



# Triple Sequence Learning for Cross-domain Recommendation

HAOKAI MA<sup>†,•</sup>, School of Software, Shandong University, China

RUOBING XIE<sup>†</sup>, TEG, Tencent, China

LEI MENG<sup>\*</sup>, School of Software, Shandong University; Shandong Research Institute of Industrial Technology, China

XIN CHEN, WeChat, Tencent, China

XU ZHANG, WeChat, Tencent, China

LEYU LIN, WeChat, Tencent, China

JIE ZHOU, WeChat, Tencent, China

Cross-domain recommendation (CDR) aims to leverage the correlation of users' behaviors in both the source and target domains to improve the user preference modeling in the target domain. Conventional CDR methods typically explore the dual-relations between the source and target domains' behaviors. However, this may ignore the informative mixed behaviors that naturally reflect the user's global preference. To address this issue, we present a novel framework, termed triple sequence learning for cross-domain recommendation (Tri-CDR), which jointly models the source, target, and mixed behavior sequences to highlight the global and target preference and precisely model the triple correlation in CDR. Specifically, Tri-CDR independently models the hidden representations for the triple behavior sequences and proposes a triple cross-domain attention (TCA) method to emphasize the informative knowledge related to both user's global and target-domain preference. To comprehensively explore the cross-domain correlations, we design a triple contrastive learning (TCL) strategy that simultaneously considers the coarse-grained similarities and fine-grained distinctions among the triple sequences, ensuring the alignment while preserving information diversity in multi-domain. We conduct extensive experiments and analyses on six cross-domain settings. The significant improvements of Tri-CDR with different sequential encoders verify its effectiveness and universality. The source code is available in <https://github.com/hulkima/Tri-CDR>.

CCS Concepts: • **Information systems** → **Recommender systems**.

Additional Key Words and Phrases: cross-domain recommendation, contrastive learning, triple learning

## 1 INTRODUCTION

Personalized recommendation aims to capture user interests and provide appropriate items [25, 35, 50]. Sequential recommendation (SR), which focuses on discovering user preferences from the essential information of users'

---

<sup>†</sup> indicates equal contributions.

<sup>\*</sup> indicates corresponding author.

• This work has been done when the author was at Tencent for internship.

Authors' addresses: Haokai Ma<sup>†,•</sup>, School of Software, Shandong University, Jinan, China, [mahaokai@mail.sdu.edu.cn](mailto:mahaokai@mail.sdu.edu.cn); Ruobing Xie<sup>†</sup>, TEG, Tencent, Beijing, China, [ruobingxie@tencent.com](mailto:ruobingxie@tencent.com); Lei Meng<sup>\*</sup>, School of Software, Shandong University; and Shandong Research Institute of Industrial Technology, Jinan, China, [lmeng@sdu.edu.cn](mailto:lmeng@sdu.edu.cn); Xin Chen, WeChat, Tencent, Beijing, China, [andrewxchen@tencent.com](mailto:andrewxchen@tencent.com); Xu Zhang, WeChat, Tencent, Beijing, China, [xuonezhang@tencent.com](mailto:xuonezhang@tencent.com); Leyu Lin, WeChat, Tencent, Beijing, China, [goshawklin@tencent.com](mailto:goshawklin@tencent.com); Jie Zhou, WeChat, Tencent, Beijing, China, [withtomzhou@tencent.com](mailto:withtomzhou@tencent.com).

---

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](mailto:permissions@acm.org).

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

ACM 1046-8188/2023/12-ART

<https://doi.org/10.1145/3638351>

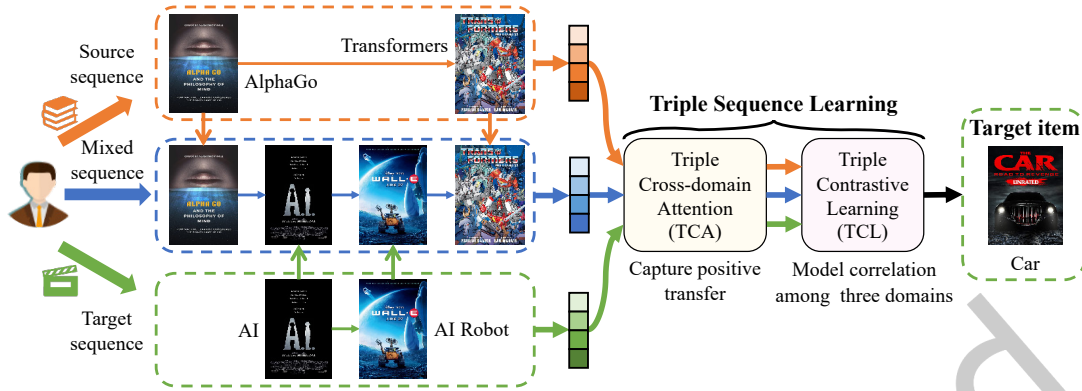


Fig. 1. Triple sequence learning on the source, target, and mixed behavior sequences in CDR.

historical behaviors, has attracted significant attention [20]. However, real-world SR models usually face the data sparsity problem, since users usually have few behaviors [29, 30, 57]. In practice, making full use of user behaviors in other domains under the user's approval is a straightforward and effective solution to the data sparsity issue in a single domain.

Cross-domain recommendation (CDR) concentrates on transferring useful information from the source domain to the target domain for performance gains in the target domain [18, 23]. Existing CDR methods mainly focus on modeling the relations between the source and target domains. EMCDR [34] and SSCDR [19] attempt to learn a mapping function across the source/target domains via aligned objects. CoNet [18] and MiNet [37] adopt explicit cross-domain information paths or attention mechanisms for knowledge transfer. Some CDR methods further build the cross-domain connections via (multi-domain) global graphs [55, 63] or feature correlations [22, 56]. However, most existing CDR models simply focus on the dual relations between the source and target behavior sequences, ignoring the rich information of the natural mixed (i.e., source+target) behavior sequence.

We define the **mixed behavior sequence** in CDR as a complete user behavior sequence containing behaviors in both the source and target domains which are ordered chronologically. In parallel, we unify all items from the source and target domains to define the **mixed domain** and construct distinct embedding representations for each item within its respective domain and the mixed domain. The left part of Figure 1 shows an example of the source, target, and mixed behavior sequences in a cross-domain sequential recommendation (CDSR) scenario. The mixed sequence can reflect a user's complete behavioral pattern and preference evolution more comprehensively and thus helps to better extract users' global interests. For example, the sequential behaviors in domain *book* and *movie* are not consistent and reasonable. Only through the complete mixed sequence containing sequential behaviors of [book: AlphaGo] → [movie: AI] → [movie: AI Robot] → [book: Transformers] → [movie: Car] can we fully understand the user's sequential action logic. We firmly believe that jointly modeling the mixed behavior sequence with the original two source/target sequences is beneficial to capture both inter-domain and intra-domain information in CDR.

In this work, we propose a new paradigm that jointly models source, target, and mixed behavior sequences in CDR. The challenges are three-fold: (1) **How to extract more informative knowledge from source and mixed sequences?** User behaviors in other domains may be good supplements, while it is also common that users have different preferences in these domains. We should maximize the cross-domain information gain while alleviating the negative transfer from possible noises. (2) **How to model the triple correlations among source, target, and mixed sequences?** The mixed sequence is built by source and target behaviors. Both the coarse-grained similarities and fine-grained distinctions among the three sequences should be carefully considered. The dual relation learning of conventional CDR models cannot be directly transferred to the triple learning

task with the additional mixed sequence. (3) *How to construct a universal CDR framework that could smoothly cooperate with different types of single-domain SR models?* Currently, lots of CDR models rely on complicated and customized networks for inter-domain interactions, which are hard to be directly adopted with other single-domain models. We aim to build a universal model-agnostic CDR framework that could be beneficial with frequently updated single-domain models.

To address these issues, we propose a novel **Triple sequence learning for cross-domain recommendation (Tri-CDR)**, serving as a model-agnostic framework to jointly model the source, target, and mixed sequences in CDSR. Specifically, we first build three sequence encoders for the source, mixed, and target domains, respectively, which model the intra-domain behavior interactions to get three hidden sequence representations. Next, to alleviate the irrelevant negative transfer, we design a **triple cross-domain attention (TCA)** method on three sequence representations to capture the informative knowledge related to users' target-domain preferences and global interests. These attention-enhanced sequence representations are then combined and fed into a Multi-Layer Perceptron (MLP) to get the final user representation. We further propose a **triple contrastive learning (TCL)** strategy to comprehensively model the correlations among three sequences. TCL adopts three CL losses to capture the coarse-grained similarities between any two sequence representations of the same user compared to other users'. More importantly, it further employs a margin-based triple loss among three sequence representations to model their fine-grained distinctions, keeping the information diversity in three domains. The advantages of Tri-CDR are: (1) the TCA enables an informative knowledge transfer related to users' target-domain preferences and global interests. (2) The TCL helps to better capture the correlations among three domains in representation learning. (3) Tri-CDR is effective, universal, and easy-to-deploy, which could be conveniently applied with different sequence encoders and additional objectives.

In experiments, we have conducted an extensive evaluation on six cross-domain settings with various base sequence encoders. As a result, Tri-CDR achieves significant improvements on all settings. We also conduct various ablation studies, universality analyses, parameter analyses, and visualization to verify the effectiveness of the proposed TCA and TCL. The contributions are summarized as follows:

- We have verified the significance of the triple sequence modeling for comprehensive user interest understanding. To the best of our knowledge, we are the first to present the triple sequence learning among source, target, and mixed behavior sequences in CDR.
- We propose a triple cross-domain attention (TCA) method to enable more positive transfer of knowledge from the source domain to the target one, which considers both the user's target-domain preferences and global interest from the source and mixed behavior sequences.
- We creatively design the triple contrastive learning (TCL) strategy, which not only models the coarse-grained similarities among multi-domain sequence representations of the same user but also detects the fine-grained distinctions via a margin-based triple loss.
- We conduct an extensive evaluation to verify the effectiveness of our Tri-CDR on multiple datasets with different base models. The proposed model is effective, universal, and easy-to-deploy.

## 2 RELATED WORK

**Cross-domain Recommendation.** Cross-domain recommendation (CDR) is a representative method to alleviate the data sparsity problem in a single domain with auxiliary information from other domains [23, 31, 34]. The basic assumption is that users' behaviors in different domains reflect the user's personal preferences to a certain extent. Classical CDR methods aim to model cross-domain knowledge transfer through directly mapping [19, 34], multi-domain interaction modeling [2, 37], meta-learning [71, 72], and transformation matrix [18, 23]. Recently, some CDR methods also leverage alignment constraint [27, 48], adversarial learning [21] and large language model [1, 8, 45] for cross-domain knowledge representation and fusion. BiTGCF [26] designs a bi-directional transfer

learning method to transfer users' single-domain preferences across two domains. CCTL [62] and AutoTransfer [7] attempt to adaptively sample informative instances from the source domain and transfer them into the target domain for leveraging the multi-domain collaborative knowledge while avoiding negative migration. As a specialized scenario, CDSR concentrates more on users' multi-domain sequential behavior modeling in CDR [22, 26, 33, 66]. Lots of CDSR methods focus on the shared-account CDSR scenario [11, 12, 33, 49].  $\pi$ -Net [33] and PSJNet [43] utilize a shared account filter unit and a cross-domain transfer unit to share information across both domains synchronously. MIFN [32] further enriches the sequence representation with knowledge graphs based on  $\pi$ -Net. TiDA-GCN [12] incorporates the time interval knowledge into attention mechanisms to further enhance the message-passing and representation learning process. RL-ISN [11] introduces a reinforcement learning-based solution to reduce the impact of irrelevant domain information in the recent share-account behaviors. DAGCN [10] designs a domain-aware graph convolution network to learn user-specific node representations on the global static graph. CDHRM [49] jointly captures users' inter-session and intra-session behavioral dynamics from different domains. DDGHM [66] builds a global dynamic graph to model source-target interactions directly, and jointly predicts via local and global information. DASL [22] proposes the dual embedding and dual attention strategies to model the correlations between source and target domains' sequences. UniSRec [17] models the transferable representations across different recommendation domains and platforms via a pre-trained universal sequence representation model, which is trained with items' associated description text. DR-MTCR [13] designs a unified disentanglement module to capture the domain-shared and domain-specific information, with the aim to transfer the trustworthy information across domains. IESRec [28] reflects the users' diverse characteristics and reduces domain discrepancy for the item embeddings with the proposed internal multi-interest exploration module and external domain alignment module. However, existing CDSR methods merely focus on dual relations of source  $\rightarrow$  target or global  $\rightarrow$  local, ignoring directly modeling the natural mixed behavior sequence with the correlations among source, target, and mixed domains. Our Tri-CDR is different from these works: (a) we directly model three mixed, source, and target behavior sequences and use them for recommendations simultaneously. (b) We emphasize the ternary relationship among three sequences for positive knowledge transfer. (c) Tri-CDR is model-agnostic and easy-to-deploy that can be applied to different base sequence encoders and even intra-domain CL tasks. By doing this, our work can not only learn both complete and independent preferences from mixed, source, and target behavior sequences but also explain users' sequential behavior comprehensively through their interactions.

**Sequential Recommendation.** Sequential Recommendation (SR) attempts to capture the user's time-aware preferences by modeling the sequential dependencies of the user's historical behavior to recommend the next item that the user may be interested in. Early works reason users' short-term preference through the Markov Chains (MCs) [15, 41]. In recent years, researchers have leveraged the Convolution Neural Network (CNN) [44], Recurrent Neural Network (RNN) [16] and Transformer [20, 42] to capture users' preference patterns from users' historical behaviors. Among them, GRU4Rec [16] leverages the Gate Recurrent Unit (GRU) as the sequential encoder to learn users' long-term dependencies. SASRec [20] introduces Transformer for behavior interaction modeling and is widely used in practice. S<sup>3</sup>Rec [68] leverage the mutual information maximization principle to employ four self-supervised objectives among item, attribute, sub-sequence, and sequence. DUVRec [59] encodes the sequential information with dual-view user representation (item-view and factor-view) to achieve enhanced SR performance. CL4SRec [57] is one of the state-of-the-art SR models that further enhance the sequential modeling with various intra-domain CL tasks. In this work, we have successfully adopted Tri-CDR with different sequence encoders, including GRU4Rec, SASRec, and CL4SRec.

**Contrastive Learning in Recommendation.** As a common self-supervised learning (SSL) method, contrastive learning (CL) has been widely used in fields of Computer Vision (CV) [5, 14], Natural Language Processing (NLP) [9, 54] and Recommendation System (RS) [51, 57]. In recommendation, CL is widely applied to session-based

recommendation [69], multi-behavior recommendation [53], sequential recommendation [39, 57, 68], and cross-domain recommendation [3, 55]. C<sup>2</sup>-CRS [69] proposes a coarse-to-fine contrastive learning method to model user preference with multi-level semantic fusion. MMCLR [53] designs three CL tasks to learn the correlations among different behavior types and modeling views. DUORec [39] proposes a contrastive regularization with the model-level augmentation to reshape and improve the embedding distribution and sequence representations. CL4SRec [57] proposes three sequence-based augmentations to build positive pairs in SSL. CDDR [55] designs an intra-domain CL task and three inter-domain CL tasks for cross-domain knowledge transfer in graph-based matching. C<sup>2</sup>DSR [3] conducts a cross-domain infomax objective to enhance the correlation between global and local representations with domain-specific global augmentations. Some recent works [52, 53, 61] have conducted certain triplet losses for more precise representation learning in recommendation. However, existing CL-based CDR models simply maximize the mutual information of representations in different domains, ignoring their conflicts that may lead to negative transfer and model collapse. To the best of our knowledge, we are the first to jointly model coarse-grained similarity and fine-grained distinction via CL in CDR.

### 3 METHOD

The primary core of this work is to alleviate the negative transfer in CDR, which refers to the phenomenon in which the knowledge captured from the source/mixed domain adversely affects the recommendation precision in the target domain. To this end, we design a triple cross-domain attention (TCA) strategy to reweight the components related to users' target-domain preferences and global interest within the triple behavior domains and a triple contrastive learning (TCL) method to accurately understand the triple correlation by modeling the coarse-grained similarities and the fine-grained distinctions among source, target, and mixed domains respectively. These two components enable existing CDR methods to alleviate the negative transfer while maintaining the positive transfer in cross-domain transfer.

#### 3.1 Problem Formulation

We first define the *source behavior sequence*  $S^S = \{\mathbf{d}_1^S, \mathbf{d}_2^S, \dots, \mathbf{d}_p^S\}$  in the source domain  $S$  and the *target behavior sequence*  $S^T = \{\mathbf{d}_1^T, \mathbf{d}_2^T, \dots, \mathbf{d}_q^T\}$  in the target domain  $T$  for each user, where  $p, q$  are the source/target historical behavior lengths, and  $\mathbf{d}_i^S$  and  $\mathbf{d}_j^T$  are behavior embeddings. In Tri-CDR, we propose a third *mixed behavior sequence*  $S^M = \{\mathbf{d}_1^M, \mathbf{d}_2^M, \dots, \mathbf{d}_{p+q}^M\}$  as a supplement to source/target sequences, which is the complete user behavior sequence containing both source and target behaviors in chronological order. Given three behavior sequences  $S^S, S^T$  and  $S^M$ , Tri-CDR tries to recommend the target item  $d_{q+1}^T$  that will be interacted by this user in the target domain.

#### 3.2 Overall Framework

In this section, we describe the proposed model-agnostic Triple sequence learning for cross-domain recommendation (Tri-CDR) framework, which jointly models source, target, and mixed behavior sequences to improve CDR. Specifically, we first model the source, target, and mixed behavior sequences through three base sequence encoders separately to generate their corresponding hidden sequence representations in three domains. Then we propose a triple cross-domain attention (TCA) method to highlight informative knowledge related to the user's target-domain preferences and global interests in building three domains' sequence representations, mitigating negative knowledge transfer. To better model the correlations among multi-domain sequence representations, we design a novel triple contrastive learning (TCL) strategy with two contrastive constraints: (a) Coarse-grained similarity modeling, which enables source/target/mixed sequence representations from the same user to be more similar than other users'. (b) Fine-grained distinction modeling, which recognizes users' diversified preferences

Table 1. Notations and Descriptions of Tri-CDR.

Notation	Description (Superscripts Represent Domains)
$p, q$	Length of the user behavior in source and target domain.
$d$	Embedding size.
$S, T, M$	Source domain, target domain and mixed domain.
$S^S, S^T, S^M$	User behavior sequence of source, target and mixed domain.
$d_i^S, d_i^T, d_i^M$	The $i$ -th behavior in source, target and mixed domain.
$\mathbf{d}_i^S, \mathbf{d}_i^T, \mathbf{d}_i^M$	The $i$ -th behavior embedding in source, target and mixed domain.
$\mathbf{H}^S, \mathbf{H}^T, \mathbf{H}^M$	The hidden behavior representation matrix in source, target and mixed domain.
$\mathbf{h}_i^S, \mathbf{h}_i^T, \mathbf{h}_i^M$	The $i$ -th hidden behavior representation in source, target and mixed domain.
$\mathbf{s}^S, \mathbf{s}^T, \mathbf{s}^M$	Sequence representation of source, target and mixed domain.
$\alpha_i^S, \alpha_i^T, \alpha_i^M$	The attention weight of source, target and mixed domain.
$P^S(\cdot), P^T(\cdot), P^M(\cdot)$	The domain-specific projector of source, target and mixed domain.
$\tilde{\mathbf{s}}_i^S, \tilde{\mathbf{s}}_i^T, \tilde{\mathbf{s}}_i^M$	Projected sequence representation of source, target and mixed domain.
$\mathbf{u}^T$	User final representation.
$y_{u,q+1}^T$	The ground-truth probability between user $u$ and item $d_{q+1}^T$ .
$\hat{y}_{u,q+1}^T$	Predicted probability between user $u$ and item $d_{q+1}^T$ .
$\mathcal{L}_{CL}$	Contrastive informative regularizer.
$\mathcal{L}_{CSM}$	Coarse-grained similarity modeling regularizer.
$\mathcal{L}_{FDM}$	Fine-grained distinction modeling regularizer.
$\mathcal{L}_{CTR}$	The binary cross-entropy loss regularizer.
$\mathcal{L}$	The overall regularizer.

in different domains to keep the information gains brought by the source and mixed sequences. TCL helps to learn more informative and accurate multi-domain representations to capture user preferences comprehensively.

The overall structure of Tri-CDR is illustrated in Fig. 2, and the notations and their corresponding descriptions of Tri-CDR are illustrated in Table. 1. In the following subsections, we first present the base sequence encoder of the proposed Tri-CDR which is implemented with the self-attention module. Subsequently, we introduce TCA, which emphasizes the user's target-domain preferences and global interest in triple domains. And then, we describe TCL to precisely model the user's coarse-grained similarities and fine-grained distinctions among triple sequences. Finally, we present the discussions on the proposed Tri-CDR.

### 3.3 Base Sequence Encoder

Inspired by the success of the self-attention mechanism in sequential recommendation, we apply SASRec [20] as our sequence encoder for all domains. Without losing generality, for the target domain sequence  $S^T$ , we build the input matrix  $\mathbf{D}^T \in \mathbb{R}^{q \times d}$ , where each behavior embedding  $\mathbf{d}_i^T$  consists of a learnable item ID embedding and a position embedding, and  $d$  is the embedding size. Then the sequence encoder transposes  $\mathbf{D}^T$  into three matrices

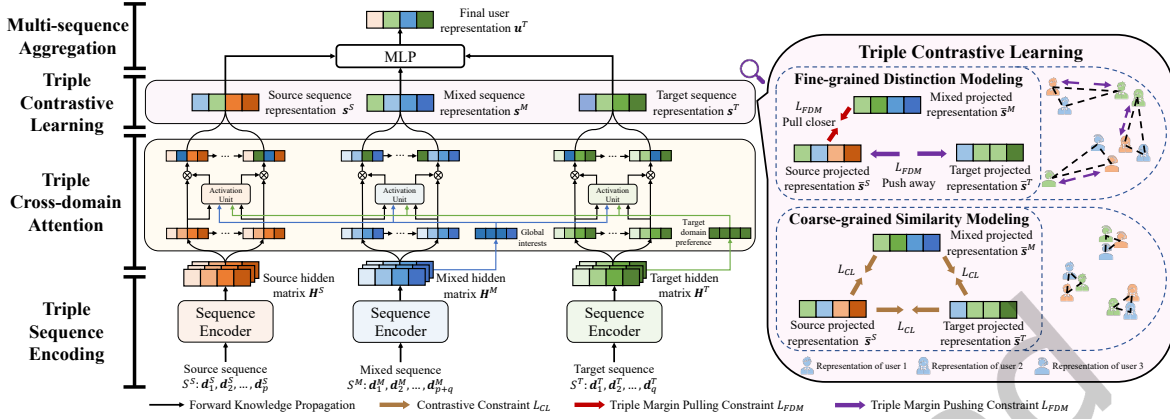


Fig. 2. The overall structure of Tri-CDR. TCA highlights informative knowledge related to users' target-domain preferences and global interests, while TCL captures the triple correlations among three domains for better cross-domain knowledge transfer.

by linear projections, and feeds them into the attention method as query, key, and value, which can be defined as:

$$\hat{H}^T = \text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V. \quad (1)$$

where  $Q = D^T W^Q$ ,  $K = D^T W^K$ ,  $V = D^T W^V$ , and  $W^Q, W^K, W^V$  denote the linear projections respectively. Meanwhile, we conduct a point-wise feed-forward network to get the hidden behavior matrix  $H^T$  of the target domain as follows:

$$H^T = \text{ReLU}\left(\hat{H}^T w_1 + b_1\right) w_2 + b_2, \quad H^T \in \mathbb{R}^{q \times d}. \quad (2)$$

where  $w_1, w_2, b_1, b_2$  denote the weight matrices and bias vectors respectively. The hidden matrices of source/mixed sequences  $H^S, H^M$  are similarly constructed with the intra-domain behavior interactions. Note that the same items in different domains are allocated with different behavior embeddings for better representation capacity, avoiding too homogeneous representations across different domains. It is also worth noting that we can conveniently adopt other sequential models as our sequence encoder or even with other intra-domain CL tasks in Tri-CDR (we have tested GRU4Rec [16] and CL4SRec [57], see the universality analysis in Sec. 4.6).

### 3.4 Triple Cross-domain Attention

It is intuitive that there exist subtle variations in preferences between different domains, even for the identical user. Consequently, a mechanical and unrefined transference of the source domain knowledge to the target domain can potentially introduce bias into the target recommender's optimization. We argue that different historical behaviors in three domains should have different importance for the target-domain prediction. Precisely, We expect that the information emphasized in three sequences should be (a) relevant to the user's target-domain preferences, so as to fit the target-domain prediction task, and (b) relevant to the user's global interests, so as to understand the user's comprehensive preferences and bring in more information gain. Hence, we propose a *Triple Cross-domain Attention (TCA)* on three hidden behavior matrices to enable the recommender to focus on the relevant and informative knowledge in the triple domains while filtering out noisy information selectively, thereby maintaining more positive transfer and alleviating the negative transfer.

In essence, TCA is proposed to highlight valuable information, accelerate positive knowledge transfer, and enable the recommender to reweight different components of the users' behaviors in triple domains. TCA functions

when we aggregate the hidden behavior embeddings in  $H^S$ ,  $H^T$ ,  $H^M$  to get three sequence representations  $s^S$ ,  $s^T$ ,  $s^M$  of the source, target, and mixed domains. Specifically, for the source domain, given  $H^S = \{h_1^S, \dots, h_p^S\}$  and the target item  $d_{q+1}^T$ , TCA calculates the attention weight  $\alpha_i^S$  of the  $i$ -th behavior's hidden embedding  $h_i^S$  in the source sequence as:

$$\alpha_i^S = f(h_q^T, h_{p+q}^M, h_i^S) = \text{MLP}^a(h_q^T \parallel h_i^S \parallel h_i^S - h_q^T \parallel h_i^S \odot h_q^T \parallel h_{p+q}^M). \quad (3)$$

Here,  $\odot$  performs element-wise vector product,  $\parallel$  denotes the concatenation operation, and  $\text{MLP}^a(\cdot)$  is a two-layer fully-connected network followed by PReLU activation functions.  $h_q^T$  and  $h_{p+q}^M$  denote the last hidden behavior embeddings in target and mixed domains. This attention setting is inspired by [67], while we adopt  $h_q^T$  to indicate the user's current target-domain preference instead of using different candidate target item's embedding  $d_{q+1}^T$ , since it is much more efficient in online serving. Moreover, we also highlight  $h_{p+q}^M$  in this attention, since we assume the hidden embedding of the last mixed behavior implies the user's global interests, which are good supplements to the target-domain preference. Based on this, we aggregate different behavior hidden embeddings  $h_i^S$  to get the source sequence representation  $s^S$  as:

$$s^S = \sum_{i=1}^p \text{Softmax}(\alpha_i^S) h_i^S. \quad (4)$$

Similarly, we also have  $s^M = \sum_{i=1}^{p+q} \text{Softmax}(f(h_q^T, h_{p+q}^M, h_i^M)) h_i^M$  and  $s^T = \sum_{i=1}^q \text{Softmax}(f(h_q^T, h_{p+q}^M, h_i^T)) h_i^T$  for the mixed and target sequence representations. With TCA, we can not only directly focus on the information related to the user's target-domain preferences, but also keep aware of the user's global interests for a more comprehensive positive transfer from the three domains. By highlighting the practical application of TCA in exploring the user's multiple preferences, it assists the recommender in mitigating the intricate negative information from source and mixed domains during cross-domain transfer, consequently ensuring precise modeling of multi-domain user preferences.

### 3.5 Triple Contrastive Learning

Compared to classical CDR models that learn dual relations, Tri-CDR faces a more challenging task to comprehensively understand triple correlations among source, target, and mixed sequences. In this work, we propose a novel *Triple Contrastive Learning (TCL)* to smartly model the correlations among three sequence representations. Precisely, we design two CL tasks in TCL, including the coarse-grained similarity modeling and the fine-grained distinction modeling. The former CL task aims to capture the coarse-grained similarities between any two sequence representations of the same user compared to others. In contrast, the latter is conducted to model the fine-grained distinctions among users' multi-domain sequence representations, keeping the diversity across different domains to enhance information gains.

**3.5.1 Coarse-grained Similarity Modeling (CSM).** It is natural that a user's behavior sequences in different domains should share common general preferences. Hence, we design a **coarse-grained similarity modeling (CSM)** to model the coarse-grained similarities among three domains' sequence representations of the same user. Specifically, we project the sequence representations  $s^S$ ,  $s^T$ ,  $s^M$  into the same space via domain-specific projectors  $P^S(\cdot)$ ,  $P^T(\cdot)$ ,  $P^M(\cdot)$  (we build these projectors via one-layer MLPs). After obtaining the projected sequence representations  $\bar{s}^S = P^S(s^S)$ ,  $\bar{s}^T = P^T(s^T)$ ,  $\bar{s}^M = P^M(s^M)$  of three domains, we calculate the contrastive loss  $\mathcal{L}_{CL}$  with any two of them as positive instances in CL. Formally, for  $\bar{s}^S$  and  $\bar{s}^T$ , we follow the classical InfoNCE [5] as:

$$\mathcal{L}_{CL}(\bar{s}^S, \bar{s}^T) = - \sum_{i \in B} \log \frac{\exp(\text{sim}(\bar{s}_i^S, \bar{s}_i^T) / \tau)}{\sum_{j \in B \setminus i} \exp(\text{sim}(\bar{s}_i^S, \bar{s}_j^T) / \tau)}. \quad (5)$$

$B$  denotes the sampled batch,  $\bar{s}_i^T$  denotes the projected sequence representation for item  $i$  in target domain,  $\tau$  denotes the temperature coefficient, and  $\text{sim}(\bar{s}_i^S, \bar{s}_i^T) = \frac{(\bar{s}_i^S)^T (\bar{s}_i^T)}{\|\bar{s}_i^S\| \cdot \|\bar{s}_i^T\|}$  denotes the cosine similarity. Finally, the CSM



loss  $\mathcal{L}_{CSM}$  is formulated with three positive pairs as:

$$\mathcal{L}_{CSM} = \sum_u (\lambda_1 \mathcal{L}_{CL}(\bar{s}^M, \bar{s}^S) + \lambda_2 \mathcal{L}_{CL}(\bar{s}^M, \bar{s}^T) + \lambda_3 \mathcal{L}_{CL}(\bar{s}^S, \bar{s}^T)), \quad (6)$$

where  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$  denote the loss weights respectively.

**3.5.2 Fine-grained Distinction Modeling (FDM).** CSM assumes that a user's domain-specific preferences should be more similar, and monotonously pulls the multi-domain representations of the same user closer. It functions well in general, while there does exist fine-grained distinctions across the user's preferences in triple domains. Over-optimizing  $\mathcal{L}_{CSM}$  may inevitably cause the model to collapse to the proximal point, where the source and mixed sequences cannot provide additional information gain for their too-similar target sequence, putting the cart before the horse.

To address this issue, we look back to the composition of the proposed mixed sequence, whose subsequences contain both source and target sequences. Hence, it is intuitive that the source-mixed and target-mixed distances should be smaller than the source-target distance. Under this intuition, we propose a new CL task named **fine-grained distinction modeling (FDM)** based on a margin-based triplet loss. We intuitively assume the distance between  $\bar{s}^S$  and  $\bar{s}^T$  should be larger than the distance between  $\bar{s}^S$  and  $\bar{s}^M$  by at least the margin  $\gamma$ . Formally, we define the FDM loss  $\mathcal{L}_{FDM}$  as:

$$\mathcal{L}_{FDM} = \sum_u \max \left\{ d(\bar{s}^S, \bar{s}^M) - d(\bar{s}^S, \bar{s}^T) + \gamma, 0 \right\}, \quad (7)$$

where  $d(\bar{s}^S, \bar{s}^T)$  measures the similarity between  $\bar{s}^S$  and  $\bar{s}^T$ , which is calculated as the  $L_2$  distance.  $\gamma$  denotes the margin to be controlled. Through jointly optimizing  $\mathcal{L}_{CSM}$  and  $\mathcal{L}_{FDM}$ , Tri-CDR could intelligently find a good balance in ensuring the alignment among a user's multi-domain preferences while keeping the fine-grained distinctions to bring in more information gain from other domains.

### 3.6 Optimization Objectives

After TCA and TCL, we concatenate the triple sequence representations  $s^M$ ,  $s^S$ ,  $s^T$  and feed them into the following multi-sequence aggregation layer to generate the user final representation  $u^T$ , which is formulated as follows:

$$u^T = \text{MLP}^f(s^M \parallel s^S \parallel s^T). \quad (8)$$

$\text{MLP}^f(\cdot)$  denotes a two-layer fully-connected network with the LeakyReLU activation. Finally, we calculate the predicted probability  $\hat{y}^T = (u^T)^\top d_{q+1}^T$  of the user  $u$  on the target item  $d_{q+1}^T$  with the final user representation  $u^T$  and the item embedding  $d_{q+1}^T$ . We formulate the binary cross-entropy loss  $\mathcal{L}_{CTR}$  as follows:

$$\mathcal{L}_{CTR} = - \sum_{(u,d) \in R^T} \left[ y_{u,d}^T \log \hat{y}_{u,d}^T + (1 - y_{u,d}^T) \log (1 - \hat{y}_{u,d}^T) \right] \quad (9)$$

where  $R^T$  is the target-domain training set,  $y_{u,d}^T = 1$  and  $y_{u,d}^T = 0$  denote the positive and negative samples respectively, and  $\hat{y}_{u,d}^T$  denotes the predicted probability of  $(u, d)$ . To optimize across triple sequences in conjunction with CL tasks, the overall objective function  $\mathcal{L}$  is a linear combination of  $\mathcal{L}_{CTR}$ ,  $\mathcal{L}_{CSM}$  and  $\mathcal{L}_{FDM}$  as:

$$\mathcal{L} = \mathcal{L}_{CTR} + \lambda_{CSM} \mathcal{L}_{CSM} + \lambda_{FDM} \mathcal{L}_{FDM} \quad (10)$$

where  $\lambda_{CSM}$  and  $\lambda_{FDM}$  denote the loss weights of  $\mathcal{L}_{CSM}$  and  $\mathcal{L}_{FDM}$ .

**Algorithm 1** The Tri-CDR Algorithm.

**Input:** User behavior sequences  $S^S, S^T$  and  $S^M$  of source, target and mixed domain; The pre-trained sequence encoder parameters  $\theta^S, \theta^T$  and  $\theta^M$ ; The temperature coefficient  $\tau$ ; The loss weight  $\lambda_{CSM}$  and  $\lambda_{FDM}$  of  $\mathcal{L}_{CSM}$  and  $\mathcal{L}_{FDM}$ ; The loss weight  $\lambda_1, \lambda_2$  and  $\lambda_3$  of three CL tasks; The triple margin  $\gamma$ ; Batch size  $B$ ; User set  $\mathcal{U}$ ; Training set  $\mathcal{R}$ .

**Output:** The Tri-CDR model parameter  $\theta^*$  and the predicted probability between user  $u$  and item  $d$  for recommendation.

**## Training Procedure**

**while** not converged **do** :

**for**  $b = \{1, 2, \dots, |\mathcal{R}|/B\}$  :

    Compute the hidden behavior representation matrices  $H^S, H^T, H^M$  via three base sequence encoders;

    Dynamically Learn the attention weights  $\alpha_i^S, \alpha_i^T, \alpha_i^M$  of each behavior according to Eq.(3);

    Generate the sequence representation  $s^S, s^T, s^M$  via Eq.(4);

    Obtain the projected sequence representations  $\bar{s}^S, \bar{s}^T, \bar{s}^M$  with three domain-specific projectors  $P^S, P^T, P^M$ ;

    Calculate the coarse-grained similarity modeling regularizer  $\mathcal{L}_{CSM}$  with Eq.(6);

    Calculate the fine-grained similarity modeling regularizer  $\mathcal{L}_{FDM}$  with Eq.(7);

    Generate the user final representation  $u^T$  with Eq.(8);

    Calculate the binary cross-entropy loss regularizer  $\mathcal{L}_{CTR}$  with Eq.(9);

    Calculate the overall regularizer  $\mathcal{L}$  via Eq.(10) and update the model parameters  $\theta$  with  $\nabla_{\theta} \mathcal{L}$ .

**end for**

**end while**

**## Inference Procedure**

**for**  $u = \{1, 2, \dots, \mathcal{U}\}$  :

  Select user  $u$ 's behavior sequences  $S_u^S, S_u^T$  and  $S_u^M$  of source, target and mixed domain;

  Compute the hidden behavior representation matrices  $H_u^S, H_u^T, H_u^M$  via three base sequence encoders of user  $u$ ;

  Dynamically Learn the attention weights  $\alpha_{u,i}^S, \alpha_{u,i}^T, \alpha_{u,i}^M$  of user  $u$ 's each behavior according to Eq.(3);

  Generate the sequence representation  $s_u^S, s_u^T, s_u^M$  via Eq.(4);

  Calculate the user final representation  $u^T$  via Eq.(8);

  Compute the predicted probability  $\hat{y}_{u,d}^T$  between user  $u$  and the target item  $d^T$ .

**end for**

**return** The predicted probability of user  $u$  and target item  $d$ .

### 3.7 Training Strategy of CDSR

Training Tri-CDR from scratch sometimes may incur the difficulty in model convergence in the early training stage, leading to unstable and unsatisfactory performance. This is mainly caused by the feature space conflicts between the source and target domains in CDR. We also observe the same phenomenon in other CDSR models. To address this issue, we conduct a two-step training strategy (i.e., single-domain pre-training + cross-domain fine-tuning) to achieve more stable and satisfactory performances. Specifically, we first pre-train two source/target SASRec models with their corresponding single-domain losses. Next, the pre-trained SASRec parameters are used as the initialization of Tri-CDR (and other CDR baselines). All parameters are then tuned via  $L$  in Eq. (10). Through this, CDR models could be trained more effectively and stably. The complete workflow of Tri-CDR is comprehensively outlined in Algorithm 1.

Table 2. Comparison among representative cross-domain sequential recommendation algorithms.

Algorithms	Domain utilization			Contrastive Learning Implementation	Sequence modeling universality
	Mixed Domain	Source Domain	Target Domain		
$\pi$ -Net [33]	×	✓	✓	×	×
DTCDR [70]	×	✓	✓	×	×
DDTCDR [23]	×	✓	✓	×	×
DASL [22]	×	✓	✓	×	×
DDGHM [66]	✓	×	✓	intra-CL	×
C <sup>2</sup> DSR [3]	✓	×	✓	inter-CL	×
Tri-CDR (ours)	✓	✓	✓	Both intra- and inter- CL	✓

### 3.8 In-depth Model Discussions

In this section, we undertake a comparison between the proposed Tri-CDR and the existing CDSR methods specifically tailored for the domain utilization, CL implementation, sequence modeling universality and complexity analyses, intending to analyze its novelty and effectiveness. The related comparison results are illustrated in Table. 2.

**3.8.1 Comparison with Existing CDR Methods in Domain Utilization.** CDSR aims to predict the next item that the user will be consumed in the target domain by leveraging the historical behavior sequence in the source domain. Therefore, the pioneer CDSR methods primarily focus on exploring how to achieve meaningful information transfer with the leverage of user’s behavior in the source domain [22, 23, 33, 70]. With the profound advancement of CDSR, some studies attempt to incorporate the mixed behavior sequence containing both source and target behaviors in chronological order to model the user’s global interests. However, these studies are specifically designed for Cross-domain Session-based Recommendation [49] and Cross-domain Share-account recommendation [33], and may not be directly applicable to CDSR scenarios. Recently, DDGHM [66] and C<sup>2</sup>DSR [3] introduce the mixed behavior sequence into CDSR through the utilization of the global graph, yielding promising performance. Nevertheless, the aforementioned approaches typically model users’ global and local interests separately with the mixed and target behavior sequences to capture their dynamic correlations. Despite the commendable performance of these methods, it overlooks the modeling of the user’s preference in the source domain, which is the coherent sequence with its internal logic and equally critical for CDSR. Consequently, it may result in sub-optimal performance.

In contrast, the proposed Tri-CDR jointly models the user’s dynamic preference from the source, mixed and target domain, which facilitates the flexible and convenient integration of all available information into CDSR scenarios. Meanwhile, Tri-CDR proposes the Triple Cross-domain Attention mechanisms and Triple Contrastive Learning strategies for negative filtering, enabling fine-grained inter-domain relationship modeling and precise cross-domain positive knowledge transfer. The source, mixed and target domains are beneficial in CDR, while it is non-trivial to make full use of them. More detailed experimental results and analyses of the domain utilization are described in Sec.4.7.

**3.8.2 Comparison with Existing CDR Methods in Contrastive Learning Implementation.** As a self-supervised learning strategy, CL has been widely applied to collaborative filtering [24], multi-media recommendation [51], and sequential recommendation [39, 57]. Recently, some studies [3, 55, 66] have also introduced CL into CDR. DDGHM [66] designs an intra-domain contrastive metric with the random item augmentation operator to enhance representation learning and alleviate data sparsity issues. However, despite employing random sequence-based CL augmentation for inferring informative representation, DDGHM’s intra-CL method does not explicitly consider the knowledge transfer and inter-domain correlation necessary for cross-domain modeling. C<sup>2</sup>DSR [3] is the most related CDR model, which develops an inter-domain contrastive infomax objective to improve the

correlation between the domain-aware prototype representations and corresponding corrupted representations. However, its inter-CL objective over-maximizes the mutual information of the individual representations between the mixed and target domain, disregarding the potential similarity conflicts that may result in optimization collapse. Both coarse-grained similarity (refer to positive transfer) and fine-grained distinction (refer to negative transfer) should be considered in inter-domain CL.

Rather than solely relying on inter-CL between two domains, Tri-CDR incorporates all three domains and carefully model their triple relationships in contrastive tasks, ensuring that the optimization does not excessively collapse and get sufficient and accurate training. Precisely, Tri-CDR creatively proposes the coarse-gained similarity modeling and the fine-gained distinction modeling to comprehensively understand the triple correlations. The former captures the coarse-gained similarities between any two single-domain sequences of the same user to enable effective representation learning, while the latter maintains the robust cross-domain positive transfer through modeling the fine-gained distinction among triple sequences (see Sec. 4.5). Moreover, as an effective and model-agnostic framework, Tri-CDR is able to bring the SOTA CL-based SR models (intra-domain CL) into the sequence modeling and achieves significant improvements in CDR by leveraging the combined interplay of both inter-CL and intra-CL strategies (see Sec. 4.6).

**3.8.3 Comparison with Existing CDR Methods in Sequence Modeling Universality.** Behavior sequence modeling is essential in SR. Lots of novel techniques such as attention mechanisms [6, 40], side information [40, 58], and contrastive learning [57, 68] have been continuously proposed and verified to improve the performance of single-domain SR models. Amounts of CDR experiments have revealed that more recent and powerful single-domain SR models are able to outperform conventional CDSR algorithms [4, 13, 21, 66], which is also observed in our experiments (see Sec. 4.4). However, most existing CDSR models commonly incorporate simple SR models as their sequence encoders to verify their CDR strategies, or design complicated and customized networks to learn multi-domain interactions tailored to their specific cross-domain settings, overlooking whether the proposed algorithm possess the flexibility to adapt to different (current or future-updated) strong sequence encoders.

Different from them, we argue that the universality of a CDSR algorithm with different sequence encoders is a crucial factor in guaranteeing the sustained effectiveness of the proposed framework, since real-world systems always prefer simple and universal methods. In this work, we try our best to enhance the universality and maintain the simplification of TCA and TCL without customized designs in the single-domain sequence modeling (see 4.6 for detailed results). Therefore, Tri-CDR could leverage possible advancements in single-domain sequence modeling (including model evolution in the future), thereby extending the lifecycle of our proposed CDSR algorithms.

**3.8.4 Complexity Analyses.** In this section, we analyze the complexity of Tri-CDR and compare it with classical SR models (SASRec [20], CL4SRec [57]) and CDSR algorithms (DASL [22], C<sup>2</sup>DSR [3]). The space complexity of the proposed Tri-CDR is largely determined by its sequence encoder. For example, apart from the mandatory space allocation for training three SASRec models (for source, mixed and target domains respectively), Tri-CDR (SASRec) only requires a limited number of MLPs (mentioned in Sec. 3.4 and 3.5) to be additionally trained. That is, Tri-CDR does not introduce excessive trainable parameters, which renders the space complexity of Tri-CDR akin to its sequence encoder.

As for the time complexity in model training, the base sequence encoders serve as the pivotal component of the CDSR tasks, and dominate its time complexity. Since different sequence encoders have different time complexity, we provide the time complexity of the used sequence encoders as follows: SASRec [20]:  $\mathcal{O}((L^2 + L)d|U|)$ , CL4SRec [57]:  $\mathcal{O}((L^2 + L + B)d|U|)$ , where  $L$  denotes the length of the behavior sequence,  $|U|$  denotes the number of users,  $B$  is the size of mini-batch. Meanwhile, we also analyze the time complexity of the state-of-the-art CDSR algorithms which are armed with dual-attention mechanism (DASL [22]), graph representation learning and contrastive learning paradigm (C<sup>2</sup>DSR [3]). That is, the time complexity of DASL and C<sup>2</sup>DSR is  $\mathcal{O}((L^2 + L)d|U|)$

Table 3. Statistics of three CDR datasets with six domains.

Dataset	Amazon Toy & Game		Amazon Book & Movie		Douban Book & Music	
Domain	Toy	Game	Book	Movie	Book	Music
Users	7,996	7,996	28,531	28,531	4,580	4,580
Items	37,868	11,735	239,042	38,185	64,340	57,586
Records	114,487	82,871	625,692	349,918	224,117	278,855
Density	0.0378%	0.0883%	0.0092%	0.0321%	0.0761%	0.1057%

and  $O((L^2 + L + B + |\mathcal{R}^M| + |\mathcal{R}^T|)d|U|)$  respectively, where  $|\mathcal{R}^M|$  and  $|\mathcal{R}^T|$  denotes the number of nonzero entries in the Laplacian matrix in the mixed and target domain respectively. In contrast, we performed a comprehensive analysis of the time complexity of Tri-CDR on different sequence encoders, revealing that both Tri-CDR (SASRec) and Tri-CDR (CL4SRec) share the identical time complexity as  $O((L^2 + L + B)d|U|)$ .

The aforementioned analyses indicate that Tri-CDR has asymptotic similar time complexity with CL4SRec in training. It is worth noting that Tri-CDR does not conduct TCL in the inference phase. Therefore, its online inference item complexity (which is the central metric considered in practical usages of recommendation models rather than the training time complexity) is comparable with DASL in magnitude. Moreover, Tri-CDR exhibits a marginal enhancement in training efficiency in the authentic recommender systems with a more extensive user and item scale, while its online serving efficiency is still parallel with other CDR models. These complexity analyses underscore Tri-CDR’s scalability on large-scale real-world recommender systems. Therefore, the proposed Tri-CDR is able to employ cross-domain positive transfer and obtain high-quality sequence representations with only tolerant additional time and space cost.

## 4 EXPERIMENTS

In this section, we conduct extensive experiments to answer the following research questions: (RQ1): How does Tri-CDR perform against the state-of-the-art SR and CDSR baselines (see Sec. 4.4)? (RQ2): How do different components of Tri-CDR benefit its performance (see Sec. 4.5)? (RQ3): Is Tri-CDR still effective with other base sequence encoders (see Sec. 4.6)? (RQ4): Does the introduction of additional domains lead to reasonable performance improvements? (see Sec. 4.7)? (RQ5): How do some important hyper-parameters affect the performance of Tri-CDR (see Sec. 4.8)? (RQ6): How do the coarse-grained similarity modeling and fine-grained distinction modeling contribute to the multi-domain representation learning in CDR(see Sec. 4.9)?

### 4.1 Datasets

To verify the effectiveness and universality of Tri-CDR, we conduct extensive experiments on six real-world CDR settings with four domains from Amazon [36] and two domains from Douban [65]. Following [22, 66], we select *Amazon Toy & Game*, *Amazon Book & Movie* and *Douban Book & Music* to form six CDR tasks. Amazon Toy & Game, Amazon Book & Movie and Douban Book & Music include the review records from October 2000 to October 2018, from December 1996 to September 2018 and from July 2005 to December 2011 respectively. These three cross-domain datasets are pre-processed via the same method following classical CDR studies [21, 22]. Specifically, we build three user behavior sequences on the source, target, and mixed domains in chronological order. We set the last interacted item of each user as the test set and the penultimate interacted item as the valid set based on the Leave-one-out splitting method [42, 64]. We randomly select the users that have behaviors in both domains, filter out users having less than 3 behaviors, and treat all rating records as interacted behaviors [19, 38]. The detailed statistics are shown in Table 3.

## 4.2 Baselines

In this section, we compare Tri-CDR with three representative single-domain SR models and seven cross-domain SR models. Note that all baselines have been sufficiently tuned on different cross-domain settings to achieve their optimal performances. These compared baselines are elaborated as follows:

- **GRU4Rec.** GRU4Rec [16] is a classical session-based recommendation method encoding the item sequence via GRU.
- **SASRec.** SASRec [20] is a widely-used sequential recommendation model. It applies the self-attention mechanism to model behavior interactions in sequential recommendation.
- **CL4SRec.** CL4SRec [57] is the SOTA CL-based SR model, which adopts three sequence augmentation approaches to generate self-supervised signals via three intra-domain CL tasks.

Benefiting from the implementation of the self-attention mechanism, SASRec achieves a significant improvement relative to RNN-based algorithms in SR. For fair comparisons with Tri-CDR, we use SASRec as the base sequence encoder for all cross-domain SR baselines (noted as *Model+*). They also share the same features and historical behaviors:

- **EMCDR.** EMCDR [34] is the classical CDR method, which utilizes a multi-layer perceptron to imitate the nonlinear mapping function across source and target domains.
- **SASRec(S+T).** SASRec(S+T) is a straightforward CDSR method based on SASRec. It first generates the source and target sequence representations respectively, then concatenates and feeds them into an MLP for the final prediction.
- **CoNet+.** CoNet [18] introduces the cross connections and a joint objective function across source and target domains to model the dual knowledge transfer.
- **DTCDR+.** DTCDR [70] is a pioneer dual-target CDR method, which proposes an adaptable embedding sharing strategy to combine and share the user embeddings across domains based on the multi-task learning paradigm.
- **DDTCDR+.** DDTCDR [23] designs a latent orthogonal mapping function for extracting and transferring user preferences between two related domains while preserving cross-domain relations across different latent spaces in an iterative manner through a dual-transfer method.
- **DASL+.** DASL [22] constructs dual embeddings to extract the user's independent preferences and captures the user's cross-domain preference through a dual-attention learning mechanism. With its inherent structural superiority, DASL facilitates smooth incorporation of different sequence encoders. After replacing the sequence encoder from GRU to SASRec, DASL further achieves the promising performance, establishing itself as a strong baseline in CDSR.
- **C<sup>2</sup>DSR+.** C<sup>2</sup>DSR [3] is the SOTA model in CDSR, which leverages a effective graph neural network to exploit the inter-domain co-occurrence collaborative relationship and proposes an contrastive infomax objective to capture and transfer the user's cross-domain preferences via the mutual information maximization mechanism.

## 4.3 Experimental Settings

We implement the above methods using PyTorch with Python 3.8.10. Leveraging the empirical knowledge, we conduct a comprehensive grid search across all hyper-parameters and select the optimal parameters for Tri-CDR and each baseline. In the parameter sensitivity analyses, we delineate the nuanced variations in the performance trend of Tri-CDR concerning several pivotal parameters (see Sec.4.8). For fair comparisons, we take Adam as the optimizing method, and the learning rate and the batch size are set as 0.0005 and 120. We initialize model parameters randomly using the Xavier method. We conducted dimension experiments in {16, 32, 64, 128} on each baseline and report their optimal performance. For Tri-CDR, we define the embedding size as 64. We adopt the same maximum sequence length for each model as 200 on all datasets. Due to the inherent diversity in user behaviors' sparsity and data distributions, the range of hyper-parameters differs across different cross-domain

Table 4. Results on cross-domain recommendation on Amazon platform. All improvements are significant ( $p < 0.05$  with paired t-tests).

Datasets	Algorithms	N@5	N@10	N@20	N@50	HR@5	HR@10	HR@20	HR@50	AUC
Game ↓ Toy	GRU4Rec	0.1305	0.1551	0.1783	0.2172	0.1860	0.2624	0.3549	0.5533	0.5613
	SASRec	0.1853	0.2097	0.2321	0.2633	0.2505	0.3260	0.4149	0.5734	0.5821
	CL4SRec	0.1898	0.2143	0.2359	0.2673	0.2577	0.3338	0.4199	0.5795	0.5828
	EMCDR	0.1397	0.1705	0.1964	0.2322	0.2015	0.2972	0.4000	0.5819	0.5969
	CoNet+	0.1971	0.2208	0.2428	0.2739	0.2657	0.3392	0.4266	0.5838	0.5840
	SASRec(S+T)	0.1948	0.2189	0.2397	0.2715	0.2610	0.3355	0.4184	0.5793	0.5815
	DTCDR+	0.1963	0.2191	0.2404	0.2724	0.2646	0.3351	0.4196	0.5812	0.5777
	DDTCDR+	0.1959	0.2185	0.2404	0.2723	0.2617	0.3316	0.4186	0.5800	0.5768
	DASL+	0.1976	0.2206	0.2418	0.2740	0.2624	0.3338	0.4183	0.5810	0.5788
	C <sup>2</sup> DSR+	0.1964	0.2213	0.2433	0.2744	0.2671	0.3442	0.4314	0.5883	0.5878
Toy ↓ Game	Tri-CDR(SASRec)	<b>0.2069</b>	<b>0.2312</b>	<b>0.2528</b>	<b>0.2832</b>	<b>0.2797</b>	<b>0.3548</b>	<b>0.4405</b>	<b>0.5945</b>	<b>0.5913</b>
	Tri-CDR(CL4SRec)	0.2062	<b>0.2313</b>	0.2525	<b>0.2836</b>	0.2794	<b>0.3576</b>	<b>0.4418</b>	<b>0.5989</b>	0.5895
	GRU4Rec	0.2682	0.3053	0.3366	0.3726	0.3697	0.4839	0.6079	0.7894	0.7510
	SASRec	0.3350	0.3726	0.4003	0.4297	0.4465	0.5631	0.6723	0.8204	0.7868
	CL4SRec	0.3362	0.3720	0.4007	0.4296	0.4474	0.5582	0.6715	0.8166	0.7807
	EMCDR	0.2395	0.2797	0.3128	0.3520	0.3403	0.4649	0.5963	0.7996	0.7513
	CoNet+	0.3371	0.3725	0.4005	0.4293	0.4479	0.5573	0.6677	0.8127	0.7784
	SASRec(S+T)	0.3436	0.3803	0.4088	0.4380	0.4566	0.5702	0.6828	0.8295	0.7931
	DTCDR+	0.3315	0.3682	0.3964	0.4253	0.4423	0.5558	0.6676	0.8129	0.7803
	DDTCDR+	0.3355	0.3708	0.3989	0.4282	0.4473	0.5566	0.6677	0.8146	0.7832
Movie ↓ Book	DASL+	0.3368	0.3734	0.4010	0.4299	0.4482	0.5613	0.6706	0.8155	0.7819
	C <sup>2</sup> DSR+	0.3292	0.3691	0.3985	0.4277	0.4475	0.5704	0.6868	0.8342	0.7951
	Tri-CDR(SASRec)	0.3514	<b>0.3892</b>	<b>0.4182</b>	<b>0.4458</b>	<b>0.4684</b>	<b>0.5854</b>	<b>0.7000</b>	<b>0.8383</b>	<b>0.8015</b>
	Tri-CDR(CL4SRec)	<b>0.3562</b>	<b>0.3915</b>	<b>0.4205</b>	<b>0.4485</b>	<b>0.4712</b>	<b>0.5806</b>	<b>0.6954</b>	<b>0.8357</b>	<b>0.8004</b>
	GRU4Rec	0.2381	0.2695	0.2994	0.3421	0.3199	0.4173	0.5362	0.7521	0.7163
	SASRec	0.2863	0.3182	0.3475	0.3851	0.3754	0.4742	0.5903	0.7804	0.7453
	CL4SRec	0.3013	0.3340	0.3627	0.4001	0.3925	0.4937	0.6078	0.7971	0.7586
	EMCDR	0.2771	0.3114	0.3417	0.3801	0.3743	0.4807	0.6008	0.7950	0.7546
	CoNet+	0.3054	0.3381	0.3668	0.4028	0.3987	0.5001	0.6140	0.7959	0.7607
	SASRec(S+T)	0.2963	0.3285	0.3576	0.3950	0.3877	0.4874	0.6029	0.7922	0.7549
Book ↓ Movie	DTCDR+	0.3034	0.3353	0.3639	0.4012	0.3963	0.4952	0.6087	0.7968	0.7587
	DDTCDR+	0.3018	0.3339	0.3629	0.3995	0.3939	0.4935	0.6081	0.7937	0.7569
	DASL+	0.3027	0.3346	0.3635	0.4004	0.3946	0.4933	0.6080	0.7948	0.7576
	C <sup>2</sup> DSR+	0.3038	0.3359	0.3651	0.4015	0.3974	0.4966	0.6126	0.7967	0.7592
	Tri-CDR(SASRec)	0.3186	<b>0.3519</b>	<b>0.3811</b>	<b>0.4171</b>	<b>0.4152</b>	<b>0.5182</b>	<b>0.6339</b>	<b>0.8156</b>	<b>0.7752</b>
	Tri-CDR(CL4SRec)	<b>0.3210</b>	<b>0.3533</b>	<b>0.3815</b>	<b>0.4171</b>	0.4140	0.5142	0.6264	0.8063	0.7693
	GRU4Rec	0.4123	0.4477	0.4735	0.5002	0.5338	0.6432	0.7448	0.8795	0.8401
	SASRec	0.4492	0.4843	0.5096	0.5343	0.5712	0.6795	0.7796	0.9034	0.8629
	CL4SRec	0.4578	0.4930	0.5177	0.5417	0.5812	0.6897	0.7872	0.9078	0.8677
	EMCDR	0.3608	0.4041	0.4334	0.4612	0.5003	0.6337	0.7496	0.8888	0.8413
Book ↓ Movie	CoNet+	0.4561	0.4911	0.5157	0.5390	0.5844	0.6922	0.7893	0.9064	0.8671
	SASRec(S+T)	0.4594	0.4944	0.5191	0.5430	0.5834	0.6913	0.7888	0.9090	0.8687
	DTCDR+	0.4508	0.4864	0.5115	0.5357	0.5779	0.6878	0.7869	0.9082	0.8671
	DDTCDR+	0.4590	0.4938	0.5187	0.5424	0.5827	0.6898	0.7883	0.9070	0.8674
	DASL+	0.4591	0.4939	0.5189	0.5425	0.5827	0.6899	0.7885	0.9072	0.8676
	C <sup>2</sup> DSR+	0.4587	0.4945	0.5189	0.5424	0.5869	0.6973	0.7936	0.9117	0.8713
	Tri-CDR(SASRec)	<b>0.4669</b>	<b>0.5015</b>	<b>0.5258</b>	<b>0.5491</b>	<b>0.5933</b>	<b>0.6999</b>	<b>0.7960</b>	<b>0.9128</b>	<b>0.8729</b>
	Tri-CDR(CL4SRec)	0.4667	0.5005	0.5244	0.5478	0.5923	0.6966	0.7914	0.9089	0.8694

settings. We have observed that optimal hyper-parameters vary across diverse datasets due to their specific behavior patterns. Specifically, we select the  $\lambda_{CSM}$  and  $\lambda_{FDM}$  in  $\{0.1, 0.5, 1, 4, 10\}$ . For the Amazon dataset, which exhibits significant disparities in sparsity between the two domains, we define the ratio between  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$  as 1:1:1 for the sparser Amazon Toy and Amazon Book and 100:1:1 and 1000:1:1 for the denser Amazon Movie and Amazon Game respectively. In contrast, we define the ratio between  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$  as 1:1:1 for the Douban dataset with similar sparsity of two domains. The temperature coefficient  $\tau$  is set to be 0.1. According to the average natural distribution of the behavior data in the Amazon dataset, we set  $\gamma$  to 4.0 in the sparser Amazon Toy and Amazon Book and 0.5 in the denser Amazon Game and Amazon Movie. Regarding the Douban dataset, we define the  $\gamma$  as 0.5 and 0.05 for Douban Book and Douban Music, respectively. We conduct five runs with different random seeds and report the average results of all models.

Table 5. Results on cross-domain recommendation on Douban platform. All improvements are significant ( $p < 0.05$  with paired t-tests).

Datasets	Algorithms	N@5	N@10	N@20	N@50	HR@5	HR@10	HR@20	HR@50	AUC
Music ↓ Book	GRU4Rec	0.4643	0.4944	0.5175	0.5441	0.5815	0.6746	0.7659	0.8998	0.8583
	SASRec	0.5428	0.5714	0.5921	0.6123	0.6591	0.7474	0.8290	0.9304	0.8942
	CL4SRec	0.5536	0.5814	0.5999	0.6193	0.6720	0.7578	0.8308	0.9279	0.8950
	EMCDR	0.4785	0.5119	0.5343	0.5555	0.6096	0.7123	0.8008	0.9069	0.8704
	CoNet+	0.5487	0.5783	0.5980	0.6161	0.6729	<u>0.7644</u>	0.8421	0.9327	0.8991
	DTCDR+	0.5413	0.5699	0.5898	0.6100	0.6600	0.7482	0.8269	0.9284	0.8927
	SASRec(S+T)	0.5518	0.5813	0.6022	0.6206	0.6687	0.7598	0.8424	<u>0.9348</u>	0.8993
	DDTCDR+	0.5374	0.5671	0.5876	0.6074	0.6544	0.7461	0.8270	0.9268	0.8918
	DASL+	0.5501	0.5806	0.6015	0.6199	0.6664	0.7602	0.8427	<b>0.9350</b>	0.8993
	C2DSR	0.5523	0.5823	0.6024	0.6205	0.6717	0.7641	<u>0.8436</u>	<u>0.9346</u>	<u>0.8996</u>
	Tri-CDR(SASRec)	<b>0.5625</b>	<b>0.5924</b>	<b>0.6118</b>	<b>0.6287</b>	<b>0.6815</b>	<b>0.7737</b>	<b>0.8503</b>	<b>0.9350</b>	<b>0.9028</b>
	Tri-CDR(CL4SRec)	<u>0.5571</u>	<u>0.5860</u>	<u>0.6043</u>	<u>0.6232</u>	<u>0.6743</u>	0.7635	0.8358	0.9301	0.8973
Book ↓ Music	GRU4Rec	0.5030	0.5338	0.5547	0.5757	0.6310	0.7258	0.8085	0.9136	0.8774
	SASRec	0.5747	0.6035	0.6219	0.6389	0.6956	0.7842	0.8565	0.9414	0.9088
	CL4SRec	0.5560	0.5851	0.6032	0.6210	0.6892	0.7787	0.8501	0.9394	0.9048
	EMCDR	0.5080	0.5411	0.5616	0.5793	0.6458	0.7478	0.8286	0.9171	0.8842
	CoNet+	0.5752	0.6053	0.6235	0.6388	0.7047	0.7970	0.8688	0.9454	0.9142
	DTCDR+	0.5689	0.5973	0.6163	0.6331	0.6937	0.7813	0.8561	0.9401	0.9082
	SASRec(S+T)	0.5771	0.6062	0.6248	0.6409	0.7015	0.7913	0.8646	<u>0.9449</u>	0.9124
	DDTCDR+	0.5720	0.6005	0.6191	0.6360	0.6935	0.7814	0.8548	0.9398	0.9076
	DASL+	0.5788	0.6089	0.6272	0.6430	0.7012	0.7936	0.8658	0.9447	0.9130
	C2DSR	0.5806	0.6101	0.6284	0.6440	0.7043	<u>0.7949</u>	<u>0.8670</u>	<u>0.9449</u>	<u>0.9136</u>
	Tri-CDR(SASRec)	<b>0.5891</b>	<b>0.6177</b>	<b>0.6354</b>	<b>0.6505</b>	<b>0.7140</b>	<b>0.8019</b>	<b>0.8717</b>	<b>0.9472</b>	<b>0.9165</b>
	Tri-CDR(CL4SRec)	<u>0.5826</u>	<u>0.6112</u>	<u>0.6291</u>	<u>0.6447</u>	<u>0.7051</u>	0.7934	0.8640	0.9421	0.9124

#### 4.4 Performance Comparison on Cross-domain Sequential Recommendation (RQ1)

We conduct our experiments on six CDR tasks, adopting three typical evaluation metrics including NDCG@k (N@k), Hit Rate@k (HR@k), and AUC with different  $k = 5, 10, 20, 50$ . Following [20], we randomly sample 99 negative items for each positive instance in testing phase. Table 4 and Table 5 shows the results on the Amazon platform and Douban platform respectively, and we can observe that:

(1) Tri-CDR achieves the best performances on all metrics and datasets compared to all baselines, with the significance level  $p < 0.05$ . The NDCG@10 improvements of Tri-CDR over the best baseline are 4.52%/2.95%/5.18%/1.42% on Amazon Toy/Game/Book/Movie, and the HR@10 improvements are 3.89%/2.63%/4.25%/0.37% on four Amazon datasets consistently. Similarly, the NDCG@10 improvements of Tri-CDR over the best baseline are 1.73%/1.25% on Douban Book/Music, and the HR@10 improvements are 1.26%/0.88%. It indicates that our triple sequence learning is beneficial in CDR. Moreover, it also demonstrates that Tri-CDR can well model the correlations among users' triple behavior sequences, and successfully capture useful information related to the target-domain prediction from all domains.

(2) Tri-CDR outperforms all CDR baselines that also consider multi-domain behaviors. It confirms the significance of (a) explicit triple sequence learning with the mixed behavior sequence that contains the user's global interests, and (b) our TCL and TCA that could better model triple correlations among three domains and combine them into user representations. Comparing with SOTA CDSR models we can know that, DDTCDR and DASL focus on the dual transfer between source and target domains. They perform worse than Tri-CDR due to the lack of triple sequence learning. C<sup>2</sup>DSR also conducts contrastive infomax objectives to improve global-local dual correlations. However, it does not consider the fine-grained distinctions in domain correlation modeling and



triple cross-domain attention in multi-domain combination. Sometimes the infomax loss is not well functioned, which may be caused by its unreal negative sequences. Tri-CDR also outperforms different ablation versions besides baselines as shown in Sec. 4.5.

(3) Comparing the improvements among different domains, we find that Tri-CDR is more beneficial on Game→Toy(sparse) and Movie→Book(sparse) settings. It not only confirms the effectiveness of Tri-CDR on different data distributions, but also reflects that Tri-CDR functions well on relatively sparser target domains, where the positive knowledge transfer from the source domain should bring in more essential information as supplements, implying the practical usage of Tri-CDR.

(4) To validate the robustness and scalability of Tri-CDR, we construct four million-level cross-domain settings on the Douban platform. These four settings encompass two instances wherein the source domain boasted a million-level interaction dataset (Movie → Book and Movie → Music), along with two instances exhibiting million-level interaction data for the target domains (Book → Movie and Music → Movie). Subsequently, we conducted comparative experiments between Tri-CDR and the pivotal baselines, thereby illuminating these models' adaptive prowess in intricately large-scale settings. Experimental results demonstrate that the NDCG@10 improvements and the HR@10 improvements of Tri-CDR over the best baseline are 1.97%/1.45%/0.87%/1.42% and 1.55%/1.22%/0.97%/0.52% on four million-level cross-domain settings on the Douban platform respectively. The above performance comparison and the time complexity analysis mentioned in Sec.3.8.4 validate the model's potential within real-world large-scale recommendation systems.

(5) To confirm its universality with different application platforms, we have also evaluated Tri-CDR on the Douban dataset (Book & Music) besides Amazon datasets. In combination with the statistics of CDR datasets in Table 5, we discover that: (a) consistent with the conventional CDSR algorithms, Tri-CDR functions well on relatively sparser target domains by transferring the positive knowledge from the denser source domain. (b) In contrast to the cross-domain settings on the Amazon platform, Tri-CDR shows relatively similar improvements in the two cross-domain settings on Douban. This may be attributed to the fact that users have longer average interactions on Douban dataset (48.93 and 60.89 for Book and Music respectively). (c) The experimental results also demonstrate that some dual-modeling CDSR methods can obtain promising performance on the Douban dataset. However, Tri-CDR outperforms these algorithms on most metrics, which provides further evidence of Tri-CDR's significant role in alleviating cross-domain negative transfer and accurately modeling the correlations among the triple domains. We further conduct a universality analysis on Tri-CDR adopted with different base sequence encoders in Sec. 4.6.

#### 4.5 Ablation Study (RQ2)

In this section, we aim to find whether different components are effective in Tri-CDR. Thus we compare Tri-CDR with Tri-CDR w/o TCA&TCL, Tri-CDR w/o TCA and Tri-CDR w/o FDM to verify the benefits of TCA, TCL and FDM. In general, most components' improvements are significant (the average error range  $\leq 0.003$ ). In Table 6 we have:

(1) Tri-CDR w/o TCA performs significantly better than Tri-CDR w/o TCA&TCL, verifying the effectiveness of TCL. TCL takes full advantage of CL's alignment and uniformity [47, 60] and extends it to triple domains, maximizing the multi-domain mutual information while remaining necessary preference diversity in knowledge transfer.

(2) Tri-CDR further improves the results of Tri-CDR w/o TCA. Our TCA highlights the information related to the target-domain preference learned from the current target-domain historical behaviors as well as the user's global interests learned from mixed behaviors via cross-domain attention, which enables more positive knowledge transfer and is beneficial for CDR.

Table 6. Results on ablation study of Tri-CDR(SASRec).

Datasets	Models	N@5	N@10	N@20	N@50	HR@5	HR@10	HR@20	HR@50	AUC
Game ↓ Toy	Tri-CDR w/o TCA and TCL	0.2005	0.2239	0.2470	0.2778	0.2720	0.3446	0.4364	0.5925	0.5833
	Tri-CDR w/o TCA	0.2022	0.2274	0.2502	0.2807	0.2745	0.3522	<b>0.4426</b>	0.5971	<b>0.5925</b>
	Tri-CDR w/o FDM	0.2030	0.2280	0.2503	0.2814	0.2757	0.3530	0.4416	<b>0.5991</b>	0.5915
	Tri-CDR	<b>0.2069</b>	<b>0.2312</b>	<b>0.2528</b>	<b>0.2832</b>	<b>0.2797</b>	<b>0.3548</b>	0.4405	0.5945	0.5913
Toy ↓ Game	Tri-CDR w/o TCA and TCL	0.3258	0.3636	0.3937	0.4234	0.4408	0.5578	0.6770	0.8262	0.7866
	Tri-CDR w/o TCA	0.3509	0.3871	0.4152	0.4430	<b>0.4698</b>	0.5816	0.6928	0.8322	0.7949
	Tri-CDR w/o FDM	0.3485	0.3861	0.4153	0.4429	0.4673	0.5837	0.6990	0.8376	<b>0.8017</b>
	Tri-CDR	<b>0.3514</b>	<b>0.3892</b>	<b>0.4182</b>	<b>0.4458</b>	0.4684	<b>0.5854</b>	<b>0.7000</b>	<b>0.8383</b>	0.8015
Movie ↓ Book	Tri-CDR w/o TCA and TCL	0.2959	0.3279	0.3577	0.3952	0.3874	0.4866	0.6047	0.7945	0.7554
	Tri-CDR w/o TCA	0.3085	0.3399	0.3679	0.4048	0.3995	0.4967	0.6076	0.7941	0.7589
	Tri-CDR w/o FDM	0.3088	0.3415	0.3706	0.4071	0.4044	0.5057	0.6211	0.8055	0.7667
	Tri-CDR	<b>0.3186</b>	<b>0.3519</b>	<b>0.3811</b>	<b>0.4171</b>	<b>0.4152</b>	<b>0.5182</b>	<b>0.6339</b>	<b>0.8156</b>	<b>0.7752</b>
Book ↓ Movie	Tri-CDR w/o TCA and TCL	0.4379	0.4754	0.5008	0.5250	0.5706	0.6860	0.7865	0.9077	0.8658
	Tri-CDR w/o TCA	0.4611	0.4953	0.5196	0.5433	0.5849	0.6902	0.7864	0.9054	0.8661
	Tri-CDR w/o FDM	0.4505	0.4858	0.5106	0.5344	0.5830	0.6916	0.7895	0.9088	0.8683
	Tri-CDR	<b>0.4669</b>	<b>0.5015</b>	<b>0.5258</b>	<b>0.5491</b>	<b>0.5933</b>	<b>0.6999</b>	<b>0.7960</b>	<b>0.9128</b>	<b>0.8729</b>
Book ↓ Music	Tri-CDR w/o TCA and TCL	0.5798	0.6092	0.6269	0.6427	0.7067	0.7970	0.8670	0.9461	0.9139
	Tri-CDR w/o TCA	0.5868	0.6160	0.6339	0.6491	0.7099	0.7997	0.8703	0.9465	0.9152
	Tri-CDR w/o FDM	0.5836	0.6138	0.6313	0.6468	0.7084	0.8014	0.8704	<b>0.9480</b>	0.9155
	Tri-CDR	<b>0.5891</b>	<b>0.6177</b>	<b>0.6354</b>	<b>0.6505</b>	<b>0.7140</b>	<b>0.8019</b>	<b>0.8717</b>	0.9472	<b>0.9165</b>
Music ↓ Book	Tri-CDR w/o TCA and TCL	0.5532	0.5833	0.6031	0.6209	0.6721	0.7649	0.8433	0.9324	0.8991
	Tri-CDR w/o TCA	0.5576	0.5881	0.6073	0.6254	0.6754	0.7693	0.8452	0.9361	0.901
	Tri-CDR w/o FDM	0.5595	0.5902	0.6100	0.6271	0.6777	0.7728	<b>0.8509</b>	<b>0.9368</b>	<b>0.9028</b>
	Tri-CDR	<b>0.5625</b>	<b>0.5924</b>	<b>0.6118</b>	<b>0.6287</b>	<b>0.6815</b>	<b>0.7737</b>	0.8503	0.9350	<b>0.9028</b>

(3) Comparing Tri-CDR with and without FDM, we further demonstrate that the fine-grained distinction modeling in TCL is indispensable. FDM keeps the domain diversity and significantly brings in additional 1.60%/0.78% average NDCG@10/HR@10 improvements on six datasets. Besides, it not only helps Tri-CDR to learn better multi-domain sequence representations, but also makes the model training more stable with different parameters.

(4) We also conduct additional experiments to verify the effectiveness of the supervised loss of the source domain signals on the Toy  $\rightarrow$  Game and Game  $\rightarrow$  Toy setting. The experimental performance demonstrates that the improvements arising from the supervised source domain loss exhibit a positive correlation with the sparsity relationship between the source and target domains. Specifically, the source domain loss fails to yield any significant performance improvement on the Toy (Sparse)  $\rightarrow$  Game (Dense) setting, whereas the improvement becomes appreciable in the Game (Dense)  $\rightarrow$  Toy (Sparse) setting. This observation aligns with the underlying mechanism of cross-domain recommendation and the intuitive understanding of the supervised source domain loss. It is beneficial in cross-domain settings where the source domain is denser, leading to more informative user preference modeling. On the other hand, the sparser source domain may introduce bias in model optimization. The integration of TCA and TCL highlights the informative knowledge while ignoring the noisy information from the source domain, thereby validating the practical application of the proposed components of Tri-CDR to some extent.

#### 4.6 Universality of Tri-CDR (RQ3)

Tri-CDR is a model-agnostic framework. We further evaluate its universality on Amazon datasets based on GRU4Rec and CL4SRec. Fig. 3 illustrates Tri-CDR models on GRU4Rec and CL4SRec, and we can find that:

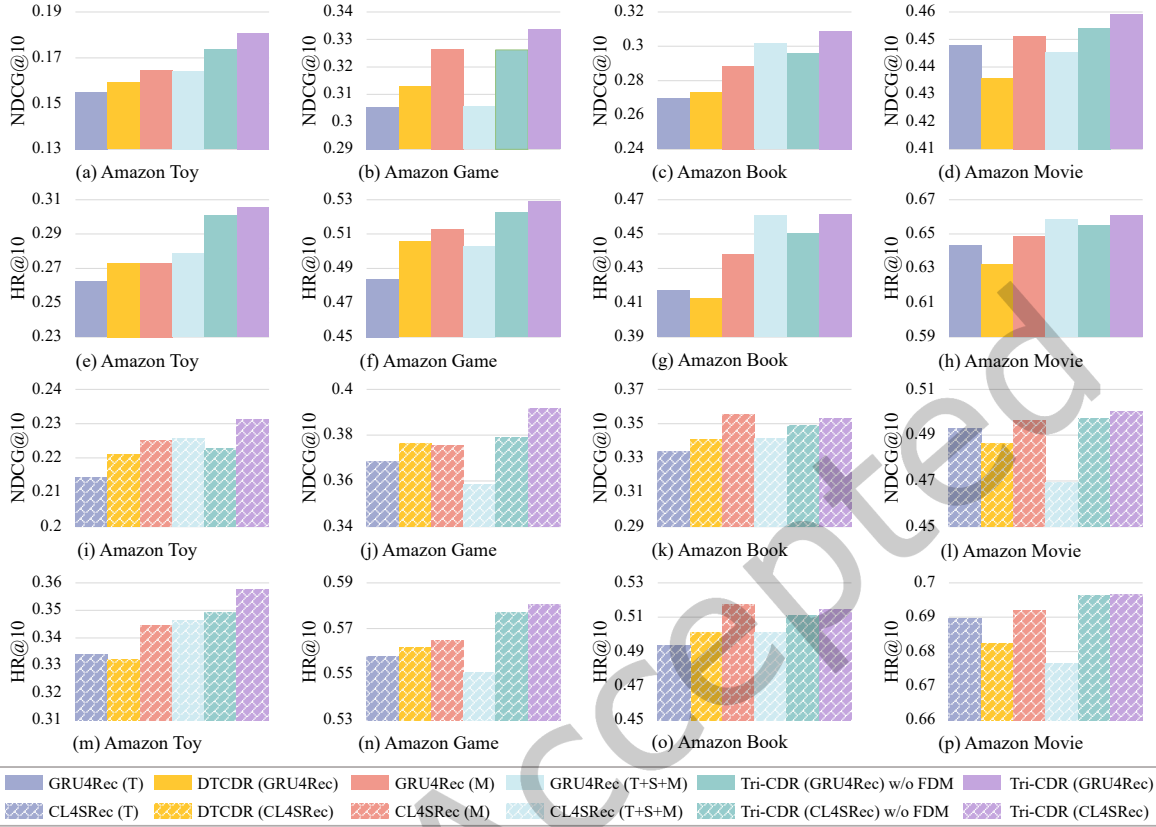


Fig. 3. Universality analysis on Tri-CDR. We show the results of different versions of Tri-CDR on GRU4Rec and CL4SRec.

(1) In general, Tri-CDR still achieves consistent and significant improvements over different ablation versions and the on all datasets with both GRU4Rec and CL4SRec, which confirms its universality when adopted with different sequential models and even other CL tasks. The improvements are consistent with other metrics.

(2) Comparing with different Tri-CDR's ablation versions, we reconfirm that (a) the mixed behavior sequence is informative, while directly combining source, target, and mixed sequences may also bring in noises, and (b) the proposed TCA and TCL with FDM are effective to make full use of information of three domains in CDR.

(3) In comparison to DTCDR, Tri-CDR accomplishes significant performance improvement on diverse sequence encoders, which highlights the practical application and the universal effectiveness of Tri-CDR. It is noteworthy that even the representative sequence encoders on the mixed domain consistently outperform DTCDR on most cross-domain settings. This reunderscores the critical importance of jointly modeling users' mixed domain behaviors in cross-domain recommendation.

(4) We should highlight that CL4SRec also conducts intra-domain CL tasks based on some sequence augmentations. The improvement brought by Tri-CDR implies that our inter-domain CL could cooperate well with various intra-domain CL. We notice that CL4SRec(M) has comparable or even better results over Tri-CDR on the Book domain after careful parameter selections. It is because that the intra-domain CL of CL4SRec on the mixed sequence (containing multi-domain behaviors) works as certain inter-domain CL tasks. Nevertheless, Tri-CDR still achieves the best results in general.

Table 7. Results on different domain utilization of Tri-CDR (SASRec).

Datasets	Models	N@5	N@10	N@20	N@50	HR@5	HR@10	HR@20	HR@50	AUC
Game ↓ Toy	SASRec (T)	0.1853	0.2097	0.2321	0.2633	0.2505	0.3260	0.4149	0.5734	0.5821
	SASRec (M)	0.1953	0.2200	0.2417	0.2744	0.2654	0.3420	0.4280	0.5941	<b>0.6017</b>
	SASRec (S+T)	0.1948	0.2189	0.2397	0.2715	0.2610	0.3355	0.4184	0.5793	0.5815
	SASRec (M+T)	0.1965	0.2210	0.2439	0.2745	0.2675	0.3434	0.4347	0.5893	0.5908
	SASRec (S+T+M)	0.2005	0.2239	0.2470	0.2778	0.2720	0.3446	0.4364	0.5925	0.5833
	Tri-CDR	<b>0.2069</b>	<b>0.2312</b>	<b>0.2528</b>	<b>0.2832</b>	<b>0.2797</b>	<b>0.3548</b>	<b>0.4405</b>	<b>0.5945</b>	0.5913
Toy ↓ Game	SASRec (T)	0.3304	0.3673	0.3962	0.4262	0.4405	0.5546	0.6691	0.8196	0.7841
	SASRec (M)	0.3408	0.3780	0.4064	0.4359	0.4553	0.5706	0.6826	0.8310	0.7966
	SASRec (S+T)	0.3436	0.3803	0.4088	0.4380	0.4566	0.5702	0.6828	0.8295	0.7931
	SASRec (M+T)	0.3191	0.3576	0.3880	0.4180	0.4359	0.5549	0.6752	0.8256	0.7878
	SASRec (S+T+M)	0.3305	0.3687	0.3988	0.4281	0.4463	0.5648	0.6836	0.8310	0.7904
	Tri-CDR	<b>0.3514</b>	<b>0.3892</b>	<b>0.4182</b>	<b>0.4458</b>	<b>0.4684</b>	<b>0.5854</b>	<b>0.7000</b>	<b>0.8383</b>	<b>0.8015</b>
Movie ↓ Book	SASRec (T)	0.2842	0.3160	0.3451	0.3839	0.3745	0.4728	0.5883	0.7844	0.7473
	SASRec (M)	0.2852	0.3183	0.3479	0.3868	0.3757	0.4783	0.5958	0.7925	0.7533
	SASRec (S+T)	0.2963	0.3285	0.3576	0.3950	0.3877	0.4874	0.6029	0.7922	0.7549
	SASRec (M+T)	0.3000	0.3326	0.3617	0.3985	0.3922	0.4930	0.6083	0.7948	0.7573
	SASRec (S+T+M)	0.2959	0.3279	0.3577	0.3952	0.3874	0.4866	0.6047	0.7945	0.7554
	Tri-CDR	<b>0.3186</b>	<b>0.3519</b>	<b>0.3811</b>	<b>0.4171</b>	<b>0.4152</b>	<b>0.5182</b>	<b>0.6339</b>	<b>0.8156</b>	<b>0.7752</b>
Book ↓ Movie	SASRec (T)	0.4492	0.4843	0.5096	0.5343	0.5712	0.6795	0.7796	0.9034	0.8629
	SASRec (M)	0.4639	0.4988	0.5234	0.5470	0.5873	0.6949	0.7924	0.9103	0.8704
	SASRec (S+T)	0.4594	0.4944	0.5191	0.5430	0.5834	0.6913	0.7888	0.9090	0.8687
	SASRec (M+T)	0.4486	0.4849	0.5100	0.5336	0.5809	0.6927	0.7918	0.9105	0.8693
	SASRec (S+T+M)	0.4469	0.4814	0.5062	0.5303	0.5749	0.6814	0.7792	0.9002	0.8611
	Tri-CDR	<b>0.4669</b>	<b>0.5015</b>	<b>0.5258</b>	<b>0.5491</b>	<b>0.5933</b>	<b>0.6999</b>	<b>0.7960</b>	<b>0.9128</b>	<b>0.8729</b>

#### 4.7 Analyses on the domain utilization of Tri-CDR (RQ4)

To inspect the effectiveness of different domains on Tri-CDR, we use S, T, and M to represent using source, target, and mixed sequences respectively, and compare Tri-CDR with five combinations. It is worth noting that SASRec (T) and SASRec (M) denote the single-domain SR models with the user's target behavior sequence and mixed behavior sequence. On the other hand, SASRec (S+T), SASRec (M+T), and SASRec (S+M+T) refer to the cross-domain SR models based on the user's source and target domain [22, 23, 33, 70], mixed and target domain [3, 66], and mixed, source and target domain (the proposed CDSR setting), respectively. From Table. 7, we can observe that:

(a) Comparing the first two versions, SASRec (M) significantly outperforms SASRec (T) on all metrics and datasets. It is reasonable since the mixed sequence is the complete user chronological behavior sequence from both the source and target domains. The sequential encoder on the mixed behavior sequence can be viewed as the cross-domain sequence encoder to some extent, yielding cross-domain information gain.

(b) Dual-modeling CDR methods (SASRec (S+T) and SASRec (M+T)) do not always achieve consistent improvement over the single-domain SR methods, and the simply triple-modeling CDR method (SASRec (S+T+M)) may even lead to further performance deterioration in some cross-domain settings. It further reinforces that both source and mixed sequences encompass cross-domain knowledge and noise information simultaneously and that simple dual- and triple-modeling strategies may be insufficient to accurately distinguish and model the complex correlations and confounding knowledge in multiple sequences.

(c) Comparing SASRec (S+T+M) and Tri-CDR on four cross-domain settings, we further demonstrate the effectiveness of the proposed TCA and TCL, which is primarily due to the fact that TCA highlights the users'

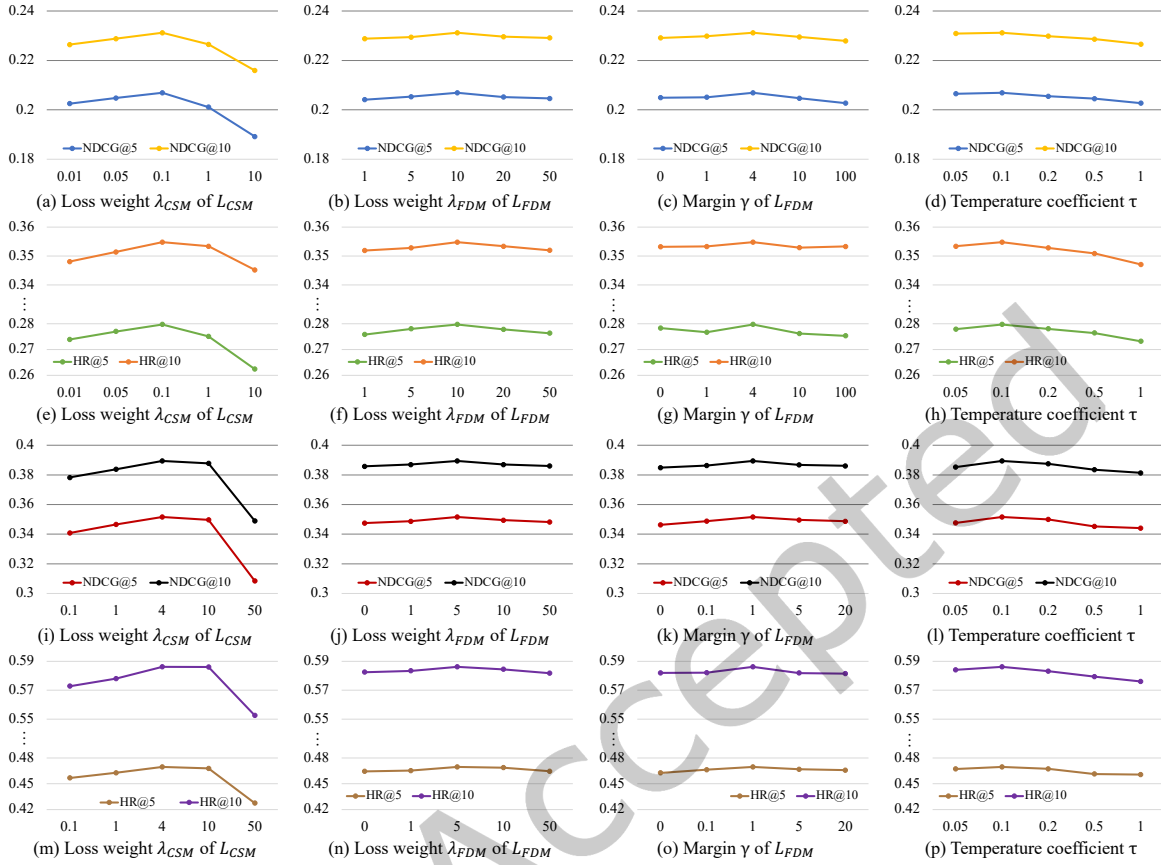


Fig. 4. Parameter analyses on loss weight  $\lambda_{CSM}$  of  $\mathcal{L}_{CSM}$ ,  $\lambda_{FDM}$  of  $\mathcal{L}_{FDM}$ , margin  $\gamma$  of  $\mathcal{L}_{FDM}$  and temperature coefficient  $\tau$ . (a)-(h) shows the results of Game→Toy, and (i)-(p) shows the results of Toy→Game.

target-domain preferences and comprehensive interests from the cross-domain knowledge transfer, and TCL precisely models the correlation among triple behavior sequences.

#### 4.8 Parameter Sensitivity Analyses (RQ5)

**4.8.1 Analyses on CL loss weights  $\lambda_{CSM}$ ,  $\lambda_{FDM}$ , margin  $\gamma$  and temperature coefficient  $\tau$ .** In Fig. 4, we conduct four parameter analyses on the Game→Toy setting (the first two rows) and the Toy→Game setting (the last two rows) to investigate the performance trends with different loss weight  $\lambda_{CSM}$  of  $\mathcal{L}_{CSM}$ , loss weights  $\lambda_{FDM}$  of  $\mathcal{L}_{FDM}$ , margin  $\gamma$  of  $\mathcal{L}_{FDM}$  and temperature coefficient  $\tau$  in Eq. (6) and Eq. (7). We observe that: (1) Tri-CDR's performance first increases and then decreases as  $\lambda_{CSM}$  and  $\tau$  gets larger. The length of behavioral sequences within the triple domain differs across diverse cross-domain settings, thereby leading to subtle distinctions in the interrelationships among triple domains. According to the behavior distribution characteristics, the loss weight  $\lambda_{CSM}$  of  $\mathcal{L}_{CSM}$  differs for different cross-domain settings.  $\lambda_{CSM} = 0.1/4.0$  achieves the best performance on Game→Toy and Toy→Game settings, respectively. In contrast, Tri-CDR is insensitive to the temperature coefficient  $\tau$ , so we set it as 0.1 for all CDSR settings with the purpose of simplifying the hyper-parameter tuning and obtaining the promising performance. (2) Too smaller  $\lambda_{FDM}$  may weaken the power of FDM in TCL, while too

Table 8. Parameter analyses on the ratio among loss weights  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$ 

Datasets	Models	Ratio among $\lambda_1, \lambda_2$ and $\lambda_3$	N@5	N@10	N@20	N@50	HR@5	HR@10	HR@20	HR@50	AUC
Game ↓ Toy	Tri-CDR w/o FDM	1:1:0.1	0.2036	0.2272	0.2496	0.2807	0.2770	0.3501	0.4387	0.5954	0.5867
		1:1:1	0.2030	0.2280	0.2503	0.2814	0.2757	0.3530	0.4416	<b>0.5991</b>	0.5915
		10:1:1	0.2016	0.2258	0.2481	0.2791	0.2730	0.3479	0.4362	0.5934	0.5861
		100:1:1	0.1992	0.2232	0.2453	0.2763	0.2682	0.3428	0.4307	0.5877	0.5835
		1000:1:1	0.1978	0.2223	0.2444	0.2754	0.2666	0.3425	0.4304	0.5873	0.5833
	Tri-CDR	1:1:0.1	<b>0.2072</b>	<b>0.2320</b>	<b>0.2538</b>	<b>0.2840</b>	<b>0.2803</b>	<b>0.3568</b>	<b>0.4433</b>	0.5961	<b>0.5916</b>
		1:1:1	0.2069	0.2312	0.2528	0.2832	0.2797	0.3548	0.4405	0.5945	0.5913
		10:1:1	0.2033	0.2275	0.2490	0.2793	0.2731	0.3479	0.4336	0.5868	0.5867
		100:1:1	0.2008	0.2246	0.2457	0.2769	0.2689	0.3429	0.4266	0.5844	0.5831
		1000:1:1	0.2003	0.2244	0.2461	0.2768	0.2676	0.3421	0.4282	0.5842	0.5832
Toy ↓ Game	Tri-CDR w/o FDM	1:1:1	0.3276	0.3664	0.3971	0.4268	0.4449	0.5647	0.6866	0.8354	0.7959
		10:1:1	0.3336	0.3722	0.4016	0.4309	0.4498	0.5688	0.6852	0.8324	0.7930
		100:1:1	0.3423	0.3801	0.4092	0.4381	0.4566	0.5737	0.6887	0.8337	0.7949
		1000:1:1	0.3485	0.3861	0.4153	0.4429	0.4673	0.5837	0.6990	0.8376	<b>0.8017</b>
		1000:1:0.1	0.3476	0.3845	0.4143	0.4424	0.4655	0.5795	0.6975	<b>0.8385</b>	0.8010
	Tri-CDR	1:1:1	0.3326	0.3708	0.4002	0.4302	0.4490	0.5671	0.6834	0.8339	0.7948
		10:1:1	0.3437	0.3809	0.4101	0.4384	0.4602	0.5752	0.6907	0.8328	0.7951
		100:1:1	0.3463	0.3842	0.4127	0.4406	0.4631	0.5807	0.6934	0.8330	0.7957
		1000:1:1	0.3514	<b>0.3892</b>	<b>0.4182</b>	<b>0.4458</b>	0.4684	<b>0.5854</b>	<b>0.7000</b>	0.8383	0.8015
		1000:1:0.1	<b>0.3503</b>	0.3877	0.4170	0.4444	<b>0.4687</b>	0.5843	0.6998	0.8375	0.8013

larger  $\lambda_{FDM}$  may also disturb the triple correlation learning in CSM. It is also natural that Tri-CDR is less sensitive to fine-grained  $\lambda_{FDM}$  compared to coarse-grained  $\lambda_{CSM}$ . The optimal performance of Tri-CDR is observed when  $\lambda_{FDM}$  is set to 10.0 and 5.0 on Game→Toy and Toy→Game settings. (3)  $\gamma = 0$  indicates that the model only wants the source-mixed distance to be smaller than the source-target distance. When  $\gamma = 100$ , Eq. (7) is always active to broaden the fine-grained cross-domain distance gap. Tri-CDR performs relatively poor under the extreme values of both sides, indicating the importance of an appropriate margin in FDM (4.0 and 1.0 for the Game→Toy and Toy→Game settings).

**4.8.2 Analyses on loss weights  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$  in CSM.** We implement a series of experiments to investigate the effect of different ratios among  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$  of  $\mathcal{L}_{CSM}$  in Tri-CDR. Table 8 shows the results of Tri-CDR(SASRec) with or without FDM on the Game→Toy and Toy→Game settings. We observe that: (1) for Tri-CDR w/o FDM, larger loss weights of  $\lambda_1$  lead to better performance on the Toy→Game setting (the relatively denser target domains). It is natural since the correlations between the source and mixed behavior sequences should be more highlighted to ensure learning informative source sequence representations from the sparser source behavior sequences. For the Game→Toy setting (the sparser target domains), we find that the loss weight ratio of 1:1:1 already achieves the best performance. (2) For Tri-CDR, we also observe the same trend on the Toy→Game setting that the performances increase as  $\lambda_1$  becomes larger. Moreover, regardless of the ratio among  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$ , Tri-CDR shows consistent improvements compared to Tri-CDR w/o FDM. Enhanced by FDM, the results are relatively satisfactory even with imperfect loss weights (average significant improvements of 1.17% and 0.51% on NDCG@10 and HR@10 respectively). It demonstrates that the proposed FDM helps to improve the effectiveness as well as the robustness of triple contrastive learning, making it less sensitive to different loss weights in CSM and more practical in real-world scenarios. (3) In comparing the different performances of Tri-CDR with different ratios between  $\lambda_2$  and  $\lambda_3$  in two cross-domain settings, we notice that Tri-CDR exhibited further improvements after reducing  $\lambda_3$  (loss weight of  $\mathcal{L}_{CSM}$  between source and target domains) on the Game→Toy setting. This improvement is interpretable as its underlying principle aligns with certain assumptions of FDM. Specifically,

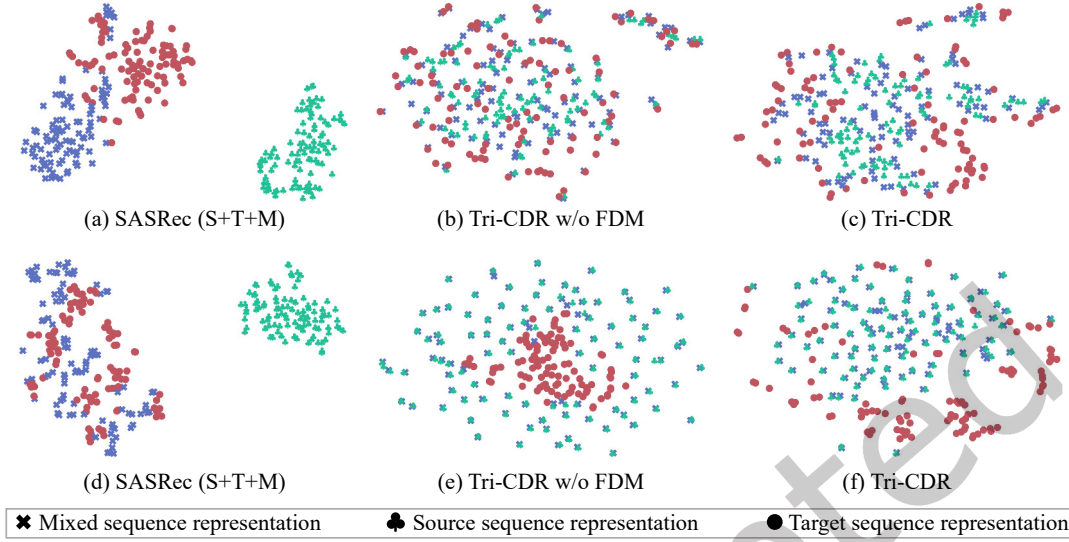


Fig. 5. Visualization of different Tri-CDR versions among triple domains, where the blue cross, green clubs, and red circle denote the mixed sequence representation, source sequence representation and target sequence representation respectively.

Tri-CDR precisely controls the similarity between the source and target domains by constraining the hyper-parameters, thereby modeling the fine-grained distinction among triple sequences. To mitigate the complexity of model training while preserving its scalability, we maintain the balanced ratio of  $\lambda_2$  and  $\lambda_3$  as 1:1 when presenting the performance of Tri-CDR. Experimental results also indicate that Tri-CDR achieves significant improvements over the current state-of-the-art baseline, even without fine-grained tuning of the ratio between  $\lambda_2$  and  $\lambda_3$ .

#### 4.9 Estimation of User Preference Modeling in Tri-CDR (RQ6)

To intuitively show the impacts of both similarity and distinction modeling in TCL, we show the visualization of randomly selected multi-domain sequence representations in SASRec (S+T+M), Tri-CDR w/o FDM, and Tri-CDR via t-Distributed Stochastic Neighbor Embedding (t-SNE) [46]. Fig. 5 and Fig. 6 illustrate the distributions of three domains and different users via different colors and shapes, respectively. In each figure, the first row refers to the visualization on the Game→Toy setting, while the second row represents the visualization on the Toy→Game setting.

**4.9.1 Visualization of the overall distribution across triple domains.** We depict the overall distribution of the randomly selected 100 users' sequential representations across different domains via t-SNE in two cross-domain settings, as illustrated in Fig. 5. We observe that: (1) As shown in Fig. 5 (a) and (d), most users' multi-domain sequence representations are naturally clustered via their domains rather than their users in SASRec (S+T+M). (2) Comparing Fig. 5 (b) and (e) with Fig. 5 (a) and (d), the multi-domain sequence representations are converted from domain-based clustering to the user-based, which indicates that the coarse-grained similarity modeling in TCL does make triple sequence representations of a user to be similar. (3) In contrast to Fig. 5 (e), the target domain sequence representations in Fig. 5 (f) do not aggregate in the approximate location but rather distributed around user-specific cross-domain information. This indicates that FDM enhances the discriminability of the target domain sequence representations in coarse-grained comparisons, enabling precise modeling of the users' target interests.

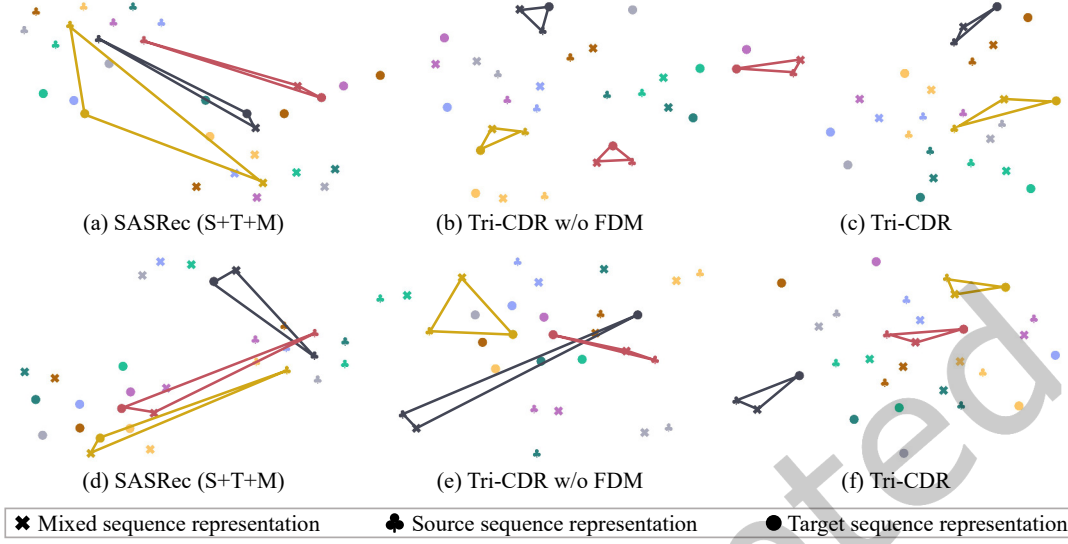


Fig. 6. Visualization of different Tri-CDR versions from the perspective of different users. We employ color differentials to distinguish different users while utilizing shapes to differentiate the same user’s sequential representations across triple domains.

**4.9.2 Visualization of the independent distribution among different users.** In order to investigate the independent distribution of different domain sequence representations among different users, we randomly select 10 users and visualized the aforementioned representations with t-SNE. The observations are as follows: (1) Similar to Fig. 5(a) and (d), the distances among the same user’s triple sequence representations are relatively large in 6(a) and (d). This leads to a failure in building adequate cross-domain correlations, and thus cannot make full use of additional source/mixed domains’ information. (2) In Fig. 6(b) and (e), irrespective of whether adjust the ratio hyper-parameters or not, the association among triple sequence representations of the same user is consistently improved with CSM. This proves the effectiveness of the information gain derived from coarse-grained similarity modeling. (3) In Fig. 6(b), some users’ multi-domain sequence representations are too close to form small acute triangles. Too homogeneous multi-domain representations may weaken the additional information gains from source/mixed sequences, which will harm the positive knowledge transfer. In contrast, armed with FDM, these users in Fig. 6(c) have more distinguishable sequence representations forming obtuse triangles, and thus Tri-CDR could achieve better results. (4) Despite the employment of sufficient tuning ratios among  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$  in CSM, Fig. 6(e) still exhibits unexpected deviations in the triple correlation. In contrast, the visualization of triangles in Fig. 6(f) provides tangible evidence of the practical significance of FDM.

## 5 CONCLUSION

In this work, we propose a model-agnostic Triple sequence learning for Cross-Domain Recommendation (Tri-CDR) framework. Conventional CDR methods mainly focus on modeling the dual-relations between the source and target domains or the mixed and target domains, failing to explore the triple correlation among the source, mixed, and target domains. Tri-CDR conducts a triple cross-domain attention method to highlight useful information and accelerate positive knowledge transfer and enables a more accurate multi-domain sequence representation learning strategy via both the coarse-grained similarity modeling and fine-grained distinction modeling. The extensive evaluation and analyses on six benchmark cross-domain settings demonstrate that Tri-CDR is able to



precisely model the similarity while preserving the information diversity among triple domains, which reveals the underlying principles of its effectiveness and universality. We believe that the triple sequence learning paradigm will provide a solid foundation for researchers and practitioners to explore new directions in cross-domain recommendation.

In the future, we will continue to explore the correlations among the source, target, and natural mixed behavior sequences, as well as more sophisticated modeling on their representation learning and multi-domain aggregation. We will also enhance Tri-CDR's capability by incorporating more modality information of item contents as semantic bridges in multi-domain recommendation.

## ACKNOWLEDGMENTS

This work is supported in part by the National Natural Science Foundation of China (Grant no. 62006141), the TaiShan Scholars Program (Grant no. tsqn202211289), the Excellent Youth Scholars Program of Shandong Province (Grant no. 2022HWYQ-048), the Oversea Innovation Team Project of the "20 Regulations for New Universities" funding program of Jinan (Grant no. 2021GXRC073). ChatGPT and Grammarly were utilized to improve grammar and correct spelling.

## REFERENCES

- [1] Keqin Bao, Jizhi Zhang, Yang Zhang, Wenjie Wang, Fuli Feng, and Xiangnan He. 2023. Tallrec: An effective and efficient tuning framework to align large language model with recommendation. *arXiv preprint arXiv:2305.00447* (2023).
- [2] Ye Bi, Liqiang Song, Mengqiu Yao, Zhenyu Wu, Jianming Wang, and Jing Xiao. 2020. A heterogeneous information network based cross domain insurance recommendation system for cold start users. In *Proceedings of International Conference on Research on Development in Information Retrieval (SIGIR)*.
- [3] Jiangxia Cao, Xin Cong, Jiawei Sheng, Tingwen Liu, and Bin Wang. 2022. Contrastive Cross-Domain Sequential Recommendation. In *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*.
- [4] Jiangxia Cao, Xixun Lin, Xin Cong, Jing Ya, Tingwen Liu, and Bin Wang. 2022. DisenCDR: Learning Disentangled Representations for Cross-Domain Recommendation. In *Proceedings of International Conference on Research on Development in Information Retrieval (SIGIR)*.
- [5] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. 2020. A simple framework for contrastive learning of visual representations. In *Proceedings of International Conference on Machine Learning (ICML)*.
- [6] Ziwei Fan, Zhiwei Liu, Yu Wang, Alice Wang, Zahra Nazari, Lei Zheng, Hao Peng, and Philip S Yu. 2022. Sequential Recommendation via Stochastic Self-Attention. In *Proceedings of International World Wide Web Conferences (WWW)*.
- [7] Jingtong Gao, Xiangyu Zhao, Bo Chen, Fan Yan, Huifeng Guo, and Ruiming Tang. 2023. AutoTransfer: Instance Transfer for Cross-Domain Recommendations. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*.
- [8] Yunfan Gao, Tao Sheng, Youlin Xiang, Yun Xiong, Haofen Wang, and Jiawei Zhang. 2023. Chat-rec: Towards interactive and explainable llms-augmented recommender system. *arXiv preprint arXiv:2303.14524* (2023).
- [9] John Giorgi, Osvald Nitski, Bo Wang, and Gary Bader. 2021. DeCLUTR: Deep Contrastive Learning for Unsupervised Textual Representations. In *Proceedings of Annual Meeting of the Association for Computational Linguistics (ACL)*.
- [10] 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. *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI)*.
- [11] 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* (2022).
- [12] Lei Guo, Jinyu Zhang, Li Tang, Tong Chen, Lei Zhu, and Hongzhi Yin. [n. d.]. Time interval-enhanced graph neural network for shared-account cross-domain sequential recommendation. *IEEE Transactions on Neural Networks and Learning Systems* ([n. d.]).
- [13] Xiaobo Guo, Shaoshuai Li, Naicheng Guo, Jiangxia Cao, Xiaolei Liu, Qiongxi Ma, Runsheng Gan, and Yunan Zhao. 2023. Disentangled Representations Learning for Multi-target Cross-domain Recommendation. *ACM Transactions on Information Systems (TOIS)* (2023).
- [14] Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, and Ross Girshick. 2020. Momentum contrast for unsupervised visual representation learning. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [15] Ruining He and Julian McAuley. 2016. Fusing similarity models with markov chains for sparse sequential recommendation. In *Proceedings of International Conference on Data Mining (ICDM)*.

- [16] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. 2016. Session-based recommendations with recurrent neural networks. In *Proceedings of International Conference on Learning Representations (ICLR)*.
- [17] Yupeng Hou, Shanlei Mu, Wayne Xin Zhao, Yaliang Li, Bolin Ding, and Ji-Rong Wen. 2022. Towards Universal Sequence Representation Learning for Recommender Systems. In *Proceedings of ACM Knowledge Discovery and Data Mining (SIGKDD)*.
- [18] Guangneng Hu, Yu Zhang, and Qiang Yang. 2018. Conet: Collaborative cross networks for cross-domain recommendation. In *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*.
- [19] SeongKu Kang, Junyoung Hwang, Dongha Lee, and Hwanjo Yu. 2019. Semi-supervised learning for cross-domain recommendation to cold-start users. In *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*.
- [20] Wang-Cheng Kang and Julian McAuley. 2018. Self-attentive sequential recommendation. In *Proceedings of International Conference on Data Mining (ICDM)*.
- [21] Chenglin Li, Mingjun Zhao, Huanming Zhang, Chenyun Yu, Lei Cheng, Guoqiang Shu, Beibei Kong, and Di Niu. 2022. RecGURU: Adversarial Learning of Generalized User Representations for Cross-Domain Recommendation. In *Proceedings of ACM International Conference on Web Search and Data Mining (WSDM)*.
- [22] Pan Li, Zhichao Jiang, Maofei Que, Yao Hu, and Alexander Tuzhilin. 2021. Dual Attentive Sequential Learning for Cross-Domain Click-Through Rate Prediction. In *Proceedings of ACM Knowledge Discovery and Data Mining (SIGKDD)*.
- [23] Pan Li and Alexander Tuzhilin. 2020. Ddtcd: Deep dual transfer cross domain recommendation. In *Proceedings of ACM International Conference on Web Search and Data Mining (WSDM)*.
- [24] Zihan Lin, Changxin Tian, Yupeng Hou, and Wayne Xin Zhao. 2022. Improving Graph Collaborative Filtering with Neighborhood-enriched Contrastive Learning. In *Proceedings of International World Wide Web Conferences (WWW)*.
- [25] Kang Liu, Feng Xue, Dan Guo, Le Wu, Shujie Li, and Richang Hong. 2023. Megcf: Multimodal entity graph collaborative filtering for personalized recommendation. *Proceedings of ACM Transactions on Information Systems (TOIS)* (2023).
- [26] Meng Liu, Jianjun Li, Guohui Li, and Peng Pan. 2020. Cross Domain Recommendation via Bi-Directional Transfer Graph Collaborative Filtering Networks. In *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*.
- [27] Weiming Liu, Jiajie Su, Chaochao Chen, and Xiaolin Zheng. 2021. Leveraging distribution alignment via stein path for cross-domain cold-start recommendation. *Proceedings of Annual Conference on Neural Information Processing Systems (NeurIPS)*.
- [28] Weiming Liu, Xiaolin Zheng, Chaochao Chen, Jiajie Su, Xinting Liao, Mengling Hu, and Yanchao Tan. 2023. Joint Internal Multi-Interest Exploration and External Domain Alignment for Cross Domain Sequential Recommendation. In *Proceedings of the ACM Web Conference 2023*.
- [29] Haokai Ma, Xiangxian Li, Lei Meng, and Xiangxu Meng. 2021. Comparative study of adversarial training methods for cold-start recommendation. In *Proceedings of ADVM*.
- [30] Haokai Ma, Zhuang Qi, Xinxin Dong, Xiangxian Li, Yuze Zheng, and Xiangxu Meng and Lei Meng. 2023. Cross-Modal Content Inference and Feature Enrichment for Cold-Start Recommendation. *Proceedings of IJCNN* (2023).
- [31] Haokai Ma, Ruobing Xie, Lei Meng, Xin Chen, Xu Zhang, Leyu Lin, and Jie Zhou. 2023. Exploring False Hard Negative Sample in Cross-Domain Recommendation. In *Proceedings of the 17th ACM Conference on Recommender Systems (RecSys '23)*.
- [32] Muiyang 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. *Proceedings of ACM Transactions on Knowledge Discovery from Data (TKDD)*.
- [33] Muiyang Ma, Pengjie Ren, Yujie Lin, Zhumin Chen, Jun Ma, and Maarten de Rijke. 2019.  $\pi$ -net: A parallel information-sharing network for shared-account cross-domain sequential recommendations. In *Proceedings of International Conference on Research on Development in Information Retrieval (SIGIR)*.
- [34] Tong Man, Huawei Shen, Xiaolong Jin, and Xueqi Cheng. 2017. Cross-domain recommendation: An embedding and mapping approach.. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI)*.
- [35] Lei Meng, Fuli Feng, Xiangnan He, Xiaoyan Gao, and Tat-Seng Chua. 2020. Heterogeneous fusion of semantic and collaborative information for visually-aware food recommendation. In *Proceedings of MM*.
- [36] Jianmo Ni, Jiacheng Li, and Julian McAuley. 2019. Justifying recommendations using distantly-labeled reviews and fine-grained aspects. In *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP)*.
- [37] Wentao Ouyang, Xiuwu Zhang, Lei Zhao, Jinmei Luo, Yu Zhang, Heng Zou, Zhaojie Liu, and Yanlong Du. 2020. Minet: Mixed interest network for cross-domain click-through rate prediction. In *Proceedings of ACM International Conference on Information and Knowledge Management (ACM International Conference on Information and Knowledge Management (CIKM))*.
- [38] Qi Pi, Guorui Zhou, Yujing Zhang, Zhe Wang, Lejian Ren, Ying Fan, Xiaoqiang Zhu, and Kun Gai. 2020. Search-based user interest modeling with lifelong sequential behavior data for click-through rate prediction. In *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*.
- [39] Ruihong Qiu, Zi Huang, Hongzhi Yin, and Zijian Wang. 2022. Contrastive learning for representation degeneration problem in sequential recommendation. In *Proceedings of ACM International Conference on Web Search and Data Mining (WSDM)*.

- [40] Ahmed Rashed, Shereen Elsayed, and Lars Schmidt-Thieme. 2022. Context and Attribute-Aware Sequential Recommendation via Cross-Attention. In *Proceedings of the 16th ACM Conference on Recommender Systems (Recsys)*.
- [41] Steffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme. 2010. Factorizing personalized markov chains for next-basket recommendation. In *Proceedings of International World Wide Web Conferences (WWW)*. 811–820.
- [42] Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. 2019. BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer. In *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*.
- [43] 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. *Proceedings of IEEE Transactions on Knowledge and Data Engineering (TKDE)*.
- [44] Jiaxi Tang and Ke Wang. 2018. Personalized top-n sequential recommendation via convolutional sequence embedding. In *Proceedings of ACM International Conference on Web Search and Data Mining (WSDM)*.
- [45] Zuoli Tang, Zhaoxin Huan, Zihao Li, Xiaolu Zhang, Jun Hu, Chilin Fu, Jun Zhou, and Chenliang Li. 2023. One Model for All: Large Language Models are Domain-Agnostic Recommendation Systems. *arXiv preprint arXiv:2310.14304* (2023).
- [46] Laurens Van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *Journal of machine learning research* (2008).
- [47] Tongzhou Wang and Phillip Isola. 2020. Understanding contrastive representation learning through alignment and uniformity on the hypersphere. In *Proceedings of International Conference on Machine Learning (ICML)*.
- [48] Tianxin Wang, Fuzhen Zhuang, Zhiqiang Zhang, Daixin Wang, Jun Zhou, and Qing He. 2021. Low-dimensional Alignment for Cross-Domain Recommendation. In *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*.
- [49] Yaqing Wang, Caili Guo, Yunfei Chu, Jenq-Neng Hwang, and Chunyan Feng. 2020. A cross-domain hierarchical recurrent model for personalized session-based recommendations. *Proceedings of Neurocomputing*.
- [50] Yiqi Wang, Chaozhao Li, Zheng Liu, Mingzheng Li, Jiliang Tang, Xing Xie, Lei Chen, and Philip S Yu. 2022. An Adaptive Graph Pre-training Framework for Localized Collaborative Filtering. *Proceedings of ACM Transactions on Information Systems (TOIS)* (2022).
- [51] Yinwei Wei, Xiang Wang, Qi Li, Liqiang Nie, Yan Li, Xuanping Li, and Tat-Seng Chua. 2021. Contrastive learning for cold-start recommendation. In *Proceedings of ACM International Conference on Multimedia (ACM MM)*.
- [52] Haolun Wu, Chen Ma, Yingxue Zhang, Xue Liu, Ruiming Tang, and Mark Coates. 2022. Adapting Triplet Importance of Implicit Feedback for Personalized Recommendation. *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*.
- [53] Yiqing Wu, Ruobing Xie, Yongchun Zhu, Xiang Ao, Xin Chen, Xu Zhang, Fuzhen Zhuang, Leyu Lin, and Qing He. 2022. Multi-view Multi-behavior Contrastive Learning in Recommendation. In *Proceedings of Database Systems for Advanced Applications (DASFAA)*.
- [54] Zhuofeng Wu, Sinong Wang, Jiatao Gu, Madian Khabsa, Fei Sun, and Hao Ma. 2020. Clear: Contrastive learning for sentence representation. *arXiv preprint arXiv:2012.15466*.
- [55] Ruobing Xie, Qi Liu, Liangdong Wang, Shukai Liu, Bo Zhang, and Leyu Lin. 2022. Contrastive cross-domain recommendation in matching. In *Proceedings of ACM Knowledge Discovery and Data Mining (SIGKDD)*.
- [56] Ruobing Xie, Zhijie Qiu, Jun Rao, Yi Liu, Bo Zhang, and Leyu Lin. 2020. Internal and Contextual Attention Network for Cold-start Multi-channel Matching in Recommendation. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI)*.
- [57] Xu Xie, Fei Sun, Zhaoyang Liu, Shiwen Wu, Jinyang Gao, Jiandong Zhang, Bolin Ding, and Bin Cui. 2022. Contrastive learning for sequential recommendation. In *Proceedings of IEEE International Conference on Data Engineering (ICDE)*.
- [58] Yueqi Xie, Peilin Zhou, and Sunghun Kim. 2022. Decoupled side information fusion for sequential recommendation. In *Proceedings of International Conference on Research on Development in Information Retrieval (SIGIR)*.
- [59] Lyuxin Xue, Deqing Yang, Shuoyao Zhai, Yuxin Li, and Yanghua Xiao. 2023. Learning Dual-view User Representations for Enhanced Sequential Recommendation. *Proceedings of ACM Transactions on Information Systems (TOIS)* (2023).
- [60] Junliang Yu, Hongzhi Yin, Xin Xia, Tong Chen, Lizhen Cui, and Quoc Viet Hung Nguyen. 2022. Are graph augmentations necessary? simple graph contrastive learning for recommendation. In *Proceedings of International Conference on Research on Development in Information Retrieval (SIGIR)*.
- [61] Shengyu Zhang, Dong Yao, Zhou Zhao, Tat-Seng Chua, and Fei Wu. 2021. Causerec: Counterfactual user sequence synthesis for sequential recommendation. In *Proceedings of International Conference on Research on Development in Information Retrieval (SIGIR)*.
- [62] Wei Zhang, Pengye Zhang, Bo Zhang, Xingxing Wang, and Dong Wang. 2023. A Collaborative Transfer Learning Framework for Cross-domain Recommendation. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*.
- [63] Cheng Zhao, Chenliang Li, and Cong Fu. 2019. Cross-domain recommendation via preference propagation graphnet. In *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*.
- [64] Wayne Xin Zhao, Junhua Chen, Pengfei Wang, Qi Gu, and Ji-Rong Wen. 2020. Revisiting Alternative Experimental Settings for Evaluating Top-N Item Recommendation Algorithms. In *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*.
- [65] Wayne Xin Zhao, Shanlei Mu, Yupeng Hou, Zihan Lin, Kaiyuan Li, Yushuo Chen, Yujie Lu, Hui Wang, Changxin Tian, Xingyu Pan, Yingqian Min, Zhichao Feng, Xinyan Fan, Xu Chen, Pengfei Wang, Wendi Ji, Yaliang Li, Xiaoling Wang, and Ji-Rong Wen. 2021.

- Recbole: Towards a unified, comprehensive and efficient framework for recommendation algorithms. In *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*.
- [66] Xiaolin Zheng, Jiajie Su, Weiming Liu, and Chaochao Chen. 2022. DDGHM: Dual Dynamic Graph with Hybrid Metric Training for Cross-Domain Sequential Recommendation. In *Proceedings of ACM International Conference on Multimedia (ACM MM)*.
- [67] 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 ACM Knowledge Discovery and Data Mining (SIGKDD)*.
- [68] Kun Zhou, Hui Wang, Wayne Xin Zhao, Yutao Zhu, Sirui Wang, Fuzheng Zhang, Zhongyuan Wang, and Ji-Rong Wen. 2020. S3-rec: Self-supervised learning for sequential recommendation with mutual information maximization. In *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*.
- [69] Yuanhang Zhou, Kun Zhou, Wayne Xin Zhao, Cheng Wang, Peng Jiang, and He Hu. 2022. C<sup>2</sup>-CRS: Coarse-to-Fine Contrastive Learning for Conversational Recommender System. In *Proceedings of ACM International Conference on Web Search and Data Mining (WSDM)*.
- [70] Feng Zhu, Chaochao Chen, Yan Wang, Guanfeng Liu, and Xiaolin Zheng. 2019. Dtdr: A framework for dual-target cross-domain recommendation. In *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*.
- [71] 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 International Conference on Research on Development in Information Retrieval (SIGIR)*.
- [72] 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 ACM International Conference on Web Search and Data Mining (WSDM)*.