Published in final edited form as:

Knowl Based Syst. 2016 August 1; 105: 83–95. doi:10.1016/j.knosys.2016.05.010.

A framework for diversifying recommendation lists by user interest expansion

Zhu Zhang^a, Xiaolong Zheng^a, and Daniel Dajun Zeng^{a,b}

^aInstitute of Automation, Chinese Academy of Sciences, Beijing, China, 100190

^bDepartment of Management Information Systems, The University of Arizona, Tucson, USA

Abstract

Recommender systems have been widely used to discover users' preferences and recommend interesting items to users during this age of information load. Researchers in the field of recommender systems have realized that the quality of a top-N recommendation list involves not only relevance but also diversity. Most traditional recommendation algorithms are difficult to generate a diverse item list that can cover most of his/her interests for each user, since they mainly focus on predicting accurate items similar to the dominant interests of users. Additionally, they seldom exploit semantic information such as item tags and users' interest labels to improve recommendation diversity. In this paper, we propose a novel recommendation framework which mainly adopts an expansion strategy of user interests based on social tagging information. The framework enhances the diversity of users' preferences by expanding the sizes and categories of the original user-item interaction records, and then adopts traditional recommendation models to generate recommendation lists. Empirical evaluations on three real-world data sets show that our method can effectively improve the accuracy and diversity of item recommendation.

Keywords

recommender systems; collaborative filtering; diversity; interest expansion; social tagging system

1. Introduction

With the fast growth of web resources such as web pages, music, movies, etc. (referred as items), lots of Internet users are confused by how to select their favorite items from huge amount of items. Since the mid-1990s, Recommender Systems (RS), as an effective information filtering tool to discover users' preferences and recommend interesting items to users, have drawn much attention from academia and industry [1,2]. In real-world applications, RS have been used to recommend web pages, books, music, movies, TV programs, places of interest, conferences, courses, etc. Generally, RS applications can be categorized into eight main domains: e-government [3], e-business [4], e-commerce/e-shopping [5], e-library [6], e-learning [7], e-tourism [8], e-resource services [9], and e-group activities [10], and the application platforms include Web, TV, radio, and mobile devices

[11]. Additionally, as users of RS applications are individuals or groups, the corresponding RS can be classified into individual-based or group-based.

No matter the input data of RS are explicit (e.g., rating and like/dislike) [12,13] or implicit (e.g., clicking and purchase) [14,15], most of RS usually provide a top-N item list for each user. A high-quality top-N recommendation list should have not only high relevance but also high diversity [16]. In this paper, diversity generally indicates the pairwise differences of the recommended items in a top-N list [17], and it has an important impact on users' experiences. For instance, if a user indicates he watched three movies directed by the famous director James Cameron and a few other movies on a movie review website, and then the RS of this site give this user a top-10 recommendation list in which nearly all of the movies are directed by this director. It is obvious that the list is not so appealing due to its monotony.

Traditional RS, including two main categories: latent factor models [14,18] and neighborhood-based models [12,19], are hard to generate a diverse item list that can cover most of one user's interests, since those algorithms mainly focus on predicting accurate items that are very similar to the dominant interests of users [20]. They are also difficult to estimate users' extensive interests due to the insufficiency of original item records' number and categories. In addition, most of traditional RS mainly focus on accuracy in terms of performance evaluation, and the involved measure metrics include precision and recall in information retrieval, and MAE and RMSE in rating prediction. To solve these problems, some studies have attempted to develop several metrics to quantify the diversity of recommendation results [17,20,21] and further to propose new recommendation models for improving the diversity. Among these studies, some of these models are composed of a relevance-oriented part and a diversity-oriented part [21], while some other models take into account relevance and diversity in a unified framework [20,22]. The two-part hybrid modes can incorporate traditional recommendation algorithms and leverage their advantages, while the unified diversification frameworks are more compact and principled in terms of formulation. However, these models don't consider expanding users' interests based on item usage records to improve recommendation diversity, although the expansion of users' interests in the process of building recommendation models may be beneficial to diversity intuitively. Additionally, the recommendation models in these studies only exploit users' item records to build users' interest profiles, so they usually ignore some semantic information such as item tags and users' interest labels. However, tags have the potential to help improve the accuracy and diversity of recommendations, since tags can represent user interests and item features effectively. Some researchers have exploited semantic tag information and the associations among users, tags, and items of the existing Social Tagging System (STS) such as Delicious¹, BibSonomy², and Flickr³, to improve recommendation accuracy. In general, the low diversity of recommendation results is still one of the major research problems in current RS studies.

¹<https://delicious.com/>

²<http://www.bibsonomy.org/>

³<https://www.flickr.com/>

To tackle this problem, we propose a novel user-interest-expansion-based recommendation framework by exploiting social tagging information, which can be adopted to enhance the diversity of recommendation lists. The proposed interest expansion strategy includes three key steps. Firstly, we represent users' features with users' tag records and compute the similarity between tags by item-tag interaction information. Then we expand users' initial tag records based on tag similarity. Secondly, we select new items with proper popularity from the items relating to each user's expanded tag sets, and add those selected items into users' initial item sets. Finally, we treat the expanded users' item sets as the input of traditional item recommendation algorithms, and then we obtain personalized and diverse recommendation lists.

The main contributions of this paper can be summarized as follows:

1. The first innovation of this study is the proposed recommendation diversification framework which mainly utilizes user interest expansion based on social tagging information. To the best of our knowledge, few studies consider enhancing recommendation diversity by exploring social tagging information in the literature concerning RS and STS. The idea of expanding user interests for the improvement of recommendation diversity is also innovative. One advantage of this innovative framework is that it can leverage different kinds of advanced traditional RS within our framework to ensure high recommendation accuracy. Another advantage is that adopting social tagging information for user interest expansion can facilitate the joint improvement of recommendation accuracy and diversity.
2. The second innovation is that we develop a novel diversity metric named tag coverage (i.e., TagCov in the following tables and figures) in the experimental evaluation. The metric can evaluate the diversity of users' interests efficiently, which enrich the diversity measure metrics regarding RS evaluation.
3. Experimental results on three real-world data sets demonstrate that our approach can outperform existing ones. Especially, our approach can improve the diversity and accuracy of item recommendations more efficiently than existing ones.

The rest of this paper is organized as follows. We present related work in Section 2 and introduce the proposed research method in Section 3. Then the experimental evaluation is presented in Section 4. In Section 5, we conclude this paper and present future work.

2. Related work

In general, the related work can be grouped into two parts: recommendation diversity and tag-based item recommendation. The former is related to how to measure and improve the diversity of recommendation list for each user, and the latter is about how to enhance the accuracy of item recommendations by leveraging tag information.

2.1. Recommendation diversity

Accuracy is not the only measure dimension of recommendation results. The importance of diversity in terms of RS performance evaluation had drawn much attention in the early

studies [16,23,24]. For example, McNee et al. [16] argued that accuracy is not always useful and proposed some new user-centric directions for recommendation evaluation. According to [25], the diversity measurement of recommendations can be classified into two types: individual diversity and aggregate diversity. Individual diversity emphasizes the recommendation diversity from an individual user's perspective, which can usually be calculated by an average dissimilarity of all pairs of items recommended to a given user [17,20,21,23]. However, aggregate diversity emphasizes the diversity of recommendations across all users, which can be measured by some metrics (e.g., coverage and diversity-in-top-N) [25]. In this paper, we only focus on the individual diversity of recommendations. For convenience, we will refer to individual diversity as diversity throughout the paper, unless explicitly specified otherwise.

Different diversity metrics were proposed in previous studies. Vargas et al. [17] summarized the state-of-the-art diversity metrics and presented a formalized framework which can generalize these metrics. The common computational method of diversity is provided in the above definition of individual diversity, and the computation of dissimilarity is often based on the semantic information of items or user-item interaction records. In addition, some strategies were proposed to improve the recommendation diversity. For instance, Ziegler et al. [21] proposed a novel model that re-ranks the recommendation list generated by traditional Collaborative Filtering (CF) according to the topic diversity metric, and the online experiments demonstrated that the recommendation accuracy decreased but users' satisfaction increased. Zhang et al. [20] presented a binary optimization method which incorporates the diversity of recommendation lists and the similarity between recommendation results and users' preferences. The experimental results shown that this model outperformed the baselines in terms of diversity and accuracy. Zhou et al. [26] proposed a hybrid method to balance the accuracy and diversity of recommendations, which combines a diffusion-based recommendation algorithm and a "heat-spreading" algorithm addressing diversity problem. Furthermore, Shi et al. [27] argued that different users should have adaptive levels of diversity because of users' different interest extents and users' difference in terms of the estimated preference uncertainty, and then they proposed a new latent factor portfolio model which combines latent factor model and mean-variance analysis in portfolio theory.

In recent years, some novel approaches regarding recommendation diversity were proposed. Noia et al. [28] modeled users' propensity toward selecting diverse items, and then presented a diversification method to re-rank the list of top-N items predicted by a recommendation algorithm. Qin et al. [29] proposed a novel approach called contextual combinatorial bandit, where diverse items that may interest a new user can be dynamically identified by a learning algorithm. However, no specific diversity measure metric was used in the evaluation. Jiang et al. [30] treated recommendation as a multilevel nested logit model from the discrete choice perspective, which can balance relevance and diversity. They then formulated the problem as a nonlinear binary integer programming problem and developed an efficient dynamic programming algorithm. Ashkan et al. [31] presented a novel recommendation diversification method which maximizes the utility of the items subject to the increase in their diversity, and then developed a greedy method to solve the optimization problem.

Some previous studies re-rank the recommendation list to diversify the recommendation results after the traditional recommendation algorithms [21,28]. Our recommendation diversification strategy based on user interest expansion is similar to the re-ranking strategy, since both of them need to combine with the traditional RS algorithms. However, our user interest expansion is performed before the traditional RS algorithms, but the re-ranking strategy is performed after the traditional RS algorithms. In our strategy, we also take into account the balance between diversity and accuracy. Additionally, some studies deal with the relation between diversity and accuracy in a unified model [20,29–31]. Compared with the unified model, one advantage of our strategy is that it can combine with different traditional RS without modifying them and leverage their advantages.

2.2. Tag-based item recommendation

Based on the mentioned association between STS and RS, a number of studies on tag-based item recommendation have been proposed. These studies can be mainly grouped into two categories: heuristic-based and model-based. The first category involves the computation of user similarities or item similarities with tag information, and the similarities can be easily incorporated into traditional user-based or item-based CF algorithms [32–35]. For example, Tso-Sutter et al. [32] extended user profile represented by item vector and item profile represented by user vector with tag information, and then fused the results of user-based and item-based CF algorithms by linear interpolation. Similarly, Zeng and Li [33] proposed two variants of neighborhood-based CF algorithms by using TF-IDF weighted tag vectors to compute user similarity and item similarity.

Except the heuristic-based methods, researchers proposed some model-based methods that utilize tag information to enhance the accuracy of item recommendation. For example, Zhen et al. [36] proposed a matrix factorization model for item recommendation, and the model is regularized by user similarity in tag space. Wetzker et al. [37] presented a recommendation model based on probabilistic latent semantic analysis, which can capture item-user and item-tag co-occurrence information. Both of them are latent factor models that represent users and items with vectors in latent factor space, but the associations among users, items, and tags are not fully exploited in these models. Peng et al. [38] presented a joint item-tag recommendation framework which can indicate users' favorite topics and make use of all interaction information among users, items, and tags. Zhang et al. [39] proposed a recommendation approach based on the information diffusion on the user-item-tag tripartite graph, which is different from the above latent factor models. Nevertheless, the sparsity problem in STS is not well addressed, so Zhang et al. [40] proposed a random walk model which can effectively leverage the associations between users and items to alleviate data sparsity and to enhance recommendation accuracy.

In recent years, Kim et al. [41] presented a hybrid item recommendation framework consisting of tag recommendation and item recommendation. Tag recommendation helps users annotate tags and enriches the data set of a STS, and item recommendation utilizes tags to recommend relevant items. Movahedian et al. proposed [42] a recommendation algorithm based on the similarities between user and item profiles. The user and item profiles are generated by using association rule mining to discover frequent tag patterns.

Pitsilis et al. [43] presented an item recommendation algorithm which is based on tag clustering and the proposed notion Annotation Competency. Zuo et al. [44] proposed an item recommendation algorithm in which users' profiles are initially represented by tags and then a deep neural network model is used to extract the in-depth features from tag space. After users' profiles are updated by the extracted features, user-based CF is used to predict the items that users may like.

All of the above studies mainly focused on how to use tag information for the improvement of recommendation accuracy, but recommendation diversity was not considered in the model designs and experimental evaluations. This paper aims to explore tag information for user interest representation and further propose a more efficient approach for enhancing recommendation accuracy and diversity simultaneously.

3. Method

3.1. Problem statement

Users' tagging behavior data in a STS are treated as the input of the proposed item recommendation method. This input includes users $\{U_i \in U, i = 1, 2, \dots, m\}$, items $\{I_j \in I, j = 1, 2, \dots, n\}$, tags $\{T_k \in T, k = 1, 2, \dots, l\}$, user-item interaction matrix UI , user-tag interaction matrix UT , and item-tag interaction matrix IT . In detail, the value of element UI_{ij} in UI equals 1 if user U_i has used item I_j , otherwise 0. Element UT_{ik} in UT indicates how many times user U_i has used tag T_k , and element IT_{jk} in IT indicates how many times item I_j has been labelled tag T_k . For convenience, we summarize the frequently used notations in Table 1. In this paper, the research question is how to generate a diverse and personalized recommendation list for each user by using social tagging data and make these lists outperform the lists produced by traditional RS algorithms without tag information in terms of accuracy and diversity.

3.2. Traditional CF algorithms

In this section, we introduce several traditional CF algorithms as the background, since these algorithms are adopted as an important step to generate final personalized recommendations in our framework. The input data in this paper are not explicit rating data, and they have no negative instances, so the corresponding solutions are those algorithms called CF with implicit data or one-class CF. As to such CF algorithms, we select the following four representative algorithms: (1) item-based CF [19]; (2) user-based CF [12]; (3) SVD-based (i.e., Singularity Value Decomposition based) RS algorithm [45]; and (4) latent factor model [14]. Since item-based and user-based CF algorithms are simple and often introduced in RS literatures, we omit the introduction of them. However, we apply cosine similarity instead of Pearson correlation coefficient when computing item and user similarity, which is different from the original item-based and user-based CF methods. The reason is that Pearson correlation coefficient is not suitable for the similarity computation between two vectors whose elements are binary. According to the empirical study [45], SVD-based RS algorithm is competitive for top-N recommendation task in terms of recommendation accuracy, so we select it as a representative of matrix factorization models. Latent factor model is a typical and effective item recommendation algorithm, often used as a benchmark in the performance

comparison of one-class CF algorithms. In the following subsections, we will elaborate the SVD-based RS algorithm and latent factor model.

3.2.1. SVD-based RS algorithm—For convenience, we adopt the original notations in the previous work [45] and refer to this algorithm as pureSVD. We treat UI or the expanded user-item interaction matrix as R , and then we estimate the approximation of R by SVD. The details of pureSVD are as follows:

$$R=U \cdot \Sigma \cdot Q^T,$$

where U is a $n \times f$ orthogonal matrix; Q is a $m \times f$ orthogonal matrix; and Σ is a $f \times f$ diagonal matrix having f singular values. The similarities between all items and item i can be obtained by q_i^T , and then we can calculate the preference of user u to item i by the following equation:

$$\hat{r}_{ui}=r_u \cdot Q \cdot q_i^T,$$

where r_u is the row vector of user u in R , and q_i is the column vector of item i in Q .

3.2.2. Latent factor model—In this model, users and items are represented as points in latent factor space, and the inner product of a user point and an item point represents this user's preference to this item. Based on the user-item interaction data, we can estimate all user vectors and item vectors and predict users' preferences. Note that the following notations are borrowed from the previous work [14]. We refer to this algorithm as wMF for convenience.

In detail, r_{ui} indicates the usage extent of user u to item i (e.g., clicks and browsing time), and p_{ui} denotes the preference of user u to item i . p_{ui} equals 1 if user u used item i , 0 otherwise. Besides, c_{ui} denotes the confidence level for p_{ui} , and the relevant formula is:

$$c_{ui}=1+\alpha r_{ui},$$

where α is a constant that can be obtained by cross validation.

$x_u \in R^f$, and x_u denotes user u . $y_i \in R^f$, and y_i denotes item i . Then we use $x_u^T y_i$ to compute the preference of user u to item i . As to the estimation of x_u and y_i , we need to adopt the following optimization objective:

$$\min_{x_u, y_i} \sum_{u, i} c_{ui} (p_{ui} - x_u^T y_i)^2 + \lambda \left(\sum_u \|x_u\|^2 + \sum_i \|y_i\|^2 \right),$$

where $\lambda(\sum_u \|x_u\|^2 + \sum_i \|y_i\|^2)$ is used to regularize the model so that the overfitting can be avoided. The parameter λ depends on the data set and can be determined by cross validation.

3.3. The proposed method

The main idea of the above-mentioned CF models is that we can estimate how much users are interested in items by modeling user-item interaction behavior with user-item interaction data as input. The data are users' item usage records and represent the users' preferences. When users begin to use RS at the early stage, most users' item records are not sufficient and contain many popular and similar items. Then traditional CF algorithms build users' preference models based on these item records of low diversity. Besides, many item recommendation models are based on user or item similarity. Therefore, the item recommendations of the above CF methods are not diverse enough to satisfy users. In this paper, we propose a novel item recommendation framework to make recommendation results accurate and diverse. Similar with Ziegler's work [21], the proposed framework contains two parts: (1) the strategy of improving recommendation diversity and (2) ordinary RS. The main idea of the framework in [21] was to re-rank the recommendation lists generated from ordinary accuracy-oriented RS algorithms according to topic diversity. Our framework, however, enhances the diversity of users' preferences by expanding the sizes and categories of the original user-item interaction records, and then adopts ordinary CF models to generate personalized recommendations. In detail, our framework has three key parts: (1) to expand the extent of users' interests represented by their used tags, (2) to add new items related to users' expanded tag sets into users' original item records, and (3) to adopt ordinary RS algorithms with the expanded user-item interaction records as input. Fig. 1 illustrates the proposed framework briefly.

3.3.1. Expanding user interests—We represent users' initial interests with the tags related to users' item records, as semantic tags are a good way to connect users and items in a STS. We can predict new tags that may interest each user, based on the user-tag interaction matrix UT and the item-tag interaction matrix IT . Then we add these new tags into the initial UT so that the item interests of each user are expanded.

According to the above view, we can use ordinary RS algorithms to solve the tag prediction problem with UT and IT as input. In this paper, we adopt item-based CF due to its effectiveness and simplicity, which is beneficial to keeping the whole recommendation model effective and uncomplicated. Nevertheless, many RS algorithms are also the possible solutions for this problem. The basic idea of item-based CF in this section has two main parts: (1) to select some tags most similar to the target tag from the initial tag records of the target user and (2) to compute the target user's preference to the target tag using the weighted average of the users' preferences to these selected tags. If we use original item-based CF, it is sufficient to treat UT as input. However, IT is generally denser than UT in a STS. Hence, IT is better than UT in the computation of tag similarity, so we use IT to compute tag similarity in our item-based CF. At last, we can predict the new user-tag interaction matrix $\widehat{UT}$. Detailed formulas are as follows:

$$\text{sim}(T_a, T_b) = \cos(IT_{\cdot a}, IT_{\cdot b}) \text{ and}$$

$$\hat{r}_{ik} = \frac{\sum_{x \in UT'(i)} \text{sim}(T_k, T_x) \cdot UT_{ix}}{\sum_{x \in UT'(i)} \text{sim}(T_k, T_x)},$$

where $\text{sim}(T_a, T_b)$ represents the similarity between tag T_a and tag T_b ; IT_a and IT_b represents the vectors of tag T_a and tag T_b respectively; $\hat{r}_{ik}$ is the estimated preference of user U_i to tag T_k ; $UT(i)$ represents the tag set consisting of the tags most similar to tag T_k and selected from the tag records of user U_i .

3.3.2. Expanding users' initial item records—After users' interests are expanded, we need to add some new items concerning users' expanded interests into users' original item records. The added items have new categories different from the categories that the original user-item interaction records have, so it can make users' item preferences more diverse. In order to decrease the risk of selecting items irrelevant to users' preferences, we consider how much users like these tags and how popular the items associated with the tags are, when we select some items linked to users' expanded tag sets. In detail, as to the tag to which a user's preference is high, we select the less popular items concerning this tag, since popular items cover a small amount of categories so that they may not make recommendations diverse. Users' high preferences to the tags can make sure that the items associated with these tags are relevant to users' preferences. On the contrary, as to the tag to which a user's preference is low, we select the more popular items relating to this tag to increase the probability of selecting relevant items. In general, our recommendation diversification strategy by user interest expansion has the potential of increasing the diversity and relevance of recommendation results simultaneously.

The procedure of expanding users' item records, summarized in Table 2, includes the following steps. First, we compute tag-item matrix TI with UI and IT . The value of each element of TI indicates the corresponding item's popularity which is the item's user number. Then we normalize TI by making the sum of each row of TI equal one, and obtain the normalized TI (i.e., $\widehat{TI}$). Secondly, we select a fixed proportion of tags from $\widehat{UT}$ for each user in the descending order of users' tag preferences. The number of the chosen tags is $(1 + r_1)|UT(i)|$, where $r_1, 0 \leq r_1$, indicates the expansion rate of users' tags, and $|UT(i)|$ is the size of the initial tag set of user U_i . Then we normalize the preferences of user U_i to the selected tags and obtain the normalized $\widehat{UT}$ where each of its rows sums up to 1. Thirdly, as to the size of the initial item set of U_i (i.e., $|UI(i)|$), we need to choose a proportion of new items where the ratio is $r_2, 0 \leq r_2$. Then we choose items according to the expanded tag set $\widehat{UT}(i)$ for user U_i , but the number of the items concerning each tag in $\widehat{UT}(i)$ is limited to $r_2|UI(i)| / ((1 + r_1)|UT(i)|)$. Finally, we add item I_j into the item set of this user, if the following conditions are met. (1) The number of the items concerning a tag does not reach the upper limit; (2) the popularity of I_j and the preference of U_i to T_k meet the condition (i.e., $\widehat{TI}_{kj} \leq 1 - \widehat{UT}_{ik}$); and (3) I_j is not in the initial item set of U_i .

3.4. Analysis of time complexity

As to the part of expanding users' interests, the time complexity of computing tag similarity with IT is $O(n^2l)$. However, in real-world applications this part can be carried out offline and implemented by adopting existing distributed and parallel computing frameworks such as Hadoop. The time complexity of computing $\widehat{UT}$ by item-based CF is $O(mlt + ml \log l)$, where t is the average number of the tags of every user. As to the part of expanding the initial item

records of users, the time complexities of computing the popularity of every item, computing $\widehat{U_T}$, and computing $\widehat{U_I}$ are $O(mn)$, $O(nl \log n)$, and $O(mnuv)$ respectively, where u is the average number of the items of every user and v is the average number of the expanded tags of every user. On the whole, the time complexity of the proposed recommendation approach is not high, so it is suitable for large-scale data.

4. Experimental evaluation

In this section, we evaluated the proposed recommendation framework with three data sets from real-world STSs. Several experiments were conducted to answer the questions: (1) can the strategy of expanding users' interests enhance the diversity of recommendation lists, compared with other baselines? (2) how do the relevant parameters impact the recommendation performance?

4.1. Experimental setup

Three different data sets from the real-world STSs are used to evaluate the proposed approach. Table 3 summarizes the key statistics of these data sets. The details of these data sets can be accessed in our previous work [40]. For the data sets, we randomly select 80% of the saved items of each user to form the training data set, and withhold the remainder as the test data set. To make the results not sensitive to the partition of each data set, 10 runs are performed for each experiment, and a different random data split is used each time. The results presented in the remainder are the average of 10 runs.

In order to evaluate the proposed strategy, we use four representative CF algorithms as benchmarks. As introduced in the above section, IB and UB indicate item-based and user-based CF algorithms respectively; pureSVD denotes the SVD-based recommendation algorithms; wMF indicates the latent factor model. Our implementation of wMF is based on the codes in MyMediaLite that is an open source library of recommender system algorithms [46]. In the following tables, the algorithms whose names begin with "UIExp" are the proposed algorithms which combine the user interest expansion strategy and traditional CF algorithms.

For all algorithms in this paper, when we tune a parameter, we keep other parameters fixed. We choose precision as the performance indicator when tuning parameters, since precision is more important than the other metrics in RS evaluation. The finally chosen parameters just make the corresponding precision reach the local optimum. The following reported results are not optimal, because it is impossible to traverse the whole parameter space for the best parameter setting.

4.2. Evaluation metrics

The evaluation metrics in this paper include precision, recall, F-measure, Intra-List Distance (ILD) [17,20], and tag coverage. The former three metrics are often used to evaluate recommendation accuracy [24]. For each user, precision equals the number of correct item recommendations divided by the number of recommendations N , where correct recommendations indicate those items in the target user's test set. Recall equals the number of correct item recommendations divided by the number of test items. F-measure is the

harmonic mean of precision and recall. ILD is widely used to measure recommendation diversity, and it indicates the average of all the distances between any two items in a list. If we use user information to represent item vector, ILD can be represented by user-based ILD (i.e., Diversity in the following tables and figures); if we use tag information to represent item vector, ILD can also be represented by tag-based ILD (i.e., TagDiv in the following tables and figures). Similar with topic coverage in [47,48], a novel diversity metric named tag coverage (i.e., TagCov in the following tables and figures) was proposed in this study, and it is the number of the tags concerning the items in a list divided by the number of all tags.

Table 4 summarizes the formal definitions of the above metrics. N_{hit} indicates the number of correct recommendations; N_{test} denotes the number of items in the target user's test set. Besides, $L(u)$ indicates the recommendation list of user u ; $T(i)$ denotes the tag set of item i . Note that these metrics are applicable for each user, and the final value in each trial is the average over all users.

4.3. Experimental results

4.3.1. Analysis of the experimental results—We compare our algorithms beginning with “UIExp” with the baselines including IB, UB, pureSVD, and wMF. Tables 5, 6, and 7 summarize the top-10 recommendation results on the data sets. Note that the unit of the values in the three tables is a percent symbol. For each metric, the bigger the value is, the better the algorithm is. As to each baseline and its counterpart, t -test is performed to detect their difference significance. If a value is bold in the tables, it means the value is bigger than its counterpart at the 5% significance level according to t -test. We can have the following observations.

1. The recommendation relevance measured by precision, recall, and F-measure are significantly improved when using tag information to expand users' interests. For example, in terms of precision, the algorithms beginning with “UIExp” improve by more than 9% over their counterpart algorithms on the BibSonomy data set, except that UIExpUB improves by more than 3% over UB.
2. As to recommendation diversity, the proposed algorithms still outperform their corresponding traditional RS algorithms in terms of Diversity and TagCov. For instance, in terms of TagCov, our algorithms improve by more than 5% over their counterpart algorithms on the three data sets except UIExpIB and UIExpWMF on the CiteULike data set. Especially, our algorithms improve by more than 30% compared with their counterparts on the BibSonomy data set. Nevertheless, our algorithms don't outperform their counterparts in terms of TagDiv, since the improvement of precision leads to more accurately predicted items in the recommendation lists and those initially chosen items are similar in terms of tag information.

In order to understand the diversification performance of our algorithms, we also compare the proposed algorithms with two kinds of widely used diversification methods (i.e., IBTopDiv [21] and IBBinOpt [20]). IBTopDiv re-ranks the recommendation list generated by traditional CF algorithms according to the topic diversity metric. IBBinOpt is a binary

optimization method which incorporates the diversity of recommendation lists and the similarity between recommendation results and users' preferences. Fig. 2–7 present the top-10 recommendation results of IB, UIExpIB, IBTopDiv, and IBBinOpt. We can have the following observations from these figures. For these figures, *t*-test is also used to test the difference significance at the 5% significance level.

1. From Fig. 2, 4, and 6, we can observe that our algorithm IBExpIB outperforms the other algorithms on all the data sets in terms of the accuracy metrics. The reason is that the proposed approach densifies the user-item interaction matrix (the input of the traditional CF algorithms) by expanding users' interests.
2. From Fig. 3, 5, and 7, we can observe that our algorithm outperforms the other algorithms on all the data sets in terms of Diversity. We can also observe that IBExpIB is better than the other algorithms on the BibSonomy and Delicious data sets in terms of TagCov. However, IBExpIB is worse than the other algorithms on all the data sets with respect to TagDiv. We think the reason is the same as the above explanation mentioned in the analysis of Tables 5, 6, and 7.

In addition, we conduct the following experiments in order to understand whether the size of recommendation lists can influence the diversification method comparison. The results are summarized in Fig. 8, 9, and 10. We can observe that the proposed UIExpIB outperforms the other algorithms on all the data sets regarding precision and UIExpIB is better than the other algorithms on the BibSonomy and Delicious sets regarding tag coverage from top-5 to top-30 recommendations.

4.3.2. Impact of parameters—In this subsection, we study two important parameters' impact on the recommendation performance. The first parameter is about the expansion of users' interests, and it equals the number of each user's added tags divided by the number of each user's original tags. The second parameter is about the expansion of users' item sets, and it equals the number of each user's added items divided by the number of each user's original items. We evaluate how the two parameters impact the top-10 recommendation performance on the Delicious data set (the largest in this study). Only F-measure is used as the accuracy metric in this part, since the performances regarding recall and precision are very similar to the performance regarding F-measure for the proposed method.

From Fig. 11 and 12, we can observe that the performances of different metrics vary differently with the increase of the ratio of added tags to original tags. This ratio has a significant impact on F-measure, TagDiv, and TagCov, respectively. However, the ratio doesn't affect Diversity significantly, since the metric Diversity uses user information rather than tag information to compute ILD. As mentioned in Section 4.1, we keep other parameters fixed when tuning the ratio of added tags to original tags. Although relevance and diversity are key aspects of recommendation quality, relevance is more important than diversity. Therefore, we select precision as the indicator when we tune the ratio of added tags to original tags. That is, the relative local optimum value for this ratio in our experiments is the value corresponding to the relative local optimum precision. Note that the relative local optimum value for the ratio of added items to original items is selected in a similar way.

From Fig. 13 and 14, we can also observe that the performances of different metrics vary differently with the increase of the ratio of added items to original items. This ratio has a significant impact on F-measure, Diversity, TagDiv, and TagCov respectively. It is obvious that TagDiv declines with the increase of the ratio of added items to original items, since the improvement of precision with the increase of this ratio results in more accurately predicted items in the recommendation lists and those predicted items are similar to each other in terms of tag information.

5. Conclusion

We study the problem of diversifying the recommendation results in RS in this paper. Our main contributions have three aspects. First, we have proposed a novel recommendation diversification framework by expanding user interests based on social tagging information. Specifically, we first expand user interests that are represented by users' used tags, and then we extend users' original item records by adding new items related to users' expanded interests. Finally, we utilize traditional RS algorithms to generate final recommendation lists with the expanded user-item interaction records as input. Second, we have also proposed a novel diversity metric named tag coverage in the experimental evaluation. Third, we have evaluated our approach on three real-world data sets. The experimental results show that our approach can significantly improve the accuracy and diversity of item recommendation.

In the future, we are interested in improving the proposed recommendation diversification framework by exploring other RS algorithms, item popularity, social network, etc. For instance, only the simple item-based CF is adopted when expanding the user-tag interaction matrix UT , so we would like to examine how the performances of other RS algorithms are in our framework. Besides, we can use other methods to deal with item popularity when expanding users' item records. We can also exploit social network to diversify recommendation results, as a user's different friends have various interests.

Acknowledgments

We would like to thank each member of the SMILES group in the Institute of Automation, Chinese Academy of Sciences. Especially, we would like to thank Kainan Cui, Chuan Luo, and Saike He for their useful discussions. This work was supported by the National Natural Science Foundation of China (Grant No. 71472175, 71103180, 71025001, and 61175040), the National Institutes of Health (NIH) of the USA (Grant No. 1R01DA037378-01), and the Ministry of Health of China (Grant No. 2012ZX10004801 and 2013ZX10004218).

References

1. Adomavicius G, Tuzhilin A. Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Trans Knowl Data Eng.* 2005; 17:734–749. DOI: 10.1109/TKDE.2005.99
2. Bobadilla J, Ortega F, Hernando A, Gutiérrez A. Recommender Systems Survey. *Knowl-Based Syst.* 2013; 46:109–132. DOI: 10.1016/j.knosys.2013.03.012
3. Shambour Q, Lu J. A hybrid trust-enhanced collaborative filtering recommendation approach for personalized government-to-business e-services. *Int J Intell Syst.* 2011; 26:814–843. DOI: 10.1002/int.20495
4. Zhang Z, Lin H, Liu K, Wu D, Zhang G, Lu J. A hybrid fuzzy-based personalized recommender system for telecom products/services. *Inf Sci.* 2013; 235:117–129. DOI: 10.1016/j.ins.2013.01.025

5. Huang Z, Zeng D, Chen H. A Comparison of Collaborative-Filtering Recommendation Algorithms for Ecommerce. *IEEE Intell Syst.* 2007; 22:68–78. DOI: 10.1109/MIS.2007.4338497
6. Porcel C, Herrera-Viedma E. Dealing with incomplete information in a fuzzy linguistic recommender system to disseminate information in university digital libraries. *Knowl-Based Syst.* 2010; 23:32–39. DOI: 10.1016/j.knosys.2009.07.007
7. Chen CM, Duh LJ. Personalized web-based tutoring system based on fuzzy item response theory. *Expert Syst Appl.* 2008; 34:2298–2315. DOI: 10.1016/j.eswa.2007.03.010
8. Lucas JP, Luz N, Moreno MN, Anacleto R, Almeida Figueiredo A, Martins C. A hybrid recommendation approach for a tourism system. *Expert Syst Appl.* 2013; 40:3532–3550. DOI: 10.1016/j.eswa.2012.12.061
9. Li HQ, Xia F, Zeng D, Wang FY, Mao WJ. Exploring Social Annotations with the Application to Web Page Recommendation. *J Comput Sci Technol.* 2009; 24:1028–1034. DOI: 10.1007/s11390-009-9292-6
10. Quijano-Sanchez L, Recio-Garcia JA, Diaz-Agudo B, Jimenez-Diaz G. Social Factors in Group Recommender Systems. *ACM Trans Intell Syst Technol.* 2013; 4:8:1–8:30. DOI: 10.1145/2414425.2414433
11. Lu J, Wu D, Mao M, Wang W, Zhang G. Recommender system application developments: A survey. *Decis Support Syst.* 2015; 74:12–32. DOI: 10.1016/j.dss.2015.03.008
12. Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J. Proc 1994 ACM Conf Comput Support Coop Work. ACM; 1994. GroupLens: an open architecture for collaborative filtering of netnews; p. 175-186.
13. Koren, Y. Proc 14th ACM SIGKDD Int Conf Knowl Discov Data Min. ACM; 2008. Factorization meets the neighborhood: a multifaceted collaborative filtering model; p. 426-434.
14. Hu Y, Koren Y, Volinsky C. Collaborative Filtering for Implicit Feedback Datasets. *IEEE Computer Society.* 2008; :263–272. DOI: 10.1109/ICDM.2008.22
15. Pan W, Zhong H, Xu C, Ming Z. Adaptive Bayesian personalized ranking for heterogeneous implicit feedbacks. *Knowl-Based Syst.* 2015; 73:173–180. DOI: 10.1016/j.knosys.2014.09.013
16. McNee, SM., Riedl, J., Konstan, JA. CHI 06 Ext Abstr Hum Factors Comput Syst. ACM; New York, NY, USA: 2006. Being Accurate is Not Enough: How Accuracy Metrics Have Hurt Recommender Systems; p. 1097-1101.
17. Vargas, S., Castells, P. Proc Fifth ACM Conf Recomm Syst. ACM; New York, NY, USA: 2011. Rank and Relevance in Novelty and Diversity Metrics for Recommender Systems; p. 109-116.
18. Koren Y, Bell R, Volinsky C. Matrix Factorization Techniques for Recommender Systems. *Computer.* 2009; 42:30–37. DOI: 10.1109/MC.2009.263
19. Sarwar, B., Karypis, G., Konstan, J., Riedl, J. Proc 10th Int Conf World Wide Web. ACM; New York, NY, USA: 2001. Item-based Collaborative Filtering Recommendation Algorithms; p. 285-295.
20. Zhang, M., Hurley, N. Proc 2008 ACM Conf Recomm Syst. ACM; New York, NY, USA: 2008. Avoiding Monotony: Improving the Diversity of Recommendation Lists; p. 123-130.
21. Ziegler, C-N., McNee, SM., Konstan, JA., Lausen, G. Proc 14th Int Conf World Wide Web. ACM; New York, NY, USA: 2005. Improving Recommendation Lists Through Topic Diversification; p. 22-32.
22. Oh, J., Park, S., Yu, H., Song, M., Park, S-T. Proc 2011 IEEE 11th Int Conf Data Min. IEEE Computer Society; Washington, DC, USA: 2011. Novel Recommendation Based on Personal Popularity Tendency; p. 507-516.
23. Smyth, B., McClave, P. Proc 4th Int Conf Case-Based Reason Case-Based Reason Res Dev. Springer-Verlag; London, UK, UK: 2001. Similarity vs. Diversity; p. 347-361.<http://dl.acm.org/citation.cfm?id=646268.758890> [accessed April 28, 2016]
24. Herlocker JL, Konstan JA, Terveen LG, Riedl JT. Evaluating Collaborative Filtering Recommender Systems. *ACM Trans Inf Syst.* 2004; 22:5–53. DOI: 10.1145/963770.963772
25. Adomavicius G, Kwon Y. Improving Aggregate Recommendation Diversity Using Ranking-Based Techniques. *IEEE Trans Knowl Data Eng.* 2012; 24:896–911. DOI: 10.1109/TKDE.2011.15

26. Zhou T, Kuscsik Z, Liu JG, Medo M, Wakeling JR, Zhang YC. Solving the apparent diversity-accuracy dilemma of recommender systems. *Proc Natl Acad Sci*. 2010; 107:4511–4515. DOI: 10.1073/pnas.1000488107 [PubMed: 20176968]
27. Shi, Y., Zhao, X., Wang, J., Larson, M., Hanjalic, A. *Proc 35th Int ACM SIGIR Conf Res Dev Inf Retr*. ACM; New York, NY, USA: 2012. Adaptive Diversification of Recommendation Results via Latent Factor Portfolio; p. 175-184.
28. Di Noia, T., Ostuni, VC., Rosati, J., Tomeo, P., Di Sciascio, E. *Proc 8th ACM Conf Recomm Syst*. ACM; New York, NY, USA: 2014. An Analysis of Users' Propensity Toward Diversity in Recommendations; p. 285-288.
29. Qin, L., Chen, S., Zhu, X. *Proc 2014 SIAM Int Conf Data Min*. Society for Industrial and Applied Mathematics; 2014. Contextual Combinatorial Bandit and its Application on Diversified Online Recommendation; p. 461-469. <http://epubs.siam.org/doi/abs/10.1137/1.9781611973440.53> [accessed April 26, 2016]
30. Jiang H, Qi X, Sun H. Choice-Based Recommender Systems: A Unified Approach to Achieving Relevancy and Diversity. *Oper Res*. 2014; doi: 10.1287/opre.2014.1292
31. Ashkan, A., Kveton, B., Berkovsky, S., Wen, Z. *Proc 24th Int Conf Artif Intell*. AAAI Press; Buenos Aires, Argentina: 2015. Optimal Greedy Diversity for Recommendation; p. 1742-1748. <http://dl.acm.org/citation.cfm?id=2832415.2832491> [accessed April 26, 2016]
32. Tso-Sutter, KHL., Marinho, LB., Schmidt-Thieme, L. *Proc 2008 ACM Symp Appl Comput*. ACM; New York, NY, USA: 2008. Tag-aware Recommender Systems by Fusion of Collaborative Filtering Algorithms; p. 1995-1999.
33. Zeng, D., Li, H. How Useful Are Tags? — An Empirical Analysis of Collaborative Tagging for Web Page Recommendation. In: Yang, CC.Chen, H.Chau, M.Chang, K.Lang, S-D.Chen, PS.Hsieh, R.Zeng, D.Wang, F-Y.Carley, K.Mao, W., Zhan, J., editors. *Intell Secur Inform*. Springer; Berlin Heidelberg: 2008. p. 320-330. http://link.springer.com/chapter/10.1007/978-3-540-69304-8_32 [accessed April 28, 2016]
34. Zhao, S., Du, N., Nauertz, A., Zhang, X., Yuan, Q., Fu, R. *Proc 13th Int Conf Intell User Interfaces*. ACM; New York, NY, USA: 2008. Improved Recommendation Based on Collaborative Tagging Behaviors; p. 413-416.
35. Zheng N, Li Q. A recommender system based on tag and time information for social tagging systems. *Expert Syst Appl*. 2011; 38:4575–4587. DOI: 10.1016/j.eswa.2010.09.131
36. Zhen, Y., Li, W-J., Yeung, D-Y. *Proc Third ACM Conf Recomm Syst*. ACM; New York, NY, USA: 2009. TagiCoFi: Tag Informed Collaborative Filtering; p. 69-76.
37. Wetzker, R., Umbrath, W., Said, A. *Proc WSDM 09 Workshop Exploit Semantic Annot Inf Retr*. ACM; New York, NY, USA: 2009. A Hybrid Approach to Item Recommendation in Folksonomies; p. 25-29.
38. Peng, J., Zeng, DD., Zhao, H., Wang, F. *Proc 19th ACM Int Conf Inf Knowl Manag*. ACM; New York, NY, USA: 2010. Collaborative Filtering in Social Tagging Systems Based on Joint Item-tag Recommendations; p. 809-818.
39. Zhang ZK, Zhou T, Zhang YC. Personalized recommendation via integrated diffusion on user-item-tag tripartite graphs. *Phys Stat Mech Its Appl*. 2010; 389:179–186. DOI: 10.1016/j.physa.2009.08.036
40. Zhang Z, Zeng DD, Abbasi A, Peng J, Zheng X. A Random Walk Model for Item Recommendation in Social Tagging Systems. *ACM Trans Manage Inf Syst*. 2013; 4:8:1–8:24. DOI: 10.1145/2490860
41. Kim H, Kim HJ. A framework for tag-aware recommender systems. *Expert Syst Appl*. 2014; 41:4000–4009. DOI: 10.1016/j.eswa.2013.12.019
42. Movahedian H, Khayyambashi MR. A tag-based recommender system using rule-based collaborative profile enrichment. *Intell Data Anal*. 2014; 18:953–972. DOI: 10.3233/IDA-140677
43. Pitsilis G, Wang W. Harnessing the power of social bookmarking for improving tag-based recommendations. *Comput Hum Behav*. 2015; 50:239–251. DOI: 10.1016/j.chb.2015.03.045
44. Zuo Y, Zeng J, Gong M, Jiao L. Tag-aware recommender systems based on deep neural networks. *Neurocomputing*. n.d; doi: 10.1016/j.neucom.2015.10.134

45. Cremonesi, P., Koren, Y., Turrin, R. Proc Fourth ACM Conf Recomm Syst. ACM; New York, NY, USA: 2010. Performance of Recommender Algorithms on Top-n Recommendation Tasks; p. 39-46.
46. Gantner, Z., Rendle, S., Freudenthaler, C., Schmidt-Thieme, L. Proc Fifth ACM Conf Recomm Syst. ACM; New York, NY, USA: 2011. MyMediaLite: A Free Recommender System Library; p. 305-308.
47. An W, Liu Q, Zhang L. Review on Diversity in Personalized Recommender Systems. Libr Inf Serv. 2013; 57
48. Cui CR, Ma J. An Image Tag Recommendation Approach Combining Relevancy with Diversity. Chin J Comput. 2013; 36

Fig. 1. Illustration of the proposed recommendation diversification framework.

Fig. 2.
Diversification method comparison regarding accuracy on the BibSonomy data set.

Fig. 3.
Diversification method comparison regarding diversity on the BibSonomy data set.

Fig. 4.

Diversification method comparison regarding accuracy on the CiteULike data set.

Fig. 5. Diversification method comparison regarding diversity on the CiteULike data set.

Fig. 6.
Diversification method comparison regarding accuracy on the Delicious data set.

Fig. 7.
Diversification method comparison regarding diversity on the Delicious data set.

Fig. 8.

Diversification method comparison regarding different recommendation sizes on the BibSonomy data set.

Fig. 9.

Diversification method comparison regarding different recommendation sizes on the CiteULike data set.

Fig. 10.

Diversification method comparison regarding different recommendation sizes on the Delicious data set.

Fig. 11.

Impact of the ratio of added tags to original tags on F-measure and Diversity.

Fig. 12.
Impact of the ratio of added tags to original tags on TagDiv and TagCov.

Fig. 13.

Impact of the ratio of added items to original items on F-measure and Diversity.

Fig. 14.

Impact of the ratio of added items to original items on TagDiv and TagCov.

Table 1

Notations.

Notation	Description
U_i	The i_{th} user in user set U
I_j	The j_{th} item in item set I
T_k	The k_{th} tag in tag set T
UI_{ij}	The element in user-item matrix UI ; if user i saves item j , it equals 1, and 0 otherwise
UT_{ik}	The element in user-tag matrix UT ; it equals the frequency which tag k has been used by user i
IT_{jk}	The element in item-tag matrix IT ; it equals the frequency which tag k has been assigned to item j
M_{norm}	The stochastic matrix generated by normalizing each row of matrix M to be of unit length
M_i	The i_{th} row vector of matrix M
M_j	The j_{th} column vector of matrix M
N	The number of recommendations for each user

Table 2

The procedure of expanding users' item sets.

Algorithm 1. The procedure of expanding users' item sets

Input: $UI, IT, \widehat{UT}$ (i.e., the expanded user-tag matrix), $UI(i)$ (i.e., the initial item set of U_i), $UT(i)$ (i.e., the initial tag set of U_i), r_1 (i.e., the ratio of added tags), and r_2 (i.e., the ratio of added items)

Output: $\widehat{UI}$ (i.e., the expanded user-item interaction matrix)

```

1:      compute the popularity of  $I_j$  from  $UI, pop(j)$  (i.e., the user number of  $I_j$ )
2:
3:      build tag-item matrix  $TI$ , where  $TI_{kj} = pop(j)$ ; then obtain  $\widehat{TI}$  by normalizing  $TI$  (i.e.,  $\widehat{TI} = TI_{norm}$ )
4:
5:      sort each row of  $\widehat{UT}$  in the descending order, and only reserve the top elements of  $\widehat{UT}$ , and then normalize it
      (i.e.,  $\widehat{UT} = \widehat{UT}_{norm}$ )
6:
7:      for  $i = 1, 2, \dots, m$ 
8:
9:          for  $k$  in  $\widehat{UT}(i) / \widehat{UT}(i)$  is the expanded and normalized tag set of  $U_i$ 
10:
11:              count = 0;
12:              for  $j = 1, 2, \dots, n$ 
13:
14:                  if  $\widehat{TI}_{kj} \leq (1 - \widehat{UT}_{ik}) \& \& I_j \notin UI(i)$ 
15:
16:                      add  $I_j$  into  $\widehat{UI}(i)$ ;
17:                      count += 1;
18:                      if count ==  $r_2 | UI(i) \cup (1 + r_1) | UT(i) |$ 
19:                          break;
20:                      end for
21:                  end for
22:              end for
23:          end for
24:      end for

```

Table 3

The descriptions of the data sets.

Data set	BibSonomy	CiteULike	Delicious
Number of users: m	125	338	1097
Number of items: n	388	392	1872
Number of selected/total tags: l	78/2311	52/2822	526/9608
Number of total transactions: p	4383	6031	44599
Data density: $p/(mn)$ (%)	9.04	4.55	2.17
Average number of items per user	35.06	17.84	40.66
Average number of users per item	11.30	15.39	23.82

Table 4

The formulas of the evaluation metrics.

Metric	Formula
Precision	N_{hit}/N
Recall	N_{hit}/N_{test}
F-measure	$2 \cdot (\text{precision} \cdot \text{recall})/(\text{precision} + \text{recall})$
Diversity	$2/(N(N-1)) \sum_{i,j \in L(u), i \neq j} (1 - \cos(UI_i, UI_j))$
Tag diversity	$2/(N(N-1)) \sum_{i,j \in L(u), i \neq j} (1 - \cos(IT_i, IT_j))$
Tag coverage	$\frac{ U_{i=1}^N T(i) }{l}$

Table 5

Top-10 recommendations on the BibSonomy data set.

Algorithm	Precision (%)	Recall (%)	F-measure (%)	Diversity (%)	TagDiv (%)	TagCov (%)
IB	10.30	4.23	6.00	88.93	93.33	30.50
UIExpIB	11.82	4.78	6.81	94.40	86.40	47.21
UB	12.59	4.97	7.13	88.36	85.88	52.13
UIExpUB	13.07	5.11	7.34	91.84	77.15	71.88
pureSVD	10.15	4.25	5.99	91.94	87.68	40.20
UIExpPureSVD	11.82	4.72	6.74	92.72	86.72	53.49
wMF	11.36	4.68	6.63	89.28	88.77	45.40
UIExpwMF	12.69	5.04	7.21	92.42	81.77	59.65

Table 6

Top-10 recommendations on the CiteULike data set.

Algorithm	Precision (%)	Recall (%)	F-measure (%)	Diversity (%)	TagDiv (%)	TagCov (%)
IB	7.09	6.07	6.54	91.29	92.91	29.10
UIExpIB	8.26	7.15	7.66	94.14	85.64	28.53
UB	9.29	7.40	8.23	91.03	89.05	40.62
UIExpUB	9.75	7.82	8.68	92.88	83.52	47.10
pureSVD	7.03	5.45	6.13	87.75	90.40	40.32
UIExpPureSVD	7.61	5.87	6.62	91.35	87.19	42.61
wMF	7.65	6.10	6.78	91.55	90.62	40.57
UIExpwMF	8.45	6.91	7.59	94.04	81.63	41.06

Table 7

Top-10 recommendations on the Delicious data set.

Algorithm	Precision (%)	Recall (%)	F-measure (%)	Diversity (%)	TagDiv (%)	TagCov (%)
IB	2.69	0.84	1.28	96.66	93.95	12.39
UIExpIB	4.53	1.42	2.17	97.93	81.95	22.65
UB	4.15	1.24	1.91	96.72	89.13	20.37
UIExpUB	4.72	1.48	2.26	97.62	76.44	23.09
pureSVD	3.58	1.10	1.68	91.72	87.91	21.22
UIExpPureSVD	4.43	1.41	2.14	97.96	82.28	23.76
wMF	4.54	1.39	2.13	95.46	86.27	22.96
UIExpwMF	4.99	1.56	2.38	97.71	79.97	24.21