Dual Interests-Aligned Graph Auto-Encoders for Cross-domain Recommendation in WeChat

Jiawei Zheng, Hao Gu, Lingling Yi, Chonggang Song, Chuan Chen, Dandan Lin
WeChat Technical Architecture Department, Tencent Inc.
{joelzheng,nickgu,chrisyi,jerrycgsong,chuanchen,danieslin}@tencent.com

ABSTRACT

Recently, cross-domain recommendation (CDR) has been widely studied in both research and industry since it can alleviate a long-standing challenge of traditional recommendation methods, *i.e.*, data sparsity issue, by transferring the information from a relative richer domain (termed *source domain*) to a sparser domain (termed *target domain*). To our best knowledge, most of (if not all) existing CDR methods focus on transferring either *the similar content information* or *the user preferences embedding* from the source domain to the target domain. However, they fail to improve the recommendation performance in the real-world recommendation scenarios where the items in the source domain are totally different from those in the target domain in terms of attributes. To solve above issue, we analyzed the historical interactions of users from different domains in WeChat platform, and found that if two users have similar interests (interactions) in one domain, they are very likely to have similar interests in another domain even though the items of these two domains are totally different in terms of attributes. Based on this observation, in this paper, we propose a novel model named Dual Interests-Aligned Graph Auto-Encoders (*DIAGAE*) by utilizing the inter-domain interest alignment of users. Besides, our proposed model DIAGAE also leverages graph decoding objectives to align *intra-domain user interests*, which makes the representation of two users who have similar interests in a single domain closer. Comprehensive experimental results demonstrate that our model DIAGAE outperforms state-of-the-art methods on both publich benchmark datasets and online A/B tests in *WeChat live-stream* recommendation scenario. Our model DIAGAE now serves the major online traffic in WeChat live-streaming recommendation scenario.

1 INTRODUCTION

Due to the explosion of machine learning techniques, personalized recommendation system has become an indispensable part in wide areas of user-oriented web services, including video streaming [5, 31], live streaming [18], news delivery [10, 41], and etc. To predict users' preferences on unobserved items, *collaborative filtering* (CF) is a widely-used recommendation method by matching users with similar tastes or interests from their historical interactions [14]. However, the CF-based methods suffer from the data sparsity issue, especially for the new users or new items (*i.e.*, the cold-start problem), thus significantly reduces recommendation accuracy [40]. To alleviating the data sparsity issue, cross-domain recommendation (CDR) [24] has been proposed as a promising solution in both industry and academia since it can utilize the information from a richer domain (*source domain*) to a sparser domain (*target domain*), which helps learn better user preferences in the target domain. To our best of knowledge, existing CDR methods can be classified into two types: (1) the content-based CDR methods [7, 8, 25, 35], which transfer similar content information from the source domain to the target domain by leveraging the content-level relevance between two domains, and (2) the mapping-based methods [2, 15, 22, 43], which adapt the user preferences learnt in the source domain to the target domain by constructing the mapping functions (where the users are overlapped in both source and target domain).

However, all of existing CDR methods fail to improve the recommendation performance in the real-world recommendation scenarios where the items in the source domain are totally different from those in the target domain in terms of attributes. For example, in Douban APP (which is a popular Chinese exchange platform for reviews on movies, books, and music [36]), the item attributes of movie domain are totally different from those of music domain, and thus, the content-level transfer strategies have biases in learning a good relevance between two domains. Although the high-dimensional representations of items could be learnt via complicated neural networks, existing content-based CDR methods still cannot provide convincing explanations on the learnt representations of items, which leads to a poor learning on the relevance of items between two domains. Besides, the interest of a user in the movie domain cannot be easily mapped into the music domain via mapping functions. For example, it is challenging to recommend a list of songs to a user who likes watching the war movies since the categories of songs are different from those of movies. Similar issues can be found in WeChat APP which is the largest social media in China by involving the article reading platform, the micro-video sharing platform, and the live-streaming platform. Hence, it is challenging to apply existing CDR methods to the WeChat recommendation scenarios.

To tackle the above issues, we conduct an analysis on the historical interactions of users in different domains in WeChat APP. Specifically, we randomly sample some overlapped users who have interactions in both micro-video platform and the live-streaming platform in WeChat. Note that all the data are masked due to the privacy policies. In the micro-video domain, we calculate an *interest similarity value* for each pair of two users. This interest similarity value is defined as the cosine similarity of the categorical distributions of the of the items that they have interacted with, as plotted in the x-axis in Figure 1. The larger the value, the more similar these two users. Similar computations are conducted in the live-streaming platform, as plotted in the y-axis in Figure 1. From this figure, a phenomena of *inter-domain interest alignment* can be observed that the distribution of interest similarity values in the micro-video platform are consistent with that in the live-streaming platform. In particular, if two users have high interest similarity in the micro-video platform, then they also preserve the same high-level in terms of interest similarity in the live-streaming platform, and vice versa. It indicates that if two users have similar interests (interactions) in one domain, they are very likely to have similar interests in another domain even though the items of these two domains are totally different in terms of properties. However, to the best of our knowledge, this phenomenon of *inter-domain interest alignment* has not been studied in literature.

In this paper, we propose a novel model named Dual Interests-Aligned Graph Auto-Encoders (*DIAGAE*) by utilizing the inter-domain interest alignment of users. Instead of mapping the single user preferences from the source domain to the target domain, our proposed model aims to map the paired interests of two users from the source domain to the target domain, which eliminates the biases revealed by the different items in two domains. Furthermore, our proposed model DIAGAE also leverages graph decoding objectives to align intra-domain user interests, which makes the representations of two users who have similar interests in a single domain closer. By exploiting the intra-domain user interests, the learnt representation of a user in the target domain can be better, which, in turn, improves the performance of inter-domain interest alignments. The major contributions of this paper are summarized as follows:

- We are the *first* one who propose the phenomena of inter-domain interest alignment with a convincing analysis by using the real-data in WeChat.
- Inside our proposed model DIAGAE, we design not only an novel inter-domain aligned module, which forces the paired interests similarity in the target domain to align those in the source domain, but also an intra-domain aligned module by utilizing the graph decoding objectives, which makes the representations of users with similar interests in a single domain closer.
- We conduct extensive experiments on both public benchmark datasets and online A/B tests in recommendations of *WeChat live-streaming* to demonstrate that DIAGAE significantly outperforms the state-of-the-art CDR models.
- We provide a case study to illustrate the explainability of our model on WeChat platform. Besides, our model has been deployed on WeChat live-streaming recommendation scenario, serving billions of users.

2 PROPOSED METHOD

2.1 Preliminaries

Under the setting of cross-domain recommendation, we have a source domain $\mathcal{S}$ with richer information and a target domain $\mathcal{T}$. We assume $\mathcal{S}$ has $\left|U^{\mathcal{S}}\right|$ users and $\left|V^{\mathcal{S}}\right|$ items, and $\mathcal{T}$ has $\left|U^{\mathcal{T}}\right|$ users and $\left|V^{\mathcal{T}}\right|$ items. Two rating matrices $\mathbf{R}^{S} \in \mathbb{R}^{|U^{\mathcal{S}}|\times|V^{\mathcal{S}}|}$ and $\mathbf{R}^{T} \in \mathbb{R}^{|U^{\mathcal{T}}|\times|V^{\mathcal{T}}|}$ represent the interactions between users and items in each domain. In general, we model the user-item interaction as a bipartite graph $\mathcal{G} = (\mathcal{U}, \mathcal{I}, \mathcal{E})$, where $\mathcal{U}, \mathcal{I}, \mathcal{E}$ denote the set of users, items and edges, respectively. Under the setting of cross-domain recommendation, there are both source domain graph $\mathcal{G}^{S}$ and target domain graph $\mathcal{G}^{T}$.

2.2 Overview of DIAGAE

The basic idea of our proposed model DIAGAE is to leverage users' interest similarities, including inter-domain and intra-domain, to model user interest relationships and thus produce better user interest representations. The framework of DIAGAE is illustrated in Fig. 2. Firstly, we learn the user preferences in the source domain during the pre-training phase by using the traditional graph neural network (GNN) recommendation methods. The source graph encoder and embedding table of users and items are optimized with the help of user interaction in the source domain. Then, we design an inter-domain aligned module to force the paired interests similarity in the target domain to align thhose in the source domain. In this way, the similar paired interests in the source domain can transfer to the target domain and thus guide the target domain interests learning process. Both source and target preferences will be integrated together to form a more complementary user interest representation. Moreover, we leverage graph decoding objectives to align intra-domain users' interests, aiming to make the features of users with similar interests closer.

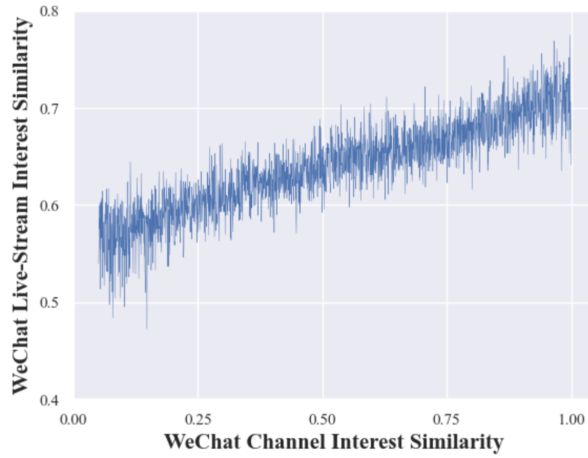


Figure 1: The distribution of user interest similarity values in the micro-video and live-streaming platform in WeChat.

2.3 Single-domain GNN Recommendation

In general, we model the user-item interaction as a bipartite graph $\mathcal{G} = (\mathcal{U}, \mathcal{I}, \mathcal{E})$, where $\mathcal{U}, \mathcal{I}, \mathcal{E}$ denote the set of users, items and edges, respectively. Considering that each neighbor node plays a different role and shows different importance for the center node, we adopt graph attention network (GAT) [27] to aggregate the neighbors' representation with different importance. Here we let the user node u as the center node, and $\mathcal{N}_u$ as its neighbors (e.g., interacted items) set to illustrate how the GAT work.

Due to the heterogeneity of nodes, nodes with different types have different feature spaces [29]. So we first project all the nodes into the same feature space and then calculate the attention scores between the center node and its neighbors.

$$\alpha_{ui} = \frac{\exp\left(\mathrm{LeakyReLU}\left(\boldsymbol{a}^{\mathrm{T}}\left[\boldsymbol{W}\boldsymbol{h}_u \| \boldsymbol{W}\boldsymbol{h}_i\right]\right)\right)}{\sum_{k\in\mathcal{N}_u}\exp\left(\mathrm{LeakyReLU}\left(\boldsymbol{a}^{\mathrm{T}}\left[\boldsymbol{W}\boldsymbol{h}_u \| \boldsymbol{W}\boldsymbol{h}_k\right]\right)\right)}, \forall i \in \mathcal{N}_u, \tag{1}$$

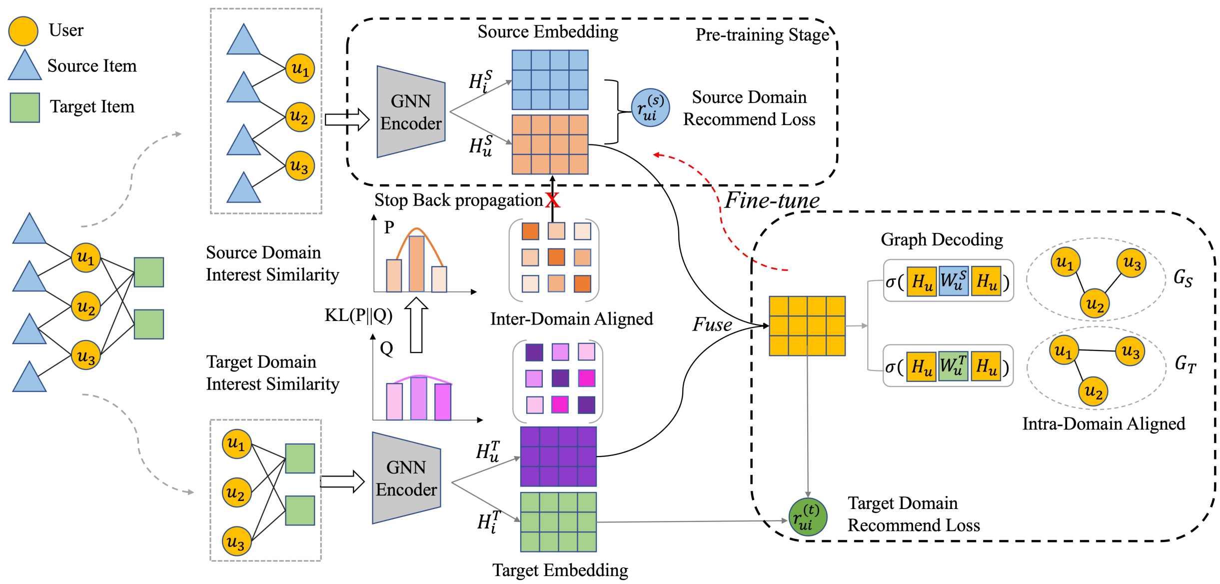


Figure 2: The framework of DIAGAE.

where $\boldsymbol{W} \in \mathbb{R}^{F' \times F}$ and $\boldsymbol{a} \in \mathbb{R}^{2F' \times 1}$ denote linear transformation. T represents transposition, and || is the concatenation operation. $\mathcal{N}_u$ denotes the neighbor set of node u. We apply LeakyReLU with negative input slope $\alpha = 0.2$ as activation function. Then we aggregate information from $\mathcal{N}_u$:

$$\boldsymbol{h}_u^{'} = \sigma\left(\sum_{i \in \mathcal{N}_u} \alpha_{ui}\boldsymbol{h}_i + \boldsymbol{h}_u\right), \tag{2}$$

where $\sigma(\cdot)$ denotes the activation function.

Through graph encoding, we obtain the representations of users and items $\boldsymbol{h}_u^{'}, \boldsymbol{h}_i^{'}$. Then we concatenate them and feed into an *MLP* to get the predicted score $\hat{r}_{ui}$. We have:

$$\hat{r}_{ui} = sigmoid(f(\boldsymbol{h}_u^{'} \oplus \boldsymbol{h}_i^{'})), \tag{3}$$

Where u and i denote the users and items, respectively. The $f(\cdot)$ is the MLP, $sigmoid(\cdot)$ denotes the sigmoid activation, and $\oplus$ is the concatenate operation.

Finally, we minimize the Mean Squared Error (MSE) loss to learn the user preferences:

$$\mathcal{L}_{rec} = \frac{1}{|\mathbf{R}|} \sum_{(u,i) \in \mathbf{R}} (r_{ui} - \hat{r}_{ui})^2 . \tag{4}$$

where $\mathbf{R}$ denotes the user-item interaction matrix, and r_{ui} is the actual rating of user u on item i.

2.4 Inter-domain Interest Aligning

In the pre-training stage, we leverage rich users' historical interaction with source items to learn users' interest representations and optimize the graph encoders in the source domain by minimizing the recommendation loss $\mathcal{L}_{rec}^S$, which is defined as Eq. 4

After obtaining the user interest representations in the source domain, we aim to transfer the users' interest relationships to the target domain. Note that the values of user's interest similarity may vary from different domains, which depend on the user's interest specificity. So directly aligning the values of user's interest similarity may suffer from the domain scale problem, which cannot reflect the true users' interest relationships. Motivated by t-SNE [26], we convert Euclidean distances between users' embeddings into probabilities that represent similarities to reduce the effect of the domain scale problem. Using the Student's t-distribution, users' interests similarities are defined as:

$$Sim_{ij} = \frac{\left(1 + \left\|h_{ui} - h_{uj}\right\|^2 / \alpha\right)^{-\frac{\alpha+1}{2}}}{\sum_{j'} \left(1 + \left\|h_{ui} - h_{uj'}\right\|^2 / \alpha\right)^{-\frac{\alpha+1}{2}}} \tag{5}$$

where h_{ui} denotes the representation of user i and α is the degree of freedom of the Student's t-distribution. Sim_{ij} measures the interest similarity of user i and j.

Here we define the users' interest similarities in source and target domain as $\mathcal{P}$ and $\mathcal{Q}$, respectively. We aim to minimize the mismatch between p_{ij} and q_{ij}. A natural measure of the faithfulness with which q_{ij} aligns p_{ij} is the Kullback-Leibler divergence. So we define our inter-domain aligned objective as KL divergence loss between the user's interest similarities in source and target domain:

$$\mathcal{L}_{inter} = \mathrm{KL}(\mathcal{P} \| \mathcal{Q}) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}} \tag{6}$$

where $\mathcal{P}$ and $\mathcal{Q}$ denote users' interest similarities in source and target domain, respectively.

Note that in our setting, users' historical interactions are much richer in the source domain, and we aim to improve the recommendation performance in the target domain. To this end, we regard users' interest similarities in the source domain as anchors to guide the learning process in the target domain and thus stop the back-propagation of $\frac{\partial \mathcal{L}_{inter}}{\partial \mathcal{P}}$.

2.5 Intra-domain Interest Aligning

In addition to act as anchors to guide the users' interest representations learning in the target domain, source user embedding are further integrated with target user embedding to form more complementary representations. Motivated by [38], we leverage graph decoding objectives to align intra-domain users' interests, aiming to make the features of users with similar interests closer.

Firstly, we transform the users' second-order relationships in the graph of both the source and target domain to first-order to describe the similarity between two users. It means that if two users interact with a lot of the same items, we consider they have similar interests and make their representations closer. As shown in the **Intra-domain Interest Align** part of Fig 2, each domain has its corresponding decoders $\left\{p\left(\hat{A} \mid H_U, W\right)\right\}$, predicting whether there is a link between two users, where $W \in \mathbb{R}^{D\times D}$ is decoder weights and D is the embedding dimension. Specifically, we reconstruct the user-user graph based on the fusion representation H_U:

$$p\left(\hat{A} \mid H_U, W\right) = \text{sigmoid}\left(H_U \cdot W \cdot H_U^{\mathrm{T}}\right). \tag{7}$$

Here we treat the graph reconstruction as the binary classification task, i.e., predict whether there is a link between two nodes. Formally, we use the binary cross entropy loss as the graph reconstruction loss. Considering we have both source and target domain graphs, the multi-domain graph decoding goals are as follows:

$$\begin{aligned} L_{intra} &= L_r^S + L_r^T \\ &= \text{loss}\left(A^S, \hat{A}^S\right) + \text{loss}\left(A^T, \hat{A}^T\right) \\ &= -\sum_{(i,j)\in A^S}\left(a_{ij}^S \log \hat{a}_{ij}^S + \left(1 - a_{ij}^S\right)\log\left(1 - \hat{a}_{ij}^S\right)\right) \\ &\quad - \sum_{(i,j)\in A^T}\left(a_{ij}^T \log \hat{a}_{ij}^T + \left(1 - a_{ij}^T\right)\log\left(1 - \hat{a}_{ij}^T\right)\right), \end{aligned} \tag{8}$$

where $a_{ij} = 1$ if there is a link between node i and j else $a_{ij} = 0$. Following [38], we calculate second-order proximity to construct user-user graphs in a mini-batch instead of in all the user nodes, which reduces the complexity from $O(N^2)$ to $O(\frac{N}{n}\times n^2) = O(N\times n)$, where $n \ll N$ denotes the batch size. We jointly optimize the overall model, and the total objective function is defined as:

$$\mathcal{L} = \mathcal{L}_{rec}^T + \lambda_{inter}\cdot\mathcal{L}_{inter} + \lambda_{intra}\cdot\mathcal{L}_{intra} \tag{9}$$

where $\mathcal{L}_{rec}^T$ is the recommendation prediction loss in the target domain, $\mathcal{L}_{inter}$ is the inter-domain aligned loss and $\mathcal{L}_{intra}$ is the intra-domain loss. λ_{inter}, λ_{intra} denote the weights of $\mathcal{L}_{inter}$ and $\mathcal{L}_{intra}$, respectively.

3 SYSTEM DEPLOYMENT

In this section, we introduce the deployment of DIAGAE in WeChat Live-Stream recommendation platform as shown in Fig. 3. In the representation generation stage, the system constructs the user-item interacted bipartite graph after extracting user's raw log data in both source and target domains. Then the system runs DIAGAE in a distributed manner to produce both user and item representations. In candidate generation stage, based on the learned representations, we generate *U2I (user-to-item) similarity maps* using inner product search. Given an item as a trigger, items with high similarity scores in the map are chosen as candidates. Note that the size of candidates is much smaller (usually thousands) compared with the size of the corpus (hundreds of millions). In ranking stage, rank models use online models to score candidates. Items with top scores are finally recommended and displayed to users. The training version of DIAGAE is online-updating.

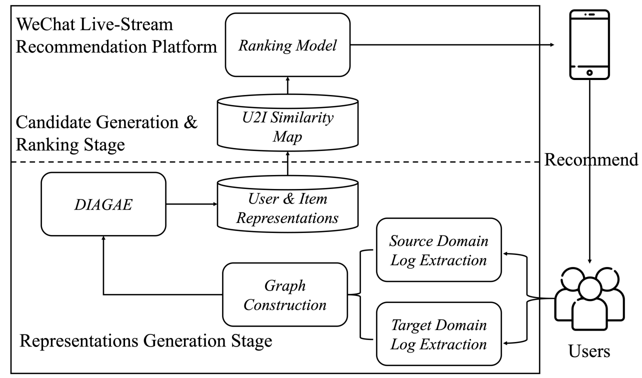


Figure 3: Deployment of DIAGAE in WeChat Live-Stream recommendation platform.

4 EXPERIMENTS

4.1 Experimental Setup

Datasets. To evaluate the benchmark performance of our proposed method DIAGAE for CDR tasks, we use one public dataset, *i.e.*, *Amazon* [43] which contains the rating scores that each user rates an item, and one production dataset (termed *WeChat*), which is extracted from our WeChat APP. Table 1 shows the statistics. Note that for all these tasks, the left (resp. right) hand side of the arrow is the source (resp. target) domain. Since the public dataset Amazon has the rating scores from 1 to 5, following previous work [43], we use two widely-used metrics to evaluate the performance of methods, *i.e.*, Mean Absolute Error (*MAE*) [12] and Root Mean Square Error (*RMSE*) [12]. For both metrics, the smaller the values, the better the performance. Besides, to evaluate the performance in terms of ranking in the Amazon dataset, we also adopt the metric *NDCG@k* with $k = 5$ by following previous work [20, 43]. Unlike the public dataset Amazon where the rating scores of each interaction are in the range of 1 to 5, the production dataset WeChat has no rating scores. Specifcially, in WeChat dataset, the interactions of users to items are binary, that is, 1 means that a user has interacted with an item (*e.g.*, watch or like), and 0 means that a user has no interaction with an item. So we adopt two standard evaluation metrics, *i.e.*, Accuracy (*ACC*) and *AUC* [21], following previous work [9, 22, 42]. For *NDCG@k*, *ACC* and *AUC*, the larger the values, the better the performance. To protect the user privacy, we preprocess all data via data masking.

Baselines. To show the superiority of our proposed method DIAGAE for CDR tasks, we compare DIAGAE with two kinds of existing methods: (1) the graph-based methods for single-domain recommendation tasks, namely, GC-MC [1], HERec [23], and NGCF [28]; and (2) the deep-learning-based methods for cross-domain recommendation tasks, namely, CoNet [13], DARec [34], BiTGCF [19], DDTCDR [17], ETL [4] and DisenCDR [2]. The details are as follows:

CDR Tasks	#Users	Source Domain			Target Domain		
		#Items	#Interactions	Density	#Items	#Interactions	Density
Book → Movie	16,738	58,987	304,540	0.031 %	191,942	638,404	0.020 %
Music → Book	80,763	35,896	963,373	0.033 %	93,799	1,323,101	0.017 %
Movie → Music	18,031	47,309	562,311	0.065 %	60,952	374,478	0.034 %
Micro-video → Live-streaming	268,502	432,749	40,220,999	0.035 %	239,482	9,779,197	0.015 %

Table 1: Statistics on Amazon and WeChat datasets for cross-domain recommendation tasks.

- GC-MC [1]: a graph-based auto-encoder framework for matrix completion. The training goal is to reconstruct the rating links in the interacted user-item bipartite graph.
- HERec [23]: firstly adopts metapath2vec [6] to generate node embeddings and then fuses multiple embedding to an extended matrix factorization model.
- NGCF [28]: a GCN-based model. It first captures the high-order connectivity information in the embedding function and then concatenates the obtained embeddings of each layer and uses the inner product to make the prediction.
- CoNet [13]: transfer dual knowledge across domains by introducing cross connections from one base network to another and vice versa.
- DARec [34]: consists of a rating pattern extractor, a domain classifier, and a predictor for rating estimation tasks, which automatically learns shared pattern representations and transfers them between two domains.
- BiTGCF [19]: a novel bi-directional transfer learning model for cross-domain recommendation. BiTGCF extends the flow of features from in-domain to inter-domain and considers the integration of uses' common features and domain-specific features.
- DDTCDR [17]: leverage the auto-encoder approach to extract the embedding and develop a novel latent orthogonal mapping to extract user preferences over multiple domains.
- ETL [4]: make an equivalent transformation assumption and model both the overlapped and domain-specific features in a joint distribution matching scheme.
- DisenCDR [2]: propose two mutual-information-based disentanglement regularizers to disentangle the domain-shared and domain-specific information.

Following previous work [19], we adapt the single-domain recommendation methods to the CDR tasks by firstly merging all interactions of both domains in datasets as a single domain and then applying the methods for recommendation in the target domain.

Implementation Details. For our proposed method DIAGAE, the embedding dimension is set to 64, the learning rate is set to 0.005, and the batch size is set to 1024 on all datasets. Besides, the hyper-parameters λ_{inter} and λ_{intra}, i.e., the weights of $\mathcal{L}_{inter}$ and $\mathcal{L}_{intra}$, are set to 0.8 and 0.6, respectively. In addition, we use the large-scale deep graph learning framework termed **PlatoDeep** [18] to handle the graph operations effectively. For fair comparison, we run the published codes of the competitors by five times and then report the average values.

4.2 Performance Comparison

In this section, we evaluate the effectiveness of our proposed method DIAGAE for solving the CDR tasks by comparing with existing state-of-the-arts. Table 2 shows the experimental results for four CDR tasks on both Amazon and WeChat datasets.

Fristly, we can observe from the results that our proposed method DIAGAE outperforms all existing graph-based single-domain recommendation methods in four CDR tasks. For example, on the WeChat dataset, DIAGAE achives the better performance than NGCF **by 10.5%** and **8.8%** in terms of ACC and AUC, respectively. It is because existing method NGCF treats all the information in both domains equally and cannot accurately capture the user interests in the target domain. Instead, our method DIAGAE can effectively transfer the knowledge from the micro-video domain (source) to the live-streaming domain (target) by distinguishing user interests in different domains. In addition, existing graph-based single-domain recommendation methods have worse performance than all the existing cross-domain recommendation methods in all CDR tasks, which demonstrates that designing concrete transferring strategies for CDR tasks is better than using one single network to model the mixed dataset.

Secondly, compared with the cross-domain recommendation methods, our method DIAGAE achieves the best performance in all CDR tasks for all metrics. For easy understanding, we define the improvement of DIAGAE in each CDR tasks as the relative improvement ratio of DIAGAE over the most competitive baseline (*i.e.*, the underlined result in each CDR tasks), termed as *%Improv.* in Table 2. For example, on the Movie → Music CDR task, our method has better performance than the state-of-the-art DisenCDR **by up to 5.45%** in terms of NDCG@5. It is because our method DIAGAE can implicitly model the relationships between users with common interests, including both inter-domain and intra-domain interest alignments, while the state-of-the-art DisenCDR explicitly learns the user domain-shared representations.

Ablation study. In order to investigate the effectiveness of each component in DIAGAE, we compare DIAGAE with two ablation models: 1) DIAGAE without inter-domain align loss, termed *w/o Inter-Domain*, and 2) DIAGAE without intra-domain loss (i.e., graph decoding loss), termed *w/o Intra-Domain* . Tabel 2 presents the results of this ablation study on four CDR tasks. From FigureTabel 2, we can observe that the full method DIAGAE achieves significant improvements for solving all four CDR tasks. For example, For Task1, the full method DIAGAE achieves the better performance than the *w/o Inter-Domain* (resp. *w/o Intra-Domain*) method**by up to 7.5%** (resp. **4.0%**) in terms of MAE. For Task4, the full method DIAGAE outperforms the *w/o Inter-Domain* (resp. *w/o Intra-Domain*)

CDR Tasks	Metrics	Single-domain Methods			Cross-domain Recommendation Methods						Ours			%Improv
		GC-MC	HERec	NGCF	CoNet	DARec	BiTGCF	DDTCDR	ETL	DisenCDR	DIAGAE	(w/o $\mathcal{L}_{inter}$)	(w/o $\mathcal{L}_{intra}$)	
Book → Movie	MAE	1.5722	0.9304	0.9855	0.7165	0.7042	0.6995	0.7068	0.7035	0.7012	**0.6684**	0.7157	0.6980	+4.45
	RMSE	2.4382	1.1509	1.1703	0.9578	0.9394	0.9102	0.9490	0.9351	0.9233	**0.8795**	0.9477	0.9113	+3.37
	NDCG@5	0.6982	0.7243	0.7192	0.7311	0.7462	0.7513	0.7316	0.7401	0.7455	**0.7847**	0.7483	0.7501	+4.45
Music → Book	MAE	1.5667	0.8046	0.7948	0.6347	0.6302	0.6257	0.6296	0.6117	0.6155	**0.5946**	0.6201	0.6134	+2.80
	RMSE	2.4449	0.9924	0.9801	0.8897	0.8714	0.8783	0.8746	0.8245	0.8163	**0.7840**	0.8194	0.8183	+3.96
	NDCG@5	0.6723	0.6899	0.6934	0.7013	0.7163	0.7188	0.7165	0.7213	0.7259	**0.7645**	0.7255	0.7248	+5.05
Movie → Music	MAE	1.0754	0.7892	0.7758	0.6568	0.6385	0.6299	0.6458	0.6294	0.6288	**0.6071**	0.6384	0.6297	+3.45
	RMSE	1.4502	1.1004	1.0052	0.9027	0.8534	0.8412	0.8677	0.8401	0.8396	**0.8257**	0.8522	0.8394	+1.66
	NDCG@5	0.6743	0.6811	0.6893	0.6901	0.7098	0.7121	0.7064	0.7183	0.7197	**0.7589**	0.7189	0.7163	+5.45
Micro-video → Live-streaming	ACC	0.6396	0.6477	0.6502	0.6599	0.6607	0.6785	0.6839	0.6831	0.6901	**0.7184**	0.6692	0.6757	+4.10
	AUC	0.6704	0.6895	0.6931	0.6983	0.7099	0.7173	0.7238	0.7206	0.7286	**0.7546**	0.6918	0.7024	+3.57

Table 2: Experimental results on Amazon and WeChat datasets. The best performance are highlighted in boldface while the second best are underlined. %Improv indicates the percentage of relative improvement of DIAGAE compared to the second best.

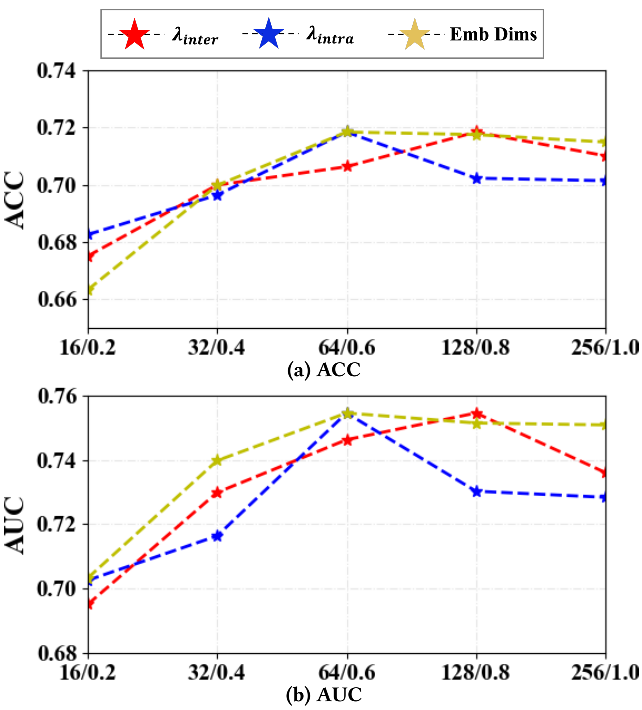


Figure 4: Analysis of parameters DIAGAE.

method**by up to 9.4%** (resp. **7.6%**) in terms of ACC. Similar results could be found in terms of RMSE (for Task1-3) and AUC (for Task4), as shown in Tabel 2. Thus, we empirically verify the effectiveness of both inter-domain and intra-domain interests alignments.

Parameter analysis.

We also study the effects of hyper-parameters on model performance, including the weight λ_{inter}, λ_{intra}, and embedding dimensions. Here we use the large-scale industry dataset *WeChat Live-Stream* as the case, shown as Fig 4.

We conduct experiments to study the effects of λ_{inter}, λ_{intra} by varying them in $\{0.2, 0.4, 0.6, 0.8, 1.0\}$. When $\lambda_{inter}, \lambda_{intra} \to 0$, the inter and intra alignment loss cannot produce positive effect. On the contrary, when λ_{inter} and λ_{intra} become too large, the interest alignment loss will suppress the recommendation loss and thus destroy the recommendation performance. Empirically, we choose $\lambda_{inter} = 0.8$ and $\lambda_{intra} = 0.6$. We also analyze the effect of the embedding dimensions by varying in $\{16, 32, 64, 128, 256\}$. DIAGAE achieves optimal performance when the dimension is set to 64. Meanwhile, around the optimal setting, the performance is generally stable, which indicates that our model is robust to the embedding dimensions.

4.3 Online A/B Test in WeChat

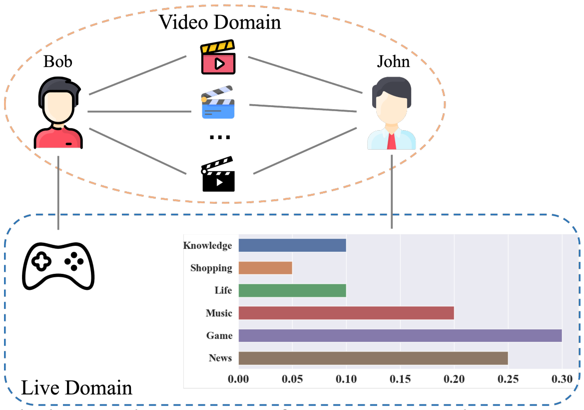
In this section, we evaluate the real online recommendation performance of our method DIAGAE by performing an online A/B test in the WeChat live-streaming recommendation service.

Settings. We choose as baseline a model NGCF [28] that has been already deployed in WeChat live-streaming recommendation service. To handle the graph-based operation for the large-scale heterogeneous graph in WeChat live-streaming recommendation system [18], the model NGCF is adaptively deployed by changing the spectral graph convolutional operators to the spatial graph convolutional operators, *i.e.*, the neighbour sampling operations. It is because the heterogenous graph used in WeChat service involves up to ten billions of edges [18], which is too large to handle the simple spectral graph convolutional operators. Due to resource limitations in company, we do not deploy other state-of-the-art CDR methods in Table 2. For both NGCF and our method DIAGAE, we utilize the information in the WeChat micro-video platform (*i.e.*, the source domain) for transferring and then improving the recommendation performance in the WeChat live-streaming scenario (*i.e.*, the target domain). Besides, following previous work [3, 18, 33, 37], three important online metrics are used to evaluate the recommendation performance in the live-streaming recommendation scenario: (1) User Click-through Rate (UCTR) , *i.e.*, $\frac{\#users-who-clicked}{\#user-who-viewed}$, which is the ratio of users who clicked at least one recommended item over the total number of users that have been recommended a list of items; (2) Page Click-through Rate (*PCTR*), *i.e.*, $\frac{\#clicks-on-items}{\#views-of-items}$, which is the ratio of the times that users clicked on an item over the total number of times that the recommended items have been viewed to the users; (3) Daily Active User (*DAU*). , *i.e.*, which is the number of users who are active on each day.

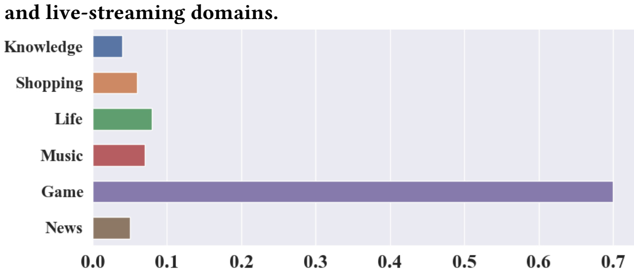
Results. We conduct the A/B test over 12 days,where nearly 50 million users are involved. Due to commercial confidentiality and secrecy agreement, we use the results of the adapted NGCF model as the base and only show the relative improvement rate of our proposed method DIAGAGE, as shown in Table 3. The results demonstrates that DIAGAE significantly outperforms the baseline in terms

Method	UCTR	PCTR	DAU
DIAGAE	+1.5732%	+2.9721%	+1.5686%

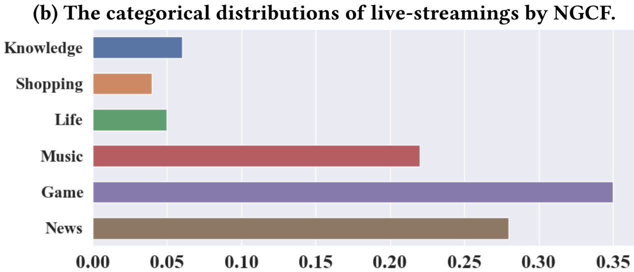
Table 3: Online A/B tests results of our method DIAGAE in WeChat live-streaming recommendation scenario.



(a) The historical interactions of two users in WeChat micro-video and live-streaming domains.



(b) The categorical distributions of live-streamings by NGCF.



(c) The categorical distributions of live-streamings by DIAGAE.

Figure 5: Case study for explanations of user interests alignments in our proposed method DIAGAE.

of all three metrics. For example, DIAGAE achieves better performance than NGCF **by up to 2.9721%** in terms of PCTR. Notice that in industrial recommendation scenarios, gain of 2% is a substantial improvement [3, 18, 33, 37]. It indicates the effectiveness of the user interest alignments in our proposed method.

Case Study for Explanations. To show that our method can provide the explanations of user interests alignments in real-world recommendation scenario, we conduct a qualitative case study by using the recommendation services in WeChat.

Specifically, we randomly select two users from WeChat who satisfy the following constraints: (1) these two users have similar historical interactions in the WeChat micro-video domain; (2) one user is an active user with rich interactions with items in live-streaming scenario; and (3) the other user is a cold-start user with a few interactions with items in live-streaming scenario. Subject to the above three constraints, two users *John* and *Bob* are selected, as shown in Figure 5a. We can see that (1) John and Bob have same historical interactions (*i.e.*, interests) in micro-video domain; (2) John has rich interactions with items from different categories in live-streaming domain, especially the items in *Music* category, *Game* category and *News* category; and (3) Bob has a few interactions with items in *Game* category in live-streaming domain. In this case study, we aim to recommend items in live-streaming domain to Bob. To show the effectiveness of recommendation performance of our proposed method, we compare with the adapted NGCF method in this case study. Figure 5b plots the categorical distributions of recommendation results returned by the adapted NGCF method. We can observe that the most of recommended items are in *Game* category, indicating that the adapted NGCF method computes a higher weights to the items that have the same category as Bob's historical interactions in live-streaming domain, which, however, leads to the information cocoon issue in recommendation system [30]. It is because the adapted NGCF method cannot build an interest connection between micro-video and live-streaming domains, which cannot effectively transfer the interest of Bob from micro-video domain to live-streaming domain. However, our proposed method DIAGAE can avoid the above information cocoon issue, as presented in Figure 5c. The results show that our method recommends to Bob the items in *Music*, *Game*, and *News* categories, which are consistent with the categorical distributions of historical interactions of John in live-streaming domain. It is because our proposed method DIAGAE can effectively transfer the similar user interests from micro-video domain to live-streaming domain by utilizing both the inter-domain and intra-domain interests alignments, which makes the representations of Bob and John in live-streaming domain closer.

5 RELATED WORK

In this section, we review the literature about the graph-based single-domain methods and the cross-domain methods.

5.1 Graph-based Recommendation Methods

Graph Neural Network (GNN) techniques are powerful in capturing connections among nodes and representation learning for graph data. In Recommender System, most data have essentially a graph structure, e.g., user-item bipartite graph or heterogeneous graph with rich side information, so GNN techniques are naturally introduced to model it. The GNN approches are various, some of which are designed for the homogeneous graphs while others for the hetero- geneous ones. In terms of the homogeneous graphs, GCN [16] and GraphSage [11]can be considered as the basic GNN algorithms by using the convolutional operations. To be specific, GCN in- corporates neighbors' feature representations using convolutional operations while GraphSage provides an inductive approach to combine structural information with node features. For the hetero- geneous graph with multiple types of vertices and/or edges, Meta- path2Vec [6] and HERec [23] formalize meta-path based random- walks to construct the heterogeneous neighborhood of a node and then leverage skip-gram models to perform node embeddings. Note that in this paper, we do not focus on in the perspective of algorithm design, e.g., the GNN approaches for the dynamic graphs, but focus on the general deep graph learning system.

5.2 Cross-domain Recommendation Methods

Most of existing CDR methods can be generally classified into two groups: (1) content-based and (2) embedding-based. The content-based approaches leverage content-level relevance and link different domains by similar contents, such as social tags [7], user-generated reviews [8, 25], item documents [35] and so on. The embedding-based methods firstly train the CF-based models (such as deep matrix factorisation [32], graphic models [39], etc.) to obtain user/item embeddings, and then transfer these embeddings through common users/items across domains by a well-designed mapping function [15, 22, 43]. Despite progress, these methods ignore the phenamenon of inter-domain interest alignment between two users. Instead, our proposed model DIAGAE can transfer the paired interests from the source domain to the target domain.

6 CONCLUSION

In this work, we propose a novel model *DIAGAE* for cross-domain recommendation in WeChat platform to better model the relationship between users with common interests, including both inter-domain and intra-domain interest aligns. Comprehensive experiments demonstrate that DIAGAE achieves significant improvements for both offline and online recommendation scenarios and is deployed on WeChat live-streaming recommendation system.

REFERENCES

[1] Rianne van den Berg, Thomas N Kipf, and Max Welling. 2017. Graph convolutional matrix completion. *arXiv preprint arXiv:1706.02263* (2017).

[2] Jiangxia Cao, Xixun Lin, Xin Cong, Jing Ya, Tingwen Liu, and Bin Wang. 2022. Disencdr: Learning disentangled representations for cross-domain recommendation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 267–277.

[3] M. Chen, A. Beutel, P. Covington, S. Jain, F. Belletti, and E. H Chi. 2019. Top-k off-policy correction for a REINFORCE recommender system. In *WSDM*. 456–464.

[4] Xu Chen, Ya Zhang, Ivor W Tsang, Yuangang Pan, and Jingchao Su. 2020. Towards equivalent transformation of user preferences in cross domain recommendation. *ACM Transactions on Information Systems (TOIS)* (2020).

[5] Zhihua Cui, Xianghua Xu, XUE Fei, Xingjuan Cai, Yang Cao, Wensheng Zhang, and Jinjun Chen. 2020. Personalized recommendation system based on collaborative filtering for IoT scenarios. *IEEE Transactions on Services Computing* 13, 4 (2020), 685–695.

[6] Yuxiao Dong, Nitesh V Chawla, and Ananthram Swami. 2017. metapath2vec: Scalable representation learning for heterogeneous networks. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*. 135–144.

[7] Ignacio Fernández-Tobías and Iván Cantador. 2014. Exploiting Social Tags in Matrix Factorization Models for Cross-domain Collaborative Filtering.. In *CBRecSys@ RecSys*. Citeseer, 34–41.

[8] Wenjing Fu, Zhaohui Peng, Senzhang Wang, Yang Xu, and Jin Li. 2019. Deeply fusing reviews and contents for cold start users in cross-domain recommendation systems. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 94–101.

[9] Chen Gao, Xiangning Chen, Fuli Feng, Kai Zhao, Xiangnan He, Yong Li, and Depeng Jin. 2019. Cross-domain recommendation without sharing user-relevant data. In *The world wide web conference*. 491–502.

[10] Yong Gao, Huifeng Guo, Dandan Lin, Yingxue Zhang, Ruiming Tang, and Xiuqiang He. 2021. Content Filtering Enriched GNN Framework for News Recommendation. *arXiv preprint arXiv:2110.12681* (2021).

[11] Will Hamilton, Zhitao Ying, and Jure Leskovec. 2017. Inductive representation learning on large graphs. *Advances in neural information processing systems* 30 (2017).

[12] Jonathan L Herlocker, Joseph A Konstan, Loren G Terveen, and John T Riedl. 2004. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)* 22, 1 (2004), 5–53.

[13] Guangneng Hu, Yu Zhang, and Qiang Yang. 2018. Conet: Collaborative cross networks for cross-domain recommendation. In *Proceedings of the 27th ACM international conference on information and knowledge management*. 667–676.

[14] Rong Hu and Pearl Pu. 2011. Enhancing collaborative filtering systems with personality information. In *Proceedings of the fifth ACM conference on Recommender systems*. 197–204.

[15] SeongKu Kang, Junyoung Hwang, Dongha Lee, and Hwanjo Yu. 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.

[16] Thomas N Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907* (2016).

[17] Pan Li and Alexander Tuzhilin. 2020. Ddtcdr: Deep dual transfer cross domain recommendation. In *Proceedings of the 13th International Conference on Web Search and Data Mining*. 331–339.

[18] Dandan Lin, Shijie Sun, Jingtao Ding, Xuehan Ke, Hao Gu, Xing Huang, Chonggang Song, Xuri Zhang, Lingling Yi, Jie Wen, et al. 2022. PlatoGL: Effective and Scalable Deep Graph Learning System for Graph-enhanced Real-Time Recommendation. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*. 3302–3311.

[19] Meng Liu, Jianjun Li, Guohui Li, and Peng Pan. 2020. 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.

[20] Weiming Liu, Xiaolin Zheng, Jiajie Su, Mengling Hu, Yanchao Tan, and Chaochao Chen. 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.

[21] Jorge M Lobo, Alberto Jiménez-Valverde, and Raimundo Real. 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global ecology and Biogeography* 17, 2 (2008), 145–151.

[22] Tong Man, Huawei Shen, Xiaolong Jin, and Xueqi Cheng. 2017. Cross-domain recommendation: An embedding and mapping approach.. In *IJCAI*, Vol. 17. 2464–2470.

[23] Chuan Shi, Binbin Hu, Wayne Xin Zhao, and S Yu Philip. 2018. Heterogeneous information network embedding for recommendation. *IEEE Transactions on Knowledge and Data Engineering* 31, 2 (2018), 357–370.

[24] Ajit P Singh and Geoffrey J Gordon. 2008. Relational learning via collective matrix factorization. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. 650–658.

[25] Shulong Tan, Jiajun Bu, Xuzhen Qin, Chun Chen, and Deng Cai. 2014. Cross domain recommendation based on multi-type media fusion. *Neurocomputing* 127 (2014), 124–134.

[26] Laurens Van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *Journal of machine learning research* 9, 11 (2008).

[27] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2017. Graph attention networks. *arXiv preprint arXiv:1710.10903* (2017).

[28] Xiang Wang, Xiangnan He, Meng Wang, Fuli Feng, and Tat-Seng Chua. 2019. Neural graph collaborative filtering. In *Proceedings of the 42nd international ACM SIGIR conference on Research and development in Information Retrieval*. 165–174.

[29] Xiao Wang, Houye Ji, Chuan Shi, Bai Wang, Yanfang Ye, Peng Cui, and Philip S Yu. 2019. Heterogeneous graph attention network. In *The World Wide Web Conference*. 2022–2032.

[30] Yifan Wang, Weizhi Ma, Min Zhang, Yiqun Liu, and Shaoping Ma. 2023. A survey on the fairness of recommender systems. *ACM Transactions on Information Systems* 41, 3 (2023), 1–43.

[31] Yinwei Wei, Xiang Wang, Liqiang Nie, Xiangnan He, Richang Hong, and Tat-Seng Chua. 2019. MMGCN: Multi-modal graph convolution network for personalized recommendation of micro-video. In *Proceedings of the 27th ACM international conference on multimedia*. 1437–1445.

[32] Hong-Jian Xue, Xinyu Dai, Jianbing Zhang, Shujian Huang, and Jiajun Chen. 2017. Deep matrix factorization models for recommender systems.. In *IJCAI*, Vol. 17. Melbourne, Australia, 3203–3209.

[33] J. Yang, X. Yi, D. Cheng, L. Hong, Y. Li, S. Wang, T. Xu, and E. H Chi. 2020. Mixed negative sampling for learning two-tower neural networks in recommendations. In *Companion Proceedings of the Web Conference 2020*. 441–447.

[34] Feng Yuan, Lina Yao, and Boualem Benatallah. 2019. DARec: Deep Domain Adaptation for Cross-Domain Recommendation via Transferring Rating Patterns. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence* (Macao, China) *(IJCAI'19)*. AAAI Press, 4227–4233.

[35] Cheng Zhao, Chenliang Li, Rong Xiao, Hongbo Deng, and Aixin Sun. 2020. CATN: Cross-domain recommendation for cold-start users via aspect transfer network. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 229–238.

[36] Junzhou Zhao, John CS Lui, Don Towsley, Xiaohong Guan, and Yadong Zhou. 2011. Empirical analysis of the evolution of follower network: A case study

on douban. In *2011 IEEE conference on computer communications workshops (INFOCOM WKSHPS)*. IEEE, 924–929.

[37] Z. Zhao, L. Hong, L. Wei, J. Chen, A. Nath, S. Andrews, A. Kumthekar, M. Sathiamoorthy, Xinyang Yi, and E. H Chi. 2019. Recommending what video to watch next: a multitask ranking system. In *Recsys*. 43–51.

[38] Jiawei Zheng, Qianli Ma, Hao Gu, and Zhenjing Zheng. 2021. Multi-view Denoising Graph Auto-Encoders on Heterogeneous Information Networks for Cold-start Recommendation. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. 2338–2348.

[39] Feng Zhu, Yan Wang, Chaochao Chen, Guanfeng Liu, and Xiaolin Zheng. 2020. A Graphical and Attentional Framework for Dual-Target Cross-Domain Recommendation.. In *IJCAI*. 3001–3008.

[40] Feng Zhu, Yan Wang, Chaochao Chen, Jun Zhou, Longfei Li, and Guanfeng Liu. 2021. Cross-domain recommendation: challenges, progress, and prospects. *arXiv preprint arXiv:2103.01696* (2021).

[41] Qiannan Zhu, Xiaofei Zhou, Zeliang Song, Jianlong Tan, and Li Guo. 2019. Dan: Deep attention neural network for news recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 5973–5980.

[42] Yongchun Zhu, Kaikai Ge, Fuzhen Zhuang, Ruobing Xie, Dongbo Xi, Xu Zhang, Leyu Lin, and Qing He. 2021. Transfer-meta framework for cross-domain recommendation to cold-start users. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1813–1817.

[43] Yongchun Zhu, Zhenwei Tang, Yudan Liu, Fuzhen Zhuang, Ruobing Xie, Xu Zhang, Leyu Lin, and Qing He. 2022. Personalized transfer of user preferences for cross-domain recommendation. In *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*. 1507–1515.