Hierarchical Personalized Federated Learning for User Modeling

Jinze Wu[1], Qi Liu[1,*], Zhenya Huang[1], Yuting Ning[1], Hao Wang[1], Enhong Chen[1]
Jinfeng Yi[2], Bowen Zhou[2]
[1]Anhui Province Key Laboratory of Big Data Analysis and Application,
School of Computer Science and Technology, University of Science and Technology of China
[2]JD AI Research
{hxwjz,ningyt,wanghao3}@mail.ustc.edu.cn,{qiliuql,huangzhy,cheneh}@ustc.edu.cn,
{yijinfeng,bowen.zhou}@jd.com

Reporter: Jinze Wu

Outline

Background

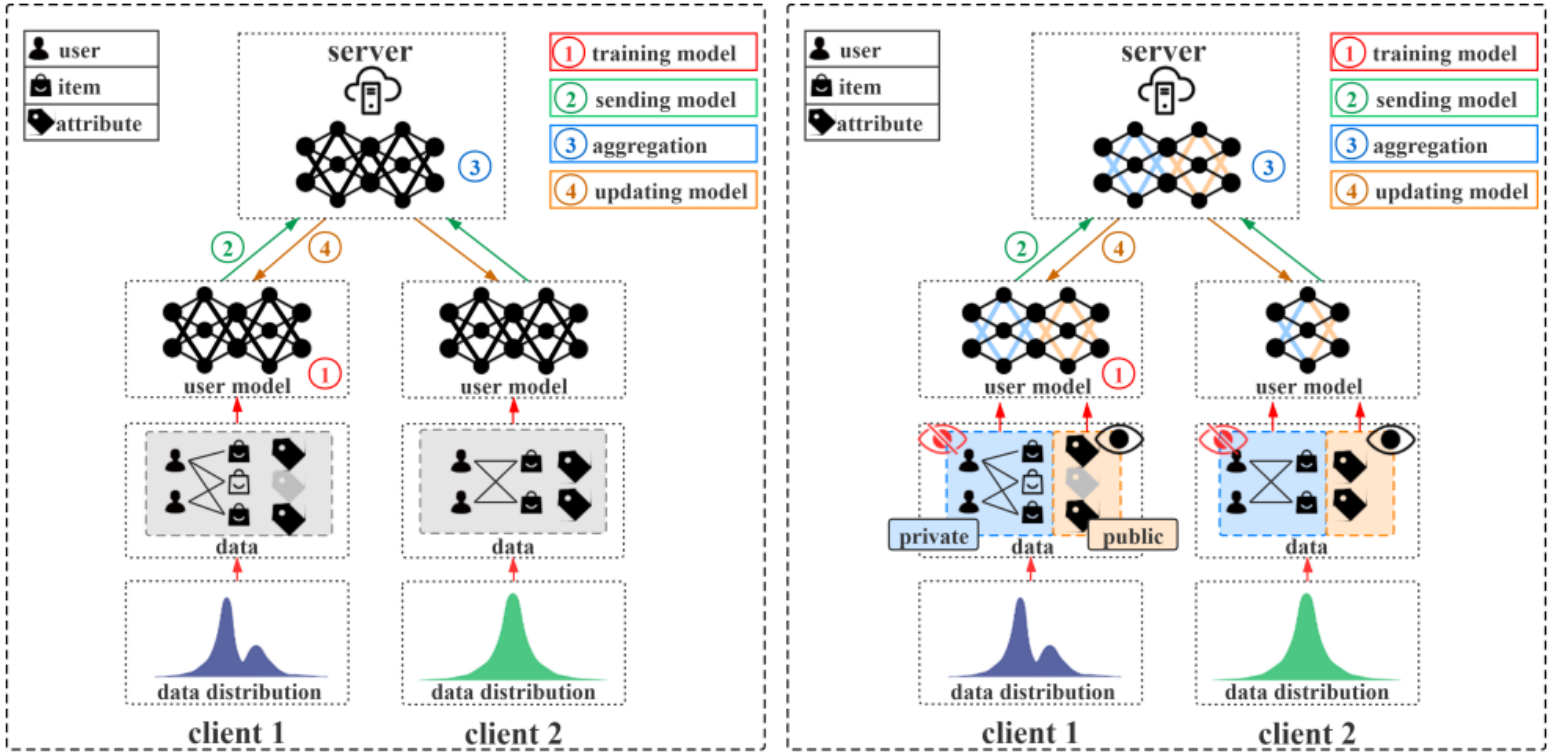
- User modeling
  - An important basis to capture useful **potential characteristics** with the reliance on **personal data**
  - User modeling has been applied to multiple typical, such as modeling capabilities or preferences from users
  - User modeling processes usually centralized training with data aggregated, which causes privacy leakage
- Federated Learning
  - Federated Learning (FL) refers to building and aggregating user models while leaving private data isolated so that preserves the data privacy
  - Federated user modeling receives widespread attention for potential of secure distributed user modeling

Background

- Problems and Challenges
  - federated user modeling for isolated scenarios with the **inconsistent clients**
  - Statistical heterogeneity: personal records for user modeling are usually non-independently identically distributed (Non-IID)
  - Privacy heterogeneity: different information in user modeling have different levels of privacy
  - Model heterogeneity: the user model structures among different clients are often different

Background

- Problems and Challenges
  - In this paper, we design hierarchical personalized federated learning to overcome the challenges



(a) standard federated learning

(b) hierarchical personalized federated learning

Outline

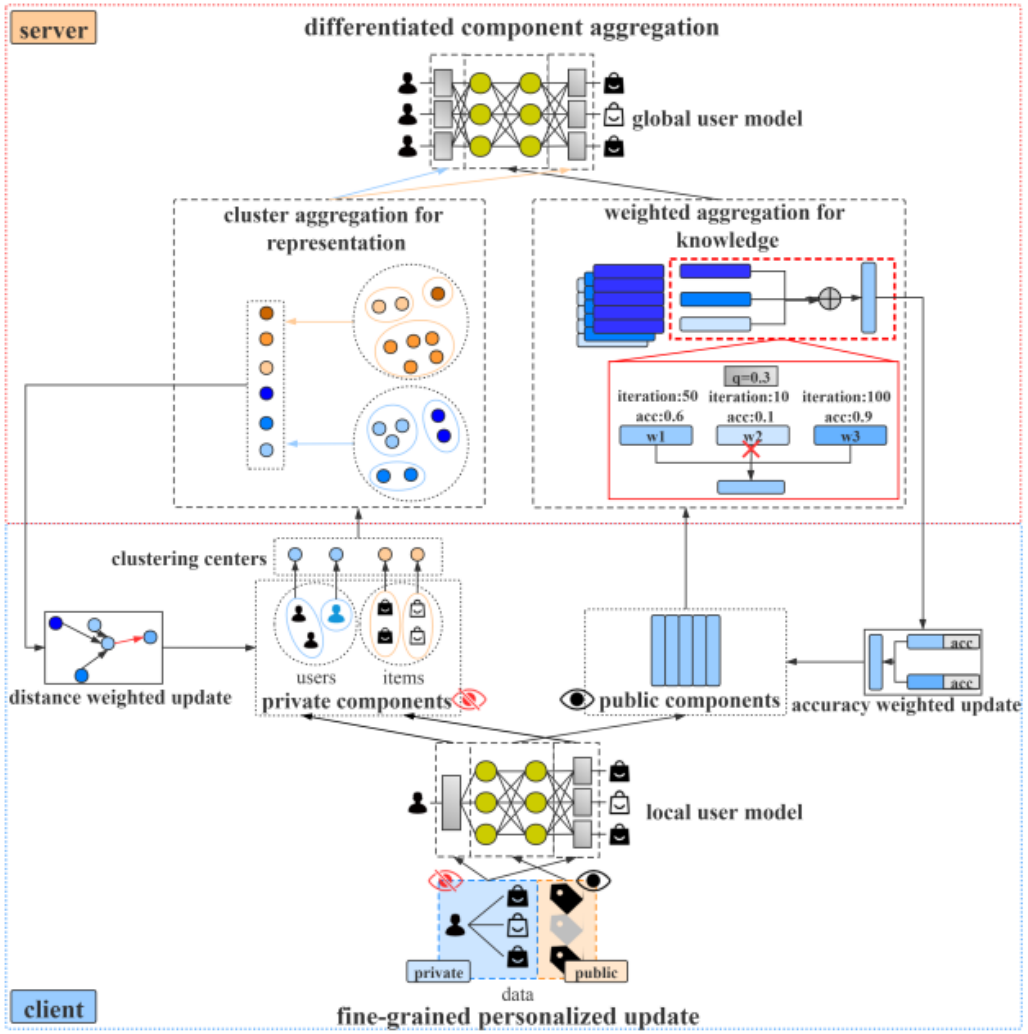
Problem Definition

- Problem Definition
  - Given:
    - C : set of clients
    - U_c: set of users
    - V_c: set of items with K attributes
    - (u, v, g): a triplet representing the interaction result between user u and item v
  - Goal: train |C| local user models, where the c-th user model can model the potential characteristics of users and predict the interaction results
  - Hierarchical information
    - *Public information: it refers to the information that contains the prior domain knowledge so that it can be shared among clients. In this case, the public information is relatively private and incompetent to expose the sensitive user information.*
    - *Private information: it is the information which is proprietary for clients and represents the unique distributions of users and items among each client. Apparently, it is with strictly privacy and needs to be protected.*

Outline

HPFL Framework

- HPFL Framework
  - Server
    - fuse heterogeneous local user models by different components with the **differentiated component aggregation strategy**.
  - Client
    - train and deliver the different components of user model
    - update a personalized user model using the **fine-grained personalized update strategy**



HPFL Framework

- Client
  - Upload phase
    - initializes and trains a general user model named GUM
    - delivers the local user model by different component
  - Update phase
    - regulate a fine-grained personalized update strategy to fuse the local personalized information and global generalized information
    - For public component: client i add the local attribute knowledge vector and the global attribute knowledge on attribute k via the corresponding accuracy

$$\mathbf{c}_{k,i}^{t} = \mathbf{c}_{k,i}^{t} \times Acc_{k,i}^{t} + \mathbf{c}_{k}^{t,g} \times (1 - Acc_{k,i}^{t}).$$

    - For private component: the client i distance weighted update new representation with the affects of all global cluster centers

$$\mathbf{Emb}^{g} = \sum_{n=1}^{N} \frac{||\mathbf{Emb}_{j,i} - \Theta_{r,n}^{g}|| \times \Theta_{r,n}^{g}}{\sum_{m=1}^{N} ||\mathbf{Emb}_{j,i} - \Theta_{r,m}^{g}||},$$

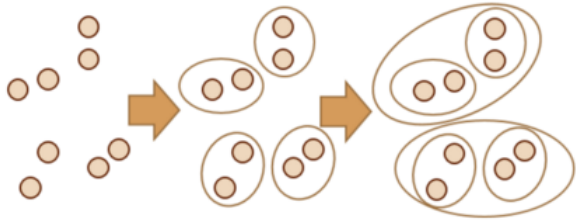
$$\mathbf{Emb}_{j,i} = \mathbf{Emb}_{j,i} \times \mathbf{Acc}_{i} + \mathbf{Emb}^{g} \times (1 - \mathbf{Acc}_{i}).$$

HPFL Framework

- Server
  - Aggregation for public components
    - weighted aggregation of the same attribute to obtain the global public components
    - fuse each knowledge vector that represents knowledge information on attribute k from clients i in reference to both the number of iterations on attribute k as well as the local validation accuracy

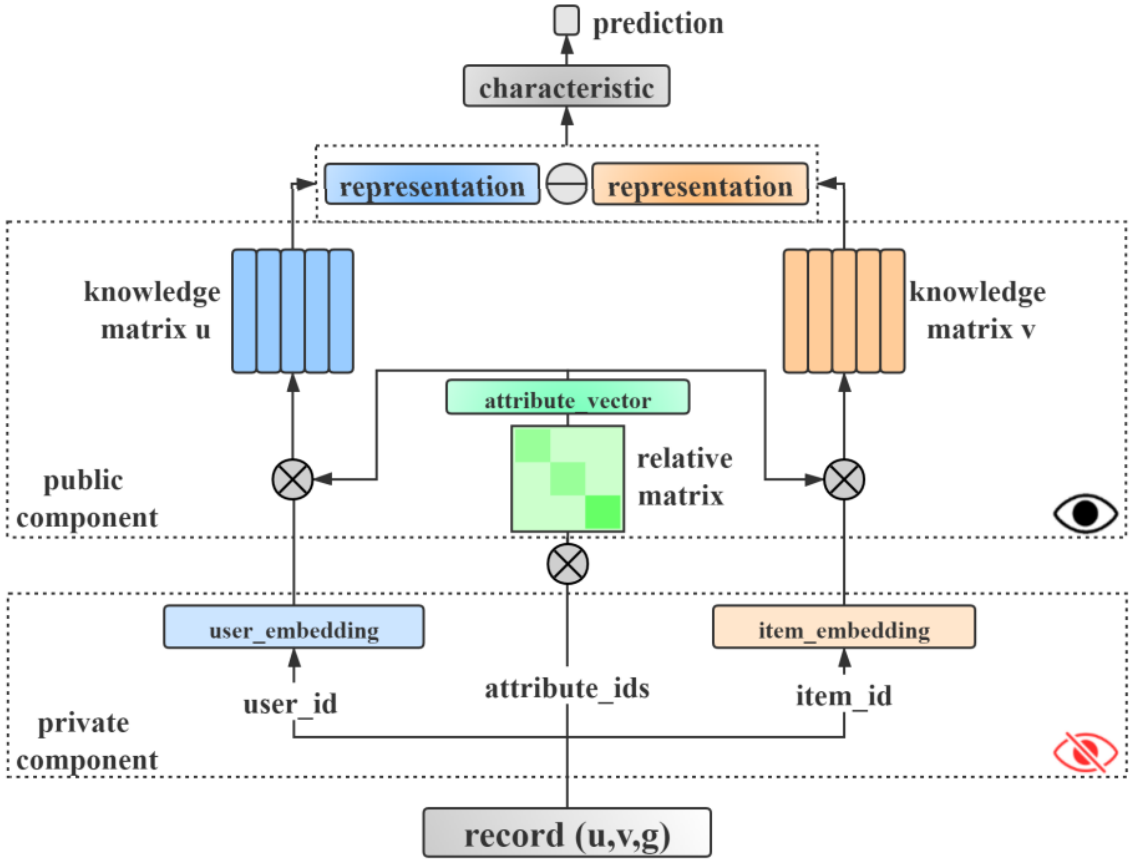
$$\mathbf{c}_k^{t,g} = \frac{\sum_{i=1}^{C} \delta(Acc_{k,i}^t, p) \times (I_{k,i}^t \times Acc_{k,i}^t) \times \mathbf{c}_{k,i}^t}{\sum_{i=1}^{C} \delta(Acc_{k,i}^t, p) \times (I_{k,i}^t \times Acc_{k,i}^t)}.$$

  - Aggregation for private components
    - further clustering with all the cluster centers



HPFL Framework

- Local User Model
  - General User Model (GUM), which is flexible, explainable and capable of deep representation

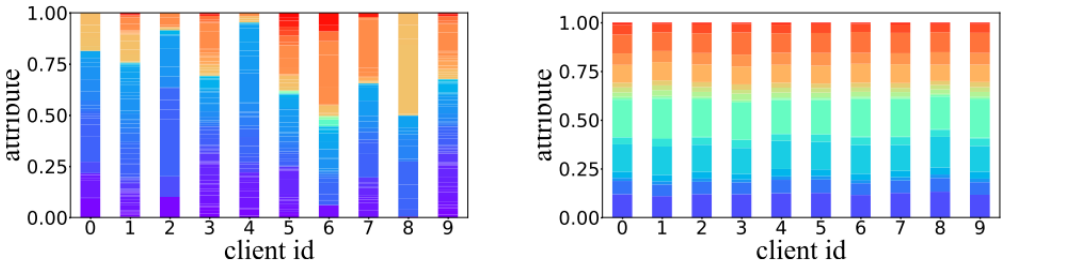


Outline

Experiment

- Dataset
  - ASSIST
  - MovieLens
- Data analysis
  - IID or Non-IID distributions on the attributes
- Baseline methods
  - centrally training methods:
    NCD, NCF
  - Traditional federated methods:
    FedSGD, FedAvg, Fednoise, FedProx, FedAtt
  - Our methods:
    HPFL-K, HPFL-R, HPFL
- Evaluation metrics
  - ACC, AUC, RMSE

Statistics	ASSIST	MovieLens
# of clients	59	10
# of records	327,058	96,538
# of users	3,477	925
# of items	17,561	1,679
# of attributes	122	19
# attributes per item	1.20	1.72
# attributes per record	1.20	2.21



(a) Distribution of attributes in ASSIST. (b) Distribution of attributes in MovieLens.

Experiment

- Accuracy performances
  - Task: student performance prediction and user rating prediction
- Observation
  - GUM model performs better than NCF and NCD: capable of deep representation
  - federated methods perform better, HPFL performs best: harness more information

Methods	ASSIST			MovieLens		
	ACC	AUC	RMSE	ACC	MAE	RMSE
NCD	0.727	0.749	0.430	-	-	-
NCF	-	-	-	0.385	0.759	0.988
GUM	**0.736**	**0.774**	**0.421**	**0.397**	**0.745**	**0.946**
Distributed	0.699	0.718	0.442	0.389	0.802	1.001
FedSGD	0.704	0.716	0.453	0.341	0.933	1.111
FedAvg	0.703	0.724	0.445	0.397	0.802	0.993
Fednoise	0.701	0.722	0.441	0.387	0.804	1.019
FedProx	0.704	0.725	0.444	0.405	0.798	0.989
FedAtt	0.715	0.727	0.438	0.404	0.796	0.989
HPFL-K	0.715	0.730	0.437	0.403	0.792	0.987
HPFL-R	0.723	0.738	0.433	0.405	0.798	0.991
HPFL	**0.726**	**0.742**	**0.431**	**0.407**	**0.786**	**0.978**

Experiment

- Ranking effectiveness
  - Task: the partial orders of user preferences
- Observation
  - GUM performs better than centralised methods: high-dimensional user model benefits
  - HPFL performs outstanding results: both components benefit
  - distributed training method performs a comparable result in NDCG: standard federated methods bring a coordination among clients

(a) Results of DOA and NDCG on ASSIST

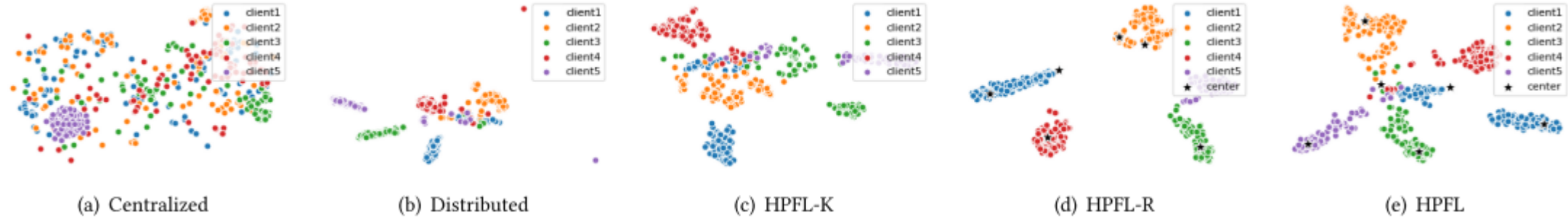
	NCD	GUM	Distributed	FedSGD	FedAvg	Fednoise	FedProx	FedAtt	HPFL-K	HPFL-R	HPFL
DOA	0.755	**0.773**	0.736	0.736	0.741	0.721	0.743	0.749	**0.756**	0.743	**0.758**
NDCG	0.826	**0.864**	0.837	0.825	0.833	0.831	0.835	0.834	0.834	**0.849**	**0.856**

(b) Results of DOA and NDCG on MovieLens

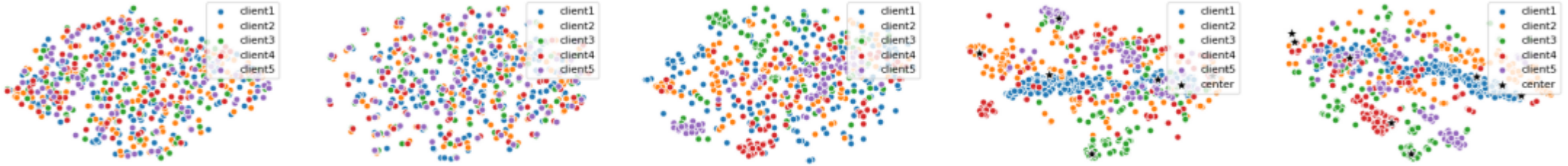
	NCF	GUM	Distributed	FedSGD	FedAvg	Fednoise	FedProx	FedAtt	HPFL-K	HPFL-R	HPFL
DOA	0.505	**0.590**	0.537	0.519	0.668	0.539	0.668	0.669	**0.678**	0.672	**0.699**
NDCG	0.855	**0.869**	0.893	0.858	0.891	0.864	0.892	0.891	0.896	**0.898**	**0.910**

Experiment

- Modeling rationality
  - Task: analyze the rationality of user models at the parameter level
- Observation
  - private components in user models are not distinguishable
  - HPFL have advantages to mine peculiarity of clients from user characteristics in user modeling



(a) Centralized (b) Distributed (c) HPFL-K (d) HPFL-R (e) HPFL



(a) Centralized (b) Distributed (c) HPFL-K (d) HPFL-R (e) HPFL

Outline

Conclusion & Future work

- Overall results
  - Design a novel client-server architecture framework to enable federated learning to be applied in user modeling tasks with inconsistent clients.
  - a fine-grained personalized update strategy and a differentiated component aggregation strategy were explored to flexibly fuse heterogeneous user models
- Future work
  - Consider the data characteristics to improve the federated strategy
  - Design a platform and apply the technical details of HPFL

Q & A

Thanks!

hxwjz@mail.ustc.edu.cn