CATN: Cross-Domain Recommendation for Cold-Start Users via Aspect Transfer Network

Cheng Zhao[1], Chenliang Li[2]★, Rong Xiao[3], Hongbo Deng[3], Aixin Sun[4]

1. State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China

2. School of Cyber Science and Engineering, Wuhan University, Wuhan, China

3. Alibaba Group, Hangzhou, China

4. School of Computer Science and Engineering, Nanyang Technological University, Singapore

{zhaocheng_whuer,cllee}@whu.edu.cn,xiaorong.xr@taobao.com,dhb167148@alibaba-inc.com,axsun@ntu.edu.sg

ABSTRACT

In a large recommender system, the products (or items) could be in many different categories or domains. Given two relevant domains (*e.g., Book* and *Movie*), users may have interactions with items in one domain but not in the other domain. To the latter, these users are considered as cold-start users. How to effectively transfer users' preferences based on their interactions from one domain to the other relevant domain, is the key issue in cross-domain recommendation. Inspired by the advances made in review-based recommendation, we propose to model user preference transfer at aspect-level derived from reviews. To this end, we propose a **c**ross-domain recommendation framework via **a**spect **t**ransfer **n**etwork for cold-start users (named CATN). CATN is devised to extract multiple aspects for each user and each item from their review documents, and learn aspect correlations across domains with an attention mechanism. In addition, we further exploit auxiliary reviews from like-minded users to enhance a user's aspect representations. Then, an end-to-end optimization framework is utilized to strengthen the robustness of our model. On real-world datasets, the proposed CATN outperforms SOTA models significantly in terms of rating prediction accuracy. Further analysis shows that our model is able to reveal user aspect connections across domains at a fine level of granularity, making the recommendation explainable.

KEYWORDS

Cold-Start Recommender Systems; Aspect-based Recommendation, Deep Learning

★Chenliang Li is the corresponding author.

1 INTRODUCTION

Recommender systems play vital roles in various e-commerce platforms. Traditional collaborative filtering methods recommend items to users mainly based on their historical feedbacks. However, these approaches become less effective for new users, *i.e.,* cold-start users, who have no historical feedbacks. Recently, cross-domain recommendation has gained wide attention [18, 40]. Given two relevant domains (*e.g., Book* and *Movie*), users may have historical interactions in one domain (*i.e.,* source domain), but not the other (*i.e.,* target domain). To the target domain, these users are considered as cold-start users. However, as the two domains are relevant, feedbacks in the source domain could be leveraged to provide meaningful recommendations in target domain.

The core task of cross-domain recommendation is user preference mapping between the two relevant domains. To achieve the mapping, existing approaches such as EMCDR [22], CDLFM [29] and RC-DFM [9] encode users' preference into single vectors, then conduct cross-domain mapping as a whole. Illustrated in Figure 1, existing solutions learn user/item representations in source domain and target domain respectively. Then, cross-domain representation mapping is learned based on the overlapping users. Note that, the direct mapping between user representations of source and target domains cannot explicitly capture users' diverse yet fine-grained preferences in different domains. For example, a user who prefers Chinese kung fu novels is more likely to be fund of Chinese ancient dramas.

In our study, we assume users' preferences are multi-faceted, *e.g., plot*, *text description*, *scene* in *Book* and *Movie* domains. Modeling these fine-grained semantic aspects and exploring their mutual relationships across domains, would lead to better user preference understanding and explainable recommendation. To this end, we aim to exploit user/item reviews for cross-domain aspect correlations. In recent years, there has been a surge of approaches utilizing user/item reviews for aspect-based recommendation (*i.e.,* rating prediction for a given user-item pair) [3–5, 19]. Inspired by their encouraging performance, we propose to explore users' preferences based on the aspects generated from reviews across domains.

In this paper, we propose a **c**ross-domain recommendation framework for cold-start users via **a**spect **t**ransfer **n**etwork, named CATN. In source domain, we represent a user by a *user document* which contains all reviews written by this user, and an item by an *item*
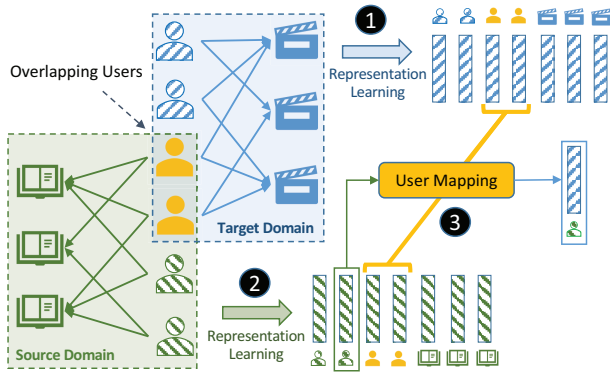
Figure 1: Existing workflow in cross-domain recommendation for cold-start users (*Best viewed in color*).

document which contains all reviews it receives. The same applies in target domain. An overlapping user therefore will have two user documents, one in source domain and the other in target domain. To extract aspects mentioned in user and item documents, we utilize an aspect-specific gate mechanism over a convolutional layer. Then, global cross-domain aspect correlations are identified and weighted through attention mechanism, for preference estimation. To support review-based knowledge transfer, we introduce a novel cross-domain review-based preference matching procedure with two learning flows. The illustration of these two learning flows is shown in Figure 2. Specifically, for a given overlapping user and an item in the target domain, we utilize the user's review document in source domain and the item's review document in target domain to perform rating prediction, and vice versa. These two learning flows are proceeded in turn with the guidance of the global cross-domain aspect correlations. Considering review scarcity [31] and the small number of overlapping users [17], we further enhance user representation by an additional *user auxiliary document* for each user. An auxiliary document contains all reviews written by the like-minded users, *i.e.,* the users who give the same rating to the same item as the current user. The auxiliary documents are also utilized in aspect extraction.

We summarize our key contributions as follows. We propose a novel deep recommendation model for cold-start users, by bridging multiple user's inherent traits via reviews in different domains. To the best of our knowledge, this is the first attempt to learn cross-domain aspect-level preference matching, in an end-to-end learning fashion. Through extensive experiments conducted on three pairs of real-world datasets, we demonstrate that CATN performs significantly better than state-of-the-art (SOTA) alternatives. We also conduct detailed analysis to validate the benefit introduced by each component of CATN, and show how CATN works at a fine-grained semantic level.

2 RELATED WORK

Our work is related to two subareas of recommender systems: cross-domain recommendation, and aspect-based recommendation. Next, we briefly review the works in each subarea.

2.1 Cross-Domain Recommendation

By leveraging relevant source domain as auxiliary information, a surge of solutions are proposed to address the data sparsity and cold-start problems for the target domain. At the very beginning, CMF [25] proposes to achieve knowledge integration across domains by concatenating multiple rating matrices and sharing user factors across domains. Then Temporal-Domain CF [18] shares the static group-level rating matrix across temporal domains. Later, CDTF [16] is proposed to capture the triadic relation of user-item-domain by tensor factorization. These collaborative filtering based works suffers severely from the data sparsity problem when considering different domains as a whole.

In recent years, with the revival of deep learning techniques, many deep learning-based models are proposed to enhance knowledge transfer. EMCDR [22] explicitly maps user representations from different domains via a multi-layer fully connected neural network. DCDCSR [40] further extends EMCDR by generating benchmark factors to solve cross-domain and cross-system problems. CoNet [15] is proposed to train a deep cross-stitch network for enhancing the recommendation on both domains simultaneously. PPGN [38] leverages the user-item interaction graph to capture the process of user preference propagation. DARec [35], equipped with an adversarial learning process, is proposed for user-item rating prediction. π-Net [21] is devised for shared-account cross-domain sequential recommendation.

To avoid the leak of user privacy, NATR [10] chooses to transfer only the item embeddings across domains. SSCDR [17] investigates the distribution of cross-domain overlapping users in real-world scenarios, and come up with a semi-supervised mapping approach to perform recommendation for cold-start users. CDLFM [29] modifies the matrix factorization and mapping process by exploiting the user neighborhoods. Another line of cross-domain recommender systems is clustering-based, which has also achieved good performance. C^3R [8] leverages users' multiple social media sources to boost the performance of venue recommendation. CDIE-C [30] enhances item embedding learning by means of cross-domain co-clustering.

Nevertheless, many of the above solutions only consider rating records while ignoring other complementary yet fertile information, *e.g.,* reviews. MVDNN [7] maps users' and items' auxiliary information to a latent space where the similarity between users and their preferred items is maximized. To combine the strength from both ratings and reviews, RB-JTF [26] transfers users' preference by a joint tensor factorization derived from the reviews. RC-DFM [9] trains user or item factors with a review-fused SDAE, which achieves the SOTA performance for cold-start user recommendation.

Existing review-based transfer solutions have earned substantial improvement over traditional interaction-based methods. However, these works still have many drawbacks to be overcome. As discussed in Section 1 and illustrated in Figure 1. Existing solutions learn users and items representations in source domain and target domain respectively (steps 1 and 2 in Figure 1). Then they learn the cross-domain representation mapping based on the overlapping users (step 3 in Figure 1). This mapping cannot explicitly distinguish the fine-grained semantic characteristics. Further, the

pipelined learning process could easily accumulate and magnify noisy information produced by the sub-optimal learning in the intermediate steps. We therefore propose a completely different network architecture, to capture and align the fine-grained user preferences between source and target domains at aspect level, through reviews, and in an end-to-end fashion.

2.2 Aspect-based Recommendation

Reviews reflect a user's purchased experience, and have shown to be effective in addressing the sparsity problem in recommendation. Nowadays, review-based recommender systems have become a pivotal building block for recommendation in single-domain [4, 19, 20, 24, 27, 32, 34, 36, 39]. Within review-based recommender systems, aspect-based recommender systems, which model the fine-grained relations between user preferences and item characteristics, have drawn great attention recently. In general, existing solutions for aspect-based recommender systems can be divided into two main categories.

Solutions in the first category extract aspects and sentiments from reviews by utilizing external NLP toolkits. Example solutions include MTER [37], TriRank [12], LRPPM [2], SULM [1] and EFM [28]. The performance of such solutions therefore are highly dependent on the quality of the external toolkits used in the process.

The second category of solutions fulfills automatic aspects extraction, with an internal model component. For example, JMARS [5] utilizes topic modeling to learn multiple aspect representations. Following JMARS, FLAME [33] and AFLM [3] are proposed to model aspect-level user preferences and item characteristics through an integrated hidden topic learning process. However, the static representations learned by these methods are incapable of modeling dynamic and complex relationships between users and items. To dynamically model the relation encoded by different user-item pairs, ANR [4] uses a co-attention mechanism to infer the importance of different aspects with respect to a given user-item pair. More recently, CARP [19] proposes a capsule network to conduct rating prediction and provide interpretability in a fine-grained manner.

Note that, solutions in both categories focus on the single-domain recommendation. These methods cannot handle the cold-start user whose historical interactions are not available in the target domain. In this work, we make the first attempt to complete this picture by designing a cross-domain aspect transfer network to achieve recommendation for cold-start users in target domain.

3 THE CATN FRAMEWORK

We start with the problem setting of cross-domain recommendation for cold-start users. Then, we provide an overview of CATN along with the motivation behind its each component. After presenting all the components, we go through the optimization process.

3.1 Problem Formulation

We use $\mathcal{D}_s$ and $\mathcal{D}_t$ to denote source domain and target domain respectively. Note that a domain includes its users, items, and interactions (*e.g.,* ratings and reviews) between users and items. Let U^o be the set of overlapping users, who have historical interactions with items in both $\mathcal{D}_s$ and $\mathcal{D}_t$. U^{cs} denotes the set of cold-start users who have interactions with items in $\mathcal{D}_s$, but not with items
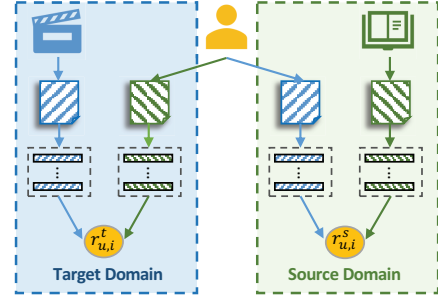

Figure 2: Two learning flows in CATN (*Best viewed in color*).

in $\mathcal{D}_t$. For a given cold-start user $u \in U^{cs}$, our task is to estimate the rating $\hat{r}_{u,i}$ that user u would give to an item i in $\mathcal{D}_t$.

3.2 Overview of CATN

The overall structure of CATN is illustrated in Figure 3. Its structure consists of three components: *Aspect Extraction*, *Auxiliary Reviews Enhancement*, and *Cross-Domain Aspect Correlation Learning*. Because our task is to achieve review-based cross-domain preference transfer, the overall procedure for rating prediction differs fundamentally from the existing review-based recommendation systems in single-domain. Here, the ratings and the reviews of the overlapping users U^o from both $\mathcal{D}_s$ and $\mathcal{D}_t$ are used for model training.

In source domain, we represent a user by a *user document D_u*, and an item by an *item document D_i*. Similarly, each user and each item in the target domain has a user document and an item document, respectively. An overlapping user will have two user documents, one from source domain D_u^s and the other from target domain D_u^t. We use superscript "s" and "t" to indicate the source domain and target domain for a clear presentation. Recall that an overlapping user has interactions with items in both source and target domains. For a given overlapping user $u \in U^o$, as shown in Figure 2, we devise a cross-domain review-based preference matching procedure with two learning flows: 1) her user document D_u^s in source domain and the item document D_i^t for item i in target domain are utilized in model training to match $r_{u,i}^t$ in target domain; and 2) her user document D_u^t in target domain and the item document D_i^s of an item i in source domain are utilized to match $r_{u,i}^s$ in source domain.

The matching of user preference between source and target domains are achieved at *aspect level*, derived from the user and item documents, as shown in Figure 3. Note that, in addition to user document, we also utilize an auxiliary review document for each user. This auxiliary review document contains reviews written by like-minded users, to be detailed shortly. Aspects derived from the two kinds of user documents are merged. Then, a cross-domain aspect correlation learning will distinguish the more correlated aspect-pair across domains to conduct rating predictions. Next, we detail the aspect extraction process.

3.3 Aspect Extraction

To extract aspects, the same process is applied to user document D_u and item document D_i, in both source and target domains. As the procedure is the same, we take D_u as a running example.
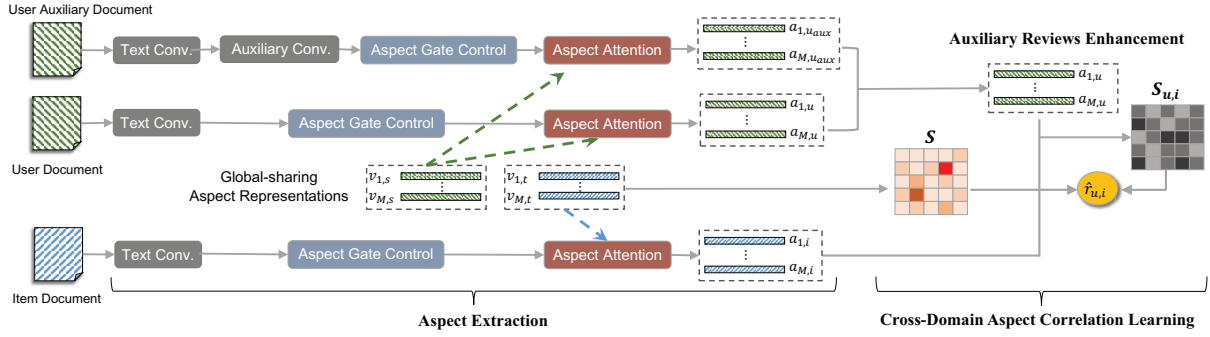
Figure 3: The architecture of CATN (*Best viewed in color*).

Text Convolution. Given a user document $D_u = [w_1, w_2, .., w_l]$, we first project each word to its embedding representation: $\mathbf{E}_u = [\mathbf{e}_1, \mathbf{e}_2, .., \mathbf{e}_l]$, $\mathbf{e}_j \in \mathbb{R}^d$, where l is the document length and d is the word embedding dimension. In order to capture the context information around each word, we perform a convolution operation with an *ReLU* activation function. Here, n convolution filters with the same sliding window of size s are applied over matrix $\mathbf{E}_u$ to extract contextual features. The resultant feature matrix is $\mathbf{C}_u = [\mathbf{c}_{1,u}, \mathbf{c}_{2,u}, .., \mathbf{c}_{l,u}]$, where $\mathbf{c}_{j,u} \in \mathbb{R}^n$ is the latent contextual feature vector for j-th word.

Aspect Gate Control. The contextual features $\mathbf{c}_{j,u}$ extracted for j-th word can be considered as a composition of multiple semantic aspects. Here, we further utilize an aspect-specific gate mechanism to identify which features are relevant to each aspect. Specifically, for m-th aspect, the aspect-specific features $\mathbf{g}_{m,j,u}$ of word w_j are extracted as follows:

$$\mathbf{g}_{m,j,u} = (\mathbf{W}_m \mathbf{c}_{j,u} + \mathbf{b}_m) \odot \sigma(\mathbf{W}_m^g \mathbf{c}_{j,u} + \mathbf{b}_m^g) \qquad (1)$$

where σ is the sigmoid activation function, $\odot$ is the element-wise product operation. $\mathbf{W}_m, \mathbf{W}_m^g \in \mathbb{R}^{k \times n}$ and $\mathbf{b}_m, \mathbf{b}_m^g \in \mathbb{R}^k$ denote the transform matrices and bias vectors respectively for the m-th aspect. k is the latent dim of aspect representation. Here, the second term on the right hand side of Equation 1 works as a soft *on-off* switch controlling which latent feature is relevant to the aspect. Consequently, we get M aspect-specific words contextual features $\mathbf{G_u}$, which are leveraged for further aspect extraction.

$$\mathbf{G_u} = [\mathbf{G}_{1,u}, \mathbf{G}_{2,u}, .., \mathbf{G}_{M,u}], \qquad (2)$$

$$\mathbf{G}_{m,u} = [\mathbf{g}_{m,1,u}, \mathbf{g}_{m,2,u}, .., \mathbf{g}_{m,l,u}] \qquad (3)$$

Aspect Attention. Reviews from different domains put emphasis on different aspects. For instance, *Book* domain tends to include *plots* and *figures*, while *Movie* domain tends to include *actors* and *special effects*. Accordingly, we design two matrices of global-sharing aspect representations in $\mathcal{D}_s$ and $\mathcal{D}_t$. They are denoted as $\mathbf{V}_s = [\mathbf{v}_{1,s}, ..., \mathbf{v}_{M,s}]$ and $\mathbf{V}_t = [\mathbf{v}_{1,t}, ..., \mathbf{v}_{M,t}]$, for source and target domains respectively. $\mathbf{V}_s$ and $\mathbf{V}_t$ serve as the query to guide the aspect extraction. Concretely, the representation $\mathbf{a}_{m,u}$ of the

m-th aspect extracted from $\mathbf{G}_{m,u}$ is derived as follows:

$$\mathbf{a}_{m,u} = \sum_{j=1}^{l} \beta_{m,j,u} \mathbf{g}_{m,j,u} \qquad (4)$$

$$\beta_{m,j,u} = \frac{exp(\mathbf{g}_{m,j,u}^\top \mathbf{v}_{m,s})}{\sum_{i=1}^{l} exp(\mathbf{g}_{m,i,u}^\top \mathbf{v}_{m,s})} \qquad (5)$$

Here, $\beta_{m,j,u}$ indicates the importance of word w_j towards the m-th aspect. Consequently, we can obtain the representations of M aspects from D_u, constituting the aspect matrices $\mathbf{A}_u = [\mathbf{a}_{1,u}, .., \mathbf{a}_{M,u}]$. Following the same precedure, we extract M aspects from D_i: $\mathbf{A}_i = [\mathbf{a}_{1,i}, .., \mathbf{a}_{M,i}]$. It is worthwhile to highlight that the parameters for aspect extraction for D_u and D_i are shared in each learning flow, though D_u and D_i are built with the reviews in different domains. Also a distinct set of parameters is used in each learning flow. Since we aim to map the aspect across domains, $\mathbf{V}_s$ and $\mathbf{V}_t$ are shared in their corresponding domains respectively.

3.4 Auxiliary Reviews Enhancement

Note that the proportion of overlapping users across domains are usually a very small number [17]. This data sparsity problem is further aggravated with review scarcity, that the user documents contain incomplete and short reviews [31].

To overcome these limitations, we choose to make full use of the interactions of similar non-overlapping users. We extract auxiliary reviews from *like-minded* users as done in [31]. Specifically, for a given user-item pair, an auxiliary review is a review written by another user with the same rating score as the target user did for this item. For user u, her auxiliary document $D_{u_{aux}}$ is formed by merging the auxiliary reviews of the historical items purchased by user u in the same domain. Note that, we only consider the auxiliary reviews from non-overlapping users, which could increase the diversity of the training data. With this data augmentation strategy, our model can still be optimized in a good shape, even when the overlapping users are very few.

One natural way to exploit the auxiliary document is to follow the same aspect extraction process and simply merge it with $\mathbf{A}_u$. However, this kind of solution ignores the fact that an auxiliary document is formed by different users who would have different language styles and different preference focuses with target user, thus may result in incompatible features. Reported in [32], stacking

a CNN network on top of the contextual matrix is effective on rating prediction, especially when the semantics in the document are incoherent. Hence, on top of *Text Convolution* used in the previous aspect extraction process, we add another convolutional layer in processing auxiliary documents, as shown in Figure 3.

$$c^i_{h,u_{aux}} = ReLU(\mathbf{W}^i_{aux} * \mathbf{H}_{u_{aux}}[h - \frac{s-1}{2} : h + \frac{s-1}{2}] + b^i_{aux}) \quad (6)$$

where $*$ is the convolution operator, $\mathbf{W}^i_{aux} \in \mathbb{R}^{s \times n}$ is the convolution weight matrix, b^i_{aux} is the bias term, and $\mathbf{H}_{u_{aux}}$ is the feature matrix extracted by *Text Convolution* in Section 3.3. Similarly, we form the abstract feature matrix $\mathbf{C}_{u_{aux}} = [\mathbf{c}_{1,u_{aux}}, \mathbf{c}_{2,u_{aux}}, \ldots, \mathbf{c}_{l,u_{aux}}]$, where $\mathbf{c}_{j,u_{aux}} \in \mathbb{R}^n$. The same *Aspect Gate Control* and *Aspect Attention* processes are conducted to get aspect matrices $\mathbf{A}_{u_{aux}}$ from $D_{u_{aux}}$. To update $\mathbf{A}_u$ with $\mathbf{A}_{u_{aux}}$ effectively, we adopt a gate mechanism based on the element-wise interactions of the corresponding aspects:

$$\mathbf{g}_{aux} = \sigma(\mathbf{W}^1_f[(\mathbf{A}_u - \mathbf{A}_{u_{aux}}) \oplus (\mathbf{A}_u \odot \mathbf{A}_{u_{aux}})] + \mathbf{b}^1_f), \quad (7)$$

$$\mathbf{A}_u = tanh(\mathbf{W}^2_f[\mathbf{A}_u \oplus (\mathbf{g}_{aux} \odot \mathbf{A}_{u_{aux}})] + \mathbf{b}^2_f) \quad (8)$$

where $\oplus$ is the concatenation operation, $\mathbf{W}^1_f, \mathbf{W}^2_f \in \mathbb{R}^{k \times 2k}$ are transform matrices, $\mathbf{b}^1_f, \mathbf{b}^2_f \in \mathbb{R}^k$ are the bias vectors. The aspect representation $\mathbf{A}_u$ is updated to better profile user u.

3.5 Cross-Domain Aspect Correlation Learning

Now, we have abstract aspect features $\mathbf{A}_u$ and $\mathbf{A}_i$ for user u and item i respectively. Intuitively, the rating prediction could be the aggregation of the semantic matchings between two aspects in $\mathbf{A}_u$ and $\mathbf{A}_i$ respectively. However, the matching scores would only reflect the semantic relatedness between two aspects for the specific user-item pair. Because not all aspect pairs are equally important, it is beneficial to identify global cross-domain aspect correlations. Then we can highlight the important aspect pairs at domain level for better rating prediction. To this end, we design a simple but effective method for cross-domain preference matching. Recall that we utilize a set of global aspect representations $\mathbf{V}_s$ and $\mathbf{V}_t$ to guide the aspect extraction. Here, we utilize these static aspect representations to calculate the global cross-domain aspect correlation matrix $\mathbf{S}$ as follow:

$$\mathbf{S} = LeakyReLU(\mathbf{V}^\top_s \mathbf{W} \mathbf{V}_t) \quad (9)$$

where $\mathbf{S}(p, q)$ reflects the importance of preference transfer based on aspect p from the source domain and aspect q from the target domain. $\mathbf{S} \in \mathbb{R}^{M \times M}$, $\mathbf{W} \in \mathbb{R}^{k \times k}$ is a learnable matrix for affinity projection. The *LeakyReLU* activation function is adopted to support the sparse aspect correlation across domains by setting the corresponding α to be a very small value (*e.g.*, 0.01).

We then calculate the semantic matching between each aspect pair between $\mathbf{A}_u$ and $\mathbf{A}_i$ as follows:

$$\mathbf{S}_{u,i} = \mathbf{A}^\top_u \mathbf{W} \mathbf{A}_i \quad (10)$$

Similar to Equation 9, $\mathbf{S}_{u,i}(p, q)$ reflects the matching degree between the corresponding aspects; $\mathbf{W}$ is shared for affinity projection.

Dataset	#users	#items	#ratings	density
Book	126,666	63,202	3,494,976	0.044%
Movie (Movies and TV)	27,822	12,287	779,376	0.228%
Music (CDs and Vinyl)	11,053	7,710	296,188	0.348%

At last, we utilize $\mathbf{S}$ as the attention weights to aggregate the pairwise aspect matchings as the final rating prediction.

$$\mathbf{S}^r_{u,i} = \mathbf{S} \odot \mathbf{S}_{u,i} \quad (11)$$

$$\hat{r}_{u,i} = \frac{1}{M * M} \sum_{p=1}^{M} \sum_{q=1}^{M} \mathbf{S}^r_{u,i}(p, q) + b_u + b_i \quad (12)$$

Here, b_u and b_i are the user bias and item bias respectively.

3.6 Optimization Strategy

For model training, we utilize the interactions made by the overlapping users in source and target domains for parameter optimization. Let O_s or O_t be a batch of observed user-item rating pairs in $\mathcal{D}_s$ or $\mathcal{D}_t$ respectively, restricted to U^o only. The loss function of $\mathcal{L}_s$ and $\mathcal{L}_t$ can be defined as follows:

$$\mathcal{L}_* = \frac{1}{|O_*|} \sum_{(u,i) \in O_*} (r_{u,i} - \hat{r}_{u,i})^2 + \lambda ||\Theta_*|| \quad (13)$$

where symbol $*$ could refer to s or t, λ is the regularization coefficient, and Θ_* are the trainable parameters. The two learning flows (*i.e.*, predicting $r_{u,i}$ in target domain by using D_u in source domain and D_i in target domain, and predicting $r_{u,i}$ in source domain by using D_u in target domain and D_i in source domain) are performed in turn batch after batch. Each training batch is composed of shuffled O_s and O_t at a fixed proportion, w.r.t. $|O_s|/|O_t| = |R_s|/|R_t|$, where $|R_s|$ and $|R_t|$ denote the number of ratings made by U^o in $\mathcal{D}_s$ and $\mathcal{D}_t$ respectively. We adopt *Adam* as the optimizer to update the parameters.

4 EXPERIMENTS

4.1 Datasets

To evaluate our model against state-of-the-art baselines, we conduct experiments on the *Amazon* review dataset [11]. Among the largest categories,[1] we choose three relevant ones as three domains, namely, *Book*, *Movie* (named "Movies and TV" in Amazon) and *Music* (named "CDs and Vinyl" in Amazon). In each domain, we remove the interaction records that are without review text, then filter out the users with fewer than 10 interactions and the items with fewer than 30 interactions following earlier studies [14, 17]. The detailed statistics of each domain is reported in Table 1.

As the three domains are relevant to each other, we construct three cross-domain scenarios in pairs. In each scenario, we choose the domain with more users as $\mathcal{D}_s$ and the other as $\mathcal{D}_t$. Following the settings in [17], we randomly sample 50% of the overlapping users to be cold-start users, *i.e.*, their interactions in $\mathcal{D}_t$ are not seen by the models, but are used for validation and testing purposes (specifically, 30% are set for test users and 20% are set for validation

---

[1]http://jmcauley.ucsd.edu/data/amazon/

Table 2: Statistics of the three cross-domain recommendation scenarios. η donotes the ratio of overlapping users included in the training set.

Scenario	domain	dataset	#overlap. users	overlap. users ratio	η=100%	η=50%	η=20%	η=10%	η=5%	#vali. users	#test users
Scenario 1	$\mathcal{D}_s$	Book	6,074	4.795%	3,037	1,518	607	303	151	1,214	1,823
	$\mathcal{D}_t$	Movie		21.832%							
Scenario 2	$\mathcal{D}_s$	Movie	2,782	9.999%	1,391	695	278	139	69	556	835
	$\mathcal{D}_t$	Music		25.170%							
Scenario 3	$\mathcal{D}_s$	Book	1,705	1.346%	853	426	170	85	42	340	512
	$\mathcal{D}_t$	Music		15.426%							

users). The remaining 50% of overlapping users are used for training purpose. In order to simulate different ratios of overlapping users, we building our training set by randomly including a certain fraction $\eta \in \{100\%, 50\%, 20\%, 10\%, 5\%\}$ of the remaining 50% overlapping users. The detailed statistics of each cross-domain scenario is reported in Table 2.

4.2 Baseline Methods

We compare against the following baselines, including the traditional ones and recent state-of-the-arts.

- **CMF** [25] is a simple and well-known method for cross-domain recommendation by sharing the user factors and factorizing joint rating matrix across domains.
- **EMCDR** [22] is the first to propose the three-step optimization paradigm by training matrix factorization in both domains successively then utilizing multi-layer perceptrons to map the user latent factors.
- **CDLFM** [29] modifies matrix factorization by fusing three kinds of user similarities as a regularization term based on their rating behaviors. A neighborhood-based mapping approach is used to replace the previous multi-layer perceptrons, by considering similar users and the gradient boosting trees (GBT) based ensemble learning method.
- **DFM** [9] is a simple version of RC-DFM [9]. It leverages the work of aSDAE [6] to generate user representations from rating matrix, with multi-layer perceptrons to conduct mapping as well.
- **R-DFM** [9] is another variant of RC-DFM [9][2]. It combines the rating records and the reviews by an extended aSDAE to enhance the user/item representations. The mapping part is also multi-layer perceptrons.
- **ANR** [4] is a state-of-the-art review-based single-domain approach by performing aspect matching for user-item pair. Here we conduct recommendation by leveraging their corresponding reviews in the source domain directly, after training the model purely on the target domain.

4.3 Experimental Setup

We preprocess both user and item documents in all datasets, following the related studies [32?]: 1) remove stop words and words with high document frequency (i.e., relative document frequency above 0.5); 2) choose the top 20,000 words as vocabulary according to their tf-idf score and remove other words from the raw documents; 3) amputate (pad) the long (short) documents to the same length of 500 words. We utilize the 300-dimension word embeddings pre-trained in Google News[3] [23] to get the initial embedding vector for each word.

We apply grid search to tune the hyper-parameters for all the methods based on the setting strategies reported in their papers. The final performances of all methods are reported over 5 runs.

For CATN[4], the number of convolution filters n = 50, window size s = 3. The batch size (number of $O_s \cup O_t$) is 256. The dropout strategy is applied to ignore a small percent of values in aspect representations randomly during the training process. The keep probability of dropout is set to be 0.8, and we choose learning rate to 0.001 for model training. The latent dimension size k is optimized from {16, 32, 64, 128}, and the aspect number M is optimized from {3, 5, 7, 9}.

For evaluation metric, we use *MSE* as performance metric, which is widely adopted in many related works for performance evaluation [4, 19, 27, 32], formulated as:

$$MSE = \frac{1}{|O|} \sum_{(u,i) \in O} (r_{u,i} - \hat{r}_{u,i})^2$$

where O is the cold-start user validation set for parameter selection or test set for performance comparison.

4.4 Results and Discussion

The overall results of all methods over the three cross-domain recommendation scenarios are reported in Table 3. We made the following observations from the results.

First of all, CATN outperforms all baselines significantly on all cross-domain recommendations, and in terms of different ratios of overlapping users in all settings. This result demonstrates the superiority of our review-based recommendation for cold-start users in cross-domain setting.

It's no surprise that CMF consistently yields the worst performance on all evaluations. Learning user representations simply

---

[2]In the RC-DFM paper, the item content is fused into another aSDAE to let it close to its review-based representation. However, the authors do not explain the detailed method to obtain item content. Moreover, the improvement of RC-DFM over R-DFM is very small, so we choose R-DFM for comparison.

[3]https://code.google.com/archive/p/word2vec/
[4]Our implementation is available at https://github.com/WHUIR/CATN

Table 3: Performance comparison on three recommendation scenarios in terms of *MSE*. The best and second best results are highlighted in boldface and underlined respectively. ▲% denotes the relative improvement of CATN over the best SOTA algorithm. All reported improvements over baseline methods are statistically significant at a 0.05 level.

| Scenario		Scenario 1					Scenario 2					Scenario 3				
$\mathcal{D}_s \to \mathcal{D}_t$		$Book \to Movie$					$Movie \to Music$					$Book \to Music$				
Method $\diagdown$ η		100%	50%	20%	10%	5%	100%	50%	20%	10%	5%	100%	50%	20%	10%	5%
CMF		1.167	1.169	1.179	1.179	1.181	1.139	1.140	1.158	1.167	1.173	0.939	0.942	0.962	0.967	0.970
EMCDR		1.129	1.138	1.142	1.140	1.148	1.116	1.138	1.144	1.172	1.175	0.924	0.927	0.934	0.936	<u>0.937</u>
CDLFM		1.126	1.130	1.135	1.138	1.144	1.115	1.133	1.145	1.169	1.171	0.918	0.925	0.930	0.931	0.951
DFM		1.141	1.143	1.149	1.150	1.156	1.136	1.158	1.162	1.166	1.175	0.923	0.929	0.933	0.941	0.952
R-DFM		1.132	1.135	1.141	1.146	1.152	1.128	1.143	1.146	<u>1.150</u>	1.166	0.911	0.917	0.928	0.936	0.943
ANR		<u>1.123</u>	<u>1.127</u>	<u>1.130</u>	<u>1.135</u>	<u>1.137</u>	<u>1.122</u>	<u>1.137</u>	<u>1.142</u>	1.155	<u>1.160</u>	<u>0.895</u>	<u>0.903</u>	<u>0.912</u>	<u>0.919</u>	0.940
CATN		**1.049**	**1.072**	**1.079**	**1.093**	**1.097**	**1.042**	**1.075**	**1.102**	**1.126**	**1.144**	**0.862**	**0.868**	**0.875**	**0.896**	**0.899**
▲%		6.59	4.88	4.51	3.70	3.52	6.55	5.45	3.50	2.09	1.38	3.69	3.88	4.06	2.50	4.06
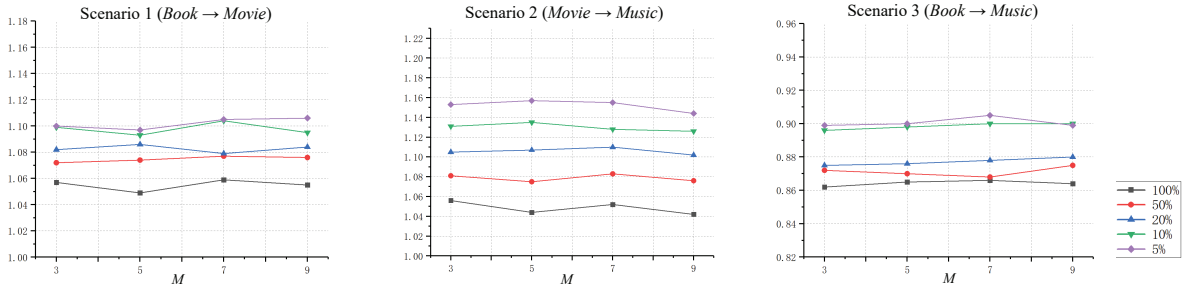
Figure 4: Impact of number of aspects M in CATN.

by factorizing a joint matrix is not adequate, which is also consistent with what has been observed in earlier studies [9, 22, 29]. CDLFM makes some improvements to the user factors learning and the cross-domain mapping processes, which leads to conspicuous improvements over EMCDR. R-DFM modifies DFM ulteriorly by fusing user reviews. However, none of them ever achieves the best result, which verifies the drawbacks of the straightforward three optimization process as shown in Figure 1.

For DFM and R-DFM, according to our experiments, the results suffer from declination compared to EMCDR. This is because, aS-DAE takes original rating vectors as input, which can be over a hundred of thousand dimensions in our dataset. In this case, millions of training parameters need to be optimized, which makes the model rather complicated to converge[5] and yield inferior results. Although ANR is not designed for cross-domain scenarios, it maintains competitive results over the other baselines, confirming the usefulness of review information for the recommendation task.

From the results, we observe that while the methods based on three-step optimization are sensitive to η, especially when the ratio is low (10% or 5%), our CATN shows more robust performance. As η gets lower, the overlapping users get fewer. Existing cross-domain mapping cannot be well trained because of the lack of training

---

[5]To avoid the issue of parameters explosion, the authors of [9] preprocess the data with the *120-cores* settings (*i.e.,* filter out the items with fewer than 120 interactions). However, it does not fit real-world sparse recommender scenarios.

instances, resulting in an inferior result. On the contrary, CATN utilizes a simple but effective way to emphasizes the transfer of cross-domain aspects, instead of the user representations directly. In this way, CATN reduces the impact by η to a large extent.

5 MODEL ANALYSIS

We now present detailed analysis of the proposed CATN model. We first investigate the impact of hyper-parameter settings (*i.e.,* M) to the performance of CATN. Next, we conduct three ablation studies to analyze how different components in our proposed model contribute to the overall results. Lastly, study cases are shown to give explainable analysis of the cross-domain aspect transfer process.

5.1 Aspect Number Sensitivity

Figure 4 plots the effect of varying $M \in \{3, 5, 7, 9\}$ for CATN across multiple evaluations settings, with different preset η values. In general, a small M leads to coarse aspects, while a large M leads to fine-grained aspects. However, as we discussed earlier, not all aspects from source and target domains would match and participate in the preference transfer, and the attention mechanism would learn optimal weights between matching aspects. In this sense, the varying of M would only affect the number of aspects in source and target domains and does not affect much of the preference transfer. As shown in the plot, given the same setting (*i.e.,* a fixed

Table 4: Performance comparison of the three model variants on three recommendation scenarios.

Scenario		Scenario 1					Scenario 2					Scenario 3				
$\mathcal{D}_s \rightarrow \mathcal{D}_t$		$Book \rightarrow Movie$					$Movie \rightarrow Music$					$Book \rightarrow Music$				
Method η		100%	50%	20%	10%	5%	100%	50%	20%	10%	5%	100%	50%	20%	10%	5%
CATN-basic		1.103	1.109	1.117	1.122	1.127	1.114	1.127	1.144	1.158	1.160	0.881	0.889	0.897	0.900	0.903
CATN-attn		1.084	1.102	1.109	1.116	1.121	1.074	1.103	1.131	1.153	1.157	0.880	0.893	0.895	0.899	0.901
CATN-separate		1.056	1.079	1.087	1.099	1.103	1.055	1.085	1.115	1.137	1.153	0.868	0.870	0.884	0.899	**0.899**
CATN		**1.049**	**1.072**	**1.079**	**1.093**	**1.097**	**1.042**	**1.075**	**1.102**	**1.126**	**1.144**	**0.862**	**0.868**	**0.875**	**0.896**	**0.899**

η in a particular cross-domain recommendation task), the performance fluctuations incurred by different M values are very small, suggesting that CATN is robust to this parameter setting.

On the other hand, the setting of η directly affects the number of overlapping users from whom the system learns the preference matching across domains. It is clear that more overlapping users lead to a better understanding of preferences across domains, hence better recommendation accuracy.

5.2 Ablation Study

Reflecting the intuition of CATN, we design a cross-domain review-based preference matching procedure with two learning flows. The learning process involves global-sharing aspect representations $\mathbf{V}_s$ and $\mathbf{V}_t$ to guide the aspect extraction. The global cross-domain aspect correlations $\mathbf{S}$ are exploited to give final predictions. In addition, auxiliary reviews from like-minded and non-overlapping users are exploited to enhance user aspect extraction, with the aim of alleviating the data sparsity issue. Accordingly, we come up with three variants of CATN as follows:

- **CATN-basic**: As the basic variant of CATN, it shares aspect extraction parameters in the two learning flows. We exclude $\mathbf{V}_s$ and $\mathbf{V}_t$ by replacing the attention mechanism with a **simple average** operation in Equation 4. The prediction only considers the aspect matchings. No auxiliary reviews are exploited in this variant.
- **CATN-attn**: In contrast to CATN-basic, we introduce the global-sharing aspect representations to fulfill aspect extraction, and the global cross-domain aspect correlations are taken into account. In other words, CATN-attn is a simplified version of CATN without domain-specific aspect extraction and auxiliary reviews.
- **CATN-separate**: In contrast to CATN-attn, we leverage two separate aspect extraction parameters in the two learning flows. In other words, CATN-separate is a simplified version of CATN by not including auxiliary reviews.

The results of the ablation studies on all evaluation settings are reported in Table 4. We make the following observations: 1) With reference to the results in Table 3, CATN-basic outperforms most baselines in all recommendation scenarios, demonstrating the effectiveness of the cross-domain aspect-based transfer approach; 2) CATN-attn gains some improvements over CATN-basic, which reveals the benefit of including global-sharing aspect representations; 3) CATN-separate outperforms the above variants, which shows the usefulness of distinct aspect extractions; and 4) As the
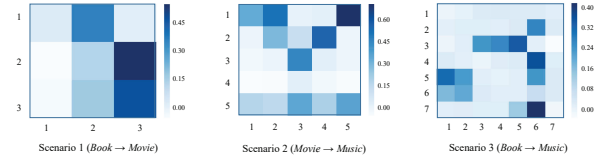
Figure 5: Global aspect correlation matrix S on three recommendation scenarios at $\eta = 50\%$.

integrated model, CATN improves the performance further by exploiting auxiliary reviews from like-minded and non-overlapping users. This observation suggests that auxiliary reviews are of vital value to alleviate the data sparsity issue.

5.3 Optimization Efficiency

Designed as an end-to-end learning framework, our proposed CATN not only overcomes the deficient three-step optimization process, but also speeds up the training time by optimizing only the ratings from overlapping users.

Specifically, in existing approaches, the third-step cross-domain transfer process cannot be conducted until the first two steps reach their optimal states, which is time-consuming. Besides, DFM and R-DFM contain massive parameters in terms of their Auto-Encoder component, thus hindering the convergence speed.

In our experiments, CATN spends about 600s (second) to reach the best validation performance in $Book \rightarrow Movie$ at $\eta = 50\%$, by using one Nvidia 1080 GPU. In contrast, it is 300s for CMF, 400s for ANR, 1000s for EMCDR, 1200s for CDLFM, over 1 hour for DFM and R-DFM. While our CATN achieves the best performance, it maintains a competitive training time over the other baselines, especially in terms of the review-based approach R-DFM.

5.4 Explainability Analysis

We further investigate whether CATN can discover meaningful aspect transfers across domains. To better visualize the aspects, we retrieve the top-5 words whose weights are the average of attention scores (*i.e.,* $\beta_{m,j,u}$ and $\beta_{m,j,i}$) in the document. Then we display the sentences containing these informative words for better understanding.

The user-item pair from each cross-domain recommendation setting at $\eta = 50\%$ is randomly sampled and displayed in Table 5 to offer semantic explanations. As Figure 5 illustrates, the global

Table 5: Example study of three user-item pairs from three recommendation scenarios at $\eta = 50\%$.

scenario 1 (Book → Movie): $r_{u,i} = 5.0$, $\hat{r}_{u,i} = 4.72$	
$A_u[2]$	...I enjoyed reading the book. It adds a lot to the movie. I think the ==biggest plot element== is that it really expands upon the Garthim-Master's ==character==...
$A_{u_{aux}}[2]$	...This is ==an interesting history==. Much of the book is ==interesting and readable==...
$A_i[3]$	...all the ==characters== playerd their ==roles== well and overall, it was a ==fun== movie to watch...
$A_u[3]$ in D_u^t	..I felt this movie was more than just a fluffy ==romantic comedy==. I was pleasantly surprised when...
target review $d_{u,i}$	...like a double romantic comedy...all of the actors are wonderful. Andy Griffith is absolutely charming as the "player" grandfather...
scenario 2 (Movie → Music): $r_{u,i} = 3.0$, $\hat{r}_{u,i} = 3.58$	
$A_u[1]$	...I think all criticism aside, this is a very ==interesting and engrossing== film...
$A_{u_{aux}}[1]$	...==tender plots== are always my favorite...
$A_i[5]$	...I like sting. And I like some ==renaissance== music...==lute== plus sting sounded like a good idea, but about half of one...
$A_u[5]$ in D_u^t	...I appreciate the ==technical merits== of the membership...
target review $d_{u,i}$	...As much as I appreciate when an artist stretches his ability, Sting should have reconsidered his venture into lute music...but enventually repetitive and unoriginal...
scenario 3 (Book → Music): $r_{u,i} = 4.0$, $\hat{r}_{u,i} = 4.33$	
$A_u[7]$	...This is a fairly entertaining novel that ==reads more like== a collection of loosely connected ==stories== held together by the common thread of...
$A_{u_{aux}}[7]$	...is the ==most grossly misogynist and sexist novel== I've read in a long time which is distressing considering...
$A_i[6]$	...because it's such a ==lovely, upbeat optimistic pop song==...what a pity that lyrics are just slightly above alternative teenage-pop...
$A_u[6]$ in D_u^t	...I've been a fan of ==trip hop== since I heard blue lines from massive attack and I've heard many bad imitations since then...
target review $d_{u,i}$	...overall it pleasantly comes together as a whole creating a lovely collection of what I could call downtempo triphop pop...

Table 6: Top-5 words for each aspect in *Example 1*. The "Aspect Labels" are manually generated based on our own interpretation.

Aspect	$A_u[1]$	$A_u[2]$	$A_u[3]$	$A_{u_{aux}}[1]$	$A_{u_{aux}}[2]$	$A_{u_{aux}}[3]$	$A_i[1]$	$A_i[2]$	$A_i[3]$
Aspect Label	**Writing**	**Plot**	**Scene**	**Writing**	**Plot**	**Scene**	**Acting**	**Feeling**	**Content**
Top-5 Words	dialogue texture complex weave nuanced	plot charactor write stereotype expedition	recommend story landscape eye looking	mislead plausible plot possibilities occurred	interesting culture readable species plot	exciting generic actions imagination title	truly playing fan wisdom chasing	hilarious entertaining fun charismatic enjoy	role classic character funny comedy

aspect correlations across domains are extremely sparse (usually focusing on one particular block). Due to space limitation, we pick the **most correlated** aspects-pair across domains (*i.e.,* corresponding to the maximum value in the matrix $\mathbf{S}$), and list the aspects extracted from user document, user auxiliary document of $\mathcal{D}_s$ and item document of $\mathcal{D}_t$ respectively. We also list the aspect information mentioned in the corresponding user document (namely D_u^t) in $\mathcal{D}_t$. The informative phrases are highlighted in orange color, including the stop words inside the context. As a reference, we underline the corresponding parts in the target review $d_{u,i}$ using red lines.

Example 1: The first example shows the aspect transfer process from *Book* domain to *Movie* domain. For better explanation, we list the top-5 words of each aspect in Table 6 and summarize the "Aspect Labels" based on our examination. We observe that the second and the third aspect of *Book* (*i.e., plot, scene*) are the most closely related to the third aspect of *Movie* (*i.e., content*). This is consistent with what is shown in Figure 5.

In detail, from $A_u[2]$ and $A_{u_{aux}}[2]$ we can infer that the user likes interesting plots or stories. According to $A_i[3]$, this item is

described as a funny movie with well-played roles. Hence the preferred aspect of the user can be transferred reasonably, leading to a high prediction accuracy. Note that, the information mentioned in D_u^t for $A_u[3]$ and the indication of user likes comedy in the target review both suggest the correctness of aspect transfer.

Example 2: The second example shows the aspect transfer from *Movie* domain to *Music* domain. According to Figure 5, the most related aspect-pair are the first aspect of $\mathcal{D}_s$ and the fifth aspect of $\mathcal{D}_t$. Correspondingly, from $A_u[1]$ and $A_{u_{aux}}[1]$ we can infer that this user prefers peaceful and tender plots in movies. According to $A_i[5]$, this item is indeed described as a soothing lute music with some negative comments. In view of the preference of the user and the reputation of the item, our model gives a moderate-score prediction, which turns out to be accurate in reference to the target review.

Example 3: The third example shows the aspect transfer from *Book* domain to *Music* domain. Similarly, according to Figure 5, the most related aspect pair is the seventh aspect of *Book* and the sixth aspect of *Music*. From $A_u[7]$ and $A_{u_{aux}}[7]$ we can infer that this user is fond of beautiful words and stories. According to $A_i[6]$, this item is a pop song praised for its melody but criticized for its lyrics.

Considering the user's preference for stories, our CATN still gives a high-score prediction in term of the item's characteristics and aspect transfer tendency.

Overall, the three sets of examples show that CATN is effective in cross-domain recommendation for cold-start user, with reasonable aspect transfer to support semantic explanation.

6 CONCLUSION

In this paper, we study the problem of review-based cross-domain recommendation for cold-start users. Our key focus is on the transfer of user preference derived from source domain to target domain, for effective and explainable recommendation. Instead of following the existing framework to first learn user/item representations in the source and target domains, then learn the mapping, we propose an end-to-end learning strategy. More importantly, we consider the fact that users' preferences are multi-faceted and only a subset of aspects in the two domains would match. In our framework, we therefore derive aspects from review documents and aim to find their correlations through a global aspect representations with attention. We show that our CATN model outperforms all existing models for cross-domain recommendation tasks. We believe that CATN offers an alternative view of the this interesting and critical task. Our study would also trigger studies on more effective ways of modeling user preference transfer across different domains. Inspired by [13], we may investigate for more possibilities on graph-based CDR in the future.

ACKNOWLEDGMENTS

REFERENCES

[1] Konstantin Bauman, Bing Liu, and Alexander Tuzhilin. 2017. Aspect Based Recommendations: Recommending Items with the Most Valuable Aspects Based on User Reviews. In *KDD*. 717–725.
[2] Xu Chen, Zheng Qin, Yongfeng Zhang, and Tao Xu. 2016. Learning to Rank Features for Recommendation over Multiple Categories. In *SIGIR*. 305–314.
[3] Zhiyong Cheng, Ying Ding, Lei Zhu, and Mohan S. Kankanhalli. 2018. Aspect-Aware Latent Factor Model: Rating Prediction with Ratings and Reviews. In *WWW*. 639–648.
[4] Jin Yao Chin, Kaiqi Zhao, Shafiq R. Joty, and Gao Cong. 2018. ANR: Aspect-based Neural Recommender. In *CIKM*. 147–156.
[5] Qiming Diao, Minghui Qiu, Chao-Yuan Wu, Alexander J. Smola, Jing Jiang, and Chong Wang. 2014. Jointly modeling aspects, ratings and sentiments for movie recommendation (JMARS). In *KDD*. 193–202.
[6] Xin Dong, Lei Yu, Zhonghuo Wu, Yuxia Sun, Lingfeng Yuan, and Fangxi Zhang. 2017. A Hybrid Collaborative Filtering Model with Deep Structure for Recommender Systems. In *AAAI*. 1309–1315.
[7] Ali Mamdouh Elkahky, Yang Song, and Xiaodong He. 2015. A Multi-View Deep Learning Approach for Cross Domain User Modeling in Recommendation Systems. In *WWW*. 278–288.
[8] Aleksandr Farseev, Ivan Samborskii, Andrey Filchenkov, and Tat-Seng Chua. 2017. Cross-Domain Recommendation via Clustering on Multi-Layer Graphs. In *SIGIR*. 195–204.
[9] 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 *AAAI*. 94–101.
[10] 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 *WWW*. 491–502.
[11] Ruining He and Julian McAuley. 2016. Ups and Downs: Modeling the Visual Evolution of Fashion Trends with One-Class Collaborative Filtering. In *WWW*. 507–517.
[12] Xiangnan He, Tao Chen, Min-Yen Kan, and Xiao Chen. 2015. TriRank: Review-aware Explainable Recommendation by Modeling Aspects. In *CIKM*. 1661–1670.
[13] Xiangnan He, Kuan Deng, Xiang Wang, Yan Li, Yongdong Zhang, and Meng Wang. 2020. Simplifying and Powering Graph Convolution Network for Recommendation. In *SIGIR*.
[14] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural Collaborative Filtering. In *WWW*. 173–182.
[15] Guangneng Hu, Yu Zhang, and Qiang Yang. 2018. CoNet: Collaborative Cross Networks for Cross-Domain Recommendation. In *CIKM*. 667–676.
[16] Liang Hu, Jian Cao, Guandong Xu, Longbing Cao, Zhiping Gu, and Can Zhu. 2013. Personalized recommendation via cross-domain triadic factorization. In *WWW*. 595–606.
[17] SeongKu Kang, Junyoung Hwang, Dongha Lee, and Hwanjo Yu. 2019. Semi-Supervised Learning for Cross-Domain Recommendation to Cold-Start Users. In *CIKM*. 1563–1572.
[18] Bin Li, Xingquan Zhu, Ruijiang Li, Chengqi Zhang, Xiangyang Xue, and Xindong Wu. 2011. Cross-Domain Collaborative Filtering over Time. In *IJCAI*. 2293–2298.
[19] Chenliang Li, Cong Quan, Li Peng, Yunwei Qi, Yuming Deng, and Libing Wu. 2019. A Capsule Network for Recommendation and Explaining What You Like and Dislike. In *SIGIR*. 275–284.
[20] Donghua Liu, Jing Li, Bo Du, Jun Chang, and Rong Gao. 2019. DAML: Dual Attention Mutual Learning between Ratings and Reviews for Item Recommendation. In *KDD*. 344–352.
[21] Muyang Ma, Pengjie Ren, Yujie Lin, Zhumin Chen, Jun Ma, and Maarten de Rijke. 2019. π-Net: A Parallel Information-sharing Network for Shared-account Cross-domain Sequential Recommendations. In *SIGIR*. 685–694.
[22] Tong Man, Huawei Shen, Xiaolong Jin, and Xueqi Cheng. 2017. Cross-Domain Recommendation: An Embedding and Mapping Approach. In *IJCAI*. 2464–2470.
[23] Tomas Mikolov, Ilya Sutskever, Kai Chen, Gregory S. Corrado, and Jeffrey Dean. 2013. Distributed Representations of Words and Phrases and their Compositionality. In *NIPS*. 3111–3119.
[24] Sungyong Seo, Jing Huang, Hao Yang, and Yan Liu. 2017. Interpretable Convolutional Neural Networks with Dual Local and Global Attention for Review Rating Prediction. In *RecSys*. 297–305.
[25] Ajit Paul Singh and Geoffrey J. Gordon. 2008. Relational learning via collective matrix factorization. In *KDD*. 650–658.
[26] Tianhang Song, Zhaohui Peng, Senzhang Wang, Wenjing Fu, Xiaoguang Hong, and Philip S. Yu. 2017. Review-Based Cross-Domain Recommendation Through Joint Tensor Factorization. In *DASFAA (Lecture Notes in Computer Science)*, Vol. 10177. 525–540.
[27] Yi Tay, Anh Tuan Luu, and Siu Cheung Hui. 2018. Multi-Pointer Co-Attention Networks for Recommendation. In *KDD*. 2309–2318.
[28] Nan Wang, Hongning Wang, Yiling Jia, and Yue Yin. 2018. Explainable Recommendation via Multi-Task Learning in Opinionated Text Data. In *SIGIR*. 165–174.
[29] Xinghua Wang, Zhaohui Peng, Senzhang Wang, Philip S. Yu, Wenjing Fu, and Xiaoguang Hong. 2018. Cross-Domain Recommendation for Cold-Start Users via Neighborhood Based Feature Mapping. In *DASFAA*. 158–165.
[30] Yaqing Wang, Chunyan Feng, Caili Guo, Yunfei Chu, and Jenq-Neng Hwang. 2019. Solving the Sparsity Problem in Recommendations via Cross-Domain Item Embedding Based on Co-Clustering. In *WSDM*. 717–725.
[31] Libing Wu, Cong Quan, Chenliang Li, and Donghong Ji. 2018. PARL: Let Strangers Speak Out What You Like. In *CIKM*. 677–686.
[32] Libing Wu, Cong Quan, Chenliang Li, Qian Wang, Bolong Zheng, and Xiangyang Luo. 2019. A Context-Aware User-Item Representation Learning for Item Recommendation. *ACM Trans. Inf. Syst.* 37, 2 (2019), 22:1–22:29.
[33] Yao Wu and Martin Ester. 2015. FLAME: A Probabilistic Model Combining Aspect Based Opinion Mining and Collaborative Filtering. In *WSDM*. 199–208.
[34] Haifeng Xia, Zengmao Wang, Bo Du, Lefei Zhang, Shuai Chen, and Gang Chun. 2019. Leveraging Ratings and Reviews with Gating Mechanism for Recommendation. In *CIKM*. 1573–1582.
[35] Feng Yuan, Lina Yao, and Boualem Benatallah. 2019. DARec: Deep Domain Adaptation for Cross-Domain Recommendation via Transferring Rating Patterns. In *IJCAI*. 4227–4233.
[36] Yang Zhang, Fuli Feng, Chenxu Wang, Xiangnan He, Meng Wang, Yan Li, and Yongdong Zhang. 2020. How to Retrain a Recommender System? A Sequential Meta-Learning Approach. In *SIGIR*.
[37] Yongfeng Zhang, Guokun Lai, Min Zhang, Yi Zhang, Yiqun Liu, and Shaoping Ma. 2014. Explicit factor models for explainable recommendation based on phrase-level sentiment analysis. In *SIGIR*. 83–92.
[38] Cheng Zhao, Chenliang Li, and Cong Fu. 2019. Cross-Domain Recommendation via Preference Propagation GraphNet. In *CIKM*. 2165–2168.
[39] Lei Zheng, Vahid Noroozi, and Philip S. Yu. 2017. Joint Deep Modeling of Users and Items Using Reviews for Recommendation. In *WSDM*. 425–434.
[40] Feng Zhu, Yan Wang, Chaochao Chen, Guanfeng Liu, Mehmet A. Orgun, and Jia Wu. 2018. A Deep Framework for Cross-Domain and Cross-System Recommendations. In *IJCAI*. 3711–3717.