

Semi-decentralized Federated Ego Graph Learning for Recommendation

Liang Qu*

qul@mail.sustech.edu.cn
Southern University of Science and
Technology
Shenzhen, China

Ningzhi Tang*

11912521@mail.sustech.edu.cn
Southern University of Science and
Technology
Shenzhen, China

Ruiqi Zheng

ruiqi.zheng@uq.net.au
The University of Queensland
Brisbane, Australia

Quoc Viet Hung Nguyen

henry.nguyen@griffith.edu.au
Griffith University
Brisbane, Australia

Zi Huang

huang@itee.uq.edu.au
The University of Queensland
Brisbane, Australia

Yuhui Shi[†]

shiyh@sustech.edu.cn
Southern University of Science and
Technology
Shenzhen, China

Hongzhi Yin[†]

db.hongzhi@gmail.com
The University of Queensland
Brisbane, Australia

ABSTRACT

Collaborative filtering (CF) based recommender systems are typically trained based on personal interaction data (e.g., clicks and purchases) that could be naturally represented as ego graphs. However, most existing recommendation methods collect these ego graphs from all users to compose a global graph to obtain high-order collaborative information between users and items, and these centralized CF recommendation methods inevitably lead to a high risk of user privacy leakage. Although recently proposed federated recommendation systems can mitigate the privacy problem, they either restrict the on-device local training to an isolated ego graph or rely on an additional third-party server to access other ego graphs resulting in a cumbersome pipeline, which is hard to work in practice. In addition, existing federated recommendation systems require resource-limited devices to maintain the entire embedding tables resulting in high communication costs.

In light of this, we propose a semi-decentralized federated ego graph learning framework for on-device recommendations, named SemiDFEGL, which introduces new device-to-device collaborations to improve scalability and reduce communication costs and innovatively utilizes predicted interacted item nodes to connect isolated ego graphs to augment local subgraphs such that the high-order

user-item collaborative information could be used in a privacy-preserving manner. Furthermore, the proposed framework is model-agnostic, meaning that it could be seamlessly integrated with existing graph neural network-based recommendation methods and privacy protection techniques. To validate the effectiveness of the proposed SemiDFEGL, extensive experiments are conducted on three public datasets, and the results demonstrate the superiority of the proposed SemiDFEGL compared to other federated recommendation methods.

CCS CONCEPTS

• Information systems → Recommender systems.

KEYWORDS

federated ego graph learning, recommender systems, semi-decentralized learning, graph neural networks

ACM Reference Format:

Liang Qu, Ningzhi Tang, Ruiqi Zheng, Quoc Viet Hung Nguyen, Zi Huang, Yuhui Shi, and Hongzhi Yin. 2023. Semi-decentralized Federated Ego Graph Learning for Recommendation. In *Proceedings of the ACM Web Conference 2023 (WWW '23)*, May 1–5, 2023, Austin, TX, USA. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3543507.3583337>

*Both authors contributed equally to this research.

[†]Corresponding Author

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

WWW '23, May 1–5, 2023, Austin, TX, USA

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-9416-1/23/04...\$15.00

<https://doi.org/10.1145/3543507.3583337>

1 INTRODUCTION

Recommender systems have been widely shown as an effective technique for helping users filter out the information they are not interested in. The typical architectures of current recommender systems [7, 8, 11] are that they collect all users' interaction data to train a powerful recommender model on a cloud server, as shown in Figure 1(a). However, with the increasing concerns for user privacy, such centralized recommendation methods that hold personal data and models in the central server are most likely to break the

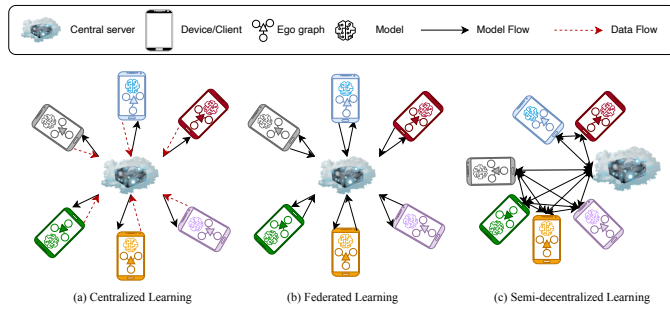


Figure 1: (a) Centralized learning (b) Federated learning (c) Semi-decentralized learning.

new privacy regulations or laws such as General Data Protection Regulation (GDPR)¹.

One natural solution to privacy issues is to keep users' data on their own devices and train models locally. However, a new problem that arises is that the local training data is very limited and thus insufficient to train an accurate model. Inspired by the effectiveness of federated learning [29, 50] in solving privacy-preserving machine learning problems, federated recommender systems (FedRecs) [1, 4, 6, 9, 21, 34, 35, 43, 45, 49] are proposed to collaboratively train recommendation models between local devices and a central server. Typically, as shown in Figure 1(b), FedRecs first train each local model on-device using the local dataset, and then sample a certain number of devices/clients to aggregate their local knowledge (e.g., model parameters/gradients) on the central server to learn a global model, which will then be redistributed back to each device for updating local models.

Nevertheless, the current federated recommendation architectures still face the following challenges. (1) **Scalability:** The star-shaped federated recommender architecture [1, 4, 34, 43] depends heavily on a central server to coordinate the device-server collaborations, making it infeasible to scale up to a large number of clients/users due to the communication bottleneck of the coordinator. There have been demonstrations with dozens and even hundreds of clients, but the complexity of a federated learning system comprising millions of mobile devices is simply inconceivable. (2) **Constrained resources on client devices:** Most existing FedRecs [1, 4] require the client devices to maintain and store the whole item embedding table, which is hard to work in practice due to the constrained on-device resources. (3) **Cumbersome pipeline:** To achieve privacy protection or data augmentation, some FedRecs require an additional third-party server to implement encryption services [21] or neighbor discovery [45], resulting in a cumbersome pipeline and also increasing the risk of privacy leakage.

In light of the above challenges, we investigate the federated recommendation problem from a novel federated ego graph learning (FedEGL) perspective, in which the training data is treated as a set of user-centered ego graphs distributed over client devices, and each client is not allowed to share their ego graphs with either the cloud server or other clients. In such a context, there is no global graph involved in the training process such that it is

non-trivial to capture the high-order collaborative information. It is worth noting that federated graph learning under the recommendation scenario differs from the traditional federated graph learning (FedGL) [13, 47, 54]. The major difference is that FedGL assumes that each device can have a large subgraph containing high-order graph information, while only first-order user-item interactions are available on each client device under the federated recommendation scenario. In addition, the subgraphs on different devices only have a few common nodes (i.e., anchor nodes), while in the context of recommendations there are often a large number of common items between users/clients, and such collaborative information is critical to recommendation performance. Therefore, those FedGL methods cannot be applied to the federated recommendation scenario.

To address the dilemma of capturing the high-order graph information from isolated ego graphs, we propose a semi-decentralized federated ego graph learning framework (as shown in Figure 1(c)) for recommendation, named SemiDFEGL, which introduces device-to-device collaborations to improve scalability as well as reduce communication costs, and also innovatively utilizes predicted interacted item nodes (called fake common items) to connect isolated ego graphs to construct local high-order subgraphs. Specifically, a client first learns an embedding representation for the local ego graph and then uploads the ego graph embedding to the central server. Based on the received ego graph embeddings and the item embedding table, the central server clusters users/items into a number of groups. To exploit the high-order graph information, items assigned to each group are called predicted interacted item nodes (also called fake common items) and are utilized to connect to each user within the group thus forming a local collaboration/communication graph on which clients/users can directly communicate with their neighbors to perform collaborative learning, making this semi-decentralized federated architecture more scalable. To be more specific, each client/user only needs to share aggregated node embeddings with other clients having common fake item nodes (i.e., neighbor clients) by the message-passing mechanism. In this way, users/clients do not need to share their own ego graphs. After finishing local collaborative training, each user/client only needs to upload the updated item embeddings (including embeddings of both positive and selected negative items) to the central server for updating the global item embedding table. Note that each client does not need to download and store the global item embedding table on their resource-constrained devices as the final recommendations are produced on the central server by calculating the relevance between the ego graph embedding and each item embedding. Furthermore, the proposed framework can accommodate almost all existing graph neural networks (GNNs) based recommenders and seamlessly integrate with existing privacy protection techniques (e.g., local differential privacy [18]).

The contributions of this paper are summarized as follows:

- We propose a semi-decentralized federated ego graph learning framework for recommendation, named SemiDFEGL, which introduces a novel device-to-device collaborative learning mechanism to improve scalability and reduce communication costs of the conventional FedRecs. The proposed

¹<https://gdpr-info.eu/>

framework is model-agnostic, which can accommodate almost all existing graph neural networks (GNNs) based recommenders and seamlessly integrate with existing privacy protection techniques such as LDP.

- We propose a novel local ego graph augmentation method to leverage predicted interacted items to connect isolated ego graphs so that the high-order collaborative information could be exploited in a privacy-preserving manner.
- Extensive experiments are conducted on three public datasets, and the experimental results demonstrate that the proposed framework could outperform other federated recommendation methods.

The rest of this paper is organized as follows. Section 2 will review the main related works, and Section 3 will detail the proposed framework. The experiments are introduced in Section 4, followed by a conclusion in Section 5.

2 RELATED WORK

2.1 Recommender Systems

Recommender systems have been shown as an effective technique for helping users filter out the information they are not interested in based on their feature information and historical interaction data (e.g., ratings and clicks). Early approaches mainly include collaborative filtering [40, 57], matrix factorization [24, 37] and factorization machines [19, 38]. Recently, deep learning [26] has achieved significant progress in computer vision [25] and natural language processing [16] due to the capabilities of capturing high-order non-linear features. Thus, there are a lot of works to apply it to recommender systems, such as deep neural network-based recommendation [7, 8, 11, 36, 52, 58], graph neural network-based recommendation [46, 53], and attention-based recommendation [20, 59]. However, as mentioned above, these methods are typically deployed in a centralized manner and collect all user data to train the model, which could suffer from data privacy, resource consumption, and response latency issues.

2.2 Federated Recommender Systems

Inspired by the effectiveness of federated learning in addressing privacy-preserving machine learning problems, federated recommender systems [17, 42, 43, 55, 56] are proposed to train on-device models in a device-to-server collaboration manner. For example, MetaMF [31] proposes a meta network to generate private item embeddings and rating prediction models deployed on user devices, and each device uploads the gradient information to the server for updating the meta network. FCF [1] extends the collaborative filter (CF) to the federated model, which achieves comparable performance against standard CF while fully preserving the user’s privacy. FedFast [34] introduces the active sampling and aggregation strategies to accelerate the convergence of local models. One similar work to ours is FedPerGNN [45] which also explores to learn privacy-preserving recommendations from a federated graph learning perspective. However, one major difference from our approach is that FedPerGNN requires a trusted third party to find users who have co-interacted items, which results in a cumbersome pipeline, and it is difficult to find such a trusted third-party server in practice. There are also some works allowing devices to collaboratively

learn with its neighbor devices in a fully device-to-device (D2D) fashion, namely decentralized recommender systems [2, 3, 5, 32]. However, such methods only rely on D2D collaborations and generally require that each device stores and maintains huge chunks of non-privacy item data, which is unnecessary and hard to work in practice due to the constrained resources.

3 PROPOSED METHOD

3.1 Preliminary

- **User-centered ego graph:** An user-centered ego graph $\mathcal{G} = \{u, V, E\}$ consists of a user node u , a set of items $V = \{v_i\}_{i=1}^n$, and a set of edges/interactions $E = \{e_j\}_{j=1}^n$, where v_i and e_j represent a item node and an interaction (e.g., clicks and ratings) between user u and item v_j , respectively. n is the number of items that user u has interacted with on a local device.
- **Federated ego graph learning (FedEGL) for recommendation:** Suppose that there are a set of devices $\mathcal{D} = \{D_k\}_{k=1}^K$, where K is the number of devices, and each device $D_k = \{\mathcal{G}_k, f(\Theta_k)\}$ consists of an ego graph \mathcal{G}_k and an on-device model $f(\Theta_k)$ parameterized by Θ_k . FedEGL aims to predict the user’s preferences based on the user-item interaction data \mathcal{G}_k in a privacy-preserving manner. It is worth noting that each device is not allowed to share its ego graph with either the central server or other devices.

3.2 SemiDFEGL

To address the dilemma of exploiting the high-order graph information from isolated ego graphs, and improve the scalability as well as reduce communication costs of current federated recommendation architectures, we propose a semi-decentralized federated ego graph learning framework for recommendation, named SemiDFEGL, which is illustrated in Figure 2. It consists of two main collaborative learning strategies: the centralized device-to-server (D2S) collaboration and the decentralized device-to-device (D2D) collaboration. Concretely, each client learns a local ego graph embedding and then uploads it to the central server. The central server clusters users/items into a number of groups based on their embeddings. Items assigned to each group are called predicted interacted item nodes (also called fake common items) and are utilized to connect to each user within the group to form a higher-order local subgraph. In this way, clients within the same group can directly communicate with their neighbors to perform D2D collaborations by sharing aggregated node embeddings with other clients rather than their ego graphs via common fake items. After that, clients only need to upload item embeddings (including embeddings of both positive and selected negative items) to the central server for updating the global item embedding table.

3.2.1 Local Ego Graph Augmentation. For the traditional GNN-based recommendation models [14] that maintains a global user-item interaction graph on the central server, the model can directly learn user/item embeddings from high-order user-item collaborative information over the global graph. However, it will inevitably lead to a high risk of user privacy leakage. In order to protect the

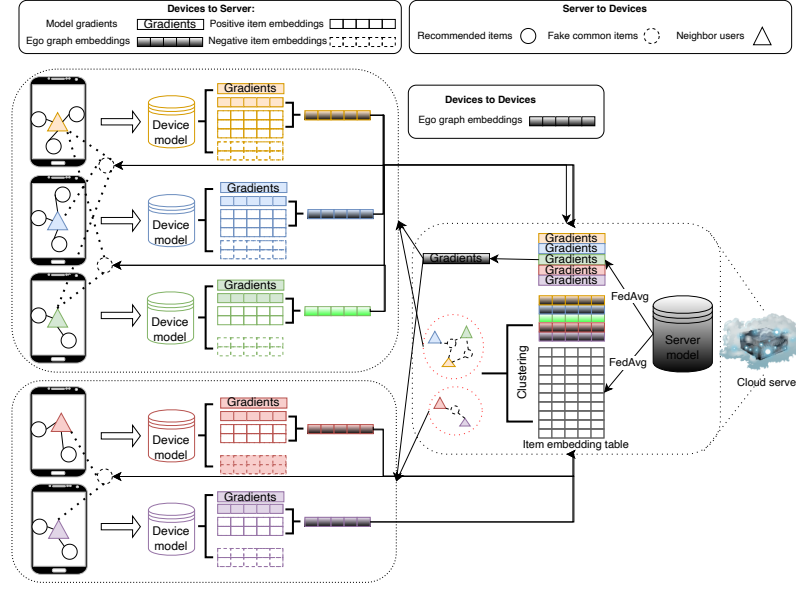


Figure 2: The overview of SemiDFEGL. The user-item interaction data, i.e., the ego graph, and item data are stored on the device side and cloud side, respectively. The device learns the embedding representation for the local ego graph and uploads it to the central server. The central server clusters users/items into a number of groups based on their embeddings. Items assigned to each group are called predicted interacted item nodes and are utilized to connect to each user within the group to form a higher-order local subgraph.

privacy of the user u , one natural solution is to keep personal interaction data that is represented as the ego graph \mathcal{G}_u on its own client. Thus, one major technical challenge is how to exploit high-order user-item collaborative information without actual data sharing. To address this problem, we propose to utilize a central server to generate a set of fake common items to connect users thus forming a local collaboration subgraph.

Specifically, each client first learns an embedding representation for the local ego graph by the message-passing mechanism of GNN as follows:

$$\mathbf{e}_u^{(k+1)} = \text{AggFun}(\mathbf{e}_u^{(k)}, \{\mathbf{e}_i^{(k)}\}_{i \in \mathcal{N}(u)}) \quad (1)$$

where \mathbf{e}_u and \mathbf{e}_i are user and item embeddings, respectively. k is the number of layers, and $\mathcal{N}(u)$ represents the set of neighbor nodes of user u . $\text{AggFun}(\cdot)$ is the aggregation function which could be defined by different strategies of GNN variants such as graph convolution network (GCN [23]), graph attention networks (GAT [41]), and Light Graph Convolution (LGC [14]). Note that the proposed framework is model-agnostic, meaning that it could be integrated with most existing GNN variants. In this section, we choose LGC as an example to introduce our method due to its superiority in the recommendation problem. Thus, the user/item embeddings could be learned as follows [14]:

$$\begin{aligned} \mathbf{e}_u^{(k+1)} &= \sum_{i \in \mathcal{N}(u)} \frac{1}{\sqrt{|\mathcal{N}(u)|} \sqrt{|\mathcal{N}(i)|}} \mathbf{e}_i^{(k)} \\ \mathbf{e}_i^{(k+1)} &= \sum_{u \in \mathcal{N}(i)} \frac{1}{\sqrt{|\mathcal{N}(i)|} \sqrt{|\mathcal{N}(u)|}} \mathbf{e}_u^{(k)} \end{aligned} \quad (2)$$

For each ego graph, as it only has first-order user-item interaction information, we can directly obtain its user embedding without layer combination. Since there is no ego graph augmentation in the initialization phase, we directly replace the ego graph embedding with the user's embedding, i.e., $\mathbf{e}_{ego} = \mathbf{e}_u$ and upload it to the central server.

Now, the central server has obtained the uploaded ego graph embeddings which represent users' historical preferences. In order to connect isolated ego graphs, we propose to employ a co-clustering method to cluster users and items into a number of groups based on ego graph embeddings $\mathcal{E}_U = \{\mathbf{e}_u\}_{u=1}^N$ and item embedding table $\mathcal{E}_V = \{\mathbf{e}_v\}_{v=1}^M$, where N and M represent the number of users and the number of items, respectively. In particular, considering the characteristics of recommendation systems, i.e., there may be an item that belongs to more than one group at the same time. We utilize the fuzzy c-means [27, 48] to learn the membership P_{ij} of entry \mathbf{x}_i (an ego graph embedding or an item embedding) in the group j by optimizing the following criterion function [48] in central server:

$$J_m(P, C) = \sum_{i=1}^{M+N} \sum_{j=1}^C (P_{ij})^l d(\mathbf{x}_i, \mathbf{c}_j)^2 \quad (3)$$

where \mathbf{c}_j is the center of group j . $d(\cdot)$ is the distance function, and we employ Euclidean distance in this paper. l is the weighting exponent to control the fuzziness of the resulting partition. During each iteration in a central server, P and C are updated as follows

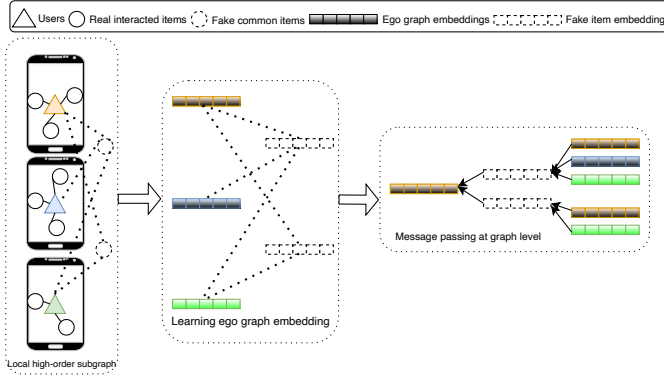


Figure 3: The illustration of device-to-device collaborations via message passing at graph level.

until the termination condition is met.

$$P_{ij} = (d(\mathbf{x}_i, \mathbf{c}_j))^{2/(1-l)} / \left[\sum_{k=1}^c (d(\mathbf{x}_i, \mathbf{c}_k))^{2/(1-l)} \right] \quad (4)$$

$$\mathbf{c}_j = \left[\sum_{i=1}^{M+N} P_{ij}^l \mathbf{x}_i \right] / \left[\sum_{i=1}^{M+N} P_{ij}^l \right] \quad (5)$$

After that, for each user’s membership, we select the element with the largest value to ensure that each user/client belongs to only one group. For the membership of each item, we pick the top- k largest elements so that the same item can belong to more than one group. Thus, the items assigned to each group are called predicted interacted item nodes (also called fake common items) and are utilized to connect to each user within the group to form a higher-order local subgraph.

3.2.2 Device-to-Device Collaborations. Although we have constructed higher-order subgraphs in each group, each device is not allowed to share its own ego graph with its neighboring clients due to privacy concerns. The information shared by the clients in the group is that each client only knows the common fake items and the clients connected to them. Therefore, we propose to perform device-to-device message passing at the graph level rather than at the node level. In particular, as shown in Figure 3, each device first learns its own ego graph embedding based on equation 2, and shares it with other neighbor clients via fake common items. In this way, the message-passing mechanism of GNN can be performed in a standard way, such as the layer combination operation of LightGCN [14] as follows:

$$\mathbf{e}_u = \sum_{k=0}^K \alpha_k \mathbf{e}_u^{(k)}; \quad \mathbf{e}_i = \sum_{k=0}^K \alpha_k \mathbf{e}_i^{(k)} \quad (6)$$

where α_k is the hyperparameter representing the importance of the k -th layer. Thus, each client can learn the final user embedding \mathbf{e}_u and the item embedding \mathbf{e}_i locally that contains the high-order graph information.

3.2.3 Device-to-Server Collaborations. After learning node embedding locally, we perform the device-to-server collaborations by sampling devices from each group and uploading item embeddings

to the central server for updating the global item embedding table. In particular, to protect the user’s interaction data, each client samples a number of items that it has not interacted with and randomly generates their embeddings \mathbf{e}_i^- with the same mean and variance as the real interacted item embedding \mathbf{e}_i^+ . Then, the embeddings of both positive and negative items are uploaded to the central server to perform the FedAvg [33]. Furthermore, since the embeddings could also leak the user’s privacy, we clip the uploaded embeddings based on L1-norm with a threshold δ , and integrate LDP with embeddings to further protect user privacy as follows:

$$\hat{\mathbf{e}}_i = \text{clip}(\hat{\mathbf{e}}_i, \delta) + \text{Laplace}(0, \lambda) \quad (7)$$

where $\hat{\mathbf{e}}_i = \{\mathbf{e}_i^+, \mathbf{e}_i^-\}$ are the embeddings that need to be uploaded to the central server, and λ is the noise scale.

3.2.4 Model Prediction and Training. The model prediction, i.e., the personalized recommendation, is produced on the central server by calculating the ranking score y_{ui} of the upload ego graph embedding of the user and each item embedding as follows:

$$y_{ui} = \mathbf{e}_{ego}^T \mathbf{e}_i \quad (8)$$

In this way, each client does not need to download and store the global item embedding table on their resource-constrained devices.

For the model training on each client, we utilize the Bayesian Personalized Ranking (BPR) [39] as the loss function \mathcal{L} as follows:

$$\mathcal{L} = - \sum_{i \in \mathcal{N}_u} \sum_{j \notin \mathcal{N}_u} \ln \sigma(y_{ui} - y_{uj}) + \lambda \|\mathbf{e}\|^2 \quad (9)$$

where λ is a hyperparameter that controls the strength of the L_2 regularization, and $\mathbf{e} = \{\mathbf{e}_i, \mathbf{e}_u\}$ is the trainable user/item embeddings. Note that since each client knows the items it has actually interacted with itself, the fake common items can be considered as negative samples in the model training process.

4 EXPERIMENTS

To validate the effectiveness of the proposed method, we conduct extensive experiments to answer the following research questions (RQs):

- **RQ1:** How does the proposed method perform compared with other federated recommendation methods?
- **RQ2:** Can the proposed method reduce communication costs?
- **RQ3:** How do different components affect the performance of the proposed method?
- **RQ4:** How do different hyperparameters influence the performance of the proposed method?
- **RQ5:** Can the proposed method be integrated with other GNN-based recommendation methods?

Table 1: The statistics of datasets.

Dataset	MovieLens-1M	Yelp2018	Gowalla
#User	6,040	31,668	29,858
#Item	3,900	38,048	40,981
#Interactions	1,000,209	1,561,406	1,027,370

Table 2: The top-K recommendation performance of SemiDFEGL and baselines on three datasets.

Method		MovieLen-1M		Yelp2018		Gowalla	
		Recall@20	NDCG@20	Recall@20	NDCG@20	Recall@20	NDCG@20
Cloud-based Recs	NeuMF	0.2402	0.3641	0.0572	0.0489	0.1506	0.1307
	LightGCN	0.2559	0.3859	0.0626	0.0511	0.18	0.1523
FedRecs	FCF	0.2301	0.3556	0.0517	0.0417	0.1434	0.1211
	FedMF	0.2397	0.3632	0.0491	0.0402	0.1432	0.1227
	MetaMF	0.2135	0.3342	0.0346	0.0283	0.1005	0.083
	FedPerGNN	0.2435	0.3686	0.0449	0.0445	0.1446	0.1127
Ours	SemiDFEGL	0.2486	0.3771	0.0583	0.0491	0.168	0.1372

4.1 Experimental Settings

4.1.1 Datasets. We validate baselines and the proposed method on three widely-used public datasets, i.e., MovieLens-1M², Yelp2018³ and Gowalla [30]. The statistics of the datasets are summarized in Table 1. All these three datasets are composed of historical user-item interactions. Following [15, 44], we filter out those users and items with less than 20/10/10 interactions for MovieLens-1M/Yelp2018/Gowalla, respectively. All three datasets are randomly split into training sets, validation sets, and test sets as the ratio 8:1:1, respectively.

4.1.2 Baselines. We compare SemiDFEGL with six baselines from two categories including the cloud-based recommendation methods and federated recommendation methods.

- **Cloud-based recommendation methods:**
 - **NeuMF** [15]: It is a representative deep recommendation method that uses DNN to replace the dot product operation to capture non-linearity in the implicit feedback.
 - **LightGCN** [14]: It utilizes graph neural networks to learn the user/item embeddings via the linear neighborhood aggregation mechanism on the user-item interaction graph, which is the state-of-the-art recommendation method.
- **Federated recommendation methods:**
 - **FCF** [1]: It extends CF to the federated model, which achieves comparable performance against standard CF while fully preserving the user’s privacy.
 - **FedMF** [4]: It proposes a user-level distributed matrix factorization framework to well protect users’ privacy by the homomorphic encryption technique, which only requires users to upload gradients to the cloud.
 - **MetaMF** [31]: It is a federated learning-based matrix factorization method that has a similar motivation to us, i.e., generating personalized models for each device. It learns a meta network on the cloud server, and the meta network utilizes the collaborative vector to generate private and personalized item embeddings and prediction models for each user.
 - **FedPerGNN** [45]: It is a GNN-based federated recommendation method, which proposes to use a trusted third-party server to assign neighbors who have co-interacted items

to each user such that the high-order graph information could be captured.

4.1.3 Evaluation metrics. We use two widely used evaluation metrics *Recall@20* and *NDCG@20* (Normalized Discounted Cumulative Gain) to measure the recommendation performance [14, 15, 51]. Following [14], we treat all items that have not interacted with a user as candidates, and the average results over all users are reported.

4.1.4 Hyper-parameter Setting. The hyperparameters of the proposed SemiDFEGL are set as follows. The user/item embeddings are initialized with the Xavier method [10], and the dimension is 64 for all methods. We utilize Adam optimizer [22] with an initial learning rate of 0.001, and the weight decay is 0.0001. The number of fake common item nodes and the number of layers are 1 and 4 for all datasets, respectively. The number of groups is 100 on MovieLens-1M and 200 on the other two datasets. The baselines are implemented by the codes provided by the authors.

4.2 Top-k recommendation (RQ1)

We first validate the effectiveness of the proposed SemiDFEGL on the top-k recommendation task. The experimental results are shown in Table 2. We can observe that:

- The cloud-based recommendation methods achieve better performance than the federated recommendation methods in most cases. A major reason is that FedRecs methods need encryption services to protect user privacy, which introduces additional noise, resulting in limiting the model capability on learning user/item representations and model optimization.
- The graph-based methods (i.e., LightGCN, FedPerGNN, and SemiDFEGL) achieve better performance than MF-based methods (i.e., NeuMF, FedMF, and MetaMF). The possible reason is that GNN could better capture the implicit collaborative information between users and items by the message-passing mechanism, which further validates the reasonability that we model the federated recommendation problem from the federated ego graph learning perspective.
- SemiDFEGL significantly outperform other federated recommendation methods on three datasets and achieve competitive performance against cloud-based methods. For example, it exceeds FedPerGNN by 2.09% and 2.3% with respect to *Recall@20* and *NDCG20* respectively on MovieLens-1M.

²<https://grouplens.org/datasets/movielens/1m/>

³<https://www.yelp.com/dataset/challenge>

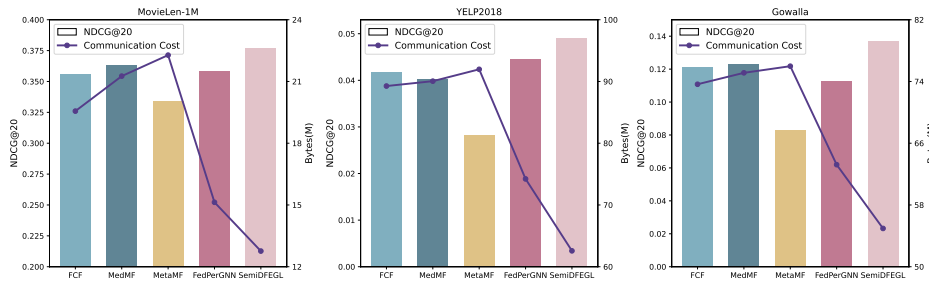


Figure 4: The model performance and communication costs for different federated recommendation methods.

Table 3: The ablation study results with respect to fake common item nodes.

Method	MovieLen-1M		Yelp2018		Gowalla	
	Recall@20	NDCG@20	Recall@20	NDCG@20	Recall@20	NDCG@20
SemiDFEGL w/o fake items	0.2314	0.3502	0.0467	0.04	0.149	0.1277
SemiDFEGL	0.2486	0.3771	0.0583	0.0491	0.168	0.1372

We attribute it to the fact that SemiDFEGL can well learn high-order user-item collaborative information via device-to-device collaborations.

4.3 Communication costs (RQ2)

Recall that we argue that the communication bottleneck of the central server makes the current federated recommendation framework infeasible to scale up to a large number of clients. Thus, we compare the model performance with their communication costs in this section. Specifically, we use the number of parameters to be uploaded and downloaded in each training round to measure the communication cost of the model. The experiments are conducted on three datasets, and the results are shown in Figure 4, from which we can observe that:

- Our model requires minimal communication costs but achieves the best performance. We attribute this to the fact that we do not need each device to maintain the global item embedding table as the final recommendations are produced on the central server.
- MF-based federated recommendation methods (i.e., FedMF, and MetaMF) require larger communication costs than GNN-based federated recommendation methods (i.e., FedPerGNN, and SemiDFEGL). It is reasonable because the former requires that the user client needs to upload the entire embedding table in each training round for protecting user interaction data.

4.4 Ablation Study (RQ3)

In this section, we target to demonstrate the effect of the fake common items. In particular, we implement the SemiDFEGL without fake common items, denoted as *SemiDFEGL w/o fake items*. In this way, each client can only use its ego graph to train the local model, and device-to-device collaborations are also not available. The experiments are conducted on three datasets, and other settings are

the same as the above experiments. The results are shown in Table 3, from which we can observe that:

- When the model has no fake common items, the performance of the model decreases substantially. This is reasonable because each isolated subgraph can only learn its first-order graph information and there is no collaboration between subgraphs, which further validates the necessity of our proposed ego graph augmentation strategy.

4.5 Hyper-parameter Sensitivity Analysis (RQ4)

To explore the effect of four crucial hyperparameters, i.e., the number of fake common items $\#FakeNodes = \{1, 5, 10, 20\}$, the number of groups $\#Groups = \{50, 100, 200, 500\}$, the number of layers $\#Layers = \{1, 2, 3, 4, 5\}$, and the number of negative items $\#NegativeItems = \{1, 5, 10, 20\}$, we conduct experiments on MovieLens-1M dataset, and report the model performance with different hyperparameter changes on Figure 5. We can observe that:

- As the number of fake common items increases, the performance of the model decreases. This is reasonable because the fake common items could connect isolated subgraphs for exploiting higher-order graph information. However, fake common items can somewhat disrupt the original data distribution and introduce additional noise.
- The model achieves the best performance when the number of groups is 100. The possible reason for this is that when the number of groups is small, the number of fake common items within each group is high, which brings in more noise and thus reduces the model performance. On the other hand, when the number of groups is large, the number of users in each cluster is small, resulting in insufficient learning of higher-order graph information.
- The model achieves the best performance when the number of layers of the network is 4. This is explained by that when

Table 4: Performance of SemiDFEGL with different GNN variants.

Methods	MovieLens-1M		Yelp2018		Gowalla	
	Recall@20	NDCG@20	Recall@20	NDCG@20	Recall@20	NDCG@20
GCN	0.2466	0.03764	0.0574	0.048	0.1663	0.1355
GAT	0.2472	0.3765	0.0569	0.0474	0.1672	0.1362
GraphSAGE	0.2467	0.3764	0.0571	0.0475	0.167	0.1358
SemiDFEGL (LightGCN)	0.2486	0.3771	0.0583	0.0491	0.168	0.1372

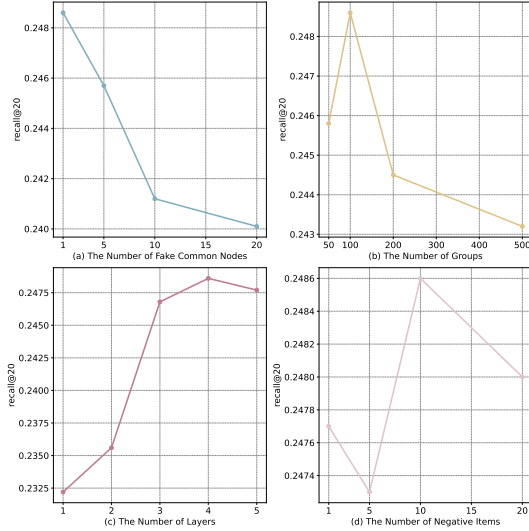


Figure 5: The performance of model with different hyperparameter settings. (a) The number of fake common nodes. (b) The number of groups. (c) The number of layers. (d) The number of negative items.

the number of layers is low, the model is unable to learn the higher-order structural information of the graph. On the other hand, when the number of layers is high, the over-smoothing problem [28] of GNNs may be introduced.

- Although the model can achieve better results when the number of negative items is small, the privacy of users may be at risk of leakage. With a slight increase in the number of negative items, additional noise is introduced which reduces the performance of the model. However, when the number of negative items is large, the negative effects may be offset.

4.6 Model-agnostic study (RQ5)

This section aims to explore if the proposed semi-decentralized federated ego graph learning framework can accommodate other GNN methods. To this end, we implement the proposed method with various GNN variants including graph convolution network (GCN [23]), graph attention networks GAT [41]), and GraphSAGE [12]. Experiments are conducted on MovieLen-1M dataset, and the one-hot ID vector is utilized as the node attributes. Experimental results are shown in Table 4. From the results, we can observe that:

- The model integrated with LightGCN achieves the best performance among the four variants. The possible reason is that the other three variants use the nonlinear layer to learn the embeddings of user and item nodes, but the nodes of the user-item bipartite graph have only one-hot ID features, leading to increased difficulty in optimizing the nonlinear layer.
- The attention-based methods (i.e., GAT and LightGCN) can achieve better performance than the other two variants. The performance improvement could be attributed to the attention mechanism that can well capture the importance of different neighbor nodes.

5 CONCLUSION

In this work, we investigated the federated recommendation problem from a novel federated ego graph learning perspective, and proposed a novel semi-decentralized federated ego graph learning framework for recommendation, named SemiDFEGL, which introduces device-to-device collaborations to improve scalability as well as reduce communication costs. To address the dilemma of capturing the high-order graph information from isolated ego graphs, we innovatively utilized fake common items to connect isolated ego graphs to construct local high-order subgraphs. Furthermore, the proposed framework can accommodate almost all existing GNN-based recommenders and seamlessly integrate with existing privacy protection techniques. To validate the effectiveness of the proposed method, comprehensive experiments have been conducted on three public datasets, and empirical results have demonstrated the superiority of the proposed method against federated recommendation methods. For the future work, we plan to explore extending our framework to non-graph-based recommendation models.

6 ACKNOWLEDGEMENT

This work is supported by Australian Research Council Future Fellowship (Grant No. FT210100624), Discovery Project (Grant No. DP190101985), Discovery Early Career Research Award (Grant No. DE200101465), the National Natural Science Foundation of China (Grant No. 61761136008), the Shenzhen Fundamental Research Program (Grant No. JCYJ20200109141235597), the Shenzhen Peacock Plan (Grant No. KQTD2016112514355531), the Guangdong Basic and Applied Basic Research Foundation (Grant No. 2021A1515110024), and the Program for Guangdong Introducing Innovative and Entrepreneurial Teams (Grant No. 2017ZT07X386).

REFERENCES

- [1] Muhammad Ahammad-Ud-Din, Elena Ivannikova, Suleiman A Khan, Were Oyomno, Qiang Fu, Kuan Eeik Tan, and Adrian Flanagan. 2019. Federated collaborative filtering for privacy-preserving personalized recommendation system. *arXiv preprint arXiv:1901.09888* (2019).
- [2] Felix Beierle and Simone Egger. 2020. MobRec—Mobile Platform for Decentralized Recommender Systems. *IEEE Access* 8 (2020), 185311–185329.
- [3] Felix Beierle and Tobias Eichinger. 2019. Collaborating with users in proximity for decentralized mobile recommender systems. In *2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI)*. IEEE, 1192–1197.
- [4] Di Chai, Leye Wang, Kai Chen, and Qiang Yang. 2020. Secure federated matrix factorization. *IEEE Intelligent Systems* (2020).
- [5] Chaochao Chen, Ziqi Liu, Peilin Zhao, Jun Zhou, and Xiaolong Li. 2018. Privacy preserving point-of-interest recommendation using decentralized matrix factorization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32.
- [6] Fei Chen, Mi Luo, Zhenhua Dong, Zhenguo Li, and Xiuliang He. 2018. Federated meta-learning with fast convergence and efficient communication. *arXiv preprint arXiv:1802.07876* (2018).
- [7] Heng-Tze Cheng, Levent Koc, Jeremiah Harmsen, Tal Shaked, Tushar Chandra, Hrishikesh Aradhye, Glen Anderson, Greg Corrado, Wei Chai, Mustafa Isipir, et al. 2016. Wide & deep learning for recommender systems. In *Proceedings of the 1st workshop on deep learning for recommender systems*. 7–10.
- [8] Paul Covington, Jay Adams, and Emre Sargin. 2016. Deep neural networks for youtube recommendations. In *Proceedings of the 10th ACM conference on recommender systems*. 191–198.
- [9] Adrian Flanagan, Were Oyomno, Alexander Grigorievskiy, Kuan Eeik Tan, Suleiman A Khan, and Muhammad Ahammad-Ud-Din. 2020. Federated multi-view matrix factorization for personalized recommendations. *arXiv preprint arXiv:2004.04256* (2020).
- [10] Xavier Glorot and Yoshua Bengio. 2010. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*. JMLR Workshop and Conference Proceedings, 249–256.
- [11] Huifeng Guo, Ruiming Tang, Yunming Ye, Zhenguo Li, and Xiuliang He. 2017. DeepFM: a factorization-machine based neural network for CTR prediction. *arXiv preprint arXiv:1703.04247* (2017).
- [12] Will Hamilton, Zitao Ying, and Jure Leskovec. 2017. Inductive representation learning on large graphs. *Advances in neural information processing systems* 30 (2017).
- [13] Chaoyang He, Keshav Balasubramanian, Emir Ceyani, Carl Yang, Han Xie, Lichao Sun, Lifang He, Liangwei Yang, Philip S Yu, Yu Rong, et al. 2021. Fedgraphnn: A federated learning system and benchmark for graph neural networks. *arXiv preprint arXiv:2104.07145* (2021).
- [14] Xiangnan He, Kuan Deng, Xiang Wang, Yan Li, Yongdong Zhang, and Meng Wang. 2020. Lightgcn: Simplifying and powering graph convolution network for recommendation. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval*. 639–648.
- [15] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural Collaborative Filtering. In *Proceedings of the 26th International Conference on World Wide Web (Perth, Australia) (WWW '17)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 173–182. <https://doi.org/10.1145/3038912.3052569>
- [16] Geoffrey Hinton, Li Deng, Dong Yu, George E Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N Sainath, et al. 2012. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal processing magazine* 29, 6 (2012), 82–97.
- [17] Mubashir Imran, Hongzhi Yin, Tong Chen, Nguyen Quoc Viet Hung, Alexander Zhou, and Kai Zheng. 2022. ReFRS: Resource-efficient Federated Recommender System for Dynamic and Diversified User Preferences. *ACM Transactions on Information Systems (TOIS)* (2022).
- [18] Matthew Joseph, Aaron Roth, Jonathan Ullman, and Bo Waggoner. 2018. Local differential privacy for evolving data. *Advances in Neural Information Processing Systems* 31 (2018).
- [19] Yuchin Juan, Yong Zhuang, Wei-Sheng Chin, and Chih-Jen Lin. 2016. Field-aware factorization machines for CTR prediction. In *Proceedings of the 10th ACM conference on recommender systems*. 43–50.
- [20] Wang-Cheng Kang and Julian McAuley. 2018. Self-attentive sequential recommendation. In *2018 IEEE International Conference on Data Mining (ICDM)*. IEEE, 197–206.
- [21] Sungwook Kim, Jinsu Kim, Dongyoung Koo, Yuna Kim, Hyunsoo Yoon, and Junbum Shin. 2016. Efficient privacy-preserving matrix factorization via fully homomorphic encryption. In *Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security*. 617–628.
- [22] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).
- [23] Thomas N Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907* (2016).
- [24] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. *Computer* 42, 8 (2009), 30–37.
- [25] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2012. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems* 25 (2012), 1097–1105.
- [26] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. 2015. Deep learning. *nature* 521, 7553 (2015), 436–444.
- [27] Jacek Łęski. 2003. Towards a robust fuzzy clustering. *Fuzzy Sets and Systems* 137, 2 (2003), 215–233.
- [28] Qimai Li, Zhichao Han, and Xiao-Ming Wu. 2018. Deeper insights into graph convolutional networks for semi-supervised learning. In *Thirty-Second AAAI conference on artificial intelligence*.
- [29] Tian Li, Anit Kumar Sahu, Ameet Talwalkar, and Virginia Smith. 2020. Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine* 37, 3 (2020), 50–60.
- [30] Dawen Liang, Laurent Charlin, James McInerney, and David M Blei. 2016. Modeling user exposure in recommendation. In *Proceedings of the 25th international conference on World Wide Web*. 951–961.
- [31] Yujie Lin, Pengjie Ren, Zhumin Chen, Zhaochun Ren, Dongxiao Yu, Jun Ma, Maarten de Rijke, and Xiuzhen Cheng. 2020. Meta Matrix Factorization for Federated Rating Predictions. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 981–990.
- [32] Jing Long, Tong Chen, Nguyen Quoc Viet Hung, and Hongzhi Yin. 2022. Decentralized Collaborative Learning Framework for Next POI Recommendation. *ACM Transactions on Information Systems (TOIS)* (2022).
- [33] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agüera y Arcas. 2017. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*. PMLR, 1273–1282.
- [34] Khalil Muhammad, Qinqin Wang, Diarmuid O'Reilly-Morgan, Elias Tragos, Barry Smyth, Neil Hurley, James Geraci, and Aonghus Lawlor. 2020. FedFast: Going Beyond Average for Faster Training of Federated Recommender Systems. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (Virtual Event, CA, USA) (KDD '20)*. Association for Computing Machinery, New York, NY, USA, 1234–1242. <https://doi.org/10.1145/3394486.3403176>
- [35] Tao Qi, Fangzhao Wu, Chuhan Wu, Yongfeng Huang, and Xing Xie. 2020. Privacy-Preserving News Recommendation Model Training via Federated Learning. (2020).
- [36] Liang Qu, Yonghong Ye, Ningzhi Tang, Lixin Zhang, Yuhui Shi, and Hongzhi Yin. 2022. Single-Shot Embedding Dimension Search in Recommender System. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval (Madrid, Spain) (SIGIR '22)*. Association for Computing Machinery, New York, NY, USA, 513–522. <https://doi.org/10.1145/3477495.3532060>
- [37] Andy Ramlatchan, Mengyun Yang, Quan Liu, Min Li, Jianxin Wang, and Yaohang Li. 2018. A survey of matrix completion methods for recommendation systems. *Big Data Mining and Analytics* 1, 4 (2018), 308–323.
- [38] Steffen Rendle. 2010. Factorization machines. In *2010 IEEE International conference on data mining*. IEEE, 995–1000.
- [39] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2012. BPR: Bayesian personalized ranking from implicit feedback. *arXiv preprint arXiv:1205.2618* (2012).
- [40] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. 2001. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web*. 285–295.
- [41] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2017. Graph attention networks. *arXiv preprint arXiv:1710.10903* (2017).
- [42] Qinyong Wang, Hongzhi Yin, Tong Chen, Zi Huang, Hao Wang, Yanchang Zhao, and Nguyen Quoc Viet Hung. 2020. Next point-of-interest recommendation on resource-constrained mobile devices. In *Proceedings of the Web conference 2020*. 906–916.
- [43] Qinyong Wang, Hongzhi Yin, Tong Chen, Junliang Yu, Alexander Zhou, and Xiangliang Zhang. 2021. Fast-adapting and Privacy-preserving Federated Recommender System. *arXiv preprint arXiv:2104.00919* (2021).
- [44] Xiang Wang, Xiangnan He, Yixin Cao, Meng Liu, and Tat-Seng Chua. 2019. KGAT: Knowledge Graph Attention Network for Recommendation. In *KDD*.
- [45] Chuhan Wu, Fangzhao Wu, Lingjuan Lyu, Tao Qi, Yongfeng Huang, and Xing Xie. 2022. A federated graph neural network framework for privacy-preserving personalization. *Nature Communications* 13, 1 (2022), 1–10.
- [46] Shiwen Wu, Fei Sun, Wentao Zhang, and Bin Cui. 2020. Graph neural networks in recommender systems: a survey. *arXiv preprint arXiv:2011.02260* (2020).

- [47] Han Xie, Jing Ma, Li Xiong, and Carl Yang. 2021. Federated graph classification over non-iid graphs. *Advances in Neural Information Processing Systems* 34 (2021), 18839–18852.
- [48] Bin Xu, Jiajun Bu, Chun Chen, and Deng Cai. 2012. An exploration of improving collaborative recommender systems via user-item subgroups. In *Proceedings of the 21st international conference on World Wide Web*. 21–30.
- [49] Liu Yang, Ben Tan, Vincent W Zheng, Kai Chen, and Qiang Yang. 2020. Federated recommendation systems. In *Federated Learning*. Springer, 225–239.
- [50] Qiang Yang, Yang Liu, Tianjian Chen, and Yongxin Tong. 2019. Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)* 10, 2 (2019), 1–19.
- [51] Jiangchao Yao, Feng Wang, KunYang Jia, Bo Han, Jingren Zhou, and Hongxia Yang. 2021. Device-Cloud Collaborative Learning for Recommendation. *arXiv preprint arXiv:2104.06624* (2021).
- [52] Hongzhi Yin, Bin Cui, Zi Huang, Weiqing Wang, Xian Wu, and Xiaofang Zhou. 2015. Joint modeling of users' interests and mobility patterns for point-of-interest recommendation. In *Proceedings of the 23rd ACM international conference on Multimedia*. 819–822.
- [53] Rex Ying, Ruining He, Kaifeng Chen, Pong Eksombatchai, William L Hamilton, and Jure Leskovec. 2018. Graph convolutional neural networks for web-scale recommender systems. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 974–983.
- [54] Ke Zhang, Carl Yang, Xiaoxiao Li, Lichao Sun, and Siu Ming Yiu. 2021. Sub-graph federated learning with missing neighbor generation. *Advances in Neural Information Processing Systems* 34 (2021), 6671–6682.
- [55] Shijie Zhang, Hongzhi Yin, Tong Chen, Zi Huang, Lizhen Cui, and Xiangliang Zhang. 2021. Graph embedding for recommendation against attribute inference attacks. In *Proceedings of the Web Conference 2021*. 3002–3014.
- [56] Shijie Zhang, Hongzhi Yin, Tong Chen, Zi Huang, Quoc Viet Hung Nguyen, and Lizhen Cui. 2022. Pipattack: Poisoning federated recommender systems for manipulating item promotion. In *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*. 1415–1423.
- [57] Zhi-Dan Zhao and Ming-Sheng Shang. 2010. User-based collaborative-filtering recommendation algorithms on hadoop. In *2010 third international conference on knowledge discovery and data mining*. IEEE, 478–481.
- [58] Ruiqi Zheng, Liang Qu, Bin Cui, Yuhui Shi, and Hongzhi Yin. 2023. AutoML for Deep Recommender Systems: A Survey. *ACM Trans. Inf. Syst.* (jan 2023). <https://doi.org/10.1145/3579355> Just Accepted.
- [59] Guorui Zhou, Xiaoqiang Zhu, Chenru Song, Ying Fan, Han Zhu, Xiao Ma, Yanghui Yan, Junqi Jin, Han Li, and Kun Gai. 2018. Deep interest network for click-through rate prediction. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1059–1068.