

Multi-Domain Recommendation to Attract Users via Domain Preference Modeling

Hyunjun Ju¹, SeongKu Kang^{1, 2}, Dongha Lee³, Junyoung Hwang¹, Sanghwan Jang¹, Hwanjo Yu^{1*}

¹ Pohang University of Science and Technology (POSTECH), Republic of Korea

² University of Illinois at Urban-Champaign (UIUC), United States

³ Yonsei University, Republic of Korea

{hyunjunju, jyhwang, s.jang, hwanjoyu}@postech.ac.kr, seongku@illinois.edu, donalee@yonsei.ac.kr

Abstract

Recently, web platforms are operating various service domains simultaneously. Targeting a platform that operates multiple service domains, we introduce *a new task*, Multi-Domain Recommendation to Attract Users (MDRAU), which recommends items from multiple “unseen” domains with which each user has not interacted yet, by using knowledge from the user’s “seen” domains. In this paper, we point out two challenges of MDRAU task. First, there are numerous possible combinations of mappings from seen to unseen domains because users have usually interacted with a different subset of service domains. Second, a user might have different preference for each of the target unseen domains, which requires recommendations to reflect users’ preference on domains as well as items. To tackle these challenges, we propose DRIP framework that models users’ preference at two levels (i.e., domain and item) and learns various seen-unseen domain mappings in a unified way with masked domain modeling. Our extensive experiments demonstrate the effectiveness of DRIP in MDRAU task and its ability to capture users’ domain-level preferences.

Introduction

Nowadays, web platforms are operating various service domains simultaneously (e.g., music streaming, game store, and eBook subscription). They allow users to experience diverse domains within a single platform and promote the mutual growth of all service domains through the Recommender System (RS). For such multi-domain platforms, recommending items from *unseen domains* with which each user has not interacted yet plays an essential role in the platform’s growth, user satisfaction, and business success. That is, users typically utilize a few domains rather than all domains, and accurate recommendations that align with user preference can attract users into unexplored domains. To this end, Cross-Domain Recommendation (CDR), which recommends items from unseen (target) domains based on user interaction history in seen (source) domains, has gained significant research attention.¹

*Corresponding Author

Copyright © 2024, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

¹In this paper, the terms “seen” and “unseen” are defined from the perspective of each user; as each user interacts with a different subset of service domains, the seen and unseen domain sets vary for

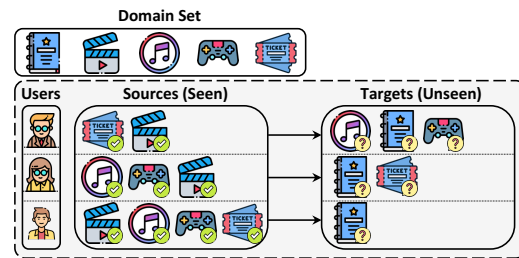


Figure 1: A conceptual illustration of MDRAU task. A platform operates five different service domains, and each user partially interacts with a subset of the entire service domains. MDRAU aims to provide recommendations from each user’s unseen domains to attract users.

Most CDR studies (Man et al. 2017; Kang et al. 2019; Zhu et al. 2022, 2021b; Fu et al. 2019) have focused on transferring user preference information from a source domain to a target domain. Given a recommender system employed for each domain, they learn a mapping function that acts as a bridge between the representation spaces of the two domains. To provide unseen domain recommendations, they transfer the user embedding from the source domain to the target domain using the mapping function and generate recommendations based on the transferred user embedding. More recently, (Cao et al. 2022b, 2023, 2022a) have achieved improved recommendation accuracy by modeling domain-specific and domain-shared information separately, and selectively transferring the domain-shared information. Despite their effectiveness, the existing studies have targeted the case of a single seen-unseen domain pair (i.e., one-to-one), and the case of multiple seen-unseen domains (i.e., many-to-many) has not been studied well. Especially, they do not consider making unified recommendations that include items from multiple domains.

In practical scenarios, it is becoming increasingly common for platforms to offer services in more than two domains. This creates a need to promote user engagement across multiple unseen domains by attracting users with accurate personalized recommendations. We refer to this prob-

each user. A user can be considered a cold-start user in the user’s unseen service domain (Man et al. 2017; Kang et al. 2019).

lem as Multi-Domain Recommendation to Attract Users (MDRAU) task (Fig. 1). Formally, MDRAU aims to provide a recommendation list that consists of items from each user’s unseen domain(s) that the user has not tried before. MDRAU brings several practical values to multi-domain platforms. It encourages users to explore new content beyond their previously interacted domains, fostering diverse and serendipitous discoveries. This diversified user experience can enhance user satisfaction and engagement, which helps provide more accurate recommendations in each domain.

Addressing MDRAU task presents two major challenges. First, since each user has interacted with different service domains, there are numerous possible combinations of seen-unseen domains. With K domains, the number of combinations can reach up to $2^K - 2$, except in the case that a user has used all services. This large number of combinations makes it difficult to apply the previous CDR methods that learn a one-to-one mapping function for each domain pair. Second, a user naturally has different preference for each unseen domain, and these varying domain preference need to be properly reflected in the recommendation process. That is, the recommendation needs to consider user preference in both *domain-level*, i.e., the inclination of a user to explore each unseen domain, and *item-level*, i.e., the inclination of a user to interact with a new item within a domain. The previous CDR methods have focused on improving item-level preference for a specific unseen domain without directly considering domain-level preference. As a result, they show limited performance when applied to MDRAU task.

To effectively solve our new task MDRAU, we propose DRIP, a new framework that learns various seen-unseen domain mappings in a unified way via masked domain modeling and models user preference at the domain and item levels. We formulate the training process of DRIP as a prediction task of missing information based on its contexts (Devlin et al. 2019; Bao et al. 2022). Then, we model the two-level preferences using a multi-domain encoder that incorporates user preference across multiple domains. The key idea is to randomly mask the user preference of some seen domains in the model input, and train the model to predict the user preference in the masked domains. During the training, we regard the masked domains as the user’s unseen domains, allowing the model to simulate and learn from *numerous scenarios* involving different combinations of seen and unseen domains. This enables the model to achieve the generalization capability of inferring user preferences in unseen domains from ones in seen domains. Furthermore, we introduce an adaptive masking scheme to make the model more focused on learning domains that a user is more likely to prefer. We validate the superiority of DRIP by extensive experiments on real-world datasets and provide a thorough comparison with various state-of-the-art methods.

Related Work

The existing CDR studies can be divided into two groups according to the type of target domain for recommendations. **CDR for Seen Domain Recommendation.** It aims to improve the recommendation quality of seen domains with

which the user has already interacted. Many studies alleviate the data sparsity problem in the sparse target domains by utilizing information from the source domain. To this end, they transfer knowledge among domains via bridging information, such as overlapping users or items. For example, CoNet (Hu, Zhang, and Yang 2018) introduces cross-connection units to transfer and integrate knowledge between source and target domains. DTCDR (Zhu et al. 2019) proposes a dual-target framework to improve the recommendation accuracy in both two involved domains simultaneously. GA-DTCDR (Zhu et al. 2020) extends DTCDR by adopting graph information. Recently, several studies have focused on multi-domain cases having more than two domains. GA-MTCDR (Zhu et al. 2023) extends GA-DTCDR with element-wise attention to integrating embeddings of overlapping users from multiple domains. CAT-ART (Li et al. 2023) proposes a contrastive autoencoder to encode a global user embedding and a mechanism to transfer user embeddings from each source domain to the target domain. UniCDR (Cao et al. 2023) introduces domain-specific and domain-shared embeddings along with aggregation schemes to make a universal model for existing CDR scenarios.

CDR for Unseen Domain Recommendation. It aims to provide recommendations in unseen domains with which the user has not yet interacted. Their focus lies on the method of obtaining user embeddings in the target domain space. For example, EMCDCR (Man et al. 2017) proposes an embedding and mapping framework, which learns a mapping function that transfers user embeddings from the source to the unseen target domain. SSCDR (Kang et al. 2019) proposes a semi-supervised embedding and mapping framework to train a mapping function, even when only a few labeled data are available. PTUPCDR (Zhu et al. 2022) uses a meta-network that generates a personalized mapping function. UniCDR (Cao et al. 2023) can also be applied to recommend unseen domains using domain-shared embeddings. They mainly focus on a single unseen target domain rather than multiple unseen target domains.

Problem Formulation

Notations

In this work, we focus on a scenario where a provider operates services for multiple domains (e.g., music streaming, game store, and eBook subscription), each of which employs a distinct recommender system. Each service domain has its own user and item set. Items of each domain are mutually exclusive, while users may use one or multiple service domain(s). Formally, given K domains $\mathcal{D} = \{d_1, \dots, d_K\}$, \mathcal{U}_k and \mathcal{V}_k denote the set of users and items for the k -th domain, respectively. The user-item interaction history for d_k is represented by a matrix $\mathbf{R}^{(k)} \in \{0, 1\}^{|\mathcal{U}_k| \times |\mathcal{V}_k|}$, where $R_{u,v} = 1$ if user u has interacted with the item v , otherwise $R_{u,v} = 0$. Without loss of generality, we define an interaction matrix of all domains $\mathbf{R} \in \{0, 1\}^{|\mathcal{U}| \times |\mathcal{V}|}$, where $\mathcal{U} = \bigcup_{k=1}^K \mathcal{U}_k$ and $\mathcal{V} = \bigcup_{k=1}^K \mathcal{V}_k$. We additionally define the user-domain relations as $\mathbf{G} \in \{0, 1\}^{|\mathcal{U}| \times |\mathcal{D}|}$, where $G_{u,k} = 1$ if the user u has interacted with items of domain

d_k , otherwise $G_{u,k} = 0$. Overlapping users indicate users who have interacted with at least two domains.

MDRAU Task

Definition 1 (Multi-Domain Recommendation to Attract Users) Given user-item interaction history from multiple service domains, MDRAU refers to the task of providing a ranking list (i.e., recommendation) that consists of items from each user’s unseen domain(s) that the user has not tried before.

In some cases, providers may have a need to promote a specific target domain within the platform. We refer to the scenario with a single target domain as MDRAU-ST, and the scenario with multiple target domains as MDRAU-MT. MDRAU provides preferable items from each user’s unseen domains, encouraging the exploration of new content beyond their previously interacted domains. By doing so, MDRAU helps to diversify user experience and facilitate serendipitous discoveries, enhancing user engagement. This enhanced engagement, in turn, helps to provide more accurate recommendations in each domain, ultimately contributing to the growth and revenue of the platform.

Proposed Framework

Overview

We present a unified framework for MDRAU task, named as DRIP (Domain pReference-aware unseen domain Item Prediction). DRIP optimizes the parameters of its model by maximizing the likelihood given the training data \mathbf{R} . Let $\mathbf{R}_u \in \{0, 1\}^{|\mathcal{V}|}$ denote a multi-hot vector representing a user’s interaction with the items over all domains. We assume the observed data are drawn from the Multinomial distribution, and the likelihood is described by

$$p(\mathbf{R}_u) = \prod_{v \in \mathcal{V}} p(v|u)^{R_{u,v}} \quad (1)$$

$$\mathbf{R}_u \sim \text{Mult}(N_u, p(v|u)),$$

where $N_u = \sum_v R_{u,v}$ and $p(v|u)$ is the probability that user u prefers the item v over the entire item set. We decompose the likelihood based on the domain-level and item-level preference as follows:

$$p(\mathbf{R}_u) = \prod_{v \in \mathcal{V}} p(v|u)^{R_{u,v}} = \prod_{d_k \in \mathcal{D}} \prod_{v \in \mathcal{V}_k} p(v, d_k|u)^{R_{u,v}} \quad (2)$$

$$= \prod_{d_k \in \mathcal{D}} \prod_{v \in \mathcal{V}_k} [p(v|u, d_k) \cdot p(d_k|u)]^{R_{u,v}},$$

where $p(v|u, d_k)$ denotes user u ’s preference for item v in domain d_k and $p(d_k|u)$ denotes the user’s preference for domain d_k . These preferences are modeled by a unified neural model with a multi-domain encoder based on self-attention. Then, to maximize the likelihood, we train the model via *masked domain modeling* that predicts the item preference of the masked domains. The recommendations are produced by considering both domain- and item-level preferences for unseen domains of each user. Fig. 2 illustrates the overall DRIP architecture.

Domain-Specific Encoder

We assume that a service platform operates multiple service domains and has deployed its own RS for each domain. An RS model contains *encoders* that encode the user and item information into representation space, where the user-item similarity is measured for recommendations. In specific, for each domain d_k , let $f_{\theta_k} : \mathbb{R}^{|\mathcal{U}_k|} \rightarrow \mathbb{R}^d$, $f_{\xi_k} : \mathbb{R}^{|\mathcal{V}_k|} \rightarrow \mathbb{R}^d$, and $\text{sim}_k(\cdot, \cdot)$ denote a user encoder, an item encoder, and similarity function, respectively. Diverse architectures can be adopted for the encoders (e.g., id-based (Koren, Bell, and Volinsky 2009; Rendle et al. 2009) and graph-based (Wang et al. 2019)), and $\text{sim}_k(\cdot, \cdot)$ can be either a simple metric or a learnable function. In this work, we use a simple id-based encoder with the inner product similarity, as done in (Rendle et al. 2009; Man et al. 2017; Zhu et al. 2022).

Multi-Domain Encoder

The main component of our unified neural model is a *multi-domain encoder* that aims to enrich user embedding from each domain-specific encoder with that from the encoders for other domains. It adopts the self-attention mechanism to aggregate information from other domains based on the similarity of user preferences across multiple domains.

Constructing Masked Input. Let $\mathbf{x}_{u,k} = f_{\theta_k}(u)$ denote user embedding of user u in domain d_k . A user u is represented as the set of the corresponding user embeddings for all domains $\{\mathbf{x}_{u,k}\}_{k=1}^K$. Note that some domains may have no interaction history with the user (i.e., $G_{u,k} = 0$), as users typically utilize a few domains rather than all domains. To handle this case, we replace the embedding for the user’s unseen domains with $\mathbf{e}_{[M]}$, which is the learnable embedding of a special mask token $[M]$ as follows:

$$\bar{\mathbf{X}}_u = \{\bar{\mathbf{x}}_{u,k}\}_{k=1}^K, \quad \bar{\mathbf{x}}_{u,k} = (1 - G_{u,k})\mathbf{e}_{[M]} + G_{u,k}\mathbf{x}_{u,k}. \quad (3)$$

Since the embeddings are obtained from independently trained domain-specific RS models, they have different distributions. We align the distributions using projectors $g_{\phi_k}(\mathbf{x}_{u,k})$, where $g_{\phi_k} : \mathbb{R}^d \rightarrow \mathbb{R}^m$. Also, we insert $\mathbf{e}_{[S]} \in \mathbb{R}^m$ at the beginning of the input for the encoder, which is the learnable embedding of a special token $[S]$; this will be used to estimate each user’s domain-level preference. The final input representation of user u is constructed as

$$\mathbf{H}_u^0 = [\mathbf{e}_{[S]}, g_{\phi_1}(\bar{\mathbf{x}}_{u,1}), \dots, g_{\phi_K}(\bar{\mathbf{x}}_{u,K})]^\top \in \mathbb{R}^{(K+1) \times m}. \quad (4)$$

We do not use position embedding because spatial position information is unnecessary for our target task.

Contextualizing User Embeddings over Multi-Domains.

The input \mathbf{H}_u^0 is forwarded into the multi-domain encoder to contextualize each domain-specific user embedding over the user’s multiple domains based on the self-attention mechanism. The multi-domain encoder is basically a stack of L transformer layers (Vaswani et al. 2017; Devlin et al. 2019). The details are described in the Appendix. The l -th transformer layer can be simply described by

$$\mathbf{H}_u^{l+1} = \text{Transformer}(\mathbf{H}_u^l), \quad \forall l \in \{0, 1, \dots, L-1\}. \quad (5)$$

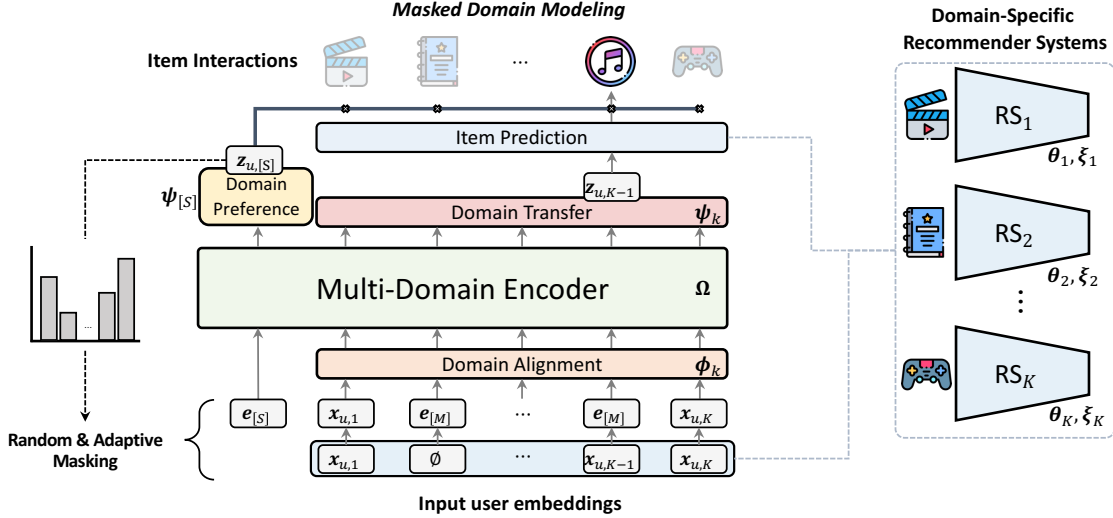


Figure 2: The overview of the proposed DRIP framework.

We denote a set of learnable parameters in the multi-domain encoder consisting of transformer layers as Ω . In the end, the final output of L -th layer is obtained by

$$\mathbf{H}_u^L = [\mathbf{h}_{u,[S]}^L, \mathbf{h}_{u,1}^L, \dots, \mathbf{h}_{u,K}^L]^\top \in \mathbb{R}^{(K+1) \times m}. \quad (6)$$

Through the transformer layers, the user embedding from each domain gets contextualized by attending to the embeddings from other domains based on the embedding similarity across the domains. As a result, $\mathbf{h}_{u,k}^L$ encodes user preference for domain d_k , being enriched by preference information from the other remaining domains.

Preference Modeling

Domain-Level Preference. The domain-level preference can be inferred by using the contextualized representation of the special token $[S]$ for user u , denoted by $\mathbf{h}_{u,[S]}^L$, which is generated by aggregating a user's preference information for multiple domains. We introduce a domain-preference predictor $q\psi_{[S]} : \mathbb{R}^m \rightarrow \mathbb{R}^K$ to predict each user's inclination to each domain: $\mathbf{z}_{u,[S]} = q\psi_{[S]}(\mathbf{h}_{u,[S]}^L)$. The domain-level preference is defined by

$$p(d_k|u; \Theta) = \frac{\exp(\mathbf{z}_{u,[S]}_k)}{\sum_{d_i \in \mathcal{D}} \exp(\mathbf{z}_{u,[S]}_i)}, \quad (7)$$

where $\mathbf{z}_{u,[S]}_k$ indicates k -th logit value in $\mathbf{z}_{u,[S]}$.

Item-Level Preference. The item-level preference in each domain is obtained from the similarity of item embeddings to the contextualized user embedding for the domain. Using a domain-specific projector $q\psi_k : \mathbb{R}^m \rightarrow \mathbb{R}^d$, we project $\mathbf{h}_{u,k}^L$ to the representation space of domain d_k : $\mathbf{z}_{u,k} = q\psi_k(\mathbf{h}_{u,k}^L)$. We compute the in-domain item-level preference based on user u 's similarity distribution over item set \mathcal{V}_k ,

$$p(v|u, d_k; \Theta) = \frac{\exp(\text{sim}_k(\mathbf{z}_{u,k}, \mathbf{x}_{v,k}))}{\sum_{\tilde{v} \in \mathcal{V}_k} \exp(\text{sim}_k(\mathbf{z}_{u,k}, \mathbf{x}_{\tilde{v},k}))}, \quad (8)$$

where $\mathbf{x}_{v,k} = f_{\xi_k}(v)$ is the embedding of item v in the domain d_k .

Model Learning

Masked Domain Modeling We formulate the training process of DRIP as a prediction task of missing domain information based on its contexts (Devlin et al. 2019; Bao et al. 2022). Our key idea is to randomly mask some of the domain-specific user embeddings (among the ones for a user's seen domains) in the input, and train the model to predict the user preference in the masked domains. That is, in the training process, we regard randomly-masked domains as the user's unseen domains, which allows our model to simulate and learn various scenarios of mapping user preference from seen domains to unseen domains. As a result, the model can capture the relations of user preferences across domains, eventually achieving the generalization capability of inferring user preferences in unseen domains from those in seen domains.

Let $\mathbf{m}_u \in \{0, 1\}^K$ denote a random masking vector for user u , where $m_{u,k} = 1$ indicates that the user embedding for domain d_k is masked. We apply the making operation to the user embedding for the seen domains (i.e., $G_{u,k} = 1$) with the probability of $p_{u,k}$. In specific, $m_{u,k}$ is drawn from a Bernoulli distribution: $m_{u,k} \sim \text{Bern}(p_{u,k})$. At the beginning of training, we set the equal masking probability for all seen domains (i.e., random masking). Then, during the training process, we gradually adjust the probability based on the domain-level preferences, i.e., $p_{u,k} \propto p(d_k|u; \Theta)$, to encourage the model to more focus on learning domains that the user is more likely to prefer (i.e, adaptive masking). The masked embedding set is represented as $\tilde{\mathbf{X}}_u = \{\tilde{\mathbf{x}}_{u,k}\}_{k=1}^K$, where its elements are obtained by

$$\tilde{\mathbf{x}}_{u,k} = (1 - G_{u,k})\mathbf{e}_{[M]} + G_{u,k}(m_{u,k}\mathbf{e}_{[M]} + (1 - m_{u,k})\mathbf{x}_{u,k}). \quad (9)$$

During the training, we use $\tilde{\mathbf{X}}_u$ (Eq. (9)) instead of \mathbf{X}_u (Eq. (3)) for the input of the multi-domain encoder, and obtain their contextualized representations. Note that we discard the case in which all seen domains are masked.

Learning Objective We train the model to predict the preference information of the masked domains (i.e., $m_{u,k} = 1$). Instead of predicting the masked embedding itself, we directly maximize the likelihood of the user’s interaction history. This makes the recommendation accuracy directly aligned with the optimization objective. Based on the negative log-likelihood of Eq. (2) and the masking process, the final loss is defined as follows:

$$\mathcal{L}_{\text{DRIP}} = - \sum_{u \in \mathcal{U}} \sum_{d_k \in \mathcal{D}} m_{u,k} \sum_{v \in \mathcal{V}_k} R_{u,v} \log[p(v|u, d_k) \cdot p(d_k|u)], \quad (10)$$

where each probability parameterized by Θ (Eq. (7), (8)) and the learning parameters Θ include the multi-domain encoder Ω , two types of projectors $\{\phi_k\}_{k=1}^K$ and $\{\psi_k\}_{k=1}^K$, and domain-preference predictor $\psi_{[S]}$.

MDRAU Recommendation

At the test phase, for each user, we construct the input representation with the user’s unseen domains masked (Eq. (4)), and calculate the domain-level and item-level preferences, $p(d_k|u; \Theta)$ and $p(v|u, d_k; \Theta)$. In the scenario where there are multiple target service domains (MDRAU-MT), the recommendation is generated by sorting the items by their score $p(v|u, d_k; \Theta) \cdot p(d_k|u; \Theta)$. Otherwise, in the scenario where we have a specific target service domain to promote (MDRAU-ST), the recommendation is generated by considering only in-domain item-level preference $p(v|u, d_k; \Theta)$.

Experiments ²

Experimental Settings

Datasets and Domain Setup We use the widely-used Amazon dataset (He and McAuley 2016; Kang et al. 2019, 2023), which consists of multiple item domains. To simulate the platform environment, we select two subsets of these domains that have been previously used in related studies (Zhu et al. 2021a). The first scenario (P1) includes Book, Movie, CD, and Game domains, and the second scenario (P2) includes Home, Health, Grocery, and Tools domains (Table 1).

Evaluation Metrics We focus on the top- K recommendation task for implicit feedback. We evaluate the recommendation accuracy of each method by using the two ranking metrics (Kang et al. 2020, 2022): Recall ($R@K$) and Normalized Discounted Cumulative Gain ($N@K$).

Compared Methods We compare DRIP with various methods from related research fields. We have modified the original methods to perform MDRAU, and the modified versions are annotated with the suffix ‘+’. The *first group* of baselines learns a recommendation model **BPRMF** (Rendle et al. 2009) while treating a union of multiple domains as a single domain (Singh and Gordon 2008). The *second group* includes multi-task learning methods (**MMOE** (Ma et al. 2018) and **PLE** (Tang et al. 2020)), which define a single task as item-level preference learning for each domain. They are widely used for RS in multi-domain cases (Cao

	Domains	#Users	#Items	#Interaction	Density
P1	Book	35,987	39,049	1,726,231	0.12%
	Movie	17,056	15,620	609,552	0.23%
	CD	5,941	9,069	211,617	0.39%
	Game	1,049	1,064	20,568	1.84%
P2	Home	10,317	7,605	127,091	0.16%
	Health	8,690	5,750	123,905	0.25%
	Grocery	4,869	2,884	71,269	0.51%
	Tools	1,804	1,422	18,040	0.70%

Table 1: Statistics of the two platform scenarios of MDRAU.

et al. 2023). For each task (i.e., domain), they employ the binary cross-entropy loss to predict user-item interactions from implicit feedback (He et al. 2017). The *third group* includes CDR methods (**EMCDR+** (Man et al. 2017) and **PTUPCDR+** (Zhu et al. 2022)) for unseen domain recommendation, which learn a mapping function for each source-target domain pair. Due to a large number of seen-unseen domain combinations, it is infeasible to directly apply them to MDRAU. For this reason, we tailor their learning task for a many-to-one mapping. The *last group* includes state-of-the-art methods (**CAT-ART+** (Li et al. 2023) and **UniCDR** (Cao et al. 2023)) that partially handle multi-domain cases. They are designed to exploit information from multiple domains to improve recommendation accuracy in each domain. For MDRAU-MT, recommendations from each unseen domain are integrated into a unified recommendation list using post-processing (e.g., normalization).

Performance Comparison for MDRAU-ST

We first compare the recommendation performance of various methods for MDRAU-ST on our two simulated platforms, P1 and P2. In this task, we set each of the domains as the target domain, and for evaluation, we consider only the users who have not tried the domain yet as test users. The results for each target domain are reported in Table 2.

We observe that DRIP consistently achieves higher recommendation performance than all the other methods in each of the target domains. Specifically, the CDR methods designed for a single pair of source-target domain based on one-to-one mapping (i.e., EMCDR+ and PTUPCDR+) show lower performance compared to the ones that can effectively handle multiple source domains based on many-to-one mapping (i.e., CAT-ART+, UniCDR, and DRIP). They are not capable of capturing the relevance among multiple source domains, whereas the multi-domain CDR methods integrate the user preferences of multiple source domains in advanced ways. In particular, the multi-domain encoder of DRIP contextualizes domain-specific user embeddings over multiple source domains with attention mechanism that is effective in capturing the inter-domain relationship; this brings a significant performance improvement for MDRAU-ST.

Furthermore, unlike the CDR methods for recommending items in a user’s unseen domain (i.e., EMCDR+, PTUPCDR+, and DRIP), the CDR baselines that aim to enhance the recommendation in the user’s seen domains (i.e., CAT-ART+ and UniCDR) have to rely on each user’s global

²More detailed experimental settings and results are in the Appendix.

Methods (P1)	Book	Movie	CD	Game
BPRMF	0.0182	0.0224	0.0371	0.0548
MMOE	0.0164	0.0310	0.0341	0.0613
PLE	0.0108	0.0296	0.0337	0.0592
EMCDR+	0.0348	0.0439	0.0610	0.0755
PTUPCDR+	0.0338	0.0427	0.0589	0.0836
CAT-ART+	0.0339	0.0445	0.0578	0.0745
UniCDR	0.0380	0.0502	0.0607	0.0823
DRIP	0.0423*	0.0517	0.0699*	0.0873*
<i>Improv.</i>	11.56%	3.13%	14.56%	4.44%
Methods (P2)	Home	Grocery	Tools	Health
BPRMF	0.0328	0.0655	0.0647	0.0415
MMOE	0.0279	0.0928	0.0678	0.0520
PLE	0.0279	0.0912	0.0697	0.0525
EMCDR+	0.0423	0.0968	0.0827	0.0639
PTUPCDR+	0.0433	0.0969	0.0836	0.0651
CAT-ART+	0.0396	0.0975	0.0758	0.0573
UniCDR	0.0436	0.1068	0.0863	0.0647
DRIP	0.0472*	0.1086*	0.0919*	0.0715*
<i>Improv.</i>	8.33%	3.23%	14.97%	9.82%

Table 2: Recommendation performance (N@20) comparison for MDRAU-ST on the platform scenario 1 (P1) and scenario 2 (P2). * denotes the improvement over the best baseline is statistically significant with $p < 0.05$, using the paired t-test.

embedding shared over multiple domains to predict the user preference for a target unseen domain. This is because they are suffering from the incapability of inferring a user’s unseen domain embedding, caused by their training process aiming at only predicting items in a user’s seen domains. In contrast, DRIP is good at inferring user embedding of a target unseen domain, conferred by masked domain modeling. To sum up, DRIP outperforms all existing methods in terms of MDRAU-ST, by (1) effectively contextualizing user embeddings over multiple domains using attention mechanism and (2) accurately inferring a user’s preference for a target unseen domain via masked domain modeling.

Performance Comparison for MDRAU-MT

We also evaluate DRIP and various methods for MDRAU-MT on two simulated platforms, P1 and P2. In this task, the set of target domains varies depending on users, because each user has different unseen domains. For each user, we assess the accuracy of recommendations for the user’s multiple unseen domains, and report the averaged accuracy over all test users as the final performance. Note that MDRAU-MT scenario differs from MDRAU-ST in that it evaluates the models’ ability to make aggregated recommendations over multiple unseen domains. In this sense, MDRAU-MT requires to properly capture *domain-level preference* with item-level preference. Table 3 presents overall performances for MDRAU-MT.

We observe that DRIP outperforms the best baseline

method more in MDRAU-MT than in MDRAU-ST. The limited MDRAU-MT performance of baselines stems from two factors. First, they focus on predicting items in single domains without considering multi-domain preferences, unlike DRIP, which captures both domain and item-level preferences. Second, most baseline methods need a post-processing step to merge recommended item lists over multiple target domains, limiting final accuracy. On the contrary, DRIP optimizes a unified model that makes recommendations for multiple target domains in an end-to-end manner, improving MDRAU-MT performance by reducing its gap from the training process. In conclusion, DRIP achieves the best performance among all the baselines with the help of its item-level preference accurately predicted for each of the target domains as well as its capability of inferring domain-level preference obtained by the training process.

Domain-Level Preference Analysis

We analyze the models’ ability to capture domain-level preference, essential for accurate recommendations in multiple unseen domains (i.e., MDRAU-MT). We compare two domain distributions obtained from (1) a user’s interaction history and (2) the recommendation list generated by each method for the user.³ Let P and Q denote the ground-truth distribution in user history and the predicted distribution, respectively. We calculate the Kullback-Leibler Divergence (KLD@ K) between the two distributions $D_{KLD}(P||Q) = \sum_i P_i \log(P_i/Q_i)$ to measure how closely the model’s prediction captures the distribution of actual domain preferences. Note that KLD only measures the domain-level accuracy, not in-domain item prediction accuracy.

In Fig. 3, we assess the KLD@ K score for the top- K recommendation and observe the following: (1) In comparison to other competing methods, BPRMF exhibits significantly better KLD scores. BPRMF treats a union of all domains as a single domain and learns the pair-wise ranking of item-level preference over the entire domains. This makes a user’s domain-level preference implicitly captured during the training process. This result can also be interpreted with the previous performance comparison for MDRAU-MT/ST. The performance of BPRMF is highly limited in MDRAU-ST due to its limited capability of capturing item-level preferences. However, BPRMF achieves comparable performance with the state-of-the-art methods in MDRAU-MT. We interpret that this improvement mainly comes from its capability of capturing domain-level preferences. (2) Other competing methods show considerably worse KLD scores compared to DRIP. As discussed earlier, they need a post-processing step to generate the unified ranking list encompassing all unseen domains (e.g., z-score normalization), which may not yield optimal recommendation accuracy. This result highlights the importance of *holistic* model training that considers both domain-level and item-level preferences in MDRAU task. (3) Among all baseline methods, DRIP achieves the best KLD scores in both scenarios. This result shows that DRIP

³In this analysis, we assume that the domain distributions in each user’s interaction history reflect the user’s actual domain-level preferences to a considerable extent.

Methods	P1		P2	
	R@20	N@20	R@20	N@20
BPRMF	0.0249	0.0256	0.0528	0.0503
MMOE	0.0188	0.0191	0.0486	0.0513
PLE	0.0190	0.0190	0.0491	0.0516
EMCDR+	0.0362	0.0374	0.0531	0.0525
PTUPCDR+	0.0367	0.0377	0.0546	0.0538
CAT-ART+	0.0395	0.0395	0.0522	0.0497
UniCDR	0.0434	0.0446	0.0588	0.0574
DRIP	0.0545*	0.0556*	0.0780*	0.0773*
Improv.	25.64%	24.55%	32.66%	34.80%

Table 3: Recommendation performance comparison for MDRAU-MT. * denotes the improvement over the best baseline is statistically significant with $p < 0.05$, using the paired t-test.

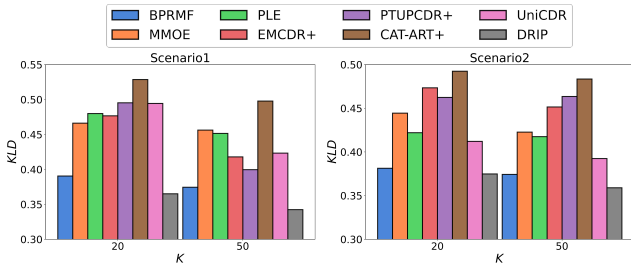


Figure 3: Domain-level preference analysis. KL-divergence scores of each method (Best viewed in color).

indeed effectively captures user’s domain-level preferences, and also supports its superior performance in MDRAU-MT.

Design Choice Analysis

We analyze alternative design choices for DRIP in Table 4 to verify the effectiveness of our design choice. We report the performance for MDRAU-MT on platform scenario P1.

First, we compare alternative training paradigms for DRIP: Single-domain learning, where all domains are treated as one, and the model is trained to maximize the likelihood of the training data; Many-to-one learning, where a single model is trained for each target domain to predict item-level preferences. To generate a unified recommendation list encompassing all domains, we apply a post-processing step. Post-processing A uses z-normalization, performing best in our test. We also consider emphasizing active domains in post-processing; specifically, for post-processing B, we multiply the ratio of the total number of interactions in each domain by the z-normalized scores and use the scores for the recommendation.

We observe that the alternative training paradigms show considerably degraded performance. Single-domain learning achieves a highly limited performance, showing the necessity of proper domain modeling in MDRAU task. Also, the many-to-one learning neglects the domain-level preference during the training, and the post-processing is applied

Designs	R@20	N@20
DRIP	0.0545	0.0556
Training Paradigm		
Single-domain learning	0.0281	0.0272
Many-to-one learning w/ post-processing A	0.0432	0.0447
Many-to-one learning w/ post-processing B	0.0409	0.0424
Domain Preference Modeling		
Uniform dist.	0.0122	0.0120
Domain activeness dist.	0.0409	0.0419
Masking Scheme		
w/o Adaptive Masking	0.0513	0.0520

Table 4: Performance comparison of different design choices. Results for MDRAU-MT on P1.

independently from the training process, which results in limited MDRAU performance. Further, in our experiments, sophisticated designs for post-processing do not bring further improvements. These results support the superiority of our training strategy that decomposes user preference into domain- and item-level preferences and jointly learns them through a unified model.

Second, we compare ablations for the domain-level preference modeling. Instead of estimating personalized domain-level preference, they use globally fixed distributions: the uniform and domain activeness distribution (the latter assumes users prefer more active domains). Both fixed domain preference approaches yield suboptimal recommendation performance. This result supports the effectiveness of our strategy that models the domain-level preference for each individual user considering their different preference.

Lastly, we provide the results without the adaptive masking. For the ablation, we use random masking with the same masking ratio as adaptive masking. The adaptive masking brings slight improvements to the final performance, indicating that our masking strategy is well-aligned with our masked domain modeling.

Conclusion

This paper highlights the importance of MDRAU task based on its practical advantages in multi-domain platforms. We propose DRIP, a new framework to provide accurate unseen domain recommendations to attract users into new service domains that they have not interacted with yet. The DRIP decomposes user preference into domain-level preference and in-domain item-level preference and then jointly learns them via a unified model with the help of a training strategy based on masked domain modeling. We conduct extensive comparisons with a wide range of CDR methods. DRIP consistently achieves superior performance compared to all competing methods in both cases of a specific target domain (MDRAU-ST) and multiple target domains (MDRAU-MT). We expect that DRIP can enhance the user experience by fostering diverse and serendipitous discoveries and potentially promote the influx of new users to each service domain, benefiting providers in multi-domain service platforms.

Acknowledgements

This work was supported by the IITP grant funded by the MSIT (No.2018-0-00584, No.2019-0-01906) and the NRF grant funded by the MSIT (No.2020R1A2B5B03097210, No.RS-2023-00217286).

References

- Bao, H.; Dong, L.; Piao, S.; and Wei, F. 2022. BEiT: BERT Pre-Training of Image Transformers. In *International Conference on Learning Representations*.
- Cao, J.; Li, S.; Yu, B.; Guo, X.; Liu, T.; and Wang, B. 2023. Towards Universal Cross-Domain Recommendation. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, 78–86.
- Cao, J.; Lin, X.; Cong, X.; Ya, J.; Liu, T.; and Wang, B. 2022a. DisenCDR: Learning Disentangled Representations for Cross-Domain Recommendation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 267–277.
- Cao, J.; Sheng, J.; Cong, X.; Liu, T.; and Wang, B. 2022b. Cross-Domain Recommendation to Cold-Start Users via Variational Information Bottleneck. In *2022 IEEE 38th International Conference on Data Engineering (ICDE)*, 2209–2223.
- Devlin, J.; Chang, M.-W.; Lee, K.; and Toutanova, K. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171–4186.
- Fu, W.; Peng, Z.; Wang, S.; Xu, Y.; and Li, J. 2019. Deeply fusing reviews and contents for cold start users in cross-domain recommendation systems. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 94–101.
- He, R.; and McAuley, J. 2016. Ups and Downs: Modeling the Visual Evolution of Fashion Trends with One-Class Collaborative Filtering. In *Proceedings of the 25th International Conference on World Wide Web, WWW '16*, 507–517. Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee. ISBN 9781450341431.
- He, X.; Liao, L.; Zhang, H.; Nie, L.; Hu, X.; and Chua, T.-S. 2017. Neural collaborative filtering. In *Proceedings of the 26th international conference on world wide web*, 173–182.
- Hu, G.; Zhang, Y.; and Yang, Q. 2018. CoNet: Collaborative Cross Networks for Cross-Domain Recommendation. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, 667–676.
- Kang, S.; Hwang, J.; Kweon, W.; and Yu, H. 2020. DE-RRD: A knowledge distillation framework for recommender system. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, 605–614.
- Kang, S.; Hwang, J.; Lee, D.; and Yu, H. 2019. Semi-supervised learning for cross-domain recommendation to cold-start users. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 1563–1572.
- Kang, S.; Kweon, W.; Lee, D.; Lian, J.; Xie, X.; and Yu, H. 2023. Distillation from Heterogeneous Models for Top-K Recommendation. In *Proceedings of the ACM Web Conference 2023*, 801–811.
- Kang, S.; Lee, D.; Kweon, W.; Hwang, J.; and Yu, H. 2022. Consensus Learning from Heterogeneous Objectives for One-Class Collaborative Filtering. In *Proceedings of the ACM Web Conference 2022*, 1965–1976.
- Koren, Y.; Bell, R.; and Volinsky, C. 2009. Matrix Factorization Techniques for Recommender Systems. *Computer*, 42(8): 30–37.
- Li, C.; Xie, Y.; Yu, C.; Hu, B.; Li, Z.; Shu, G.; Qie, X.; and Niu, D. 2023. One for All, All for One: Learning and Transferring User Embeddings for Cross-Domain Recommendation. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, 366–374.
- Ma, J.; Zhao, Z.; Yi, X.; Chen, J.; Hong, L.; and Chi, E. H. 2018. Modeling task relationships in multi-task learning with multi-gate mixture-of-experts. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, 1930–1939.
- Man, T.; Shen, H.; Jin, X.; and Cheng, X. 2017. Cross-Domain Recommendation: An Embedding and Mapping Approach. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, 2464–2470.
- Rendle, S.; Freudenthaler, C.; Gantner, Z.; and Schmidt-Thieme, L. 2009. BPR: Bayesian Personalized Ranking from Implicit Feedback. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, 452–461.
- Singh, A. P.; and Gordon, G. J. 2008. Relational learning via collective matrix factorization. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, 650–658.
- Tang, H.; Liu, J.; Zhao, M.; and Gong, X. 2020. Progressive layered extraction (ple): A novel multi-task learning (mtl) model for personalized recommendations. In *Proceedings of the 14th ACM Conference on Recommender Systems*, 269–278.
- Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L. u.; and Polosukhin, I. 2017. Attention is All you Need. In *Advances in Neural Information Processing Systems 2017*, volume 30.
- Wang, X.; He, X.; Wang, M.; Feng, F.; and Chua, T.-S. 2019. Neural Graph Collaborative Filtering. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 165–174.
- Zhu, F.; Chen, C.; Wang, Y.; Liu, G.; and Zheng, X. 2019. Dtdcr: A framework for dual-target cross-domain recommendation. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 1533–1542.

- Zhu, F.; Wang, Y.; Chen, C.; Liu, G.; and Zheng, X. 2020. A Graphical and Attentional Framework for Dual-Target Cross-Domain Recommendation. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*, 3001–3008.
- Zhu, F.; Wang, Y.; Chen, C.; Zhou, J.; Li, L.; and Liu, G. 2021a. Cross-Domain Recommendation: Challenges, Progress, and Prospects. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, 4721–4728.
- Zhu, F.; Wang, Y.; Zhou, J.; Chen, C.; Li, L.; and Liu, G. 2023. A Unified Framework for Cross-Domain and Cross-System Recommendations. *IEEE Transactions on Knowledge and Data Engineering*, 35(2): 1171–1184.
- Zhu, Y.; Ge, K.; Zhuang, F.; Xie, R.; Xi, D.; Zhang, X.; Lin, L.; and He, Q. 2021b. Transfer-Meta Framework for Cross-Domain Recommendation to Cold-Start Users. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1813–1817.
- Zhu, Y.; Tang, Z.; Liu, Y.; Zhuang, F.; Xie, R.; Zhang, X.; Lin, L.; and He, Q. 2022. Personalized transfer of user preferences for cross-domain recommendation. In *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, 1507–1515.