This article has been accepted for publication in IEEE Transactions on Multimedia. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TMM.2024.3374577

Inter- and Intra-Domain Potential User Preferences for Cross-Domain Recommendation

Jing Liu, *Member, IEEE,* Lele Sun, Weizhi Nie, *Member, IEEE,* Yuting Su, *Member, IEEE,* Yongdong Zhang, *Senior Member, IEEE,* Anan Liu*, *Senior Member, IEEE,*

***Abstract*—Data sparsity poses a persistent challenge in Recommender Systems (RS), driving the emergence of Cross-Domain Recommendation (CDR) as a potential remedy. However, most existing CDR methods often struggle to circumvent the transfer of domain-specific information, which are perceived as noise in the target domain. Additionally, they primarily concentrate on inter-domain information transfer, disregarding the comprehensive exploration of data within intra-domains. To address these limitations, we propose SUCCDR (Separating User features with Compound samples), a novel approach that tackles data sparsity by leveraging both cross-domain knowledge transfer and comprehensive intra-domain analysis. Specifically, to ensure the exclusion of noisy domain-specific features during the transfer process, user preferences are separated into domain-invariant and domain-specific features through three efficient constraints. Furthermore, the unobserved items are leveraged to generate compound samples that intelligently merge observed and unobserved potential user-item interaction, utilizing a simple yet efficient attention mechanism to enable a comprehensive and unbiased representation of user preferences. We evaluate the performance of SUCCDR on two real-world datasets, Douban and Amazon, and compare it with state-of-the-art single-domain and cross-domain recommendation methods. The experimental results demonstrate that SUCCDR outperforms existing approaches, highlighting its ability to effectively alleviate data sparsity problem.**

***Index Terms*—Cross-Domain Recommendation, Transfer Learning, Attention Mechanism.**

I. Introduction

THE rapid development of the Internet inevitably brings the problem of information overload to people's life. And recommender systems (RS) were born to effectively solve the problem by quickly matching potential items of interest to users in a large information flow. At present, recommender systems have been widely used in various application scenarios, such as TikTok (Short Videos), Douban (Movies), and Amazon (E-commerce). However, as the user-item interactions become increasingly sparse, the performance drops dramatically [1], which means existing recommendation approaches are still significantly hindered by the inherent issue of data sparsity [2]. Fortunately, the same user in real life tends to have consistent preferences in different domains to some extend.

This work is supported in part by the National Key Research and Development Program of China (2021YFF0901600)

J. Liu, L. Sun, W. Nie, Y. Su, A.-A. Liu are with the School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China. A.-A. Liu is also with the Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei 230088, China. Y. Zhang is with the University of Science and Technology of China, Hefei, China.

Corresponding author: Anan Liu (e-mail: anan0422@gmail.com)

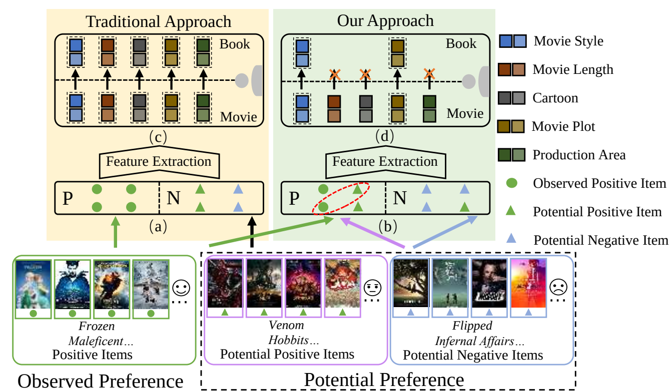


Fig. 1. Comparison of our SUCCDR with traditional CDR methods in terms of inter-domain and intra-domain potential preferences. (a) Unobserved interaction as negative items. (b) The utilization of potential intra-domain preferences with the proposed compound samples. (c) Total knowledge transfer. (d) Inter-domain knowledge transfer with domain-invariant features.

So the cross-domain recommendation (CDR) was proposed to alleviate the data sparsity problem [3]–[5] by introducing the potential inter-domain user preference . As shown in Fig. 1, Alex likes magic movies such as *Frozen* and *Maleficent*. Therefore, we can transfer Alex's movie preferences to the book domain and thus recommending more magic books to Alex. Vice versa, transferring Alex's book preferences can also help the recommendation of movies.

To facilitate knowledge transfer, CoNet [6] proposed cross-connections network to continuously integrate dual-domain features. DDTCDR [7] and DML [8], on the other hand, leveraged orthogonal mapping functions to achieve knowledge transfer between domains. GA-DTCDR [9] utilized attention mechanisms to accomplish knowledge transfer across domains. However, these approaches made the strong assumption that users' interests remain the same across domains so that they primarily focus on generating user representations for direct knowledge transfer. Unfortunately, it is overlooked that users may have interest discrepancy in different domains that is non-transferable. We call this portion of interest as domain-specific preferences. While domain-invariant features represent user preferences that are shared across domains, domain-specific features are only relevant to the source domain and can be considered as noise or even negatively impact the target domain. Consequently, it becomes crucial to distinguish only the domain-invariant features while disregarding the domain-specific features during the knowledge transfer process. As shown in Fig. 1 (d), the preference of story style and movie plot is good for recommending a book whereas the preferences such as movie duration, and production region are irrelevant and expected to be discarded from knowledge transfer.

Recently, some approaches have considered extracting and transferring the domain-invariant features. ATLRec [10] utilized MLPs to extract both domain-invariant and domain-specific features, without imposing significant constraints on them. DisenCDR [11] utilized variational inference techniques and amplified the Kullback-Leibler (KL) divergence to separate the two features from each other. MADD [12], being one of the most advanced cross-domain recommendation methods, leveraged an orthogonal loss to regularize the separation process. However, we argue that solely relying on orthogonal loss constraints or drawn-out distance constraints cannot guarantee full extraction of domain-invariant features from the highly entangled features. In the proposed SUCCDR, the separation of domain-specific features and domain-invariant features is accomplished by three well-designed constraints: separation constraint and peeling constraint enforce the separation of both features, and similarity constraint impels the efficient transfer of domain-invariant features.

As discussed above, existing CDR approaches primarily prioritize information transfer within inter-domains, paying little attention to exploring a more effective and comprehensive understanding of intra-domain user preferences. Actually, the challenges of generating an efficient and transferable user preference within each domain persists and should be regarded equally important as the inter-domain preference transfer. In terms of intra-domain modeling, most existing methods [13], [14] follow traditional RSs to take observed items as positive samples and unobserved items as negative samples for preferences learning. However, as CDR is introduced to eliminate the dilemma of data sparsity in a single domain, the extracted intra-domain user preference in this scenario is highly likely to be incomplete. More comprehensive user preferences can be captured as the number of user-interacted items increases. Fig. 2 validates this commonsense where an obvious improvement in the recommendation performance is observed as the number of interactions to learn a recommender increases.

This intuitive conclusion motivates us to take advantage of potential positive items that users might be interested in from a vast pool of uninteracted items, distinct from the conventional strategy of considering all unobserved items as negative samples. The potential user-item interactions can effectively enrich the items used to learn user preference, facilitating a more unbiased and comprehensive representation of user preferences. As a result, the performance of the model is expected to improve. For example, in Fig. 1, Alex likes magic movies but has no interaction with *Venom*. Based on the interaction histories of *Venom* with other users and its property of magic, Alex will probably also like *Venom*, which is obviously unwise to be taken as a negative sample. To investigate the potential positive items, in the proposed SUCCDR, we stochastically combine one observed item with a random number of items from co-liked items to generate one positive compound sample. With a simple yet efficient attention mechanism, the potential positive items are utilized as positive samples to alleviate the data sparsity problem.

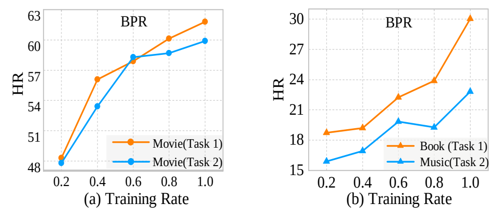


Fig. 2. The HR@10 performance of the BPR [15] model is evaluated on the Douban dataset (Task 1 and Task 2 in Table I), where different proportions of randomly selected user-item interactions from the training set are used as positive samples for training and the test set is same for each dataset.

Our main contributions are summarized as follows:

- We propose SUCCDR which focuses on investigating potential user preferences from both inter-domain and intra-domain perspectives.
- User preferences are separated into domain-invariant and domain-specific features with the former being transferred between different domains and the latter being excluded.
- We innovatively propose compound samples that contain a random number of items to take the advantage of potential positive items within each domains. Besides, a simple yet efficient attention mechanism is proposed for compound feature fusion.
- Extensive experiments are carried out on Douban and Amazon datasets. The results show that our model outperforms several state-of-the-art single-domain and cross-domain recommendation models.

The rest of this paper is organized as follows. In Sec.2, we introduce the related works on CDR and feature separation. Sec.3 presents the details of the proposed method. In Sec.4, we present the experiments to demonstrate the effectiveness of the model. Finally, we conclude this paper in Sec.5.

II. Related Work

A. Cross-Domain Recommendation

CDR is proposed to alleviate the data sparsity problem in single user-item interaction domain. Existing CDR works can be boardly classified into two branches: unidirectional CDR and bidirectional CDR. Unidirectional CDR emphasize the transfer of valuable information, such as textual information [16]–[18], user social relationships [19], and item attributes [20], from the source domain to the target domain. This transfer strategies serves to address challenges such as the cold start problem [21]–[23] and data sparsity problems in the target domain.

Bidirectional CDR [24], [25] has gained significant attention in recent years, drawing the inspiration from unidirectional CDR. The objective of Bidirectional CDR is to enhance the recommendation efficiency simultaneously in both two domains by transferring relevant information to each other. To achieve this objective, researchers have primarily focused on developing directed or undirected pairwise transfer strategies. Several approaches, such as CMF, CDTF, and CDFM [26]–[28], utilized Matrix Factorization (MF) to leverage interactions in the source domain as valuable information. However, these methods encounter challenges in capturing highly non-linear user-item interaction patterns. As a result, researchers have proposed models based on neural networks to address this limitation. CoNet [6] adopted a cross-connections network between Multi-Layer Perceptrons (MLPs) to facilitate knowledge transfer. In DARec [29], the first step is to learn users' rating patterns within each domain, and subsequently transfer these rating patterns to other domains. GA-DTCDR [9] utilized element-wise attention mechanisms to transfer user features across domains. DDTCDR and DML [7], [8] employed a latent orthogonal mapping function to transfer user preferences. BITGCF [30], a method based on graph neural networks, effectively incorporated a fusion layer to integrate cross-domain features during the process of graph convolution. This integration showcases promising performance results. Our work also focuses on bidirectional knowledge transfer. In contrast to these methods that transfer all information as a whole, we separate user features into domain-invariant features and domain-specific features, and only transfer the domain-invariant features that are shared across domains.

B. Feature Separation in Recommendation

Feature separation is a technique that aims to learn distinct representations of different factors from available interaction data. This approach has proven to be highly effective in improving the accuracy of recommendation systems within a single domain [31]. Building upon this success, recent research has extended the concept of feature separation to CDR. By incorporating feature separation, these studies have demonstrated its potential in enhancing the performance of recommendation models. ATLRec [10] utilized two separate Multi-Layer Perceptron (MLP) layers to extract both domain-invariant features and domain-specific features unique to each domain from the user's original representations. CDAML [32] employed a similar operation. Apart from that, there are no constraints imposed on the two features within the domain. DisenCDR [11] introduced variational inference and amplified the Kullback-Leibler (KL) divergence to distinguish between domain-invariant features and domain-specific features. MADD [12], a state-of-the-art Cross-Domain Recommendation (CDR) method, enhanced the separation of domain-specific features and domain-invariant features by applying an orthogonal loss after feature extraction. Nonetheless, we argue that relying solely on orthogonality constraints or feature distance operations does not guarantee the effective separation and enhanced information of domain-invariant and domain-specific representations. This limitation often results in sub-optimal recommendation performance. In contrast, our proposed model incorporates multiple constraints to maximize the transferable information in domain-invariant features as well as excluding the transferable information from the domain-specific features.

III. The Proposed Model

Fig. 3 mainly sketches the architecture of the proposed SUCCDR model, including the compound sample module, feature separation module, and constraint module. For clarity, we only illustrate the paradigm in domain A and the paradigm in domain B can be easily inferred accordingly. In this section, we first describe the notations used in this paper and then introduce the details of each model component.

A. Notations and Problem Definition

We consider a general CDR scenario where users in different domains are common. $\mathcal{D}^A$ and $\mathcal{D}^B$ are used to denote two domains. $\mathcal{U}$ denotes the common user set. $\mathcal{V}^A$ and $\mathcal{V}^B$ denote item set of $\mathcal{D}^A$ and $\mathcal{D}^B$, respectively. Two binary interaction matrices $Y^A \in \{0,1\}^{|\mathcal{U}|\times|\mathcal{V}^A|}$ and $Y^B \in \{0,1\}^{|\mathcal{U}|\times|\mathcal{V}^B|}$ are used to represent implicit user feedback, where the element r_{ui} is 1 if user u is observed to have interacted with item i and 0 otherwise (i.e., unobserved). $I \in \mathbb{R}^{|\mathcal{V}|\times d}$ denotes the item embedding, where d is the feature dimension. The overall user preferences exhibited in each single domain, which is named as domain-holistic features, are denoted as $U^{HA} \in \mathbb{R}^{|\mathcal{U}|\times d}$ and $U^{HB} \in \mathbb{R}^{|\mathcal{U}|\times d}$, respectively. Noticing that a partial of user preference remains invariant across different domains, in this work, we additionally separate user preferences into domain-invariant ones and domain-specific ones. Specifically, $U^I \in \mathbb{R}^{|\mathcal{U}|\times d}$ denoted the domain-invariant user preferences while $U^{PA} \in \mathbb{R}^{|\mathcal{U}|\times d}$ and $U^{PB} \in \mathbb{R}^{|\mathcal{U}|\times d}$ denote domain-A-specific and domain-B-specific user preferences, respectively.

B. Compound Samples for Intra-domains Potential Preferences

Most existing methods utilize observed items as positive samples and randomly select some unobserved items as negative samples to train the model. However, in scenarios with sparse user-item interaction, this setting often tends to learn largely incomplete or biased user preferences due to the presence of numerous potential positive samples hidden among the unobserved items. To tackle this issue, we discard the conventional sampling strategy and propose a novel compound sampling strategy where each compound sample consists of a random number of items within each domain. In particular, the positive compound samples involve both observed items and unobserved items. We further employ a simple yet efficient attention mechanism to identify the potentially liked items among the unobserved items and incorporate them with the observed item to supplement user preferences from the intra-domain perspective.

Compound Sampling Generation. As a basic hypothesis of the collaborative filtering, users with similar preferences are assumed to like common items [2], [33]. Motivated by this, we construct the co-liked item set for each user, which is regarded as the possibly liked item set so as to investigate user's potential preferences. As illustrated in compound sample module in Fig. 3, the target user in black is observed to interact with a few items in green circles. The subsequent task entails identifying similar users (depicted as black circles) who have also interacted with these observed items. The co-liked item set is then defined as a collection of all items (triangles) that have been interacted with by these similar users.

Afterward, we stochastically select an observed item and a random number of items from the co-liked item set to generate

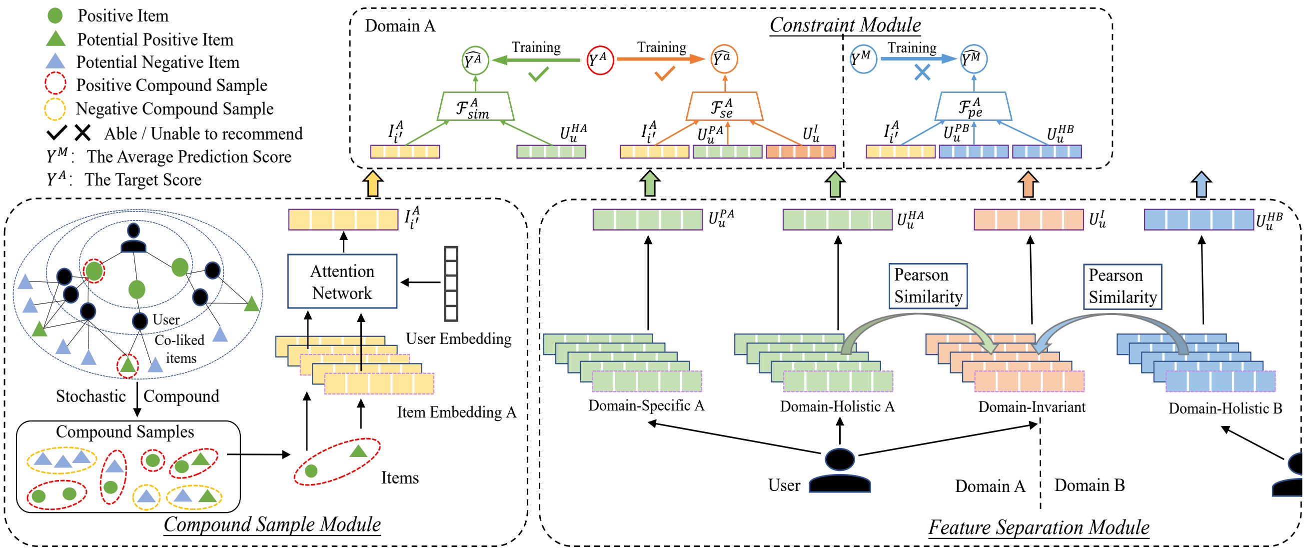


Fig. 3. The paradigm of SUCCDR, which contains the compound sample module, feature separation module, and constraint module. It mainly describes the paradigm in domain A and the paradigm in domain B can be easily inferred accordingly.

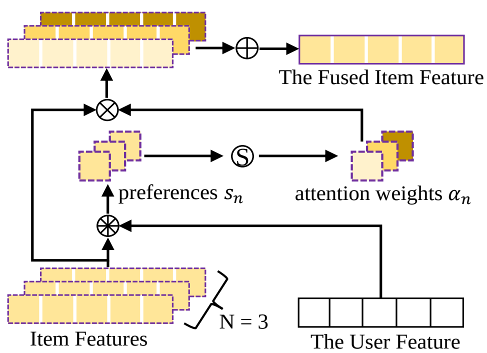


Fig. 4. Illustration of attention mechanism for compound samples of $\mathcal{D}^A$ ($N = 3$). The latent feature of potentially liked items are weighted involved with the guidance of user latent feature so that the uncertainty in the enriched user preference is avoided as much as possible. The user feature exhibits variations under different prediction heads, with domain-specific features (U_u^{PA}) under the separation head, domain-holistic features (U_u^{HA}) under the similarity head, domain-specific features of $\mathcal{D}^B$ (U_u^{PB}) under the peeling head.

one positive compound sample. Among the co-liked items, the item being co-liked by more similar users is more likely to be a potential positive item. Therefore, we sample the unobserved items with probability proportional to the number of similar users who liked them.

In the case of a negative compound sample, it simply selects a random number of items from the set of unobserved items. While this procedure may appear similar to existing methods [7], [34], [35], the key distinction lies in the selection of multiple unobserved samples. This deliberate choice significantly enhances the likelihood of including at least one true negative item within the negative sample. In comparison to conventional single negative items, the negative compound sample effectively diminishes bias during the optimization process of user preferences.

Feature Fusion with Attention Mechanism. In line with conventional Recommender Systems (RSs), the user's preference for a particular item is typically represented by the product of the user's latent feature and the item's latent feature. Inspired by this, we propose a simple yet efficient attention mechanism that fuses the features of different items based on their correlation with user features. Fig. 4 visually illustrates the process of feature fusion for a compound sample comprising two items in domain A. The features of these items, denoted as $I_{i_n}^A$ for $n = 1, 2$, are multiplied with user features in different prediction heads, for example, the domain-holistic features U_u^H employed in the similarity head, which will be further explained later. Consequently, the element-wise products between the user features and the respective item features are calculated as follows:

$$s_n = U_u (I_{i_n}^A)^T, n = 1, ..., N, \tag{1}$$

where N represents the number of items in a compound sample, and in this specific scenario, $N = 2$. The variable s_n signifies the user's individual preferences for the different items. To obtain a normalized score, reflecting the relative preference, these values are passed through a softmax activation layer, resulting in a probability distribution:

$$\alpha_n = \frac{\exp(s_n)}{\sum_{i=1}^N \exp(s_i)}, n = 1, 2, ..., N. \tag{2}$$

Afterwards, the attention weights α_n are employed to fuse the features of different items. The fused item features, denoted as $I_{i'}^A$, are obtained through the following equation:

$$I_{i'}^A = \sum_{n=1}^{N} \alpha_n I_{i_n}^A. \tag{3}$$

The advantages of compound sampling strategy are as follows. As the user and item features are gradually optimized, the attention weight α_n gradually converges towards the user's preference score for item i_n. Consequently, in a compound sample containing only one observed positive item, the attention weight assigned to that item will be high, while the

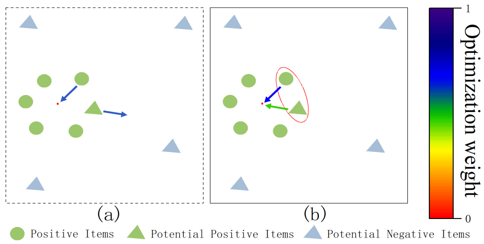


Fig. 5. Optimization differences for samples in (a) conventional sampling strategy and (b) compound sampling strategy.

weights for the remaining negative items will be low. This means that the negative items have minimal impact, and the compound sample is optimized in a similar manner to a single positive item sampled using conventional strategies. However, when a compound sample consists of one or more potential positive items in addition to the observed positive item, the attention weights for all positive items will be high. This leads to a significant enhancement of the positive compound sample features through the inclusion of the potential positive items. This enhancement is beneficial in learning less biased user preferences. Fig. 5 illustrates a comparison of the optimization process between the conventional sampling strategy and the proposed compound sampling strategy. In previous methods, potential positive items (depicted as green triangles) are considered negative and are pushed away from the user preference (represented by the red point). In contrast, with the compound sampling strategy, the potential positive item also contributes to the user preference model, thereby alleviating the data sparsity problem.

C. Feature Separation with Constraints for Inter-domains Preferences

As illustrated in Fig. 3, the user preferences are represented using three different forms. The domain-holistic features encapsulate the overall user preferences exhibited in single domain while the domain-invariant features and domain-specific features stand for the components of user preferences that can and cannot be shared across different domains, respectively. To optimize these features, our approach incorporates three specifically designed constraints, each tailored to a specific prediction head.

Following conventional NCF [36] paradigm, the user features are learned with simple multi-layer perception (MLP) structures, which is formulated as:

$$f_{MLP}(\boldsymbol{x}|\boldsymbol{\theta}) = \psi(\boldsymbol{W}_L\phi(...(\phi(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{b}_1))...) + \boldsymbol{b}_L), \tag{4}$$

where $\boldsymbol{x}$ denotes the concatenated representations of user features and fused item features. ϕ, ψ, $\boldsymbol{W}$ and $\boldsymbol{b}$ denote ReLU activation function, Sigmoid activation function, the weight matrix and the bias vector respectively. $\boldsymbol{\theta}$ is used to represent the parameters in MLP head for clarity. The MLP prediction head is trained by binary cross-entropy loss, which can be formulated as:

$$\mathcal{L}_{bce}(r, \widehat{r}, \mathcal{R}) = -\sum_{\mathcal{R}} r \log \widehat{r} + (1-r)\log(1-\widehat{r}), \tag{5}$$

where r and $\hat{r}$ are the label and prediction score for each sample, respectively; $\mathcal{R}$ is the sample set. Here the samples refer to the compound samples proposed in Sec. III-B.

Separation Constraint. The separation constraint stems from the motivation that user features is made up of domain-invariant features and domain-specific features. Therefore, it is reasonable to constrain that the combination of domain-invariant features and the domain-specific features is expected to reconstruct user-item interactions in different domains. To implement this constraint, we design an MLP head that combines the domain-invariant features and domain-specific features. The prediction can be expressed as follows:

$$\begin{aligned} \widehat{r}^A_{se} &= \mathcal{F}^A_{se}(U^I_u \oplus U^{PA}_u \oplus I^A_{i'}), \\ \widehat{r}^B_{se} &= \mathcal{F}^B_{se}(U^I_u \oplus U^{PB}_u \oplus I^B_{i'}), \end{aligned} \tag{6}$$

where $\mathcal{F}^A_{se}(\boldsymbol{x}) = \boldsymbol{f}_{MLP}(\boldsymbol{x}|\boldsymbol{\theta}^A_{se})$ and $\mathcal{F}^B_{se}(\boldsymbol{x}) = \boldsymbol{f}_{MLP}(\boldsymbol{x}|\boldsymbol{\theta}^B_{se})$ represent two independent MLP models. The variables U^I_u, U^P_u, and $I_{i'}$ denote the domain-invariant features, domain-specific features, and fused compound sample features, respectively. The operator $\oplus$ represents the concatenation operation. For the sake of clarity, the user index u and compound sample index i' are omitted in the prediction score $\widehat{r}$.

The separation constraint can be expressed as cross-entropy loss as follow:

$$\mathcal{L}^A_{se} = \mathcal{L}_{bce}(r^A, \widehat{r}^A_{se}, \mathcal{R}^A), \quad \mathcal{L}^B_{se} = \mathcal{L}_{bce}(r^B, \widehat{r}^B_{se}, \mathcal{R}^B), \tag{7}$$

where $\mathcal{R}^A$ and $\mathcal{R}^B$ represent the compound sample sets in $\mathcal{D}^A$ and $\mathcal{D}^B$, respectively. The variables r^A and r^B denote the corresponding ground-truth scores in $\mathcal{D}^A$ and $\mathcal{D}^B$, respectively.

Similarity Constraint. The primary objective of Cross-Domain Recommendation is to facilitate the transfer of valuable user preferences across different domains. In order to achieve this goal, it is crucial to maximize the effectiveness of the transferred knowledge. However, relying solely on the separation constraint does not ensure that the extracted domain-invariant features contain full information that is transferable. In extreme cases, the domain-invariant features might lack substantial effective information, resulting in the model resembling two Single-Domain Recommendation (SDR) systems. To overcome this limitation, we propose the similarity constraint, which is designed to maximize the amount of domain-invariant information, and thereby boost the effectiveness of the transferred knowledge.

To obtain as many transferable user preferences as possible, we first introduce domain-holistic features that capture the complete user preferences within each individual domain. The corresponding MLP heads can be expressed as follows:

$$\widehat{r}^A_{sim} = \mathcal{F}^A_{sim}(U^{HA}_u \oplus I^A_{i'}), \quad \widehat{r}^B_{sim} = \mathcal{F}^B_{sim}(U^{HB}_u \oplus I^B_{i'}), \tag{8}$$

where U^H_u denotes the domain-holistic features. $\mathcal{F}^A_{sim}(\boldsymbol{x})$ and $\mathcal{F}^B_{sim}(\boldsymbol{x})$ are independent MLP heads, i.e., $\mathcal{F}^A_{sim}(\boldsymbol{x}) = \boldsymbol{f}_{MLP}(\boldsymbol{x}|\boldsymbol{\theta}^A_{sim})$ and $\mathcal{F}^B_{sim}(\boldsymbol{x}) = \boldsymbol{f}_{MLP}(\boldsymbol{x}|\boldsymbol{\theta}^B_{sim})$.

Next, we are motivated by that domain-invariant features are transferable parts of the domain-holistic features, and propose to simultaneously maximize the similarity between the domain-invariant features with each domain-holistic features. Based on experimental results, the *pearson* similarity has been found to outperform cosine similarity. Therefore, the *pearson* similarity is chosen in our model, and its formulation is as follows:

$$\rho_{sim}^{A} = \frac{cov(U_u^I, U_u^{HA})}{\sigma(U_u^I)\sigma(U_u^{HA})}, \quad \rho_{sim}^{B} = \frac{cov(U_u^I, U_u^{HB})}{\sigma(U_u^I)\sigma(U_u^{HB})}, \tag{9}$$

$cov(\cdot,\cdot)$ and $\sigma(\cdot)$ calculate covariance and standard deviation, respectively. Combining with cross entropy losses for holistic MLP head, we get the overall similarity constraints as:

$$\begin{aligned} \mathcal{L}_{sim}^{A} &= \mathcal{L}_{bce}(r^A, \widehat{r}_{sim}^{A}, \mathcal{R}^A) - \lambda_p \rho_{sim}^{A}, \\ \mathcal{L}_{sim}^{B} &= \mathcal{L}_{bce}(r^B, \widehat{r}_{sim}^{B}, \mathcal{R}^B) - \lambda_p \rho_{sim}^{B}, \end{aligned} \tag{10}$$

where λ_p is a hyper-parameter.

Peeling Constraint. To further ensure the complete separation of domain-invariant fetures and domain-specific features, we further impose a peeling constraint on domain-specific features. As domain-specific features represent the user preferences irrelevant to the other domain, these features are considered useless or even harmful for predicting user-item interactions in the other domain. Motivated by this intuition, we strengthen the domain-specific features upon domain-holistic features and enforce the combined features to be less predictive of user-item interactions when transferred to other domains. In this way, the beneficial domain-invariant information is peeled from domain-specific features. The prediction is implemented by another set of MLP heads defined as:

$$\begin{aligned} \widehat{r}_{pe}^{A} &= \mathcal{F}_{pe}^{A}(U_u^{PB} \oplus U_u^{HB} \oplus I_{i'}^{A}), \\ \widehat{r}_{pe}^{B} &= \mathcal{F}_{pe}^{B}(U_u^{PA} \oplus U_u^{HA} \oplus I_{i'}^{B}), \end{aligned} \tag{11}$$

where $\mathcal{F}_{pe}^{A}(\boldsymbol{x}) = \boldsymbol{f}_{MLP}(\boldsymbol{x}|\boldsymbol{\theta}_{\boldsymbol{pe}}^{\boldsymbol{A}})$, $\mathcal{F}_{pe}^{B}(\boldsymbol{x}) = \boldsymbol{f}_{MLP}(\boldsymbol{x}|\boldsymbol{\theta}_{\boldsymbol{pe}}^{\boldsymbol{B}})$. In the peeling constraint, it is important to note that the user features from $\mathcal{D}^B$ and the item features from $\mathcal{D}^A$ are used to predict interaction score for compound samples belonging to $\mathcal{D}^A$. The similar situation occurs in $\mathcal{D}^B$.

The incapacity to predict correct items is implemented by optimizing the prediction score towards a uniform value, which indicates that user has no clear preference for certain item. The cross-entropy losses can be formulated as:

$$\mathcal{L}_{pe}^{A} = \mathcal{L}_{bce}(\bar{r}^A, \widehat{r}_{pe}^{A}, \mathcal{R}^A), \quad \mathcal{L}_{pe}^{B} = \mathcal{L}_{bce}(\bar{r}^B, \widehat{r}_{pe}^{B}, \mathcal{R}^B), \tag{12}$$

where $\bar{r}$ denotes the uniform prediction score of a batch of samples in our model, outperforming the fixed value of 0.5 suggested in [37], [38] based on experimental results.

The overall loss function is the combination of separation constraint, similarity constraint and peeling constraint in different domains, written as:

$$\mathcal{L} = \mathcal{L}_{se}^{A} + \mathcal{L}_{sim}^{A} + \lambda_s \mathcal{L}_{pe}^{A} + \mathcal{L}_{se}^{B} + \mathcal{L}_{sim}^{B} + \lambda_s \mathcal{L}_{pe}^{B}, \tag{13}$$

where λ_s is a hyper-parameter to balance the loss.

IV. Experiments and Analysis

In this section, we conduct extensive experiments on four datasets to answer the following research questions (RQs):

- RQ1: How does the proposed SUCCDR model perform in comparison with other state-of-the-art recommender?
- RQ2: Can the feature separation and compound sample module improve recommendation performance?
- RQ3: Does our model achieve the separation of domain-invariant user features and domain-specific user features?
- RQ4: How do different hyperparameter settings influence the recommendation performance of our method?
- RQ5: How helpful is additional text information in improving recommendation performance?

TABLE I
Statistics of four CDR tasks on *Douban* and *Amazon* datasets.

Datasets	Tasks		Users	Items	Interactions	Density
Douban	Task 1	Movie	2,106	9,555	969,937	4.82%
		Book		6,777	95,974	0.67%
	Task 2	Movie	1,665	9,555	833,676	5.24%
		Music		5,567	69,680	0.75%
Amazon	Task 3	Movie	1,051	4,261	80,560	1.80%
		Book		5,562	68,448	1.46%
	Task 4	Movie	695	4,350	32,480	1.07%
		Music		2,981	13,604	0.66%

A. Datasets

We conduct extensive evaluations of the proposed model on popularly used real-world datasets, i.e., Douban[1] and Amazon[2].

The Douban dataset encompasses three distinct domains: *Douban Movie*, *Douban Book* and *Douban Music*. These domains have been acquired through a private crawl of douban.com, and each of them comprises ratings, user profiles (personal information of users and self-evaluation), reviews (helpfulness text), and item details (descriptions, brands, category information). Notably, several prestigious CDR works (e.g., [39], [9], and [40]) have adopted this comprehensive dataset for evaluation.

The Amazon dataset utilized in this study is derived from a publicly accessible compilation of Amazon reviews. It encompasses an extensive assortment of 42 distinct domains or categories, encompassing ratings, user profiles, reviews, and item specifics. As part of our evaluation process, we deliberately choose pertinent domains to serve as experimental benchmarks for our model, specifically *Amazon Movie*, *Amazon Book*, and *Amazon Music*.

As this work focuses on the common user scenario in CDR, we preprocess the datasets to remain only common users. In addition, we normalize user ratings to 1 for those who have rated, and 0 for those who have not. Based on the above settings, we end up with four cross-domain tasks i.e., *Douban Movie* & *Douban Book* (task 1), *Douban Movie* & *Douban Music* (task 2), *Amazon Movie* & *Amazon Book* (task 3) and *Amazon Movie* & *Amazon Music* (task 4). The detailed statistics of datasets are shown in Table I.

B. Experimental Settings

1) Evaluation Protocols: In cross-domain recommendation scenarios, the leave-one-out (LOO) evaluation technique enjoys broad adoption, and we employ the same approach in this work. To be specific, we randomly select one item from

[1]https://github.com/FengZhu-Joey/GA-DTCDR/tree/main/Data
[2]http://jmcauley.ucsd.edu/data/amazon/index_2014.html

the set of items that a user has interacted with as the test target item, while the remaining interacted items are used for training to learn the user's preferences. Consistent with the literatures [7], [30], [39], we randomly sample 99 uninteracted items for the user, which serve as the test negative items. During the testing phase, the CDR model generates scores for all 100 items per user and the top K items are considered as the primary recommendations. Moreover, we employ two widely adopted evaluation metrics, namely **HR** (Hit Ratio) and **NDCG** (Normalized Discounted Cumulative Gain), to evaluate the model's efficacy in making recommendation. The metrics measuring whether the target item is in the K primary recommendation are denoted as HR@K and NDCG@K, accordingly. To reduce the interference of randomness, we conduct five times of training and testing for each model and report the average results.

2) Parameter Settings: In our experiments, for each positive compound sample, we randomly generate five negative compound samples that exclusively contain unobserved items. We set the maximum number of epochs to 50, the learning rate to 0.001, and the l_2 regularization to 0.0005. The batch size for Douban dataset is set to 4096, while for Amazon dataset, it is set to 256. We use λ_p = 0.5, λ_s = 0.5. The dimension (d) of the embedding varies within the set {32, 64}. The structure of the MLP is defined as "$8d - 4d - d - 1$". To optimize the model, we employ the Adam optimizer [41].

C. Baseline

We compare the proposed SUCCDR model with the following baselines, including several state-of-the-art single-domain and cross-domain recommendation methods.

- **DeepMF** [34]. Deep Matrix Factorization (DeepMF) employs a neural network-based deep matrix decomposition model to effectively leverage both implicit and explicit feedback information.
- **NCF** [36]. Neural Collaborative Filtering (NCF) combines matrix factorization with multilayer perceptrons (MLPs) to capture intricate nonlinear interactions between users and items.
- **CoNet** [42]. Co-occurrence Neural Network (CoNet) is designed to effectively capture the intricate relationships among items, thereby enabling a more comprehensive understanding of user preference features.
- **DTCDR** [39]. Dual-Target Cross-Domain Recommendation (DTCDR) uses max-pooling operation on common user embeddings to transfer user preference features across domains. It should be noted that for a fair comparison, the text information of this model has been excluded in this section.
- **ACDR** [43]. Adversarial Cross Domain Recommendation (ACDR) utilizes the idea of adversarial learning to intelligently incorporate global user preferences and domain-specific user preferences.
- **DDTCDR** [7]. Deep Dual Transfer Cross Domain Recommendation (DDTCDR) incorporates a latent orthogonal mapping function to facilitate the transfer of user preferences across different domains.
- **DML** [8]. Dual Metric Learning (DML) employs a novel latent orthogonal mapping function based on DDTCDR to align the representations of shared users and facilitate the transfer of user preferences across different domains.
- **BITGCF** [30]. Bi-directional Transfer Graph Collaborative Filtering Network (BITGCF) first generates users/items representations based on two graph encoders and then uses a bi-directional transfer layer to fuse user features.
- **ETL** [44]. Equivalent Transformation Learner (ETL) introduces a novel approach by proposing an equivalent transformation learner that models the joint distribution of user behaviors in two domains.

D. Performance Comparisons (for RQ1)

Comprehensive experiments have been conducted on various dimensions of the latent embedding representation (d = 32 and d = 64). The corresponding experimental results on HR@10 and NDCG@10 are provided in Table II and Table III for different tasks as defined in Sec. IV-A. Note that for single-domain recommendation (SDR) models, we train these models individually for each domain and report their respective results. Several observations can be drawn from the experimental results in Table II to III:

- **For single-domain methods:** (1) The majority of single-domain recommendation methods perform inferior to cross-domain recommendation methods. For *Douban Movie* in task 1, the average performance of cross-domain methods improves over the single-domain methods by 4.18% and 4.17% in HR with regard to d = 32 & 64. The corresponding performance gain in NDCG is 6.37%, and 3.83% with regard to d = 32 & 64. This discrepancy in performance can be attributed to the effective knowledge transfer across different domains. (2) The comparable performance of NCF and CoNet highlights the validity and rationality of learning item relationships.
- **For cross-domain methods:** (1) The varying performance of different cross-domain methods indicates that the choice of transferring strategies can have a significant impact on the recommended performance. (2) The results obtained from DML and DDTCDR indicate that the utilization of an orthogonal mapping function can contribute to enhancing the recommendation performance. (3) BiTGCF outperforms other baselines significantly across all four tasks, demonstrating the effectiveness of using Graph Neural Networks (GNNs) for knowledge transfer across domains. (4) ETL and ACDR, both incorporating domain-specific features, achieve comparable results on the Douban datasets. This observation suggests the feasibility of separating user features into domain-specific and domain-invariant features.
- **For our SUCCDR:** (1) Across all four cross-domain recommendation tasks, SUCCDR consistently surpasses that of all baseline methods in terms of recommendation performance. This consistent superiority in performance demonstrates the effectiveness of utilizing potential item

TABLE II
PERFORMANCE COMPARISON OF DIFFERENT METHODS ON DOUBAN DATASETS (TASK 1 AND TASK 2) WITH d = 32 & 64. THE BEST PERFORMANCE IS **BOLDFACED** AND THE SECOND BEST PERFORMANCE IS UNDERLINED.

	Method		**SUCCDR**	BITGCF	ETL	DML	DDTCDR	ACDR	DTCDR	CoNet	NeuMF	DeepMF
Douban d=32	Movie	HR	**70.13**	68.42	67.95	68.66	68.80	67.43	67.95	66.70	67.76	62.73
		NDCG	**45.61**	44.19	43.09	43.37	43.30	42.61	42.95	42.22	42.25	38.47
	Book	HR	**50.76**	47.72	47.82	46.34	46.01	46.11	46.05	46.43	46.87	46.87
		NDCG	**31.43**	29.55	28.94	28.21	28.01	27.96	26.18	27.86	28.80	28.26
	Movie	HR	**69.90**	67.61	67.85	68.53	67.99	66.55	66.55	65.77	67.09	59.70
		NDCG	**44.31**	41.91	42.10	42.26	41.99	41.18	41.26	41.10	42.61	36.03
	Music	HR	**46.13**	43.12	43.79	42.56	41.74	39.42	35.56	40.88	41.98	41.68
		NDCG	**26.42**	23.61	24.54	23.54	23.38	21.36	18.76	23.04	23.54	23.46
Douban d=64	Movie	HR	**69.75**	69.04	69.15	69.09	68.71	67.68	68.80	66.81	67.38	64.20
		NDCG	**45.33**	44.32	43.00	43.81	43.60	42.47	43.93	42.22	43.60	40.67
	Book	HR	**51.03**	48.29	48.48	48.34	46.68	47.15	47.10	47.50	46.77	47.34
		NDCG	**31.95**	30.15	29.46	29.11	28.87	28.33	28.69	28.01	29.28	29.32
	Movie	HR	**69.30**	67.87	68.15	67.97	67.45	66.85	66.93	66.37	66.73	62.22
		NDCG	**44.17**	42.29	42.74	42.24	42.12	41.85	42.70	41.14	42.74	38.44
	Music	HR	**46.54**	42.98	43.12	43.30	43.42	39.96	37.48	42.40	42.40	42.22
		NDCG	**26.54**	24.95	24.54	24.73	24.93	22.99	20.83	23.49	24.13	23.32

TABLE III
PERFORMANCE COMPARISON OF DIFFERENT METHODS ON AMAZON DATASETS (TASK 3 AND TASK 4) WITH d = 32 & 64. THE BEST PERFORMANCE IS **BOLDFACED** AND THE SECOND BEST PERFORMANCE IS UNDERLINED.

	Method		**SUCCDR**	BITGCF	ETL	DML	DDTCDR	ACDR	DTCDR	CoNet	NeuMF	DeepMF
Amazon d=32	Movie	HR	**52.90**	50.62	46.15	50.33	50.71	40.72	48.72	49.06	49.48	38.63
		NDCG	**34.25**	29.34	25.03	27.92	27.61	21.31	26.30	27.55	28.16	19.59
	Book	HR	**69.93**	67.27	63.28	65.84	65.84	49.00	65.18	65.18	64.99	54.71
		NDCG	**52.72**	45.48	39.80	41.84	41.12	31.74	41.70	41.61	43.94	33.94
	Movie	HR	**46.04**	45.54	45.45	42.59	42.58	32.66	40.86	41.73	43.59	35.25
		NDCG	**30.48**	27.42	26.56	24.11	23.67	17.30	23.09	23.87	25.29	19.32
	Music	HR	**36.55**	35.42	35.62	34.96	35.11	22.88	31.37	33.67	32.16	28.06
		NDCG	**22.54**	20.01	21.89	19.42	19.96	11.64	18.45	18.93	19.75	14.46
Amazon d=64	Movie	HR	**54.42**	52.90	49.10	48.91	49.00	41.30	48.72	48.61	49.86	39.96
		NDCG	**36.32**	33.84	27.47	27.36	27.83	22.98	27.72	27.12	29.02	20.94
	Book	HR	**72.12**	70.98	64.26	66.41	66.41	51.38	66.03	66.08	67.70	56.80
		NDCG	**57.61**	49.80	40.38	43.52	43.87	33.71	46.78	43.47	45.60	36.22
	Movie	HR	**48.63**	48.49	47.59	44.46	43.74	32.52	40.86	42.59	45.33	36.26
		NDCG	**31.32**	29.96	29.17	25.78	25.57	17.15	24.10	24.81	25.98	18.89
	Music	HR	**38.30**	37.11	37.33	37.41	37.70	23.96	31.80	35.18	33.96	31.08
		NDCG	**24.06**	21.92	22.44	21.83	21.92	12.88	19.69	21.05	21.65	15.81

information and only transferring domain-invariant features. For example, in *Douban Movie* (Task 1), our model surpasses the runner-up results by improving HR and NDCG by 1.93% and 3.21% respectively at d = 32. In *Douban Book* (Task 1), our model achieves a 6.15% HR improvement and a 6.36% NDCG improvement over the runner-up with regard to d = 32. (2) Certain methods exhibit sensitivity to specific datasets, as exemplified by a notable decline in the performance of ACDR on Amazon datasets. This sensitivity could be attributed to the limited size of the Amazon interaction data, which may hinder effective model optimization. However, our proposed model maintains outstanding stability and performance consistency across datasets of varying sizes, validating its remarkable robustness.

TABLE IV
ABLATION STUDY OF KEY MODULES OF SUCCDR ON TWO TASKS WITH d = 32. *w/o* CS REMOVES THE COMPOUND SAMPLE MODULE TO SAMPLE ONLY ONE ITEM. *w/o* SAM SAMPLES UNOBSERVED ITEMS RANDOMLY IN THE WHOLE ITEM SET. V-SIM SAMPLES UNOBSERVED ITEMS BASED ON THE SIMILARITY OF PRETRAINED ITEM EMBEDDING. *w/o* ATT DIRECTLY SUMS UP THE ITEM FEATURES IN THE COMPOUND SAMPLE.

Datasets	Metrics	*w/o* Att	V-Sim	*w/o* Sam	*w/o* CS	SUCCDR
Movie	HR	69.32	68.69	69.58	69.25	70.13
	NDCG	45.01	44.77	45.09	44.64	45.61
Book	HR	49.98	49.59	50.15	48.18	50.76
	NDCG	30.78	30.88	30.98	29.32	31.43
Movie	HR	69.39	68.99	69.24	68.87	69.90
	NDCG	44.07	43.53	43.79	43.66	44.31
Music	HR	45.64	44.61	45.57	43.78	46.13
	NDCG	26.01	25.84	26.27	24.64	26.42

E. Ablation Study (for RQ2)

1) Impact of Compound Sample Module: To validate the efficacy of our proposed compound sample module, which aims to enhance the discovery of potential user preferences, we have conducted a series of ablation studies from the perspective of conventional single sample, variant sampling strategy of the unobserved items, and sample feature fusion. Specifically, *w/o* CS directly discards the compound sample module and follows the previous model settings of sampling one item for one user. *w/o* Sam samples unobserved items randomly in the whole item set instead of in the co-liked item set while V-Sim samples unobserved items based on the similarity of pretrained item embedding by NCF. Both two variant sampling strategies generate the same number of items as our model for fair comparison. *w/o* Att directly sums up the item features in the compound sample without the attention mechanism. The experimental results are presented in Table IV. We can observe that: (1) The SUCCDR consistently outperforms *w/o* CS, indicating that our compound samples effectively leverage potential positive items to learn user features with reduced bias. (2) *w/o* Sam performs worse than the proposed method, indicating that our co-liked sampling strategy can find more potential positive items than direct randomly sampling. Moreover, in comparison to V-Sim, our method exhibits a performance improvement. This can be attributed to the fact that sampling based on the item similarity has a higher probability of selecting very similar items, which does not significantly contribute to refining user features. On the other hand, our algorithm incorporates sampling based on co-liked items, ensuring a high likelihood of user preferences while also introducing certain unobserved types of items. This approach enhances the diversity of recommended items, consequently refining the potential user features. (3) The observed performance degradation of *w/o* Att implies that utilizing the attention mechanism to assign varying optimization weights based on user favorability assist in effectively leveraging potential positive items.

2) Impact of Feature Separation: In order to demonstrate the effectiveness of feature separation, we have conducted additional experiments by transferring domain-holistic features using an element-wise attention mechanism rather than transferring only only domain-invariant features. The corresponding model is denoted as EWCCDR. In addition, to show the efficiency of the proposed three constrains, we carry out ablation studies by removing or replacing certain constraints. *w/o* Peel denotes the elimination of peeling constraint, *w/o* Sim denotes the removal of similarity constraint. Furthermore, we introduce orthogonal constraints (V-orth) on domain-invariant and domain-specific features within each domain, a widely adopted technique by existing models [7], [12]. The experimental results, as depicted in Fig. 6, validate the effectiveness of our proposed method. We have the following key observations: (1) Our model consistently outperforms EWCCDR, highlighting the necessity of separating user features and limiting cross-domain transfer solely to domain-invariant features. (2) The observed decline in performance for both *w/o* Sim and *w/o* Peel compared to SUCCDR suggests that the presence of similarity

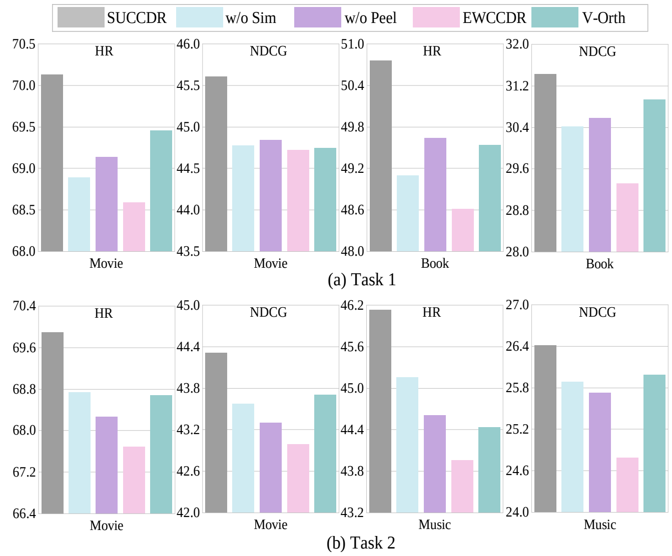


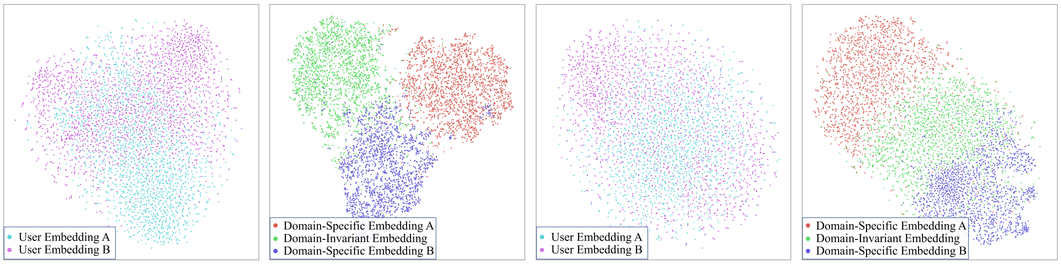
Fig. 6. Performance comparison of different variants of SUCCDR on two tasks with d = 32. *w/o* Peel and *w/o* Sim remove the peeling constraint and similarity constraint, respectively. EWCCDR transfers domain-holistic features. V-orth applies orthogonal constraints on domain-invariant and domain-specific features.

constraint plays a crucial role in maximizing domain-invariant information within the transferred features. Simultaneously, the peeling constraint contribute to minimizing the inclusion of domain-invariant information within the domain-specific features, thereby promoting enhanced feature separation. This phenomenon highlights the significance of both similarity and separation constraints in optimizing the overall performance of the model. (3) The superior performance of our model compared to V-orth underscores the limitations of relying on orthogonal constraints.

F. Latent User Factor Visualization (for RQ3)

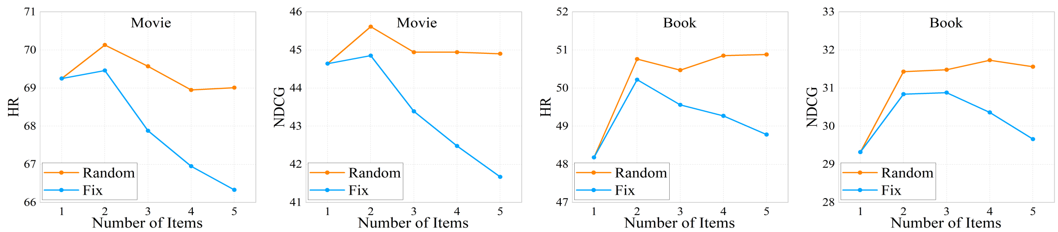
In line with our research objective, we separate user features into domain-invariant features and domain-specific features. To validate the effectiveness of our model in separating domain-invariant features from domain-specific features, we employ the t-SNE [45] nonlinear dimensionality reduction technique to transform the separated high-dimensional user features into two-dimensional representations. Specifically, we apply the t-SNE dimensionality reduction process to both our proposed method and the EWCCDR method, separately for Task 1 and Task 2. The resulting visualizations are presented in Fig. 7.

The results depicted in Fig. 7 (a) and (c) highlight the significant overlap of user features between the two domains in the EWCCDR method, making it difficult to differentiate between them. This finding suggests that there are inherent similarities in the user features across the two domains. Conversely, Fig. 7 (b) and (d) provide compelling evidence that our proposed SUCCDR model successfully separates user features within each domain while ensuring distinct separation of domain-specific features for both domains.

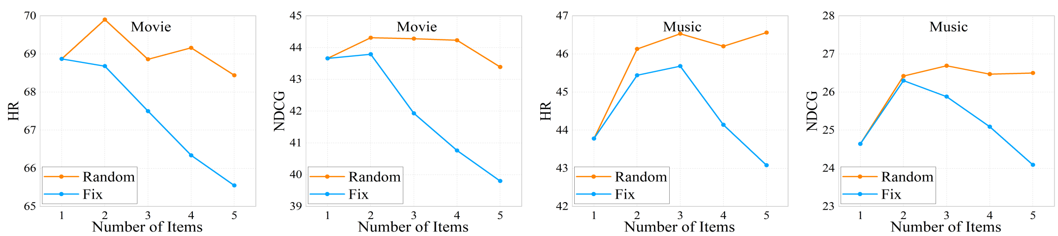


(a) EWCCDR (task 1) (b) SUCCDR (task 1) (c) EWCCDR (task 2) (d) SUCCDR (task 2)

Fig. 7. t-SNE visualization of user embeddings in EWCCDR and SUCCDR. (a) and (b) are the results of *Douban Movie* and *Douban Book* (task 1). (c) and (d) are the results of *Douban Movie* and *Douban Music* (task 2).

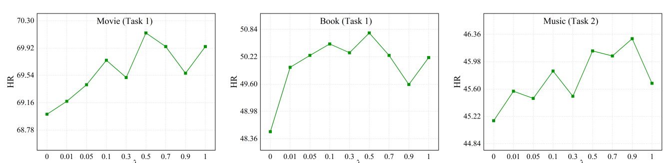


(a) Performance Variation on the *Douban Movie* and *Douban Book* (Task 1).

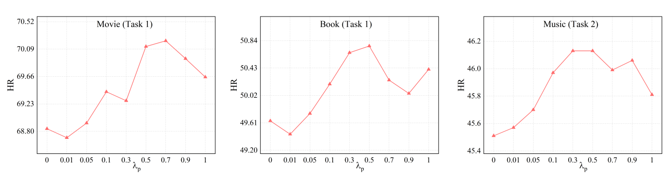


(b) Performance Variation on the *Douban Movie* and *Douban Music* (Task 2).

Fig. 8. Unraveling the Impact of Item Numbers N in Compound Samples with d = 32. We investigate two distinct approaches employed for the generation of compound samples. The first method involves selecting a fixed number of N items for each compound sample. Conversely, the second method involves a random selection process where the number of items in each compound sample is chosen to be not more than N. Please note that when N = 1, the two cases are identical.



(a) The Effect of Similarity coefficient λ_s.



(b) The Effect of Peeling coefficient λ_p.

Fig. 9. Investigating the Impact of Key Hyperparameters on Model Performance using Douban dataset, with the feature dimension d = 32.

G. *Parameter Sensitivity (for RQ4)*

In this section, we study the impact of key hyper-parameters on model performance, including N, λ_p and λ_s.

1) Impact of Item Numbers in Compound Samples: The compound sample plays a critical role in SUCCDR, facilitating the model's ability to delve into users' potential preferences. To gain a deeper understanding of the compound sample's significance, this section extensively examines the influence of item numbers in compound samples. By thoroughly investigating the impact of varying item numbers, we aim to unravel the intricate relationship between the compound sample and its capacity to uncover and explore the potential preferences of users. Based on the experimental results presented in Fig. 8, several notable phenomena and conclusions can be derived: (1) Remarkably, irrespective of the scenarios, the model consistently exhibits superior recommendation performance when N = 2 compared to N = 1. This compelling finding suggests that our proposed compound sample module successfully leverage items that users may also like, thereby exemplifying its ability to tap into their potential preferences effectively. (2) The models employing a random number of items up to N consistently outperform the corresponding models with fixed N items. This can be attributed to the increased robustness resulting from the uncertainty introduced in compound samples. (3) In the case of random generation, the sparser book and music domains

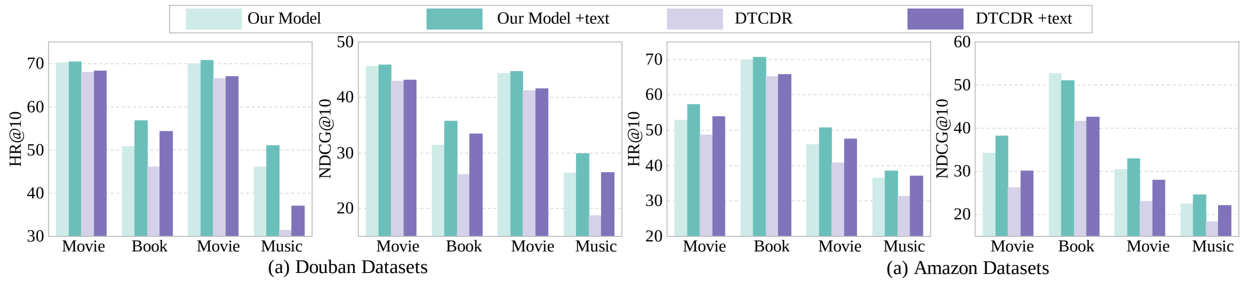


Fig. 10. The effectiveness of text information with d = 32. The light histogram represents performance under normal conditions, while the dark histogram shows performance after the introduction of text information.

tend to achieve optimal performance with larger values of N, while the denser movie domain shows a different trend. This phenomenon can be attributed to the nature of dataset density. In denser datasets, users' existing interactions already provide a comprehensive representation of their preferences. Therefore, employing excessively large compound samples in such cases may introduce noise and unnecessary information. On the other hand, in sparser datasets, the existing interactions may not sufficiently capture users' complete preferences. Consequently, the use of compound samples becomes crucial for exploring and uncovering users' potential preferences in such scenarios.

2) Impact of Similarity coefficient λ_s and peeling coefficient λ_p: We conduct experiments to study the effects of λ_p and λ_s by varying them in $\{0, 0.01, 0.05, 0.1, 0.3, 0.5, 0.7, 0.9, 1\}$ and Fig. 9 (a)-(b) show results of λ_s and λ_p respectively. All results exhibit a general pattern of initially increasing and then decreasing recommendation performance. An appropriately chosen value for λ_s maximizes the information captured by the domain-invariant features. However, if λ_s is set to an excessively large value, it introduces noise during knowledge transfer, leading to suboptimal solutions. The impact of λ_p exhibits similar trends, where suitable values of λ_p effectively extract the transferable information from domain-specific features. However, excessively large λ_p shifts the model's focus towards domain-specific features, leading to performance decline.

H. Effectiveness of additional text information (for RQ5)

To address the issue of data sparsity, other approaches [9], [40] leverage multi-source information, including text, images, and more. In this section, we investigate the influence of text information, which encompasses reviews, tags, user profiles, and item details. We extract text embeddings using the Doc2vec [46] and utilize them as user and item features. For simplicity, we directly concatenate the extracted textual features with the learnable interaction features.

Fig. 10 shows the effectiveness of text information. For fair comparison, we also report the results of original DTCDR which contains text information. We can observe that (1) Across different methods and datasets, the introduction of additional text information consistently improves the recommendation performance of the model, providing clear evidence that text information is helpful in addressing the data sparsity problem. (2) The impact of text information varies across various datasets. For instance, the improvement observed in the *Douban Movie* domain is relatively smaller compared to the enhancement seen in the *Douban book* and *Douban Music* domains. This difference can be attributed to the density of the *Douban Movie* domain, where the average number of user interactions is higher. As a result, the model can already capture user preferences comprehensively through existing interactions alone, making the role of text information less pronounced in this particular domain.

V. Conclusion

In this paper, we propose a novel cross-domain recommendation algorithm named SUCCDR to address data sparsity problem. A novel sampling strategy called compound sampling is proposed to alleviate the biased user preferences caused by sparse user-item interaction. The compound samples containing both observed and unobserved items are optimized with an attention mechanism to take advantage of potential positive items within intra-domains. In addition, in contrast to blindly transferring the inter-domain knowledge, we separate user preferences into domain-invariant features and domain-specific features, which are accomplished by three well-designed constraints. Finally, SUCCDR is validated by extensive experiments on real-world datasets, showing better performance than state-of-the-art SDR and CDR methods.

References

[1] F. Zhu, Y. Wang, C. Chen, J. Zhou, L. Li, and G. Liu, "Cross-domain recommendation: Challenges, progress, and prospects," in *International Joint Conferences on Artificial Intelligence Organization (IJCAI)*, 8 2021, pp. 4721–4728.

[2] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *International World Wide Web Conference (WWW)*, 2001, p. 285–295.

[3] S. Berkovsky, T. Kuflik, and F. Ricci, "Cross-domain mediation in collaborative filtering," in *International conference on User Modeling (UM)*, 2007, p. 355–359.

[4] P. Cremonesi, A. Tripodi, and R. Turrin, "Cross-domain recommender systems," in *International Conference on Data Mining Workshops (ICDMW)*, 2011, pp. 496–503.

[5] I. Cantador, I. Fernández-Tobías, S. Berkovsky, and P. Cremonesi, "Cross-domain recommender systems," in *Recommender Systems Handbook*, 2015, pp. 919–959.

[6] G. Hu, Y. Zhang, and Q. Yang, "Conet: Collaborative cross networks for cross-domain recommendation," in *International Conference on Information and Knowledge Managemen (CIKM)*, 2018, p. 667–676.

[7] P. Li and A. Tuzhilin, "Ddtcdr: Deep dual transfer cross domain recommendation," in *International Conference on Web Search and Data Mining (WSDM)*, 2020, pp. 331—339.

[8] P. Li and A. Tuzhilin, "Dual metric learning for effective and efficient cross-domain recommendations," *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, vol. 35, no. 1, pp. 321–334, 2021.

[9] F. Zhu, Y. Wang, C. Chen, G. Liu, and X. Zheng, "A graphical and attentional framework for dual-target cross-domain recommendation," in *International Joint Conferences on Artificial Intelligence Organization (IJCAI)*, 2020, pp. 3001–3008.

[10] Y. Li, J.-J. Xu, P.-P. Zhao, J.-H. Fang, W. Chen, and L. Zhao, "Atlrec: An attentional adversarial transfer learning network for cross-domain recommendation," *Journal of Computer Science and Technology (JCST)*, vol. 35, no. 4, p. 794–808, 2020.

[11] J. Cao, X. Lin, X. Cong, J. Ya, T. Liu, and B. Wang, "Disencdr: Learning disentangled representations for cross-domain recommendation," in *International Conference on Research and Development in Information Retrieval (SIGIR)*, 2022, p. 267–277.

[12] X. Zhang, J. Li, H. Su, L. Zhu, and H. T. Shen, "Multi-level attention-based domain disentanglement for bcdr," *ACM Transactions on Information Systems (TOIS)*, vol. 41, no. 4, 2023.

[13] F. Liu, H. Chen, Z. Cheng, A. Liu, L. Nie, and M. Kankanhalli, "Disentangled multimodal representation learning for recommendation," *IEEE Transactions on Multimedia (TMM)*, pp. 1–11, 2022.

[14] H. Tang, G. Zhao, J. Gao, and X. Qian, "Personalized representation with contrastive loss for recommendation systems," *IEEE Transactions on Multimedia (TMM)*, pp. 1–11, 2023.

[15] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, "Bpr: Bayesian personalized ranking from implicit feedback," in *the Conference on Uncertainty in Artificial Intellige (UAI)*, 2009, pp. 452–461.

[16] S. Tan, J. Bu, X. Qin, C. Chen, and D. Cai, "Cross domain recommendation based on multi-type media fusion," *Neurocomputing*, vol. 127, pp. 124–134, 2014.

[17] G. Hu, Y. Zhang, and Q. Yang, "Transfer meets hybrid: A synthetic approach for cross-domain collaborative filtering with text," in *International World Wide Web Conference (WWW)*, 2019, p. 2822–2829.

[18] Z. Cheng, X. Chang, L. Zhu, R. C. Kanjirathinkal, and M. Kankanhalli, "Mmalfm: Explainable recommendation by leveraging reviews and images," *ACM Transactions on Information Systems (TOIS)*, vol. 37, no. 2, 2019.

[19] M. Kaminskas and F. Ricci, "Location-adapted music recommendation using tags," in *The International Conference on User Modeling, Adaption, and Personalization (UMAP)*, 2011, p. 183–194.

[20] S. Berkovsky, T. Kuflik, and F. Ricci, "Mediation of user models for enhanced personalization in recommender systems," *User Modeling and User-Adapted Interaction*, vol. 18, no. 3, p. 245–286, 2008.

[21] C. Zhao, C. Li, R. Xiao, H. Deng, and A. Sun, "Catn: Cross-domain recommendation for cold-start users via aspect transfer network," in *International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 2020, p. 229–238.

[22] H. Su, Y. Zhang, X. Yang, H. Hua, S. Wang, and J. Li, "Cross-domain recommendation via adversarial adaptation," in *International Conference on Information and Knowledge Managemen (CIKM)*, 2022, p. 1808–1817.

[23] Y. Zhu, Z. Tang, Y. Liu, F. Zhuang, R. Xie, X. Zhang, L. Lin, and Q. He, "Personalized transfer of user preferences for cross-domain recommendation," in *International Conference on Web Search and Data Mining (WSDM)*, 2022, pp. 1507–1515.

[24] J. Liu, P. Zhao, F. Zhuang, Y. Liu, V. S. Sheng, J. Xu, X. Zhou, and H. Xiong, "Exploiting aesthetic preference in deep cross networks for cross-domain recommendation," in *International World Wide Web Conference (WWW)*, 2020, p. 2768–2774.

[25] C. Zhao, C. Li, and C. Fu, "Cross-domain recommendation via preference propagation graphnet," in *International Conference on Information and Knowledge Managemen (CIKM)*, 2019, p. 2165–2168.

[26] A. P. Singh and G. J. Gordon, "Relational learning via collective matrix factorization," in *International conference on Knowledge discovery and data mining (KDD)*, 2008, pp. 650–658.

[27] L. Hu, J. Cao, G. Xu, L. Cao, Z. Gu, and C. Zhu, "Personalized recommendation via cross-domain triadic factorization," in *International World Wide Web Conference (WWW)*, 2013, pp. 595–606.

[28] B. Loni, Y. Shi, M. Larson, and A. Hanjalic, "Cross-domain collaborative filtering with factorization machines," in *Advances in Information Retrieval*, 2014, pp. 656–661.

[29] F. Yuan, L. Yao, and B. Benatallah, "Darec: Deep domain adaptation for cross-domain recommendation via transferring rating patterns," in *International Joint Conferences on Artificial Intelligence Organization (IJCAI)*, 2019, pp. 4227—4233.

[30] M. Liu, J. Li, G. Li, and P. Pan, "Cross domain recommendation via bi-directional transfer graph collaborative filtering networks," in *International Conference on Information and Knowledge Managemen (CIKM)*, 2020, p. 885–894.

[31] X. Wang, H. Jin, A. Zhang, X. He, T. Xu, and T.-S. Chua, "Disentangled graph collaborative filtering," in *International Conference on Research and Development in Information Retrieval (SIGIR)*, 2020, p. 1001–1010.

[32] J. Xu, J. Song, Y. Sang, and L. Yin, "Cdaml: A cluster-based domain adaptive meta-learning model for cross domain recommendation," *World Wide Web*, vol. 26, no. 3, p. 989–1003, jun 2022.

[33] A. I. Schein, A. Popescul, L. H. Ungar, and D. M. Pennock, "Methods and metrics for cold-start recommendations," in *International Conference on Research and Development in Information Retrieval (SIGIR)*, 2002, p. 253–260.

[34] H.-J. Xue, X. Dai, J. Zhang, S. Huang, and J. Chen, "Deep matrix factorization models for recommender systems," in *International Joint Conferences on Artificial Intelligence Organization (IJCAI)*, 2017, pp. 3203–3209.

[35] W. Nie, X. Wen, J. Liu, J. Chen, J. Wu, G. Jin, J. Lu, and A.-A. Liu, "Knowledge-enhanced causal reinforcement learning model for interactive recommendation," *IEEE Transactions on Multimedia (TMM)*, pp. 1–14, 2023.

[36] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural collaborative filtering," in *International World Wide Web Conference (WWW)*, 2017, p. 173–182.

[37] Y. Fu, Y. Fu, J. Chen, and Y.-G. Jiang, "Generalized meta-fdmixup: Cross-domain few-shot learning guided by labeled target data," *IEEE Transactions on Image Processing*, vol. 31, pp. 7078–7090, 2022.

[38] M. Li, P.-Y. Huang, X. Chang, J. Hu, Y. Yang, and A. Hauptmann, "Video pivoting unsupervised multi-modal machine translation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 3, pp. 3918–3932, 2023.

[39] F. Zhu, C. Chen, Y. Wang, G. Liu, and X. Zheng, "Dtcdr: A framework for dual-target cross-domain recommendation," in *International Conference on Information and Knowledge Managemen (CIKM)*, 2019, pp. 1533—1542.

[40] C. Zhao, H. Zhao, M. HE, J. Zhang, and J. Fan, "Cross-domain recommendation via user interest alignment," in *International World Wide Web Conference (WWW)*, 2023, p. 887–896.

[41] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *International Conference on Learning Representations (ICLR)*, 2015, pp. 1–15.

[42] M. Chen, Y. Li, and X. Zhou, "Conet: Co-occurrence neural networks for recommendation," *Future Generation Computer Systems*, vol. 124, pp. 308–314, 2021.

[43] P. Li, B. Brost, and A. Tuzhilin, "Adversarial learning for cross domain recommendations," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 14, no. 1, pp. 1–25, 2022.

[44] X. Chen, Y. Zhang, I. W. Tsang, Y. Pan, and J. Su, "Toward equivalent transformation of user preferences in cross domain recommendation," *ACM Transactions on Information Systems (TOIS)*, vol. 41, no. 1, pp. 1–31, 2023.

[45] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, "Decaf: A deep convolutional activation feature for generic visual recognition," in *International Conference on International Conference on Machine Learning (ICML)*, 2014, pp. 647–655.

[46] Q. Le and T. Mikolov, "Distributed representations of sentences and documents," in *International Conference on International Conference on Machine Learning (ICML)*, 2014, pp. 1188–1196.