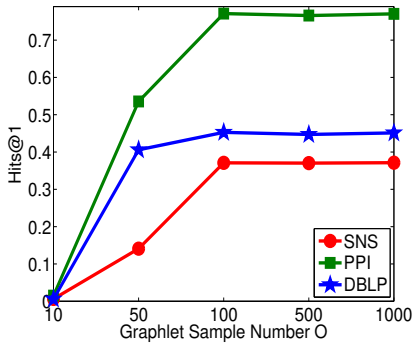
(a) Δ^s (b) k (c) Rounds

Figure 3: Final $Hits@1$ with varying parameters on three datasets

Sensitivity of subgraph size. Figure 3 (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 3 (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 make 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.

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



(a) O

Figure 4: Final $Hits@1$ with varying parameters on three datasets

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.

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.

Acknowledgements

This research is partially sponsored by the National Science Foundation (NSF) under Grant No. OAC- 2313191.

Impact Statement

In this work, the three graph datasets are all open-released datasets (Zhang et al., 2015; Zitnik & Leskovec, 2017; DBL), which allow researchers to use for non-commercial research and educational purposes. These three datasets are widely used in training/evaluating the graph matching. All baseline codes are open-accessed resources that are from the GitHub and 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.

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. Supervised graph matching methods that use the pre-matched node pairs as the training data is improper for the federated graph matching (FGM) scenarios due to privacy risks of direct cross-client information exchange when the graph data are distributed over different clients. Our method combines the unsupervised FGM and the encryption of local raw graph data to provide strong privacy protection for sensitive local data. The proposed separate trust region for pseudo supervised FGM is helpful to enhance the efficiency while maintaining the privacy constraints. Our framework can play an important building block for a wide variety of cross-graph analysis applications that usually require near-zero tolerance of data leaking, such as financial crime detection and user account linking. This paper is primarily of a theoretical nature. We expect our findings to produce positive impact, i.e, significantly improve the privacy of FGM models by utilizing the pseudo supervised FGM. To our best knowledge, we do not envision any immediate negative societal impacts of our results, such as security, privacy, and fairness issues.

An important product of this paper is to explore the possibility of capturing 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. Due to the expensive cost of graphlet enumeration one by one on large-scale graphs. The Monte Carlo Markov Chain (MCMC) technique is leveraged to sample a small number of graphlets for dramatically improving 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.

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.

Baek, J., Jeong, W., Jin, J., Yoon, J., and Hwang, S. J. Personalized subgraph federated learning. *CoRR*, abs/2206.10206, 2022. doi: 10.48550/arXiv.2206.10206. URL https://doi.org/10.48550/arXiv.2206.10206.

Bai, Y., Ding, H., Gu, K., Sun, Y., and Wang, W. 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.

Bao, X., Liu, L., Xiao, N., Zhou, Y., and Zhang, Q. Policy-driven autonomic configuration management for nosql. In *Proceedings of the 2015 IEEE International Conference on Cloud Computing (CLOUD'15)*, pp. 245–252, New York, NY, June 27-July 2 2015.

Bayram, H. C. and Rekik, I. A federated multigraph integration approach for connectional brain template learning. In Syeda-Mahmood, T. F., Li, X., Madabhushi, A., Greenspan, H., Li, Q., Leahy, R. M., Dong, B., and Wang, H. (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.

Caldarola, D., Mancini, M., Galasso, F., Ciccone, M., Rodolà, E., and Caputo, B. 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.

Carlini, N., Tramèr, F., Wallace, E., Jagielski, M., Herbert-Voss, A., Lee, K., Roberts, A., Brown, T. B., Song, D., Erlingsson, Ú., Oprea, A., and Raffel, C. Extracting training data from large language models. In Bailey, M. and Greenstadt, R. (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.

Chakrabarti, S., Singh, H., Lohiya, S., Jain, P., and Mausam. Joint completion and alignment of multilingual knowledge graphs. In Goldberg, Y., Kozareva, Z., and Zhang, Y. (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.

Che, T., Zhang, Z., Zhou, Y., Zhao, X., Liu, J., Jiang, Z., Yan, D., Jin, R., and Dou, D. Federated fingerprint learning with heterogeneous architectures. In *Proceedings of the 22nd IEEE International Conference on Data Mining (ICDM'22)*, pp. 31–40, Orlando, FL, November 28-December 1 2022.

Che, T., Liu, J., Zhou, Y., Ren, J., Zhou, J., Sheng, V. S., Dai, H., and Dou, D. Federated learning of large language models with parameter-efficient prompt tuning and adaptive optimization. In *Proceedings of the 28th Conference on Empirical Methods in Natural Language Processing (EMNLP'23)*, Singapore, December 6-10 2023a.

Che, T., Zhou, Y., Zhang, Z., Lyu, L., Liu, J., Yan, D., Dou, D., and Huan, J. Fast federated machine unlearning with nonlinear functional theory. In *Proceedings of the 40th International Conference on Machine Learning (ICML'23)*, pp. 4241–4268, Honolulu, HI, July 23-29 2023b.

Chen, C., Hu, W., Xu, Z., and Zheng, Z. Fedgl: Federated graph learning framework with global self-supervision. *CoRR*, abs/2105.03170, 2021.

Chen, C., Zhou, J., Zheng, L., Wu, H., Lyu, L., Wu, J., Wu, B., Liu, Z., Wang, L., and Zheng, X. Vertically federated graph neural network for privacy-preserving node classification. In Raedt, L. D. (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.

Chen, F., Long, G., Wu, Z., Zhou, T., and Jiang, J. Personalized federated learning with a graph. In Raedt, L. D. (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.

Chen, H., Yin, H., Sun, X., Chen, T., Gabrys, B., and Musial, K. 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.

Chen, L., Zhao, Y., Lyu, B., Jin, L., Chen, Z., Zhu, S., and Yu, K. 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.

Chen, M., Zhang, W., Yao, Z., Chen, X., Ding, M., Huang, F., and Chen, H. Meta-learning based knowledge extrapolation for knowledge graphs in the federated setting. In Raedt, L. D. (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.

Chen, X., Heimann, M., Vahedian, F., and Koutra, D. Conealign: 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.

Chen, Z., Ta, X., Zhou, Z., and Zhou, Y. A channel aggregation based dynamic pruning method in federated learning. In *Proceedings of the 24th IEEE Global Communications Conference (GLOBECOM'23)*, Kuala Lumpur, Malaysia, December 4-8 2023.

Cheng, H., Zhou, Y., and Yu, J. X. Clustering large attributed graphs: A balance between structural and attribute similarities. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 5(2):1–33, 2011.

Cheng, H., Zhou, Y., Huang, X., and Yu, J. X. Clustering large attributed information networks: An efficient incremental computing approach. *Data Mining and Knowledge Discovery (DMKD)*, 25(3):450–477, 2012.

Chu, X., Fan, X., Yao, D., Zhu, Z., Huang, J., and Bi, J. 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.

Du, X., Yan, J., and Zha, H. 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.

Fan, Z., Mao, C., Wu, Y., and Xu, J. 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.

Farid, M., Leong, W. J., and Hassan, M. A. A new two-step gradient-type method for large-scale unconstrained optimization. *Comput. Math. Appl.*, 59(10):3301–3307, 2010.

Feng, J., Zhang, M., Wang, H., Yang, Z., Zhang, C., Li, Y., and Jin, D. 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.

Fey, M., Lenssen, J. E., Morris, C., Masci, J., and Kriege, N. M. Deep graph matching consensus. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*, 2020.

Gao, Q., Wang, F., Xue, N., Yu, J., and Xia, G. 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.

Guimu Guo, Da Yan, L. Y. J. K. C. L. Z. J. and Zhou, Y. Maximal directed quasi-clique mining. In *Proceedings of the 38th IEEE International Conference on Data Engineering (ICDE'22)*, pp. 1900–1913, Kuala Lumpur, Malaysia, May 9-12 2022.

Guo, L., Zhang, Q., Sun, Z., Chen, M., Hu, W., and Chen, H. Understanding and improving knowledge graph embedding for entity alignment. In Chaudhuri, K., Jegelka, S., Song, L., Szepesvári, C., Niu, G., and Sabato, S. (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.

Guzzi, P. H. and Milenkovic, T. 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.

Haller, S., Feineis, L., Hutschenreiter, L., Bernard, F., Rother, C., Kainmüller, D., Swoboda, P., and Savchynskyy, B. A comparative study of graph matching algorithms in computer vision. In Avidan, S., Brostow, G. J., Cissé, M., Farinella, G. M., and Hassner, T. (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.

He, C., Balasubramanian, K., Ceyani, E., Rong, Y., Zhao, P., Huang, J., Annavaram, M., and Avestimehr, S. Fedgraphnn: A federated learning system and benchmark for graph neural networks. *CoRR*, abs/2104.07145, 2021a. URL https://arxiv.org/abs/2104.07145.

He, C., Balasubramanian, K., Ceyani, E., Rong, Y., Zhao, P., Huang, J., Annavaram, M., and Avestimehr, S. Fedgraphnn: A federated learning system and benchmark for graph neural networks. *CoRR*, abs/2104.07145, 2021b.

He, C., Ceyani, E., Balasubramanian, K., Annavaram, M., and Avestimehr, S. Spreadgnn: Serverless multi-task federated learning for graph neural networks. *CoRR*, abs/2106.02743, 2021c.

He, C., Ceyani, E., Balasubramanian, K., Annavaram, M., and Avestimehr, S. 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.

He, J., Huang, Z., Wang, N., and Zhang, Z. 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.

Heimann, M., Shen, H., Safavi, T., and Koutra, D. 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.

Hong, J., Zhu, Z., Lyu, L., Zhou, Y., Boddeti, V. N., and Zhou, J. International workshop on federated learning for distributed data mining. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'23)*, pp. 5861–5862, Long Beach, CA, August 6-10 2023.

Hu, Y., Liang, W., Wu, R., Xiao, K., Wang, W., Li, X., Liu, J., and Qin, Z. Quantifying and defending against privacy threats on federated knowledge graph embedding. In Ding, Y., Tang, J., Sequeda, J. F., Aroyo, L., Castillo, C., and Houben, G. (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.

Hutschenreiter, L., Haller, S., Feineis, L., Rother, C., Kainmüller, D., and Savchynskyy, B. 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.

Huynh, T. T., Toan, N. T., Tong, V. V., Hoang, T. D., Thang, D. C., Hung, N. Q. V., and Sattar, A. A comparative study on network alignment techniques. *Expert Syst. Appl.*, 140, 2020a.

Huynh, T. T., Tong, V. V., Nguyen, T. T., Yin, H., Weidlich, M., and Hung, N. Q. V. 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.

J. E. Dennis, J. and Wolkowicz, H. Sizing and least-change secant methods. *SIAM Journal on Numerical Analysis*, 30(5):1291–1314, 1993.

Jiang, E. and Koyejo, O. O. Weakly-supervised domain adaptation in federated learning. In *ICLR 2023 Submission*, 2023.

Jiang, M., Jung, T., Karl, R., and Zhao, T. Federated dynamic GNN with secure aggregation. *CoRR*, abs/2009.07351, 2020.

Jiang, Y., Perng, C.-S., Sailer, A., Silva-Lepe, I., Zhou, Y., and Li, T. Csm: A cloud service marketplace for complex service acquisition. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 8(1):1–25, 2016.

Jiang, Z., Rahmani, H., Angelov, P. P., Black, S., and Williams, B. M. 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.

Jin, D., Wang, L., Zheng, Y., Li, X., Jiang, F., Lin, W., and Pan, S. CGMN: A contrastive graph matching network for self-supervised graph similarity learning. In Raedt, L. D. (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, 2022a. doi: 10.24963/ijcai.2022/292. URL https://doi.org/10.24963/ijcai.2022/292.

Jin, J., Ke, Z. T., and Luo, S. Network global testing by counting graphlets. In Dy, J. G. and Krause, A. (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.

Jin, J., Ren, J., Zhou, Y., Lv, L., Liu, J., and Dou, D. Accelerated federated learning with decoupled adaptive optimization. In *Proceedings of the 39th International Conference on Machine Learning (ICML'22)*, pp. 10298–10322, Baltimore, MD, July 17-23 2022b.

Jin, R., Li, D., Gao, J., Liu, Z., Chen, L., and Zhou, Y. Towards a better understanding of linear models for recommendation. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'21)*, pp. 776–785, Virtual Event, Singapore, August 14-18 2021.

Kang, H., Li, Z., and Zhang, Q. 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.

Karimireddy, S. P., Kale, S., Mohri, M., Reddi, S. J., Stich, S. U., and Suresh, A. T. 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.

Karimireddy, S. P., Jaggi, M., Kale, S., Mohri, M., Reddi, S. J., Stich, S. U., and Suresh, A. T. 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.

Ke, C. and Honorio, J. Federated myopic community detection with one-shot communication. *CoRR*, abs/2106.07255, 2021.

Kelley, B. P., Sharan, R., Karp, R. M., Sittler, T., Root, D. E., Stockwell, B. R., and Ideker, T. 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.

Khalil, J., Yan, D., Yuan, L., Han, J., Adhikari, S., Long, C., and Zhou, Y. Fsm-explorer: An interactive tool for frequent subgraph pattern mining from a big graph. In *Proceedings of the 40th IEEE International Conference on Data Engineering (ICDE'24)*, Utrecht, Netherlands, May 13-16 2024.

Kipf, T. N. and Welling, M. 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.

Kondor, R., Shervashidze, N., and Borgwardt, K. M. The graphlet spectrum. In Danyluk, A. P., Bottou, L., and Littman, M. L. (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.

Kong, X., Zhang, J., and Yu, P. S. 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.

Lalitha, A., Kilinc, O. C., Javidi, T., and Koushanfar, F. Peer-to-peer federated learning on graphs. *CoRR*, abs/1901.11173, 2019. URL http://arxiv.org/abs/1901.11173.

Lee, K., Liu, L., Tang, Y., Zhang, Q., and Zhou, Y. Efficient and customizable data partitioning framework for distributed big rdf data processing in the cloud. In *Proceedings of the 2013 IEEE International Conference on Cloud Computing (CLOUD'13)*, pp. 327–334, Santa Clara, CA, June 27-July 2 2013.

Lee, K., Liu, L., Schwan, K., Pu, C., Zhang, Q., Zhou, Y., Yigitoglu, E., and Yuan, P. Scaling iterative graph computations with graphmap. In *Proceedings of the 27th IEEE international conference for High Performance Computing, Networking, Storage and Analysis (SC'15)*, pp. 57:1–57:12, Austin, TX, November 15-20 2015.

Lee, K., Liu, L., Ganti, R. L., Srivatsa, M., Zhang, Q., Zhou, Y., and Wang, Q. Lightwieight indexing and querying services for big spatial data. *IEEE Transactions on Services Computing (TSC)*, 12(3):343–355, 2019.

Li, C., Wang, S., Yu, P. S., Zheng, L., Zhang, X., Li, Z., and Liang, Y. 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.

Li, C., Wang, S., Wang, H., Liang, Y., Yu, P. S., Li, Z., and Wang, W. 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.

Li, C., Wang, S., Wang, Y., Yu, P. S., Liang, Y., Liu, Y., and Li, Z. 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.

Li, W., Liu, X., and Yuan, Y. 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.

Li, X., Zhang, W., Li, R.-H., Zhao, Y., Zhu, Y., and Wang, G. A new paradigm for federated structure non-iid subgraph learning. In *ICLR 2023 Submission*, 2023.

Liu, C., Zhang, S., Yang, X., and Yan, J. Self-supervised learning of visual graph matching. In Avidan, S., Brostow, G. J., Cissé, M., Farinella, G. M., and Hassner, T. (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.

Liu, J., Huang, J., Zhou, Y., Li, X., Ji, S., Xiong, H., and Dou, D. From distributed machine learning to federated learning: A survey. *Knowledge and Information Systems (KAIS)*, 64(4):885–917, 2022b.

Liu, J., Jia, J., Ma, B., Zhou, C., Zhou, J., Zhou, Y., Dai, H., and Dou, D. Multi-job intelligent scheduling with cross-device federated learning. *IEEE Transactions on Parallel and Distributed Systems (TPDS)*, 34(2):535–551, 2023.

Liu, J., Che, T., Zhou, Y., Jin, R., Dai, H., Dou, D., and Valduriez, P. Aedfl: Efficient asynchronous decentralized federated learning with heterogeneous devices. In *Proceedings of the 24th SlAM International Conference on Data Mining (SDM'24)*, Houston, TX, April 18-20 2024a.

Liu, J., Jia, J., Che, T., Huo, C., Ren, J., Zhou, Y., Dai, H., and Dou, D. Fedasmu: Efficient asynchronous federated learning with dynamic staleness-aware model update. In *Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI'24)*, Vancouver, Canada, February 20-27 2024b.

Liu, L., Cheung, W. K., Li, X., and Liao, L. 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.

Liu, L., Hughes, M. C., Hassoun, S., and Liu, L. Stochastic iterative graph matching. In Meila, M. and Zhang, T. (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.

Liu, S., Wang, S., Zhu, F., Zhang, J., and Krishnan, R. 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.

Liu, W., Chen, L., Chen, Y., and Zhang, W. Accelerating federated learning via momentum gradient descent. *IEEE Trans. Parallel Distributed Syst.*, 31(8):1754–1766, 2020.

Liu, W., Qian, H., Zhang, C., Xie, J., Shen, Z., and Zheng, N. 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, 2022c. URL https://ojs.aaai.org/index.php/AAAI/article/view/20329.

Liu, X., Hong, H., Wang, X., Chen, Z., Kharlamov, E., Dong, Y., and Tang, J. Selfkg: Self-supervised entity alignment in knowledge graphs. In Laforest, F., Troncy, R., Simperl, E., Agarwal, D., Gionis, A., Herman, I., and Médini, L. (eds.), *WWW '22: The ACM Web Conference 2022, Virtual Event, Lyon, France, April 25 - 29, 2022*, pp. 860–870. ACM, 2022d. doi: 10.1145/3485447.3511945. URL https://doi.org/10.1145/3485447.3511945.

Liu, Y., Ding, H., Chen, D., and Xu, J. 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.

Lyu, G., Wu, Y., and Feng, S. Deep graph matching for partial label learning. In Raedt, L. D. (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.

Malmi, E., Gionis, A., and Terzi, E. 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.

Man, T., Shen, H., Liu, S., Jin, X., and Cheng, X. 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.

Mei, G., Guo, Z., Liu, S., and Pan, L. 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.

Meng, C., Rambhatla, S., and Liu, Y. Cross-node federated graph neural network for spatio-temporal data modeling. In Zhu, F., Ooi, B. C., and Miao, C. (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.

Meng, L., Ren, Y., Zhang, J., Ye, F., and Yu, P. S. 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.

Mitra, A., Jaafar, R., Pappas, G., and Hassani, H. 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.

Mu, X., Zhu, F., Lim, E., Xiao, J., Wang, J., and Zhou, Z. 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.

Nassar, H., Veldt, N., Mohammadi, S., Grama, A., and Gleich, D. F. 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.

Ni, X., Xu, X., Lyu, L., Meng, C., and Wang, W. 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.

Park, J., Tran, C., Shin, W., and Cao, X. Gradalign+: Empowering gradual network alignment using attribute augmentation. In Hasan, M. A. and Xiong, L. (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.

Pei, S., Yu, L., Yu, G., and Zhang, X. Graph alignment with noisy supervision. In Laforest, F., Troncy, R., Simperl, E., Agarwal, D., Gionis, A., Herman, I., and Médini, L. (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.

Peng, H., Li, H., Song, Y., Zheng, V. W., and Li, J. Differentially private federated knowledge graphs embedding. In Demartini, G., Zuccon, G., Culpepper, J. S., Huang, Z., and Tong, H. (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.

Peng, X., Huang, Z., Zhu, Y., and Saenko, K. 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.

Qiao, Y., Wu, Y., Duo, F., Lin, W., and Yang, J. Siamese neural networks for user identity linkage through web browsing. *IEEE Trans. Neural Networks Learn. Syst.*, 31 (8):2741–2751, 2020.

Qin, K. K., Salim, F. D., Ren, Y., Shao, W., Heimann, M., and Koutra, D. 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.

Qu, J., Ling, H., Zhang, C., Lyu, X., and Tang, Z. Adaptive edge attention for graph matching with outliers. In Zhou, Z. (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.

Qu, L., Tang, N., Zheng, R., Nguyen, Q. V. H., Huang, Z., Shi, Y., and Yin, H. Semi-decentralized federated ego graph learning for recommendation. In Ding, Y., Tang, J., Sequeda, J. F., Aroyo, L., Castillo, C., and Houben, G. (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.

Qu, W., Yan, D., Guo, G., Wang, X., Zou, L., and Zhou, Y. Parallel mining of frequent subtree patterns. In *Software Foundations for Data Interoperability and Large Scale Graph Data Analytics - 4th International Workshop, SFDI 2020, and 2nd International Workshop, LSGDA 2020, held in Conjunction with VLDB 2020*, pp. 18–32, Tokyo, Japan, September 4 2020.

Rácz, M. Z. and Sridhar, A. Correlated stochastic block models: Exact graph matching with applications to recovering communities. In Ranzato, M., Beygelzimer, A., Dauphin, Y. N., Liang, P., and Vaughan, J. W. (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.

Randall, D. Efficient generation of random nonsingular matrices. *Random Struct. Algorithms*, 4(1):111–118, 1993.

Reddi, S. J., Charles, Z., Zaheer, M., Garrett, Z., Rush, K., Konečný, J., Kumar, S., and McMahan, H. B. Adaptive federated optimization. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*, 2021.

Ren, F., Zhang, Z., Zhang, J., Su, S., Sun, L., Zhu, G., and Guo, C. 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.

Ren, J., Zhou, Y., Jin, R., Zhang, Z., Dou, D., and Wang, P. Dual adversarial learning based network alignment. In *Proceedings of the 19th IEEE International Conference on Data Mining (ICDM'19)*, pp. 1288–1293, Beijing, China, November 8-11 2019a.

Ren, J., Zhang, Z., Jin, J., Zhao, X., Wu, S., Zhou, Y., Shen, Y., Che, T., Jin, R., and Dou, D. Integrated defense for resilient graph matching. In *Proceedings of the 38th International Conference on Machine Learning (ICML'21)*, pp. 8982–8997, Virtual Event, July 18-24 2021.

Ren, Q., Bao, Q., Wang, R., and Yan, J. 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.

Ren, Y., Aggarwal, C. C., and Zhang, J. 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, 2019b.

Rizk, E. and Sayed, A. H. 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.

Shervashidze, N., Vishwanathan, S. V. N., Petri, T., Mehlhorn, K., and Borgwardt, K. M. Efficient graphlet kernels for large graph comparison. In Dyk, D. A. V. and Welling, M. (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.

Shu, K., Wang, S., Tang, J., Zafarani, R., and Liu, H. User identity linkage across online social networks: A review. *SIGKDD Explorations*, 18(2):5–17, 2016.

Singh, R., Xu, J., and Berger, B. 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.

Soufiani, H. A. and Airoldi, E. M. Graphlet decomposition of a weighted network. In Lawrence, N. D. and Girolami, M. A. (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.

Su, Z., Liu, L., Li, M., Fan, X., and Zhou, Y. Servicetrust: Trust management in service provision networks. In *Proceedings of the 10th IEEE International Conference on Services Computing (SCC'13)*, pp. 272–279, Santa Clara, CA, June 27-July 2 2013.

Su, Z., Liu, L., Li, M., Fan, X., and Zhou, Y. Reliable and resilient trust management in distributed service provision networks. *ACM Transactions on the Web (TWEB)*, 9(3): 1–37, 2015.

Sun, Y., Chong, N. S. T., and Ochiai, H. 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.

Sun, Z., Wang, C., Hu, W., Chen, M., Dai, J., Zhang, W., and Qu, Y. 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.

Suzumura, T., Zhou, Y., Barcardo, N., Ye, G., Houck, K., Kawahara, R., Anwar, A., Stavarache, L. L., Klyashtorny, D., Ludwig, H., and Bhaskaran, K. Towards federated graph learning for collaborative financial crimes detection. *CoRR*, abs/1909.12946, 2019.

Tan, H., Wang, C., Wu, S., Wang, T., Zhang, X., and Liu, C. 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.

Tan, Y., Liu, Y., Long, G., Jiang, J., Lu, Q., and Zhang, C. 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.

Tian, Z., Ding, Y., Zhang, R., Liu, J., and Ren, K. Towards federated learning of deep graph neural networks. In *ICLR 2023 Submission*, 2023.

Tu, K., Li, J., Towsley, D., Braines, D., and Turner, L. D. gl2vec: learning feature representation using graphlets for directed networks. In Spezzano, F., Chen, W., and Xiao, X. (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.

Vijayan, V. and Milenkovic, T. Multiple network alignment via multimagna++. *IEEE/ACM Trans. Comput. Biology Bioinform.*, 15(5):1669–1682, 2018.

Vijayan, V., Gu, S., Krebs, E. T., Meng, L., and Milenkovic, T. Pairwise versus multiple global network alignment. *IEEE Access*, 8:41961–41974, 2020.

Wang, B., Li, A., Li, H., and Chen, Y. Graphfl: A federated learning framework for semi-supervised node classification on graphs. *CoRR*, abs/2012.04187, 2020a.

Wang, B., Li, A., Pang, M., Li, H., and Chen, Y. Graphfl: A federated learning framework for semi-supervised node classification on graphs. In Zhu, X., Ranka, S., Thai, M. T., Washio, T., and Wu, X. (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.

Wang, C., Wang, Y., Zhao, Z., Qin, D., Luo, X., and Qin, T. Credible seed identification for large-scale structural network alignment. *Data Min. Knowl. Discov.*, 34(6): 1744–1776, 2020b.

Wang, C., Chen, B., Li, G., and Wang, H. FL-AGCNS: federated learning framework for automatic graph convolutional network search. *CoRR*, abs/2104.04141, 2021a.

Wang, D., Qi, Y., Lin, J., Cui, P., Jia, Q., Wang, Z., Fang, Y., Yu, Q., Zhou, J., and Yang, S. A semi-supervised graph attentive network for financial fraud detection. In Wang, J., Shim, K., and Wu, X. (eds.), *2019 IEEE International Conference on Data Mining, ICDM 2019, Beijing, China, November 8-11, 2019*, pp. 598–607. IEEE, 2019a.

Wang, F., Xue, N., Yu, J., and Xia, G. 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.

Wang, J., Xu, Z., Garrett, Z., Charles, Z., Liu, L., and Joshi, G. 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.

Wang, R., Yan, J., and Yang, X. 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.

Wang, S., Li, X., Ye, Y., Feng, S., Lau, R. Y. K., Huang, X., and Du, X. Anchor link prediction across attributed networks via network embedding. *Entropy*, 21(3):254, 2019b.

Wang, T., Jiang, Z., and Yan, J. 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, 2020e.

Wang, T., Liu, H., Li, Y., Jin, Y., Hou, X., and Ling, H. 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, 2020f.

Wang, Y., Feng, C., Chen, L., Yin, H., Guo, C., and Chu, Y. User identity linkage across social networks via linked heterogeneous network embedding. *World Wide Web*, 22 (6):2611–2632, 2019c.

Wang, Z., Kuang, W., Xie, Y., Yao, L., Li, Y., Ding, B., and Zhou, J. Federatedscope-gnn: Towards a unified, comprehensive and efficient package for federated graph learning. In Zhang, A. and Rangwala, H. (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.

Weighill, D. A., Guebila, M. B., Lopes-Ramos, C. M., Glass, K., Quackenbush, J., Platig, J., and Burkholz, R. 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.

Wu, C., Wu, F., Cao, Y., Huang, Y., and Xie, X. Fedgnn: 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 2021a.

Wu, S., Li, Y., Zhang, D., Zhou, Y., and Wu, Z. Diverse and informative dialogue generation with context-specific commonsense knowledge awareness. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL'20)*, pp. 5811–5820, Online, July 5-10 2020a.

Wu, S., Li, Y., Wang, M., Zhang, D., Zhou, Y., and Wu, Z. More is better: Enhancing open-domain dialogue generation via multi-source heterogeneous knowledge. In *Proceedings of the 26th Conference on Empirical Methods in Natural Language Processing (EMNLP'21)*, pp. 2286–2300, Virtual Event / Punta Cana, Dominican Republic, November 7-11 2021b.

Wu, S., Li, Y., Zhang, D., Zhou, Y., and Wu, Z. Topicka: Generating commonsense knowledge-aware dialogue responses towards the recommended topic fact. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI'20)*, pp. 3766–3772, Online, January 7-15 2021c.

Wu, S., Wang, M., Zhang, D., Zhou, Y., Li, Y., and Wu, Z. Knowledge-aware dialogue generation via hierarchical infobox accessing and infobox-dialogue interaction graph network. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI'21)*, pp. 3964–3970, Virtual Event / Montreal, Canada, August 19-27 2021d.

Wu, Y., Liu, X., Feng, Y., Wang, Z., and Zhao, D. 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, 2020b.

xiang Yuan, Y. A modified bfgs algorithm for unconstrained optimization. *IMA Journal of Numerical Analysis*, 11(3): 325–332, 1991.

Xie, H., Ma, J., Xiong, L., and Yang, C. Federated graph classification over non-iid graphs. In Ranzato, M., Beygelzimer, A., Dauphin, Y. N., Liang, P., and Vaughan, J. W. (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.

Xie, H., Xiong, L., and Yang, C. Federated node classification over graphs with latent link-type heterogeneity. In Ding, Y., Tang, J., Sequeda, J. F., Aroyo, L., Castillo, C., and Houben, G. (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.

Xie, W., Mu, X., Lee, R. K., Zhu, F., and Lim, E. 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.

Xin, K., Sun, Z., Hua, W., Hu, W., Qu, J., and Zhou, X. Large-scale entity alignment via knowledge graph merging, partitioning and embedding. In Hasan, M. A. and Xiong, L. (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.

Xu, H., Luo, D., Zha, H., and Carin, L. 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.

Xu, K., Wang, L., Yu, M., Feng, Y., Song, Y., Wang, Z., and Yu, D. 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.

Yan, D., Qu, W., Guo, G., Wang, X., and Zhou, Y. Prefixfpm: A parallel framework for general-purpose mining of frequent and closed patterns. *The VLDB Journal (VLDBJ)*, 31(2):253–286, 2022a.

Yan, D., Zhou, Y., and Guo, G. Think-like-a-task programming model. In Albert Zomaya, J. T. and Sakr, S. (eds.), *Encyclopedia of Big Data Technologies*. Springer, 2022b.

Yan, D., Zhou, Y., Guo, G., and Liu, H. Parallel graph processing. In Albert Zomaya, J. T. and Sakr, S. (eds.), *Encyclopedia of Big Data Technologies*. Springer, 2022c.

Yan, J., Yang, S., and Hancock, E. R. 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.

Yang, X., Liu, Z., and Qiaoxu, H. Incorporating discrete constraints into random walk-based graph matching. *IEEE Trans. Syst. Man Cybern. Syst.*, 50(4):1406–1416, 2020.

Yasar, A. and Çatalyürek, Ü. V. 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.

Ye, Z., Yenamandra, T., Bernard, F., and Cremers, D. 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.

Yu, L., Xu, J., and Lin, X. 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.

Yu, T., Wang, R., Yan, J., and Li, B. 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.

Yu, T., Wang, R., Yan, J., and Li, B. Deep latent graph matching. In Meila, M. and Zhang, T. (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.

Yuan, L., Ahmad, A., Yan, D., Han, J., Adhikari, S., Yu, X., and Zhou, Y. G2-aimd: A memory-efficient subgraph-centric framework for efficient subgraph search on gpus. In *Proceedings of the 40th IEEE International Conference on Data Engineering (ICDE'24)*, Utrecht, Netherlands, May 13-16 2024a.

Yuan, L., Yan, D., Han, J., Ahmad, A., Zhou, Y., and Jiang, Z. Faster depth-first subgraph matching on gpus. In *Proceedings of the 40th IEEE International Conference on Data Engineering (ICDE'24)*, Utrecht, Netherlands, May 13-16 2024b.

Zhang, H., Shen, T., Wu, F., Yin, M., Yang, H., and Wu, C. Federated graph learning - A position paper. *CoRR*, abs/2105.11099, 2021a.

Zhang, H., Liu, J., Jia, J., Zhou, Y., Dai, H., and Dou, D. Fedduap: Federated learning with dynamic update and adaptive pruning using shared data on the server. In *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI'22)*, pp. 2776–2782, Messe Wien, Vienna, Austria, July 23-29 2022a.

Zhang, J. and Yu, P. S. 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.

Zhang, J., Kong, X., and Yu, P. S. 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, 2013a.

Zhang, K., Yang, C., Li, X., Sun, L., and Yiu, S. Subgraph federated learning with missing neighbor generation. In Ranzato, M., Beygelzimer, A., Dauphin, Y. N., Liang, P., and Vaughan, J. W. (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.

Zhang, K., Wang, Y., Wang, H., Huang, L., Yang, C., Chen, X., and Sun, L. Efficient federated learning on knowledge graphs via privacy-preserving relation embedding aggregation. In Goldberg, Y., Kozareva, Z., and Zhang, Y. (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, 2022b. URL https://aclanthology.org/2022.findings-emnlp.43.

Zhang, Q., Liu, L., Ren, Y., Lee, K., Tang, Y., Zhao, X., and Zhou, Y. Residency aware inter-vm communication in virtualized cloud: Performance measurement and analysis. In *Proceedings of the 2013 IEEE International Conference on Cloud Computing (CLOUD'13)*, pp. 204–211, Santa Clara, CA, June 27-July 2 2013b.

Zhang, Q., Liu, L., Lee, K., Zhou, Y., Singh, A., Mandagere, N., Gopisetty, S., and Alatorre, G. Improving hadoop service provisioning in a geographically distributed cloud. In *Proceedings of the 2014 IEEE International Conference on Cloud Computing (CLOUD'14)*, pp. 432–439, Anchorage, AK, June 27-July 2 2014.

Zhang, S. and Tong, H. 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.

Zhang, S. and Tong, H. 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.

Zhang, S., Tong, H., Maciejewski, R., and Eliassi-Rad, T. Multilevel network alignment. In *The World Wide Web Conference, WWW 2019, San Francisco, CA, USA, May 13-17, 2019*, pp. 2344–2354, 2019.

Zhang, S., Tong, H., Xia, Y., Xiong, L., and Xu, J. Nettrans: Neural cross-network transformation. In Gupta, R., Liu, Y., Tang, J., and Prakash, B. A. (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, 2020a. doi: 10.1145/3394486.3403141. URL https://doi.org/10.1145/3394486.3403141.

Zhang, S., Tong, H., Jin, L., Xia, Y., and Guo, Y. Balancing consistency and disparity in network alignment. In Zhu, F., Ooi, B. C., and Miao, C. (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.

Zhang, X., Hong, M., and Chen, J. 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.

Zhang, Y., Tang, J., Yang, Z., Pei, J., and Yu, P. S. 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.

Zhang, Z., Zhang, Z., Zhou, Y., Shen, Y., Jin, R., and Dou, D. Adversarial attacks on deep graph matching. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020 (NeurIPS'20)*, Virtual, December 6-12 2020b.

Zhang, Z., Jin, J., Zhang, Z., Zhou, Y., Zhao, X., Ren, J., Liu, J., Wu, L., Jin, R., and Dou, D. Validating the lottery ticket hypothesis with inertial manifold theory. In *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021 (NeurIPS'21)*, pp. 30196–30210, Virtual, December 6-14 2021d.

Zhang, Z., Zhang, Z., Zhou, Y., Wu, L., Wu, S., Han, X., Dou, D., Che, T., and Yan, D. Adversarial attack against cross-lingual knowledge graph alignment. In *Proceedings of the 26th Conference on Empirical Methods in Natural Language Processing (EMNLP'21)*, pp. 5320–5337, Virtual Event / Punta Cana, Dominican Republic, November 7-11 2021e.

Zhao, K., Tu, S., and Xu, L. 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.

Zheng, D., Wang, M., Gan, Q., Zhang, Z., and Karypis, G. Learning graph neural networks with deep graph library. In Seghrouchni, A. E. F., Sukthankar, G., Liu, T., and van Steen, M. (eds.), *Companion of The 2020 Web Conference 2020, Taipei, Taiwan, April 20-24, 2020*, pp. 305–306. ACM / IW3C2, 2020.

Zheng, V. W., Sha, M., Li, Y., Yang, H., Fang, Y., Zhang, Z., Tan, K., and Chang, K. C. 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.

Zhong, Z., Cao, Y., Guo, M., and Nie, Z. 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.

Zhou, C., Liu, J., Jia, J., Zhou, J., Zhou, Y., Dai, H., and Dou, D. Efficient device scheduling with multi-job federated learning. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI'22)*, pp. 9971–9979, Vancouver, Canada, February 22-March 1 2022a.

Zhou, F., Liu, L., Zhang, K., Trajcevski, G., Wu, J., and Zhong, T. 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.

Zhou, F., Wen, Z., Trajcevski, G., Zhang, K., Zhong, T., and Liu, F. 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, 2019a.

Zhou, F., Cao, C., Trajcevski, G., Zhang, K., Zhong, T., and Geng, J. 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.

Zhou, J. and Fan, J. 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.

Zhou, J., Chen, C., Zheng, L., Wu, H., Wu, J., Zheng, X., Wu, B., Liu, Z., and Wang, L. Vertically federated graph neural network for privacy-preserving node classification. *CoRR*, abs/2005.11903, 2020b.

Zhou, T., Lim, E., Lee, R. K., Zhu, F., and Cao, J. 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.

Zhou, X., Liang, X., Du, X., and Zhao, J. Structure based user identification across social networks. *IEEE Trans. Knowl. Data Eng.*, 30(6):1178–1191, 2018b.

Zhou, Y. *Innovative Mining, Processing, and Application of Big Graphs*. PhD thesis, Georgia Institute of Technology, Atlanta, GA, USA, 2017.

Zhou, Y. and Liu, L. Clustering analysis in large graphs with rich attributes. In Holmes, D. E. and Jain, L. C. (eds.), *Data Mining: Foundations and Intelligent Paradigms: Volume 1: Clustering, Association and Classification*. Springer, 2012.

Zhou, Y. and Liu, L. Social influence based clustering of heterogeneous information networks. In *Proceedings of the 19th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'13)*, pp. 338–346, Chicago, IL, August 11-14 2013.

Zhou, Y. and Liu, L. Activity-edge centric multi-label classification for mining heterogeneous information networks. In *Proceedings of the 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'14)*, pp. 1276–1285, New York, NY, August 24-27 2014.

Zhou, Y. and Liu, L. Social influence based clustering and optimization over heterogeneous information networks. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 10(1):1–53, 2015.

Zhou, Y. and Liu, L. Approximate deep network embedding for mining large-scale graphs. In *Proceedings of the 2019 IEEE International Conference on Cognitive Machine Intelligence (CogMI'19)*, pp. 53–60, Los Angeles, CA, December 12-14 2019.

Zhou, Y., Cheng, H., and Yu, J. X. Graph clustering based on structural/attribute similarities. *Proceedings of the VLDB Endowment (PVLDB)*, 2(1):718–729, 2009.

Zhou, Y., Cheng, H., and Yu, J. X. Clustering large attributed graphs: An efficient incremental approach. In *Proceedings of the 10th IEEE International Conference on Data Mining (ICDM'10)*, pp. 689–698, Sydney, Australia, December 14-17 2010.

Zhou, Y., Liu, L., Perng, C.-S., Sailer, A., Silva-Lepe, I., and Su, Z. Ranking services by service network structure and service attributes. In *Proceedings of the 20th International Conference on Web Service (ICWS'13)*, pp. 26–33, Santa Clara, CA, June 27-July 2 2013.

Zhou, Y., Liu, L., and Buttler, D. Integrating vertex-centric clustering with edge-centric clustering for meta path graph analysis. In *Proceedings of the 21st ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'15)*, pp. 1563–1572, Sydney, Australia, August 10-13 2015a.

Zhou, Y., Liu, L., Lee, K., Pu, C., and Zhang, Q. Fast iterative graph computation with resource aware graph parallel abstractions. In *Proceedings of the 24th ACM Symposium on High-Performance Parallel and Distributed Computing (HPDC'15)*, pp. 179–190, Portland, OR, June 15-19 2015b.

Zhou, Y., Liu, L., Lee, K., and Zhang, Q. Graphtwist: Fast iterative graph computation with two-tier optimizations. *Proceedings of the VLDB Endowment (PVLDB)*, 8(11): 1262–1273, 2015c.

Zhou, Y., Liu, L., Pu, C., Bao, X., Lee, K., Palanisamy, B., Yigitoglu, E., and Zhang, Q. Clustering service networks with entity, attribute and link heterogeneity. In *Proceedings of the 22nd International Conference on Web Service (ICWS'15)*, pp. 257–264, New York, NY, June 27-July 2 2015d.

Zhou, Y., Liu, L., Seshadri, S., and Chiu, L. Analyzing enterprise storage workloads with graph modeling and clustering. *IEEE Journal on Selected Areas in Communications (JSAC)*, 34(3):551–574, 2016.

Zhou, Y., Amimeur, A., Jiang, C., Dou, D., Jin, R., and Wang, P. Density-aware local siamese autoencoder network embedding with autoencoder graph clustering. In *Proceedings of the 2018 IEEE International Conference on Big Data (BigData'18)*, pp. 1162–1167, Seattle, WA, December 10-13 2018c.

Zhou, Y., Wu, S., Jiang, C., Zhang, Z., Dou, D., Jin, R., and Wang, P. Density-adaptive local edge representation learning with generative adversarial network multi-label edge classification. In *Proceedings of the 18th IEEE International Conference on Data Mining (ICDM'18)*, pp. 1464–1469, Singapore, November 17-20 2018d.

Zhou, Y., Jiang, C., Zhang, Z., Dou, D., Jin, R., and Wang, P. Integrating local vertex/edge embedding via deep matrix fusion and siamese multi-label classification. In *Proceedings of the 2019 IEEE International Conference on Big Data (BigData'19)*, pp. 1018–1027, Los Angeles, CA, December 9-12 2019b.

Zhou, Y., Ling Liu, Qi Zhang, K. L., and Palanisamy, B. Enhancing collaborative filtering with multi-label classification. In *Proceedings of the 2019 International Conference on Computational Data and Social Networks (CSoNet'19)*, pp. 323–338, Ho Chi Minh City, Vietnam, November 18-20 2019c.

Zhou, Y., Ren, J., Wu, S., Dou, D., Jin, R., Zhang, Z., and Wang, P. Semi-supervised classification-based local vertex ranking via dual generative adversarial nets. In *Proceedings of the 2019 IEEE International Conference on Big Data (BigData'19)*, pp. 1267–1273, Los Angeles, CA, December 9-12 2019d.

Zhou, Y., Liu, L., Lee, K., Palanisamy, B., and Zhang, Q. Improving collaborative filtering with social influence over heterogeneous information networks. *ACM Transactions on Internet Technology (TOIT)*, 20(4):36:1–36:29, 2020d.

Zhou, Y., Ren, J., Jin, R., Zhang, Z., Dou, D., and Yan, D. Unsupervised multiple network alignment with multinominal gan and variational inference. In *Proceedings of the 2020 IEEE International Conference on Big Data (BigData'20)*, pp. 868–877, Atlanta, GA, December 10-13 2020e.

Zhou, Y., Zhang, Z., Wu, S., Sheng, V., Han, X., Zhang, Z., and Jin, R. Robust network alignment via attack signal scaling and adversarial perturbation elimination. In *Proceedings of the 30th Web Conference (WWW'21)*, pp. 3884–3895, Virtual Event / Ljubljana, Slovenia, April 19-23 2021.

Zhou, Y., Ren, J., Jin, R., Zhang, Z., Zheng, J., Jiang, Z., Yan, D., and Dou, D. Unsupervised adversarial network alignment with reinforcement learning. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 16(3): 50:1–50:29, 2022b.

Zhu, J., Koutra, D., and Heimann, M. CAPER: coarsen, align, project, refine - A general multilevel framework for network alignment. In Hasan, M. A. and Xiong, L. (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.

Zhu, Q., Zhou, X., Wu, J., Tan, J., and Guo, L. 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.

Zhu, R., Luo, X., Ma, M., and Wang, P. Adaptive graph convolutional network for knowledge graph entity alignment. In Goldberg, Y., Kozareva, Z., and Zhang, Y. (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.

Zhu, X., Li, G., and Hu, W. Heterogeneous federated knowledge graph embedding learning and unlearning. In Ding, Y., Tang, J., Sequeda, J. F., Aroyo, L., Castillo, C., and Houben, G. (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.

Zitnik, M. and Leskovec, J. Predicting multicellular function through multi-layer tissue networks. *Bioinform.*, 33(14): i190–i198, 2017.

A. Appendix

A.1. Related Work

Centralized Graph Matching. Graph machine learning has attracted active research in recent years (Zhou et al., 2009; 2010; Cheng et al., 2011; Zhou & Liu, 2012; Cheng et al., 2012; Zhou & Liu, 2013; 2014; Zhou et al., 2015d; Zhou & Liu, 2015; Zhou et al., 2015a; 2013; 2016; Zhou, 2017; Zhou et al., 2019d; 2018d;c; 2019b; Zhou & Liu, 2019; Zhou et al., 2019c; 2020d; Wu et al., 2020a; 2021c; Jin et al., 2021; Wu et al., 2021b;d). 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., 2020e; 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; 2019a; Zhou et al., 2020e; Zhang et al., 2020b; Zhou et al., 2021; Ren et al., 2021; Zhang et al., 2021e; Zhou et al., 2022b; Yuan et al., 2024b). Graph matching 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). 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., 2020e), 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., 2022a), FOTA (Liu et al., 2022c), 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., 2013a), 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., 2019b), NAME (Zhou et al., 2019a), 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;f; 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., 2020b; Zhu et al., 2022b; Chakrabarti et al., 2022; Liu et al., 2022d; Guo et al., 2022; Zhu et al., 2022b; Xin et al., 2022) and text matching (Chen et al., 2020b).

Federated Graph Learning. Parallel, distributed, and federated learning have been extensively studied in recent years (Zhang et al., 2013b; Lee et al., 2013; Su et al., 2013; Zhang et al., 2014; Su et al., 2015; Zhou et al., 2015b; Bao et al., 2015; Zhou et al., 2015c; Lee et al., 2015; Jiang et al., 2016; Lee et al., 2019; Qu et al., 2020; Zhang et al., 2021d; Zhou et al., 2022a; Guimu Guo & Zhou, 2022; Jin et al., 2022b; Zhang et al., 2022a; Che et al., 2022; Yan et al., 2022a; Liu et al., 2022b; Yan et al., 2022b;c; Che et al., 2023b; Hong et al., 2023; Chen et al., 2023; Che et al., 2023a; Liu et al., 2023; 2024b;a; Yuan et al., 2024a; Khalil et al., 2024). 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., 2022b; 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 workThis work is the first to has 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 (1.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 $x_k = \{v_1, v_2, \cdots, v_k\}$ be the original node ordering of g_k via the subgraph expansion, $\mathcal{S}(g_k) = [v_1, v_2, \cdots, v_k]$ be the set of all possible node sequences of x_k, $x_k[i]$ be the ith node in x_k, F be a user-defined normalizing factor in the subgraph expansion, and $h_l(x_k) = \{v_1, v_2, \cdots, v_l, x_k, G\}$ be an induced subgraph of graph G with the first l nodes $\{v_1, v_2, \cdots, v_l\}$ in x_k, then the probability of getting a k-subgraph g_k via the subgraph expansion is*

$$p(g_k) = \sum_{x_k \in \mathcal{S}(g_k)} \frac{f(\deg(x_k[1]))}{F} \prod_{l=1}^{k-1} \frac{\left|E_{h_{l+1}(x_k)}\right| - \left|E_{h_l(x_k)}\right|}{\sum_{i=1}^{l} \deg(x_k[i]) - 2\left|E_{h_l(x_k)}\right|} \tag{23}$$

Proof. We can consider a subgraph expansion process as a way of sampling a sequence $x_k = \{v_1, v_2, \cdots, v_k\}$, 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 $h_l = \{v_1, v_2, \cdots, v_l\}$ 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 $h_{l+1} = \{v_1, v_2, \cdots, v_l, v_{l+1}\}$ is equal to

$$\mathbb{P}\left(v_{l+1} \mid h_l\right) = \frac{\deg_{h_l}\left(v_{l+1}\right)}{\left|\mathcal{N}_e\left(h_l\right)\right|} = \frac{\left|E_{h_{l+1}}\right| - \left|E_{h_l}\right|}{\sum_{i=1}^{l} \deg(v_i) - 2\left|E_{h_l}\right|} \tag{24}$$

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 $x_k = \{v_1, v_2, \cdots, v_k\} \in \mathcal{S}(g_k)$ in the subgraph expansion process is

equal to

$$\tilde{p}(x_k) = q(v_1) \prod_{l=1}^{k-1} \mathbb{P}(v_{l+1} \mid h_l) = \frac{f(\deg(v_1))}{F} \prod_{l=1}^{k-1} \frac{\left|E_{h_{l+1}}\right| - |E_{h_l}|}{\sum_{i=1}^{l} \deg(v_i) - 2|E_{h_l}|} \tag{25}$$

Notice that

$$p(g_k) = \sum_{x_k \in \mathcal{S}(g_k)} \tilde{p}(x_k) \tag{26}$$

Since

$$\mathcal{S}(g_k) = \{[v_1, \ldots, v_k] | \{v_1, \ldots, v_k\} = V_{g_k}, g_k | \{v_1, \ldots, v_l\} \textit{is connected}\} \tag{27}$$

,

then we have

$$p(g_k) = \sum_{x_k \in \mathcal{S}(g_k)} \frac{f(\deg(x_k[1]))}{F} \prod_{l=1}^{k-1} \frac{\left|E_{h_{l+1}(x_k)}\right| - \left|E_{h_l(x_k)}\right|}{\sum_{i=1}^{l} \deg(x_k[i]) - 2\left|E_{h_l(x_k)}\right|} \tag{28}$$

where $x_k = \{v_1, v_2, \cdots, v_k\}$ be the original node ordering of g_k via the subgraph expansion process. $x_k[i]$ be the ith node in x_k. $h_l(x_k) = \{v_1, v_2, \cdots, v_l, x_k, G\}$ be an induced subgraph of graph G with the first l nodes $\{v_1, v_2, \cdots, v_l\}$ 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, $d_1, \cdots, d_k$ be the k highest degrees of nodes in G, and denote $D = \prod_{l=2}^{k-1}(d_1 + \cdots + d_k)$. If q for sampling the starting node is the stationary distribution of the node random walk, then the upper bound of the variance $\mathrm{Var}(\tilde{n}_{kr}(G))$ is*

$$\mathrm{Var}(\tilde{n}_{kr}(G)) \leq \frac{1}{O} n_{kr}(G) \frac{2|E_G|}{|\mathcal{S}(\mathcal{G}_r)|} D \tag{29}$$

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

$$\mathrm{Var}(\tilde{n}_{kr}(G)) = \frac{1}{O} \mathrm{Var}\left(\frac{\mathbb{I}(g_{kO} \sim \mathcal{G}_r)}{p(g_{kO})}\right) = \frac{1}{O}\left(\sum_{g_k \in G_k} \frac{\mathbb{I}(g_k \sim \mathcal{G}_r)}{p(g_k)} - n_{kr}(G)^2\right) \tag{30}$$

It is observed that the variance $\mathrm{Var}(\tilde{n}_{kr}(G))$ 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

$$\phi_o = \frac{\mathbb{I}(g_{ko} \sim \mathcal{G}_r)}{p(g_{ko})} \tag{31}$$

The variance can be rewritten as follows.

$$\operatorname{Var}\left(\tilde{n}_{kr}(G)\right)=\frac{1}{O}\operatorname{Var}\left(\phi_o\right) \tag{32}$$

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 variancein Eq.(30) by the second moment, which is bounded by,

$$\mathbb{E}\phi_o^2\leq\mathbb{E}\phi_o\max\phi_o=n_{kr}(G)\max\phi_o \tag{33}$$

By seeking to control the the maximum of ϕ_o*, we have*

$$\max_{g_k}\frac{1}{p(g_k)}\leq\max_{x_k}\frac{1}{|\mathcal{S}(g_k)|\tilde{p}(x_k)}\leq\max\frac{\prod_{l=1}^{k-1}(d_1+\cdots+d_l)}{|\mathcal{S}\left(\mathcal{G}_r\right)|\,q\left(d_1\right)} \tag{34}$$

and we obtain

$$\max_{x}\frac{|\mathcal{N}_v(x)|}{\mathcal{S}(\mathcal{G}_r)|\tilde{p}(x)}\leq\max\frac{\prod_{l=1}^{k-1}(d_1+\cdots+d_l)}{|\mathcal{S}\left(\mathcal{G}_r\right)|\,q\left(d_1\right)} \tag{35}$$

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

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*

$$\mathcal{A}_{b+1}\otimes z_b^3=\sum_{i,j,k=1}^{d}\frac{\partial^3\mathcal{L}^s(M_{b+1})}{\partial M^i\partial M^j\partial M^k}z_b^iz_b^jz_b^k \tag{36}$$

and

$$\mathcal{B}_{b+1}\otimes z_b^4=\sum_{i,j,k,l=1}^{d}\frac{\partial^4\mathcal{L}^s(M_{b+1})}{\partial M^i\partial M^j\partial M^k\partial M^l}z_b^iz_b^jz_b^kz_b^l. \tag{37}$$

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

$$z_b^T\nabla^2\mathcal{L}^s(M_{b+1})z_b-\alpha_{b+1}(\omega)z_b^Tz_b=\left(\frac{1}{2}-\frac{\omega}{6}\right)\mathcal{A}_{b+1}\otimes z_b^3-\left(\frac{1}{6}-\frac{\omega}{12}\right)\mathcal{B}_{b+1}\otimes z_b^4+\mathcal{O}\left(\|z_b\|^5\right) \tag{38}$$

Proof. By utilizing the Taylor expansion, we obtain

$$\begin{aligned}\mathcal{L}^s(M_b)=&\mathcal{L}^s(M_{b+1})-(\nabla\mathcal{L}^s(M_{b+1}))^Tz_b+\frac{1}{2}z_b^T\nabla^2\mathcal{L}^s(M_{b+1})z_b-\\&\frac{1}{6}\mathcal{A}_{b+1}\otimes z_b^3+\frac{1}{24}\mathcal{B}_{b+1}\otimes z_b^4+\mathcal{O}\left(\|z_b\|^5\right)\end{aligned} \tag{39}$$

and

$$\begin{aligned}(\nabla\mathcal{L}^s(M_b))^Tz_b=&(\nabla\mathcal{L}^s(M_{b+1}))^Tz_b-z_b^T\nabla^2\mathcal{L}^s(M_{b+1})z_b+\\&\frac{1}{2}\mathcal{A}_{b+1}\otimes z_b^3-\frac{1}{6}\mathcal{B}_{b+1}\otimes z_b^4+\mathcal{O}\left(\|z_b\|^5\right)\end{aligned} \tag{40}$$

In addition, we have

$$\alpha_{b+1}(\omega) = \frac{z_b^T y_b + \omega \left[2\left(\mathcal{L}^s(M_b) - \mathcal{L}^s(M_{b+1})\right) + \left(\nabla\mathcal{L}^s(M_b) + \nabla\mathcal{L}^s(M_{b+1})\right)^T z_b\right]}{z_b^T z_b} \tag{41}$$

By combining Eqs.(39), (40), and (41), we get

$$\begin{aligned}
&z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b - \alpha_{b+1}(\omega) z_b^T z_b \\
=&z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b - z_b^T y_b - \\
&\omega \left[2\left(\mathcal{L}^s(M_b) - \mathcal{L}^s(M_{b+1})\right) + \left(\nabla\mathcal{L}^s(M_b) + \nabla\mathcal{L}^s(M_{b+1})\right)^T z_b\right] \\
=&\left(\frac{1}{2} - \frac{\omega}{6}\right)\mathcal{A}_{b+1} \otimes z_b^3 - \left(\frac{1}{6} - \frac{\omega}{12}\right)\mathcal{B}_{b+1} \otimes z_b^4 + \mathcal{O}\left(\|z_b\|^5\right)
\end{aligned} \tag{42}$$

Therefore, the proof is concluded.

Sensitivity analysis of weight ω. Based on Eq.(41), if $\omega = 0$, we have

$$\alpha_{b+1}(0) = \frac{z_b^T y_b}{z_b^T z_b} \tag{43}$$

Then, it derives the following equation based on Eq.(38).

$$z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b - \alpha_{b+1}(\omega) z_b^T z_b = \frac{1}{2}\mathcal{A}_{b+1} \otimes z_b^3 - \frac{1}{6}\mathcal{B}_{b+1} \otimes z_b^4 + \mathcal{O}\left(\|z_b\|^5\right) \tag{44}$$

According to Eq.(44) and Theorem 3, it is reasonable to believe that if the weight parameter ω is chosen such that

$$\left|\frac{1}{2} - \frac{\omega}{6}\right| < \frac{1}{2} \tag{45}$$

and

$$\left|\frac{1}{6} - \frac{\omega}{12}\right| < \frac{1}{6} \tag{46}$$

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 ω and the corresponding formulas for $\alpha_{b+1}(\omega)$.

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

$$\alpha_{b+1}(1) = \frac{z_b^T y_b + 2\left(\mathcal{L}^s(M_b) - \mathcal{L}^s(M_{b+1})\right) + \left(\nabla\mathcal{L}^s(M_b) + \nabla\mathcal{L}^s(M_{b+1})\right)^T z_b}{z_b^T z_b} \tag{47}$$

The resulting matrix $\alpha_{b+1}(1)\mathbb{I}$ satisfies the weak quasi-Newton equation in Eq.(17). Based on Eq.(38), we have

$$z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b - \alpha_{b+1}(1) z_b^T z_b = \frac{1}{3}\mathcal{A}_{b+1} \otimes z_b^3 - \frac{1}{12}\mathcal{B}_{b+1} \otimes z_b^4 + \mathcal{O}\left(\|z_b\|^5\right) \tag{48}$$

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

$$\alpha_{b+1}(2) = \frac{z_b^T y_b + 4\left(\mathcal{L}^s(M_b) - \mathcal{L}^s(M_{b+1})\right) + 2\left(\nabla\mathcal{L}^s(M_b) + \nabla\mathcal{L}^s(M_{b+1})\right)^T z_b}{z_b^T z_b} \tag{49}$$

The following equation is derived from Eq.(38).

$$z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b - \alpha_{b+1}(2) z_b^T z_b = \frac{1}{6}\mathcal{A}_{b+1} \otimes z_b^3 + \mathcal{O}\left(\|z_b\|^5\right) \tag{50}$$

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

$$\alpha_{b+1}(3) = \frac{z_b^T y_b + 6\left(\mathcal{L}^s(M_b) - \mathcal{L}^s(M_{b+1})\right) + 3\left(\nabla\mathcal{L}^s(M_b) + \nabla\mathcal{L}^s(M_{b+1})\right)^T z_b}{z_b^T z_b} \tag{51}$$

According to Eq.(38), we obtain

$$z_b^T \nabla^2 \mathcal{L}^s(M_{b+1}) z_b - \alpha_{b+1}(3) z_b^T z_b = \frac{1}{12}\mathcal{B}_{b+1} \otimes z_b^4 + \mathcal{O}\left(\|z_b\|^5\right) \tag{52}$$

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

$$u_b(0) - u_b(z_b) \geq \frac{1}{2}\|\nabla\mathcal{L}^s(M_b)\| \min\left\{\Delta^s, \frac{\|\nabla\mathcal{L}^s(M_b)\|}{\alpha_b}\right\} \tag{53}$$

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

$$z^* = \operatorname{argmin} u_b(z) \approx \mathcal{L}^s(M_b) + (\nabla\mathcal{L}^s(M_b))^T z + \frac{1}{2}\alpha_b z^T z,\ s.t. \|z\| \leq \Delta^s \tag{54}$$

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.(54) 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.

$$z^* = \operatorname{argmin} u_b(z) \approx \mathcal{L}^s(M_b) + (\nabla\mathcal{L}^s(M_b))^T z + \frac{1}{2}\alpha_b z^T z,\ s.t. \|z\| = \Delta^s \tag{55}$$

From Eq.(55), 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.(54) can be rewritten as follows.

$$z_b = -\frac{1}{\tilde{\alpha}_b}\nabla\mathcal{L}^s(M_b), \quad \textit{where } \tilde{\alpha}_b = \max\left\{\alpha_b, \frac{\|\nabla\mathcal{L}^s(M_b)\|}{\Delta^s}\right\} \tag{56}$$

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*

$$\begin{aligned} u_b(0) - u_b(z_b) =& -\left(\nabla\mathcal{L}^s(M_b)\right)^T\left(-\frac{1}{\alpha_b}\nabla\mathcal{L}^s(M_b)\right) \\ & -\frac{1}{2}\left(-\frac{1}{\alpha_b}\nabla\mathcal{L}^s(M_b)\right)^T\alpha_b I\left(-\frac{1}{\alpha_b}\nabla\mathcal{L}^s(M_b)\right) \\ =& \frac{\|\nabla\mathcal{L}^s(M_b)\|^2}{\alpha_b} - \frac{1}{2}\frac{\|\nabla\mathcal{L}^s(M_b)\|^2}{\alpha_b} \\ =& \frac{1}{2}\frac{\|\nabla\mathcal{L}^s(M_b)\|^2}{\alpha_b} \end{aligned} \tag{57}$$

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*

$$\begin{aligned} u_b(0) - u_b(z_b) =& -\left(\nabla\mathcal{L}^s(M_b)\right)^T\left(-\frac{\Delta^s}{\|\nabla\mathcal{L}^s(M_b)\|}\nabla\mathcal{L}^s(M_b)\right) \\ & -\frac{1}{2}\left(-\frac{\Delta^s}{\|\nabla\mathcal{L}^s(M_b)\|}\nabla\mathcal{L}^s(M_b)\right)^T\alpha_b I\left(-\frac{\Delta^s}{\|\nabla\mathcal{L}^s(M_b)\|}\nabla\mathcal{L}^s(M_b)\right) \\ =& \Delta^s\|\nabla\mathcal{L}^s(M_b)\| - \frac{1}{2}\alpha_b(\Delta^s)^2 \\ >& \Delta^s\|\nabla\mathcal{L}^s(M_b)\| - \frac{1}{2}\Delta^s\|\nabla\mathcal{L}^s(M_b)\| \\ =& \frac{1}{2}\Delta^s\|\nabla\mathcal{L}^s(M_b)\| \end{aligned} \tag{58}$$

By integrating Eqs.(57) and (58), we obtain Eq.(53).

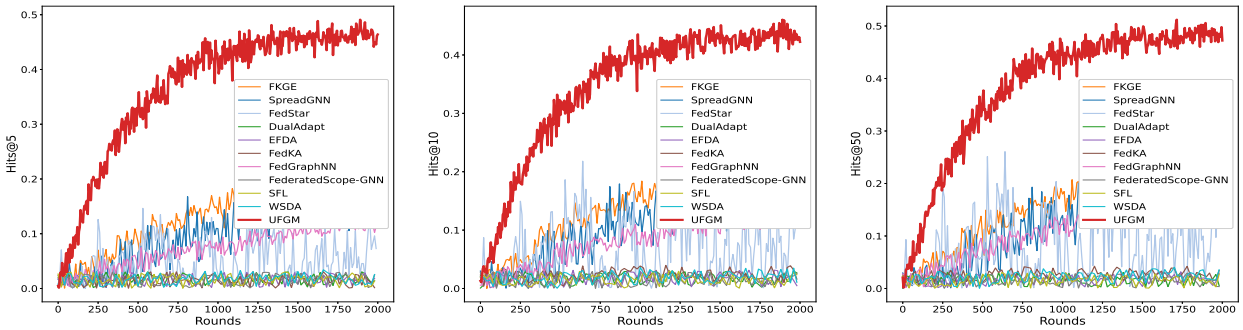
Therefore, the proof is concluded.

A.3. Additional Experiments

Table 3: Final performance on DBLP

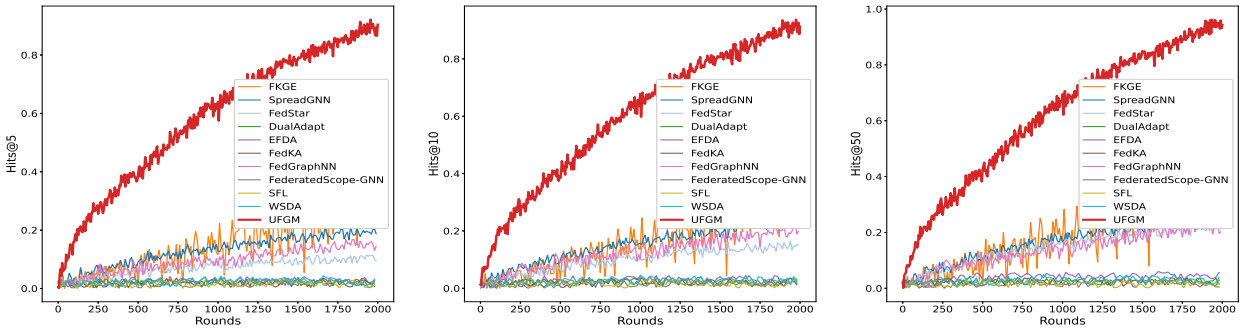
Type	Algorithm	$Hits@1$	$Hits@5$	$Hits@10$	$Hits@50$	$Loss$
Centralized Graph Matching	NextAlign	**0.572**	**0.609**	0.632	0.690	0.222
	NetTrans	0.529	0.592	0.616	0.632	1.881
	CPUGA	0.136	0.199	0.276	0.296	2.232
	ASAR-GM	0.172	0.237	0.260	0.271	2.052
	SIGMA	0.276	0.360	0.378	0.421	1.992
	SeedGNN	0.530	0.582	**0.637**	**0.702**	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
Federated Graph Learning	FedGraphNN	0.072	0.117	0.134	0.166	5.373
	FKGE	0.272	0.375	0.402	0.449	1.482
	SpreadGNN	0.092	0.140	0.176	0.197	**0.052**
	SFL	0.000	0.000	0.000	0.000	4.120
	FederatedScope-GNN	0.000	0.001	0.002	0.002	4.020
	FedStar	0.048	0.092	0.120	0.152	3.145
	UFGM	0.453	0.552	0.591	0.659	0.332

Final Performance and Convergence on SNS, PPI, and DBLP. Table 3 and Figures-5-8 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).



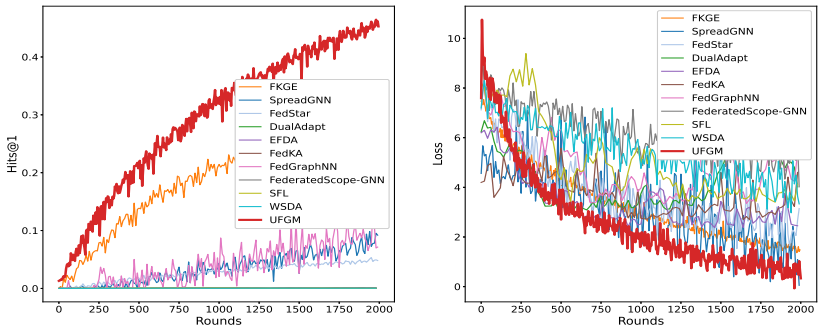
(a) $Hits@5$ (b) $Hits@10$ (c) $Hits@50$

Figure 5: Convergence on SNS



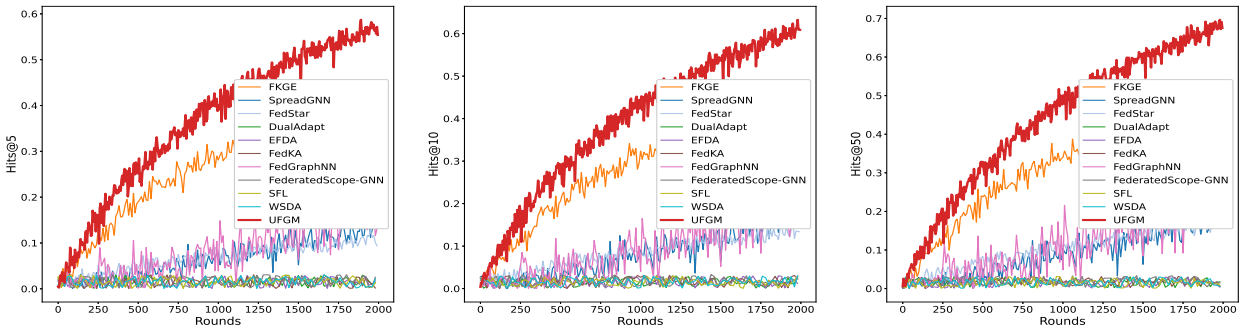
(a) $Hits@5$ (b) $Hits@10$ (c) $Hits@50$

Figure 6: Convergence on PPI



(a) $Hits@1$ (b) $Loss$

Figure 7: Convergence on DBLP



(a) $Hits@5$ (b) $Hits@10$ (c) $Hits@50$

Figure 8: 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.

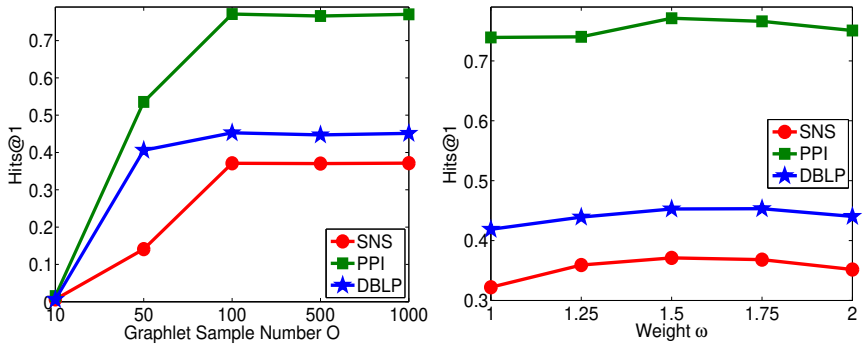
Table 6: Final performance of quasi-Newton approximation on three datasets

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
GMA-PGD	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. GMA-PGD is the exact version with the exact Hessian computation. The experiment results in Table 6 exhibit that the approximate version UFGM achieves slightly lower performance than the exact version GMA-PGD 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.



(a) O (b) ω

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.

Impact of weight ω between two types of weak quasi-Newton conditions. Figures 9 (b) 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 ω increases. Intuitively, a large ω 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 ω continuously increases. A reasonable explanation is that a too large ω 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 ω for separate trust region.

Table 7: 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 7 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 8: Statistics of the datasets

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

Datasets. We study federated graph learning tasks on three representative graph learning benchmark datasets: social networks (SNS) [1], protein-protein interaction networks (PPI) [2], and DBLP coauthor graphs (DBLP) [3]. The above three graph datasets are all public datasets, which allow researchers 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., 2020a). **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 keypoints 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 under- lying structure information for inter-graph federated learning tasks (Tan et al., 2022). (3) **Federated domain adaption baselines.** We compare the model performance with four recent federated domain adaption methods. **DualAdapt** aims to align the represen- tations 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).

[1] https://www.aminer.cn/cosnet
[2] http://snap.stanford.edu/ohmnet/
[3] http://dblp.uni-trier.de/xml/

Implementation. For six state-of-the-art centralized graph matching models of NextAlign [4], NetTrans [5], CPUGA [6], ASAR-GM [7], SeedGNN [8], and SIGMA [9], 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 [10], FKGE [11], SpreadGNN [12], SFL [13], FederatedScope-GNN [14], and FedStar [15], we also use the default parameters in the authors' implementation. For four recent federated domain adaption methods of DualAdapt [16], EFDA [17], WSDA [18], and FedKA [19], 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.

[4] https://github.com/sizhang92/NextAlign-KDD21
[5] https://github.com/sizhang92/NetTrans-KDD20
[6] https://github.com/scpei/CPUGA
[7] https://github.com/Thinklab-SJTU/ThinkMatch
[8] https://openreview.net/forum?id=iYvbPx8GTta
[9] https://github.com/CityU-AIM-Group/SIGMA
[10] https://github.com/FedML-AI/FedGraphNN
[11] https://github.com/HKUST-KnowComp/FKGE
[12] https://github.com/FedML-AI/SpreadGNN
[13] https://github.com/dawenzi098/SFL-Structural-Federated-Learning
[14] https://github.com/alibaba/federatedscope
[15] https://github.com/yuetan031/fedstar
[16] https://drive.google.com/file/d/1OekTpqB6qLfjlE2XUjQPm3F110KDMFc0/view?usp=sharing
[17] https://github.com/yuetan031/fedstar
[18] https://openreview.net/forum?id=_1gu0EX0mM3
[19] https://github.com/yuweisunn/federated-knowledge-alignment