Demographic Inference via Knowledge Transfer in Cross-Domain Recommender Systems

Jin Shang, Mingxuan Sun
Division of Computer Science and Engineering
Louisiana State University
Email:jshang2@lsu.edu, msun@csc.lsu.edu

Kevyn Collins-Thompson
School of Information
University of Michigan
Email:kevynct@umich.edu

***Abstract*—User demographics such as age and gender are very useful in recommender systems for applications such as personalization services and marketing, but may not always be available for individual users. Existing approaches can infer users' private demographics based on ratings, given labeled data from users who share demographics. However, such labeled information is not always available in many e-commerce services, particularly small online retailers and most media sites, for which no user registration is required. We introduce a novel probabilistic matrix factorization model for demographic transfer that enables knowledge transfer from the source domain, in which users' ratings and the corresponding demographics are available, to the target domain, in which we would like to infer unknown user demographics from ratings. Our proposed method is based on two observations: (1) Items from different but related domains may share the same latent factors such as genres and styles, and (2) Users who share similar demographics are likely to prefer similar genres across domains. This approach can align latent factors across domains that share neither common users nor common items, associating user demographics with latent factors in a unified framework. Experiments on cross-domain datasets demonstrate that the proposed method consistently improves demographic classification accuracy over existing methods.**

***Index Terms*—Demographic inference, Recommender systems, Matrix factorization**

I. Introduction

User demographics are important attributes for enriching online services that include personalization, marketing, and targeted advertisement. However, demographic information is not always available for online users, typically because either they decline to provide it [1], [2] or the online service is not designed to collect it. Instead, user interactions such as ratings, clicks, and purchases in recommender systems can sometimes provide sufficient information to infer user demographic attributes. For example, a Netflix user's preference for family-oriented and occasional children's movies may indicate that the user is a parent. Existing attempts [3]–[5] suggest that it is possible to infer user genders based on ratings with as high as 80% accuracy given labeled data from users who share demographic information in recommender systems. However, such labeled user demographic information is not always available in many e-commerce services, particularly small online retailers (e.g., an outdoor shopping site) and most media sites (e.g., Yahoo News), in which no user registration is required.

It is therefore very useful to transfer knowledge from the source domain, in which users' ratings and the corresponding demographics are available, to the target domain, in which we would like to predict demographics from ratings. Initial work in this area includes de-anonymization of movie ratings datasets [6] by matching rating patterns between the source domain IMDB database and the target domain Netflix, in which the user identities are inferred. The success of the approach is based on the assumption that a subset of common items is rated by common users in both domains. Other approaches such as [7] require that different domains have some auxiliary information such as item content features for linking and grouping users or items. However, due to privacy concerns, the sharing of user and item information may be limited in practice.

Our task is to infer demographics from ratings in a target domain by transferring the knowledge from a source domain with both ratings and user demographics. Note that this task is totally different from traditional transfer learning via user modeling, since no demographic information in the target domain is available and no entities (e.g., users, items) can be linked across domains. Demographic transfer learning under this scenario is possible based on two observations. First, when two domains such as movies and books are related, different items may share the same latent factors such as genres and styles. For example, the "The Matrix" movie and the "Neuromancer" book by William Gibson both belong to the 'cyberpunk' science fiction genre. Second, users who share similar demographics are more likely to prefer similar genres across domains. For example, multiple studies (e.g., [8], [9]) have identified group-level differences in movie and book preferences between men and women (e.g., male preference for action-adventure and sports themes vs. female preference for relationship-based themes).

Inspired by these observations, we propose a probabilistic cross-domain matrix factorization model called *Transfer Matrix Factorization* (TMF), which can infer the demographics in the target domain via transfer learning across domains where no entities can be linked. Our proposed method is based on the joint matrix factorization of two user-item rating matrices from different domains with an important twist: it characterizes a user profile as an integration of both a group-level profile that captures the preference of users within the

same demographic group, and a personal profile that captures the personal preference of each user. The group-level profile is further decomposed into the product of two components: the user membership of demographic groups and the association between demographic groups and latent factors. Since both the latent factors and the association between demographic groups and latent factors are shared across domains, the knowledge from the source domain can be used to improve the demographic inference in the target domain.

II. Related Work

Transfer learning has been used in cross-domain recommender systems to predict ratings. For example, Collective matrix factorization (CMF) [10] can be applied in cross-domain recommendation assuming that entities such as users and items are shared across domains. A recent study [7] has integrated auxiliary content information, such as user and item features, to improve recommendation accuracy. Another group of work has improved rating prediction in domains where neither items nor users are shared. Some representative methods such as [11], [12] are rating generative models based on the assumption that ratings are drawn from a shared cluster-level model. Our work focuses on a different perspective of recommender systems where we would like to infer private user traits from ratings.

The idea of transferring group-level knowledge has also been applied to cross-domain document categorization [13], [14]. Specifically, these types of approaches extend previous work [11], [15] with document class labels and transfer the association between word clusters and document classes based on nonnegative matrix factorization. In comparison with document modeling, our work models each user with both group-level preference related to demographics and individual preference, which is more suitable for recommender systems.

Our model is inspired by constrained probabilistic matrix factorization (CPMF) [16] and its extension [17], but our approach is different in the following ways. First, the CPMF models user preferences in a single domain with observed metadata such as demographics, but ours models user preferences across domains where the associations between demographics and user latent features are shared and the demographics of the target domain are unknown. Second, in CPMF the demographic indicator of a user is assumed to be an observed binary variable. However, we generalize this to be the latent probability of a user belonging to one of the demographic clusters, which follows a normal distribution.

III. Model

In this section, we introduce our TMF model for inferring demographics from rating matrices in cross-domain recommender systems.

A. Basic Concept and Notation

Throughout this paper, we denote the set of real numbers and the set of nonnegative real numbers as $\mathbb{R}$ and $\mathbb{R}_+$, respectively. The entry at the *i*-th row and *j*-th column of a matrix A is referred to as $A_{(ij)}$.

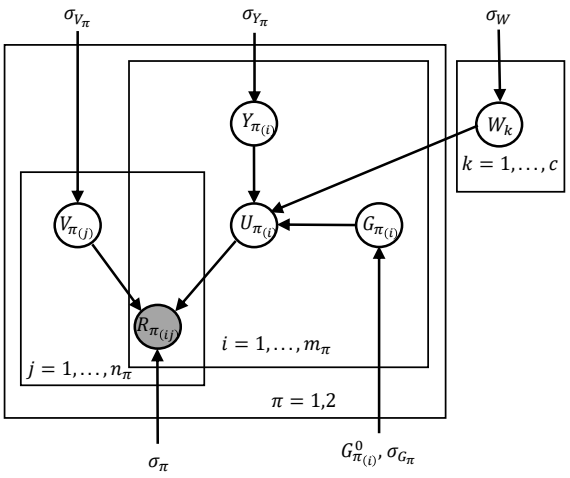


Fig. 1. The graphical model of Transfer Matrix Factorization (TMF).

Given user ratings and demographic labels in the source domain, our goal is to predict the user demographic labels from ratings in the target domain. Note that domains share neither common users nor common items. Specifically, the source domain rating matrix is denoted by $R_1 \in \mathbb{R}^{m_1 \times n_1}$ with m_1 users rating on n_1 items, and the target rating matrix is denoted by $R_2 \in \mathbb{R}^{m_2 \times n_2}$ with m_2 users rating on n_2 items. Assume there are a total of c demographic categories in both domains, e.g., $c = 2$ for binary categories (e.g., married vs. not married) and $c > 2$ for multi-class categories. Let $G_1^0 \in \mathbb{R}_+^{m_1 \times c}$ represent the true demographic label indicator matrix in the source domain. The column of matrix G_1^0 indicates the class membership, that is $G^0_{1_{(ij)}} = 1$ if the *i*-th user is in the *j*-th demographic category and $G^0_{1_{(ij)}} = 0$ otherwise.

B. Transfer Matrix Factorization (TMF) Model

In traditional Matrix Factorization (MF), the rating matrix R is approximated with the product of two low-rank matrices: $U \in \mathbb{R}^{m \times k}$ that represents the latent user feature matrix, and $V \in \mathbb{R}^{k \times n}$ that represents the latent movie feature matrix. Each entry in R is approximated by the inner product of a row vector in U and a column vector in V: $R \approx UV$.

The key innovation of our model is to associate demographic information with latent user features as:

$$U_i = Y_i + \frac{\sum_{k=1}^{c} G_{ik} W_k}{\sum_{k=1}^{c} G_{ik}}, \tag{1}$$

where $Y \in \mathbb{R}^{m \times k}$ is the user personal feature matrix, $G \in \mathbb{R}^{m \times c}$ is the latent demographic matrix, and $W \in \mathbb{R}^{c \times k}$ is the association matrix between demographic categories and latent user features. In particular, we assume that the association matrix W can be shared in different but related domains. Informally, the row of the matrix W models the effect that a user with a specific demographic label has on the prior mean of the corresponding feature vector. Therefore, users with similar demographics will have feature vectors with similar

prior distributions. The final feature vector of user i is obtained by adding offset Y_i to the mean of the prior distribution, which is important since ratings rely on not only demographics but also users' individual preferences, and usually the latter is much more crucial in predicting ratings. Without Y, the user feature matrix would only rely on demographics, meaning that, for example, users of the same gender will give the same rating score on one movie, and that is unrealistic. Our probabilistic graphical model is shown in Figure 1.

The likelihood of the observed ratings in each domain $\pi \in \{1, 2\}$ is as follows:

$$p(R_\pi|Y_\pi, V_\pi, W, G_\pi, \sigma_\pi^2) = \prod_{i=1}^{m_\pi}\prod_{j=1}^{n_\pi}[\mathcal{N}(R_{\pi_{(ij)}}\Big|\left(Y_{\pi_{(i\cdot)}}+\frac{\sum_{k=1}^{c}G_{\pi_{(ik)}}W_k}{\sum_{k=1}^{c}G_{\pi_{(ik)}}}\right)V_{\pi_{(\cdot j)}}, \sigma_\pi^2)]^{I_{\pi_{(ij)}}}, \tag{2}$$

where $\mathcal{N}(x|\mu, \sigma^2)$ is the density function of a normal distribution with mean μ and variance σ^2. The indicator function I_{ij} is equal to 1 if user i rated movie j and is equal to 0 otherwise. We also regularize all latent vectors by imposing Gaussian priors as follows: $\mathcal{N}(Y_{\pi_{(i\cdot)}}|\mathbf{0}, \sigma_{Y_\pi}^2\boldsymbol{I})$, $\mathcal{N}(V_{\pi_{(j\cdot)}}^T|\mathbf{0}, \sigma_{V_\pi}^2\boldsymbol{I})$, $\mathcal{N}(W_k|\mathbf{0}, \sigma_W^2\boldsymbol{I})$, $\mathcal{N}(G_{2_{(i\cdot)}}|\mathbf{0}, \sigma_{G_2}^2\boldsymbol{I})$, and $\mathcal{N}(G_{1_{(i\cdot)}}|G_{1_{(i\cdot)}}^0, \sigma_{G_1}^2\boldsymbol{I})$. Note that G_1^0 contains the true demographic label in the source domain and is the mean of the prior distribution of G_1.

We can obtain the *maximum a posteriori* (MAP) estimates of model parameters Y_π, V_π, G_π, and W with hyperparameters such as the prior variance σ_{Y_π} and the observation variance σ_π kept fixed by minimizing the following sum-of-squared-error objective function E:

$$\begin{aligned}E=&\frac{1}{2}\sum_{\pi=1}^{2}\sum_{i=1}^{m_\pi}\sum_{j=1}^{n_\pi}I_{\pi_{(ij)}}[R_{\pi_{(ij)}}-\left(Y_{\pi_{(i\cdot)}}+\frac{\sum_{k=1}^{c}G_{\pi_{(ik)}}W_k}{\sum_{k=1}^{c}G_{\pi_{(ik)}}}\right)V_{\pi_{(\cdot j)}}]^2\\&+\sum_{\pi=1}^{2}\lambda_{Y_\pi}\sum_{i=1}^{m_\pi}\|Y_{\pi_{(i\cdot)}}\|^2+\sum_{\pi=1}^{2}\lambda_{V_\pi}\sum_{j=1}^{n_\pi}\|V_{\pi_{(\cdot j)}}\|^2+\lambda_W\sum_{k=1}^{c}\|W_k\|^2\\&+\alpha\sum_{i=1}^{m_1}\|G_{1_{(i\cdot)}}-G_{1_{(i\cdot)}}^0\|^2+\gamma\sum_{i=1}^{m_2}\|G_{2_{(i\cdot)}}\|^2,\end{aligned} \tag{3}$$

where regularization parameters $\lambda_{Y_\pi} = \sigma_\pi^2/2\sigma_{Y_\pi}^2$, $\lambda_{V_\pi} = \sigma_\pi^2/2\sigma_{V_\pi}^2$, $\lambda_W = \sum_{\pi=1}^{2}\sigma_\pi^2/2\sigma_W^2$, $\alpha = \sigma_1^2/2\sigma_{G_1}^2$, and $\gamma = \sigma_2^2/2\sigma_{G_2}^2$.

Furthermore, since matrices G_1 and G_2 indicate the probabilities that users belong to demographic classes, we revise the objective function by adding non-negative constraints to model parameters. To make it simpler, we minimize the loss function as follows:

$$\begin{aligned}&\min_{Y_\pi, V_\pi, G_\pi, W}\|[R_1-(Y_1+G_1W)V_1]\circ I_1\|^2+\beta\|[R_2-(Y_2+G_2W)V_2]\circ I_2\|^2\\&+\sum_{\pi=1}^{2}\lambda_{Y_\pi}\|Y_\pi\|^2+\sum_{\pi=1}^{2}\lambda_{V_\pi}\|V_\pi\|^2+\lambda_W\|W\|^2+\alpha\|G_1-G_1^0\|^2+\gamma\|G_2\|^2\\&s.t.\quad\sum_{j=1}^{c}G_{\pi_{(ij)}}=1, G_\pi, V_\pi, Y_\pi, W\geq 0, \pi\in\{1,2\},\end{aligned} \tag{4}$$

where $\circ$ denotes element-wise product and β is the non-negative trade-off factor controlling the balance between the number of observations in the source and the target domains. Since G_1^0 contains the true demographic label information in the source domain, the regularization term α enforces the similarity between G_1 and the prior G_1^0 in the source domain. We will then present an efficient algorithm to learn model parameters Y_π, V_π, G_π, and W to minimize the objective function in eq. (4). The probability matrix G_2 in the target domain obtained through optimization will be used to predict user demographic labels. Specifically, the predicted demographic class label of the *i*-th user in the target domain is the index of the category with the largest probability. In addition, regularization parameters such as λ_{Y_π}, λ_{V_π}, λ_W, α, β, and γ provide significant flexibility in how the model is regularized. To determine these parameters, we consider a set of reasonable parameter values for each of them, train the model for each setting, and choose the ones that perform best on the validation data.

C. Learning Algorithm

We now present the learning algorithm to find the optimal solution to our optimization problem in eq. (4), which is achieved through the following theorem. As we use sparse matrices in the experiments, we ignore the indicator function matrix I_π in the following equations.

Theorem 1. *Updating Y_1, Y_2, V_1, V_2, G_1, G_2, W with eqs.* (5) *to* (9) *and normalizing G_1, G_2 to satisfy the equality constraints with eq.* (10) *in each iteration will monotonically decrease the objective function in eq.* (4) *until convergence.*

$$Y_\pi \leftarrow Y_\pi \circ \sqrt{\frac{[R_\pi V_\pi^T - G_\pi W V_\pi V_\pi^T]}{[Y_\pi(V_\pi V_\pi^T+\lambda)]}}, \tag{5}$$

$$V_\pi \leftarrow V_\pi \circ \sqrt{\frac{[(Y_\pi + G_\pi W)^T R_\pi]}{[((Y_\pi + G_\pi W)^T(Y_\pi + G_\pi W)+\lambda)V_\pi]}}, \tag{6}$$

$$G_1 \leftarrow G_1 \circ \sqrt{\frac{[(R_1 - Y_1V_1)V_1^T W^T + \alpha G_1^0]}{[G_1(WV_1V_1^T W^T+\alpha)]}}, \tag{7}$$

$$G_2 \leftarrow G_2 \circ \sqrt{\frac{[(R_2 - Y_2V_2)V_2^T W^T]}{[G_2(WV_2V_2^T W^T+\gamma)]}}, \tag{8}$$

$$W \leftarrow W \circ \frac{[\sqrt{\beta G_2^T(R_2V_2^T - Y_2V_2V_2^T)+G_1^T(R_1V_1^T - Y_1V_1V_1^T)}]}{[\sqrt{\beta G_2^T G_2 W V_2V_2^T + G_1^T G_1 W V_1V_1^T + \lambda W}]}, \tag{9}$$

$$G_{\pi_{(i\cdot)}} \leftarrow \frac{G_{\pi_{(i\cdot)}}}{\sum_{j=1}^{c}G_{\pi_{(ij)}}}, \tag{10}$$

where $\circ$ denotes element-wise product, $\frac{[\cdot]}{[\cdot]}$ denotes element-wise division, and $\sqrt{\cdot}$ denotes element-wise square root.

The learning algorithm for the model optimization is summarized in Algorithm 1. The proof of Theorem 1 is omitted.

Algorithm 1: Transfer Matrix Factorization (TMF) for Cross-domain Recommender Systems

Input: Source domain rating matrix R_1 and true demographic label matrix G_1^0; target domain rating matrix R_2

Output: The demographic probability matrix in the target domain G_2

begin

Initialize the matrix variables as Y_1, Y_2, V_1, V_2, W, G_1, G_2 and set parameters α, β, λ and γ. The details of the initialization method are in the experimental section.

for *iter* $\leftarrow$ *1* ***to*** *maxIter* **do**

1. update Y_1, Y_2, V_1, V_2, W, G_1, G_2 by eqs. (5) to (9).
2. normalize G_1, G_2 by eq. (10).

end

Output the matrix G_2 containing demographic labels.

end

D. Computational Complexity

We measure the computational complexity for eqs. (5) to (10) in a similar way as [18]. The computational complexity for TMF in each iteration is of order $3m_1n_1k + cm_1k + k^2m_1 + k^2n_1$ for eq. (5). In general, the latent dimension k and the number of categories c are much smaller than the size of rating matrices, that is, $k, c \ll min\{m, n\}$. Suppose $N = max\{m, n\}$, so the computational complexity is $O(N^2)$ in each iteration. Similarly, the computational complexity is $O(N^2)$ for eqs. (6) to (9) and is $O(N)$ for eq. (10) in each iteration. We assume this algorithm needs $maxIter$ iterations to converge. Therefore, multiplying these orders by $maxIter$ and then summating all the orders, we have the overall computational complexity as $O(maxIter \cdot N^2)$. Considering this is the worst case and the matrices are usually sparse in experiments, these matrix multiplications can be computed more efficiently in most cases.

IV. Experiments

We evaluate the proposed TMF approach on the union of three real-world rating datasets.

A. Datasets and Evaluation Criteria

MovieLens[1]: The MovieLens dataset contains 3952 movies, 6040 users, and about 1 million ratings (scales 1-5). Each user has more than 20 ratings. We select 3461 movies with more than 3 ratings for the experiment. There are 999K ratings and the density is 4.78%. The fraction of male users is 71.1%.

Flixster: The Flixster movie dataset is collected by Jamali et al. [19]. We randomly select users with more than 200 ratings and movies with more than 100 ratings, which results in a subset of 2608105 ratings for 3500 movies by 6000 users. The rating density is 12.43% and the fraction of males is 38.3%. Note that the skew in gender distribution is the opposite of the one in MovieLens. In the task of age prediction, we follow an arbitrary convention of setting 25 years as the threshold between 'young' and 'old'. People who are below the threshold are labeled as 'young', and otherwise are labeled as 'old'. The fraction of users with an 'old' label is 41.0%.

BookCrossing[2]: For the BookCrossing dataset, for consistency across the evaluation datasets, we normalize the rating scales from 1 to 5 and select 6461 users and 3680 books with more than 15 and 20 ratings, respectively. There are 170134 ratings and the rating density is 0.72%. The fraction of users with an 'old' label is 79.6%.

To evaluate, we withheld the ground truth labels in the target domain and measured the classification accuracy using weighted-*precision*, *recall*, and *f-score*. In addition, we measured the *f-score* in each demographic category. Note that demographic labels in target domains are only used for evaluation, and not for training.

B. Baseline Methods and Parameter Settings

Our baselines include **MF-Logistic**, which uses matrix factorization [20], [21] to decompose the rating matrices to latent vectors of size k in the source and target domains independently, and trains a logistic regressor to predict the user demographic labels given the low-rank user feature matrix in the source domain. Finally, the regressor is used to predict the user labels in the target domain. The drawback of this model is that the latent vectors in both domains may not be well aligned and in fact may represent quite different latent characteristics in source and target domains.

RMGM [11] is a rating matrix generative model based on the assumption that ratings are drawn from a shared cluster-level model. The core idea of this method is that each rating matrix R_i can be decomposed via tri-factor matrix factorization. The mixture generative model can be applied to demographic prediction with a simple modification in which the demographic label of the majorities in the source domain can be used as the predicted labels for users in the target domain. This approach transfers demographic knowledge in an unsupervised fashion.

MTrick is applied to cross-domain text document classification by Zhuang [13] based on tri-factor matrix factorization. The decomposition consists of document membership matrices, word membership matrices, and an association matrix between word clusters and document classes that is shared between source and target domains. One drawback is that the model assumes users in the same demographic cluster will give the same rating scores on the same item and ignores the individual preferences of each user. Regularization terms for MTrick were later added to improve model generalization, a baseline we call **DKT** [14].

In addition, we consider our **TMF** model with a full selection of regularization terms λ, γ, which we call **Reg-TMF**. **TMF** involves fewer regularization terms, i.e., no

[1]https://grouplens.org/datasets/movielens/

[2]http://www2.informatik.uni-freiburg.de/~cziegler/BX/

TABLE I
AVERAGE PERFORMANCE AND PER-GROUP FSCORE FOR MULTIPLE METHODS INFERRING DIFFERENT TYPES OF DEMOGRAPHICS.

Data	Metrics	Methods					
		MF-Logistic	RMGM	MTrick	DKT	TMF	Reg-TMF
F to M infer gender	Precision	0.3750	0.4672	0.5659	0.5639	0.7194	**0.7195**
	Recall	0.4015	0.4860	0.5440	0.5418	0.7304	**0.7311**
	Fscore	0.3878	0.4764	0.5547	0.5526	0.7249	**0.7253**
	Female-Fscore	0.2835	0.3554	0.3405	0.3398	**0.4767**	0.4759
	Male-Fscore	0.4457	0.5460	0.6765	0.6744	0.8083	**0.8085**
M to F infer gender	Precision	0.4278	0.5098	0.6338	0.6337	**0.6980**	0.6978
	Recall	0.6214	0.5350	**0.7692**	0.7689	0.6963	0.6961
	Fscore	0.5068	0.5221	0.6950	0.6948	**0.6971**	0.6970
	Female-Fscore	0.3315	0.6224	0.7468	0.7467	**0.7519**	0.7517
	Male-Fscore	0.4999	0.3016	0.3389	0.3387	**0.6141**	0.6140
F to B infer age	Precision	0.2843	0.4728	0.4337	0.4331	**0.6282**	0.6262
	Recall	**0.7696**	0.5461	0.6173	0.6179	0.5921	0.5898
	Fscore	0.4152	0.5068	0.5095	0.5092	**0.6096**	0.6074
	Young-Fscore	0.3310	0.3271	**0.3504**	**0.3504**	0.2545	0.2502
	Old-Fscore	0.2306	0.5667	0.4980	0.4970	**0.7524**	0.7511

regularization parameter λ, γ for latent vectors Y_π, V_π, G_2 and W. We also explore regularization parameter sensitivities. For most of the experiments, $\gamma = \lambda = 0.01$ is used for all latent variables. The trade-off parameters are $\alpha = 0.2$ and $\beta = 1$. We evaluate the objective function under different numbers of latent dimensions from 5 to 50 and choose the best latent dimension k. At the beginning, we randomly initialize Y_1, Y_2, V_1, V_2, W with non-negative values. We randomly initialize the demographic information matrices G_1 and G_2 with entries in the range of 0 to 1. The maximum number of iterations *maxIter* used in the optimization is 300.

C. Demographic Prediction Results

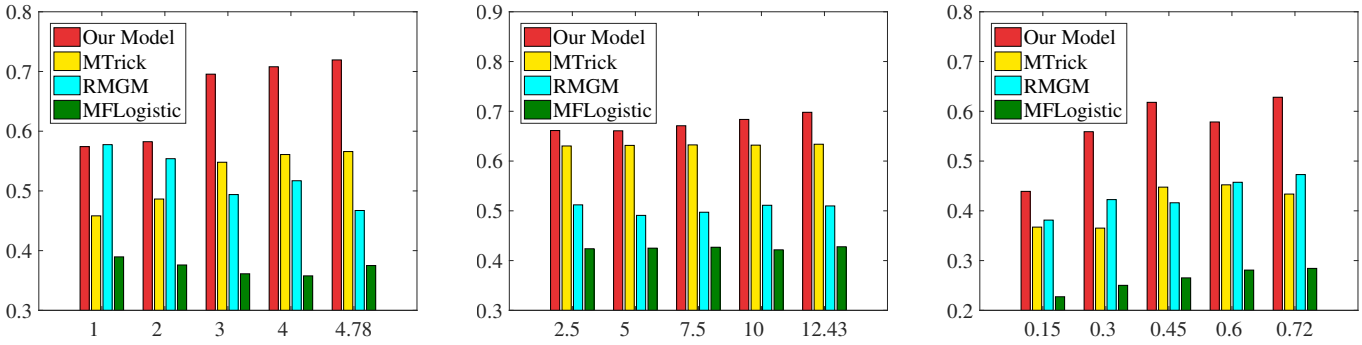
The results are reported in Table I, where F denotes the Flixster dataset, M denotes the MovieLens dataset and B denotes the BookCrossing dataset.

First, comparing all six models, we can see that our model TMF and its variation Reg-TMF consistently outperform others in all three types of demographic prediction. According to Weinsberg's [3] work, gender prediction in a single domain has reached 80% while our method achieves up to 73% across certain domains. The MF-Logistic approach performs the worst since it does not align the two domains jointly. RMGM is better than MF-Logistic but the weighted F-score is just slightly above 0.5. The reason for the poor performance is that the demographic labels are not correlated with the generation of the ratings. MTrick and TMF perform much better than others because they both correlate demographic labels with rating generation in a supervised fashion. The crucial reason for TMF achieving the highest performance is that it characterizes a user profile with both an individual preference profile and a non-demographic user preference.

For all three types of demographic prediction, the distribution of the demographic labels in the source domain is totally opposite to that in the target domain, which makes the prediction tasks very challenging. In particular, the MovieLens dataset has a majority of males while the Flixster dataset has many more female users. In age prediction, the proportion of youth and old in the source domain is also opposite to that in the target domain. As shown in Table I, the MF-Logistic approach has extremely high recall but extremely low precision when we predict ages from Flixster to the BookCrossing dataset. This is because the method has a strong tendency to predict the majority class in the source domain for most of the users when the source domain is unbalanced. A good method should balance between recall and precision. Compared to other methods, our model TMF is more robust and consistently outperforms others regardless of the difference in demographic distributions in the source and target domains.

We also evaluated the sensitivity of our model TMF with respect to different levels of rating sparsity in the target domain. In addition to the original datasets, we subsampled datasets with an additional 4 levels of density in the target domain, for each type of demographic prediction task. Specifically, due to different density levels in the original data, for the target domain datasets, we randomly dropped out some ratings from the original matrix and obtained subsets with 1%, 2%, 3% and 4% rating ratios for Movielens, subsets with 2.5%, 5.0%, 7.5% and 10.0% rating ratios for Flixster, and subsets with 0.15%, 0.30%, 0.45% and 0.60% rating ratios for BookCrossing.

As shown in Figure 2, the prediction accuracies of all models increase as the density of the target domain increases. Moreover, the figures demonstrate that our model TMF is more sensitive to the density when the target domain data is not dense enough. However, when the first level is 2.5% in Figure 2(b), all the models do not improve much further with respect to the density. This shows TMF can reach good demographic prediction performance using less data, which also means higher efficiency in practice. Finally, it is obvious that the factor models (MTrick and TMF) perform much better than the mixture model (RMGM), which indicates the advantages of factor models that integrate demographic information from users' labels.



(a) F to M infer Gender (b) M to F infer Gender (c) F to B infer Age

Fig. 2. Accuracy with respect to densities of testing data using different methods.

V. Conclusions

We introduced a novel method called **Transfer Matrix Factorization (TMF)** to solve the problem of predicting user demographics using ratings in a target domain, through knowledge transfer from the source domain, in which users' ratings and the corresponding demographics are available. Our main contributions are: (1) Our model explores effectively the correlation between demographics and ratings across different domains that share neither common users nor common items. (2) Extensive experiments using real-world datasets demonstrate that our model can achieve higher classification accuracy. Our approach can be used as an analytical tool to assess the privacy impact for users of providing specific kinds of user information in one or more source domains, in the context of the existence of complementary data in a target domain. In future work, we would like to use our results to investigate effective strategies for operations such as obfuscating ratings that better protect user privacy.

VI. Acknowledgement

This work was supported in part by the Louisiana Board of Regents under Grant LEQSF(2017-20)-RD-A-29.

References

[1] N. F. Awad and M. Krishnan, "The personalization privacy paradox: an empirical evaluation of information transparency and the willingness to be profiled online for personalization," *MIS Quarterly*, vol. 30, no. 1, pp. 13–28, 2006.

[2] J. Calandrino, A. Kilzer, A. Narayanan, E. Felten, and V. Shmatikov, ""you might also like:" privacy risks of collaborative filtering," in *Proc. of the IEEE Symposium on Security and Privacy*, 2011, pp. 231–246.

[3] U. Weinsberg, S. Bhagat, S. Ioannidis, and N. Taft, "Blurme: inferring and obfuscating user gender based on ratings," in *Proc. of the ACM Conference on Recommender Systems (RECSYS)*, 2012, pp. 195–202.

[4] S. Bhagat, U. Weinsberg, S. Ioannidis, and N. Taft, "Recommending with an agenda: active learning of private attributes using matrix factorization," in *Proc. of the ACM Conference on Recommender Systems (RECSYS)*, Oct. 2014, pp. 65–72.

[5] M. Sun, C. Li, and H. Zha, "Inferring private demographics of new users in recommender systems," in *Proc. of the ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWIM)*, Nov. 2017, pp. 237–244.

[6] A. Narayanan and V. Shmatikov, "Robust de-anonymization of large sparse datasets," in *Proc. of the IEEE Symposium on Security and Privacy*, 2008, pp. 111–125.

[7] A. M. Elkahky, Y. Song, and X. He, "A multi-view deep learning approach for cross domain user modeling in recommendation systems," in *Proc. of the International World Wide Web Conference (WWW)*, 2015, pp. 278–288.

[8] M. B. Oliver, J. B. Weaver, III, and S. L. Sargent, "An examination of factors related to sex differences in enjoyment of sad films," *Journal of Broadcasting & Electronic Media*, vol. 44, no. 2, pp. 282–300, 2000.

[9] M. Thelwall, "Reader and author gender and genre in goodreads," *Journal of Librarianship and Information Science*, p. 0961000617709061, 2017.

[10] A. P. Singh and G. J. Gordon, "Relational learning via collective matrix factorization," in *Proc. of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2008, pp. 650–658.

[11] B. Li, Q. Yang, and X. Xue, "Transfer learning for collaborative filtering via a rating-matrix generative model," in *Proc. of the International Conference on Machine Learning (ICML)*, 2009, pp. 617–624.

[12] T. Iwata and K. Takeuchi, "Cross-domain recommendation without shared users or items by sharing latent vector distributions," in *Proc. of the International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2015, pp. 379–387.

[13] F. Zhuang, P. Luo, H. Xiong, Z. Shi, Q. He, and Y. Xiong, "Exploiting associations between word clusters and document classes for cross-domain text categorization," in *Proc. of the SIAM International Conference on Data Mining (SDM)*, 2010, pp. 13–24.

[14] H. Wang, H. Huang, F. Nie, and C. Ding, "Cross-language web page classification via dual knowledge transfer using nonnegative matrix tri-factorization," in *Proc. of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2011, pp. 933–942.

[15] B. Li, Q. Yang, and X. Xue, "Can movies and books collaborate?: cross-domain collaborative filtering for sparsity reduction," in *Proc. of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2009, pp. 2052–2057.

[16] R. Salakhutdinov and A. Mnih, "Probabilistic matrix factorization," in *Proc. of the Annual Conference on Neural Information Processing Systems (NIPS)*, 2008.

[17] C. Severinski, "Augmenting probabilistic matrix factorization models for rare users," Ph.D. dissertation, University of Toronto (Canada), 2016.

[18] C. Ding, T. Li, and M. I. Jordan, "Convex and semi-nonnegative matrix factorizations," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 32, no. 1, pp. 45–55, 2010.

[19] M. Jamali and M. Ester, "A matrix factorization technique with trust propagation for recommendation in social networks," in *Proc. of the ACM Conference on Recommender Systems (RECSYS)*, 2010, pp. 135–142.

[20] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, 2009.

[21] J. Lee, M. Sun, and G. Lebanon, "PREA: Personalized recommendation algorithms toolkit," *Journal of Machine Learning Research (JMLR)*, vol. 13, no. 1, pp. 2699–2703, Sept. 2012.