

Received March 11, 2021, accepted March 30, 2021, date of publication April 14, 2021, date of current version April 27, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3073196

A CCA-Based Item-Side Alignment Method for Cross-Domain Recommendation System

LANTING WANG¹ AND YU XIN

Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China

Corresponding author: Yu Xin (xinyu@nbu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61602133, in part by the Natural Science Foundation of Ningbo under Grant 2019A610093, in part by The 3315 Plan Foundation of Ningbo under Grant 2019B-18-G, and in part by The Major Project of Ningbo Technological Innovation 2025 under Grant 2019B10125.

ABSTRACT For cross-domain recommendation, it can be divided into strong correlation and weak correlation problems according to the consistency between auxiliary domain and target domain. The weak correlation problem is more practical than the strong correlation problem, and the solution is more difficult. The difficulty lies in how to establish an effective transfer model, to make sure the auxiliary domain and the target domain can perform effective collaborative training. For weak correlation problem, if the item-side of auxiliary domain and the target domain are not aligned, or the transfer model has a strong dependency on the user-side of the auxiliary domain, it will seriously affect the effect of cross-domain recommendation. To solve these problems mentioned above, we propose a CCA-based item-side alignment method (CIAM) by introducing: (1) item side alignment method. We use CCA to align the item side between auxiliary and target domain, to intensify the weak correlation between 2 domains. (2) the transfer model of retaining the user feature of target domain. The CIAM retained user features of target domain in UV decomposition, that makes the transfer model could not destroy the user feature between 2 domains. The proposed CIAM can improve the assistance of auxiliary domain, and can avoid the influence of the needless user-side of the auxiliary domain on cross-domain recommendation. By experimental analysis, it can be verified that the proposed CIAM algorithm has a better performance than general cross-domain recommendation methods.

INDEX TERMS CCA, cross-domain recommendation, item-side alignment, transfer learning.

I. INTRODUCTION

With the development of social information industry, the volume of movies, music, online shopping and other industries continues to grow. In contrast, it is extremely difficult for individuals to experience the vast products and services one after another. In order to achieve accurate sales of products and improve the personal service quality, it is necessary to have a good algorithm to give a personally recommendation. The accuracy of recommendation method not only determines the individual's life experience, but also determines the commercial benefits. Therefore, research on recommendation systems in recent years has gradually become a hot topic in academic research.

The recommendation question contains 3 entities: user, item, and rating. The 3-tuple represents the rating of an item by a user in a particular domain (such as movie). The

objective is to predict the rating of the unknown user-item based on the observed user, item, and rating. According to the research, recommendation methods can be divided into content-based recommendation [1], collaborative filtering recommendation [2] and factorization based recommendation [3]. The content-based recommendation methods take context content of items as the research object, and detect similar content items by content semantic analysis. The items with similar contents can be regarded as the recommended items. The collaborative filtering recommendation methods, such as UserCF and ItemCF regard the relationship between users or items as object to determine the similarity between users or items. The ratings of similar user-user or item-item can be used to predict the unknown rating for a typical user or item. The factorization-based recommendation methods regard the latent factors between user-item as object. The work of these methods is to find an efficient factorization method on sparse rating matrix, and use the obtained latent factors to rebuild a full rating matrix.

The associate editor coordinating the review of this manuscript and approving it for publication was Sathish Kumar¹.

The challenge for the recommendation system is the sparseness of the observed ratings in a particular domain (such as movie), where the number of observed 3-tuple (user, item, and rating) is small, which will lead to the bias in rating prediction for a particular user-item. Therefore, if synthesize the rating data in multiple domains (such as music, comics, games), and perform analysis in a certain single domain (such as movies), it can improve the recommendation accuracy in the single domain. By this principle, many cross-domain recommendation methods are born, and gradually become the research focus of recommendation system [4]. The main problem of cross-domain recommendation research is how to accurately detect the relationship or correlation between the auxiliary domain and the target domain, and how to effectively use the relationship or correlation. So more and more researches utilize transfer learning models to solve the cross-domain recommendation problem. The main process is to find the correlation factors between the auxiliary domain and the target domain, and cooperative train these correlation factors cross the domains by data mining methods. The trained correlation factors can be transferred into target domain, and predict the void ratings in target domain. There are some difficulties in establishing a multi-domain transfer learning model:

1) Meanings. For cross-domain recommendation, the key issue is to detect common attributes between domains. If the established transfer model can accurately describe attributes of the auxiliary or target domain, the transfer model can be effectively performed on cross-domain training. Therefore, the meaning of transfer model is the first principle of transfer model selection.

2) Commonality. The objective of transfer learning model is to express the commonality, correlation and association among multiple domains. It is necessary to give a factor combination method, to make sure these combined factors can express the same attributes between auxiliary domain and target domain. The more similar the combination factors between domains, the better the effect of cross-domain recommendation.

3) Completeness. Since auxiliary and target domain are different domain, it is necessary to ensure the transfer model can fully express the features of auxiliary and target domains. So the selected transfer model needs to outline the auxiliary and target domains. Therefore, how to concisely express the profile of domains feature is a key issue on transfer model selection.

In this paper, we proposed a CCA-based item-side alignment method (CIAM), which provides a transfer learning and factor alignment method. The contributions of CIAM are the following:

1) Item-side alignment. We use CCA to align the item side between auxiliary and target domain, to intensify the weak correlation between 2 domains. The adopted CCA can ensure that the aligned item has the greatest commonality on transfer learning, to improve the rationality of cross-domain recommendations.

2) Item-side cross-domain transfer. The proposed CIAM only transfer the aligned item features from auxiliary domain to target domain. So the user feature of auxiliary domain cannot interfere that of target domain, to improve the collaborative recommendation effect.

3) Retain user features of target domain. The CIAM retained user features of target domain in UV decomposition, that makes the transfer model could not destroy the user feature between 2 domains. That makes the recommendation result closer to the truth.

The rest of the paper is organized as follows: in Section 2 the work related to cross-domain recommendation system is discussed. The preliminary work on CCA-based item-side alignment method is presented in Section 3. Section 4 discusses the proposed CIAM method. The experiments to verify the effectiveness of the method are presented in Section 5 and the conclusion is made in Section 6.

II. RELATED WORK

The principle of cross-domain recommendation system (CDRS) is to collaboratively recommend items to users, by model transfer or correlation analysis crossing multiple domains, according to the correlation among multiple domains. Therefore, the CDRS methods need to make an improvement on single domain recommendation methods, to establish a collaborative training mechanism among multiple domains. At present, the mainstream CDRS solving model can be divided into latent factor model and side information model.

The latent factor models (LFM) are based on the factorization method to make cross domain recommendation. These methods establish the association between auxiliary domain and target domain by extracting and transferring latent factor obtained by matrix decomposition. The LFMs express the user-item association in form of user-factor-item. The matrix decomposition methods for LFMs can be divided into linear decomposition and probabilistic decomposition. The objective of the linear decomposition methods, such as PMF [5], SVDFeature [6], Funk-SVD [3], ListRankMF [7], is to establish a linear mapping for user-factor and item-factor. That can be interpreted as mining user-factor and item-factor transition matrices from user-item association. The objective of the probabilistic decomposition methods, such as FMM [8], is to establish relationship between factor-user and item-factor by probability graph model.

In the linear decomposition field, Li et al proposed CBT [9]. The CBT uses Codebook matrix to record the clustering relationship between user and item in auxiliary domain, and transfer the Codebook matrix to target domain, to train user and item features and predict void ratings in target domain. Pan *et al.* proposed TCF [10] and CST [11] based on the CBT model. The TCF regularized the transfer of Codebook matrix, and use the Codebook as constraint to control the target domain decomposition. The CST uses the user-factor and item-factor in auxiliary domain as the

constraint of target domain decomposition, to make sure the user features in target domain cannot be destroyed by decomposition.

In the probabilistic decomposition field, probability graph model is used to associate the relationship among user, factor and item. It assumes the factors between auxiliary and target domain are the same. If the same factors are obtained by collaborative training between the 2 domains, the ratings in target domain can be predicted according to the obtained factors. The FMM is a basic probabilistic graph model for 2 sides clustering on users and items. Li *et al.* proposed RMGM [13] based on FMM model. The RMGM performs 2 sides clustering on users and items, and regards the 2 side clusters as the latent factors between users and items. The clusters indicate the relationship of users and items in auxiliary domain, so they can also be observed as Codebook for cross domain transfer. Zhang *et al.* proposed the CIT [14] based on FMM. The CIT obtains latent factors and coordinates of users and items by clustering the users and items in auxiliary domain, then utilize GFK [15] to predict void ratings in target domain. Furthermore, CPT-v and Logit-vd [16] adopted LDA [17] to extract latent factors from rating, by treating the ratings as generative samples of users and items in 2 domains. It establishes a probabilistic association between user-item and rating, to predict the void rating in target domain by sampling method.

A. LATENT FACTOR MODELS

The latent factor models (LFM) are based on the factorization method to make cross domain recommendation. These methods establish the association between auxiliary domain and target domain by extracting and transferring latent factor obtained by matrix decomposition. The LFM express the user-item association in form of user-factor-item. The matrix decomposition methods for LFM can be divided into linear decomposition and probabilistic decomposition. The objective of the linear decomposition methods, such as PMF [3], SVDFeature [6], Funk-SVD [3], ListRank-MF [7], is to establish a linear mapping for user-factor and item-factor. That can be interpreted as mining user-factor and item-factor transition matrices from user-item association. The objective of the probabilistic decomposition methods, such as FMM [8], is to establish relationship between factor-user and item-factor by probability graph model.

In the linear decomposition field, Li *et al.* proposed CBT [9]. The CBT uses Codebook matrix to record the clustering relationship between user and item in auxiliary domain, and transfer the Codebook matrix to target domain, to train user and item features and predict void ratings in target domain. Pan *et al.* proposed TCF [10] and CST [11] based on the CBT model. The TCF regularized the transfer of Codebook matrix, and use the Codebook as a constraint to control the target domain decomposition. The CST uses the user-factor and item-factor in auxiliary domain as the constraint of the target domain decomposition, to make the decomposition retain the feature of target domain.

Tseuif [12] employs Funk-SVD decomposition to extract extra user and item features from user- and item-side auxiliary domains, respectively, with the purpose of expanding the two-dimensional location feature vector.

In the probabilistic decomposition field, probability graph model is used to associate the relationship among user, factor and item. It assumes the factors between auxiliary and target domain are the same. If the same factor is obtained by collaborative training between the 2 domains, the ratings in target domain can be predicted according to the obtained factors. The FMM is a basic probabilistic graph model for 2 sides clustering on users and items. Li *et al.* proposed RMGM [13] based on the FMM model. The RMGM performs 2 sides clustering on users and items, and regards the 2 side clusters as the latent factors between users and items. The clusters indicate the relationship of users and items in auxiliary domain, so it can also be observed as Codebook for cross domain transfer. Zhang *et al.* proposed the CIT [14] based on FMM. The CIT obtains latent factors and coordinates of users and items by clustering the users and items in auxiliary domain, then utilize GFK [15] to predict void ratings in target domain. Furthermore, CPT-v and Logit-vd [16] adopted LDA [17] to extract latent factors from rating, by treating the ratings as generative samples of users and items in 2 domain. It establishes a probabilistic association between user-item and rating, to predict the void rating in target domain by sampling method.

B. SIDE INFORMATION MODELS

The side information models (SIM) are based on the same or similar side information to make cross domain recommendation. If the users of auxiliary domain and target domain are both from the same user set or the users have common features, that can be denoted as user-side problem. Similarly, if the items of auxiliary domain and target domain are both from the same user set or the items have common features, it can be denoted as user-side problem. According to different application scenarios, the SIM based cross domain recommendation problems can be divided into user-side problem [18]–[20], item-side problem [21], two sides problem [10], [11], [22], [23], and side information overlapping problem [24], [25].

In use or item-side information model field, SIM is regarded as a crossing association between auxiliary domain and target domain. The SIM can integrate the separated auxiliary domain with the target domain, to predict the unknown ratings in target domain by SVM [27] and other methods. For two sides problems, Loni *et al.* [23] adopts factorization machine [26] to quantify the user-item pairs in auxiliary domains. The quantified user-item pairs by factorization machine in auxiliary domain are used as a training sample, to predict the rating by the user-item pairs in the target domain. CTSIF-SVMs [22] adopts ternary factorization method to establish latent factor space, and extracts coordinates for users and items to give a two sides metric policy.

For the side information overlapping problem, Zhang *et al.* proposed KerKT [24] based on the CIT model. KerKT uses

GFK [15] to align the extracted user and item features, and utilizes the Laplace method to measure the similarity for non-overlapping users and items. That can convert edge overlap to general recommendation problem. Users and items in different domains may have the same label, resulting in overlapping information between domains. For this problem, Shi *et al.* proposed CTagCDCF [25]. The CTagCDCF established user tag indicator matrix using semantic tags, to measure the similarity of users and items between 2 domains. Hao *et al.* proposed the CTagCDR [25] based on CTagCDCF. The CTagCDR utilizes TF-IDF to quantify the shared tags between 2 domains, and utilizes Jaccard coefficient to measure the similarity of shared tags and specific tags in 2 domains.

III. PRELIMINARY WORK

A. UV FACTORIZATION BASED RECOMMENDATION METHODS

Rating matrix $R_{n \times c}$ can be decomposed into the product of two low dimensional matrices U and V , where $U \in \mathbb{R}^{n \times f}$, $V \in \mathbb{R}^{c \times f}$. A rating r_{ui} of user u on item i can be approximated by the product of U and V .

$$\hat{r}_{ui} = \sum_{k=1}^f U_{u,k} V_{i,k} \quad (1)$$

So, rating matrix R can be approximated as following.

$$R \approx \hat{R} = UV' \quad (2)$$

Funk-SVD uses RMSE (root mean square error) to minimize the target and optimize the values of U and V .

$$\min F = \|R - \hat{R}\|_2^2 + \lambda (\|U\|^2 + \|V\|^2) \quad (3)$$

$$\min F = \sum_{u,i} \left\{ \left(r_{ui} - \sum_{k=1}^f U_{u,k} V_{i,k} \right)^2 + \lambda (\|U\|^2 + \|V\|^2) \right\} \quad (4)$$

where λ is regularization parameter.

$$\frac{\partial F}{\partial U_{u,k}} = -2V_{i,k}e_{u,k} + 2\lambda U_{u,k} \quad (5)$$

$$\frac{\partial F}{\partial V_{i,k}} = -2U_{u,k}e_{u,k} + 2\lambda V_{i,k} \quad (6)$$

The following recursive formula can be obtained by SGD (stochastic gradient descent) method.

$$U_{u,k} = U_{u,k} + \alpha (V_{i,k}e_{u,k} - \lambda U_{u,k}) \quad (7)$$

$$V_{i,k} = V_{i,k} + \alpha (U_{u,k}e_{u,i} - \lambda V_{i,k}) \quad (8)$$

where $e_{ui} = r_{ui} - \sum_{k=1}^f U_{u,k} V_{i,k}$, U and V contains f columns.

This process is UV decomposition, and the f columns in U and V represent f latent factors between users and items.

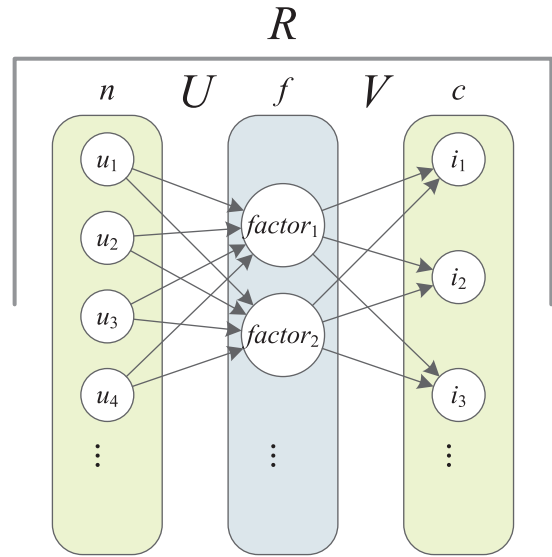


FIGURE 1. UV decomposition with f latent factors.

Fig.1 shows the UV decomposition process of a rating matrix R with f latent factors. The f latent factors determine the decomposition results of matrices U and V , which can further determine the prediction effect of the rating matrix R . For cross-domain recommendation, $R^{(1)} \in \mathbb{R}^{n \times c}$ represents target domain, and $R^{(2)} \in \mathbb{R}^{n \times d}$ represents auxiliary domain. If the item-side of $R^{(1)}$ and $R^{(2)}$ are aligned, the UV decomposition of auxiliary domain $R^{(2)}$ can adjust the UV decomposition of target domain $R^{(1)}$. By this way, the UV decomposition of target domain $R^{(1)}$ can be optimized, and make the recommendation take into account both features of auxiliary and target domain.

B. COLLABORATIVE FILTERING BASED RECOMMENDATION METHODS

In the recommendation system, a user only gives ratings to a limited number of items, so the rating matrix R is a sparse matrix. Traditional UserCF-based and ItemCF-based recommendation methods perform collaborative filtering on sparse matrix R , based on user and item similarity. The collaborative filtering can convert sparse matrix R into a full matrix \tilde{R} , and the full matrix \tilde{R} can be used to predict the rating of user to item.

(1) UserCF-based methods. These methods need to use user similarity as the basis for recommendation. The similarity between user u and user v can be expressed as cosine similarity:

$$W_{u,v} = \frac{|N(u) \cap N(v)|}{\sqrt{|N(u)| |N(v)|}} \quad (9)$$

where $N(u)$ represents the collection of items that user u has rated, $|N(u)|$ indicates the number of items in $N(u)$, and $|N(u) \cap N(v)|$ represents the number of items in the intersection of $N(u)$ and $N(v)$.

$$\tilde{R}_{u,i} = \sum_{v \in S(u,k) \cap N(i)} w_{u,v} R_{v,i} \quad (10)$$

where $S(u, K)$ represents K users with the highest similarity to user u , $N(i)$ indicates the user who has rated item i .

(2) ItemCF-based methods. These methods need to use item similarity as the basis for recommendation. The similarity between item i and item j can be expressed as cosine similarity:

$$W_{i,j} = \frac{|N(i) \cap N(j)|}{\sqrt{|N(i)||N(j)|}} \quad (11)$$

where $N(i)$ represents the collection of users who have rated item i .

$$\tilde{R}_{u,i} = \sum_{j \in S(i,K) \cap N(u)} w_{i,j} R_{u,i} \quad (12)$$

where $S(j, K)$ represents K items with the highest similarity to item j , $N(u)$ indicates the items which have been rated by user u .

IV. CCA-BASED ITEM-SIDE ALIGNMENT METHOD (CIAM)

A. ITEM-SIDE ALIGNMENT

For cross-domain recommendation problem, the role of the auxiliary domain depends on its correlation with target domain. If the auxiliary domain has a strong correlation with the item-side of the target domain, it can effectively improve the recommendation accuracy. For example, the item tags of movies, novels and music are as follows:

Movie: (Love, Biography, History, Science, Fantasy, Suspense)

Novel: (Love, Biography, History, Science, Fantasy, Suspense)

Music: (Electronic, Rock, R&B, Jazz, Classical, Metal)

It can be seen that movies and novels have the same tags (love, biography, history, science fiction, fantasy, suspense). If novels are used as auxiliary domain on movie recommendation, it will significantly improve the recommendation accuracy. There are no clear correlative tags on movie and music domain. So, directly using music domain to recommend movies would have not effective performance. However, the tags in movie and music domain may have latent correlation. For instance, the tag 'Love' in movie domain has a tight correlation with the tags 'R&B' and 'Classical' in music domain. Furthermore the tag 'Love' in movie domain has a tight correlation with the tags 'Electronic' and 'Metal' in music domain. Therefore, a cross-domain problem with strong item-side correlation can be defined as strong correlation, such as movie-novel. Correspondingly, a cross-domain problem with weak item-side correlation can be defined as weak correlation, such as movie-music and novel-music. It is easy to deal with strong correlation recommendation problem, such as CTagCDCF and CTagCDR. For the weak correlation problem, the key is how to align the auxiliary domain with the items in the target domain. That can also be defined as finding the correspondence of tags between two domains. For example, find a map or function $fun1$, $fun2$, to make sure 'Love'= $fun1$ ('R&B', 'Classical') and 'Fantasy'= $fun2$ ('Electronic', 'Metal').

For the cross-domain recommendation problem with weak correlation, the user⁽¹⁾, item⁽¹⁾ in target domain $D^{(1)}$ and the user⁽²⁾, item⁽²⁾ in auxiliary domain $D^{(2)}$ have weak correlation. There is no intersection between item⁽¹⁾ and item⁽²⁾, but there exists a certain correlation between them. So item⁽¹⁾ and item⁽²⁾ can be mapped into the same alignment space by alignment factors. In alignment space, the item⁽¹⁾ and item⁽²⁾ will have a strong correlation. So, how to find out the alignment factors of item-side between auxiliary and target domain is the key to solve cross-domain recommendation problem.

In this paper, the rating matrix $R^{(1)}$ of $D^{(1)}$ and the rating matrix $R^{(2)}$ of $D^{(2)}$ are mapped into t -dimensional alignment space by means of CCA. It can be denoted that $R^{(1)} \in \mathbb{R}^{n \times c}$ and $R^{(2)} \in \mathbb{R}^{n \times d}$ are converted into $H^{(1)} \in \mathbb{R}^{n \times t}$ and $H^{(2)} \in \mathbb{R}^{n \times t}$ by t alignment factors. The spatial mapping of CCA requires that all elements in matrix $R^{(1)}$ and $R^{(2)}$ are full matrix.

However, for recommendation issue, the rating matrix $R^{(1)}$ and $R^{(2)}$ are sparse matrix. So we can use the UserCF-based or ItemCF-based algorithm to convert the $R^{(1)}$ and $R^{(2)}$ into a full matrix. That can be denoted as $R^{(1)} = \text{UserCF}(R^{(1)})$, $R^{(2)} = \text{UserCF}(R^{(2)})$ or $R^{(1)} = \text{ItemCF}(R^{(1)})$, $R^{(2)} = \text{ItemCF}(R^{(2)})$. The obtained $\tilde{R}^{(1)}$ and $\tilde{R}^{(2)}$ can be denoted as the following.

$$\begin{aligned} \tilde{R}^{(1)} &= [r_1^{(1)}, r_2^{(1)}, \dots, r_c^{(1)}], \\ \tilde{R}^{(1)} &\in \mathbb{R}^{n \times c}, r_i \in \mathbb{R}^{n \times 1} \\ \tilde{R}^{(2)} &= [r_1^{(2)}, r_2^{(2)}, \dots, r_d^{(2)}], \\ \tilde{R}^{(2)} &\in \mathbb{R}^{n \times d}, r_i \in \mathbb{R}^{n \times 1} \end{aligned} \quad (13)$$

$H^{(1)}$ and $H^{(2)}$ are the aligned $\tilde{R}^{(1)}$ and $\tilde{R}^{(2)}$. The alignment process is a linear transformation from $\tilde{R}^{(1)}$ and $\tilde{R}^{(2)}$ to $H^{(1)}$ and $H^{(2)}$.

$$\begin{aligned} H^{(1)} &= [h_1^{(1)}, h_2^{(1)}, \dots, h_t^{(1)}], \\ h_i^{(1)} &= a_{1,i}r_1^{(1)} + a_{2,i}r_2^{(1)} + \dots + a_{c,i}r_c^{(1)} = \tilde{R}^{(1)}A \\ H^{(2)} &= [h_1^{(2)}, h_2^{(2)}, \dots, h_t^{(2)}], \\ h_i^{(2)} &= a_{1,i}r_1^{(2)} + a_{2,i}r_2^{(2)} + \dots + a_{d,i}r_d^{(2)} = \tilde{R}^{(2)}A \end{aligned} \quad (14)$$

where

$$\begin{aligned} A &\in \mathbb{R}^{c \times t}, A = [a_1, a_2, \dots, a_t], a_i \in \mathbb{R}^{c \times 1} \\ B &\in \mathbb{R}^{d \times t}, A = [b_1, b_2, \dots, b_t], b_i \in \mathbb{R}^{d \times 1} \\ h_i^{(1)} &= \tilde{R}^{(1)}a_i, h_i^{(2)} = \tilde{R}^{(2)}b_i \end{aligned} \quad (15)$$

The correlation coefficient ρ_i between $h_i^{(1)}$ and $h_i^{(2)}$ is as the following:

$$\rho_i = \text{corr} \left(h_i^{(1)}, h_i^{(2)} \right) = \frac{a_i' \Sigma_{\tilde{R}^{(1)}\tilde{R}^{(2)}} b_i'}{\sqrt{a_i' \Sigma_{\tilde{R}^{(1)}\tilde{R}^{(2)}} a_i b_i' \Sigma_{\tilde{R}^{(1)}\tilde{R}^{(2)}} b_i}} \quad (16)$$

ρ_i represents the correlation between $h_i^{(1)}$ and $h_i^{(2)}$. Given t ρ s, there will be t corresponding $[a_1, a_2, \dots, a_t]$ and $[b_1, b_2, \dots, b_t]$. So $[a_1, a_2, \dots, a_t]$ and $[b_1, b_2, \dots, b_t]$ corresponding to t largest ρ s can realize alignment on item⁽¹⁾ and item⁽²⁾. According to the principles of CCA, ρ_i^2 is the common eigenvalues of Σ_A and Σ_B , a_i and b_i are eigenvectors

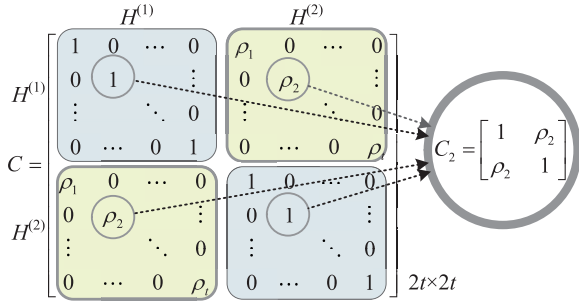


FIGURE 2. The correlation matrix between $H^{(1)}$.

corresponding to ρ_i^2 . The relation of Σ_A , Σ_B , ρ_i , a_i and b_i can be described as follows:

$$\begin{aligned}\Sigma_A &= \Sigma_{\tilde{R}^{(1)}}^{-1} \Sigma_{\tilde{R}^{(1)}\tilde{R}^{(2)}} \Sigma_{\tilde{R}^{(2)}}^{-1} \Sigma_{\tilde{R}^{(2)}\tilde{R}^{(1)}} \\ \Sigma_B &= \Sigma_{\tilde{R}^{(2)}}^{-1} \Sigma_{\tilde{R}^{(2)}\tilde{R}^{(1)}} \Sigma_{\tilde{R}^{(1)}}^{-1} \Sigma_{\tilde{R}^{(1)}\tilde{R}^{(2)}} \\ \Sigma_A a_i &= \rho_i^2 a_i \\ \Sigma_B b_i &= \rho_i^2 b_i\end{aligned}\quad (17)$$

To obtain the best $h_i^{(1)}$ and $h_i^{(2)}$, it needs to select $A = [a_1, a_2, \dots, a_t]$, $B = [b_1, b_2, \dots, b_t]$ corresponding to t largest ρ_s . $H^{(1)} \in \mathbb{R}^{n \times t}$ and $H^{(2)} \in \mathbb{R}^{m \times t}$ can be obtained by Eq(14). The t columns of $H^{(1)}$ and $H^{(2)}$ are linear combination of $\tilde{R}^{(1)}$ and $\tilde{R}^{(2)}$ by A and B . So the t columns corresponding to $H^{(1)}$ and $H^{(2)}$ have the largest correlation. That means that A and B can align the item-side of $\tilde{R}^{(1)}$ and $\tilde{R}^{(2)}$. Fig.2 shows a canonical correlation matrix C between $H^{(1)}$ and $H^{(2)}$. The canonical correlation matrix C consists of 4 sub-correlation matrices, where the correlation matrices of $H^{(1)}$ - $H^{(1)}$ and $H^{(2)}$ - $H^{(2)}$ are identity matrices, and the correlation matrices of $H^{(1)}$ - $H^{(2)}$ and $H^{(2)}$ - $H^{(1)}$ are diagonal matrices. For instance, in Fig.2 the correlation matrix C_2 between $h_i^{(1)}$ and $h_i^{(2)}$ consists of $[C_{2,2}, C_{2,t+2}, C_{t+2,2}, C_{t+2,t+2}]$, and the correlation coefficient is ρ_2 .

B. ITEM-SIDE TRANSFER

Performing item-side alignment on $\tilde{R}^{(1)}$ and $\tilde{R}^{(2)}$ can improve the accuracy of cross-domain recommendation. To consider the impact of auxiliary domain $H^{(2)}$ on target domain $H^{(1)}$, it needs to modify the decomposition process of $H^{(1)}$. That can make the UV decomposition of $H^{(1)}$ contain the influence of auxiliary domain $H^{(2)}$. The UV decomposition of $H^{(1)}$ and $H^{(2)}$ can be denoted as $H^{(1)} \approx U^{(1)}V^{(1)'}$, $H^{(2)} \approx U^{(2)}V^{(2)'}$, where $U^{(1)} \in \mathbb{R}^{n \times f}$ represents users' coordinates in f dimensional user-item latent space of target domain, and $V^{(2)} \in \mathbb{R}^{t \times f}$ represents items' coordinates in f dimensional user-item latent space of auxiliary domain. The t dimensional items in $H^{(2)}$ have been aligned, so $V^{(2)}$ can be directly transferred into target domain to constrain the item-side decomposition. Correspondingly, to preserve the user feature in the target domain, it needs to constraint the user-side decomposition by $U^{(1)}$. So the UV decomposition of $H^{(1)}$ can

be expressed as the following.

$$\min_{q^*, p^*} \left\{ \sum_{u,i} (H_{ui}^{(1)} - U_u V_i')^2 + \lambda \left(\|U_u - U_u^{(1)}\|^2 + \|V_i - V_i^{(2)}\|^2 \right) \right\} \quad (18)$$

$$\hat{H}^{(1)} = UV' \quad (19)$$

Partial derivatives of Eq(18) are

$$\begin{aligned}\frac{\partial F}{\partial V_i} &= -2[e_{u,i}U_u - \lambda(V_i - V_i^{(2)})] \\ \frac{\partial F}{\partial U_u} &= -2[e_{u,i}V_i - \lambda(U_u - U_u^{(1)})]\end{aligned}\quad (20)$$

where $e_{u,i} = \hat{h}_{u,i} - U_u V_i'$, the iterative process as follows can be obtained by SGD (stochastic gradient descent).

$$\begin{aligned}V_i &\leftarrow V_i + \gamma[e_{u,i}U_u - \lambda(V_i - V_i^{(2)})] \\ U_u &\leftarrow U_u + \gamma[e_{u,i}V_i - \lambda(U_u - U_u^{(1)})]\end{aligned}\quad (21)$$

where γ are step-size control factor. According to Eq(19), $\hat{H}^{(1)}$ can be obtained, which is the modified $H^{(1)}$. By transition matrix A , $H^{(1)}$ can be obtained from $\tilde{R}^{(1)}$ to $H^{(1)}$ as shown in Eq(15). So \hat{R} can also be obtained by least square inverse from $\hat{H}^{(1)}$ as Eq(22).

$$\hat{R} = \hat{H}^{(1)}A'(AA')^{-1} \quad (22)$$

C. OVERALL PROCESS

After CCA alignment and UV decomposition on target domain, the prediction matrix \hat{R} of target domain can be obtained according to Eq(22). The overall process of CIAM can be illustrated in Fig.3. In Fig.3, $\tilde{R}^{(1)}$ and $\tilde{R}^{(2)}$ are original target domain and original auxiliary domain. $H^{(1)}$ and $H^{(2)}$ are aligned $\tilde{R}^{(1)}$ and $\tilde{R}^{(2)}$ by CCA. After UV decomposition on $H^{(1)}$ and $H^{(2)}$, $U^{(1)} \in \mathbb{R}^{n \times f}$, $V^{(1)} \in \mathbb{R}^{t \times f}$, $U^{(2)} \in \mathbb{R}^{m \times f}$ and $V^{(2)} \in \mathbb{R}^{t \times f}$ can be obtained. As $H^{(1)}$ and $H^{(2)}$ have been aligned, $V^{(1)}$ and $V^{(2)}$ are homomorphic item feature matrices.

$V^{(2)}$ can be directly transferred into target domain to constrain the item-side decomposition. Correspondingly, to preserve the user feature in the target domain, it needs to constraint the user-side decomposition by $U^{(1)}$. To consider the item-side influence of auxiliary domain $H^{(2)}$ on target domain $H^{(1)}$, $V^{(2)}$ and $U^{(1)}$ are used to constrain the UV decomposition on target domain. Then \hat{R} can be obtained by least square inverse from $\hat{H}^{(1)}$ as Eq(22).

The key of CIAM is the similarity between $V^{(1)}$ and $V^{(2)}$, which is the correlation between the target domain and the auxiliary domain. If the similarity between $V^{(1)}$ and $V^{(2)}$ is weaker, the bias of \hat{R} to $\tilde{R}^{(1)}$ is larger. Conversely, if the similarity between $V^{(1)}$ and $V^{(2)}$ is stronger, the bias of \hat{R} to $\tilde{R}^{(1)}$ is less.

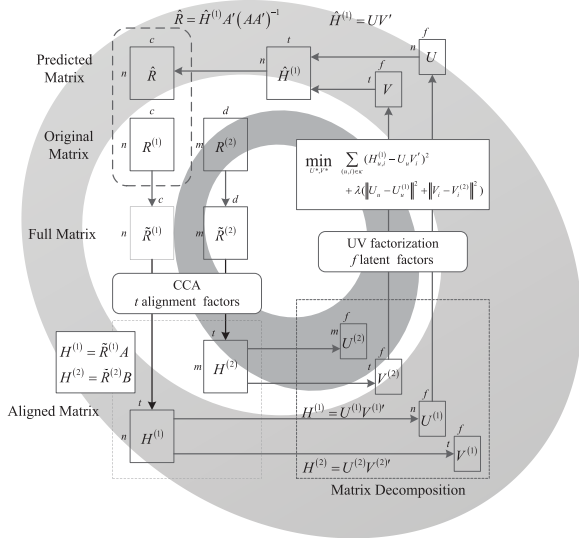


FIGURE 3. The overall process of CIAM.

V. EXPERIMENTS

A. DATASETS

We use MovieLens (user×item=71567×10681), Netflix (user×item=480,000×18,000) and EachMovie (user×item=72916×1628) as experimental datasets. In order to effectively compare the performance of CIAM with other methods, we perform the following preprocessing on the four datasets.

(1) MovieLens 10M. This dataset contains more than 10 million ratings of user-movies (user×item=71567×10681). The levels of ratings are in (0.5-5). We normalize the levels of ratings to (1-5), and randomly extract 30,000 users and 4,000 items, to obtain 2,230,000 ratings. This data is denoted by MovieLens 2M.

(2) Netflix. This dataset contains more than 10^8 ratings of user-movies (user×item=480,000×18,000). The levels of ratings are in (1-5). We randomly extract 30,000 users and 2,000 items, to obtain 3760,000 ratings. This data is denoted by Netflix 3M.

(3) EachMovie. This dataset contains more than 28,000 ratings user-movies (user×item=72916×1628). The levels of ratings are in (1-6). We randomly extract 70,000 users and 3,000 items, to obtain 23,000 ratings. This data is denoted by EachMovie 2M.

B. CCA BASED ITEM-SIDE ALIGNMENT

In order to compare the importance of CCA based alignment, this experiment designed the following comparison methods, denoted by PCA-based method.

(1) Obtain t principal component of $\tilde{R}^{(1)}$ and $\tilde{R}^{(2)}$ by PCA. $\lambda_1^{(1)} \geq \lambda_2^{(1)} \geq \dots \geq \lambda_t^{(1)}$ and $\lambda_1^{(2)} \geq \lambda_2^{(2)} \geq \dots \geq \lambda_t^{(2)}$ are t principal components of $\tilde{R}^{(1)}$ and $\tilde{R}^{(2)}$. $\tilde{A} = [\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_t]$ and $\tilde{B} = [\tilde{b}_1, \tilde{b}_2, \dots, \tilde{b}_t]$ are eigenvectors corresponding to $\lambda_1^{(1)} \geq \lambda_2^{(1)} \geq \dots \geq \lambda_t^{(1)}$ and $\lambda_1^{(2)} \geq \lambda_2^{(2)} \geq \dots \geq \lambda_t^{(2)}$.

(2) Transform $\tilde{R}^{(1)}$ and $\tilde{R}^{(2)}$ into $H^{(1)}$ and $H^{(2)}$ by \tilde{A} , \tilde{B} , where $H^{(1)} = \tilde{R}^{(1)}\tilde{A}$ and $H^{(2)} = \tilde{R}^{(2)}\tilde{B}$.

TABLE 1. 7 sampling polices on MovieLens 2M.

sampling policies	auxiliary domain	target domain	Jaccard
U1	50,000 users	60,000 users	0
U2	50,000 users	50,000 users	0.2
U3	60,000 users	50,000 users	0.4
I1	5,000 items	6,000 items	0
I2	6,000 items	5,000 items	0.1
I3	5,000 items	5,000 items	0.2
I4	6,000 items	6,000 items	0.4

(3) According to Eq(18)-(22), the prediction matrix of A can be obtained with $H^{(1)}$ and $H^{(2)}$ as inputs.

The difference between PCA-based and CCA-based (CIAM) method is that the transformation process of PCA-based method does not consider the correlation of t principal components between the 2 domains. CCA-based uses t correlation coefficients to indicator of the item-side correlation between the 2 domains. The item-side of the 2 domains are linearly combined by A and B to achieve the alignment.

We choose MovieLens 2M to carry out the experiment on PCA-based and CCA-based method. In order to compare the performance of PCA-based and CCA-based method, we designed 7 sampling polices U1, U2, U3, I1, I2, I3, I4 as shown in Table 1. For example, U1 means that randomly select 50,000 users in MovieLens 2M as users of auxiliary domain, and 60,000 users as users of target domain, with the constraint that the intersection of users in auxiliary and target domain is empty, namely Jaccard coefficient is 0. I2 means that randomly select 6,000 items in MovieLens 2M as items of auxiliary domain, and 5,000 items as items of target domain, and the Jaccard coefficient of items between auxiliary and target domain is 0.1.

In this experiment, we crossly adopt U1~U3 and I1~I4 to sample the experimental data, and obtain 12 data sets such as U1-I1, U1-I2, ..., U3-I4.

The user-side overlapping degree of auxiliary and target domain from U1 to U3 becomes larger, and the item-side overlapping degree of auxiliary and target domain from I1 to I4 also becomes larger. That means correlation of 2 domains is becomes larger from U1 to U3, or from I1 to I4. The purpose of this experiment is to analyze the alignment effect between PCA-based and CCA-based methods. So it is necessary to anonymize the item columns of the 12 data sets obtained by U1-I1, U1-I2, ..., U3-I4. Then carry out PCA-based and CCA-based methods on the anonymised 12 data sets, and record absolute errors (AE) of all the predicted ratings. Fig.4 shows the AE distribution of prediction ratings with PCA-based and CCA-based methods in 12 data sets. As can be seen from Fig.4, if ignored items sampling policies I1~I4, the AE distributions of the PCA-based and the CCA-based methods tend to be 0 from U1 to U3. That means when the user-side overlapping degree between 2 domains increases, the recommendation accuracy can be improved. If ignored users sampling policies U1~U3, the AE distribution of PCA-based methods has a more obvious trend toward 0 compared with CCA-based method, from I1 to I4.

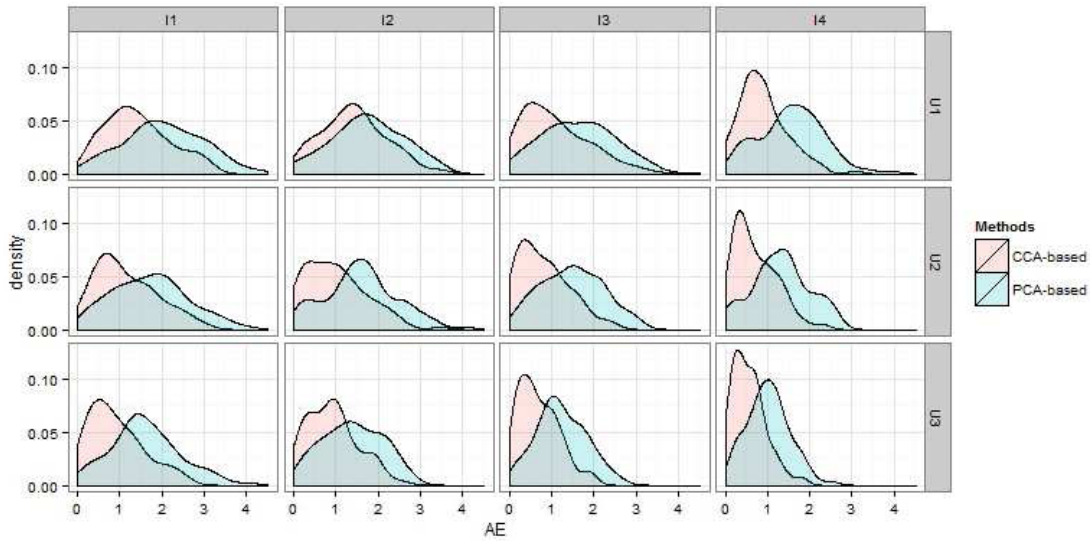


FIGURE 4. AE distribution obtained by PCA-based and CCA-based methods.

That means when the correlation between the 2 domains on item-side is stronger, the cross-domain recommendation methods perform better, and the advantage of CCA-based method is more obvious. Therefore, for the cross-domain recommendation problem, if the item edges are aligned, the accuracy of the cross-domain recommendation can be effectively improved.

C. METRIC MODELS

This experiment uses MAE (Mean Absolute Error) and RMSE (Root Mean Square Error) to evaluate the recommendation effect. The MAE is the average of absolute errors, and the RMSE is the expectation of the Square Error between the estimated ratings and the real ratings. The expression of MAE and RMSE are as follows:

$$MAE = \frac{1}{T} \sum_{i \in T} |r_i - \tilde{r}_i| \tag{23}$$

$$RMSE = \sqrt{\frac{\sum_{i \in T} (r_i - \tilde{r}_i)^2}{|T|}} \tag{24}$$

where T is the number of ratings in the test data, r_i is the real ratings in test data, \tilde{r}_i is the estimated ratings. If MAE and RMSE are small, the recommendation method is accurate. The estimated ratings by Funk-SVD, CBT, RMGM and CIAM are non-discrete. So we discretized the estimated ratings by each recommendation method. That makes MAE and RMSE test more realistic.

D. METHODS COMPARISON

In this experiment, Funk-SVD, CBT and RMGM are selected as the contrast methods.

(1) Funk-SVD. The objective of this method is to minimize the RMSE on rating matrix decomposition. After rating matrix decomposition, U and V can be obtained, where U and V represent the coordinates of users and items in latent

space. The recommendation performance of Funk-SVD is affected by the dimension f of latent space. The empirical value of dimension f is [30, 70]. In this experiment, we set $f=50$ for Funk-SVD.

(2) CBT. This method uses codeblock as the transfer model. The trained codeblock in auxiliary domain is transferred into target domain, to predict the rating of a specific user-item in target domain. For CBT, the number of clusters of users and items, denoted by K and L , needs to be set in advance. According to the empirical values of K and L in [9], and we set $K = 50$ $L = 50$.

(3) RMGM. This method is a cross-domain recommendation algorithm based on probability graph model. RMGM uses the FMM [8] to detect the correlation matrices of user-cluster and item-cluster on auxiliary domain. The obtained user-item cluster correlation matrix is transferred into the target domain, by which the rating values in target domain can be estimated. Similar to CBT, RMGM also needs to set the clusters of users and items, namely K and L , and we set $K = 50$ $L = 50$.

(4) CIAM. This is the proposed method in this paper. The parameters of CIAM contain t and f , where t is the number of alignment factors and f is the number of latent factors. t and f need to satisfy $f \leq t, t \leq c, t \leq d$, where c is the number of items in the target domain and d is the number of items in the auxiliary domain. For CIAM, we set $t = \text{floor}(\min\{c, d\})$. The choice of f needs to consider correlation coefficients $\rho_1, \rho_2, \dots, \rho_t$. So we utilize the following equation to set f .

$$f = \text{argmax} \left| \frac{\sum_{i=1}^f \rho_i}{\sum_{i=1}^t \rho_i} - 80\% \right| \tag{25}$$

E. EXPERIMENTAL RESULTS

According to the sampling policies in Table 1, 4 data sets of cross-domain test are generated by U1-I1, U2-I2, U2-I3, U3-I4 on MovieLens 10M, Netflix 3M and EachMovie 2M.

TABLE 2. The comparison of 4 methods on MAE and RMSE.

sampling policies	auxiliary domain	target domain		
		MovieLens 10M	Netflix 3M	EachMovie 2M
U1-I1	Funk-SVD	1.387	1.399	1.151
	CBT	0.827	0.714	0.745
	RMGM	0.798	0.698	0.775
	CIAM	0.692	0.62	0.65
U2-I2	Funk-SVD	1.309	1.327	1.145
	CBT	0.775	0.689	0.726
	RMGM	0.687	0.639	0.692
	CIAM	0.63	0.628	0.589
U2-I3	Funk-SVD	1.302	1.327	1.107
	CBT	0.715	0.607	0.641
	RMGM	0.615	0.591	0.658
	CIAM	0.625	0.55	0.579
U3-I4	Funk-SVD	1.265	1.311	1.015
	CBT	0.695	0.469	0.591
	RMGM	0.587	0.565	0.628
	CIAM	0.564	0.492	0.503

Sampling Policies	Methods	RMSE		
		MovieLens 10M	Netflix 3M	EachMovie 2M
U1-I1	Funk-SVD	1.574	1.583	1.512
	CBT	0.83	0.938	1.014
	RMGM	0.884	0.92	0.768
	CIAM	0.842	0.805	0.782
U2-I2	Funk-SVD	1.48	1.494	1.428
	CBT	0.825	0.9	0.937
	RMGM	0.855	0.908	0.765
	CIAM	0.794	0.748	0.737
U2-I3	Funk-SVD	1.479	1.472	1.37
	CBT	0.822	0.852	0.864
	RMGM	0.828	0.904	0.751
	CIAM	0.706	0.698	0.709
U3-I4	Funk-SVD	1.414	1.4	1.286
	CBT	0.785	0.744	0.797
	RMGM	0.723	0.849	0.661
	CIAM	0.662	0.55	0.565

We carry out Funk-SVD, CBT, RMGM and CIAM on these 4 data sets, and compare MAE and RMSE on the 4 methods as shown in Table 2.

From the comparison between MAE and RMSE in Table 2, the correlation between the auxiliary domain and the target domain increases from U1-I1 to U3-I4, and the accuracy of CBT, RMGM, and CIAM algorithms gradually improves. The reason is that the overlap between auxiliary and target domain is higher, cross-domain recommendation is close to single domain recommendation. So the correlation of user-item is stronger, each recommendation method performs more accurate. For U1-I1, the difference between auxiliary and target domain is large, and the performances of Funk-SVD, CBT, RMGM, and CIAM are similar. This shows that for the cross-domain recommendation problem with less correlation, the cross-domain recommendation methods CBT, RMGM, CIAM have no obvious advantage over the single domain recommendation method Funk-SVD. For U2-I2, the correlation between auxiliary and target domain is enhanced. The cross-domain recommendation methods CBT, RMGM, CIAM perform better than the single domain recommendation method Funk-SVD, but there is little difference between CBT, RMGM, and CIAM. This shows that when auxiliary and target domains are weakly related, the performance of each cross-domain recommendation method is similar. For U2-I3 and U3-I4, the cross-domain recommendation methods CBT, RMGM, and CIAM perform significantly better than

the single-domain recommendation method Funk-SVD, and the CIAM performs better than CBT and RMGM. This shows that when correlation between auxiliary and target domain is high, the cross-domain recommendation methods can fully utilize auxiliary domain to optimize the recommendation on target domain.

In summary, through the analysis of this experiment, it can be seen that CIAM considering item-side alignment performs better than other methods such as Funk-SVD, CBT and RMGM in cross-domain recommendation. The item-side alignment strategy can fully utilize auxiliary domain on cross-domain recommendation, so this strategy can improve the performance of cross-domain transfer learning.

VI. CONCLUSION

In this paper, a cross-domain recommendation algorithm CIAM based on item-side alignment is proposed. This method linearly aligns the item-side of auxiliary and target domain by correlation coefficients. The aligned item of auxiliary domain is transferred into target domain, avoiding the interference of the user features in auxiliary domain to the user features in target domain. In addition, this paper preserves the user features in the target domain during the UV decomposition process. The item transfer process takes into account the item features in the auxiliary domain and emphasizes on retaining user features in target domain. Therefore, the proposed CIAM not only considers the transfer learning problem from auxiliary domain to target domain, but also considers the user features in target domain. That can alleviate the overfitting problem in cross-domain transfer learning.

By experimental comparison, it can be seen that the aligned item-side can sufficiently extract the correlation between auxiliary and target domain, to improve the performance of transfer learning based cross-domain recommendation method. The proposed CIAM can provide a research approach on transfer learning for cross-domain recommendation problems. The side information alignment can strengthen the rationality of transfer learning models. The disadvantage of this study is that CCA requires a full rating matrix. Therefore, it is necessary to pre-process the auxiliary and target domain by UserCF or ItemCF methods, which increases the algorithm complexity. So the future work of this research is to reduce the dependency of item alignment on full matrix, and strengthen the dependence on original rating to enhance the effectiveness of cross-domain recommendation.

REFERENCES

- [1] S. Boutemedjet and D. Ziou, "Predictive approach for user long-term needs in content-based image suggestion," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 8, pp. 1242–1253, Aug. 2012.
- [2] X. Luo, M. Zhou, S. Li, Z. You, Y. Xia, and Q. Zhu, "A nonnegative latent factor model for large-scale sparse matrices in recommender systems via alternating direction method," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 3, pp. 579–592, Mar. 2016.
- [3] (2006). *Funk*. [Online]. Available: <http://sifter.org/~simon/journal/200612/11.html>
- [4] I. Fernandez-Tobas, I. Cantador, M. Kaminskas, and F. Ricci, "Cross-domain recommender systems: A survey of the state of the art," in *Proc. Spanish Conf. Inf. Retr.*, 2012, pp. 1–12.

- [5] R. Salakhutdinov and A. Mnih, "Probabilistic matrix factorization," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2007, pp. 1257–1264.
- [6] T. Chen, W. Zhang, Q. Lu, K. Chen, Z. Zheng, and Y. Yu, "Svdfeature: A toolkit for feature-based collaborative filtering," *J. Mach. Learn. Res.*, vol. 13, pp. 3619–3622, Dec. 2012.
- [7] Y. Shi, M. Larson, and A. Hanjalic, "List-wise learning to rank with matrix factorization for collaborative filtering," in *Proc. 4th ACM Conf. Recommender Syst.*, 2010, pp. 269–272.
- [8] L. Si and R. Jin, "Flexible mixture model for collaborative filtering," in *Proc. 20th Int. Conf. Mach. Learn. (ICML)*, 2003, pp. 704–711.
- [9] B. Li, Q. Yang, and X. Xue, "Can movies and books collaborate? cross-domain collaborative filtering for sparsity reduction," in *Proc. 21st Int. Joint Conf. Artif. Intell.*, 2009, pp. 1–6.
- [10] W. Pan, N. N. Liu, E. W. Xiang, and Q. Yang, "Transfer learning to predict missing ratings via heterogeneous user feedbacks," in *Proc. 22nd Int. Joint Conf. Artif. Intell.*, 2011, pp. 1–7.
- [11] W. Pan, E. W. Xiang, N. N. Liu, and Q. Yang, "Transfer learning in collaborative filtering for sparsity reduction," in *Proc. 24th AAAI Conf. Artif. Intell.*, 2010, pp. 230–235.
- [12] X. Yu, F. Jiang, J. Du, and D. Gong, "A cross-domain collaborative filtering algorithm with expanding user and item features via the latent factor space of auxiliary domains," *Pattern Recognit.*, vol. 94, pp. 96–109, Oct. 2019.
- [13] B. Li, Q. Yang, and X. Xue, "Transfer learning for collaborative filtering via a rating-matrix generative model," in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, 2009, pp. 617–623.
- [14] Q. Zhang, D. Wu, J. Lu, F. Liu, and G. Zhang, "A cross-domain recommender system with consistent information transfer," *Decis. Support Syst.*, vol. 104, pp. 49–63, Dec. 2017.
- [15] B. Gong, K. Grauman, and F. Sha, "Learning kernels for unsupervised domain adaptation with applications to visual object recognition," *Int. J. Comput. Vis.*, vol. 109, nos. 1–2, pp. 3–27, Aug. 2014.
- [16] B. M. Marlin, R. S. Zemel, S. T. Roweis, and M. Slaney, "Recommender systems: Missing data and statistical model estimation," in *Proc. 22nd Int. Joint Conf. Artif. Intell.*, 2011, p. 2968.
- [17] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Mar. 2003.
- [18] X. Yu, F. Jiang, J. Du, and D. Gong, "A user-based cross domain collaborative filtering algorithm based on a linear decomposition model," *IEEE Access*, vol. 5, pp. 27582–27589, 2017.
- [19] L. Hu, J. Cao, G. Xu, L. Cao, Z. Gu, and C. Zhu, "Personalized recommendation via cross-domain triadic factorization," in *Proc. 22nd Int. Conf. World Wide Web*, 2013, pp. 595–606.
- [20] S. Berkovsky, T. Kuflik, and F. Ricci, "Cross-domain mediation in collaborative filtering," in *Proc. Int. Conf. User Modeling*. Springer, 2007, pp. 355–359.
- [21] A. P. Singh and G. J. Gordon, "Relational learning via collective matrix factorization," in *Proc. 14th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2008, pp. 650–658.
- [22] X. Yu, Y. Chu, F. Jiang, Y. Guo, and D. Gong, "SVMs classification based two-side cross domain collaborative filtering by inferring intrinsic user and item features," *Knowl.-Based Syst.*, vol. 141, pp. 80–91, Feb. 2018.
- [23] B. Loni, Y. Shi, M. Larson, and A. Hanjalic, "Cross-domain collaborative filtering with factorization machines," in *Proc. Eur. Conf. Inf. Retr.* Springer, 2014, pp. 656–661.
- [24] Q. Zhang, J. Lu, D. Wu, and G. Zhang, "A cross-domain recommender system with kernel-induced knowledge transfer for overlapping entities," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 7, pp. 1998–2012, Jul. 2019.
- [25] Y. Shi, M. Larson, and A. Hanjalic, "Tags as bridges between domains: Improving recommendation with tag-induced cross-domain collaborative filtering," in *Proc. Int. Conf. User Modeling, Adaptation, Personalization*, pp. 305–316, Springer, 2011.
- [26] S. Rendle, "Factorization machines with libFM," *ACM Trans. Intell. Syst. Technol.*, vol. 3, no. 3, pp. 1–22, May 2012.
- [27] X. Yu, "Training SVMs on a bound vectors set based on Fisher projectio," *Frontiers Comput. Sci.*, vol. 8, no. 5, pp. 793–806, 2014.



LANTING WANG received the B.S. degree in Internet of Things from Hunan Normal University. She is currently pursuing the M.S. degree in computer technology with Ningbo University. Her current research interests include multiple classifier and prediction systems, processing and modeling of uncertainty in predictive modeling, recommendation systems, and decision support systems.



YU XIN received the Ph.D. degree in computer science and technology from Harbin Engineering University, in 2015. He is currently an Associate Professor with the Faculty of Electrical Engineering and Computer Science, Ningbo University, China. He engaged in the researches on social network analysis. His current research interests include multiple classifier and prediction systems, processing and modeling of uncertainty in predictive modeling, recommendation systems, diagnostic analysis, and decision support systems.

• • •