Dual-Perspective Disentanglement: Learning Symmetric Group-Aware Representations for Cross-Domain Recommendation

Borui Wu[12], **Yuanbo Xu**[12*]

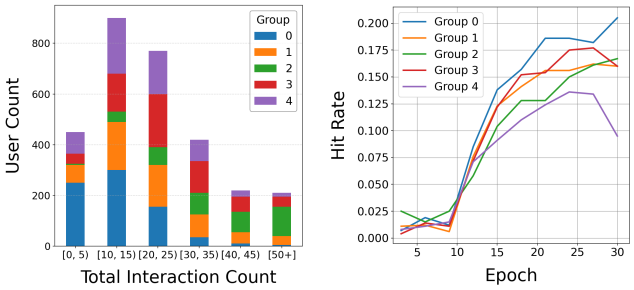
[1]College of Computer Science and Technology, Jilin University
[2]MIC Lab, College of Computer Science and Technology, Jilin University
wubr2122@mails.jlu.edu.cn, yuanbox@jlu.edu.cn

Abstract

Cross-Domain Recommendation (CDR) transfers user preferences from a source domain to alleviate data sparsity in a target domain. While disentangling representations into domain-specific and shared components is a common method, existing methods overlook user preference heterogeneity and item appeal heterogeneity. To this end, we propose **DPGCDR**, a **D**ual-**P**erspective **G**roup-aware **CDR** method that learns symmetric group-aware representations from both user and item. Conceptually, DPGCDR dynamically clusters users into groups and items into themes, then symmetrically disentangles user preferences into group-specific and cross-group shared components, and item attributes into theme-specific and cross-theme shared components. We propose a two-stage training scheme: 1) an initial warm-up stage learns preliminary representations to dynamically cluster users and items into group and theme structures which generalize cross-domain scenarios into multi-group disentanglement analogous to multi-domain settings; 2) a fusion-based aggregation stage integrates these group/theme-specific components into unified global representations. Additionally, an information-theoretic alignment regularizer further ensures consistency and discriminability between global shared and group/theme-specific representations, facilitating effective knowledge transfer by explicitly modeling and preserving the inherent multi-group structure within cross-domain interactions. Extensive experiments show DPGCDR achieves state-of-the-art performance, with significant gains of up to 25% in HR@10 over baselines on datasets with heterogeneous interaction structures. Further analyses confirm our dynamic clustering mechanism effectively adapts to underlying data patterns, enabling fine-grained cross-domain knowledge transfer.

Introduction

Recommender systems often suffer from data sparsity and cold-start issues, which degrade accuracy. Cross-domain recommendation (CDR) alleviates these problems by transferring knowledge across domains (e.g., movies to books) to enrich user preference models (Zang et al. 2022; Zhu et al. 2021a, 2022). Early CDR techniques relied on content-based aggregation using profiles or item attributes (Tang et al. 2012), then evolved to transfer-learning methods that share latent user/item factors (Man et al. 2017). Advances in deep learning introduced neural models like DCDCSR (Zhu et al. 2018) and unified frameworks such as Magnetic Metric Learning (Xu et al. 2021), both enhancing representation expressiveness. More recently, disentanglement approaches—DisenCDR (Cao et al. 2022) and DIDA-CDR (Zhu et al. 2023)—explicitly separate domain-specific, shared, and even domain-independent factors, reducing negative transfer and boosting performance on Amazon datasets.



(a) User group distribution. (b) HR@10 of each user group

Figure 1: The pilot experiments on user group structure and corresponding performance under DisenCDR (Cao et al. 2022) settings on Amazon Dataset.

Despite these gains, most CDR methods assume user homogeneity, ignoring critical user-group heterogeneity and causing uneven knowledge transfer (Song et al. 2024; Zhu et al. 2023). Moreover, they often employ asymmetric architectures that overlook the symmetric grouping of users (by preference) and items (by appeal). Addressing this gap requires a dual-perspective symmetric disentanglement over the symmetric user and item subgroup structures to achieve more robust and generalizable cross-domain transfer.

We conducted a pilot study on Amazon datasets using a standard CDR model, segmenting users via K-means clustering on learned embeddings. Visual analyses revealed uneven distributions along total interaction volume (Figure 1a), reflecting distinct user behavioral differences. Evaluating performance over epochs (Figure 1b) showed substantial variability among groups, highlighting that uniform disen-

*Corresponding Author.

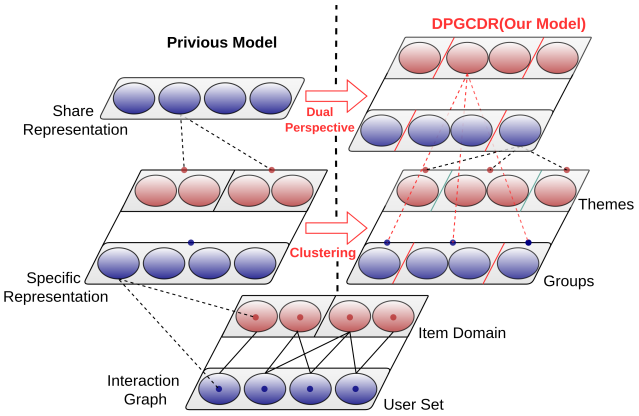


Figure 2: The idea illustration of DPGCDR architecture compared to previous models.

tangle model without explicit subgroup modeling tend to suffer from overfitting or negtive transfer. Thus, explicitly modeling user groups is crucial for fine-grained disentanglement for cross-domain transfer.

Motivated by these insights, we propose **DPGCDR** to learn **D**ual-**P**erspective **G**roup-Aware Representations for **CDR**. As depicted in Figure 2, comparing with previous models, DPGCDR introduces dynamic clustering to adaptively identify meaningful user groups and item themes directly from data and performs symmetric dual-perspective disentanglement from user and item: both sides modeling *specific* and *share* representations similarly. This facilitates precise bidirectional interactions between users and items, enhancing generalizable knowledge transfer. Our main contributions include:

- A two-stage framework that performs warm-up training then **dynamically clusters latent user groups and item themes beyond predefined domain boundaries**.
- **Dual-perspective disentanglement**, which symmetrically separates user and item representations into group/theme-specific and cross-group/theme-shared components, enabling fine-grained knowledge transfer.
- **Comprehensive validation of DPGCDR** on multiple real-world datasets, supported by visual analyses demonstrating the effectiveness of the learned structures.

Preliminaries

Problem Definition

Let $\mathcal{U} = \{u_1, \ldots, u_{N_u}\}$ and $\mathcal{I} = \{i_1, \ldots, i_{N_i}\}$ denote the sets of users and items respectively, where N_u and N_i are the number of users and items. We consider a cross-domain setting with two domains $\mathcal{D}^{(A)}$ and $\mathcal{D}^{(B)}$. Each observed interaction is a tuple $(u, i) \in \mathcal{U} \times \mathcal{I}$, indicating user u consumed item i. The binary implicit-feedback matrix $\mathbf{R} \in \{0,1\}^{N_u \times N_i}$ is defined by $R_{ui} = 1$ if (u, i) is observed in either domain and 0 otherwise.

To model structural relations, we build a bipartite interaction graph $G = (\mathcal{U} \cup \mathcal{I}, \mathcal{E})$ with adjacency matrix $\mathbf{A}$ and its transpose $\mathbf{A}^\top$ encodes the reverse edges.

Latent Representation Space

Our model defines four types of latent representations. First, the warm-up representation $(z_u^{\text{warm}}, z_i^{\text{warm}})$ serves as the basis for dynamically clustering users into groups and items into themes. Next, the group/theme-specific representation $(z_u^{\text{spec}}, z_i^{\text{spec}})$ captures fine-grained preferences or attributes unique to each user group or item theme. The group/theme-perspective representation $(\tilde{z}_u^{\text{persp}}, \tilde{z}_i^{\text{persp}})$ anchors the disentanglement process by providing Gaussian parameters that guide the separation of shared components. Finally, the group/theme-shared representation $(z_u^{\text{share}}, z_i^{\text{share}})$ encodes knowledge useful across multiple groups or themes, supporting effective transfer and generalization.

Methodology

Figure 3 gives an overview of the whole pipeline: a *warm-up* stage that discovers latent user groups and item themes, followed by a *dynamic* stage that performs symmetric group-aware disentanglement.

Initial Embedding & Graph Construction

Adjacency matrix. The interactions are encoded in the block matrix:

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{R} \\ \mathbf{R}^\top & \mathbf{0} \end{bmatrix}, \quad \mathbf{R}_{ui} = 1 \text{ iff } (u, i) \in \mathcal{E}. \tag{1}$$

Learnable initial embeddings. Every user u and item i is associated with a d-dimensional trainable vector:

$$\mathbf{e}_u^{(0)},\ \mathbf{e}_i^{(0)} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}),$$

where the variance σ^2 is fixed at initialisation. Optional side features (e.g., textual descriptions or images) can be concatenated. The tuple $\left(\mathbf{e}_u^{(0)}, \mathbf{e}_i^{(0)}, \mathbf{A}\right)$ forms the input to the VBGE.

Variational Bipartite Graph Encoder (VBGE)

We build on the Variational Autoencoder (VAE) framework (Kingma, Welling et al. 2013) and *2-hop* variational propagation scheme introduced in DisenCDR (Cao et al. 2022). Let $\mathbf{e}_u^{(l)} \in \mathbb{R}^d$ and $\mathbf{e}_i^{(l)} \in \mathbb{R}^d$ be the layer-l embeddings of user u and item i. We perform:

(1) Homogeneous-neighborhood aggregation. First, each node aggregates information from its *same-type* neighbors via a single GCN step:

$$\begin{aligned} \widehat{\mathbf{e}}_u &= \text{LeakyReLU}\big(\text{Norm}(\mathbf{A}^\top)\, \mathbf{U}^{(l)}\, \mathbf{W}_u^{(0)}\big), \\ \widehat{\mathbf{e}}_i &= \text{LeakyReLU}\big(\text{Norm}(\mathbf{A})\, \mathbf{I}^{(l)}\, \mathbf{W}_i^{(0)}\big), \end{aligned} \tag{2}$$

where $\text{Norm}(\cdot)$ denotes row-wise normalization, and $\mathbf{W}^{(0)}$ are trainable weight matrices. This step effectively propagates information along *2-hop* paths of the original bipartite graph. $\mathbf{U}^{(l)} \in \mathbb{R}^{N_u \times d}$ (resp. $\mathbf{I}^{(l)} \in \mathbb{R}^{N_i \times d}$) is the matrix whose rows are the d-dimensional embeddings of all users (resp. items) at layer l. By multiplying the normalized adjacency with $\mathbf{U}^{(l)}$ or $\mathbf{I}^{(l)}$, we update every *same-type* node in one batched operation.

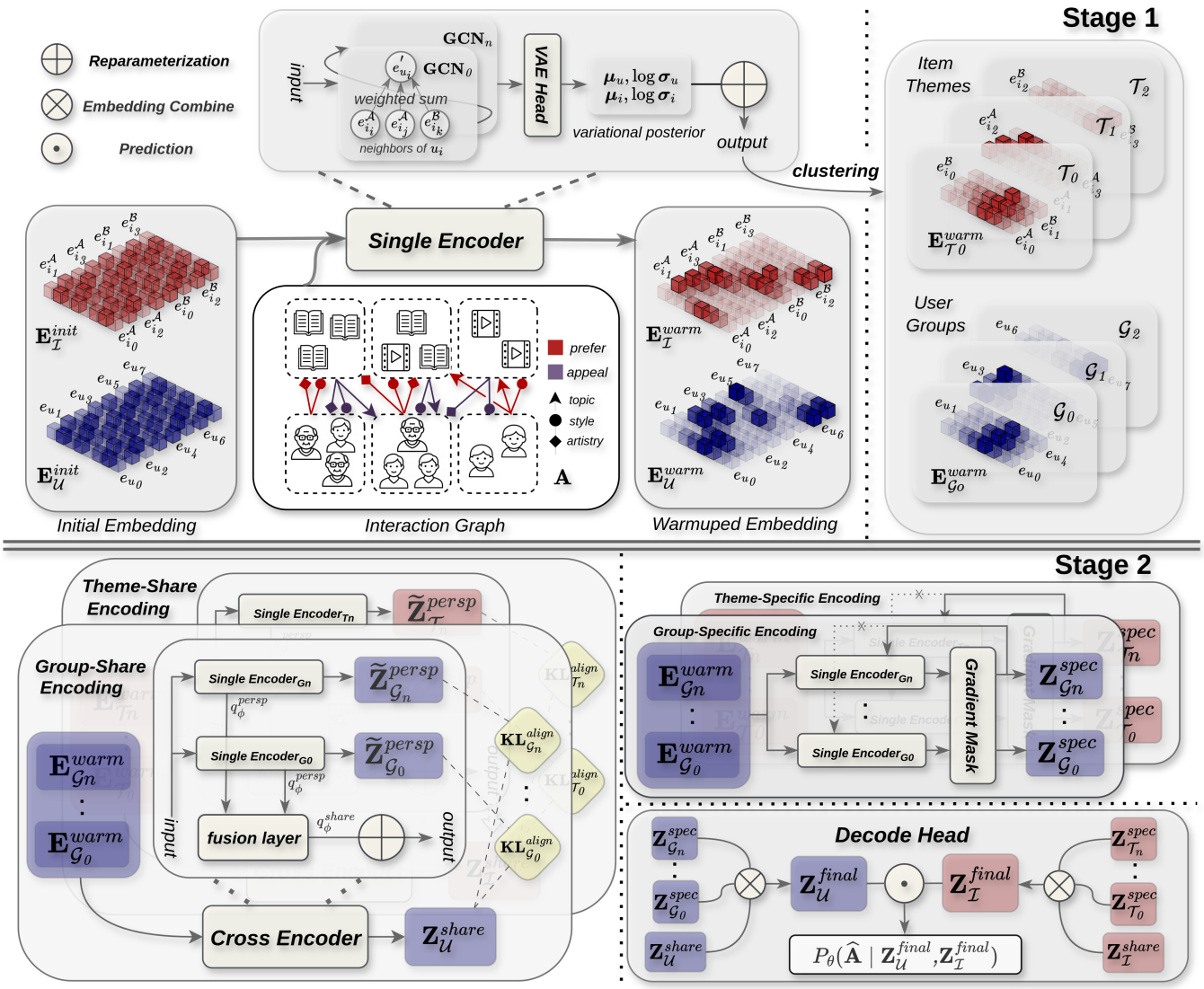


Figure 3: The framework of DPGCDR.

(2) Variational parameter head (VAE head). Next, we form the Gaussian posterior parameters by concatenating each node's own embedding and its intermediate representation:

$$\begin{aligned} \boldsymbol{\mu}_u &= \mathrm{LeakyReL}\big([\widehat{\mathbf{e}}_u \,\|\, \mathbf{e}_u^{(l)}]\,\mathbf{W}_{u,\mu}\big), \\ \log \boldsymbol{\sigma}_u &= \mathrm{Softplus}\big([\widehat{\mathbf{e}}_u \,\|\, \mathbf{e}_u^{(l)}]\,\mathbf{W}_{u,\sigma}\big), \end{aligned} \tag{3}$$

(and analogously for $\boldsymbol{\mu}_i$ and $\log \boldsymbol{\sigma}_i$) where $[\widehat{\mathbf{e}}_u \,\|\, \mathbf{e}_u^{(l)}] \in \mathbb{R}^{1\times 2d}$ denotes the concatenation, yielding a $2d$ vector. Multiplying by $\mathbf{W}_{u,\mu} \in \mathbb{R}^{2d\times d}$ (resp. $\mathbf{W}_{u,\sigma}$) and applying the nonlinearity produces the mean $\boldsymbol{\mu}_u \in \mathbb{R}^d$ and log-variance $\log \boldsymbol{\sigma}_u \in \mathbb{R}^d$ of the approximate posterior.

(3) Reparameterization and KL regularisation. We then sample the latent code via reparameterization trick (Kingma, Welling et al. 2013)

$$\mathbf{z}_u = \boldsymbol{\mu}_u + \exp\big(\tfrac{1}{2}\log \boldsymbol{\sigma}_u\big) \odot \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}), \tag{4}$$

(and analogously for $\mathbf{z}_i$). Each layer contributes the KL term $\mathrm{KL}\big(\mathcal{N}(\boldsymbol{\mu}, \mathrm{diag}\ \exp(\log \boldsymbol{\sigma})) \,\|\, \mathcal{N}(\mathbf{0}, \mathbf{I})\big)$ to the overall loss.

Two-Stage Learning Framework

Stage 1: Warm-up & Dynamic Clustering We first train a *single* VBGE on the entire interaction graph. Maximizing the standard Evidence Lower Bound (ELBO) (Liang et al. 2018) produces warm-up embeddings $\{\tilde{\mathbf{z}}_u, \tilde{\mathbf{z}}_i\}$.

The warm-up embeddings $\{\tilde{\mathbf{z}}_u\}$ and $\{\tilde{\mathbf{z}}_i\}$ are then clustered via K-Means. The number k is searched from $K_{\min}$ to $K_{\max}$ with the highest *silhouette score* (Rousseeuw 1987) above a given threshold. If no candidate k exceeds this threshold, it falls back to 1. The final group assignments $g(u) \in \{1, \ldots, G\}$ are then obtained. An identical procedure is performed to obtain $t(i) \in \{1, \ldots, T\}$.

Stage 2: Dual-Perspective Disentanglement Guided by the discovered clusters, we construct *four* parallel encoding branches—two for users, two for items— that each side simultaneously models *specific* and *share* factors.

Gradient Masking for Specific Paths. We allow VBGE to aggregate information from *all* neighbors, yet restrict back-propagation to update parameters and embeddings only for the nodes in group g or theme t. Concretely, we form the masked input:

$$\widehat{\mathbf{h}}_u = m_g \odot \mathbf{h}_u \; + \; (1 - m_g) \odot \mathbf{h}_u^{\text{detach}}, \tag{5}$$

(and analogously for $\widehat{\mathbf{h}}_i$) where $\mathbf{h}_u^{\text{detach}}$ is the same embedding but with its gradient detached and $m_g \in \{0,1\}^{|\mathcal{U}|}$ with $m_g = 1$ if u belongs to group g, and 0 otherwise.

Fusion Layer for Shared Paths. To build the *share* representation, we first run VBGE for each group g (or theme t) to obtain perspective parameters $(\mu^g,\ \log\sigma^g)$. We then fuse these perspective parameters into one share distribution via an attention-based weighted sum, which learns a scalar weight for each group/theme, applies softmax to obtain attention scores, and compute the weighted sum of means and log-variance across perspectives. Finally, we sample the shared embedding from the share distribution via the reparameterization trick, yielding $(\mathbf{z}_u^{\text{share}}, \mathbf{z}_i^{\text{share}})$.

Embedding Combine. To obtain the final embed- dings, we concatenate the *specific* and *share* embeddings with a learnable gate:

$$\mathbf{z}_u^{\text{final}} = [\mathbf{z}_u^{\text{share}} \parallel \mathbf{z}^{\text{spec},\, g(u)}], \tag{6}$$

(and analogously for $\mathbf{z}_i^{\text{final}}$) where the gate adapts the weight between transferable knowledge and subgroup nuances.

Prediction. The interaction score is computed as:

$$s_{ui} = \big(\mathbf{Z}_u^{\text{final}}\big)^\top \mathbf{Z}_i^{\text{final}}. \tag{7}$$

For implicit feedback, the probability of observing an interaction (u, i) in the bipartite graph is:

$$P_\theta\big(\widehat{\mathbf{A}} \;\big|\; \mathbf{Z}_u^{\text{final}}, \mathbf{Z}_i^{\text{final}}\big) = \text{sigmoid}\big(s_{ui}\big). \tag{8}$$

Optimization & Disentanglement Objective

To ensure that *shared* factors truly capture transferable knowledge while *specific* factors preserve group- or theme-level nuances, we optimize a composite loss

$$\mathcal{L} \;=\; \underbrace{\mathcal{L}_{\text{rec}}}_{\text{reconstruction}} \;+\lambda_{\text{KLD}}\; \underbrace{\mathcal{L}_{\text{KLD}}}_{\text{VAE regularizer}} \;+\beta_{\text{align}}\; \underbrace{\mathcal{L}_{\text{align}}}_{\text{disentanglement}}. \tag{9}$$

Reconstruction loss $\mathcal{L}_{\text{rec}}$. For implicit feedback, a bilinear decoder predicts $\hat{r}_{ui} = s_{ui}$ and we minimize the binary cross-entropy:

$$\mathcal{L}_{\text{rec}} = -\sum_{(u,i)\in\mathcal{E}^+\cup\mathcal{E}^-} \Big[R_{ui}\, \log\sigma(s_{ui}) + (1 - R_{ui})\, \log(1 - \sigma(s_{ui}))\Big]. \tag{10}$$

Variational regularizer $\mathcal{L}_{\text{KLD}}$. For each VBGE path we compute a regularizer penalty. Let $q(z) = \mathcal{N}(\boldsymbol{\mu}, \text{diag}(\boldsymbol{\sigma}^2))$ with elementwise log-variance $\boldsymbol{\ell} = \log\boldsymbol{\sigma}^2$. The KLD to the standard normal prior $p(z) = \mathcal{N}(\mathbf{0}, \mathbf{I})$ has the closed form:

$$\big(q\|p\) = \frac{1}{2}\sum_{j=1}^{d}\Big(\mu_j^2 + \exp(\ell_j) - \ell_j - 1\Big). \tag{11}$$

Alignment loss $\mathcal{L}_{\text{align}}$. The *share* distribution aim to *summarise* information contained in all group/theme *perspective* distributions, while remaining *distinct* from the corresponding specific distribution. For a user u, let $q_u^{\text{share}} = \mathcal{N}(\boldsymbol{\mu}_u^s, \text{diag}(\boldsymbol{\sigma}_u^s)^2)$ be the user-shared posterior, and $q_u^{\text{persp},g} = \mathcal{N}(\boldsymbol{\mu}_u^g, \text{diag}(\boldsymbol{\sigma}_u^g)^2)$ the g-th user-perspective posterior ($g = 1{:}G$). The KL divergence between two diagonal Gaussians admits

$$\big(q_u^{\text{share}} \,\|\, q_u^{\text{persp},g}\big) = \frac{1}{2}\sum_{j=1}^{d}\Bigg[\log\frac{(\sigma_{u,j}^g)^2}{(\sigma_{u,j}^s)^2} - 1 + \frac{(\sigma_{u,j}^s)^2}{(\sigma_{u,j}^g)^2} + \frac{(\mu_{u,j}^s - \mu_{u,j}^g)^2}{(\sigma_{u,j}^g)^2}\Bigg]. \tag{12}$$

Aggregating over all groups and users gives the *alignment–aggregation* term

$$\mathcal{L}_{\text{user}}^{\text{agg}} = \frac{1}{|\mathcal{U}_\mathcal{B}|}\sum_{u\in\mathcal{U}_\mathcal{B}}\frac{1}{G}\sum_{g=1}^{G}\big(q_u^{\text{share}} \,\|\, q_u^{\text{persp},g}\big), \tag{13}$$

where $\mathcal{U}_\mathcal{B}$ is the set of users in the batch. The item side is analogous.

During training, the three losses are computed for every mini-batch and back-propagated jointly. The detailed algorithm is presented in the appendix.

Experiments

In this section, we aim to answer the following questions:

- **RQ1**: How does DPGCDR perform compared to single-domain and cross-domain recommendation methods?
- **RQ2**: Can the dynamically identified user groups effectively improve cross-domain transfer performance, and how do different groups benefit from our method?
- **RQ3**: What is the contribution of each component of the proposed dual-perspective disentanglement?
- **RQ4**: How sensitive is DPGCDR to disentanglement hyperparameters.

Experimental Setup

Datasets To make a fair comparison, we evaluate our model on four real-world benchmark datasets from Amazon. We follow the preprocessing of BiTGCF (Liu et al. 2020a), combining the datasets into four CDR scenarios: **Elec&Phone**, **Elec&Cloth**, **Sport&Phone** and **Sport&Cloth**. Moreover, we follow the preprocessing of DisenCDR (Cao et al. 2022), removing the cold-start items from the test set.

Datasets	Metrics@10	single domain methods			cross domain methods			
		BPRMF	**LightGCN**	**DPGCDR***	**PPGN**	**BiTGCF**	**DisenCDR**	**DPGCDR**
Elec	HR	20.65	24.60	24.56	20.14	21.64	24.88	**29.21** (+17.4%)
	NDCG	11.66	14.32	14.16	11.37	12.23	14.60	**18.27** (+25.1%)
Cloth	HR	9.47	13.22	10.43	12.45	13.11	**15.14**	13.63 (–9.9%)
	NDCG	5.07	7.55	5.18	6.46	6.80	**8.44**	7.95 (–5.8%)
Elec	HR	15.71	21.60	20.62	17.14	19.04	23.20	**23.26** (+0.2%)
	NDCG	9.19	12.97	12.24	9.42	10.47	13.45	**13.59** (+1.0%)
Phone	HR	16.32	25.43	26.46	20.78	21.62	26.45	**27.78** (+5.0%)
	NDCG	8.53	13.88	15.11	12.28	12.79	14.72	**15.15** (+0.3%)
Sport	HR	10.43	14.84	16.32	13.64	14.83	16.77	**19.26** (+14.8%)
	NDCG	5.41	8.78	9.15	7.31	7.95	9.27	**10.86** (+17.1%)
Cloth	HR	11.53	14.03	11.91	14.24	14.68	**15.70**	13.65 (–13.0%)
	NDCG	6.25	8.38	6.38	7.69	7.93	**8.51**	7.30 (–14.2%)
Sport	HR	9.89	18.60	15.49	18.53	18.63	**20.15**	18.74 (–7.0%)
	NDCG	5.16	10.58	8.80	10.51	10.11	**11.23**	10.88 (–3.1%)
Phone	HR	13.60	22.88	23.73	20.04	21.10	22.03	**27.46** (+24.6%)
	NDCG	7.27	13.09	13.53	10.69	11.25	11.87	**15.97** (+34.5%)

Table 1: Overall comparison on four cross-domain tasks (HR@10 / NDCG@10).

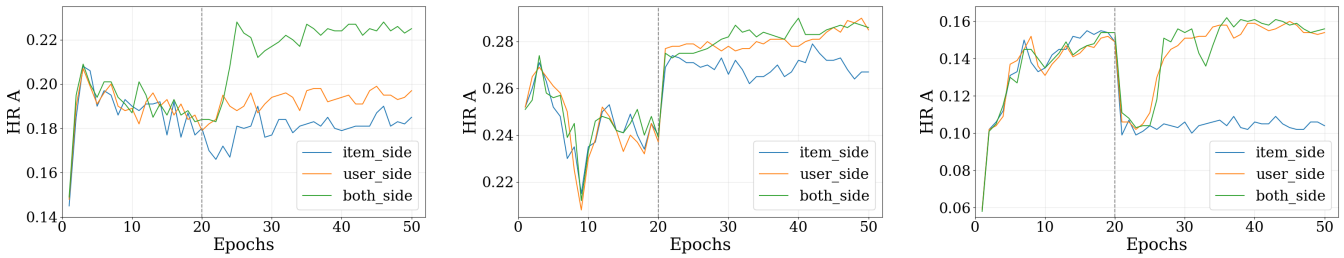
Baselines We compare our method against three *single-domain* baselines and *cross-domain* baselines. For single-domain recommendation, we consider **BPRMF** (Rendle et al. 2009), which employs Bayesian Personalized Ranking with matrix factorization, **LightGCN** (He et al. 2020), a simplified graph convolutional network for collaborative filtering, and **DPGCDR***, which denotes our model trained only in the warm-up stage. For cross-domain recommendation, we compare against **PPNG** (Zhao, Li, and Fu 2019), a model employs two distinct GCNs to learn user and item representations, while sharing an initialized user embedding layer, **BiTGCF** (Liu et al. 2020a), a bidirectional transfer GCN with feature-transfer layers, and **DisenCDR** (Cao et al. 2022), which applies disentangled VAE on bipartite graphs.

Evaluation Metrics We adopt the *leave-one-out* evaluation protocol following Zhu (Liu et al. 2020b): for each test user, one positive and 999 randomly sampled negatives are ranked among 1000 candidates. We report two widely used metrics: Hit Rate at 10 (HR@10), which measures the fraction of users whose ground-truth item appears in the *top-10* positions, and Normalized Discounted Cumulative Gain at 10 (NDCG@10), which evaluates the position-aware gain for the ground-truth item within the *top-10* ranked list. All reported results are averaged over three independent runs with different random seeds.

Implementation Details Our implementation is built upon the DisenCDR framework (Cao et al. 2022), with extensions for dynamic grouping and dual-perspective fusion. The model is trained on a single NVIDIA 3090 GPU. We set the embedding and GNN dimensions to $d = 64$, and apply two VBGE layers with a dropout rate of 0.3. Training is divided into two stages: a warm-up stage of 20 epochs with a learning rate of 1×10^{-3}, and a dynamic stage of 30 epochs with a learning rate of 5×10^{-4}. Optimization is performed using Adam with $\beta_1 = 0.9$, $\beta_2 = 0.999$, and weight decay of 5×10^{-4}. The batch size is set to 1024 user–item pairs.We use a silhouette threshold of $\tau = 0.5$ and set the the KLD weight $\lambda_{\text{KLD}} = 0.01$ and the alignment weight $\beta_{\text{align}} = 0.1$. The source code is presented in the appendix.

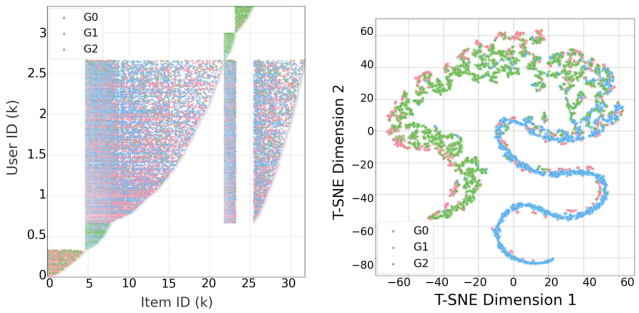
Performance Comparisons (RQ1)

Table 1 presents the overall results on four cross-domain recommendation scenarios, evaluated by HR@10 and NDCG@10. From the experimental results, we draw several key observations: (1) Compared to traditional single-domain approaches (e.g., BPRMF), the graph-based models (LightGCN and DPGCDR with our base VBGE encoder) achieve consistently better performance across all datasets. This demonstrates that effectively modeling higher-order collaborative signals significantly enhances recommendation quality. (2) Cross-domain methods consistently outperform their single-domain counterparts. Specifically, methods explicitly designed for cross-domain transfer, such as PPGN and BiTGCF, show clear advantages over graph-based single-domain methods. This highlights the importance of utilizing dedicated cross-domain transfer mechanisms instead of directly merging multiple domains into a single model. (3) Within the cross-domain methods, disentanglement-based models (DisenCDR and DPGCDR) consistently outperform other methods, including GNN-based transfer models (PPGN and BiTGCF). This suggests that explicitly modeling disentangled domain-specific and shared representations effectively reduces negative transfer, facilitating superior knowledge transfer across domains. (4) Compared to DisenCDR, DPGCDR out performances in most scenarios, particularly in datasets with clearer subgroup structures. However, performance gains become smaller or slightly negative in more homogeneous or sparse domains, indicating that the advantage of our model is most pronounced when distinct subgroup structures and rich cross-domain interaction patterns exist.



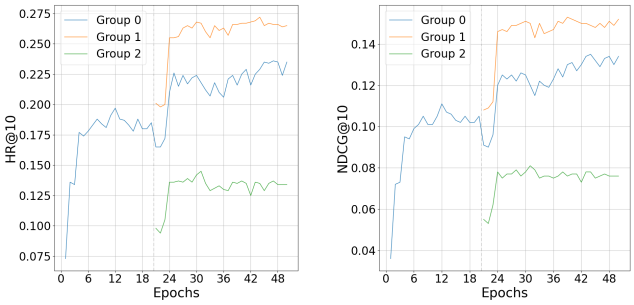
(a) HR@10 on domain A (Elec) (b) HR@10 on domain A (Elec) (c) HR@10 on domain A (Sport)

Figure 4: Ablation study of item-side, user-side, and both-sides disentanglement settings on scenario Elec&Phone (a), Elec&Cloth (b) and Sport&Phone (c). The dashed line indicates the division of stages.



(a) User group clustering. (b) User share embeddings.

Figure 5: Visualization on Elec&Phone scenario.



(a) HR@10 on domain A (b) NDCG@10 on domain A

Figure 6: Groups performance on domain Sport (A) &Phone (B) scenario. The dashed line indicates the division of stages.

Clustering & Transfer Effectiveness (RQ2)

Clustering quality. Figure 5a demonstrates that the three automatically identified groups (G0–G2) exhibit well-separated interaction patterns with limited overlap, indicating meaningful behavioural segmentation in the interaction matrix. Correspondingly, the t-SNE visualization shown in Figure 5b reveals three clearly separated cluster with minimal intermixing, indicating that our dynamic grouping captures distinct behavioural patterns. These two views demonstrate that the learned clusters reflect meaningful, behaviorally coherent user segments.

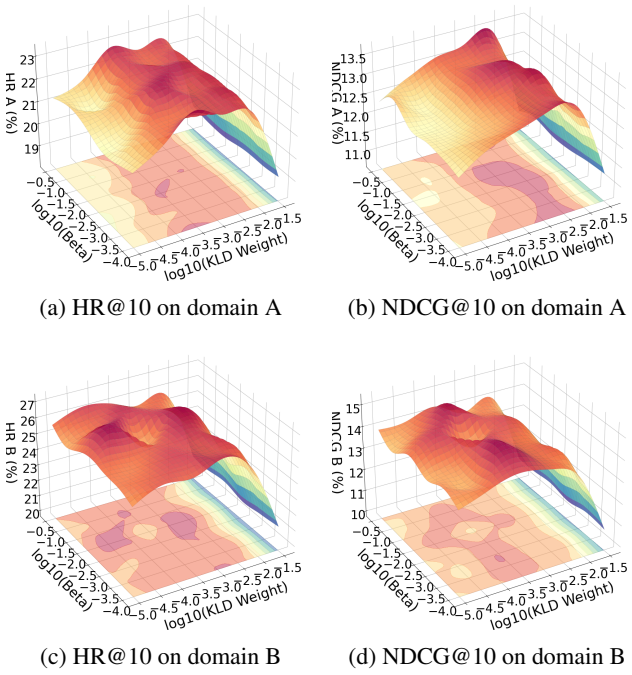
Per-group transfer effectiveness. Figure 6 shows HR@10 and NDCG@10 for each user group (all users start in a single group during the warm-up stage). Once the model transitions to dynamic training, every groups experiences an immediate improvement in recommendation accuracy, but the benefit is most pronounced for the group that was weakest initially. The cold-start group, which lagged behind during warm-up, sees its hit rate increase by roughly forty percent after clustering, demonstrating that group-specific parameters can significantly help users whose cross-domain affinities were previously under-modelled. The balanced group also enjoys a substantial gain (nearly 30%) as finer-grained alignment enhances its performance without compromising generality. Even the high-transfer group, already strong at the end of warm-up, achieves an additional lift in HR@105, indicating that disentangling item factors benefits all user segments. Overall, these results confirm that the automatically discovered clusters capture meaningful behavioral distinctions and that our group-aware transfer mechanism delivers sizable improvements across every groups, particularly rescuing those with the lowest initial accuracy.

Ablation Study (RQ3)

To examine the contributions of different disentanglement components, we evaluate three settings: (1) **Item-side**, which retains only the item branch; (2) **User-side**, which retains only the shared user branch; and (3) **Both-sides**, which activates both branches. Figure 4 plots HR on three scenarios:

Scenario I – Clear Dual-Perspective Advantage. In the *Phone→Elec* scenario (Figure 4a), where the two domains differ markedly in catalogue structure, (**Both-sides**) achieves a substantially higher hit rate compared to the user-only variant once dynamic grouping begins, underscoring that leveraging both user and item signals is crucial when cross-domain heterogeneity is pronounced.

Scenario II – User-Side Dominates but Dual View Still Helps. In *Cloth→Elec* scenario (Figure 4b), the much



(a) HR@10 on domain A (b) NDCG@10 on domain A

(c) HR@10 on domain B (d) NDCG@10 on domain B

Figure 7: Hyperparameter Sensitivity Analysis (KLD Weight & Beta) on Elec (A) &Phone (B) scenario.

larger target catalogue makes modeling user preferences essential. Although the user-only model (**User-side**) already outperforms the item-only baseline, the dual-perspective approach delivers a modest additional gain and converges more quickly, demonstrating that even a weaker item perspective can refine group boundaries and speed up training.

Scenario III – Similar Ceilings, Faster Convergence. In *Phone→Sport* Scenario (Figure 4c), all three variants reach comparable performance plateaus, but the dual-perspective model attains its peak several epochs earlier and maintains a slight lead. When interaction patterns are relatively homogeneous, the primary benefit of the dual view is improved training efficiency rather than a large boost in final accuracy.

Conclusion. Our dual-perspective disentanglement consistently beats its single-view counterparts while adapting to the structural characteristics of each dataset: the larger the divergence between user and item transfer patterns, the larger the gain; when divergence is low, dual views mainly speed up convergence without hurting final accuracy.

Hyperparameter Sensitivity (RQ4)

Figure 7 presents the sensitivity analysis regarding two key hyperparameters—KLD weight (λ) and alignment weight (β). Both metrics show consistent patterns across domains, revealing distinct performance regions. Specifically, excessively large values of λ lead to sharp performance drop due to the posterior collapse (Bowman et al. 2016). Meanwhile, optimal performance occurs within an median range of β, indicating the importance of balancing alignment strength. Overall, the performance plateaus are broad, suggesting robustness of the model's effectiveness within moderate ranges of λ and β.

Related Work

Cross-Domain Recommendation. CDR aims to mitigate data sparsity by transferring knowledge between domains (Zang et al. 2022). Early methods leveraged content and side information to identify domain commonalities, such as user profiles or item attributes (Hou et al. 2022; Li et al. 2023). Later, latent factor transfer approaches, like collective matrix factorization (Singh and Gordon 2008) and embedding mapping (e.g. EMCDR (Man et al. 2017)), enabled sharing of learned representations across domains either asymmetrically or symmetrically.

Recent methods increasingly use graph neural networks (GNNs) and meta-learning. GNN-based models, such as PPGN (Zhao, Li, and Fu 2019) and BiTGCF (Liu et al. 2020a), leverage heterogeneous graphs to propagate embeddings across domains. Meta-learning frameworks like TMCDR (Zhu et al. 2021b) pre-train domain-specific models and adapt to cold-start scenarios using meta-adaptation. Recognizing the challenge of negative transfer, models like TrineCDR (Song et al. 2024) explicitly filter irrelevant cross-domain signals. Recent studies (Jiang et al. 2023) emphasize the benefit of modeling such intervals, suggesting potential for context-aware enhancements in CDR.

Disentangled Representation Learning for Recommendation. Disentangled representation learning separates latent factors underlying user-item interactions to enhance generalization and interpretability (Ma et al. 2019; Wang et al. 2022). In the context of single-domain recommendation, methods such as DGCF (Wang et al. 2022) and DICE (Zheng et al. 2021) are designed to capture diverse user intents and isolate genuine preferences from external influences.

In cross-domain settings, recent disentangled methods aim to separate domain-specific from domain-shared factors. For example, DisenCDR (Cao et al. 2022) employs variational bipartite graph encoders with mutual information regularizers to encourage factor separation. Moreover, unified embedding approaches like MML (Xu et al. 2021) further emphasize preserving collaborative signals across domains through metric learning.

Conclusion and Future

In this paper, we emphasize that inherent user group structures analogously to item domain in CDR. Motivated by this dual symmetry, we propose **DPGCDR** to learn the Dual-Perspective Disentanglement on the dynamically discovered both user and item subgroup structures.

For future work, we hope to explore more robust and efficient clustering methods to better adapt to diverse dataset structures. Additionally, we hope to extend our framework to multi-domain scenarios, investigating the generalizability of our approach.

Acknowledgments

This work is supported by the Natural Science Foundation of China No. 62472196, Jilin Science and Technology Research Project 20230101067JC

References

Bowman, S.; Vilnis, L.; Vinyals, O.; Dai, A.; Jozefowicz, R.; and Bengio, S. 2016. Generating sentences from a continuous space. In *Proceedings of the 20th SIGNLL conference on computational natural language learning*, 10–21.

Cao, J.; Lin, X.; Cong, X.; Ya, J.; Liu, T.; and Wang, B. 2022. 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.

He, X.; Deng, K.; Wang, X.; Li, Y.; Zhang, Y.; and Wang, M. 2020. Lightgcn: Simplifying and powering graph convolution network for recommendation. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval*, 639–648.

Hou, Y.; Mu, S.; Zhao, W. X.; Li, Y.; Ding, B.; and Wen, J.-R. 2022. Towards universal sequence representation learning for recommender systems. In *Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining*, 585–593.

Jiang, Y.; Yang, Y.; Xu, Y.; and Wang, E. 2023. Spatial-temporal interval aware individual future trajectory prediction. *IEEE Transactions on Knowledge and Data Engineering*, 36(10): 5374–5387.

Kingma, D. P.; Welling, M.; et al. 2013. Auto-encoding variational bayes.

Li, J.; Wang, M.; Li, J.; Fu, J.; Shen, X.; Shang, J.; and McAuley, J. 2023. Text is all you need: Learning language representations for sequential recommendation. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 1258–1267.

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.

Liu, M.; Li, J.; Li, G.; and Pan, P. 2020a. Cross domain recommendation via bi-directional transfer graph collaborative filtering networks. In *Proceedings of the 29th ACM international conference on information & knowledge management*, 885–894.

Liu, Q.; Hu, Z.; Jiang, R.; and Zhou, M. 2020b. DeepCDR: a hybrid graph convolutional network for predicting cancer drug response. *Bioinformatics*, 36(Supplement_2): i911–i918.

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.

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*, UAI '09, 452–461. Arlington, Virginia, USA: AUAI Press. ISBN 9780974903958.

Rousseeuw, P. J. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20: 53–65.

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.

Song, Z.; Zhang, W.; Deng, L.; Zhang, J.; Wu, Z.; Bian, K.; and Cui, B. 2024. Mitigating negative transfer in cross-domain recommendation via knowledge transferability enhancement. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2745–2754.

Tang, J.; Wu, S.; Sun, J.; and Su, H. 2012. Cross-domain collaboration recommendation. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, 1285–1293.

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.

Xu, Y.; Wang, E.; Yang, Y.; and Chang, Y. 2021. A unified collaborative representation learning for neural-network based recommender systems. *IEEE Transactions on Knowledge and Data Engineering*, 34(11): 5126–5139.

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.

Zhao, C.; Li, C.; and Fu, C. 2019. Cross-domain recommendation via preference propagation graphnet. In *Proceedings of the 28th ACM international conference on information and knowledge management*, 2165–2168.

Zheng, Y.; Gao, C.; Li, X.; He, X.; Li, Y.; and Jin, D. 2021. Disentangling user interest and conformity for recommendation with causal embedding. In *Proceedings of the web conference 2021*, 2980–2991.

Zhu, F.; Wang, Y.; Chen, C.; Liu, G.; Orgun, M.; and Wu, J. 2018. A deep framework for cross-domain and cross-system recommendations. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, IJCAI'18, 3711–3717. AAAI Press. ISBN 9780999241127.

Zhu, F.; Wang, Y.; Chen, C.; Zhou, J.; Li, L.; and Liu, G. 2021a. Cross-Domain Recommendation: Challenges, Progress, and Prospects. In Zhou, Z.-H., ed., *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, 4721–4728. International Joint Conferences on Artificial Intelligence Organization. Survey Track.

Zhu, J.; Wang, Y.; Zhu, F.; and Sun, Z. 2023. Domain disentanglement with interpolative data augmentation for dual-

target cross-domain recommendation. In *Proceedings of the 17th ACM Conference on Recommender Systems*, 515–527.

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.