

“© 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

Enhancing Cross Domain Recommendation with Domain Dependent Tags

Peng Hao

Centre for Quantum Computation and
Intelligent Systems
Faculty of Engineering and
Information Technology
University of Technology Sydney
Sydney, Australia 2007
Email: peng.hao@student.uts.edu.au

Guangquan Zhang

Centre for Quantum Computation and
Intelligent Systems
Faculty of Engineering and
Information Technology
University of Technology Sydney
Sydney, Australia 2007
Email: guangquan.zhang@uts.edu.au

Jie Lu

Centre for Quantum Computation and
Intelligent Systems
Faculty of Engineering and
Information Technology
University of Technology Sydney
Sydney, Australia 2007
Email: jie.lu@uts.edu.au

Abstract—One challenge in recommender system is to deal with data sparsity. To handle this issue, social tags are utilized to bring disjoint domains together for knowledge transfer in cross-domain recommendation. The most intuitive way is to use common tags that present in both source and target domains. However, it is difficult to obtain a strong domain connection by exploiting a small amount of common tags, especially when the tagging data in target domain is too scarce to share enough common tags with source domain. In this paper we propose a novel framework, called *Enhanced Tag-induced Cross Domain Collaborative Filtering* (ETagiCDCF), to integrate the rich information contained in domain dependent tags into recommendation procedure. We perform experiments on two public datasets and compare with several single and cross domain recommendation approaches, the results demonstrate that ETagiCDCF can effectively address data sparseness and improve recommendation performance.

1. Introduction

Collaborative filtering (CF) is one of fundamental techniques in making most recommender systems [1], [2], [3], [4]. It only utilizes historical rating data to predict the interests of an active user on unseen items and as a result can be applied in most recommendation scenarios. However, CF approach is widely known to be suffered from data sparsity problem [5], which is caused since most of users usually rated a small number of items in reality. To address data sparseness, some researchers adopted the idea of transfer learning [6], [7] and developed some cross domain collaborative filtering (CDCF) approaches [8], [9], [10]. The underlying idea of CDCF is to employ the shared knowledge in the other domains to facilitate the recommendation making in target domain.

Performance of CDCF depends on whether an effective relation can be built as bridge to connect different domains [11]. Without any overlapping of users or items, user-contributed tags provide a possible way to achieve this goal. A typical method called TagCDCF is proposed in [9], in which the

model is based on an intuitive idea that users with same tastes tend to attach similar tags to items. For example, in figure 1, we want to guess the preferences of users on books (indicated by domain B) by borrowing recommendation data of movies (indicated by domain A). It is difficult to infer Bob’s preference on the book “The matrix” with CF methods because his preference is not similar to any of users in domain B. But in domain A we have Alex rated 4.5 and tagged ‘*sci-fi*’ to the movie “Avengers”. In this case, it is possible to assume that Bob will also enjoy the book since he used the same tag with Alex. However, there are two limitations in above assumption. First, if two disjoint domains share limited tags, which may happen to a new system without holding enough tagging data. A weak domain connection will be formed since most of users and items in both domains are not covered by those limited common tags. Consider the same example shown in figure 1, it is not possible to infer Bob’s preference on the book “The girl on the train” because the only tag he had given to that book is a domain dependent tag, which means it is domain exclusive and does not appear in domain A. Second, it is a waste to discard domain dependent tags, which are effective to reflect the characteristics of user preference in the individual domain. If we can take them into consideration, more linkage will be built up to enhance the domain connection. For example in figure 1, we discover that users in domain A are inclined to assign ‘*crime*’ and ‘*action*’ to “fast and furious 7”, while ‘*crime*’ and ‘*thriller*’ are both used to label “The girl on the train” in domain B. Under such observation, we can merge ‘*action*’ and ‘*thriller*’ into same cluster by their co-occurrences with the common tag ‘*crime*’. Then Jimmy and Bob will be linked up and Jimmy’s rating knowledge can be borrowed to predict Bob’s preference on the book “The girl on the train” as well, though no direct relationship between their tastes exists. Motivated by above example, here we explore alternative approach to exploit user-contributed tags in connecting different domains. Specifically, we propose a novel framework, called *Enhanced Tag-induced Cross Domain Collaborative Filtering* (ETagiCDCF), to explore the role of domain de-

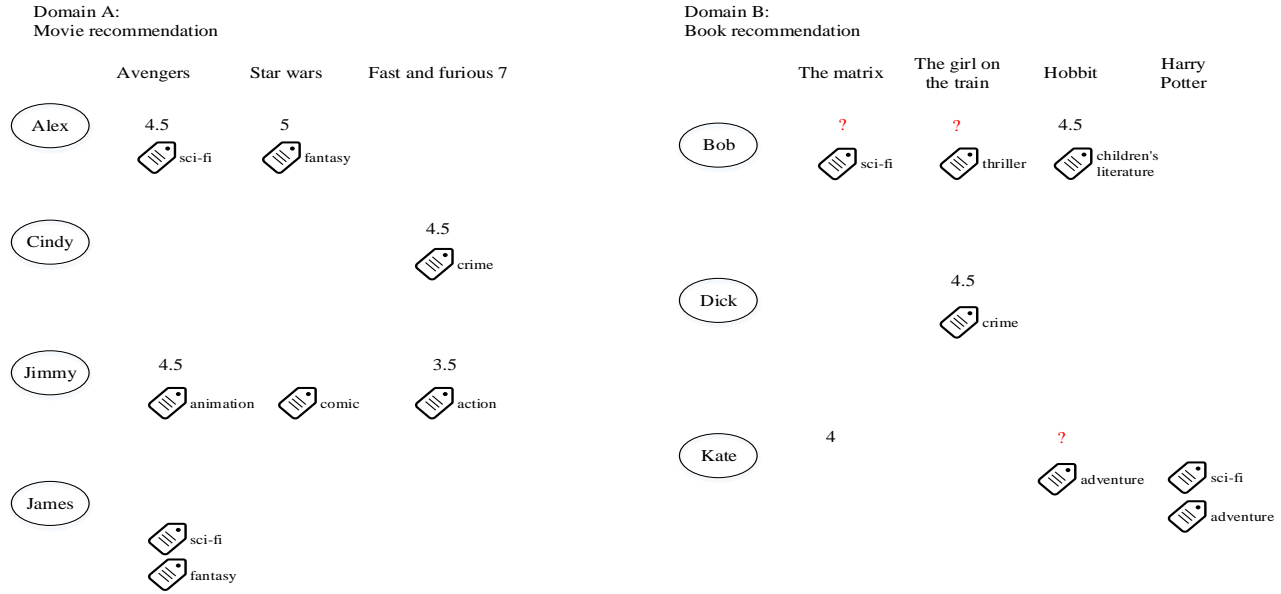


Figure 1. The scenario of cross-domain recommendation

pendent tags in improving cross domain recommendation quality. The key challenge lies in the fact that domain dependent tags are domain specific. To overcome this problem, adapted spectral clustering is applied to group the domain dependent tags based on their co-occurrences with common tags, so that similar tags in different domains will be gathered together to form a compact tag representation. We then build user and item profiles with the tag representation to map cross domain users and items to the same subspace for evaluation. As a result, different domains can be connected from the view of both users and items, and thus useful knowledge will be transferred among domains.

The contribution of this work is three folds: First, to the best of our knowledge, ETagiCDCF is the first model that incorporates the information of domain dependent tags into cross domain recommendation system. Second, ETagiCDCF outperforms its counterpart TagCDCF, which shows domain dependant tags contains more useful information than limited common tags and ETagiCDCF can utilize this kind of information very effectively. Third, ETagiCDCF can alleviate data sparseness and increase the quality of recommendation though there are limited common tags available.

The remainder of this paper is organised as follows. Section 2 discusses related work. Section 3 gives some basic definitions and notations that are used in this paper. The proposed method will then be presented in Section 4. A series of experiments are conducted in Section 5. Finally we draw conclusions and propose some research directions for future in Section 6.

2. Related Work

Existing cross domain recommendation approaches mainly exploit two categories of relations in establishing

domain connection, which are content-based relations and rating-based relations, respectively [10].

The research proposed in rating-based relations focus on mining implicit rating patterns and latent features from multiple rating matrices as the shared knowledge to link different domains. Specially, rating pattern sharing aims to build bridge between two rating matrices at user and item cluster level simultaneously. In [8], the authors proposed to extract compact cluster-level rating pattern that is called codebook from auxiliary domain first, and then merge this codebook into target domain to reconstruct rating matrix. Similar ideas were also adopted in the following works [12], [13]. Latent feature sharing aims to factorize rating matrix to obtain latent user and item feature matrices and shares those matrices across domains as the transferred knowledge. Pan et al. [14] assumed that a set of tastes, which refer to common coordinate systems for representing preferences of users and main factors of items, are shared between auxiliary domain and target domain. Compared to the adaptive model in codebook based approach [8], whose knowledge transfer process is conducted in a sequence, coordinate system transfer method proposes to construct the latent feature space for both auxiliary and target domains collectively and integrates the implicit feedback (in binary form) from auxiliary domain to predict numerical ratings in target domain [15]. ETagiCDCF is substantially different from all the aforementioned approaches since it exploits social tags for establishing domain connection.

For content-based relations, social tags are most widely studied to bridge up different domains. In this direction, Wang et al. [16] proposed a tag transfer learning (TTL) model, which clusters tags in auxiliary domain based on tag topics and transfers those tag topics to target domain to facilitate user similarity calculation. Though TTL applies

tags in linking different domains, its performance is based on tag clustering rather than matrix factorization, which is a fundamental technique adopted in ETagiCDCF for generating recommendations. The most similar work to ETagiCDCF is proposed in [9], where a tag-induced cross domain collaborative filtering (TagCDCF) algorithm was developed. TagCDCF computes cross domain user and item similarities by modelling user and item profiles with common tags, then those similarities are integrated into collective matrix factorization process to regularize the factorization process. However, their approach requires that target users to tag several items in order to gather enough common tags to cover most pairs of users and items in both domains. ETagiCDCF extends this model and exploits abundant domain dependent tags to link different domains for more effective knowledge transfer when common tags are limited.

3. Problem Setting

In this section, we are going to present some definitions and give the basic notations that will be used in the rest of the paper.

Definition 1 (Domain): According to the survey [17], the term of domain refers to types of items.

We adopt the same definition in our paper. For the ease of explanation, we only use two domains: one source domain D^s and one target domain D^t , $I(I = \{s, t\})$ denotes domain indices. In the π -th ($\pi \in I$) domain, all the ratings are denoted by $R^\pi = \{r_{ij}\}$, where $i \in \{1, 2, \dots, m_\pi\}$ is the user index and $j \in \{1, 2, \dots, n_\pi\}$ is the item index. In addition, we also utilize tagging information. Each observed tagging data is a 3-tuple (i, j, t_{ij}) where t_{ij} is the tag that is assigned by user i to item j . All the unique tag t_{ij} consists of tag set T^π .

Definition 2 (Common tags): Given unique tag sets T^s and T^t , we define common tags as a set of tags that appear both in D^s and D^t , and they can be indicated by $T_c = T^s \cap T^t$. As to our example in figure 1, we have two common tags and therefore $T_c = \{\text{'sci-fi'}, \text{'crime'}\}$.

Definition 3 (Domain dependent tags): We define domain dependent tags as a set of domain exclusive tags. Under this definition, domain dependent tags for both source and target domains are denoted by $T_d^s = T^s \setminus T_c = T^s - T_c$ and $T_d^t = T^t \setminus T_c = T^t - T_c$, respectively.

For our example, if we treat movie as source domain and book as target domain, then we have $T_d^s = \{\text{'fantasy'}, \text{'action'}, \text{'animation'}, \text{'comic'}\}$ and $T_d^t = \{\text{'thriller'}, \text{'children's literature'}, \text{'adventure'}\}$.

4. Enhanced Tag-induced Cross Domain Collaborative Filtering

Given two domains D^s and D^t , suppose there is limited tagging data T^t in D^t , while T^s in D^s is comparatively rich. As a result common tags T_c are not enough in building an effective domain connection between D^s and D^t . The task of ETagiCDCF is to group similar tags in T_d^s and T_d^t into

same clusters and model a new user/item profile based on those tag clusters, so that cross domain users and items can be mapped to the same space for comparison and useful knowledge can be transferred from D^s to D^t .

ETagiCDCF mainly consists of following three steps: 1) Mapping domain dependent tags into pre-defined k clusters to reduce the difference between domain dependent tags. 2) Modelling user and item profiles on those tag clusters and compute cross domain user-to-user and item-to-item similarities. 3) Integrating computed cross domain similarities into collective matrix factorization to regularize factorization and generate recommendation.

4.1. Domain dependent Tags Alignment

We apply adapted spectral clustering [18] to align domain dependent tags. More specifically, the clustering of domain dependent tags is implemented by the co-occurrence relationship between common and domain dependent tags. As pointed out in [19], both user-tag and item-tag interactions should be fully explored to maximize the role of tagging information in recommendation. So we propose to model tag co-occurrence relationship on user and item side, respectively.

For users, user-based tag co-occurrence is defined as following:

$$C_u(t_d, t_c) = \begin{cases} 1 & \text{if } UN(t_d, t_c) \geq 1 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where $UN(t_d, t_c)$ denotes the number of users who have given both domain dependent tag t_d and common tag t_c . $C_u(t_d, t_c) = 1$ means t_d and t_c are both favored in the user's tagging behavior, otherwise $C_u(t_d, t_c) = 0$. Based on user-based tag co-occurrence, we can pick out user-based domain dependent tags, which are denoted by $DS_u = \{t_d | \forall t_d \in T_d^s \cup T_c, \exists t_c \in T_c \Rightarrow C_u(t_d, t_c) = 1\}$.

Similarly, item-based tag co-occurrence is defined below:

$$C_v(t_d, t_c) = \begin{cases} 1 & \text{if } VN(t_d, t_c) \geq 1 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where $VN(t_d, t_c)$ denotes the number of items which are labeled by both domain dependent tag t_d and common tag t_c . Based on item-based tag co-occurrence, item-based domain dependent tags can be filtered and indicated by $DS_v = \{t_d | \forall t_d \in T_d^t \cup T_c, \exists t_c \in T_c \Rightarrow C_v(t_d, t_c) = 1\}$.

For our example in Figure 1, given $T_c = \{\text{'sci-fi'}, \text{'crime'}\}$, we select $\{\text{'fantasy'}\}$ and $\{\text{'thriller'}, \text{'children's literature'}, \text{'adventure'}\}$ from movie and book domain respectively due to user-based tag co-occurrence and then the union of these two sets forms $DS_U = \{\text{'fantasy'}, \text{'thriller'}, \text{'children's literature'}, \text{'adventure'}\}$. Similarly, we select $\{\text{'fantasy'}, \text{'animation'}, \text{'action'}\}$ and $\{\text{'thriller'}, \text{'adventure'}\}$ from Movie and Book domain respectively due to item-based tag co-occurrence and the union of these two sets results in $DS_V = \{\text{'fantasy'}, \text{'animation'}, \text{'action'}, \text{'thriller'}, \text{'adventure'}\}$.

After filtering domain dependent tags from both user and

Algorithm 1 Domain Dependent Tags Alignment

- Input:** original domain dependent tags T_d^s and T_d^t , common tags T_c , number of clusters k .
- Output:** user-based domain dependent tags DS_u , item-based domain dependent tags DS_v , user-based domain dependent tag membership matrix M_u , item-based domain dependent tag membership matrix M_v .
- 1: Apply Equation 1 and 2 on the union of $T_s^d \cup T_t^d$ to select l_u domain dependent tags DS_u from user side and l_v domain dependent tags DS_v from item side, respectively.
 - 2: Based on DS_u and T_c , calculate user domain dependent tag-common tag co-occurrence matrix $Q^u \in \mathbb{R}^{l_u \times l_c}$, where l_c is the number of common tags, and $Q_{i,j}^u = 1$ if $C_u(t_{d_i}, t_{c_j}) = 1$, otherwise $Q_{i,j}^u = 0$. Similarly, based on DS_v and T_c , calculate item domain dependent tag-common tag co-occurrence matrix $Q^v \in \mathbb{R}^{l_v \times l_c}$.
 - 3: **for** each $z \in [u, v]$ **do**
 - 4: Construct matrix $L_z = (D^z)^{-1/2} \times A^z \times (D^z)^{-1/2}$, where $A^z = \begin{bmatrix} \mathbf{0}_{l_z \times l_z} & Q^z \\ (Q^z)^T & \mathbf{0}_{l_c \times l_c} \end{bmatrix}$, $\mathbf{0}$ is a zero matrix, D^z is a diagonal matrix, and $D_i^z = \sum_j A_{i,j}^z$.
 - 5: Find the k largest eigenvectors of L_z , u_1, u_2, \dots, u_k , and form the matrix $E_z = [u_1, u_2, \dots, u_k]$.
 - 6: **end for**
 - 7: Extract the first l_u rows of matrix $E_u \in \mathbb{R}^{(l_u+l_c) \times k}$ to obtain $M_u \in \mathbb{R}^{l_u \times k}$ and first l_v rows of matrix $E_v \in \mathbb{R}^{(l_v+l_c) \times k}$ to obtain $M_v \in \mathbb{R}^{l_v \times k}$.
 - 8: **return** DS_u, DS_v, M_u and M_v
-

item perspectives, our next goal is to group corresponding domain dependent tags into k clusters by spectral clustering. The complete procedure of our domain dependent tags alignment is described in Algorithm 1.

As proved in [20], the k principle components, which refers to the k largest eigenvectors u_1, u_2, \dots, u_k in step 5, can be used to cluster original data in the subspace spanned by those k principle components. For our case, we consider those k principle components as the high-level tag representation by clustering domain dependent tags. The learned tag representation can effectively eliminate the difference among domain dependent tags and serves as common features shared by different domains to define new user and item profiles.

4.2. Cross Domain Similarities Refinement

In this section, we will first construct mapping functions to map across domain users and items to the new subspace spanned by domain dependent tag clusters and then compute cross domain user-to-user and item-to-item similarities for linking different domains.

Given our motivation of modelling user and item profiles with domain dependent tag clusters that are shared by different domains as high-level features, we need to independently learn the role of domain dependent tag clusters in describing

Algorithm 2 Cross Domain Similarities Refinement

- Input:** user-based domain dependent tag membership matrix M_u , item-based domain dependent tag membership matrix M_v .
- Output:** cross domain user-to-user similarity matrix $S^u \in \mathbb{R}^{m_s \times m_t}$ and item-to-item similarity matrix $S^v \in \mathbb{R}^{n_s \times n_t}$.
- 1: Apply Equation 3 to generate source domain user-tag cluster relation matrix $UT^s \in \mathbb{R}^{m_s \times k}$ and target domain user-tag cluster relation matrix $UT^t \in \mathbb{R}^{m_t \times k}$.
 - 2: Apply Equation 4 to generate source domain item-tag cluster relation matrix $VT^s \in \mathbb{R}^{n_s \times k}$ and target domain item-tag cluster relation matrix $VT^t \in \mathbb{R}^{n_t \times k}$.
 - 3: **for** user $i \in [1, 2, \dots, m_s]$ **do**
 - 4: **for** user $p \in [1, 2, \dots, m_t]$ **do**
 - 5:
$$\rho_{i,p} = \frac{\sum_{f=1}^k (UT_{if}^s - \overline{UT}_i^s)(UT_{pf}^t - \overline{UT}_p^t)}{\sqrt{\sum_{f=1}^k (UT_{if}^s - \overline{UT}_i^s)^2} \sqrt{\sum_{f=1}^k (UT_{pf}^t - \overline{UT}_p^t)^2}},$$
 where $\overline{UT}_i^s = \frac{\sum_{f=1}^k UT_{if}^s}{k}$
 - 6: $S_{i,p}^u = \frac{1}{1 + \exp(-\rho_{i,p})}$
 - 7: **end for**
 - 8: **end for**
 - 9: **for** item $j \in [1, 2, \dots, n_s]$ **do**
 - 10: **for** item $q \in [1, 2, \dots, n_t]$ **do**
 - 11:
$$\rho_{j,q} = \frac{\sum_{f=1}^k (VT_{jf}^s - \overline{VT}_j^s)(VT_{qf}^t - \overline{VT}_q^t)}{\sqrt{\sum_{f=1}^k (VT_{jf}^s - \overline{VT}_j^s)^2} \sqrt{\sum_{f=1}^k (VT_{qf}^t - \overline{VT}_q^t)^2}},$$
 where $\overline{VT}_j^s = \frac{\sum_{f=1}^k VT_{jf}^s}{k}$
 - 12: $S_{j,q}^v = \frac{1}{1 + \exp(-\rho_{j,q})}$
 - 13: **end for**
 - 14: **end for**
 - 15: **return** S^u and S^v
-

user preferences and encoding item properties. In order to capture such relationship, we propose following linear mapping functions:

$$\Phi_U(X^\pi) = X^\pi \times M_u \quad (3)$$

$$\Phi_V(Y^\pi) = Y^\pi \times M_v \quad (4)$$

where $X^\pi \in \mathbb{R}^{m_\pi \times l_u}$ is user-domain dependent tag frequency matrix and X_{ix}^π indicates the tagging frequency of user i in using domain dependent tag x in domain π . Similarly, $Y^\pi \in \mathbb{R}^{n_\pi \times l_v}$ is item-domain dependent tag frequency matrix and Y_{jy}^π indicates the tagging frequency of item j labeled by domain dependent tag y in domain π . Based on Equations 3 and 4, we are able to define user and item profiles with the new tag representation, then we can adopt Pearson similarity in [19] to compute cross domain user-to-user and item-to-item similarities directly. The overall procedure of cross-domain similarities refinement is summarised in Algorithm 2.

4.3. Model Formulation and Learning

Similar to [9], given cross-domain similarity matrices S^u and S^v , we integrate those two matrices as additional

constraints into the collective matrix factorization process, serving as a bridge to bring source and target domains together. Hence useful knowledge can be transferred from source domain to benefit the rating prediction in target domain. Overall objective function of ETagiCDCF can be formulated as below:

$$\begin{aligned}
F = & \frac{1}{2} \sum_{i=1}^{m_s} \sum_{j=1}^{n_s} I_{ij}^s (R_{ij}^s - (U_{*i}^s)^T V_{*j}^s)^2 \\
& + \frac{1}{2} \sum_{p=1}^{m_t} \sum_{q=1}^{n_t} I_{pq}^t (R_{pq}^t - (U_{*p}^t)^T V_{*q}^t)^2 \\
& + \frac{\alpha}{2} \sum_{i=1}^{m_s} \sum_{p=1}^{m_t} (S_{ip}^U - (U_{*i}^s)^T U_{*p}^t)^2 \\
& + \frac{\beta}{2} \sum_{j=1}^{n_s} \sum_{q=1}^{n_t} (S_{jq}^V - (V_{*j}^s)^T V_{*q}^t)^2 \\
& + \frac{\lambda}{2} (\|U^s\|_F^2 + \|V^s\|_F^2 + \|U^t\|_F^2 + \|V^t\|_F^2)
\end{aligned} \tag{5}$$

where $U^s \in \mathbb{R}^{d \times m_s}$, whose i th column U_{*i}^s denotes the d -dimensional latent feature vector for user i in source domain, while $V^s \in \mathbb{R}^{d \times n_s}$, whose j th column V_{*j}^s denotes the d -dimensional latent feature vector for item j in source domain. Similarly we have $U^t \in \mathbb{R}^{d \times m_t}$ and $V^t \in \mathbb{R}^{d \times n_t}$ representing latent feature matrices of users and items in target domain respectively. α and β are two trade-off parameters, which control the relative importance of cross-domain user-to-user and item-to-item similarity, respectively. λ is the regularization parameter used to avoid over-fitting. Our objective is to estimate four optimal variables U^s, V^s, U^t, V^t . The local minimum solution for Equation 5 can be found by performing gradient descent on four variables U^s, V^s, U^t, V^t with alternatively updating one while fixing the other three unchanged. Specifically, the gradient with respect to each variable is computed as below:

$$\begin{aligned}
\frac{\partial F}{\partial U_{*i}^s} = & \sum_{j=1}^{n_s} I_{ij}^s ((U_{*i}^s)^T V_{*j}^s - R_{ij}^s) V_{*j}^s \\
& + \alpha \sum_{p=1}^{m_t} ((U_{*i}^s)^T U_{*p}^t - S_{ip}^U) U_{*p}^t + \lambda U_{*i}^s
\end{aligned} \tag{6}$$

$$\begin{aligned}
\frac{\partial F}{\partial V_{*j}^s} = & \sum_{i=1}^{m_s} I_{ij}^s ((V_{*j}^s)^T U_{*i}^s - R_{ij}^s) U_{*i}^s \\
& + \beta \sum_{q=1}^{n_t} ((V_{*j}^s)^T V_{*q}^t - S_{jq}^V) V_{*q}^t + \lambda V_{*j}^s
\end{aligned} \tag{7}$$

$$\begin{aligned}
\frac{\partial F}{\partial U_{*p}^t} = & \sum_{q=1}^{n_t} I_{pq}^t ((U_{*p}^t)^T V_{*q}^t - R_{pq}^t) V_{*q}^t \\
& + \alpha \sum_{i=1}^{m_s} ((U_{*i}^s)^T U_{*p}^t - S_{ip}^U) U_{*i}^s + \lambda U_{*p}^t
\end{aligned} \tag{8}$$

TABLE 1. STATISTICS OF DATASETS USED IN THE EXPERIMENTS

	MovieLens 10M	LibraryThing
#users	5000	5000
#items	5000	5000
#ratings	584628	179419
rating sparsity	97.70%	99.30%
#unique tags	672	5408
#tag assignments	1322	179419
ratio of overlapping tags	32.44%	4.03%

$$\begin{aligned}
\frac{\partial F}{\partial V_{*q}^t} = & \sum_{p=1}^{m_t} I_{pq}^t ((U_{*p}^t)^T V_{*q}^t - R_{pq}^t) U_{*p}^t \\
& + \beta \sum_{j=1}^{n_s} ((V_{*j}^s)^T V_{*q}^t - S_{jq}^V) V_{*j}^s + \lambda V_{*q}^t
\end{aligned} \tag{9}$$

In the training phrase, we update the variables according to the following rules:

$$\begin{aligned}
U^s &= U^s - \varepsilon \frac{\partial F}{\partial U^s} & V^s &= V^s - \varepsilon \frac{\partial F}{\partial V^s} \\
U^t &= U^t - \varepsilon \frac{\partial F}{\partial U^t} & V^t &= V^t - \varepsilon \frac{\partial F}{\partial V^t}
\end{aligned} \tag{10}$$

The learning rate ε determines the updating extent of variables during each iteration. We apply binary search to adjust learning rate automatically and set initial $\varepsilon = 0.001$.

5. Experiment

In this section, a series of experiments are conducted to evaluate the proposed ETagiCDCF approach, especially under the condition that there are a small number of common tags shared between source and target domains. We first describe the dataset and experimental settings that are used in our experiments. Then we study the impact of two trade-off parameters on the final recommendation performance. Further comparison with single and cross domain recommendation algorithms examines the effectiveness of ETagiCDCF in exploiting domain dependent tags to improve recommendation quality.

5.1. Dataset and Set up

We evaluate ETagiCDCF on two publicly available datasets: the MovieLens 10M data set¹ and LibraryThing data set². The MovieLens 10M (ML) dataset contains over 10 million ratings and 100,000 tag applications applied to 10,681 movies by 71,567 users. The LibraryThing (LT) dataset contains over 700,000 ratings and 2 million tag applications assigned by 7,564 users to 39,519 books. Ratings in both datasets are represented on 1-5 star scale with half star increment. Those two datasets contain both rating and tagging information for our experiments.

To make the fair evaluation of our model, we need to prune original dataset for our analysis. Specifically, we sample a

1. <http://www.grouplens.org/node/73>
2. <http://www.macle.nl/tud/LT/>

portion of the original dataset by keeping both rating records and tagging assignments that are made by first 5000 users on first 5000 items according to the identifiers in the original dataset. Under such selection, we can collect enough rating data to train our model and further limit the number of common tags shared between two datasets. Since in the original ML dataset, not every user had left a tag assignment when he or she made a rating record. Thus in our ML dataset not every movie has a tag applied. TABLE 1 shows the details of our final datasets.

In our experiments, we evaluate all the recommendation methods in two settings, namely set ML as source domain and LT as target domain, and vice versa. In both cases, we randomly select 20 percent of users and their corresponding rating records plus tag assignments to items as test set. The validation set contains 20 percent of randomly select users and their corresponding ratings and tag assignments. The training set contains the remaining 60 percent of users and their corresponding ratings and tag assignments. In addition, all the ratings and tag assignments from source domain are also added into training set as training data to build the model. Note that we only utilize the existing tag assignments in the training set for model learning. The validation set is used to tune parameters and experiments are performed on the test set for 5 times. We report average results as our final results

For consistency with experiments reported in the literature, all the results are evaluated by Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), which are widely adopted to measure the performance of rating prediction. A smaller value of MAE or RMSE indicates a better performance.

For the experiment setting of ETagiCDCF, we set tag cluster number k to 30 and dimensionality of latent features d to 10 since experiments on validation set reveal that this combination of parameters can achieve the best performance. The regularization parameter λ in Equation 5 is set to 0.01 after tuning on the validation set. The selection of two trade-off parameters α and β will be discussed later in the next section.

5.2. Impact of Parameters

In this subsection, we conduct experiments to investigate the impact of two trade-off parameters α and β in ETagiCDCF, which respectively controls the relative importance of cross domain user-to-user and item-to-item similarity to the final objective function in Equation 5.

Similar to [9], we adopt the same strategy to tune these parameters on the validation set of each experiment setting. We first fix $\beta = 0$ unchanged and vary the value of α alternatively to check the impact of parameter α in terms of MAE and RMSE. The changing results are reported in Fig. 2. Based on the results in Fig. 2, we set $\alpha = 0.01$ as it gets optimal results on both LT and ML datasets. By adopting $\alpha = 0.01$, we then tune the value of β to investigate the impact of β in terms of MAE and RMSE. The corresponding results are shown in Fig. 3. We finally choose $\alpha = 0.01$ and $\beta =$

0.01 as the optimal values in our experiment since under this setting ETagiCDCF can achieve best results in both datasets.

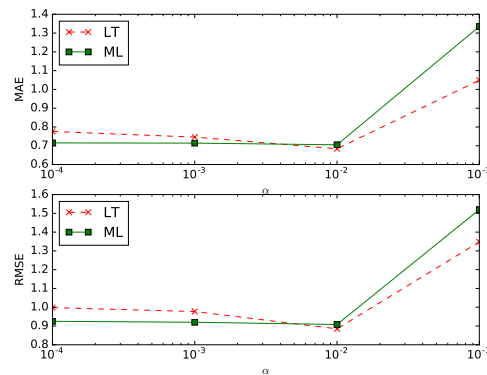


Figure 2. MAE and RMSE variations via changing α with $\beta = 0$ fixed

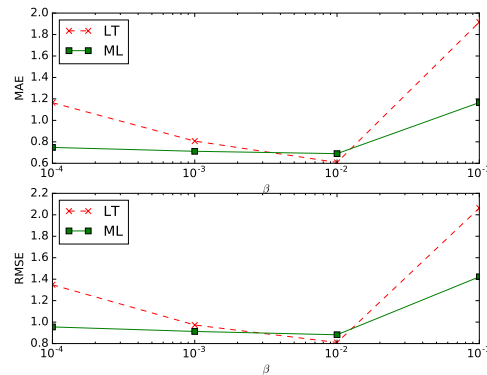


Figure 3. MAE and RMSE variations via changing β with $\alpha = 0.01$ fixed

5.3. Performance Comparison

We compare the performance of ETagiCDCF with some single and cross domain recommendation methods listed below:

UCF: User-based collaborative filtering [21] is a conventional memory-based single domain recommendation approach. It looks for the preferences of users who share similar tastes with the target user to make a recommendation. The key challenge is to compute similarities between pairs of users. In our implementation, the similarity is computed by Pearson correlation coefficient and the neighbourhood size is set to 50.

ICF: Item-based collaborative filtering [22] is another memory-based single domain recommendation approach. It generates recommendations for a user by finding potential items that are similar to the ones that the user had consumed before. Since the relationship between items is relatively static, item-item similarity model does not have to be built so often and as a result it reduces more computation. In our

TABLE 2. PERFORMANCE COMPARISON IN TERMS OF MAE

dataset	UCF	ICF	SVD	TagCDCF	ETagiCDCF
LT	0.749±0.003	0.731±0.009	0.741±0.002	0.730±0.084	0.628±0.006
ML	0.755±0.001	0.779±0.001	0.897±0.002	0.717±0.001	0.708±0.003

TABLE 3. PERFORMANCE COMPARISON IN TERMS OF RMSE

dataset	UCF	ICF	SVD	TagCDCF	ETagiCDCF
LT	0.954±0.004	0.941±0.009	0.945±0.002	0.944±0.105	0.810±0.007
ML	0.959±0.001	0.986±0.002	1.086±0.002	0.928±0.002	0.903±0.004

implementation, Pearson correlation coefficient is adopted to compute item-item similarity.

SVD: Singular value decomposition [23] is a well-known model-based single domain recommendation approach, which relies on matrix factorization technique and decomposes a rating matrix into three sub-matrices with reducing the dimensionality of the product space. It maps users and items into a low dimensional space and discovers the intrinsic relationship among the latent features of users and items for making recommendation. In our implementation, the dimension of latent feature space is set to 10.

TagCDCF: Tag-induced cross domain collaborative filtering [9] is a recently proposed cross domain recommendation approach, which exploits user-contributed tags that are common to multiple domains to establish the domain link for knowledge transfer. However, only common tags are utilized in this model. Note that all the parameters involved in TagCDCF are tuned based on the observation from validation set.

The performance of ETagiCDCF and other baseline approaches are shown in TABLE 2 and TABLE 3, where the number behind \pm denotes the standard deviations.

The results on the test set reveal several interesting points and can be applied to answer following three questions:

Q1. Can knowledge obtained from other auxiliary domains be useful to improve recommendation performance in the target domain?

To answer this question, we need to compare the results of cross domain recommendation approaches (i.e. TagCDCF, ETagiCDCF) with those classical single domain recommendation benchmarks (i.e. UCF, ICF) to check whether the recommendation performance in target domain can be greatly improved by transferring knowledge from the source domain. Table 2 and Table 3 show that both TagCDCF and ETagiCDCF can consistently outperform single domain baselines in terms of MAE and RMSE. For example, when evaluated by MAE, ETagiCDCF improves over ICF by 14.1% on LT dataset and 9.1% on ML dataset, and over UCF by 16.2% on LT and 6.7% on ML. Similarly, TagCDCF improves over ICF by 0.1% on LT and 8.0% on ML, and over UCF by 2.5% on LT and 5.0% on ML. These results indicate that exploiting common knowledge between domains can effectively improve the recommendation quality in target domain.

Q2. Can social tags provide additional information to benefit personalised recommendation?

To identify the role of social tags in improving recommendation performance, we need to compare the results of SVD with TagCDCF and ETagiCDCF since they are all built on matrix factorization model. The only difference lies in the fact that SVD only relies on ratings to make recommendation while TagCDCF and ETagiCDCF also integrate tag information into their models. The results in Table 2 and Table 3 reveal that both ETagiCDCF and TagCDCF perform better than SVD. Based on these observations, we can summarise that social tags indeed offer promising new information that goes beyond ratings to the factorization process.

Q3. Can domain dependent tags provide richer information than limited common tags?

To check the effectiveness of domain dependent tags over limited common tags in providing information to recommendation, we need to compare ETagiCDCF with TagCDCF. From Table 2 and Table 3, we observe that ETagiCDCF can consistently perform better than TagCDCF. The improvement on ML dataset is up to 1.3% in MAE and 2.7% in RMSE, while on LT dataset the improvement is up to 14.0% in MAE and 14.2% in RMSE. We find ETagiCDCF can significantly outperform TagCDCF on LT dataset because in our experiment ML dataset contains sparse tagging data, as a result the tags shared by both LT and ML are greatly reduced especially when ML is set as source domain. Without enough common tags to establish domain linkage will undermine performance of TagCDCF, which had already been proved in [9]. In contrast, ETagiCDCF are designed to utilize abundant domain dependent tags to link up two domains in this situation. The observed results clearly indicates the benefit of applying domain dependent tags to link different domains when limited number of common tags are given and ETagiCDCF can effectively utilize the information contained in domain dependent tags.

6. Conclusion and Future Work

In this paper, we address a challenge problem of transferring knowledge between non-overlapping domains via user-contributed tags. Specifically, we propose to utilize abundant domain dependent tags to build the domain link. To this end, we first form new tag representation by clustering domain dependent tags into pre-defined groups. Then, we build user and item profiles with derived tag representation and compute cross domain user-to-user and item-to-item similarities, which are further integrated into collective

matrix factorization to guide factorization process and act as the bridge to link different domains for knowledge transfer. The experimental results demonstrate that our algorithm is capable of establishing strong domain connection for knowledge transfer when limited common tags are available. Besides, we also observe that domain dependent tags can bring more information about user preferences into cross domain recommendation.

As future work, we are going to test our proposed method on a larger dataset and compare with more advanced approaches, such as [24], etc. We are also interested in discovering the critical boundary when adding more domain dependent tags will bring no further improvement on performance.

Acknowledgments

This work is supported by the Australian Research Council (ARC) under discovery grant DP140101366 and the China Scholarship Council.

References

- [1] P. Resnick and H. R. Varian, "Recommender systems," *Communications of the ACM*, vol. 40, no. 3, pp. 56–58, 1997.
- [2] J. Lu, D. Wu, M. Mao, W. Wang, and G. Zhang, "Recommender system application developments: A survey," *Decision Support Systems*, vol. 74, no. 0, pp. 12 – 32, 2015.
- [3] W. Wang, J. Lu, and G. Zhang, "A new similarity measure-based collaborative filtering approach for recommender systems," in *Foundations of Intelligent Systems*, 2014, pp. 443–452.
- [4] Q. Shambour and J. Lu, "A hybrid trust-enhanced collaborative filtering recommendation approach for personalized government-to-business e-services," *International Journal of Intelligence System.*, vol. 26, no. 9, pp. 814–843, 2011.
- [5] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Analysis of recommendation algorithms for e-commerce," in *Proceedings of the 2nd ACM Conference on Electronic Commerce*, 2000, pp. 158–167.
- [6] J. Lu, V. Behbood, P. Hao, H. Zuo, S. Xue, and G. Zhang, "Transfer learning using computational intelligence: A survey," *Knowledge-Based Systems*, vol. 80, no. 0, pp. 14 – 23, 2015.
- [7] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [8] B. Li, Q. Yang, and X. Xue, "Can movies and books collaborate?: Cross-domain collaborative filtering for sparsity reduction," in *Proceedings of the 21st International Joint Conference on Artificial Intelligence*, 2009, pp. 2052–2057.
- [9] Y. Shi, M. Larson, and A. Hanjalic, "Tags as bridges between domains: Improving recommendation with tag-induced cross-domain collaborative filtering," in *User Modeling, Adaption and Personalization*, 2011, pp. 305–316.
- [10] I. Fernández-Tobías, I. Cantador, M. Kaminskas, and F. Ricci, "Cross-domain recommender systems: A survey of the state of the art," in *2nd Spanish Conference on Information Retrieval*, 2012, pp. 3–14.
- [11] Y. Shi, M. Larson, and A. Hanjalic, "Collaborative filtering beyond the user-item matrix: A survey of the state of the art and future challenges," *ACM Comput. Surv.*, vol. 47, no. 1, pp. 1–45, 2014.
- [12] B. Li, Q. Yang, and X. Xue, "Transfer learning for collaborative filtering via a rating-matrix generative model," in *Proceedings of the 26th Annual International Conference on Machine Learning*, 2009, pp. 617–624.
- [13] B. Li, X. Zhu, R. Li, C. Zhang, X. Xue, and X. Wu, "Cross-domain collaborative filtering over time," in *Proceedings of the Twenty-Second international joint conference on Artificial Intelligence*, 2011, pp. 2293–2298.
- [14] W. Pan, E. W. Xiang, N. N. Liu, and Q. Yang, "Transfer learning in collaborative filtering for sparsity reduction," in *Twenty-Fourth AAAI Conference on Artificial Intelligence*, vol. 10, 2010, pp. 230–235.
- [15] W. Pan and Q. Yang, "Transfer learning in heterogeneous collaborative filtering domains," *Artificial Intelligence*, vol. 197, no. 0, pp. 39 – 55, 2013.
- [16] W. Wang, Z. Chen, J. Liu, Q. Qi, and Z. Zhao, "User-based collaborative filtering on cross domain by tag transfer learning," in *Proceedings of the 1st International Workshop on Cross Domain Knowledge Discovery in Web and Social Network Mining*, 2012, pp. 10–17.
- [17] B. Li, "Cross-domain collaborative filtering: A brief survey," in *Proceedings of the 23rd IEEE International Conference on Tools with Artificial Intelligence*, 2011, pp. 1085–1086.
- [18] S. J. Pan, X. Ni, J.-T. Sun, Q. Yang, and Z. Chen, "Cross-domain sentiment classification via spectral feature alignment," in *Proceedings of the 19th international conference on World wide web*, 2010, pp. 751–760.
- [19] Y. Zhen, W.-J. Li, and D.-Y. Yeung, "Tagicofi: Tag informed collaborative filtering," in *Proceedings of the Third ACM Conference on Recommender Systems*, 2009, pp. 69–76.
- [20] C. Ding and X. He, "K-means clustering via principal component analysis," in *Proceedings of the twenty-first international conference on Machine learning*, 2004, pp. 29–35.
- [21] J. L. Herlocker, J. A. Konstan, A. Borchers, and J. Riedl, "An algorithmic framework for performing collaborative filtering," in *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, 1999, pp. 230–237.
- [22] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the 10th international conference on World Wide Web*, 2001, pp. 285–295.
- [23] B. M. Sarwar, G. Karypis, J. A. Konstan, and J. T. Riedl, "Application of dimensionality reduction in recommender system – a case study," in *ACM WebKDD Workshop*, 2000, pp. 1–12.
- [24] I. Fernández-Tobías and I. Cantador, "Exploiting social tags in matrix factorization models for cross-domain collaborative filtering," in *Proceedings of the 1st Workshop on New Trends in Content-based Recommender Systems*, 2014, pp. 34–41.