Domain-Level Disentanglement Framework Based on Information Enhancement for Cross-Domain Cold-Start Recommendation

Nian Rong[*1], Fei Xiong[*1†], Shirui Pan[2], Guixun Luo[3], Jia Wu[4], Liang Wang[5]

[1]School of Electronic and Information Engineering, Beijing Jiaotong University
[2]Griffith University
[3]School of Computer Science & Technology, Beijing Jiaotong University
[4]Macquarie University
[5]Northwestern Polytechnical University
{nianrong, xiongf}@bjtu.edu.cn, s.pan@griffith.edu.au,
gxluo@bjtu.edu.cn, jia.wu@mq.edu.au, liangwang0123@gmail.com

Abstract

Recommender systems in various applications often encounter the challenge of cold-start, which refers to how to provide recommendations for completely new users. Cross-domain recommendation offers a solution to address this cold-start issue by leveraging user interaction information from other domains and providing recommendations for users in the target domain. However, applying the classic two-tower model in cross-domain scenarios for zero interaction cold-start users proves challenging, and most existing cross-domain cold-start recommendation models adopt an embedding-mapping framework that lacks end-to-end efficiency. In this paper, we propose a generalized framework that Domain-level Disentanglement framework based on information enhancement for Cross-domain Cold-start Recommendation. On one hand, we achieve deep utilization of domain-level information through independent extraction of domain knowledge and fusion using heuristic strategies. On the other hand, our model is incorporated with an information enhancement network based on user attention and a user personalized adaptor. We introduce measures to assess user variability and immutability in cross-domain recommendation, aiming to eliminate inter-domain bias and highlight individual user preferences. Experimental results on widely used cross-domain recommendation datasets demonstrate that our proposed model outperforms state-of-the-art methods, validating its effectiveness.

1 Introduction

Recommender systems(Koren, Bell, and Volinsky 2009; Mnih and Salakhutdinov 2007; Rendle et al. 2012) are extensively employed in various real-world applications, such as e-commerce, audio and video software, and social media platforms, to enable users to independently discover relevant contents and products, thereby enhancing user experience and generating substantial economic and practical value. With the increasing number of user interactions across diverse domains, the exploration of leveraging information between different domains has become a subject of extensive contemplation. Cross-Domain Recommendation (CDR) (Zang et al. 2022; Xu et al. 2021; Kamani, Kumar, and Kagita 2023)has gained significant attention in recent years. For instance, in movie-book cross-domain scenario, there is typically an overlapping user named Alice. Leveraging Alice's information to establish inter-domain relationships, we utilize a combination of book and movie domain data to enhance recommendation performance in both domains. However, in the cold-start scenario, when dealing with a new user like Bob who lacks interactive information from one domain (e.g. movie-domain), only the interactive information from the other domain (e.g. book-domain) is available(Wei et al. 2021).

We differentiate between cold-start cross-domain recommendation and regular cross-domain recommendation models in this study because many existing cross-domain recommendation(Zhu et al. 2020b; Cao et al. 2023; Zhu et al. 2019; Chen et al. 2023) methods require overlapping entities to integrate information from different domains, with a focus on enhancing the recommendation performance of the target domain. However, these models are not capable of effectively addressing the zero-interaction cold-start scenario as outlined in this paper, i.e., providing recommendations to cold-start users in the absence of historical interaction data within the target domain. On the other hand, among the existing approaches applicable to cold-start cross-domain recommendation, the embedding mapping-based methods(Man et al. 2017) severely hampers recommendation efficiency and lacks an end-to-end framework. Similarly, multi-domain co-training methods(Hu, Zhang, and Yang 2018; Jiang et al. 2016; Li and Tuzhilin 2021) encounter the challenge of ineffective information filtration and the introduction of unnecessary negative transfer knowledge during cross-domain processing. Although some methods have acknowledged the bias in user preferences across different domains(Zhao et al. 2023; Liu et al. 2022), such as individuals who enjoy suspense novels may not necessarily appreciate suspense movies, no effective and adaptive solution has been proposed to address this issue.

Considering the potential correlation between different

[*]These authors contributed equally.
[†]Corresponding author.

domains and their respective cross-domain features, we posit that it would be efficacious to propose an integrated framework for mining and fusing domain-level information, while concurrently emphasizing personalization of users across different domains is advantageous. To this end, we propose D2C2R, **D**omain-level **D**isentanglement framework based on information enhancement for **C**ross-domain **C**old-start **R**ecommendation.

We propose a domain meta-knowledge extraction module and a cross-domain fusion module. In the meta-knowledge extraction module, we employ Variational Graph Auto-Encoders (VGAE) and Fusion Variational Graph Auto-Encoders (FVGAE) to independently construct domain-specific information and extract shared information from the fusion domain. The cross-domain fusion module incorporates an information enhancement network that considers users' dependence on shared information while achieving domain information fusion through its own bias attention mechanism. Relevant domains will be connected and mapped in subsequent stages, employing our proposed heuristic fusion strategy. Overall, the proposed method leverages independent domain information to constrain the extraction of fusion domain knowledge, and employs an integrated fusion approach in the cross-domain fusion module to obtain a comprehensive representation of user behaviors within a specific domain. In contrast to conventional cross-domain mapping approaches, we propose a personalized user adaptation method that leverages contrastive learning loss functions to minimize disparities in user representations.

The contributions of this paper are as follows:

- We propose a novel framework for cross-domain cold-start recommendation, which comprises a fusion graph variational encoder, an information enhancement network based on user attention, and a matching heuristic cross-domain fusion network that leverages domain-level disentanglement to capture the correlation between domains.
- We propose a personalized adaptation optimizer that effectively aligns diverse dimensional representations of individual users, mitigates bias, and emphasizes the importance of personalization.
- The proposed model demonstrates superior performance compared to state-of-the-art approaches on three widely used real datasets, and we substantiate its effectiveness through ablation experiments.

2 Methodology

2.1 Domain-meta Knowledge Extraction

We initially establish a cross-domain cold-start scenario, encompassing two distinct domains, namely X domain and Y domain. We provide entity recommendations in Y domain for each cold-start user who has interactions in X domain but no interaction in Y domain. The respective interaction graph of the two domains $G^X = \{S_u^X, S_v^X, E^X\}$, $G^Y = \{S_u^Y, S_v^Y, E^Y\}$is constructed for the initial two domains based on the user set S_u^X, S_u^Y, item set S_v^X, S_v^Y, and interaction set E^X, E^Y. The initial embeddings of users and items are denoted as U^X, U^Y and V^X, V^Y.

Note that, we merge the interaction sets of the two domains into one set when obtaining the initialization of the fusion graph. The user set $S_u^O = S_u^X \cup S_u^Y$, the item set $S_v^O = S_v^X \cup S_v^Y$, and the interaction set $E^O = E^X \cup E^Y$ are constructed to form the initial fusion domain interaction graph G^O. Instead of separately building initial embeddings for users in the fusion graph, we directly use the mean of initial embeddings from both independent domains to construct a portion of the initial embedding in the fusion graph. For each user, we compute the mean of the initial user representations of the two independent domains $u_i^O = \{(u_i^X + u_i^Y)/2\}$. For items, there is $V^O = V^X \oplus V^Y$.

The connectivity matrix A can be derived from the interaction set in both domains i.e. $\mathbf{A}^X \in \mathbb{R}^{N_U^X \times N_V^X}$, $\mathbf{A}^Y \in \mathbb{R}^{N_U^Y \times N_V^Y}$, $\mathbf{A}^O \in \mathbb{R}^{N_U^O \times N_V^O}$. It consists of 0,1, where 1 indicates that the user has an interaction with the item. The variables N_U^X, N_V^X, N_U^Y, N_V^Y and N_U^O, N_V^O represent the quantities of users and items in each respective domain.

Independent Information Construction In the context of recommendation scenarios, Variational Graph Autoencoders(VGAE) based on bipartite graph generation excels at capturing comprehensive user-item interaction information with minimal links. The fundamental element of VGAE is the Graph Convolutional Network (GCN). We initially aggregate neighboring information through the first layer in the bipartite user-item recommendation graph to generate the intermediate representation information

$$\hat{U}^X = f\left[\left(\mathbf{A}^X\right)^T U^X W_u^X\right] \tag{1}$$

for users. W_u^X is the parameter to be learned. $f(\cdot)$ is LeakyReLU function.

The final representations of the user node vectors are not obtained deterministically through GCN, but rather sampled from a multidimensional Gaussian distribution. The mean and standard deviation of this distribution are derived from the next two convolutional layers and the intermediate representation of the users

$$\begin{aligned} \mu_u^X &= f\left[\left(f\left(\mathbf{A}^X \hat{U}^X \hat{W}_{u,\mu}^X\right) \oplus U^X\right) W_{u,\mu}^X\right] \\ \sigma_u^X &= f\left[\left(f\left(\mathbf{A}^X \hat{U}^X \hat{W}_{u,\sigma}^X\right) \oplus U^X\right) W_{u,\sigma}^X\right] \end{aligned} \tag{2}$$

where $\oplus$ denotes the contact operation. Respectively, through the structure of the two-layer GCN, we obtain the mean and standard deviation of the Gaussian distribution by $\hat{U}^X$ that aggregating neighbor information. With the given mean and standard deviation, we can uniquely determine the Gaussian distribution(Kipf and Welling 2016) by resampling techniques commonly used in neural networks, as shown in the following equation, which includes $\varepsilon \sim N(0,1)$:

$$z_u^X = \mu_u^X + \varepsilon \odot \sigma_u^X \tag{3}$$

where $\odot$ represents the element-wise dot product. The user representations in the X domain are taken as an illustrative example in the formula. Likewise, we employ analogous methods for handling Y domain and we also get Z_u^Y, Z_v^X, Z_v^Y.

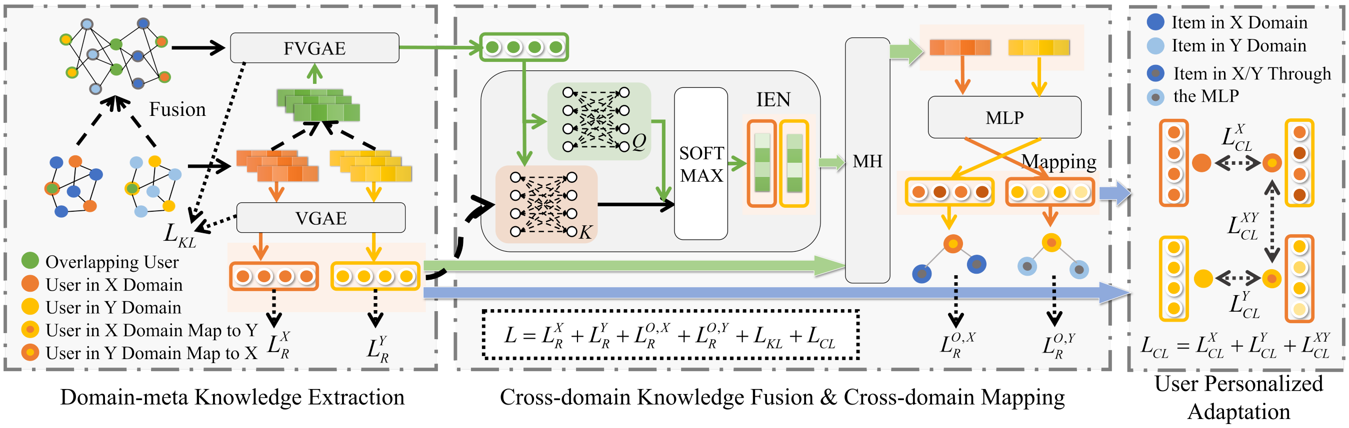

Figure 1: An overview of the D2C2R model. a) Domain-meta knowledge extraction(DKE) extracts information from both independent and fused domains. b) The cross-domain knowledge fusion (CDKF) component leverages the information enhancement network(IEN) to fuse independent domain and fusion domain information, enabling cross-domain fusion network for matching heuristics (MH). c) The User Personalized Adaptation (UPA), integrated as a loss component in contrastive learning, refines the user representations by emphasizing personalized features.

Shared Knowledge Acquisition Our objective is to effectively extract shared information from the fusion graph in the training process of two independent domains. In addition to incorporating information from these domains during initialization, we further fuse their respective means and standard deviations throughout the training process. Specifically, after aggregating neighbor information of the fusion graph users, we utilize the means and standard deviations of each independent domain as guidance for constructing corresponding values in the fusion domain. Since sparsity of the two domains is different, a hyperparameter β_{XY} is employed to regulate the extent of capturing initial fusion information in both X and Y domains simultaneously

$$\begin{aligned}\mu_u^O =& f\left\{\left[\left(f\left(A^O\hat{U}^O\hat{W}^O_{u,\mu}\right)\right)\right.\right.\\ &\left.\left.\oplus\left(\left(\beta_{XY}\odot\mu_u^X\right)+\left(\left(1-\beta_{XY}\right)\odot\mu_u^Y\right)\right)\right]W^O_{u,\mu}\right\}\\ \sigma_u^O =& f\left\{\left[\left(f\left(A^O\hat{U}^O\hat{W}^O_{u,\sigma}\right)\right)\right.\right.\\ &\left.\left.\oplus\left(\left(\beta_{XY}\odot\sigma_u^X\right)+\left(\left(1-\beta_{XY}\right)\odot\sigma_u^Y\right)\right)\right]W^O_{u,\sigma}\right\}\end{aligned} \tag{4}$$

where $Z_u^O \sim \mathcal{N}\left(\mu_u^O, \left[diag\left(\sigma_u^O\right)\right]^2\right)$. It should be noted that the construction of item fusion domain does not involve the parameter β_{XY}, but it necessitates a stitching process.

2.2 Cross-domain Knowledge Fusion

User-based Information Enhancement Networks It is not feasible for every user to input 100% of information into the extraction module for inter-domain mapping relations. We propose an information enhancement network that utilizes the attention mechanism to evaluate the actual impact of node representations from both independent and fusion domains on different users, as well as measure the real influence of fusion domain on independent domain.

$$Att\_v\left(Q,K\right) = \text{Softmax}\left(QK^T {d_k}^{-1/2}\right) \tag{5}$$

In this model, $Q = Z_u^O$ represents the target shared information for which we need to query representation, while $K^T = \left(\left(Z_u^X\right)^T \oplus \left(Z_u^O\right)^T\right)$ encompasses all independent and shared information of a certain user pertaining to the queried item. d_k is the dimension of Q, K and is used for gradient stabilization during training. This algorithm does not directly incorporate the attention mechanism, as it aims to capture the users' individual reliance on shared domain information in order to enhance the personalized characteristics of the cross-domain model and mitigate negative transfer effects resulting from merging irrelevant information. After calculating the weight of the shared information in X domain, we obtain the final value $Z_u^{O,X}$ of the shared information for the corresponding user's dependency in X domain.

$$\tilde{Z}_u^{O,X} = Att\_v\left(Z_u^O, \left(\left(Z_u^X\right)^T \oplus \left(Z_u^O\right)^T\right)^T\right) \odot Z_u^O \tag{6}$$

From the perspective of domains, the two domains exhibit distinct dependencies on shared information. Considering micro individual users, each user also possesses unique personalized characteristics. This network exemplifies the augmentation of information from both macro and micro perspectives.

Matching Heuristics Cross-domain Fusion Network We propose a heuristic cross-domain fusion network that establishes connections between the nodes' independent and fused domains by capturing the sub and dot product values of the fused domains. Essentially, concatenation, subtractions, and dot products all quantify the correlation between two domains. This network extends the evaluation of information enhancement networks and serving as a module for fusion in domain-level disentanglement mapping. The extraction-fusion process enables a generalized disentanglement for shared information, facilitating the maximization

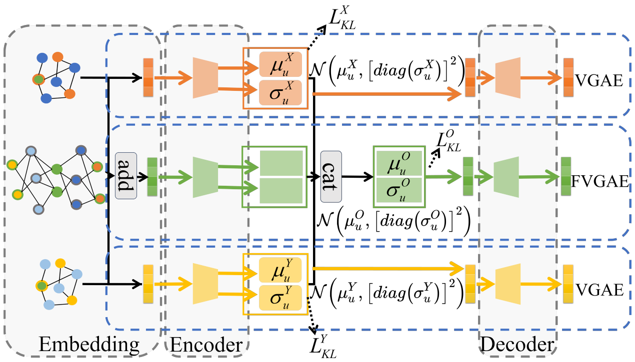

Figure 2: The two Variational Graph Autoencoders (VGAE) independently capture domain-specific information in two domains, while the Fusion Variational Graph Autoencoder (FVGAE) extracts fusion domain information during graph construction and variational parameter generation.

of independent domain characteristics for each node and the construction of shared information at the domain level. Subsequently, multilayer perceptron (MLP) is employed to train the mapping relationships using enhanced information representation in this module, thereby obtaining representations of the user mapping to the other domain.

$$
\begin{aligned}
T_u^Y &= MLP_u^{X2Y}\left[Z_u^X \oplus \tilde{Z}_u^{O,X} \oplus \left(Z_u^X - \tilde{Z}_u^{O,X}\right)\right.\\
&\qquad \left.\oplus \left(Z_u^X \odot \left(\tilde{Z}_u^{O,X}\right)^T\right)\right]\\
T_u^X &= MLP_u^{Y2X}\left[Z_u^Y \oplus \tilde{Z}_u^{O,Y} \oplus \left(Z_u^Y - \tilde{Z}_u^{O,Y}\right)\right.\\
&\qquad \left.\oplus \left(Z_u^Y \odot \left(\tilde{Z}_u^{O,Y}\right)^T\right)\right]
\end{aligned} \tag{7}
$$

Given the subsequent alignment and other reconstruction representations, we need the MLP to satisfy, $input \in \mathbb{R}^{4d}$, $output \in \mathbb{R}^{d}$. Item representations differ from user representations. In domain X, the item representation is solely derived from domain X information. The alignment of items and users in the same dimension is achieved by applying similar operations for cross-domain mapping on the items involved in cross-domain recommendation.

2.3 User Personalized Adaptation

In the context of cross-domain recommendation for overlapping users, each user possesses dual representations in X domain and Y domain. Considering potential biases in information across different domains, it becomes crucial to address both individual consistency within the same users and inter-domain bias. To achieve this goal, we propose employing contrastive learning to construct a personalized adaptive module, with the aim of mitigating domain-level bias while facilitating user-level adaptive expression. We set three user personalized adaptation relations as follows $Z_u^X \leftrightarrow T_u^X$, $T_u^Y \leftrightarrow Z_u^Y$, $T_u^Y \leftrightarrow T_u^X$, to underscore the consistent user invariance across X and Y domains, as well as to harmonize the user embedding representations within different domains. For each user, it is implied that there exists a relationship between t_{ui}^X and t_{ui}^Y:

$$
l\left(t_{ui}^X, t_{ui}^Y\right) = -\log \frac{\exp\left(\frac{sim\left(t_{ui}^X, t_{ui}^Y\right)}{\tau}\right)}{\sum_{k=1}^{M} 1_{[k\neq i]} \exp\left(\frac{sim\left(t_{ui}^X, t_{uk}^Y\right)}{\tau}\right)} \tag{8}
$$

where $sim\left(t_{ui}^X, t_{ui}^Y\right) = t_{ui}^X \left(t_{ui}^Y\right)^T \left(\left\|t_{ui}^X\right\| \left\|t_{ui}^Y\right\|\right)^{-1}$ and $1_{[k\neq i]}$ denotes 1 when $k \neq i$ and 0 otherwise. During training, it is important to note that this component solely focuses on overlapping users, where M represents the number of randomly sampled participants from all overlapping users. For user i in domain X, all other users in domain Y are considered as negative samples.

$$
L_{CL}^{XY} = \frac{1}{2N_U^O} \sum_{i=0}^{N_U^O} \left[l\left(t_{ui}^X, t_{ui}^Y\right) + l\left(t_{ui}^Y, t_{ui}^X\right)\right] \tag{9}
$$

This model focuses on three pairs of contrasting relations, aiming to streamline the representations of the same users within and across domains, thereby enhancing personalization for individual users while mitigating domain-specific biases.

$$
L_{CL} = L_{CL}^X + L_{CL}^Y + L_{CL}^{XY}
$$

2.4 Loss Construction

During vector construction, the posterior probability distribution of the embedding vectors for domain X is determined by unique A^X, U^X, and V^X. We assume that the posterior probability $q\left(Z_u^X | U^X, V^X, A^X\right)$ conforms to a Gaussian distribution. During prediction, accurate reflection of the original interaction graph by nodes is desirable. Specifically, VGAE and FVGAE are employed to generate a unique posterior probability distribution. Our objective is to maximize the likelihood of reconstructing the original graph in order to enhance node representation fidelity in capturing underlying interactions. According to maximum likelihood estimation method, this problem can be equivalently expressed as $\max L = \max \log p\left(U^X, V^X, A^X\right)$.

$$
\begin{aligned}
&\log p\left(U^X, V^X, A^X\right) \geq ELBO\\
&= -\left(\underbrace{KL\left[q\left(Z^X | U^X, V^X, A^X\right) || p\left(Z^X\right)\right]}_{L_{KL}^X}\right.\\
&\left.\underbrace{-\mathbb{E}_{q\left(Z^X | U^X, V^X, A^X\right)}\left[\log\left(p\left(U^X, V^X, A^X | Z^X\right)\right)\right]}_{L_R^X}\right)
\end{aligned} \tag{10}
$$

The evidence lower bound (ELBO) with respect to maximization can be translated to minimization $L_{KL}^X + L_R^X$. Where L_{KL}^X has the Kullback-Leibler (KL) divergence formula obtained as follows:

$$
\begin{aligned}
L_{KL}^X &= \frac{\beta_{uv}}{2}\left[-\log\left(\sigma_u^X\right) + \left(\sigma_u^X\right)^2 + \left(\mu_u^X\right)^2 - 1\right]\\
&+ \frac{\left(1-\beta_{uv}\right)}{2}\left[-\log\left(\sigma_v^X\right) + \left(\sigma_v^X\right)^2 + \left(\mu_v^X\right)^2 - 1\right]
\end{aligned} \tag{11}
$$

In the construction of an independent domain graph, we hope that the final results of the independently trained variational encoder can well represent the independent domain features, and finally reconstruct the edge matrix by the following formula $p\left(\hat{A}^X_{ij}=1|z^X_{ui},z^X_{vj}\right)=z^X_{ui}\left(z^X_{vj}\right)^T$. By constructing the edges using the dot product operation, the final loss can be evaluated through binary cross-entropy to measure the disparity between the generated and actual outcomes. The fusion graph is uniquely determined by A^X, A^Y, U^X, U^Y and V^X, V^Y. Although the generation process of FVGAE differs from that of general VGAE, the fundamental derivation process remains consistent due to the independence. The distinction lies in the fact that FVGAE incorporates a subsequent cross-domain knowledge fusion module into the reconstruction process, which entails reconstructing the user and item interaction graph of the target domain after cross-domain mapping as the equation $p\left(\hat{A}^{O,X}_{ij}=1|t^X_{ui},t^X_{vj}\right)=t^X_{ui}\left(t^X_{vj}\right)^T$. Similarly, binary cross-entropy is employed as our chosen loss function.

$$\begin{aligned}
&L_R^{O,X}+L_R^{O,Y}= \\
&-\left[A^X\log\left(\hat{A}^{O,X}\right)+\left(1-A^{O,X}\right)\log\left(1-\hat{A}^{O,X}\right)\right] \\
&-\left[A^Y\log\left(\hat{A}^{O,Y}\right)+\left(1-A^Y\right)\log\left(1-\hat{A}^{O,Y}\right)\right]
\end{aligned} \tag{12}$$

The final overall objective function is

$$L=L^X_{KL}+L^Y_{KL}+L^O_{KL}+L^X_R+L^Y_R+L^{O,X}_R+L^{O,Y}_R+L_{CL}$$

3 Experiments

3.1Experimental Settting

Dataset Based on previous research, we use three cross-domain recommendation datasets and preprocessing settings to create cold-start scenarios across domains. Specifically, the Amazon dataset(Dhage et al. 2023) includes ToysG-VideoG, Music-Movie, Phones-Ele pairs. After preprocessing, the minimum number of interactions per user is set to 5 or higher. The number of overlapping users in each scenario is shown in Table 1. This study focuses specifically on the cold-start (zero shot) scenario(Huan et al. 2022), where 10% of users from each domain are selected as cold-start users and only the interaction information from one domain is retained for testing. For example, 10% of users are recommended from ToysG to VideoG and another 10% from VideoG to ToysG, while the rest are used for training purposes. For each actual interaction in the set, 999 negative samples are randomly chosen from the target domain for calculation and ranking of 1000 records. The performance of top-k recommendation is evaluated using three widely adopted metrics(Fouss and Saerens 2008).The experiments were conducted using a computing infrastructure equipped with a NVIDIA GeForce RTX 4090 includes PyTorch 1.13.1 with CUDA 11.6 support.

Baseline To validate the efficacy of the proposed model in the cross-domain cold-start scenarios, we compare it against state-of-the-art recommendation models. The two recommendation models for single domain (1) **BPRMF**(Rendle et al. 2012) and (2) **NGCF**(Wang et al. 2019) were selected. We have chosen multiple cross-domain cold-start recommendation models that exhibit exceptional performance to comparative analysis: (1) **EMCDR**(Man et al. 2017) (2) **DCDCSR**(Zhu et al. 2020a) (3) **SSCDR**(Kang et al. 2019) (4) **TMCDR**(Zhu et al. 2021) (5) **SA-VAE**(Salah, Tran, and Lauw 2021) (6) **CDRIB**(Cao et al. 2022b) (7)**UniCDR**(Cao et al. 2023). In our experiments, all models were trained individually to obtain their respective optimal parameters in the three scenarios. To ensure a fair comparison, we set the learning rate to 0.001, batch size to 1024, number of mapping neural network layers to 3, and node embedding dimension to 128 for these models.

scenarios	user	item	rating	overlapping user	overlapping item
ToysG	16,412	11,924	167,597	1,380	6,773
VideoG	24,303	10,672	231,780		6,667
Music	5,541	3,568	64,706	2,431	3,545
Movie	123,960	50,052	1,697,533		32,135
Phones	157,212	48,186	1,128,437	83,232	47,758
Ele	728,719	160,052	6,739,590		136,249

Table 1: Data statistics of three scenarios of CDR

3.2 Comparative Experiment

The experimental results for the ToysG-VideoG scenario are presented in Table 2. These show that NGCF outperforms BPRMF in two single-domain recommendation scenarios, indicating that integrating graph neural networks enhances user-item interaction characterization. Overall, cross-domain methods generally outperform single-domain cold-start approaches, suggesting the benefits of leveraging information from other domains to enhance cold-start recommendations. DCDCSR and SSCDR outperform EMCDR by incorporating deeper mining of user representations and aggregating neighbor information respectively. The application of VAE in SA-VAE yields notably favorable outcomes. CDRIB introduces the concept of an information bottleneck and improves performance. The UniCDR model employs a mask mechanism to enhance data, contributing to improved recommendation performance at the cost of significant time overhead. Our D2C2R model exhibits significant enhancement in cold-start recommendation performance across three datasets by placing greater emphasis on incorporating domain-specific characteristic information and exploring associations between domains for more effective utilization in cross-domain cold-start scenarios. Unlike conventional cross-domain mapping recommendation frameworks, D2C2R focuses on personalized adaptation representation for users by leveraging a personalized adaptor which highlights both user invariance and variability between domains-ultimately leading to improved recommendation performance.

model	ToysG						VideoG					
	mrr	ndcg@5	ndcg@10	hr@1	hr@5	hr@10	mrr	ndcg@5	ndcg@10	hr@1	hr@5	hr@10
BPRMF	1.98±0.13	1.12±0.11	1.78±0.15	0.57±0.07	1.77±0.23	3.64±0.04	2.37±0.21	1.69±0.13	2.28±0.31	0.67±0.12	2.74±0.41	4.33±0.54
NGCF	2.87±0.12	2.01±0.06	2.68±0.17	1.01±0.13	3.09±0.28	5.07±0.26	3.66±0.30	2.79±0.24	3.74±0.14	1.44±0.08	4.13±0.36	7.30±0.41
EMCDR	2.23±0.09	1.40±0.16	2.08±0.07	0.67±0.05	2.31±0.20	4.49±0.42	4.09±0.35	3.18±0.19	4.16±0.08	1.55±0.24	4.67±0.03	7.85±0.51
DCDCSR	2.31±0.16	1.56±0.10	1.69±0.28	0.70±0.03	2.34±0.13	3.58±0.29	3.37±0.30	2.65±0.19	3.35±0.17	0.98±0.06	4.21±0.42	6.58±0.32
SSCDR	1.80±0.11	1.07±0.02	1.43±0.06	0.30±0.01	1.74±0.16	3.28±0.20	2.40±0.31	1.76±0.05	2.53±0.34	0.58±0.02	3.09±0.18	5.47±0.34
TMCDR	2.43±0.21	1.87±0.10	2.55±0.23	0.52±0.02	3.21±0.17	5.22±0.31	3.79±0.08	3.06±0.22	4.30±0.08	1.40±0.13	4.87±0.32	8.40±0.25
SA-VAE	2.66±0.09	1.74±0.05	2.69±0.12	1.02±0.04	3.48±0.23	5.76±0.28	3.62±0.20	2.58±0.16	3.64±0.30	1.05±0.04	4.09±0.30	7.38±0.39
CDRIB	4.52±0.22	3.51±0.42	4.34±0.16	2.04±0.09	4.97±0.71	7.72±0.27	5.61±0.18	4.69±0.20	6.08±0.11	2.17±0.20	7.23±0.25	11.61±0.29
UniCDR	4.35±0.20	3.41±0.32	4.49±0.34	1.79±0.09	5.46±0.40	7.59±0.51	6.27±0.19	5.41±0.39	7.25±0.27	2.24±0.30	8.25±0.78	11.90±0.54
D2C2R	**4.68±0.31**	**3.83±0.35**	**4.68±0.48**	**2.13±0.18**	**5.59±0.63**	**8.42±0.98**	**6.35±0.36**	**5.37±0.57**	**6.93±0.36**	**2.31±0.25**	**8.41±0.94**	**13.62±0.64**

Table 2: Experimental results (%) for the ToysG-VideoG scene. Optimal results are indicated by bold font while suboptimal ones are underlined.

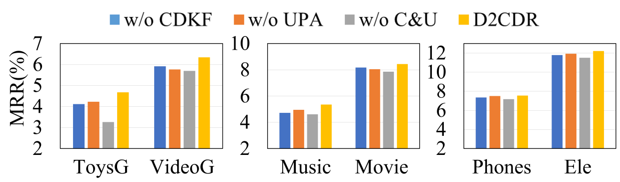

Figure 3: MRR values(%) of ablation experiments result.

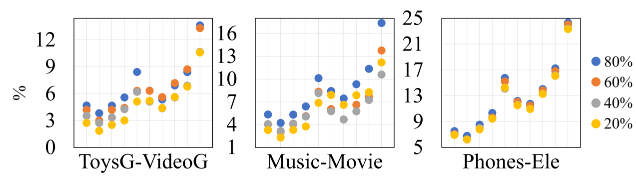

Figure 4: The evaluation metrics (%) with varying degrees of overlap users are as follows, from left to right is mrr, ndcg@5, ndcg@10, hr@5, hr@10 in domain X and mrr, ndcg@5, ndcg@10, hr@5, hr@10 in domain Y.

3.3 Ablation Study

To validate the efficacy of each module in our model, we evaluated three variations of D2C2R: one without CDKF; another without UPA; and a third without C&U, which does not preserve either of the mentioned modules. The key evaluation of MRR is presented in Figure 3, while the remaining indicators can be found in Table 3. Based on experimental results, it is evident that the comprehensive D2C2R model outperforms all three variants, highlighting the effectiveness of its constituent components. Employing a user-based information augmentation network and a cross-domain heuristic decision fusion strategy in CDKF proves superior to direct mapping as it effectively leverages disparities between independent and shared information across domains for establishing inter-domain relationships. Our proposed user personalized adaptor effectively aligns users, accentuates their individuality, mitigates domain-induced bias, and enhances cross-domain cold-start recommendation through three contrasting learning functions. Moreover, compared to variant w/o C&U, variants w/o CDKF and w/o UPA exhibit relatively superior results.

3.4 Analysis of the Impact of Overlap User Proportion

The core of cross-domain recommendation is to extract the relationship between two domains by identifying overlapping users. We conducted experiments to observe how different proportions of overlapping users impact the performance. Dividing three real cross-domain datasets into training, validation, and testing sets, we used 80%, 60%, 40%, and 20% of overlap users for training respectively, while using the remaining user information for testing and validation. Figure 4 shows relevant performance metrics derived from these datasets. Overall, increasing the number of overlapping users during training improves effectiveness in cross-domain cold-start recommendations. The blue color represents the effect of 80% training users across all metrics. However, in larger datasets with a significant amount of proprietary data like Phones-Ele dataset, this relative advantage may not be very noticeable but still outperforms situations with a smaller proportion of overlapping users. This observation underscores the benefits of cross-domain recommendation as comprehensive data from diverse domains facilitates efficient feature extraction within a specific domain.

3.5 Analysis of the Impact of User Interaction Sparsity

We designed experimental scenarios (Table 4) to investigate cold-start performance with different interaction sparsity levels in different cross-domain datasets. Test users from the Y domain were grouped based on varying numbers of interactions from the X domain to observe cold-start impacts. Increasing interactions in the X domain lead to decreased sparsity but doesn't necessarily improve cold-start recommendation in the Y domain. For instance, in the ToysVideoG domain, performance improves with more interactions. However, this may not apply to every dataset or cross-

scenarios	value	w/o CDKF	w/o UPA	w/o C&U	D2C2R
ToysG	ndcg	4.03	3.90	2.98	4.68
	hr	7.27	6.56	5.85	8.43
VideoG	ndcg	6.42	6.41	6.06	6.93
	hr	12.49	12.49	11.60	13.62
Music	ndcg	4.69	4.90	4.66	5.33
	hr	8.64	9.29	8.74	10.13
Movie	ndcg	9.03	8.89	8.77	9.34
	hr	16.75	16.56	16.67	17.38
Phones	ndcg	7.80	8.30	7.97	8.54
	hr	14.96	15.00	14.73	15.77
Ele	ndcg	13.58	13.64	13.21	14.05
	hr	23.83	23.66	23.06	24.32

Table 3: Results (%) of ablation experiments.

scenarios	rating	mrr	ndcg@5	ndcg@10	hr@5	hr@10
ToysG ↓ VideoG	0-10	5.47	4.57	5.92	7.67	11.87
	10-20	5.61	4.85	7.19	8.82	15.97
	20-30	7.05	4.94	6.90	6.06	12.12
	30-	19.22	16.50	19.56	18.18	27.27
Muscic ↓ Movie	0-10	7.71	6.51	8.70	10.27	17.07
	10-20	9.05	7.84	9.95	12.11	18.81
	20-30	11.44	10.42	11.81	15.62	20.31
	30-	7.24	6.31	9.87	10.00	20.00
Phones ↓ Ele	0-10	12.27	11.64	13.97	16.80	24.02
	10-20	11.44	10.97	13.08	16.11	22.63
	20-30	6.83	6.12	7.45	8.97	13.14
	30-40	6.61	6.32	7.82	10.31	14.80
	40-	31.50	1.99	2.93	2.91	5.81

Table 4: Recommendation performance(%) in the Y domain corresponding to varying numbers of interactions in the X domain across three scenarios.

domain scenario due to varying domain correlations and user biases. In the Phones-Ele scenario, recommendation effectiveness declines after 20 to 40 interactions. To address this issue to some extent, our proposed User Preference Adaptation (UPA) module offers a solution.

4 Related Work

Classical models like EMCDR(Man et al. 2017) use matrix factorization for distinct domains and employ overlapping users for training source-target domain mappings. Similarly, DCDCSR(Zhu et al. 2020a) utilizes overlapping entities mapped onto a standardized domain for providing recommendations while MeKB-Rec(Su et al. 2023) uses personal knowledge graphs to enhance user information. Although CDLFM(Wang et al. 2020) improves representation methods using user neighbor information, it has limitations including low overall efficiency and neglecting inter-domain relationship mining with respect to domain alignment and variability. While the DisAlign model(Liu et al. 2021) integrates distributed alignment, it overlooks the integration of domain-specific shared information. The MAN(Lin et al. 2024) model uses hybrid attention to enable cold-start recommendations in cross-domain scenarios without overlapping users, though it requires stricter sequential temporal consistency. The variational encoder, proposed by Liang et al.(Liang et al. 2018), offers a microscopic disentanglement approach; however, its direct applicability to cross-domain scenarios remains limited(Wang et al. 2022).DCCF(Ren et al. 2023) proposes a novel approach to disentangle representation learning from recommendation. The CurCoDis(Wang et al. 2023) model analyzes user behaviors in social recommendation by disentangling consumption links and social environments. MacridVAE(Ma et al. 2019) captures user intentions and preferences with controllability. However, these approaches lack a comprehensive cross-domain observation. DisenCDR(Cao et al. 2022a) aims to enhance recommendation by learning domain-specific and shared information but focuses more on user representations.

Contrastive learning(Chen et al. 2020, 2019) is increasingly used in recommender systems. The AdaGCL framework(Jiang, Huang, and Huang 2023) introduces supplementary information to address the issue of data sparsity in the recommendation domain by enhancing the contrasting perspective of the graph model. SITN(Sun et al. 2023) uses a two-level approach to refine domain representations. CLHG(Li et al. 2021) introduces a global-level model to handle noisy knowledge in heterogeneous graphs. CCTL(Zhang et al. 2023) mitigates negative transfer effects through contrastive learning. CCDR(Xie et al. 2022) uses triple contrastive learning tasks. CCFCRec(Zhou, Zhang, and Yang 2023) leverages contrastive learning for cross-domain co-embedding. This paper effectively utilizes contrastive learning to eliminate user deviations across domains when extracting personalized features.

5 Conclusion

In this paper, we propose a domain-level generalized disentanglement framework for cross-domain cold-start recommendation, called D2C2R. The present study and proposes a domain-level perspective disentanglement analysis approach. Firstly, we design a FVGAE to extract shared information by leveraging the training process of two independent domains. Subsequently, we propose a heuristic decision-making method based on cross-domain fusion of user attention. Through the extraction and fusion process, relevant user information from different domains is captured at the domain level. The D2C2R model also incorporates a user personalized adaptor to refine the representations of individual users, emphasizing their personalized preferences and mitigating domain bias issues. Extensive experimental results demonstrate that our approach surpasses the current state-of-the-art methods in challenging cross-domain cold-start scenarios. In future work, we will delve into exploring multi-domain recommendation techniques and assess the applicability of this concept in enhancing recommendation performance for target domains.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 62472024, and Grant 72004009.

References

Cao, J.; Li, S.; Yu, B.; Guo, X.; Liu, T.; and Wang, B. 2023. Towards Universal Cross-Domain Recommendation. In *ACM International Conference on Web Search and Data Mining (WSDM)*.

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. IEEE.

Chen, Q.; Zhao, H.; Li, W.; Huang, P.; and Ou, W. 2019. Behavior sequence transformer for e-commerce recommendation in alibaba. In *Proceedings of the 1st international workshop on deep learning practice for high-dimensional sparse data*, 1–4.

Chen, T.; Kornblith, S.; Norouzi, M.; and Hinton, G. 2020. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, 1597–1607. PMLR.

Chen, X.; Cheng, Z.; Xiao, S.; Zeng, X.; and Huang, W. 2023. Cross-domain augmentation networks for click-through rate prediction. *arXiv preprint arXiv:2305.03953*.

Dhage, R.; Nehete, S.; Hon, S.; Patankar, T.; and Kale, L. 2023. Implementation of Recommendation System's Service Model Using Amazon E-commerce Dataset. In *International Conference on ICT for Sustainable Development*, 415–426. Springer.

Fouss, F.; and Saerens, M. 2008. Evaluating performance of recommender systems: An experimental comparison. In *2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, volume 1, 735–738. IEEE.

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.

Huan, Z.; Zhang, G.; Zhang, X.; Zhou, J.; Wu, Q.; Gu, L.; Gu, J.; He, Y.; Zhu, Y.; and Mo, L. 2022. An industrial framework for cold-start recommendation in zero-shot scenarios. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 3403–3407.

Jiang, M.; Cui, P.; Yuan, N. J.; Xie, X.; and Yang, S. 2016. Little is much: Bridging cross-platform behaviors through overlapped crowds. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30.

Jiang, Y.; Huang, C.; and Huang, L. 2023. Adaptive graph contrastive learning for recommendation. In *Proceedings of the 29th ACM SIGKDD conference on knowledge discovery and data mining*, 4252–4261.

Kamani, R.; Kumar, V.; and Kagita, V. R. 2023. Cross-domain Recommender Systems via Multimodal Domain Adaptation. *arXiv preprint arXiv:2306.13887*.

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.

Kipf, T. N.; and Welling, M. 2016. Variational graph auto-encoders. *arXiv preprint arXiv:1611.07308*.

Koren, Y.; Bell, R.; and Volinsky, C. 2009. Matrix factorization techniques for recommender systems. *Computer*, 42(8): 30–37.

Li, H.; Luo, X.; Yu, Q.; and Wang, H. 2021. Session-based recommendation via contrastive learning on heterogeneous graph. In *2021 IEEE International Conference on Big Data (Big Data)*, 1077–1082. IEEE.

Li, P.; and Tuzhilin, A. 2021. Dual metric learning for effective and efficient cross-domain recommendations. *IEEE Transactions on Knowledge and Data Engineering*, 35(1): 321–334.

Liang, D.; Krishnan, R. G.; Hoffman, M. D.; and Jebara, T. 2018. Variational autoencoders for collaborative filtering. In *Proceedings of the 2018 world wide web conference*, 689–698.

Lin, G.; Gao, C.; Zheng, Y.; Chang, J.; Niu, Y.; Song, Y.; Gai, K.; Li, Z.; Jin, D.; Li, Y.; et al. 2024. Mixed Attention Network for Cross-domain Sequential Recommendation. In *Proceedings of the 17th ACM International Conference on Web Search and Data Mining*, 405–413.

Liu, W.; Su, J.; Chen, C.; and Zheng, X. 2021. Leveraging distribution alignment via stein path for cross-domain cold-start recommendation. *Advances in Neural Information Processing Systems*, 34: 19223–19234.

Liu, W.; Zheng, X.; Su, J.; Hu, M.; Tan, Y.; and Chen, C. 2022. Exploiting variational domain-invariant user embedding for partially overlapped cross domain recommendation. In *Proceedings of the 45th International ACM SIGIR conference on research and development in information retrieval*, 312–321.

Ma, J.; Zhou, C.; Cui, P.; Yang, H.; and Zhu, W. 2019. Learning disentangled representations for recommendation. *Advances in neural information processing systems*, 32.

Man, T.; Shen, H.; Jin, X.; and Cheng, X. 2017. Cross-domain recommendation: An embedding and mapping approach. In *IJCAI*, volume 17, 2464–2470.

Mnih, A.; and Salakhutdinov, R. R. 2007. Probabilistic matrix factorization. *Advances in neural information processing systems*, 20.

Ren, X.; Xia, L.; Zhao, J.; Yin, D.; and Huang, C. 2023. Disentangled contrastive collaborative filtering. In *Proceedings of the 46th International ACM SIGIR Conference on*

Research and Development in Information Retrieval, 1137–1146.

Rendle, S.; Freudenthaler, C.; Gantner, Z.; and Schmidt-Thieme, L. 2012. BPR: Bayesian personalized ranking from implicit feedback. *arXiv preprint arXiv:1205.2618*.

Salah, A.; Tran, T. B.; and Lauw, H. 2021. Towards source-aligned variational models for cross-domain recommendation. In *Proceedings of the 15th ACM Conference on Recommender Systems*, 176–186.

Su, X.; Zhou, Y.; Shan, Z.; and Chen, Q. 2023. MeKB-Rec: Personal Knowledge Graph Learning for Cross-Domain Recommendation. *arXiv preprint arXiv:2310.11088*.

Sun, G.; Shen, Y.; Zhou, S.; Chen, X.; Liu, H.; Wu, C.; Lei, C.; Wei, X.; and Fang, F. 2023. Self-supervised interest transfer network via prototypical contrastive learning for recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, 4614–4622.

Wang, X.; Chen, H.; Zhou, Y.; Ma, J.; and Zhu, W. 2022. Disentangled representation learning for recommendation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(1): 408–424.

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.

Wang, X.; Pan, Z.; Zhou, Y.; Chen, H.; Ge, C.; and Zhu, W. 2023. Curriculum co-disentangled representation learning across multiple environments for social recommendation. In *International Conference on Machine Learning*, 36174–36192. PMLR.

Wang, X.; Peng, Z.; Wang, S.; Yu, P. S.; Fu, W.; Xu, X.; and Hong, X. 2020. CDLFM: cross-domain recommendation for cold-start users via latent feature mapping. *Knowledge and Information Systems*, 62: 1723–1750.

Wei, Y.; Wang, X.; Li, Q.; Nie, L.; Li, Y.; Li, X.; and Chua, T.-S. 2021. Contrastive learning for cold-start recommendation. In *Proceedings of the 29th ACM International Conference on Multimedia*, 5382–5390.

Xie, R.; Liu, Q.; Wang, L.; Liu, S.; Zhang, B.; and Lin, L. 2022. Contrastive cross-domain recommendation in matching. In *Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining*, 4226–4236.

Xu, T.; Chen, W.; Wang, P.; Wang, F.; Li, H.; and Jin, R. 2021. Cdtrans: Cross-domain transformer for unsupervised domain adaptation. *arXiv preprint arXiv:2109.06165*.

Zang, T.; Zhu, Y.; Liu, H.; Zhang, R.; and Yu, J. 2022. A survey on cross-domain recommendation: taxonomies, methods, and future directions. *ACM Transactions on Information Systems*, 41(2): 1–39.

Zhang, W.; Zhang, P.; Zhang, B.; Wang, X.; and Wang, D. 2023. A Collaborative Transfer Learning Framework for Cross-domain Recommendation. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 5576–5585.

Zhao, C.; Zhao, H.; He, M.; Zhang, J.; and Fan, J. 2023. Cross-domain recommendation via user interest alignment. In *Proceedings of the ACM Web Conference 2023*, 887–896.

Zhou, Z.; Zhang, L.; and Yang, N. 2023. Contrastive collaborative filtering for cold-start item recommendation. In *Proceedings of the ACM Web Conference 2023*, 928–937.

Zhu, F.; Chen, C.; Wang, Y.; Liu, G.; and Zheng, X. 2019. Dtcdr: 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.; Orgun, M.; and Wu, J. 2020a. A deep framework for cross-domain and cross-system recommendations. *arXiv preprint arXiv:2009.06215*.

Zhu, F.; Wang, Y.; Chen, C.; Liu, G.; and Zheng, X. 2020b. A graphical and attentional framework for dual-target cross-domain recommendation. In *IJCAI*, volume 21, 39.

Zhu, Y.; Ge, K.; Zhuang, F.; Xie, R.; Xi, D.; Zhang, X.; Lin, L.; and He, Q. 2021. Transfer-meta framework for cross-domain recommendation to cold-start users. In *Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval*, 1813–1817.