Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)

Capturing Semantic Correlation for Item Recommendation in Tagging Systems

Chaochao Chen, Xiaolin Zheng
College of Computer Science
Zhejiang University
{zjuccc,xlzheng}@zju.edu.cn

Yan Wang
Department of Computing
Macquarie University
yan.wang@mq.edu.au

Fuxing Hong, Deren Chen
College of Computer Science
Zhejiang University
{cstur4,drc}@zju.edu.cn

Abstract

The popularity of tagging systems provides a great opportunity to improve the performance of item recommendation. Although existing approaches use topic modeling to mine the semantic information of items by grouping the tags labelled for items, they overlook an important property that tags link users and items as a bridge. Thus these methods cannot deal with the *data sparsity without commonly rated items* (*DS-WO-CRI*) problem, limiting their recommendation performance. Towards solving this challenging problem, we propose a novel tag and rating based collaborative filtering (CF) model for item recommendation, which first uses topic modeling to mine the semantic information of tags for each user and for each item respectively, and then incorporates the semantic information into matrix factorization to factorize rating information and to capture the bridging feature of tags and ratings between users and items. As a result, our model captures the semantic correlation between users and items, and is able to greatly improve recommendation performance, especially in *DS-WO-CRI* situations. Experiments conducted on two popular real-world datasets demonstrate that our proposed model significantly outperforms the conventional CF approach, the state-of-the-art social relation based CF approach, and the state-of-the-art topic modeling based CF approaches in terms of both precision and recall, and it is an effective approach to the *DS-WO-CRI* problem.

Introduction

In recent years, tagging systems, such as *Delicious* (social bookmarking), *Last.fm* (social music), *Flickr* (photo sharing), and *YouTube* (video sharing), provide effective ways for users to organize, manage, share, and search various kinds of items (resources). For example, one may tag Lady Gaga with "pop" and "female vocalist" when he listens to her music on *Last.fm*. These valuable tags, which appear along with the tagging and rating behaviors, strongly suggest the need to use such information to provide personalized recommendation services (Zheng and Li, 2011).

The increasing popularity of tagging systems has promoted the development of recommender systems, especially collaborative filtering (CF) approaches, in tagging systems. So far, two main types of CF on tagging systems exist: *tag recommendation* (Wang, Chen, and Li, 2013; Fang et al., 2015), which aims to recommend appropriate tags for items, and *tag-based item recommendation* (Zhou et al., 2009; Xu et al., 2011; Zhou et al., 2010), which focuses on recommending similar users or items to the target user based on tag and other information (e.g., rating).

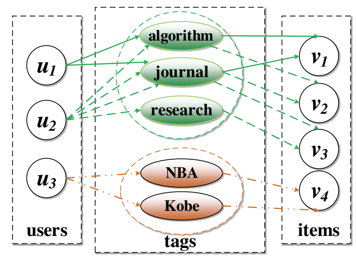


Figure 1: An example of tagging system. There is a rating behind each tagging behavior, and we omit the ratings for conciseness.

Currently, a trend in the literature is the use of topic modeling in CF to handle tag information (Agarwal and Chen, 2010; Wang and Blei, 2011; Purushotham, Liu, and Kuo, 2012; Wang, Chen, and Li, 2013; Chen et al., 2014). For example, Wang and Blei (2011) proposed a collaborative topic regression (CTR) model that can be used for tag-based item recommendation. Chen et al. (2014) proposed another item recommendation method that combined CTR with social matrix factorization (Ma et al., 2011) to make a better prediction. However, the existing approaches just associate tags with users or items, and overlook an important property of tags that tags link users and items as a bridge, as what ratings do. But tags can reflect the semantic correlation between users and items, which ratings cannot do.

When a user has tagged some items, these tags clearly represent the user's preference for the items. The more frequently a tag has been used by a user, the more likely this user is interested in the group of items that can be labeled by this tag (Zheng and Li, 2011). Similarly, the more frequently a tag has been given to an item by users, the more likely this item matches the tag. Thus, tags contain the semantic information of both users and items, not just one of them.

A motivating example. Figure 1 depicts an example of a tagging system, which consists of three users (u_1, u_2, and u_3), four items (v_1, v_2, v_3, and v_4), and five tags ("algorithm", "journal", "research", "NBA", and "Kobe").

In this example, user u_1 labelled item v_1, and user u_2 labelled items v_2 and v_3. Thus users u_1 and u_2 have no commonly rated items. We term this situation *data sparsity without commonly rated items* (*DS-WO-CRI*). DS-WO-CRI is a typical subset of the standard data sparsity problem (i.e., the known user-item actions are rare comparing with all the user-item pairs). In *DS-WO-CRI* situations, the existing CF approaches, e.g., PMF and CTR, cannot essentially recommend item v_1 to user u_2 because they cannot capture any relation between them. However, a good recommender system should recommend items v_2 and v_3 to user u_1 and recommend item v_1 to user u_2, because in this example, users u_1 and u_2 are probably researchers on algorithms, and items v_1, v_2, and v_3 are probably related to algorithms.

The existing studies have shown that a user's action on an item, e.g., tag and rate, has already indicated this user's interests in this item, regardless of how the user rated this item (Koren, 2008; Koren, Bell, and Volinsky, 2009). In other words, a user implicitly expresses his preferences by voicing his opinion through tagging and voting a (high or low) rating (Koren, 2008). Thus, a user and the items that he has tagged and rated tend to share similar latent features, and we term it *implicit preference* in this paper. In the above example, in particular, user u_2 gives item v_3 tags "journal" and "research" and this not only shows that user u_2 is likely to be a researcher, but also indicates that item v_3 is likely to be a research journal or something related to it. Thus, semantically, the latent features of user u_2 and item v_3 should be similar to some extent.

However, the existing approaches fail to capture the semantic correlation between users and items, and thus their recommendation performance is limited, especially in *DS-WO-CRI* situations.

Our proposal. To deal with the above mentioned *DS-WO-CRI* problem, in this paper, we propose a novel CF model. We first use topic modeling to mine the semantic information of tags for each user and for each item respectively, and then incorporate the semantic information into matrix factorization to factorize rating information and capture the bridging feature of tags and ratings between users and items (i.e., implicit preference). As a result, our model captures the semantic correlation between users and items, and can recommend an item to a user if they have similar semantic information, though they are in a *DS-WO-CRI* situation.

Contributions. The main contributions of our work are summarized as follows: (1) We first point out the important feature of tags, namely, they link users and items as a bridge, outlining the semantic correlation between users and items, and then we illustrate that utilizing this feature can help deal with the *DS-WO-CRI* problem. To the best of our knowledge, this is the first study in the literature to identify this problem; (2) We propose a novel tag and rating based CF model, which can capture the semantic correlation between users and items and thus can greatly improve the recommendation performance, especially in *DS-WO-CRI* situations. We also propose our parameter learning method based on coordinate ascent algorithm. To the best of our knowledge, this study is the first attempt in the literature to capture the semantic correlation between users and items provided by tags in tag-based item recommendation; (3) Experiments conducted on two popular real-world datasets demonstrate that our proposed model significantly outperforms the state-of-the-art approaches in terms of both precision and recall. The experiments also demonstrate that our proposed model is an effective approach to the *DS-WO-CRI* problem.

Related Work

In this section, we review the existing item recommendation methods in tagging systems in three groups, including (1) the conventional CF approaches, (2) the topic modeling based CF approaches, and (3) the social relation based CF approaches.

Based on the existing research (Shi, Larson, and Hanjalic, 2014), the conventional CF approaches, which only use user-item rating information to make recommendations, are in two major categories: the memory-based CF (Deshpande and Karypis, 2004) and model-based CF (Koren, Bell, and Volinsky, 2009; Koren, 2008; Zhou et al., 2009; Xu et al., 2015), both of which can be used to make recommendations in tagging systems.

The conventional CF approaches, e.g., TagRec (Zhou et al., 2009), cannot capture content (e.g., tag) information. Thus, some hybrid approaches were proposed to combine content-based approach and CF to do item recommendation (Melville, Mooney, and Nagarajan, 2002) and (Basilico and Hofmann, 2004). However, these methods take content simply as a vector of words, and thus cannot mine their semantic information. To take advantage of semantic information provided by content (e.g., tag), researchers use topic modeling to improve recommendation performance. Agarwal et al., proposed fLDA (Agarwal and Chen, 2010), which combines RLFM (Agarwal and Chen, 2009) with latent Dirichlet allocation (LDA) by assigning item factors through a richer prior learnt from LDA. Both RLFM and fLDA incorporate additional covariates that are obtained from additional meta-feature information, e.g., user age and item category, which, however, are out of the scope of this paper. Later on, Wang et al. (Wang and Blei, 2011) proposed CTR to combine probabilistic matrix factorization (PMF) (Mnih and Salakhutdinov, 2007) with LDA (Blei, Ng, and Jordan, 2003) to make recommendations. It has been proven in (Wang and Blei, 2011) that CTR performs better than fLDA in a similar setting, since fLDA largely ignores the other users' ratings.

Moreover, social information between users and between items is considered valuable to improve recommendation performance (Chen et al., 2013). First, user social information is incorporated into conventional CF models (Jamali and Ester, 2010; Ma et al., 2011). For example, Ma et al. (2011) proposed Soreg to constrain the difference between the user latent factors of connected users. Second, neighbor user or item social information is incorporated into topic modeling based CF models (e.g., CTR) to further improve recommendation performance. For example, Purushotham, Liu, and Kuo (2012) and Chen et al. (2014) proposed two models (i.e., CTR-SMF and CTR-SMF2) to incorporate user social network into CTR to further improve item recommendation performance. Wang, Chen, and Li (2013) proposed a model to incorporate item social relationship into CTR to further improve tag recommendation performance in social tagging systems.

However, all the above approaches not only overlook the semantic information between both users and items embedded in tags, but also neglect the bridging feature of tags and ratings between users and items (i.e., implicit preference). Therefore, they cannot capture the semantic correlation between users and items, and suffer from *DS-WO-CRI* problem. To overcome these shortcomings, in this paper, we propose a novel tag and rating based CF model, which can capture the semantic correlation between users and items. Hence, our model can help deal with the *DS-WO-CRI* problem and improve recommendation performance.

The Proposed Model–TRCF

In this section, we present a novel tag and rating based CF (TRCF) model. We first formalize the tag-based item recommendation problem and define notations. Then, we present TRCF, which is a hierarchical Bayesian model. Finally, we propose our parameter learning method based on coordinate ascent algorithm.

Preliminaries

Assume that we have a set of users $\mathbb{U} = \{u_1, ..., u_I\}$, who have labelled a set of items $\mathbb{V} = \{v_1, ..., v_J\}$ with a set of tags $\mathbb{T} = \{t_1, ..., t_N\}$ and a set of ratings $\mathbb{R} = \{R_1, ..., R_O\}$, where *I*, *J*, *N*, and *O* denote the numbers of users, items, tags, and ratings, respectively. Each *user-item-tag-rating* (U-I-T-R) observe data is a 4-tuple (u_i, v_j, T_{ij}, R_{ij}), where $u_i \in \mathbb{U}$, $v_j \in \mathbb{V}$, T_{ij} is a set of tags that user u_i gives to item v_j, and $T_{ij} \subseteq \mathbb{T}$. R_{ij} is the rating that user u_i gives to item v_j based on the extent to which he likes the item and tags it at the same time; however, the user-item (U-I) rating set $\mathbb{R}$ is typically of integers, e.g., in the range [1, 5] in *MovieLens*. Let $U \in \mathbb{R}^{K\times I}$ denote the latent user feature matrices, where the column vector U_i represents the *K*-dimensional user-specific latent feature vector of user u_i. Let $V \in \mathbb{R}^{K\times J}$ denote the latent item feature matrices, where the column vector V_j represents the *K*-dimensional item-specific latent feature vector of item v_j.

For tag and rating based item recommendation, given the existing U-I-T-R 4-tuples, our goal is to predict the unknown rating from a user u_i to an item v_j.

Tag and Rating based Collaborative Filtering

TRCF is a novel hierarchical Bayesian model, and its graphical model is shown in Figure 2, where N_u and N_v denote the number of tags for user u_i and for item v_j, respectively. TRCF first groups the tags for each user and for each item respectively, and then it uses latent Dirichlet allocation (LDA) to mine the semantic information of tags for each user and each item respectively (plotted in red in Figure 2). Finally it incorporates these semantic information into matrix factorization to factorize rating information (plotted in purple in Figure 2) and capture the implicit preference provided by tags and ratings (plotted in blue in Figure 2).

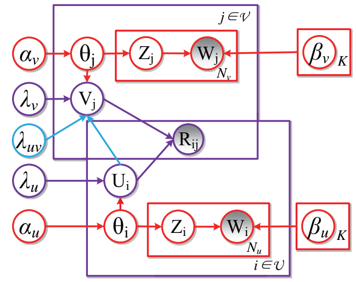


Figure 2: Graphical model of TRCF. The LDA part is plotted in red, the implicit preference part is shown in blue, and the PMF part is plotted in purple.

TRCF performs LDA on both the user side and the item side, and thus can capture the semantic information for both users and items, not just items as the existing works do. In addition, in TRCF, when a user and an item are linked by tags and ratings, their latent features are similar to each other to some extent, which is referred to as implicit preference. In contrast, the existing topic modeling based CF approaches, e.g., CTR, CTR-SFM, and CTR-SMF2, assume users and items are independent, and neglect the bridge feature of tags and ratings between users and items. Thus, TRCF can capture the semantic correlation between users and items, and is able to deal with the *DS-WO-CRI* problem. Assuming there are *K* topics for both users and items, the generative process of TRCF works as follows:

1. **Mining semantic information of tags for users**. For each user u_i,
   (a) Draw topic proportions $\theta_i \sim Dirichlet(\alpha_u)$;
   (b) Draw user latent vector as $U_i \sim \mathcal{N}(\theta_i, \lambda_u^{-1} I_K)$;
   (c) For each tag w_{in_u} of user u_i,
      i. Draw topic assignment $z_{in_u} \sim Mult(\theta_i)$;
      ii. Draw tag $w_{in_u} \sim Mult(\beta_{z_{in_u}})$;
2. **Mining semantic information of tags for items, and capturing implicit preference between users and items**. For each item v_j,
   (a) Draw topic proportions $\theta_j \sim Dirichlet(\alpha_v)$;
   (b) Draw item latent vector as $V_j \sim \mathcal{N}(\theta_j, \lambda_v^{-1} I_K) \times \prod_i I_{ij}^R \mathcal{N}(U_i, \lambda_{uv}^{-1} I_K)$;
   (c) For each tag w_{jn_v} of user v_j,
      i. Draw topic assignment $z_{jn_v} \sim Mult(\theta_j)$;
      ii. Draw tag $w_{jn_v} \sim Mult(\beta_{z_{jn_v}})$;
3. **Drawing the rating**. For each user-item pair (*i*, *j*),

$$R_{ij} \sim \mathcal{N}({U_i}^T V_j, c_{ij}^{-1}).$$

In the above generative process, $\mathcal{N}(x|\mu, \sigma^2)$ is a Gaussian distribution with a mean μ and a variance σ^2, and I_K is an identity matrix with K rows and K columns. I_{ij}^R is an indicator function the value of which equal to 1 if user u_i

rated item v_j, 0 otherwise. C is a rating confidence matrix with element c_{ij} denotes the rating confidence. Please refer to (Wang and Blei, 2011) for more details.

The parameter λ_u balances the contribution of user semantic information provided tags and rating information to the model performance. Similarly, the parameter λ_v balances the contribution of item semantic information provided by tags and rating information to the recommendation performance. The parameter λ_{uv} balances the contribution of implicit preference on model performance, i.e., the degree of the latent feature similarity of a user and an item linked by a rating and tags.

The conditional distribution of the observed ratings can be formalized as

$$p\left(R|U,V,C\right)=\prod_i\prod_j\mathcal{N}(R_{ij}|U_i^TV_j,c_{ij}).$$

The user and latent vectors U_i and V_j are generated in a similar way to CTR, which can be formalized as

$$p\left(U|\lambda_u\right)\sim\prod_i\mathcal{N}(\theta_i,\lambda_u^{-1}I_K),$$

$$p\left(V|U,\lambda_v,\lambda_{uv}\right)\sim\prod_j\mathcal{N}(\theta_j,\lambda_v^{-1}I_K)\times\prod_i I_{ij}^R\mathcal{N}(U_i,\lambda_{uv}^{-1}I_K).$$

Given the U-I-T-R information, by using Bayesian inference, we can obtain the following equation for the posterior probability of latent feature vectors of TRCF:

$$\begin{aligned}&p\left(U,V|R,C,\lambda_u,\lambda_v,\lambda_{uv}\right)\\ &\propto p\left(R|U,V,C\right)p\left(U|\lambda_u\right)p\left(V|U,\lambda_v,\lambda_{uv}\right).\end{aligned}\tag{1}$$

Parameter Learning of TRCF

Given topic parameters β_u and β_v, computing the full posterior of U_i, V_j, θ_i, and θ_j directly is intractable. We use coordinate ascent algorithm to learn the maximum a posteriori estimates. Maximizing the posterior over the two latent features with fixed hyper-parameters in Equation (1) is equivalent to maximizing the following complete log likelihood of U, V, $\theta_{1:I}$, $\theta_{1:J}$ and R, given λ_u and λ_v:

$$\begin{aligned}L=&-\frac{\lambda_u}{2}\sum_i\left(U_i-\theta_i\right)^T\left(U_i-\theta_i\right)\\ &-\frac{\lambda_v}{2}\sum_j\left(V_j-\theta_j\right)^T\left(V_j-\theta_j\right)\\ &-\sum_{ij}\frac{c_{ij}}{2}\left(R_{ij}-{U_i}^TV_j\right)^2\\ &+\sum_i\sum_{n_u}\log\left(\sum_k\theta_{ik}\beta_{k,w_{in_u}}\right)\\ &+\sum_j\sum_{n_v}\log\left(\sum_k\theta_{jk}\beta_{k,w_{jn_v}}\right)\\ &-\frac{\lambda_{uv}}{2}I_{ij}^R\sum_{ij}\left(U_i-V_j\right)^T(U_i-V_j).\end{aligned}\tag{2}$$

We omit a constant and set the Dirichlet priors $\alpha_u=\alpha_v=1$. This function can be optimized by using coordinate ascent. That is, we fix β_u and β_v, and iteratively optimize the MF variables U_i, V_j and the topic proportions θ_i and θ_j. Specifically, we first update U_i and V_j, given the current estimate of θ_i,θ_j. We take the gradient of L in Equation (2) with respect to U_i and V_j, and set it to zero,

$$\frac{\partial L}{\partial U_i}=0,\frac{\partial L}{\partial V_j}=0.\tag{3}$$

Solving the above equations will lead to the following update equation,

$$\begin{aligned}U_i\leftarrow&\left(VC_iV^T+\lambda_uI_K+\lambda_{uv}\sum_j I_{ij}^RI_K\right)^{-1}\\ &(VC_iR_i+\lambda_u\theta_i+\lambda_{uv}\sum_j I_{ij}^RV_j),\\ V_j\leftarrow&\left(UC_jU^T+\lambda_vI_K+\lambda_{uv}\sum_i I_{ij}^RI_K\right)^{-1}\\ &(UC_jR_j+\lambda_v\theta_j+\lambda_{uv}\sum_i I_{ij}^RU_i),\end{aligned}\tag{4}$$

where C_i is a diagonal matrix with $c_{ij}, j=1,...,J$ as its diagonal elements, and $R_i={R_{ij}}_{j=1}^J$ for user u_i. For item v_j, C_j and R_j are similarly defined.

Equation (4) shows how parameters λ_u, λ_v, and λ_{uv} affect the user latent feature and the item latent feature. A bigger λ_u corresponds to a bigger proportion of the user latent feature from the user tags rather than the rating information. Similarly, a bigger λ_v indicates a bigger proportion of the item latent feature from the item tags, rather than the rating information. Also, a bigger λ_{uv} means a stronger constraint that the paired user and item linked by tags and ratings should have a similar latent feature, i.e., implicit preference. From Equation (4), we can see that probabilistic matrix factorization (PMF) and collaborative topic regression (CTR) are all special cases of TRCF.

Then, we update the topic proportions θ_i and θ_j given the current MF variables U_i and V_j. For θ_i, we first define $q\left(z_{in_u}=k\right)=\Phi_{in_uk}$, and then separate the users that contain θ_i and apply Jensen's inequality,

$$\begin{aligned}L(\theta_i)\geq&-\frac{\lambda_u}{2}(U_i-\theta_i)^T\left(U_i-\theta_i\right)\\ &+\sum_{n_u}\sum_k\Phi_{in_uk}(log\theta_{ik}\beta_{k,w_{in_u}}-log\Phi_{in_uk})\\ &=L(\theta_i,\Phi_i).\end{aligned}$$

Here, $\Phi_i={\Phi_{in_uk}}_{n_u=1,k=1}^{N_u\times K}$. Obviously $L(\theta_i,\Phi_i)$ is a tight lower bound of $L(\theta_i)$, and we can use projection gradient (Bertsekas, 1995) to optimize θ_i. The optimal Φ_{in_uk} is $\Phi_{in_uk}\propto\theta_{ik}\beta_{k,w_{in_u}}$. For θ_j, it is similarly updated.

As for β_u, we update the same M-step for topics as in LDA (Blei, Ng, and Jordan, 2003),

$$\beta_{kw_i}\propto\sum_i\sum_{n_u}\Phi_{in_uk}1[w_{in_u}=w].$$

For β_v [1], it is similarly updated. After the optimal param-

[1] When using TF-CTR, a useful tactic is to fuse the user tags and item tags as the input of LDA, which ensures that users and items have the same semantic information in each element of *K*, that is, to make $\beta_u=\beta_v$.

latex表格

Dataset	users	items	tags	user-tags-items	ratings
Delicious	1,867	69,226	53,388	437,593	104,799
Lastfm	1,892	17,632	11,946	186,479	92,834

Table 1: Dataset description

eters U^*, V^*, ${\theta_{1:I}}^*$, ${\theta_{1:J}}^*$, ${\beta_u}^*$, and ${\beta_v}^*$ have been learned, our model can predict ratings:

$$R_{ij}{}^* \approx (U_i^*)^T V_j^*.$$

Experiments and Analysis

In this section, we introduce the experiments conducted on two popular real-world datasets, which aim to answer the following questions: (1) How does our model perform comparing the state-of-the-art approaches? (2) How does our approach deal with the *DS-WO-CRI* problem? (3) How do parameters λ_u, λ_v, and λ_{uv} affect the performance of TRCF?

Datasets

We use two real-world datasets in our experiments: hetrec2011-delicious-2k (*Delicious*) and hetrec2011-lastfm-2k (*Lastfm*) (Cantador, Brusilovsky, and Kuflik, 2011). Both datasets have been widely used to conduct experiments in tagging systems (Bellogín, Cantador, and Castells, 2013), and they are described in Table 1.

For each of the two datasets, we consider a user rating for an item as '1' if the user has bookmarked (or listened) the item; otherwise, the user rating for the item is '0'.

In our experiments, we split each dataset into three parts — a training dataset (80%), a held-out validation dataset (10%), and a test dataset (10%). We train our model on the training dataset, obtain the optimal parameters on the validation dataset, and evaluate our model on the test dataset.

Comparison and Evaluation

As stated in related works, there are many kinds of recommendation approaches, e.g., memory-based approach and hybrid approach. Here, we compare the proposed TRCF with the following three kinds of state-of-the-art approaches, i.e., the conventional CF approach, the social relation based CF approach, and the topic modeling based approach:

SVD++ (Koren, 2008) is a classic conventional CF approach that only uses U-I rating information.

Soreg (Ma et al., 2011) is a state-of-the-art social relation based CF approach, which uses U-I rating and user social information.

CTR (Wang and Blei, 2011) is a state-of-the-art topic modeling based CF approach, which uses U-I-T-R information similar to the 4-tuple used in TRCF.

CTR-SMF (Purushotham, Liu, and Kuo, 2012) combines user social matrix factorization with CTR. It incorporates additional user social information additional to the U-I-T-R 4-tuple used in TRCF.

CTR-SMF2 (Chen et al., 2014) improves CTR-SMF, and it also incorporates user social information additional to the U-I-T-R 4-tuple used in TRCF.

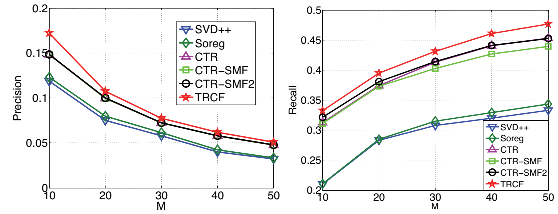


Figure 3: *Precision* and *Recall* comparison with different *M* and the best parameters of each method on *Delicious*.

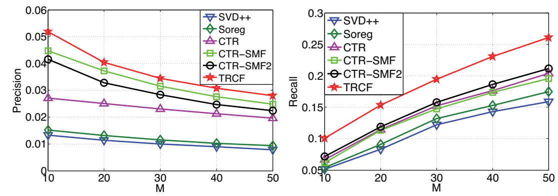


Figure 4: *Precision* and *Recall* comparison with different *M* and the best parameters of each method on *Lastfm*.

Precision and *Recall* have been widely used as the metrics to evaluate recommendation performance (Herlocker et al., 2004). Thus, we use both *Precision* and *Recall* to evaluate the recommendation performance, and note that the same way of computing recall is also used in CTR, CTR-SMF, and CTR-SMF2. For each user, *Precision* and *Recall* are defined as follows:

$$Precision@M = \frac{\#\, items\, the\, user\, likes\, in\, Top\; M}{M},$$

$$Recall@M = \frac{\#\, items\, the\, user\, likes\, in\, Top\; M}{\#\, total\, items\, the\, user\, likes},$$

where *M* is the number of returned items. We compute the average of all the items' precision and recall in the test dataset as the final result.

Performance Comparison and Analysis

During the comparison, we have used the best parameters for SVD++, CTR, and CTR-SMF that are set in CTR-SMF (Purushotham, Liu, and Kuo, 2012), which uses the same datasets. For each of Soreg, CTR-SMF2, and our model, we have used grid search to obtain the best parameters.

Results: Figures 3 and 4 show the overall performance for each recommendation approach on *Delicious* dataset and *Lastfm* dataset, in which we set *M* = 10, 20, 30, 40, 50 and fix the parameters of each approach to the best values. The results show that the conventional CF approach (i.e., SVD++) and the social relation based CF approach (i.e., Soreg) have the similar performance. The three topic modeling based CF approaches (i.e., CTR, CTR-SMF, and CTR-SMF2) significantly outperform SVD++ and Soreg, and also have similar performance, which shows the importance of tag information in recommendations.

Our proposed method, TRCF, significantly outperforms each of SVD++, Soreg, CTR, CTR-SMF, and CTR-SMF2

sub-dataset name	*LC1*	*LC2*	*LC3*	*LC4*
DS-WO-CRI degree	20%	40%	60%	80%
users	1,837	1,709	1,718	1,706
items	11,584	8,018	7,650	7,431

Table 2: Statistics of each *DS-WO-CRI* sub-dataset

on the two datasets in terms of different *M*. Specifically, on average, TRCF improves SVD++, Soreg, CTR, CTR-SMF, and CTR-SMF2 by 46.75%, 39.74%, 8.62%, 8.78%, and 8.65%, in terms of precision, and by 44.99%, 42.40%, 5.23%, 7.27%, and 4.21%, in terms of recall, on the *Delicious* dataset. On average, TRCF improves SVD++, Soreg, CTR, CTR-SMF, and CTR-SMF2 by 259.80%, 210.27%, 57.96%, 11.64%, and 23.85%, in terms of precision, and by 73.03%, 60.39%, 34.18%, 39.04%, and 28.12%, in terms of recall, on the *Lastfm* dataset.

Analysis and summary: The comparison demonstrates the effectiveness of our proposed method which captures the semantic correlation between users and items. Experimental results also indicate that though user social information (e.g., adopted in Soreg, CTR-SMF, and CTR-SMF2) can improve recommendation performance, considering user and item semantic correlation is a more effective way to improve item recommendation performance.

DS-WO-CRI Experiments

All four topic modeling based CF approaches, including TRCF, can improve recommendation performance by capturing the semantic information of items. To study their capability of handing the *DS-WO-CRI* problem, we conduct the following experiments.

We first randomly filter the original *Lastfm* datasets into four sub-datasets based on the degree of *DS-WO-CRI*, and each sub-dataset is described in Table . The degree of *DS-WO-CRI* is defined as follows:

$$x\% = \frac{\#\, users\, without\, commonly\, rated\, items}{\#\, total\, users}.$$

We then conduct comparison experiments on each of the sub-dataset. Figure 5 shows that our model always achieves the best performance under different *DS-WO-CRI* degrees. The average improvements of our model over other three topic modeling based approaches on *LC1*, *LC2*, *LC3*, *LC4* are 49.67%, 75.81%, 345.77%, and 428.97% respectively, in terms of precision, and are 33.94%, 65.00%, 383.63%, and 458.00% respectively, in term of recall. The experimental results show that a greater degree of *DS-WO-CRI* corresponds to a higher improvement of our model against other models on each of the two datasets.

Summary: The *DS-WO-CRI* experiments demonstrate the effectiveness of our method in dealing with the *DS-WO-CRI* problem: a greater degree of *DS-WO-CRI* corresponds to a higher improvement of our model against other models. This is due to the ability of our method to capture the semantic correlation between users and items provided by tags.

Parameter Effect Analysis

Figure 6(a) shows the effect of λ_u and λ_v in Eq.(2) on the performance of TRCF by fixing $\lambda_{uv} = 0$ on *Lastfm*. We can see that TRCF achieves the best performance when $\lambda_u = \lambda_v = 10$, which means that both user and item semantic information contribute significantly to model performance. Figure 6(b) shows how the performance of TRCF is affected by parameter λ_{uv} in Eq.(2) on each *DS-WO-CRI* sub-dataset with the best λ_u and λ_v. As we can see, the performance of TRCF first increases with λ_{uv} and then starts to decrease at a certain threshold. The best value of λ_{uv} on *LC1*, *LC2*, *LC3*, and *LC4* is 0.0001, 0.01, 0,01, and 0.1, respectively. These results demonstrate that a greater degree of *DS-WO-CRI* corresponds to a greater λ_{uv}. In other words, the feature of bridging users and items by tags and ratings (i.e., the implicit preference) is more important when the *DS-WO-CRI* problem is severer, which explains why our model performs well in severe *DS-WO-CRI* situations.

Conclusions

In this paper, we first introduce the *DS-WO-CRI* problem that exists in item recommendation in real tagging systems. Then, we present a novel tag and rating based CF model to deal with this problem. The proposed model uses topic modeling to mine the semantic information of tags for users and items respectively, and incorporates the semantic information into matrix factorization to factorize the rating information and to capture the bridging feature of tags and ratings between users and items. To the best of our knowledge, this is the first attempt in the literature to introduce the *DS-WO-CRI* problem and propose a model to deal with it. Finally, the experiments conducted on two well-known datasets have demonstrated that our model significantly outperforms the state-of-the-art approaches in terms of both precision and recall, especially in *DS-WO-CRI* situations.

Acknowledgments

We thank the anonymous reviewers for their helpful suggestions. This work was supported in part by the National Natural Science Foundation of China (No. 61379034), the National Key Technology R&D Program (No. 2014BAH28F05), the Guangdong Province Science and Technology Program (No. 2013B040100004 and No. 2013B040403002), and the Zhejiang Provincial Natural Science Foundation of China (LQ14F010006).

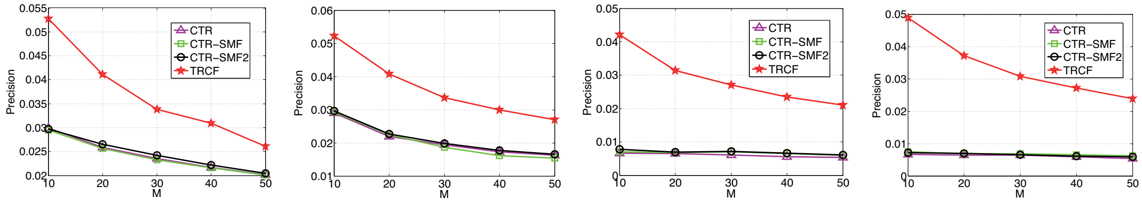
References

Agarwal, D., and Chen, B.-C. 2009. Regression-based latent factor models. In *SIGKDD*, 19–28.

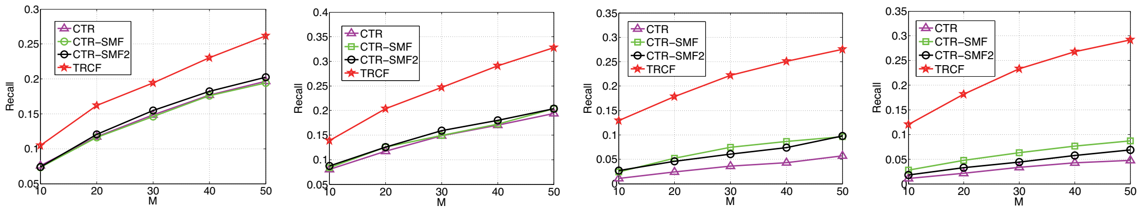
Agarwal, D., and Chen, B.-C. 2010. flda: matrix factorization through latent dirichlet allocation. In *WSDM*, 91–100.

Basilico, J., and Hofmann, T. 2004. Unifying collaborative and content-based filtering. In *ICML*, 65–72.

Bellogín, A.; Cantador, I.; and Castells, P. 2013. A comparative study of heterogeneous item recommendations in social systems. *Information Sciences* 221:142–169.

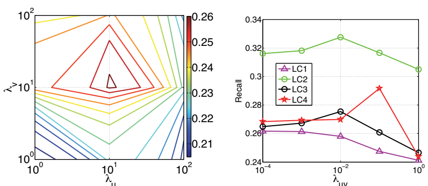


(a) *LC1* (20% *DS-WO-CRI*) (b) *LC2* (40% *DS-WO-CRI*) (c) *LC3* (60% *DS-WO-CRI*) (d) *LC4* (80% *DS-WO-CRI*)



(e) *LC1* (20% *DS-WO-CRI*) (f) *LC2* (40% *DS-WO-CRI*) (g) *LC3* (60% *DS-WO-CRI*) (h) *LC4* (80% *DS-WO-CRI*)

Figure 5: *Precision* and *Recall* comparison on each *DS-WO-CRI* sub-dataset with the best parameters of each method.



(a) Effect of λ_u and λ_v (b) Effect of λ_{uv}

Figure 6: Effect of λ_u, λ_v, and λ_{uv} on the recall performance of TRCF by fixing *M*=50.

Bertsekas, D. P. 1995. Nonlinear programming. *Athena Scientific*.

Blei, D. M.; Ng, A. Y.; and Jordan, M. I. 2003. Latent dirichlet allocation. *JMLR* 3:993–1022.

Cantador, I.; Brusilovsky, P.; and Kuflik, T. 2011. Second workshop on information heterogeneity and fusion in recommender systems (hetrec2011). In *RecSys*, 387–388.

Chen, C.; Zeng, J.; Zheng, X.; and Chen, D. 2013. Recommender system based on social trust relationships. In *ICEBE*, 32–37.

Chen, C.; Zheng, X.; Wang, Y.; Hong, F.; and Lin, Z. 2014. Context-aware collaborative topic regression with social matrix factorization for recommender systems. In *AAAI*. 9–15.

Deshpande, M., and Karypis, G. 2004. Item-based top-n recommendation algorithms. *TOIS* 22(1):143–177.

Fang, X.; Pan, R.; Cao, G.; He, X.; and Dai, W. 2015. Personalized tag recommendation through nonlinear tensor factorization using gaussian kernel. In *AAAI*, 439–445.

Herlocker, J. L.; Konstan, J. A.; Terveen, L. G.; and Riedl, J. T. 2004. Evaluating collaborative filtering recommender systems. *TOIS* 22(1):5–53.

Jamali, M., and Ester, M. 2010. A matrix factorization technique with trust propagation for recommendation in social networks. In *RecSys*, 135–142.

Koren, Y.; Bell, R.; and Volinsky, C. 2009. Matrix factorization techniques for recommender systems. *Computer* 42(8):30–37.

Koren, Y. 2008. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In *SIGKDD*, 426–434.

Ma, H.; Zhou, D.; Liu, C.; Lyu, M. R.; and King, I. 2011. Recommender systems with social regularization. In *WSDM*, 287–296.

Melville, P.; Mooney, R. J.; and Nagarajan, R. 2002. Content-boosted collaborative filtering for improved recommendations. In *AAAI*, 187–192.

Mnih, A., and Salakhutdinov, R. 2007. Probabilistic matrix factorization. In *NIPS*, 1257–1264.

Purushotham, S.; Liu, Y.; and Kuo, C.-c. J. 2012. Collaborative topic regression with social matrix factorization for recommendation systems. In *ICML*, 759–766.

Shi, Y.; Larson, M.; and Hanjalic, A. 2014. Collaborative filtering beyond the user-item matrix: A survey of the state of the art and future challenges. *CSUR* 47(1):3.

Wang, C., and Blei, D. M. 2011. Collaborative topic modeling for recommending scientific articles. In *SIGKDD*, 448–456.

Wang, H.; Chen, B.; and Li, W.-J. 2013. Collaborative topic regression with social regularization for tag recommendation. In *IJCAI*, 2719–2725.

Xu, G.; Gu, Y.; Dolog, P.; Zhang, Y.; and Kitsuregawa, M. 2011. Semrec: a semantic enhancement framework for tag based recommendation. In *AAAI*. 1267–1272.

Xu, J.; Yao, Y.; Tong, H.; Tao, X.; and Lu, J. 2015. Ice-breaking: Mitigating cold-start recommendation problem by rating comparison. In *IJCAI*, 3981–3987.

Zheng, N., and Li, Q. 2011. A recommender system based on tag and time information for social tagging systems. *ESWA* 38(4):4575–4587.

Zhou, T. C.; Ma, H.; King, I.; and Lyu, M. R. 2009. Tagrec: Leveraging tagging wisdom for recommendation. In *CSE*, volume 4, 194–199.

Zhou, T. C.; Ma, H.; Lyu, M. R.; and King, I. 2010. Userrec: A user recommendation framework in social tagging systems. In *AAAI*. 1486–1491.