FedCSR: A Federated Framework for Multi-Platform Cross-Domain Sequential Recommendation with Dual Contrastive Learning

Dongyi Zheng[1,2], Hongyu Zhang[3], Jianyang Zhai[1,2], Zhong Lin[3], Lingzhi Wang[3], Jiyuan Feng[1,3], Xiangke Liao[2,1], Yonghong Tian [4,1], Nong Xiao[2,1], Qing Liao[3,1]*

[1]Pengcheng Laboratory, [2]Sun Yat-sen University,
[3]Harbin Institute of Technology (Shenzhen), [4]Peking University

zhengdy23@mail2.sysu.edu.cn, {orion-orion, 24b951017, fengjy}@stu.hit.edu.cn
zhaijy01@pcl.ac.cn, lzwang1120@gmail.com, {xkliao, nongxiao}@nudt.edu.cn
yhtian@pku.edu.cn, liaoqing@hit.edu.cn

Abstract

Cross-domain sequential recommendation (CSR) has garnered significant attention. Current federated frameworks for CSR leverage information across multiple domains but often rely on user alignment, which increases communication costs and privacy risks. In this work, we propose FedCSR, a novel federated cross-domain sequential recommendation framework that eliminates the need for user alignment between platforms. FedCSR fully utilizes cross-domain knowledge to address the key challenges related to data heterogeneity both inter- and intra-platform. To tackle the heterogeneity of data patterns between platforms, we introduce Model Contrastive Learning (MCL) to reduce the gap between local and global models. Additionally, we design Sequence Contrastive Learning (SCL) to address the heterogeneity of user preferences across different domains within a platform by employing tailored sequence augmentation techniques. Extensive experiments conducted on multiple real-world datasets demonstrate that FedCSR achieves superior performance compared to existing baseline methods[1].

1 Introduction

Cross-domain sequential recommendation (CSR) has attracted wide attention in various online platforms (Liu et al., 2023b; Lu et al., 2023; Song et al., 2024; Li et al., 2024; Zhai et al., 2023). By leveraging information from multiple domains, CSR enhances recommendation, providing users with a richer discovery experience. To enable collaboration across platforms while ensuring data security, federated learning (Zhang et al., 2024c; Yu et al., 2024; Cong et al., 2023; Guo et al., 2024b; Liu et al., 2024) has been applied to CSR, allowing multiple platforms to work together without sharing raw data.

In this work, we tackle the problem of federated cross-domain sequential recommendation, enabling different platforms to collaboratively train a robust CSR model without sharing raw data. Unlike previous methods (Zhang et al., 2024a; Liu et al., 2023a; Guo et al., 2024a), which require user alignment across platforms to establish connections between domains, resulting in increased communication costs and privacy concerns, our approach focuses on fully utilizing intra-platform cross-domain knowledge and eliminates the need for user alignment between platforms.

However, incorporating cross-domain knowledge between platforms without user alignment presents a significant challenge due to the heterogeneity of data across platforms, which means that the data patterns, such as recommendation item distribution and user preferences, can differ greatly between platforms. For instance, an e-commerce company operating globally may have multiple platforms with distinct data characteristics due to cultural differences. This sequence pattern heterogeneity will lead to an obvious drift between each platform's local model and the aggregated global model. As training progresses, this drift intensifies, making it difficult for the global model to perform well across all platforms. Furthermore, even within a single platform, users often display distinct preferences across various domains, which leads to heterogeneity in sequence representations. This issue becomes more pronounced when certain domains have sparse data, further complicating the modeling of user preferences and reducing the effectiveness of cross-domain knowledge transfer.

To address these challenges, we introduce a novel framework called FedCSR, which incorporates dual contrastive learning for local training on platforms, eliminating the need for additional user alignment engineering across platforms. Specifi-

*Corresponding author.

[1]We release our code and configuration files at https://github.com/zdy769243418/FedCSR-v1

Proceedings of the 31st International Conference on Computational Linguistics, pages 8699–8713
January 19–24, 2025. ©2025 Association for Computational Linguistics

cally, to mitigate the sequence pattern heterogeneity between platforms, we develop Model Contrastive Learning (MCL) to reduce the information gap between local and global models. Additionally, to tackle the issue of intra-platform sequence representation heterogeneity across domains, we design Sequence Contrastive Learning (SCL). In SCL, we first implement a tailored cross-domain sequence augmentation method that employs replace and reorder operations to enhance the diversity and balance of inter-domain sequences. We then encourage the original and augmented user sequence representations to predict each other, maximizing their agreement and improving the overall quality of the sequential representations.

Furthermore, we conduct extensive experiments on multiple datasets to demonstrate the superiority of our FedCSR framework. We analyze algorithm complexity, communication efficiency, and perform ablation studies to further evaluate its effectiveness. Additional experiments from various perspectives provide deeper insights into the framework's performance and its applicability to CSR.

In brief, our contributions are summarized as:

- We propose FedCSR, a novel federated multi-platform cross-domain sequential recommendation framework, enabling platforms with multiple domains to collaboratively train an effective CSR model without requiring user alignment between platforms.
- We design MCL and SCL mechanisms to address the inter- and intra-platform data pattern heterogeneity, respectively, improving the quality and robustness of data representation modeling across different platforms and domains.
- Extensive experiments on multiple real-world datasets demonstrate that FedCSR consistently outperforms baseline methods.

2 Related Work

Cross-Domain Recommendation. The cross-domain recommendation (Zheng et al., 2022; Li et al., 2020a) refers to the process of leveraging knowledge from multiple distinct domains to improve recommendation accuracy. It involves transferring insights or preferences from one domain (e.g., movies) to generate better recommendations in another domain (e.g., books). With advancements in transfer learning, several methods (Man et al., 2017; Zhu et al., 2022, 2021) utilize two separate neural networks and a transfer module to map user representations between domains. Additionally, other approaches (Zhu et al., 2019; Hu et al., 2018; Xie et al., 2022) focus on identifying shared information across domains by sharing parts of the network architecture.

Cross-Domain Sequential Recommendation. The cross-domain sequential recommendation captures the temporal order of user behaviors to improve recommendations. Earlier works (Ma et al., 2019; Sun et al., 2021) introduced parallel information-sharing networks to encode cross-domain behavior sequences. Subsequent methods (Guo et al., 2021; Cao et al., 2022a; Ma et al., 2022) utilized graph-based approaches to establish sequential connections between items across domains. For instance, Zhang et al. (2023a) proposed a graph convolutional network to capture collaborative filtering signals from both domains. Recently, techniques such as reinforcement learning (Guo et al., 2022), triple learning (Ma et al., 2024), and mixed-interest networks (Lin et al., 2024; Liu et al., 2023c; Xu et al., 2024) have been applied to enhance cross-domain representation.

Federated Cross-Domain Sequential Recommendation. To address privacy concerns in cross-domain recommendation, several methods (Cai et al., 2022; Zheng et al., 2023; Wang et al., 2021; Zhang et al., 2024b) have been proposed, using techniques like differential privacy (Chen et al., 2023) and generative adversarial networks (Liao et al., 2023) for secure data transfer. Besides, to incorporate sequential information, Zhang et al. (2024a) and Liu et al. (2023a) have been proposed for federated cross-domain sequential recommendation. Guo et al. (2024a) further utilizes prompt learning to enhance cross-domain representation.

Despite these advances, challenges like user alignment, which complicates implementation and poses privacy risks, and the limited use of single-domain data persist. Our method eliminates user alignment and fully utilizes multi-domain data, enhancing scalability and practicality.

3 Methodology

In this section, we introduce the overall framework of FedCSR, as illustrated in Figure 1.

3.1 Problem Formulation

Federated Training. We consider a central server and K local platforms. All platforms share

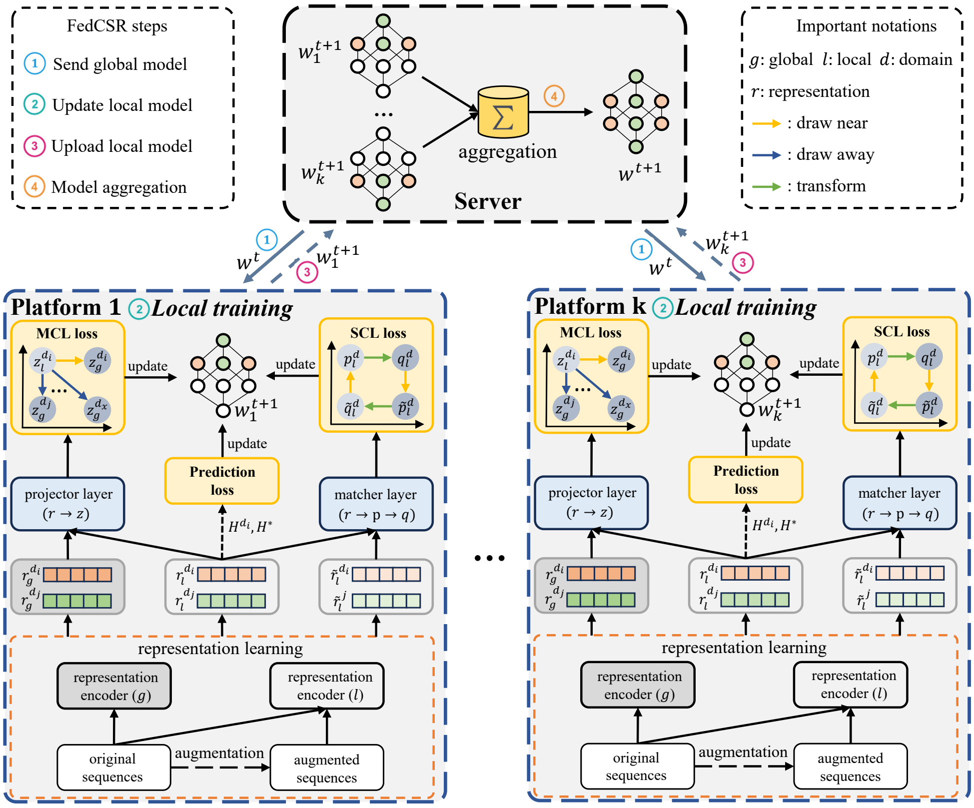


Figure 1: Overview of FedCSR. Each platform leverages its local cross-domain dataset to learn sequential representations and compute SCL, MCL, and prediction loss. The local model is optimized through multi-task learning, while the server aggregates the local models through weighted updates to form the latest global model.

the same CSR model structure, denoted as w. In each communication round t, the central server distributes the global model w^t to all platforms. Each platform k then receives the global model w^t and trains it using its local private dataset $\mathcal{D}_k$, resulting in a locally trained model denoted as w_k^t.

Cross-Domain Sequential Recommendation. Each platform k holds a private cross-domain sequence dataset $\mathcal{D}_k$, which contains m domains $d_1, d_2, ..., d_m$. Each domain d_i includes an item set $\mathcal{V}^{d_i}$, and the combined item set for all domains is represented as $\mathcal{V}^*$. The item feature dimension is n. Each user in $\mathcal{D}_k$ has $m+1$ interaction sequences $s^{d_1}, s^{d_2}, ..., s^{d_m}, s^*$. Here, s^* represents the cross-domain sequence, which is formed by merging the individual sequences $s^{d_1}, s^{d_2}, ..., s^{d_m}$ in chronological order. The CSR model leverages these $m+1$ sequences for each user to generate more accurate recommendations within each domain d_i.

3.2 Base Representation Encoder

Graph neural networks combined with self-attention have proven effective in capturing both inter-sequence item relationships and intra-sequence order patterns in user sequences (Cao et al., 2022a; Zhang et al., 2023a; Guo et al., 2021). We use an attentional graph neural network, referred to as GAEncoder, to extract sequential representations for each domain. Each GAEncoder consists of an item embedding layer, a graph neural network, and a self-attention layer.

For each domain d_i, we construct a binary item-item matrix $\boldsymbol{A}^{d_i} \in \mathbb{R}^{|\mathcal{V}^{d_i}|*|\mathcal{V}^{d_i}|}$ based on the temporal order of item occurrences in the sequences. If item i appears before item j, then $\boldsymbol{A}^{d_i}_{ij} = 1$; otherwise $\boldsymbol{A}^{d_i}_{ij} = 0$. We also generate a cross-domain item-item matrix $\boldsymbol{A}^* \in \mathbb{R}^{|\mathcal{V}^*|*|\mathcal{V}^*|}$ from the combined cross-domain sequences. For a sequence $s \in s^{d_1}, s^{d_2}, \dots, s^{d_m}, s^*$ and its associated item-item matrix $\boldsymbol{A}$, the encoding process is:

$$\boldsymbol{H} = \text{GAEncoder}(\text{Embedding}(s), \boldsymbol{A}) \quad (1)$$

where $\boldsymbol{H} \in \mathbb{R}^{|s|*n}$ denotes the representations at different timestamps, $|s|$ represents the sequence length and n is the feature dimension.

As each user has m single-domain sequences, along with a cross-domain sequence, we employ $m+1$ encoders to encode them separately, yielding $\{\boldsymbol{H}^{d_1}, \boldsymbol{H}^{d_2}, ..., \boldsymbol{H}^*\}$. Then, we aggregate and average all timestamp representations to obtain the sequential representation for each domain d_i: $\boldsymbol{r}^{d_i} = \text{Avg}(\boldsymbol{H}^{d_i} + \boldsymbol{H}^*)$. For simplicity, the entire sequential representation learning process is:

$$\boldsymbol{r}^{d_i} = f(s^{d_i}, s^*|w), \tag{2}$$

where $\boldsymbol{r}^{d_i} \in \mathbb{R}^{1*n}$ represents the sequential representation for domain d_i and f denotes the complete representation learning function for CSR model w.

3.3 Model Contrastive Learning

To address sequence pattern heterogeneity among platforms, we propose Model Contrastive Learning (MCL). MCL creates contrastive signals between local and global user representations to mitigate model drift between platforms and central server.

For each domain, MCL treats the domain's local and global user representations as positive samples. Since irrelevant domain information can trigger negative transfer issues (Cao et al., 2022b, 2023), MCL treats other domain's global sequential representations as negative samples. This helps disentangle user representations across domains and improves the representation quality for each domain. We describe the MCL process as follows.

Firstly, let's consider platform k conducting local training. It receives the global model w^t from the server and initializes its local model as w_k^t. For each domain d_i, given one user's sequences s^{d_i} and s^*, we extract the sequential representation $\boldsymbol{r}_l^{d_i}$ from the updating local model w_k^t. Additionally, we obtain the sequential representation $\boldsymbol{r}_g^{d_i}$ from the global model w^t.

$$\boldsymbol{r}_l^{d_i} = f(s^{d_i}, s^*|w_k^t), \boldsymbol{r}_g^{d_i} = f(s^{d_i}, s^*|w^t) \tag{3}$$

where $\boldsymbol{r}_l^{d_i} \in \mathbb{R}^{1*n}$ and $\boldsymbol{r}_g^{d_i} \in \mathbb{R}^{1*n}$ respectively represent the local and global sequential representations for domain d_i.

Subsequently, to enhance contrastive learning's ability to distinguish different representations, we define a projector to project them to a low-dimensional space. Due to feature dimension compression, features in the low-dimensional space are more representative and robust. Moreover, the distance between features can better reflect the correlation between real sequences. We employ a nonlinear layer as the projector, denoted as $\varphi(\cdot)$.

$$\boldsymbol{z}_l^{d_i} = \varphi(\boldsymbol{r}_l^{d_i}), \boldsymbol{z}_g^{d_i} = \varphi(\boldsymbol{r}_g^{d_i}) \tag{4}$$

where $\boldsymbol{z}_l^{d_i} \in \mathbb{R}^{1*|\frac{n}{2}|}$ and $\boldsymbol{z}_g^{d_i} \in \mathbb{R}^{1*|\frac{n}{2}|}$ are the local and global sequential representations in the low-dimensional space for domain d_i.

Next, in order to enable the local model to capture the broader sequence patterns from a global view and generalize well to unseen data, we reduce the information gap of $(\boldsymbol{z}_l^{d_i}, \boldsymbol{z}_g^{d_i})$ for each domain d. The information gap function $ig(\cdot)$ is defined as:

$$ig(a, b) = -\exp(\text{sim}(a, b)/\tau), \tag{5}$$

where $\text{sim}(\cdot)$ measures the cosine similarity, and τ denotes a temperature. If a and b are similar, the information gap between them will be small.

Finally, similar to NT-Xent (Sohn, 2016) loss, we define our model contrastive loss as:

$$\mathcal{L}_{MCL} = \sum_{i=1}^{m} \log \frac{ig(\boldsymbol{z}_l^{d_i}, \boldsymbol{z}_g^{d_i})}{g(\boldsymbol{z}_l^{d_i}, \boldsymbol{z}_g^{d_i}) + \sum_j g(\boldsymbol{z}_l^{d_i}, \boldsymbol{z}_g^{d_j})} \tag{6}$$

where $j \in \{x|1 \leq x \leq m, x \neq i\}$. During the local training, we consider $(\boldsymbol{z}_l^{d_i}, \boldsymbol{z}_g^{d_i})$ as positive pair. These positive pairs are drawn near each other to encourage the local model to reduce the information gap to the global model. Additionally, we treat all $(\boldsymbol{z}_l^{d_i}, \boldsymbol{z}_g^{d_j})$ as negative pairs. These negative pairs are intentionally drawn away during the local training to encourage the local model to distinguish between representations from different domains.

3.4 Sequence Contrastive Learning

In this subsection, we introduce Sequence Contrastive Learning (SCL) to tackle domain-specific sequence bias. SCL employs a novel cross-domain sequence augmentation technique to boost diversity and balance in inter-domain sequences, and then aligns original and augmented representations to enhance their agreement and overall quality.

Cross-Domain Sequence Augmentation. At this stage, we utilize the replace and reorder operations to augment the user sequences.

Since the cross-domain sequence is more informative than single-domain sequences, we first operate cross-domain sequence s^* to obtain the augmented $\widetilde{s}^*$. Then, we utilize $\widetilde{s}^*$ to obtain the corresponding augmented single-domain sequence $\widetilde{s}^{d_i}$ for each domain d_i.

Firstly, we represent the cross-domain sequence as $s^* = [v_1, v_2, ..., v_l, v_{l+1}, v_{l+2}, ..., v_{2l}]$. Based on its length, we evenly split it into two subsequences. One subsequence is denoted as $s_x^* = [v_1, v_2, ..., v_l]$, and the other subsequence is denoted as $s_y^* = [v_{l+1}, v_{l+2} ..., v_{2l}]$.

Then, for the augmentation of s_x^*, we randomly select some items in s_x^* with a ratio α. For each item in selected items, it is replaced by an item sampled from another domain. The augmented sequence of s_x^* is denote as:

$$\widetilde{s}_x^* = \{\widetilde{v}_1, \widetilde{v}_2, ..., \widetilde{v}_l\} \tag{7}$$

where each item $\widetilde{v}_i$ in $\widetilde{s}_x^*$ can be expressed as:

$$\widetilde{v}_i = \begin{cases} v_i, & \text{if } v_i \notin \{s_x^* | \alpha\} \\ \text{sample}(\mathcal{V}^{d_j}), & \text{if } I_i \in \{s_x^* \cap \mathcal{V}^{d_i}\} \end{cases} \tag{8}$$

Next, for the augmentation of s_y^*, we select the latest interaction items with a ratio α. We randomly reorder the selected items to obtain the augmented sequence of s_y^*:

$$\widetilde{s}_y^* = [v_{l+1}, v_{l+2}, \ldots, \widetilde{v}_{|l\cdot(1+\alpha)|}, \ldots, \widetilde{v}_{2l}] \tag{9}$$

The augmented cross-domain sequence of s^* is formed by combining $\widetilde{s}_x^*$ and $\widetilde{s}_y^*$, expressed as:

$$\widetilde{s^*} = [\widetilde{v}_1, \ldots, \widetilde{v}_l, v_{l+1}, \ldots, \widetilde{v}_{|l\cdot(1+\alpha)|}, \ldots, \widetilde{v}_{2l}] \tag{10}$$

Finally, we derive the corresponding augmented single-domain sequences $\{\widetilde{s}^{d_1}, \widetilde{s}^{d_2}, ..., \widetilde{s}^{d_m}\}$ from $\widetilde{s}^*$. For domain d_i, each item in $\widetilde{s}^{d_i}$ can be expressed as follows:

$$\widetilde{s}^{d_i} = \{v_i \mid v_i \in \widetilde{s}^*, \text{ if } v_i \text{ belongs to } \mathcal{V}^{d_i}\} \tag{11}$$

We use both replacement and reordering to augment sequences. Replacement boosts item count in sparse domains, while reordering disrupts recent item order to enhance sequence diversity and balance across domains, thereby enhancing the diversity and proportionality of sequences between domains.

Sequence Contrastive Signal Construction. In this stage, we construct a symmetric contrastive objective in each domain to draw near the similarity between the original and augmented sequential representations, thereby improving the quality of sequential representation.

We first feed the augmented sequences $\{\widetilde{s}^{d_i}, \widetilde{s}^{d_2}, ..., \widetilde{s}^{d_m}, \widetilde{s}^*\}$ into the local representation encoders to obtain the augmented sequential representation for each domain d_i.

$$\widetilde{\boldsymbol{r}}_l^{d_i} = f(\widetilde{s}^{d_i}, \widetilde{s}^* | w_k^t) \tag{12}$$

where $\widetilde{\boldsymbol{r}}_l^{d_i} \in \mathbb{R}^{1*n}$ is the augmented sequential representation for domain d_i and w_k^t is the local updating CSR model of platform k.

In order to break the full symmetry in the model and avoid the model learning a simple identity mapping, thus collapsing to a simple solution, we introduce a nonlinear transformation layer. Different from the bottleneck structure (Chen and He, 2021), we define the transformation layer as a broadened structure. The definition of the nonlinear transformation layer is as follows:

$$\boldsymbol{p}_l^{d_i} = t(\boldsymbol{r}_l^{d_i}), \widetilde{\boldsymbol{p}}_l^{d_i} = t(\widetilde{\boldsymbol{r}}_l^{d_i}) \tag{13}$$

where t is a two-layer nonlinear network with a broadened structure of $n \rightarrow 2n \rightarrow n$. $\boldsymbol{p}_l^{d_i} \in \mathbb{R}^{1*n}$ and $\widetilde{\boldsymbol{p}}_l^{d_i} \in \mathbb{R}^{1*n}$ are the transformed sequential representations of domain d_i.

Next, we employ an nonlinear layer as a matcher, denoted as $g(\cdot)$ for each domain d_i to transform the representation of one view to the other view.

$$\boldsymbol{q}_l^{d_i} = g(\boldsymbol{p}_l^{d_i}), \widetilde{\boldsymbol{q}}_l^{d_i} = g(\widetilde{\boldsymbol{p}}_l^{d_i}) \tag{14}$$

where $\boldsymbol{q}_l^{d_i}$ and $\widetilde{\boldsymbol{q}}_l^{d_i}$ are the mapped features.

Finally, we define the symmetric contrastive loss function by mutual prediction to draw near $\{\boldsymbol{q}_l^{d_i}, \widetilde{\boldsymbol{p}}_l^{d_i}\}$ and $\{\widetilde{\boldsymbol{q}}^{d_i}, \boldsymbol{p}_l^{d_i}\}$, so as to maximize the consistency of the original and augmented sequential representations, as follows:

$$\begin{aligned} \mathcal{L}_{\text{SCL}} = \sum_{i=1}^{m} &\text{sim}(\boldsymbol{q}_l^{d_i}, \text{stopgrad}(\widetilde{\boldsymbol{p}}_l^{d_i})) \\ &+ \text{sim}(\widetilde{\boldsymbol{q}}_l^{d_i}, \text{stopgrad}(\boldsymbol{p}_l^{d_i})) \end{aligned} \tag{15}$$

where $\text{sim}(\cdot)$ is the negative cosine similarity, the stop-gradient operation means vector in stopgrad$(\cdot)$ is treated as constant, and it does not generate gradient information.

3.5 Federated Training

The overall federated learning algorithm is shown in Appendix A.3, and its training process consists of the platform update and server aggregation stage.

Platform Update. Given user behavior sequences $\{s^{d_1}, s^{d_2}, ..., s^{d_m}, s^*\}$, the training of CSR is to predict the next item $v_{j+1}^{d_i}$ for domain d based on the interactions before timestamp j.

We define three prediction functions as single-domain (sd), unified-domain (ud) and cross-domain (cd) to calculate the prediction probability of next-item in each domain d_i.

$$\begin{aligned} &\mathrm{P}_{sd}^{d_i}(v_{j+1}^{d_i} | s^{d_i}) = \text{softmax}(\boldsymbol{h}_j^{d_i} \boldsymbol{W}^{d_i}) \\ &\mathrm{P}_{ud}^{d_i}(v_{j+1}^{d_i} | s^*) = \text{softmax}(\boldsymbol{h}_j^* \boldsymbol{W}^{d_i}) \\ &\mathrm{P}_{cd}^{d_i}(v_{j+1}^{d_i} | s^{d_i}, s^*) = \text{softmax}(\boldsymbol{h}_j^{d_i} \boldsymbol{W}^{d_i} + \boldsymbol{h}_j^* \boldsymbol{W}^{d_i}) \end{aligned} \tag{16}$$

where the $\boldsymbol{h}_j^d \in \mathbb{R}^{1\times n}$ and $\boldsymbol{h}_j^* \in \mathbb{R}^{1\times n}$ are the sequential representations at timestep j in the learned $\boldsymbol{H}^{d_i}$ and $\boldsymbol{H}^*$. $\boldsymbol{W}^{d_i} \in \mathbb{R}^{n\times|\mathcal{V}^{d_i}|}$ is the parameter matrix for prediction layer of domain d_i.

We use the negative log-likelihood loss function and summarize the prediction function:

$$\mathcal{L}_{Pre} = -\sum_{i=1}^{m} \log\mathrm{P}_{sd}^{d_i} + \log\mathrm{P}_{ud}^{d_i} + \log\mathrm{P}_{cd}^{d_i} \quad (17)$$

Then, we combine the $\mathcal{L}_{Pre}$, $\mathcal{L}_{MCL}$ and $\mathcal{L}_{SCL}$ to train the local CSR model by a multi-task learning manner. The total loss is defined as:

$$\mathcal{L}_{Total} = \lambda_1 \mathcal{L}_{Pre} + \lambda_2 \mathcal{L}_{MCL} + \lambda_3 \mathcal{L}_{SCL} \quad (18)$$

where $\lambda_1, \lambda_2, \lambda_3$ is the harmonic factor.

Server Aggregation. The global model w^{t+1} is updated with the weighted sum of all received local models. The aggregation process is as follows:

$$w^{t+1} = \sum_{k=1}^{K} \frac{|\mathcal{D}_k|}{|\mathcal{D}|} w_k^{t+1} \quad (19)$$

where $|\mathcal{D}_k|$ is the total number of users for platform k and $|\mathcal{D}|$ is total number users of all platforms.

4 Experimental Setup

Datasets. To simulate federated CSR scenarios, we employ six public datasets from Amazon[2], which are commonly used in cross-domain recommendation (Lu et al., 2023; Cao et al., 2022a). Following (Cao et al., 2022a), we first use ratings with n-core filtering to retain users at least 10 interactions in both domains. Then, we construct three cross-domain datasets, "Food-Kitchen", "Movie-Book" and "Entertainment-Education". Cross-domain sequences have at least 3 items per domain. Each dataset has a training set and a testing set. The training set is non-uniformly partitioned into 10 segments to ensure the heterogeneity across platforms. Each segment is then allocated to a distinct platform, ensuring that the user sets among the platforms are non-overlapping, and moreover, the item popularity distributions and average sequence lengths are also distinct. The testing set, stored on the server, is only accessible for evaluating the performance of the global CSR model. The statistics for the two CSR scenarios are in Table 1.

Domain	#Users	#Items	#Train/Valid/Test	Avg.Len	Sparsity
Food Kitchen	16,579	29,207 34,886	34,117 / 8,173 / 8,406	9.91	99.953%
Movie Book	15,352	36,845 63,937	58,515 / 7,644 / 7,708	11.98	99.943%

Table 1: Statistics of two CSR scenarios.

[2] https://jmcauley.ucsd.edu/data/amazon/

Compared Baselines and Implementation. We compare our proposed method with various baselines from three categories, including sequential recommendation (TiSASRec (Li et al., 2020a), CLSR (Zheng et al., 2022)), cross-domain recommendation (CoNet (Hu et al., 2018), CCDR (Xie et al., 2022)) and cross domain recommendation (π-Net (Ma et al., 2019), PSJNet (Sun et al., 2021), MIFN (Ma et al., 2022), C^{2}DSR (Cao et al., 2022a)). The details of their implementation and how we adapt them to the FL setting are listed in Appendix A.1.

Evaluation Metrics. We adopt MRR, NDCG@k, and HR@k (k=5, 10) to evaluate recommendation performance. Additionally, we report the performance for each domain in the cross-domain setting to assess how well the model generalizes across different domains.

Algorithm Complexity and Communication Efficiency. We conducted an analysis of the time complexity and communication efficiency of FedCSR. The results indicate that FedCSR boasts low time complexity and high communication efficiency, which effectively supports its expansion in aspects such as platforms, domains, and users. Consequently, FedCSR holds great promise for practical applications in actual scenarios. More detailed analysis can be found in Appendix A.4 and A.5.

5 Experimental Results

5.1 Main Comparison Results

We report the recommendation performance on the "Food-Kitchen" and "Movie-Book" cross-domain test set in Table 2. The full version with more cross-domain scenarios is discussed in Appendix B.1. From Table 2, we have the following observations. **(1)** Our proposed method FedCSR generally outperforms all baselines across multiple metrics in the listed two cross-domain scenarios. **(2)** Contrastive learning-based models, such as CLSR and CCDR, outperform TiSASRec in single-domain sequential recommendation and CoNet in cross-domain recommendation, respectively. However, cross-domain recommendation (e.g., CCDR) still lags behind single-domain sequential models (e.g., CLSR) due to the importance of employing sequence characteristics. **(3)** Cross-domain sequential models like MIFN, and C^{2}DSR show strong performance, emphasizing the value of incorporating cross-domain sequential information.

Methods	Food-domain			Kitchen-domain			Movie-domain			Book-domain		
	MRR	NDCG@5	HR@5	MRR	NDCG@5	HR@5	MRR	NDCG@5	HR@5	MRR	NDCG@5	HR@5
FedTiSASRec	4.27	3.60	5.71	2.79	2.07	3.19	4.35	3.39	5.50	2.41	2.00	2.99
FedCLSR	5.88	5.36	6.51	2.82	2.12	3.26	4.46	3.49	4.47	2.26	2.05	2.83
FedCoNet	2.95	2.53	4.24	2.48	1.79	2.59	2.95	2.30	3.96	1.47	0.98	1.27
FedCCDR	3.35	2.97	6.13	2.60	1.99	2.85	3.48	2.76	4.39	1.82	1.36	1.98
Fedπ-Net	6.40	6.12	8.51	2.83	2.14	3.35	4.59	3.53	5.53	1.99	1.52	2.11
FedPSJNet	6.88	6.75	8.75	3.45	2.87	4.41	4.83	3.72	5.88	2.57	1.93	2.80
FedMIFN	7.04	6.24	8.53	4.00	3.07	4.70	5.07	3.95	6.12	**2.80**	2.20	3.05
FedC^{2}DSR	6.98	6.80	9.18	3.63	3.14	4.64	4.08	3.28	4.76	1.86	1.50	2.11
FedCSR	**9.15**	**9.07**	**12.21**	**4.93**	**4.44**	**6.35**	**6.73**	**5.84**	**8.81**	2.74	**2.32**	**3.10**

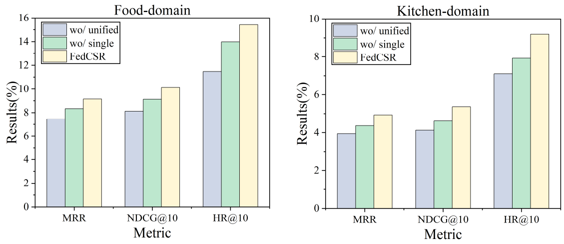
Table 2: Federated experimental results (in %) for the "Food-Kitchen" and "Movie-Book" scenarios. The **best** (second best) results are in bold (underlined). The reported scores are the average of 5 runs.

Ablation of MLC and SCL. We conduct an ablation study on the Food-Kitchen and Movie-Book datasets to assess the impact of the proposed contrastive learning methods, MCL and SCL. We report the results in Table 3 and 4. We can find that: **(1)** Both SCL and MCL improve recommendation performance, which may due to their ability to enhance representation learning. **(2)** The performance gains from SCL and MCL are consistent across different datasets and domains, demonstrating the robustness of the proposed mechanisms.

Ablation of Prediction Targets. As introduced in Section 3.5, we utilize three prediction targets: single-domain, unified-domain, and cross-domain. To further investigate their effectiveness, we conducted an ablation study comparing the performance of FedCSR (the full model with all three prediction targets) against its variants without unified-domain and single-domain predictions, abbreviated as "w/o unified" and "w/o single", respectively. The results, shown in Figure 2, reveal that: (1) Removing either unified- or single-domain predictions results in a decline in recommendation performance, demonstrating the positive contribution of both. (2) The model without unified-domain prediction shows a more significant performance drop across all metrics compared to the model without single-domain prediction, indicating that unified-domain prediction provides more valuable information and underscoring the importance of developing cross-domain algorithms.

5.2 Further Analysis

Convergence Analysis. Figure 3 shows the trend of MRR scores over communication rounds in the Food-Kitchen scenario. We observe that adding MCL and SCL (i.e., the FedCSR method) does



(a) Food Domain (b) Kitchen Domain

Figure 2: The impact of single and unified domain predictions of FedCSR.

Component		Food-domain			Kitchen-domain		
SCL	MCL	MRR	HR@10	NDCG@10	MRR	HR@10	NDCG@10
✗	✗	6.48	10.90	7.02	3.57	6.35	3.69
✓	✗	8.48	13.53	9.09	4.77	8.45	5.06
✓	✓	**9.15**	**15.42**	**10.11**	**4.93**	**9.19**	**5.36**

Table 3: Ablation study on Food-Kitchen scenario.

not increase the number of communication rounds required compared to the method without these mechanisms. FedCSR converges within approximately 25 rounds. Additionally, the inclusion of MCL and SCL consistently improves performance as the number of communication rounds increases, demonstrating the superiority of the FedCSR method.

FedCSR vs Individual Training. We further investigate whether FedCSR achieves better recommendation performance compared to individual training. We compare the performance of FedCSR and individual training through experiments in the Food-Kitchen scenario. In the individual training approach, platforms independently train CSR models, denoted as "Individual", while in FedCSR, all platforms collaboratively train a centralized CSR model. The results, summarized in Figure 4, show that FedCSR outperforms all individual platforms,

Component		Movie-domain			Book-domain		
SCL	MCL	MRR	HR@10	NDCG@10	MRR	HR@10	NDCG@10
✗	✗	3.94	6.02	3.41	1.82	2.81	1.79
✓	✗	5.31	11.02	5.21	2.33	3.89	2.42
✓	✓	**6.73**	**12.71**	**7.12**	**2.74**	**4.32**	**2.72**

Table 4: Ablation study on Movie-Book scenario.

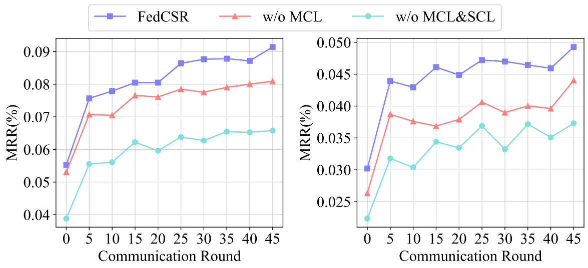


Figure 3: MRR (Left: Food, Right: Kitchen) over communication rounds in the Food-Kitchen scenario.

demonstrating the superiority of collaborative training and the effectiveness of the proposed dual contrastive learning mechanisms. A more detailed discussion can be found in Appendix B.2.

Hyper-Parameter Discussion. The hyperparameter α, introduced in Section 3.4, controls the degree of data augmentation applied to input sequences, balancing augmented data with original information for recommendation tasks. To evaluate its impact, we conducted a hyperparameter experiment, with results shown in Table 5. As α increases from 0.1 to 0.3 in the Food and Kitchen domains, performance improves, but further increases to 0.5 and 0.7 result in declines, which may due to excessive augmentation introducing noise. The best performance is achieved with α set to 0.3.

Case Study. We present a selected example from the test set to demonstrate that our method provides more accurate item recommendations than the baselines. Figure 5 shows the ranked prediction lists, from highest to lowest probability, for FedCLSR, FedC^{2}DSR, and FedCSR. Our method, FedCSR, successfully ranks the ground truth item at the top. This may be because it effectively captures the "Adventure" genre from the historical sequence (row (a)) across both domains (Movie and Book). Additionally, by observing the top 6 items predicted by FedCSR, we notice that "Adventure" is the most common genre. This indicates that FedCSR effectively captures genre information from the "Movie" domain, without being restricted to the "Book" domain, where the historical sequence primarily reflects the "History" genre.

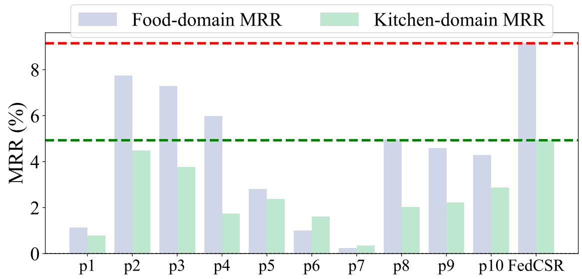


Figure 4: Comparison between FedCSR and individual training in the Food-Kitchen scenario. p# refers to the individual training on platform #.

α	Food-domain			Kitchen-domain		
	MRR	HR@10	NDCG@10	MRR	HR@10	NDCG@10
0.1	8.82	14.85	9.71	4.76	8.51	5.04
0.3	**9.15**	**15.42**	**10.11**	**4.93**	**9.19**	**5.36**
0.5	8.68	15.17	9.67	4.77	8.82	5.12
0.7	8.56	14.45	9.32	4.69	8.61	5.01

Table 5: Performance (%) of FedCSR in Food-Kitchen scenario over hyper-parameter α.

6 Conclusion

We introduce FedCSR, a novel federated multi-platform cross-domain sequential recommendation framework that enhances recommendation accuracy while addressing data heterogeneity and privacy. By bypassing the need for user alignment, FedCSR reduces communication overhead and preserves privacy. Our dual contrastive learning mechanisms, MCL and SCL, effectively tackle inter-platform and intra-platform data pattern heterogeneity. Experiments on real-world datasets show that FedCSR consistently outperforms existing methods, proving its effectiveness and scalability in collaborative recommendation systems.

Limitations

Despite the promising results of FedCSR, there are two limitations to consider. (1) While FedCSR demonstrates strong performance on the evaluated datasets, scaling the framework to accommodate very large datasets with numerous platforms and domains may introduce computational challenges that require further optimization. (2) Our current approach relies on overlapping users within a platform to facilitate inter-domain information transfer. However, in cases where user overlap is sparse, the effectiveness of this transfer may be limited. Future work will focus on addressing this issue by exploring alternative methods for inter-domain knowledge sharing.

Figure 5: Case study. The red box marks the ground truth, corresponding to the book *A Wizard Abroad* (Fantasy, Adventure). ■ represents the domain, and ● represents the style for each item.

Ethical Considerations

In this work, we prioritize data privacy and responsible research practices. All datasets used are open-sourced and anonymized, ensuring a low risk of leaking personally identifiable information. We carefully selected these datasets to comply with ethical standards for data privacy and protection.

Furthermore, the design and application of our FedCSR framework are guided by privacy-preserving principles. By leveraging federated learning, we avoid sharing raw user data across platforms, which helps mitigate the risks of data leakage and unauthorized access. Federated learning facilitates collaborative model training while keeping sensitive data decentralized and secure, aligning with global privacy regulations such as GDPR.

Acknowledgements

We gratefully acknowledge the support and sponsorship that have been crucial in enabling this research. This work is supported by the Major Key Project of PCL (Grant No. PCL2024A05). This work is supported by Guangdong Major Project of Basic and Applied Basic Research (No. 2019B03030200245). This work is supported by Shenzhen Science and Technology Program (Grant No.ZDSYS20210623091809029).

References

Jianping Cai, Yang Liu, Ximeng Liu, Jiayin Li, and Hongbin Zhuang. 2022. Privacy-preserving federated cross-domain social recommendation. *arXiv preprint arXiv:2206.03200*.

Jiangxia Cao, Xin Cong, Jiawei Sheng, Tingwen Liu, and Bin Wang. 2022a. Contrastive cross-domain sequential recommendation. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, pages 138–147.

Jiangxia Cao, Shaoshuai Li, Bowen Yu, Xiaobo Guo, Tingwen Liu, and Bin Wang. 2023. Towards universal cross-domain recommendation. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, pages 78–86.

Jiangxia Cao, Xixun Lin, Xin Cong, Jing Ya, Tingwen Liu, and Bin Wang. 2022b. Disencdr: Learning disentangled representations for cross-domain recommendation. In *Proceedings of the 45th International ACM SIGIR conference on research and development in information retrieval*, pages 267–277.

Gaode Chen, Xinghua Zhang, Yijun Su, Yantong Lai, Ji Xiang, Junbo Zhang, and Yu Zheng. 2023. Win-win: A privacy-preserving federated framework for dual-target cross-domain recommendation. In *Thirty-Seventh AAAI Conference on Artificial Intelligence, AAAI 2023*, pages 4149–4156. AAAI Press.

Xinlei Chen and Kaiming He. 2021. Exploring simple siamese representation learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 15750–15758.

Yue Cong, Jing Qiu, Kun Zhang, Zhongyang Fang, Chengliang Gao, Shen Su, and Zhihong Tian. 2023. Ada-ffl: Adaptive computing fairness federated learning. *CAAI Transactions on Intelligence Technology*, 9(3):573–584.

Lei Guo, Ziang Lu, Junliang Yu, Quoc Viet Hung Nguyen, and Hongzhi Yin. 2024a. Prompt-enhanced federated content representation learning for cross-domain recommendation. In *Proceedings of the ACM Web Conference 2024*, WWW '24, page 3139–3149, New York, NY, USA. Association for Computing Machinery.

Lei Guo, Li Tang, Tong Chen, Lei Zhu, Quoc Viet Hung Nguyen, and Hongzhi Yin. 2021. Da-gcn: A domain-aware attentive graph convolution network for shared-account cross-domain sequential recommendation. In *International Joint Conference on Artificial Intelligence*.

Lei Guo, Jinyu Zhang, Tong Chen, Xinhua Wang, and Hongzhi Yin. 2022. Reinforcement learning-enhanced shared-account cross-domain sequential recommendation. *IEEE Transactions on Knowledge and Data Engineering*.

Shengnan Guo, Xibin Wang, Shigong Long, Hai Liu, Liu Hai, and Toong Hai Sam. 2023. A federated learning scheme meets dynamic differential privacy. *CAAI Transactions on Intelligence Technology*, 8(3):1087–1100.

Wei Guo, Fuzhen Zhuang, Xiao Zhang, Yiqi Tong, and Jin Dong. 2024b. A comprehensive survey of federated transfer learning: Challenges, methods and applications. *Frontiers of Computer Science*, 18(6):186356.

Guangneng Hu, Yu Zhang, and Qiang Yang. 2018. Conet: Collaborative cross networks for cross-domain recommendation. In *Proceedings of the 27th ACM international conference on information and knowledge management*, pages 667–676.

Jiacheng Li, Yujie Wang, and Julian McAuley. 2020a. Time interval aware self-attention for sequential recommendation. In *Proceedings of the 13th international conference on web search and data mining*, pages 322–330.

Mingming Li, Songlin Hu, Fuqing Zhu, and Qiannan Zhu. 2024. Few-shot learning for cold-start recommendation. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation, LREC/COLING 2024, 20-25 May, 2024, Torino, Italy*, pages 7185–7195. ELRA and ICCL.

Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. 2020b. Federated optimization in heterogeneous networks. *Proceedings of Machine learning and systems*, 2:429–450.

Xinting Liao, Weiming Liu, Xiaolin Zheng, Binhui Yao, and Chaochao Chen. 2023. Ppgencdr: A stable and robust framework for privacy-preserving cross-domain recommendation. In *Thirty-Seventh AAAI Conference on Artificial Intelligence, AAAI 2023*, pages 4453–4461. AAAI Press.

Guanyu Lin, Chen Gao, Yu Zheng, Jianxin Chang, Yanan Niu, Yang Song, Kun Gai, Zhiheng Li, Depeng Jin, Yong Li, and Meng Wang. 2024. Mixed attention network for cross-domain sequential recommendation. WSDM '24, page 405–413, New York, NY, USA. Association for Computing Machinery.

Fengxia Liu, Zhiming Zheng, Yexuan Shi, Yongxin Tong, and Yi Zhang. 2024. A survey on federated learning: a perspective from multi-party computation. *Frontiers of Computer Science*, 18(1):181336.

Weiming Liu, Chaochao Chen, Xinting Liao, Mengling Hu, Jianwei Yin, Yanchao Tan, and Longfei Zheng. 2023a. Federated probabilistic preference distribution modelling with compactness co-clustering for privacy-preserving multi-domain recommendation. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI 2023*, pages 2206–2214. ijcai.org.

Weiming Liu, Xiaolin Zheng, Chaochao Chen, Mengling Hu, Xinting Liao, Fan Wang, Yanchao Tan, Dan Meng, and Jun Wang. 2023b. Differentially private sparse mapping for privacy-preserving cross domain recommendation. In *Proceedings of the 31st ACM International Conference on Multimedia*, MM '23, page 6243–6252, New York, NY, USA. Association for Computing Machinery.

Weiming Liu, Xiaolin Zheng, Chaochao Chen, Jiajie Su, Xinting Liao, Mengling Hu, and Yanchao Tan. 2023c. Joint internal multi-interest exploration and external domain alignment for cross domain sequential recommendation. WWW '23, page 383–394, New York, NY, USA. Association for Computing Machinery.

Jinhu Lu, Guohao Sun, Xiu Fang, Jian Yang, and Wei He. 2023. A contrastive learning framework for dual-target cross-domain recommendation. In *Proceedings of the 31st ACM International Conference on Multimedia*, MM '23, page 6332–6339, New York, NY, USA. Association for Computing Machinery.

Hongqin Lyu, Yongxiong Zhang, Chao Wang, Shigong Long, and Shengnan Guo. 2023. Federated learning privacy incentives: Reverse auctions and negotiations. *CAAI Transactions on Intelligence Technology*, 8(4):1538–1557.

Haokai Ma, Ruobing Xie, Lei Meng, Xin Chen, Xu Zhang, Leyu Lin, and Jie Zhou. 2024. Triple sequence learning for cross-domain recommendation. *ACM Trans. Inf. Syst.*, 42(4).

Muyang Ma, Pengjie Ren, Zhumin Chen, Zhaochun Ren, Lifan Zhao, Peiyu Liu, Jun Ma, and Maarten de Rijke. 2022. Mixed information flow for cross-domain sequential recommendations. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 16(4):1–32.

Muyang Ma, Pengjie Ren, Yujie Lin, Zhumin Chen, Jun Ma, and Maarten de Rijke. 2019. π-net: A parallel information-sharing network for shared-account cross-domain sequential recommendations. In *Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval*, pages 685–694.

Tong Man, Huawei Shen, Xiaolong Jin, and Xueqi Cheng. 2017. Cross-domain recommendation: An embedding and mapping approach. In *IJCAI*, volume 17, pages 2464–2470.

Kihyuk Sohn. 2016. Improved deep metric learning with multi-class n-pair loss objective. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*, NIPS'16, page 1857–1865, Red Hook, NY, USA. Curran Associates Inc.

Xiaotong Song, Huiping Lin, Jiatao Zhu, and Xinyi Gong. 2024. CAGK: collaborative aspect graph enhanced knowledge-based recommendation. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources* *and Evaluation, LREC/COLING 2024, 20-25 May, 2024, Torino, Italy*, pages 2612–2621. ELRA and ICCL.

Wenchao Sun, Muyang Ma, Pengjie Ren, Yujie Lin, Zhumin Chen, Zhaochun Ren, Jun Ma, and Maarten De Rijke. 2021. Parallel split-join networks for shared account cross-domain sequential recommendations. *IEEE Transactions on Knowledge and Data Engineering*.

Li-e Wang, Yihui Wang, Yan Bai, Peng Liu, and Xianxian Li. 2021. Poi recommendation with federated learning and privacy preserving in cross domain recommendation. In *IEEE INFOCOM 2021-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 1–6. IEEE.

Ruobing Xie, Qi Liu, Liangdong Wang, Shukai Liu, Bo Zhang, and Leyu Lin. 2022. Contrastive cross-domain recommendation in matching. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 4226–4236.

Wujiang Xu, Qitian Wu, Runzhong Wang, Mingming Ha, Qiongxu Ma, Linxun Chen, Bing Han, and Junchi Yan. 2024. Rethinking cross-domain sequential recommendation under open-world assumptions. In *Proceedings of the ACM Web Conference 2024*, WWW '24, page 3173–3184, New York, NY, USA. Association for Computing Machinery.

Sixing Yu, Juan Pablo Muñoz, and Ali Jannesari. 2024. Federated foundation models: Privacy-preserving and collaborative learning for large models. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation, LREC/COLING 2024, 20-25 May, 2024, Torino, Italy*, pages 7174–7184. ELRA and ICCL.

Jianyang Zhai, Xiawu Zheng, Chang-Dong Wang, Hui Li, and Yonghong Tian. 2023. Knowledge prompt-tuning for sequential recommendation. In *Proceedings of the 31st ACM International Conference on Multimedia*, pages 6451–6461.

Hongyu Zhang, Dongyi Zheng, Xu Yang, Jiyuan Feng, and Qing Liao. 2024a. Feddcsr: Federated cross-domain sequential recommendation via disentangled representation learning. In *Proceedings of the 2024 SIAM International Conference on Data Mining (SDM)*, pages 535–543. SIAM.

Hongyu Zhang, Dongyi Zheng, Lin Zhong, Xu Yang, Jiyuan Feng, Yunqing Feng, and Qing Liao. 2024b. Fedhcdr: Federated cross-domain recommendation with hypergraph signal decoupling. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 350–366. Springer.

Jinyu Zhang, Huichuan Duan, Lei Guo, Liancheng Xu, and Xinhua Wang. 2023a. Towards lightweight cross-domain sequential recommendation via external attention-enhanced graph convolution network. *arXiv preprint arXiv:2302.03221*.

Yong Zhang, Ying Hu, Xiaozhi Gao, Dunwei Gong, Yinan Guo, Kaizhou Gao, and Wanqiu Zhang. 2023b. An embedded vertical-federated feature selection algorithm based on particle swarm optimisation. *CAAI Transactions on Intelligence Technology*, 8(3):734–754.

Zhuo Zhang, Jintao Huang, Xiangjing Hu, Jingyuan Zhang, Yating Zhang, Hui Wang, Yue Yu, Qifan Wang, Lizhen Qu, and Zenglin Xu. 2024c. Revisiting data reconstruction attacks on real-world dataset for federated natural language understanding. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation, LREC/COLING 2024, 20-25 May, 2024, Torino, Italy*, pages 14080–14091. ELRA and ICCL.

Dongyi Zheng, Yeting Guo, Fang Liu, Nong Xiao, and Lu Gao. 2023. Metaem: Meta embedding mapping for federated cross-domain recommendation to cold-start users. In *Collaborative Computing: Networking, Applications and Worksharing: 18th EAI International Conference, CollaborateCom 2022, Hangzhou, China*, pages 154–172. Springer.

Yu Zheng, Chen Gao, Jianxin Chang, Yanan Niu, Yang Song, Depeng Jin, and Yong Li. 2022. Disentangling long and short-term interests for recommendation. In *Proceedings of the ACM Web Conference 2022*, pages 2256–2267.

Feng Zhu, Chaochao Chen, Yan Wang, Guanfeng Liu, and Xiaolin Zheng. 2019. Dtcdr: A framework for dual-target cross-domain recommendation. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pages 1533–1542.

Yongchun Zhu, Kaikai Ge, Fuzhen Zhuang, Ruobing Xie, Dongbo Xi, Xu Zhang, Leyu Lin, and Qing He. 2021. Transfer-meta framework for cross-domain recommendation to cold-start users. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1813–1817.

Yongchun Zhu, Zhenwei Tang, Yudan Liu, Fuzhen Zhuang, Ruobing Xie, Xu Zhang, Leyu Lin, and Qing He. 2022. Personalized transfer of user preferences for cross-domain recommendation. In *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, pages 1507–1515.

A Detailed Implementation

A.1 Compared Baselines

In this subsection, we introduce three categories of baselines and their critical operations for adaptation. To be fair, we will compare the performance of these baselines under federated learning environment (Lyu et al., 2023; Guo et al., 2023; Zhang et al., 2023b). Specifically, we use the widely used FedProx(Li et al., 2020b) to make baselines better adapt to the non-i.i.d. data distributions.

Sequential recommendation baselines:

- TiSASRec(Li et al., 2020a) combines the advantages of absolute position and relative time interval coding for self-attention.
- CLSR(Zheng et al., 2022) is proposed to disentangle long and short interests with a contrastive learning method.

Cross-Domain Recommendation baselines:

- CoNet(Hu et al., 2018) first models behaviors of two domains and then transfers information by a cross-network.
- CCDR(Xie et al., 2022) designs the intra-domain and inter-domain contrastive learning tasks for representation learning.

Cross-Domain Sequential Recommendation baselines:

- π-Net(Ma et al., 2019) devises a novel gating recurrent module to model and transfer knowledge across different domains.
- PSJNet(Sun et al., 2021) introduces a parallel split-join scheme to transfer the different user intentions across domains.
- MIFN(Ma et al., 2022) proposes constructing a knowledge graph to guide the connection between items from other domains to transfer information across domains.
- C^{2}DSR(Cao et al., 2022a) constructs a contrastive learning loss to supervise the learning of cross-domain representations.

A.2 Our Method's Implementation and Setting

We set the same federated learning parameters for FedCSR and all baseline methods for fair comparisons. Specifically, we set 10 platforms, 100 training rounds, 3 local epochs in each round, early stopping with the patience of 20.

For the CSR model, we set the following model parameters: embedding size and mimi-batch size are set to 256, the dropout rate is selected from to {0.1, 0.2, 0.3}, the learning rate is selected from {0.001, 0.0005, 0.0001}, the harmonic factors $\lambda_1, \lambda_2, \lambda_3$ are fixed at {0.5, 1.5, 0.5}. The data augmentation ratio α is selected from {0.1, 0.3, 0.5, 0.7}, and Adam optimizer is used to update the model parameters. We set the temperature parameter τ to 0.5. For other baselines, we adopt the FedProx as the base federated algorithm and configure all other hyperparameters based on the suggestions provided in their original papers. We employ three evaluation metrics to assess the performance of the recommendation methods: Mean Reciprocal Rank (MRR), Normalized Discounted Cumulative Gain (NDCG), and Hit Rate (HR).

A.3 The Algorithm of FedCSR

In the FedCSR setting, the CSR model is denoted as w. Platform k updates its local model w_k using the local private dataset $\mathcal{D}_k$. During local training, we employ sequence augmentation and two contrastive learning objectives to improve the representation quality of users in every domain. The global CSR model of the central server is represented as w, which is obtained by aggregating all the individual w_k. By iteratively repeating local training and aggreagation, the global model gradually improves and captures insights from diverse platforms datasets while preserving the privacy of each platform's raw data.

Due to privacy concerns, only model parameters can be transmitted between the server and platforms. The user's sequence data is stored locally on the platform, and the server cannot access each platform's dataset. Moreover, there is no data sharing or exchange among platforms. Hence, our goal is to utilize the datasets from all platforms to train a more robust CSR model for each platform compared to models trained solely on their individual datasets.

A.4 Algorithm complexity

The time complexity analysis is as follows. Initially, for a given sequence of data, the primary contributor to the time complexity of its feature encoding is the self-attention mechanism, resulting in a complexity of $\mathcal{O}(n \times l^2)$, where we assume the length of each sequence is l and n represents the feature dimension. Next, in the SCL of FedCSR, each user possesses $m + 1$ sequences

Algorithm 1: FedCSR

Input: The platform k has local private dataset $\mathcal{D}_k$; $|\mathcal{D}|$ is the total samples of all platforms; B are the local minibatch sets of $\mathcal{D}_k$, E is the number of local epochs, and η is the learning rate.

Output: The optimal model parameter w.

Server executes:
initialize w_0;
for *each round t = 1,2,...* **do**
    $K \leftarrow$ (set of platforms)
    for *platform $k \in K$ in parallel do* **do**
        w_k^{t+1} = Platform_Update(k, w^t)
    end
    $w^{t+1} = \sum_{k=1}^{K} \frac{|\mathcal{D}_k|}{|\mathcal{D}|} w_k^{t+1}$ ▷ Eq.(19)
end

Platform_Update(k, w**):** // Run on platform k
save w as w_g
$B \leftarrow$ (split $\mathcal{D}_k$ into batches of set)
for *each local epoch from 1 to E* **do**
    `// sequence augmentation`
    draw one augmentation function $t \sim \mathcal{T}$
    for *all batch $b \in B$* **do**
        $\tilde{b} = t(b)$ ▷ Eq.11
    end
    $\widetilde{B}$ = set of all $\tilde{b}$
    `// local train`
    for *batch $b, \tilde{b} \in B, \widetilde{B}$* **do**
        $w \leftarrow w - \eta \nabla_w \mathcal{L}_{Total}(w, w_g, b, \tilde{b})$ ▷ Eq.(18)
    end
end
return w to server

$\{s^{d_1}, s^{d_2}, ..., s^{d_m}, s^*\}$ and $m+1$ corresponding augmented sequences $\{\widetilde{s}^{d_i}, \widetilde{s}^{d_2}, ..., \widetilde{s}^{d_m}, \widetilde{s}^*\}$. Encoding the features for such a user incurs a time complexity of $\mathcal{O}(n \times l^2 \times 2 \times (lm + lm)) = \mathcal{O}(4lm)$, where both the length of s^* and $\widetilde{s}^*$ is lm. In MCL of FedCSR, the global model computing time complexity is $\mathcal{O}(n \times l^2 \times 1 \times (lm + lm)) = \mathcal{O}(n \times l^2 \times 2lm)$. Therefore, the total time complexity of local training is $\mathcal{O}(n \times l^2 \times 4lm) + \mathcal{O}(n \times l^2 \times 2lm) = \mathcal{O}(n \times 6ml^3)$. Subsequently, considering a platform k that comprises $|\mathcal{D}_k|$ users, the time complexity for that platform is $\mathcal{O}(|\mathcal{D}_k| \times n \times l^2 \times 6lm) = \mathcal{O}(|\mathcal{D}_k| \times n \times 6ml^3)$. During federated training, owing to parallel processing among platforms, the total time complexity is $\mathcal{O}(T \times K \times E \times |\mathcal{D}_k| \times n \times 6ml^3)$, where T and E represent the number of training iterations and epochs, respectively. Notably, the number of platforms K can be omitted due to parallelization. Finally, in our work, T, E, and n remain fixed as small hyperparameters throughout the training process. Additionally, l is both bounded by the maximum sequence length within all domains, which is also a fixed value. Consequently, the total time complexity of federated training scales linearly with the number of users per platform $(i.e., |\mathcal{D}_k|)$ and the number of domains $i.e., m$.

A.5 Communication efficiency

In the FedCSR framework, the communication load between the server and platforms plays a significant factor. This load primarily stems from the transmission requirements of the global item embedding matrices from m different domains. In a 64-bit floating-point computing environment, the data volume for a single communication can be calculated as $(\sum_{i=1}^{m} |\mathcal{V}^{d_i}| \times n \times 64)/8$, where the $\sum_{i=1}^{m} |\mathcal{V}^{d_i}|$ are the total items for all domains. For example, even when the total item numbers of all domains reaches 100K, and the feature dimension n is set to 32, the data volume required for a single communication is only 24.41MB. Considering that our work focuses on recommendation platforms with good data transmission capabilities, and the entire training process requires only approximately 25 rounds of communication to achieve convergence, such communication load does not pose a significant communication pressure on the entire system.

B Further Experimental Results

B.1 Full Main Results

Tables 6, 7, and 8 show the performance of the compared methods in the CSR scenarios "Food-Kitchen", "Movie-Book" and "Entertainment-Education". The best (second best) results are in bold (underlined).

From the experimental results, several key observations can be made. Firstly, among the single-domain recommendation baselines, CLSR outperforms TiSASRec, validating the benefit of contrastive learning for better sequential representation learning. Secondly, for the cross-domain baselines, CCDR performs better than CoNet, indicating the promising advantages of contrastive learning in cross-domain recommendation. However, cross-domain recommendation lags behind single-domain sequential recommendation, highlighting the need for fully mining user behavior sequence information. Thirdly, cross-domain sequential baselines such as π-Net, PSJNet, MIFN, and C^{2}DSR surpass single-domain sequential and cross-domain recommendation baselines, demonstrating the benefit of cross-domain sequential information for enhancing recommendation performance. Finally, compared with all baselines, FedCSR achieves the best performance in many indicators in three cross-domain scenarios.

B.2 Further Analysis

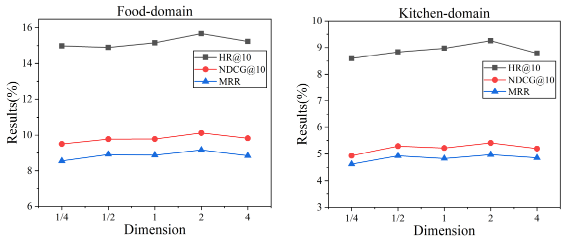
Full Results of Comparison Between FedCSR and Individual Training. We have three critical observations based on the experimental results: (1) FedCSR surpassed individual training, as evident in Table 9. It consistently outperformed models trained individually by each data owner. This highlights FedCSR's ability to leverage collective platform knowledge for enhanced overall model performance. (2) Data sparsity significantly impacts model accuracy. As previously set, platform 2 has the most significant proportion of training samples, achieved the highest scores among individual models in both domains. platform 7 has the least training data and the worst recommendation performance.

platform	Food-domain			Kitchen-domain		
	MRR	HR@10	NDCG@10	MRR	HR@10	NDCG@10
1	1.13	2.17	0.72	0.79	2.25	0.89
2	7.74	13.07	8.39	4.47	7.76	4.65
3	7.28	11.33	7.52	3.76	6.73	3.85
4	5.99	9.96	6.39	1.73	4.12	1.80
5	2.81	5.79	2.99	2.37	3.17	1.96
6	0.99	1.85	0.79	1.61	1.23	1.24
7	0.24	0.42	0.19	0.35	0.26	0.21
8	4.95	7.69	5.11	2.03	3.75	1.92
9	4.58	8.77	5.15	2.21	3.51	2.03
10	4.28	8.07	4.53	2.88	6.17	3.09
FedCSR	**9.15**	**15.42**	**10.11**	**4.93**	**9.19**	**5.36**

Table 9: The results of FedCSR, individual and centralized training on Food-Kitchen dataset.

Hidden Layer Dimension Analysis of the transformation in SCL. To determine the influence of the hidden layer dimension of the transformation $t(\cdot)$ in SCL on recommendation performance, we conduct an experiment in the food-kitchen scenario. In this experiment, we sequentially set the hidden layer dimensions to 1/4, 1/2, 1, 2, and 4 times the dimension of the output layer. The experimental results are shown in Fig 6.

Our experimental results show that setting the hidden layer dimension to twice the output layer leads to a slight improvement in all metrics in both domains. This is different from SimSaim (Chen and He, 2021), which suggests a bottleneck structure. We found that a larger hidden layer dimension helps to learn more complex representations, leading to stable improvement in our method. It's important to consider other factors when determining the optimal hidden layer dimension, such as dataset characteristics and model architecture.



(a) Food performance (b) Kitchen performance

Figure 6: The results (in %) of different hidden layer dimensions on the Food-Kitchen scenario.

Experimental Environment. All methods are replicated based on the PyTorch framework (1.8.1+cu11) and experimented on Tesla V100 and Intel(R) Xeon(R) Platinum 8268 CPU @ 2.90GHz.

Methods	Food-domain recommendation						Kitchen-domain recommendation					
	MRR	NDCG		HR			MRR	NDCG		HR		
		@5	@10	@1	@5	@10		@5	@10	@1	@5	@10
FedTiSASRec	4.27	3.60	4.60	1.62	5.71	8.79	2.79	2.07	2.85	0.99	3.19	5.64
FedCLSR	5.88	5.36	6.49	2.95	6.51	10.52	2.82	2.12	2.86	1.10	3.26	5.71
FedCoNet	2.95	2.53	3.18	0.88	4.24	6.28	2.48	1.79	2.48	1.01	2.59	4.76
FedCCDR	3.35	2.97	3.96	1.20	6.13	7.75	2.60	1.99	2.53	1.06	2.85	5.12
Fedπ-Net	6.4	6.12	7.14	3.52	8.51	11.66	2.83	2.14	2.89	0.97	3.35	5.64
FedPSJNet	6.88	6.75	7.29	3.63	8.75	11.83	3.45	2.87	3.33	1.45	4.41	6.25
FedMIFN	7.04	6.24	7.00	3.75	8.53	10.91	4.00	3.07	3.73	1.38	4.70	6.75
FedC^{2}DSR	6.98	6.80	7.47	4.27	9.18	11.31	3.63	3.14	3.82	1.62	4.64	6.72
FedCSR	**9.15**	**9.07**	**10.11**	**5.65**	**12.21**	**15.42**	**4.93**	**4.44**	**5.36**	**2.41**	**6.35**	**9.19**

Table 6: Federated experimental results (in %) on the "Food-Kitchen" scenario. The **best** (second best) results are in bold (underlined). The reported scores are the average of 5 runs.

Methods	Movie-domain recommendation						Book-domain recommendation					
	MRR	NDCG		HR			MRR	NDCG		HR		
		@5	@10	@1	@5	@10		@5	@10	@1	@5	@10
FedTiSASRec	4.35	3.39	4.51	1.40	5.50	9.00	2.41	2.00	2.48	1.10	2.99	**4.50**
FedCLSR	4.46	3.49	4.76	1.49	4.47	8.77	2.46	2.05	2.51	1.16	2.83	4.32
FedCoNet	2.95	2.30	2.99	0.57	3.96	6.09	1.47	0.98	1.21	0.71	1.27	2.01
FedCCDR	3.48	2.76	3.67	0.98	4.39	7.41	1.82	1.36	1.73	0.86	1.98	2.90
Fedπ-Net	4.59	3.53	5.00	1.55	5.53	10.03	1.99	1.52	1.94	0.93	2.11	3.39
FedPSJNet	4.83	3.72	5.14	1.66	5.88	10.17	2.57	1.93	2.36	1.19	2.80	3.98
FedMIFN	5.07	3.95	5.28	1.82	6.12	10.23	**2.80**	2.20	2.63	1.33	3.05	4.29
FedC^{2}DSR	4.08	3.28	3.99	1.82	4.76	6.97	1.86	1.50	1.75	0.90	2.11	2.91
FedCSR	**6.73**	**5.84**	**7.12**	**3.54**	**8.81**	**12.71**	2.74	**2.32**	**2.72**	**1.62**	**3.10**	4.32

Table 7: Federated experimental results (in %) on the "Movie-Book" scenario. The **best** (second best) results are in bold (underlined). The reported scores are the average of 5 runs.

Methods	Entertainment-domain recommendation						Education-domain recommendation					
	MRR	NDCG		HR			MRR	NDCG		HR		
		@5	@10	@1	@5	@10		@5	@10	@1	@5	@10
FedTiSASRec	29.16	30.06	34.42	16.81	42.60	56.02	38.93	40.28	44.35	26.53	52.67	65.21
FedCLSR	32.49	35.99	38.40	23.22	47.51	59.65	40.44	42.79	49.17	29.68	54.36	68.45
FedCoNet	18.11	17.70	22.83	6.73	28.28	44.10	25.38	26.63	31.42	11.56	40.62	55.03
FedCCDR	23.42	24.30	27.56	10.14	34.19	48.03	30.38	33.90	37.56	18.27	45.15	59.36
Fedπ-Net	40.92	42.98	46.61	27.14	57.29	68.49	46.43	48.20	51.35	34.71	60.12	69.78
FedPSJNet	41.78	43.96	48.08	27.99	57.43	68.76	47.97	50.04	52.16	35.79	60.72	72.46
FedMIFN	43.89	46.10	49.84	29.92	60.62	72.15	49.04	51.08	54.00	36.45	63.76	72.76
FedC^{2}DSR	43.88	45.70	49.34	30.93	58.98	70.25	52.28	54.79	57.68	39.95	68.00	76.87
FedCSR	**47.34**	**49.21**	**52.85**	**34.39**	**62.51**	**73.63**	**55.65**	**58.05**	**60.54**	**44.03**	**70.24**	**77.89**

Table 8: Federated experimental results (in %) on the "Entertainment-Education" scenario. The **best** (second best) results are in bold (underlined). The reported scores are the average of 5 runs.