Pers Ubiquit Comput (2014) 18:1855–1869
DOI 10.1007/s00779-014-0785-0

ORIGINAL ARTICLE

On improving aggregate recommendation diversity and novelty in folksonomy-based social systems

Hao Wu · Xiaohui Cui · Jun He · Bo Li · Yijian Pei

Received: 29 December 2013 / Accepted: 29 April 2014 / Published online: 10 August 2014

Abstract Benefit from technical advances in the Internet of Things, many social media applications relative to folksonomy have become ubiquitous. The size and complexity of folksonomy-based systems can unfortunately lead to information overload and reduced utility for users. Consequentially, the increasing need for recommender services from users has arisen. Many efforts have been made to address recommendation accuracy as well as other issues with respect to personalized recommendation in such systems. A key challenge facing these systems is that the most useful individual recommendations are to be found among diverse niche resources while increasing diversity most often compromises accuracy. In this paper, we introduce a simple yet elegant method—Diversity-aware Personalized PageRank (DaPPR)—to address this challenge from the aggregate perspective. DaPPR exploits a balance factor to adjust the influence of a personalized ranking vector and a unified non-personalized ranking vector based on PageRank. By this, it can reduce the impact of resource popularity on recommendations and then generate more diverse and novel recommendations to users. A hybrid DaPPR model that combines two ranking processes on the user–resource and the resource–tag bipartite graphs is specifically designed to meet the requirements in folksonomy-based systems. According to solid experiments, our proposed method yields better results balancing both aggregate accuracy and aggregate diversity (novelty). Improvements of all performance metrics are also obtained compared with the existing algorithms.

Keywords Folksonomy · Recommender systems · Diversity-aware · Personalized PageRank · Ubiquitous computing

H. Wu (✉) · B. Li · Y. Pei
School of Information Science and Engineering, Yunnan University, Kunming 650091, China
e-mail: haowu@ynu.edu.cn

B. Li
e-mail: libo@ynu.edu.cn

Y. Pei
e-mail: pei3p@ynu.edu.cn

X. Cui (✉)
School of Software, Wuhan University, Wuhan 420073, China
e-mail: cuixhui@gmail.com

J. He (✉)
Nanjing University of Information Science and Technology, Nanjing 210044, China
e-mail: hejun.zz@gmail.com

1 Introduction

Collaborative tagging systems [10], such as Delicious, Flickr, Youtube, Lastfm, Connotea, CiteUlike and MovieLens, have become a kind of booming business on the Internet. These systems provide a wealth of information, where any persons can freely find, annotate, organize various resources of interest and share their findings (this practice is coined as *Folksonomy* by Thomas Vander Wal). As an information carrier, the tags play a key role in such systems, since they cannot only express the main features of the resources, but also cover relationships of users–resources/items(we use them alternatively) and items–items. Benefit from technical advances in the Internet of Things, many social media applications relative to folksonomy have become ubiquitous [27]. Users are easily and conveniently accessing rich multimedia content, with the rapid popularity of smart devices. They can obtain information of Web pages from the Delicious, music from the Lastfm and photos from the Flickr at any time, to name a few.

The size and complexity of folksonomy-based systems can unfortunately lead to information overload and reduced utility for users. Too many resources can make users helpless in their process of finding useful contents. Consequentially, the increasing need for recommender services from users has arisen. For these reasons, researchers have sought to apply the techniques of recommender systems to deliver personalized views. The current researches of personalization in such systems can be subdivided into tag recommendation and item/resource recommendation. Given a user and a resource, the former predicts what and how tags will be adopted by the user to explain the resource, whereas the latter emphasizes suggesting unseen items of interest to the user. Many works based on different principles, such as network-based, tensor-decomposition and collaborative filtering, have been proposed to address recommendation accuracy as well as the problems of data sparsity, cold start and so on [32].

Nevertheless, there still remain many issues [11, 32], among which, one key problem is how to improve diversity while maintaining accuracy of recommendation (this issue exists in all kinds of social systems, not just folksonomy-based systems). Accuracy is only one property of recommender system that decides whether generated recommendations are accurately identified to user's interests and likes–dislikes or not [24]. However, the quality of recommendations can be evaluated along a number of dimensions, and relying on the accuracy of recommendations alone may not be enough to find the most useful items for each user [1, 7, 13, 16, 32]. In particular, the importance of diverse recommendations has been previously emphasized in many studies [4, 15, 22, 26, 30, 34], due to a general observation that the most useful individual recommendations are to be found among diverse niche items [13, 33]. Most of these studies measure recommendation diversity from an individual user's perspective (i.e., *individual diversity*) and seek to find out the best diverse subset of items (which often maximizes an average dissimilarity between all pairs of recommended items) suggested to the individual user. In contrast to individual diversity, which has been intensively explored [24], some recent studies [1, 7, 9, 23, 33] start examining the impact of recommender systems on diversity by considering aggregate diversity of recommendations across all users. According to them, the benefits of recommender systems that provide higher aggregate diversity would be apparent to both users and business providers, since it can essentially improve user experience and provide a better coverage of the solution space, furthermore be beneficial for some business models as well.

On the other hand, as stated in prior works [1, 33], there also exists an apparent accuracy–diversity dilemma from the aggregate perspective that increasing recommendation diversity most often compromises recommendation accuracy, and vice versa. Therefore, how to develop effective algorithms to balance both aggregate recommendation accuracy and diversity still remains as an open problem. To this end, this paper first proposes a novel algorithm named Diversity-aware Personalized PageRank (DaPPR) grounding on the well-known PageRank [5]. Specially, to solve the recommendation accuracy–diversity dilemma, a balance factor is introduced to adjust the influence of a personalized ranking vector and a unified non-personalized ranking vector. Then, we adapt the DaPPR to form a hybrid model against folksonomy schema and fulfill personalized recommendation for folksonomy-based systems. We use real data collected from three folksonomy-based systems to verify the validity of the proposed approach.

The rest of this paper is structured as follows. Section 2 presents basic definitions of metrics to characterize recommendation performance. In Sect. 3, we first explain the concepts of the folksonomy schema and Personalized PageRank. Based on this, we present how to introduce a balance factor to Personalized PageRank to create the DaPPR, and how to adapt the DaPPR to form a hybrid model for personalized recommendation in folksonomy-based systems. In Sect. 4, we present and analyze the experimental results according to three folksonomy-based data collections. In Sect. 5, some discussions concerning our methods and related works are given. Finally, we conclude the paper and point out future directions.

2 Basic definitions

2.1 Accuracy of recommendations

Along with the development of recommendation techniques, various metrics have been employed for measuring the accuracy of recommendations including statistical accuracy metrics and decision-support measures [13]. Mean absolute error (MAE) and root mean squared error (RMSE) metrics, the two basic statistical accuracy metrics, have been extensively used in the evaluation of performance of item rating prediction techniques. Since we focus on top-k item recommendation instead of rating prediction of items, *Precision*, a most basic decision-support metric, is employed to characterize the accuracy of recommendations in the later sections. Precision represents the probability that a selected item is relevant. Given a candidate list L ($|L|$ is the list length) for the active user u, and suppose N_k represents the number of recovered resources in the top-k

($k \leq |L|$) places of L, the precision at the ranking position k is defined as following ($P@k$ is short for $Precision@k$),

$$Precision@k = N_k/k \tag{1}$$

2.2 Diversity and novelty of recommendations

As mentioned above, the diversity of recommendations can be measured in two ways: individual and aggregate. The metrics of *individual diversity* usually characterize an accumulated dissimilarity between all pairs of items within a recommendation list, e.g., *intra-list similarity* [4, 34] and *item novelty* [15]. In contrast to individual diversity, there have been few studies that explore *aggregate diversity* in recommender systems. Herlocker et al. [13] experimented with the *coverage* defined as the percentage of items for which the recommender system is able to make recommendations. Adomavicius and Kwon [1] defined *diversity-in-top-N* as an aggregate diversity measure, to count the total number of distinct items recommended to all users. Since we intend to evaluate recommender systems based on the top-k recommended lists of items that the system provides to its users, in this paper, we take *hamming distance* as our aggregate diversity measure. Given two candidate lists L_u and L_v for user u and v, the difference in the top-k ($k \leq \max(|L_u|, |L_v|)$) places can be measured using hamming distance as following [33],

$$HD@k = 1 - overlap@k/k \tag{2}$$

where $overlap@k$ is the number of shared resources in the top-k places of the two lists. Averaging over all pairs of users, we can obtain the aggregate diversity of the system. Clearly, $HD@k$ can characterize the uniqueness of different user's recommendation lists, higher diversity means higher personalization of users' recommendation lists, $HD@k = 1$ points to the fact that every user receives his/her own unique top-k items [33].

Novelty and diversity are different though related notions. The novelty of a piece of information generally refers to how different it is with respect to "what has been previously seen" by a specific user, or a community as a whole [29]. To evaluate *aggregate novelty* of recommendations, we use popularity-based item novelty proposed by Vargas & Castells [29],

$$Nov@k = -log p(k|u) \tag{3}$$

where k indicates the k-th resource in the recommendation list L of u. Here, the posterior $p(k|u)$ cannot be estimated directly, since it assumes no observation of u accessing k before recommendations are made. Instead, Vargas & Castells suggested estimating $p(k|u)$ based on other items the user has accessed. Let us assume the observed data consist of a set S of user-based assignment of tags to resources, reflecting item access by users; then, $p(k|u)$ can be inferred as followings,

$$\begin{aligned} p(k|u) &\propto \sum_{i \in I_u} p(k|i)p(i|u) \\ p(k|i) &\propto \frac{|U_k \cap U_i|}{|U_i|} \\ p(i|u) &\propto \frac{|(u,t,i) \in S|}{|(u,t,i') \in S|} \end{aligned} \tag{4}$$

where I_u denotes the set of items collected by u, and U_k denotes the set of users who have accessed k. We average $Nov@k$ across all users, to measure the capability of a recommender system provides novel items to its users. Since the novelty measure is not normalized and universal for different systems, it is considered as a complementary measure to the diversity. Note that there still exists an apparent dilemma between recommendation accuracy and diversity (novelty). Thus, in this paper, we aim to find new techniques to improve aggregate diversity and novelty while maintaining accuracy of recommendations, especially for folksonomy-based systems.

3 Methodology

3.1 Folksonomy schema

A folksonomy describes the users, the resources, the tags and the user-based assignment of tags to resources. It is a tuple (U, T, R, S) where U, T and R are finite sets whose elements are called users, tags and resources, respectively, and S is a ternary relation between them, i.e., $S = U \times T \times R$, whose elements are called tag assignments [17]. Users are typically described by their user ID, and tags may be arbitrary strings. What is perceived as a resource depends on the type of social application system. For instance, in Delicious, the resources are URLs, and in Lastfm, the resources are artists and tracks. There are also other homogenous links among users or resources, such as friendship among users, hyperlinks among Web pages and citation links among papers. In addition, resources may have various attributes information. A typical folksonomy schema can be described as Fig. 1, which can be also formulated as an association link network [21].

3.2 Personalized PageRank

PageRank [5] is an ordering nodes technique by a random surfer model in a directed graph $G = (V, E)$, where $V(|V| = n)$ is the set of nodes and E is the set of edges. The random surfer performs a *Markovian* walk on G. The surfer jumps from node to node following a link with uniform probability d (called as *damping factor*) or gets bored and

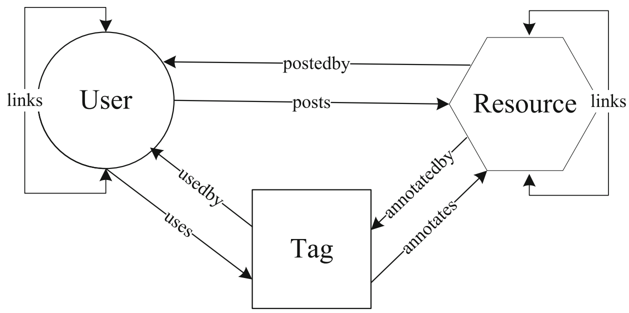

Fig. 1 A typical folksonomy schema

jumps to a random node with likelihood $1-d$. Let r be the ranking vector of all nodes in G; the PageRank value r_i is the probability that the surfer is at the given node i. A fast and easy way to compute the PageRank is using the *power iteration method*, shown as following,

$$r^{(t+1)} = dAr^{(t)} + (1-d)e \tag{5}$$

where A is a stochastic matrix $(n \times n)$. $A_{ij} = o_j^{-1}$ if the node j links to the node i, and o_j is the outgoing degree of the node j; otherwise, $A_{ij} = 0$. Starting with an arbitrary vector $r^{(0)}$, the solving of Eq.5 is equivalent to apply the operator $\widehat{M} = dA + (1-d)e$ in succession, i.e., $r^{(t+1)} = \widehat{M} r^{(t)}$, until $|r^{(t+1)} - r^{(t)}| < \epsilon$. The vector $e(|e| = n)$ is a *preference* (or *personalized*) vector that may represent the interests of a particular user. When setting e to prefer a subset of V, the PageRank model is usually called as Personalized PageRank (PPR [12]).

PPR has recently attracted many attentions in various recommendation scenes and has been proved to achieve superior performance with ability to alleviate data sparsity [8, 14, 18]. It can be adapted to recommend resources for users against the folksonomy schema as following: a) Given a folksonomy data graph, we treat the heterogeneous edges with different types as a single bidirectional link (namely all of them take a weight value of 1); b) set e to prefer the node representing a certain user u, or the nodes representing the resources collected by (or in the profile of) the user u ; and c) find the PPR vector $r_u^{(t)}$ (where t is the state after convergence) using Eq. 5. Since $r_u^{(t)}$ give us the long-term visit rate of each node, given a bias toward the user u, it can be considered as a measure of relatedness between any node i and the user u in the folksonomy data graph [18].

3.3 Diversity-aware Personalized PageRank

When applying the PPR to personalized recommendation, we find that it is trivial to diversify the recommendations while maintaining the accuracy through the adjustment of the damping factor d (see later sections). To cope with this problem, we introduce a new method to estimate the relatedness between an active user u and resources against a data graph G. The detail of the proposed method is indicated by Algorithm 1.

Algorithm 1 *Diversity-aware Personalized PageRank(DaPPR)*

Input: The column-normalized adjacency matrix A of the graph G, the damping factor d, the personalized vector e of the user u, the balance factor b;
Output: The ranking values of all nodes in G;
Let $d = 1$ and compute the ranking vector r_0: $r_0 = \mathbf{1}Ar_0$;
Let $d < 1$ and compute the ranking vector r_1: $r_1 = dAr_1 + (1-d)e$;
Compute the diversity-aware ranking vector $\bar{r} = r_1 - br_0$;
return $\bar{r}$;

The logic behind this method is quite straightforward. Since recommending the most popular resources to each user typically leads to diversity reduction, and the personalized ranking vector r_1 is inevitably influenced by the popularity of nodes (The PPR tends to suggest popular items localized around starting nodes, which combines both the factors of similarity and popularity), if we can reduce the effects of node's popularity on the PPR and recommend less popular resources to each user, the recommendation diversity intuitively should be increased. As we know, the ranking vector r_0 is non-personalized for all users and highly correlated with the degree of nodes (we calculated the *Pearson's correlation coefficient* for the PPR vector r_0 and the in-degree vector of top-1000 popular nodes against user–resource relation, the correlation coefficient is larger than 0.8); thus, it can be seen as an indicator of popularity of nodes. Correspondingly, increasing the *balance factor* b can weaken those resources with higher popularity and minimize the impact of resource popularity on recommendations. Hence, the improvement on diversity and novelty of recommendation can be achieved. We call $\bar{r}$ as Diversity-aware Personalized PageRank (DaPPR) to distinguish from the PPR.

3.4 Hybrid DaPPR model

For personalized recommendation, a popular approach is to exploit hybrid models, for example, a hybrid Collaborative Filtering (CF) method usually combines a user-based algorithm and an item-based algorithm [13]. Following this general idea, we introduce a hybrid version of DaPPR to fulfill personalized recommendation in folksonomy-based systems. For this, we use the projection of folksonomy schema, which has been deeply explored in the research field of tag-aware recommendation [32]. And to investigate the fundamental properties of the model, we only use user–

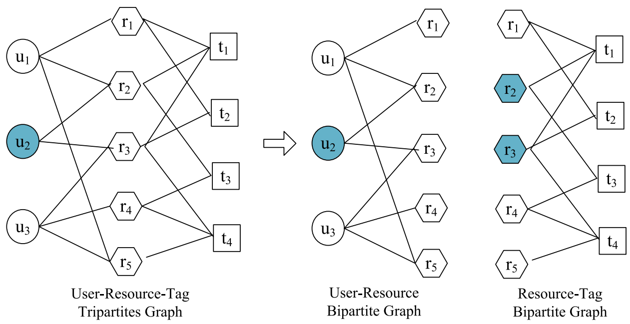

Fig. 2 An example for projection of folksonomy data graph

resource relation and resource–tag relation to create the corresponding bipartite graphs (also known as bigraph). Also, the weights of links are assigned a unified value of 1 in both bigraphs. In the first step, the DaPPR performs independently on the two bigraphs, and in the second step, the two ranking vectors of resources are aggregated using a simple linear combination as following,

$$\bar{r} = \lambda \bar{r}_{UR} + (1 - \lambda)\bar{r}_{RT} \quad (6)$$

where $\bar{r}_{UR}$ and $\bar{r}_{RT}$, respectively, represent the DaPPR-based ranking vectors of resources in the UR bigraph and the RT bigraph. λ is a weighting factor to control the effects of two relations. Correspondingly, for the UR bigraph, we set e to prefer the node representing u, and for the RT bigraph, we set e to prefer the set of resource nodes I_u in the profile of u. The configuration of the preference vector e is described as Eq.7,

$$e_i = \begin{cases} 1 & \texttt{if}\ i = u(\texttt{for}\ \bar{r}_{UR}) \\ \frac{1}{|I_u|} & \texttt{if}\ i \in I_u(\texttt{for}\ \bar{r}_{RT}) \\ 0 & \quad \texttt{otherwise} \end{cases} \quad (7)$$

Figure 2 shows an example for projection of folksonomy data graph, where we consider suggesting items to user u_2. Suppose the vector of nodes for the UR bigraph and the RT bigraph is $[u_1, u_2, u_3, r_1, r_2, r_3, r_4, r_5]$ and $[r_1, r_2, r_3, r_4, r_5, t_1, t_2, t_3, t_4]$, respectively; the configurations of e is correspondingly described as [0, 1, 0, 0, 0, 0, 0, 0] and [0, 1/2, 1/2, 0, 0, 0, 0, 0, 0].

Essentially, the DaPPR in combination with the user–resource relation and the resource–tag relation resembles, respectively, the user-based and the item-based collaborative model, consequently incorporating them, should improve recommendation effectiveness. Note that, our model can be easily extended to take advantage of additional semantic information, for example, we can add the friendships to the users in the UR bipartite graph, and specific attributes information to the resources in the RT bipartite graph. We will demonstrate this in a later section.

Table 1 The statistics of the datasets

Dataset	Delicious	Lastfm	MovieLens
Users	1,867	1,892	2,113
Items	69,223	12,523	5,908
Tags	40,897	9,749	9,079
Tag assignments	437,593	186,479	47,957
Tas per user	234.38	98.56	22.70
Tas per item	6.32	14.89	8.12
Items per user	56.13	37.56	13.11
Users per item	1.51	5.67	4.69
Tags per item	5.93	8.76	6.35
Items per tag	10.04	11.26	4.13
Density of $U \times R$	$8.1 * 10^{-4}$	$3.0 * 10^{-3}$	$2.2 * 10^{-3}$
Density of $R \times T$	$1.5 * 10^{-4}$	$9.0 * 10^{-4}$	$7.0 * 10^{-4}$
Training users	1,839	1,821	1,598
Testing users	798	358	151

4 Experiments and discussions

4.1 Datasets

For experiments, we use the real datasets collected from three well-known social media systems: Delicious[1], Last-fm[2] and MovieLens[3]. Delicious is one of the most popular social bookmarking Web sites, which allow users not only to store and organize personal bookmarks (URLs), but also to look into other users' collections and find what they might be interested in by simply, keeping track of the baskets with tags or resource. Lastfm is the world's largest online music catalogue and allows user tagging music tracks and artists. In this dataset, we take artists as resources. MovieLens is a recommender system and virtual community Web site that recommends films for its users to watch, based on their film preferences and using collaborative filtering. The Web site is maintained by the laboratory of GroupLens Research[4]. The collaborative tagging function had been added to the Web site; thus, researchers can gather tag-aware data for research purpose. For these three systems, we use their data collections released in the framework of the 2nd International Workshop on Information Heterogeneity and Fusion in Recommender Systems [6] to make an evaluation. Statistics of datasets are listed in Table 1, and more detailed descriptions of these datasets can be found in [2].

[1] http://www.delicious.com
[2] http://www.lastfm.com
[3] http://www.imdb.com, http://www.rottentomatoes.com
[4] http://www.grouplens.org

To test the algorithmic performance, both the Delicious dataset and the Lastfm dataset are divided into two parts according to the tag assignment timestamp: the training set contains 90% past entries of tas and the remaining 10% future entries of tas makes up the testing set. Because test cases for the MovieLens dataset are relatively small, we separate this dataset by the ratio of 80%:20%. This policy follows the common observation as known information used for recommending, while no information in the testing set is allowed to be used for recommending. Also, it meets the online operation principle of recommender systems, that is, the recommenders periodically provides active users with resources of interest, at a certain point of time, using the historical data of the systems. Note that, since we do not focus on the cold-start problem in this paper, new users and new resources are eliminated from the testing dataset. The finally selected test cases are also presented in Table 1. Also, when generating the recommendation candidate list for a certain user, the resources already collected by the user are excluded from the list. Note that, the density of a bipartite graph is defined as the ratio of the total number of undirected edges existing between vertices ($|E|$), to the product of cardinality of the two disjoint sets of vertices, for instance, it is $|E|/|U| * |R|$ in the UR bigraph.

4.2 Experiments with DaPPR on user–resource and resource–tag bigraphs

As above mentioned, the DaPPR processing on the user–resource and the resource–tag bigraphs can be seen as, respectively, a user-based and an item-based collaborative model. Here, we first see the recommendation effectiveness of the PPR based on both bigraphs in three datasets. And we observe the values of top-20 recommendations for all performance metrics. The damping factor is settled as $d \in [0.02, 1]$, since as d is close to 0, the adjacent matrix of the PageRank is annihilated, resulting in the meaningless uniform process [3]. For the Delicious dataset (see Fig. 3), all the performance metrics basically remain unchanged in a wide range of $d \in [0.1, 0.9]$. For both the Lastfm and the MovieLens datasets (see Figs. 4 and 5), the recommendation effectiveness also has no significant changes in the range of $d \in [0.1, 0.5]$; however, the recommendation diversity degrades as d increases. When d is close to 1, the personalized vector e decreases in importance, and the recommendation diversity declines remarkably for all the datasets. Basically, the novelty is contradictory to the accuracy, but the diversity is not. Based on these observations, it is obviously trivial to diversify the recommendations while maintain the accuracy by adjusting the damping factor.

In the next, we see the recommendation effectiveness of the DaPPR on both the UR and the RT bigraphs against three datasets. The experimental results are presented in Figs. 6, 7 and 8. For simplicity, the damping factor in each dataset is configured as a fixed value. For the Delicious dataset, d is taken as 0.15, and for both the Lastfm and the MovieLens datasets, the setting is $d = 0.5$. Here, the damping factor can take the same values in both bigraphs. The logic behind this lies in that both the $U \times R$ relation and the $R \times T$ relation are simultaneously derived from the hyper-relation $S = U \times T \times R$; thus, they have inherently the semantic conformity. Also, the performance of recommendations on the UR bigraph and the RT bigraph is different, but it has basically the same tendency as b increases. In the following experiments, we can find that the setting of b also follows the same principle.

For the Lastfm dataset and the MovieLens dataset, shared improvements are obtained for both the diversity and the novelty as b increases (see Figs. 7b, c, e, f, 8b, c, e, f). However, maximizing b has a negative effect on accuracy metric. Fortunately, it is possible to choose a value of balance factor ($b \in (0, 0.1]$ for the Lastfm dataset, and $b \in (0, 0.05]$ for the MovieLens dataset) to improve diversity and novelty without or with few losses of accuracy. For the Delicious dataset (see Fig. 6), there is a little difference. Maximizing b not only leads to the best recommendation accuracy in both bigraphs(see Fig. 6a, d), but also slightly

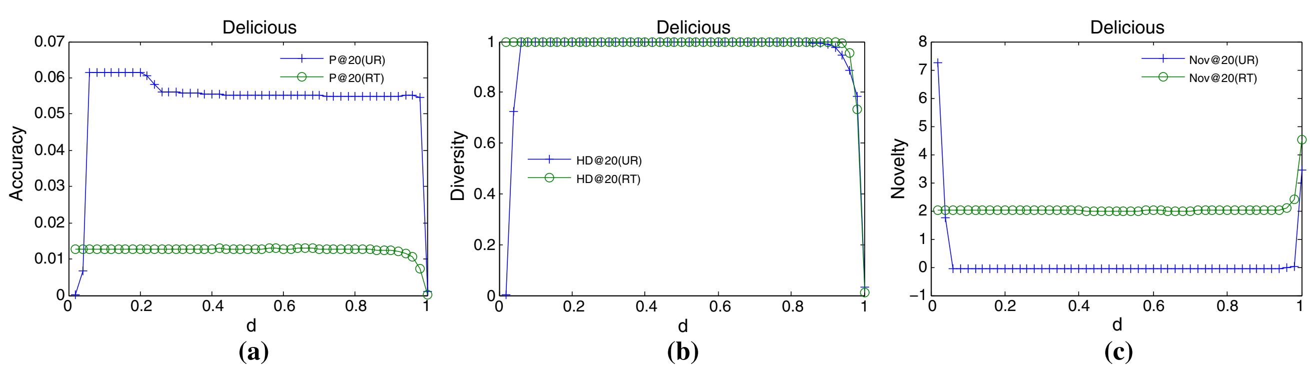

Fig. 3 Recommendation performance of the PPR as varying damping factor $d \in [0.02, 1]$ on the Delicious dataset

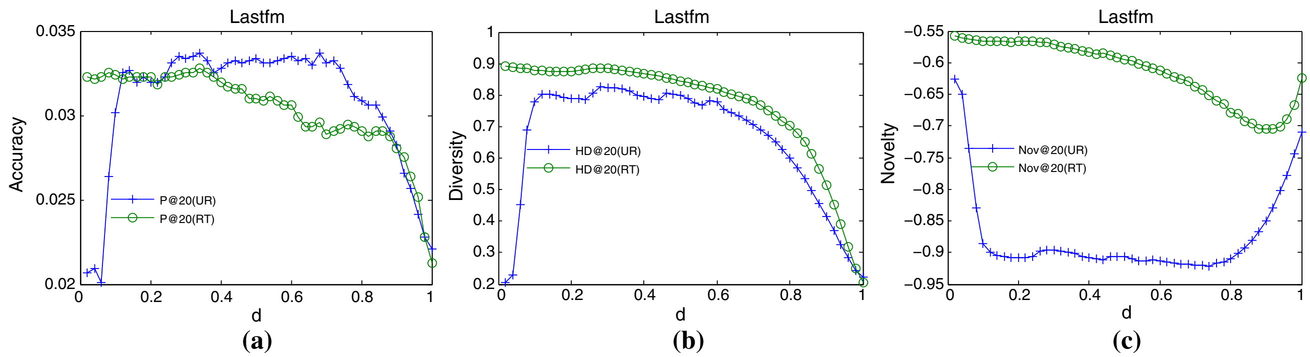

Fig. 4 Recommendation performance of the PPR as varying damping factor $d \in [0.02, 1]$ on the Lastfm dataset

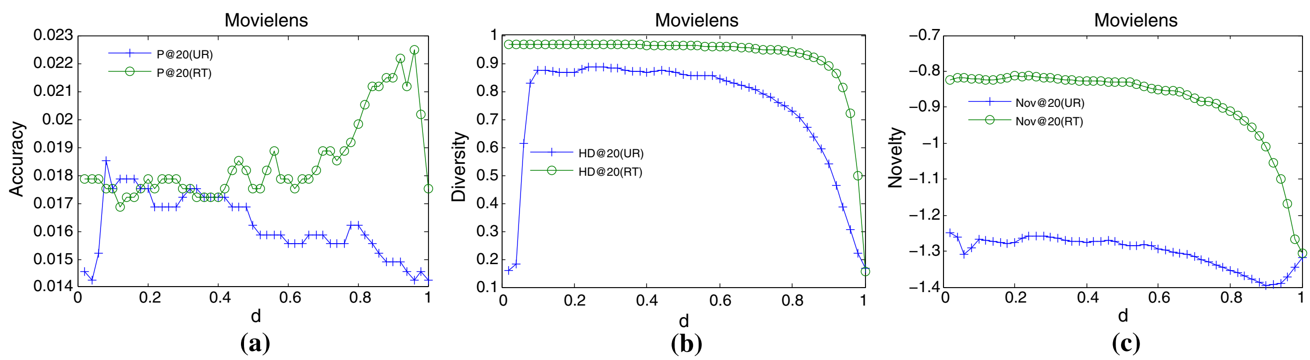

Fig. 5 Recommendation performance of the PPR as varying damping factor $d \in [0.02, 1]$ on the MovieLens dataset

increases recommendation diversity (see Fig. 6e) and novelty(see Fig. 6c). A reasonable explanation to this phenomenon is that the data in the Delicious are much sparser than those of the Lastfm and the MovieLens, and the behaviors of users accessing resources in the Delicious system are dominated by the interests of the users instead of the popularity of the resources, increasing balance factor consequentially brings positive impact to all indicators.

Additionally, the use of resource–tag bigraph typically results in more diverse and niche commendations; this is in accordance with the observation in the work [2]; on the contrary, exploiting user–resource relation most often achieves better recommendations in accuracy. Incorporating two relations, however, allows us to have the best of both worlds. Moreover, we observe that the setting of b yields basically the same tendency of the recommendation performance in the user–resource bigraph and the resource–tag bigraph. This helps to simplify the parameter settings of hybrid recommendation model, where we can always let b (and d) take the same values in both cases.

4.3 Experiments with hybrid recommendation method

In this section, we study the effectiveness of hybrid DaPPR in personalized recommendation and make detailed comparisons with other counterpart methods. For this, we select two methods, which are also based on the network-theory. Although CF-based models are prevalent in rating-based recommender systems, they do not perform well as the network-based methods in folksonomy-based systems [18, 32]; accordingly, we omit them in our experiments.

Hybrid PPR: This method is adapted from the hybrid DaPPR model using $b = 0$. We take the hybrid PPR as the baseline, as it can be seen as a non-diversity-aware model.

Hybrid ProbS+HeatS: Zhou et al. [33] developed an approach that combines an accuracy-focused algorithm (ProbS) and a diversity-focused algorithm (HeatS) to improve aggregate recommendation diversity. According to authors, such collaborations can yield best results balancing both accuracy and diversity, without relying on any semantic- or context-specific information. Given a user–resource bigraph, and suppose o_i and o_u to represent, respectively, the number of users who have collected resource i and the number of resources collected by user u. The ProbS+HeatS algorithm works by assigning items an initial level of "*resources*" (generally, it can be seen as "*energies*") denoted by the vector f (where f_i is the "*energy*" possessed by item i), and then redistributing it via the transformation $\tilde{f} = Wf$, where

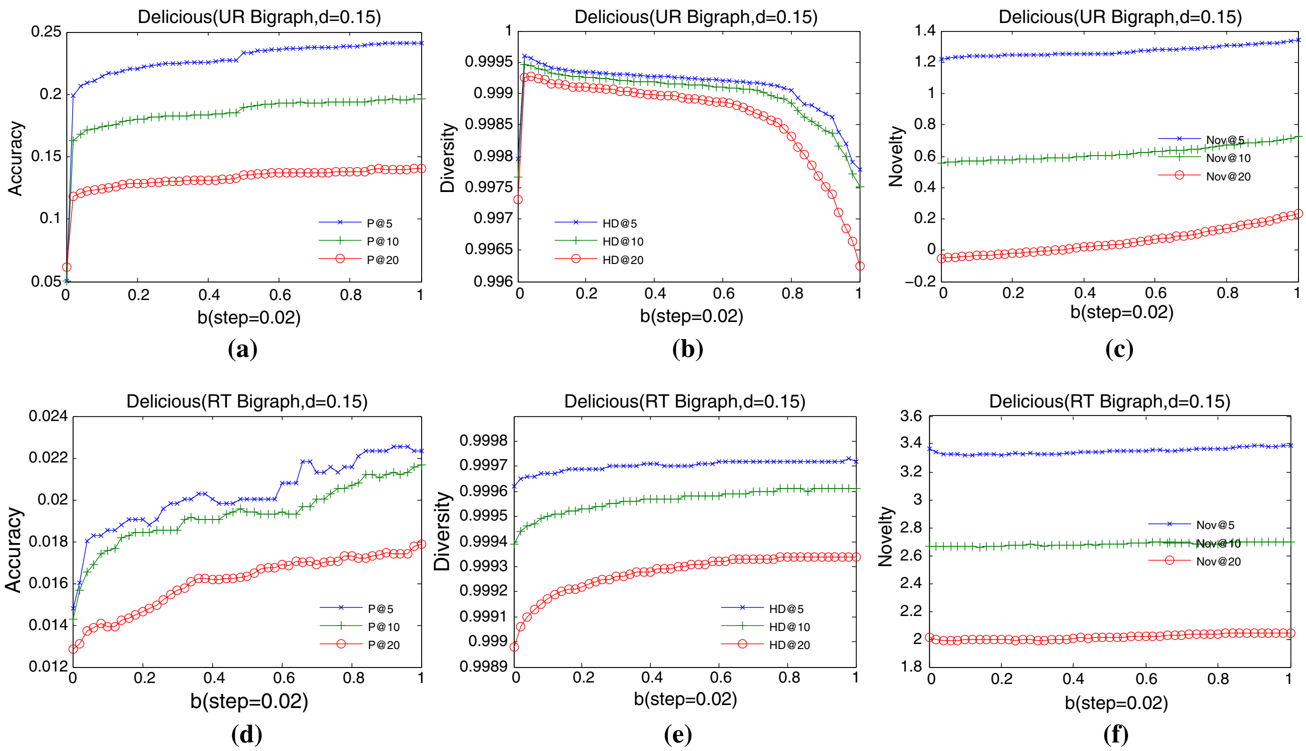

Fig. 6 Recommendation performance of the DaPPR as varying balance factor $b \in [0, 1]$ on the Delicious dataset

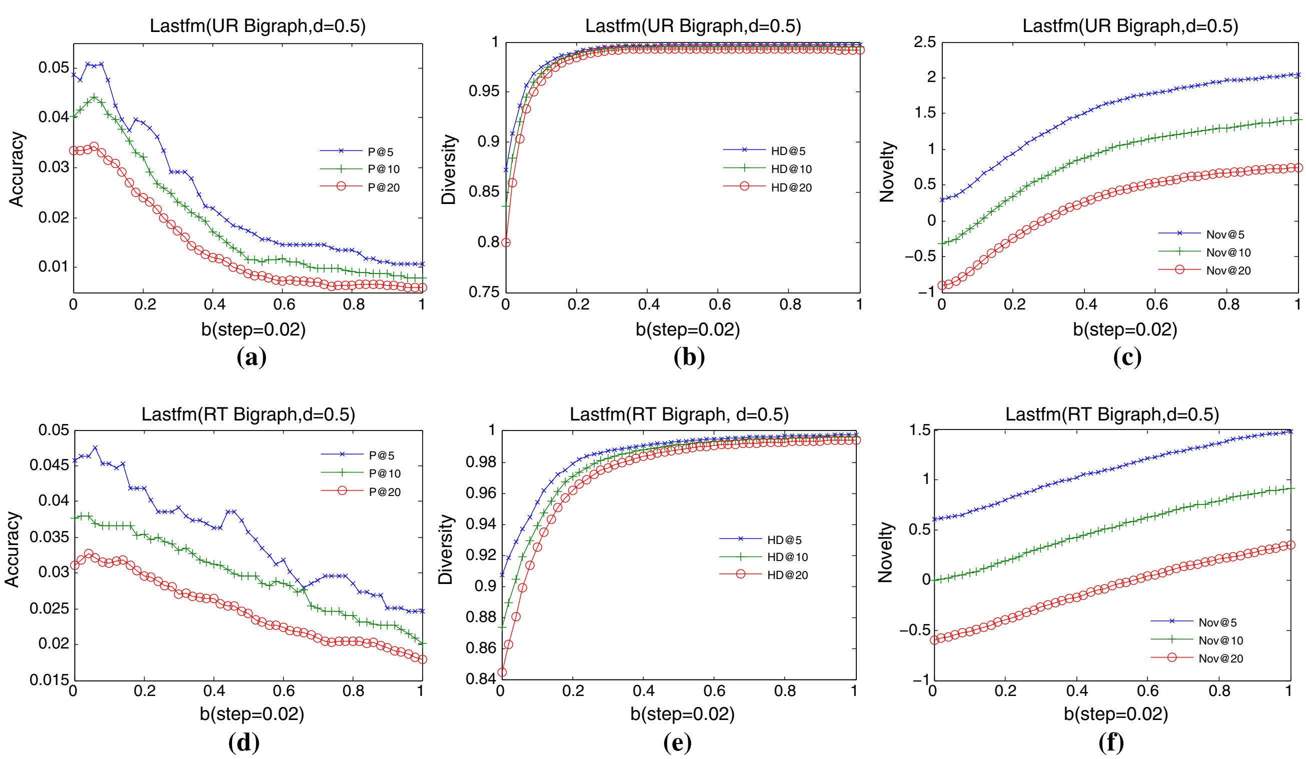

Fig. 7 Recommendation performance of the DaPPR as varying balance factor $b \in [0, 1]$ on the Lastfm dataset

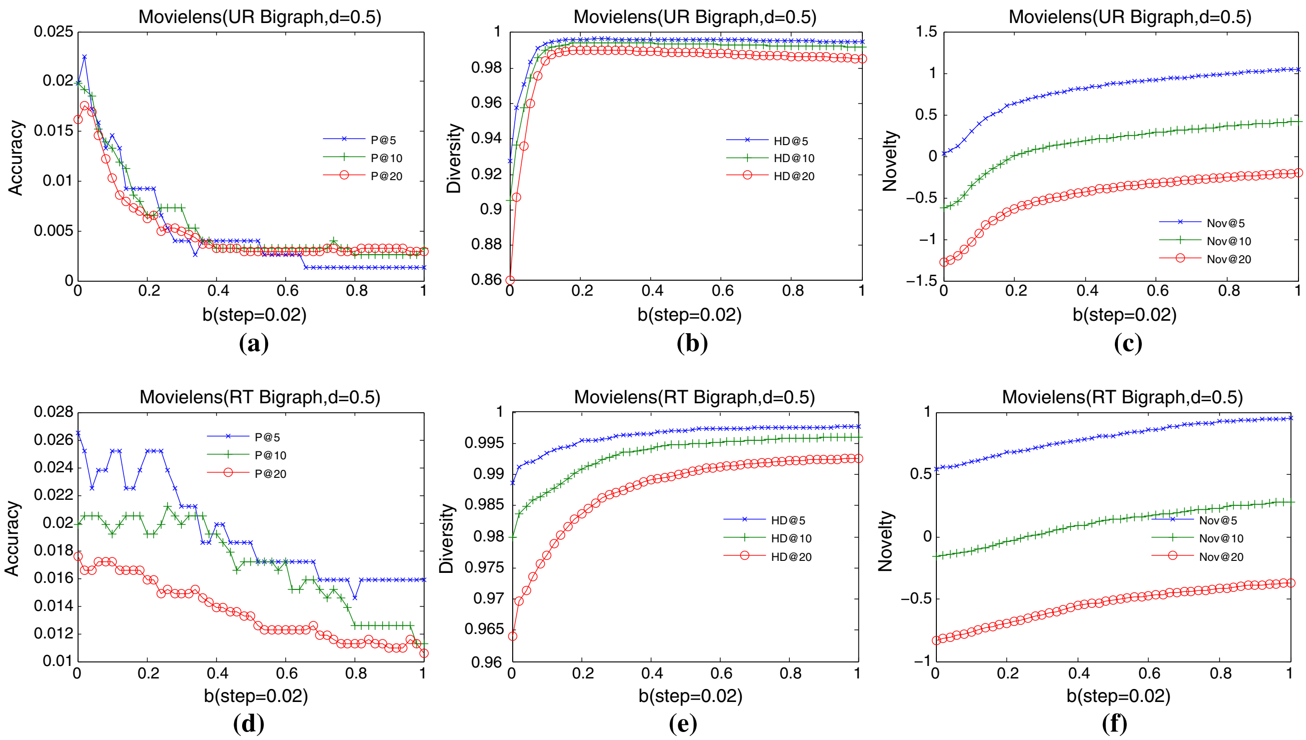

Fig. 8 Recommendation performance of the DaPPR as varying balance factor $b \in [0, 1]$ on the MovieLens dataset

$$W_{ij} = \frac{1}{o_i^b o_j^{1-b}} \sum_u \frac{A_{iu} A_{ju}}{o_u} \tag{8}$$

is a row-normalized $n \times n$ matrix(n is the total number of items in the training set). The adjacency matrix A corresponds to the user–resource bigraph, where $A_{iu} = 1$ if item i is collected by the user u, and $A_{iu} = 0$ otherwise. b is taken as a balance factor, where $b = 1$ gives us the pure HeatS algorithm, and $b = 0$ gives us pure ProbS. Recommendations for a given user u are obtained by setting the initial energy vector f in accordance with the items the user has already collected, that is, by setting $f_i = A_{iu}$. The resulting recommendation list of uncollected items is then sorted according to $\tilde{f}_j$ in descending order. This method can be easily adapted to the resource–tag bigraph by replacing user nodes with tag nodes. Accordingly, we can build a hybrid model by incorporating two processes of ProbS+HeatS on both bigraphs in the same way as the hybrid DaPPR (shown as Eq.9).

$$\tilde{f} = \lambda \tilde{f}_{UR} + (1 - \lambda) \tilde{f}_{RT} \tag{9}$$

To make a quantitative comparison of selected algorithms, we define an indicator-*improvement(L)*, to measure how a new method improves the performance of the baseline. Given a ranking list with length L, *improvement(L)* is defined as following,

$$improvement(L) = \frac{1}{L} \sum_{k=1}^{L} \frac{v^k - v^k_{base}}{abs(v^k_{base})} \tag{10}$$

where v^k and v^k_{base} represent the value of performance metric at the ranking position k for the new method and the baseline, respectively. Improvement averages the gain over L and can be used to estimate the gain as a whole of different performance metrics. Generally, a smaller L gives more meaningful outcomes, since the final results do not differ significantly with the increment of L. A value of $L = 20$ is chosen for the results displayed here to reflect the probable length of a practical recommendation list.

For three hybrid methods, the optimal configuration of λ (which would result in the best accuracy) for each dataset is estimated. We surprisedly find that the hybrid ProbS+HeatS almost shares the same optimal settings of λ with the hybrid DaPPR, and the setting of b of the UR bigraph is also in accord with that of the RT bigraph; consequently, we let b in both cases take the same values.

We first investigate the experimental results against the Delicious dataset. As showed in Fig. 9, both the hybrid ProbS+HeatS and the hybrid DaPPR perform much better than the hybrid PPR in the accuracy metric, and the accuracy continues to grow as b increases. There are no improvements for both two algorithms in terms of the diversity, because the average diversity of top-20

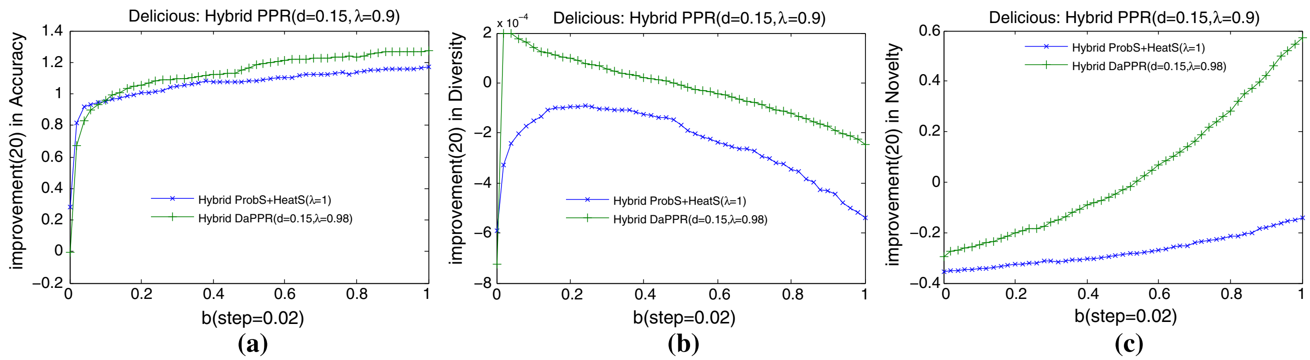

Fig. 9 Comparisons of selected algorithms based on the Delicious dataset

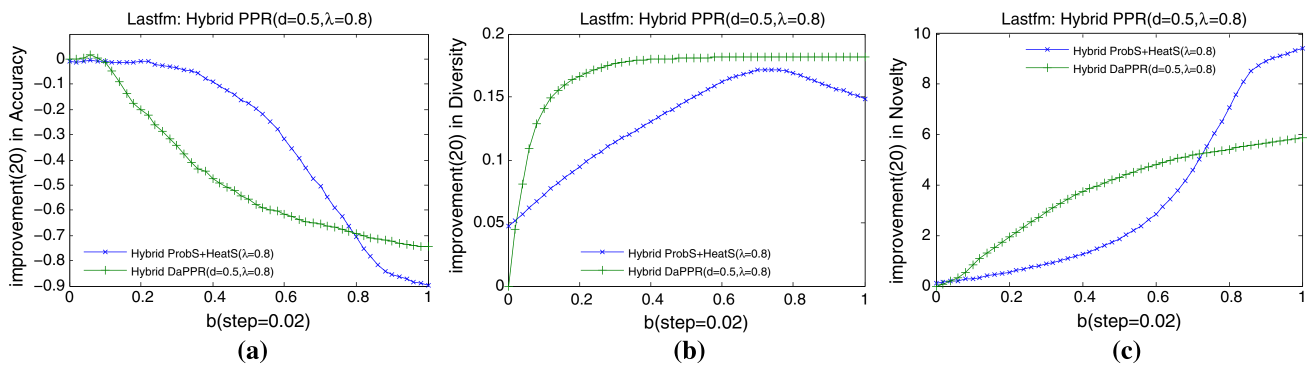

Fig. 10 Comparisons of selected algorithms based on the Lastfm dataset

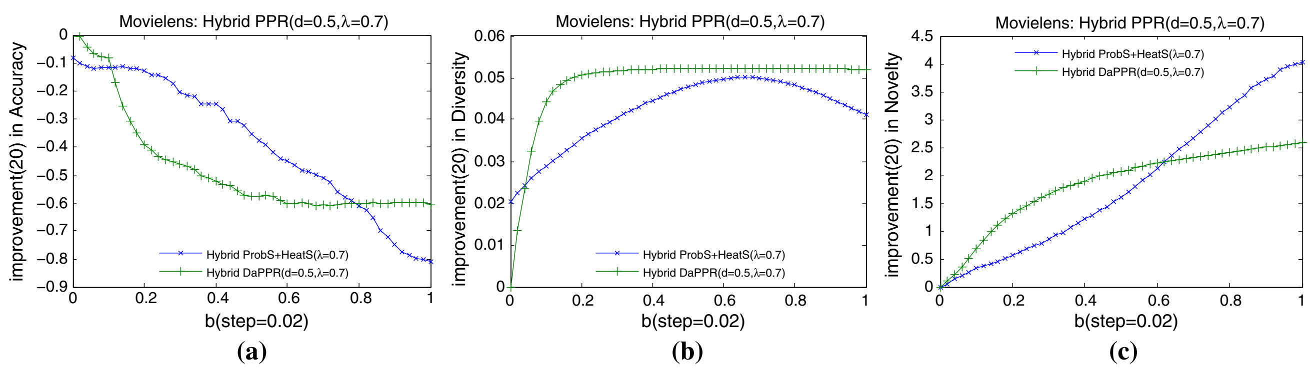

Fig. 11 Comparisons of selected algorithms based on the MovieLens dataset

recommendations of the hybrid PPR has reached as high as 99.95%; however, the losses of the diversity are so tiny (the value is less than one-thousandth of the baseline) that we can think the diversity remains unchanged. In term of the novelty, the hybrid DaPPR performs much better than the hybrid ProbS+HeatS as b takes 1.

We next see the experimental results based on both the Lastfm dataset and the MovieLens dataset. According to Figs. 10 and 11, both the hybrid DaPPR and the hybrid ProbS+HeatS possess the capability to improve the diversity and the novelty while maintaining the accuracy of recommendation. However, the potential trade-offs between the accuracy and the diversity/novelty of recommendation for both algorithms are quite different. In this sense, it is difficult to compare the two algorithms. If we consider an ideal trade-off as the case where a minimized loss of accuracy can be achieved as b increases, for the hybrid DaPPR, such a trade-off can be achieved as the balance factor goes around 0.1 in both datasets, while for the hybrid ProbS +HeatS, it can be obtained as b goes around 0.2. We present these trade-offs of two algorithms in Table 2. By this, we can find the hybrid DaPPR

Table 2 The ideal trade-offs of recommendation accuracy and diversity/novelty of two selected algorithms

Delicious		Improvement (20)	
Baseline:Hybrid PPR(d $= 0.15$, $\lambda = 0.9$)	Accuracy (%)	Diversity (%)	Novelty (%)
Hybrid ProbS+HeatS($\lambda = 1$, $b = 1$)	+117	−0.05	−13.9
Hybrid DaPPR($d = 0.15$, $\lambda = 0.98$, $b = 1$)	+127	−0.02	+57.3
Lastfm		Improvement(20)	
Baseline:Hybrid PPR($d = 0.5$, $\lambda = 0.8$)	Accuracy (%)	Diversity (%)	Novelty (%)
Hybrid ProbS+HeatS($\lambda = 0.8$, $b = 0.2$)	−1.0	+9.9	+60.5
Hybrid DaPPR($d = 0.5$, $\lambda = 0.8$, $b = 0.1$)	−1.2	+14.1	+81.2
MovieLens		Improvement (20)	
Baseline:Hybrid PPR($d = 0.5$, $\lambda = 0.7$)	Accuracy (%)	Diversity (%)	Novelty (%)
Hybrid ProbS+HeatS($\lambda = 0.7$, $b = 0.18$)	−11.8	+3.4	+51.4
Hybrid DaPPR($d = 0.5$, $\lambda = 0.7$, $b = 0.1$)	−8.0	+4.4	+67.4

outperforms the hybrid ProbS+HeatS, since the former generates more diverse and novel recommendations under comparable levels of accuracy. Note that the definition of ideal comprises is not rigorous. Both the proposed algorithms need to be tailored to different custom situations and requirements.

In addition, there is one point worth emphasizing that the marginal improvements of diversity metric do not reflect all nature of things. As b takes 0.1, the hybrid DaPPR seems merely to improve the baseline method by 14.1% on the Lastfm dataset and 4.4% on the MovieLens dataset. However, things are quite different if we consider the total number of distinct resources (NDR, same as *diversity-in-top-N* defined in the work [1]) recommended to all users. In this case, baseline method returns totally 1200 items to 358 users in the Lastfm dataset, and 1109 items to 151 users in the MovieLens dataset. In contrast, the hybrid DaPPR returns, respectively, 2304 items and 1506 items, and the corresponding gains reach 92.0% and 35.8%. Obviously, the gains will be more significant as the number of users increases.

According to all above experiments, although the paradoxes always exist between the accuracy and the diversity (also novelty), our proposed balancing factor is more sensitive and powerful to allow an adequate compromise between two imperatives. On the one hand, either the accuracy or the diversity (novelty) of recommendation can continue to grow as it increases. On the other hand, a small increment of the balance factor (e.g., $b \in (0, 0.1]$) can make the aggregate diversity approach the upper bound, where the diversity is close to 1.

4.4 Experiments with DaPPR in combination with augmented bipartite graphs

In this section, we investigate the recommendation effectiveness of DaPPR in combination with augmented bipartite graphs. Here, the augmented UR bigraph (AUR) is obtained by adding social relations to the users, and the augmented RT bigraph (ART) is built by adding the attributes information to the resources (as shown in Fig. 12). Since attributes can be treated as another kind of tag, adding attributes to resources would not change the characteristics of the RT bigraph. In three experimental datasets, both the Delicious dataset and the Lastfm dataset

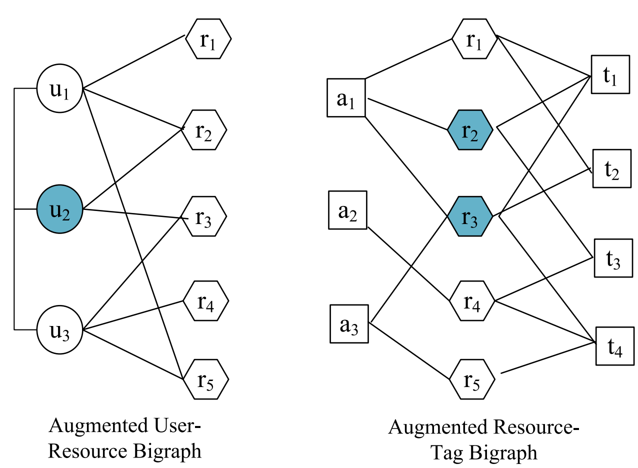

Fig. 12 Augmented user–resource bigraph and resource–tag bigraph

Table 3 The statistics of augmented bipartite graphs in experiments

AUR	Users	Relations	Density
Delicious	1,867	15,328	$8.7 * 10^{-4}$
Lastfm	1,892	25,434	$3.5 * 10^{-3}$
ART	Resources	Attributes	Density
MovieLens	5,908	Actors Genres Directors	$4.2 * 10^{-4}$

provide social relations among users, while the MovieLens dataset provides rich attribute information for the movies. Statistics of augmented bipartite graphs are presented in Table 3, where three kinds of attributes: actors, genres and directors, are specially used for movies.

We experiment the DaPPR in the AUR bigraph based on both the Delicious dataset and the Lastfm dataset and perform the DaPPR in combination with the ART bigraph according to the MovieLens dataset. For simplicity, we fix the settings of damping factor (d) as previous and select the balance factors (b) for three datasets considering the situation when the best accuracy of recommendation is observed in accordance with Figs. 6a, 7a and 8d. The experimental results are summarized as Table 4.

We firstly analyze the performance of recommendation in accuracy metric. For both the Delicious dataset and the MovieLens dataset, the recommendation accuracy of the PPR (recall that the PPR is a special case of the DaPPR where b always equals to 0) in the augmented bigraph is much worse than that of the basic bigraph; however, improvements on recommendation accuracy are observed in the Lastfm dataset, where the PPR with the AUR bigraph outperforms the basic one in P@1-5, in particular, the gain reaches 20.9% in P@1 metric (such a gain reaches 25.1% w.r.t the DaPPR). As grounding on the PPR, the performance DaPPR has the same tendency to that of the PPR. The reason causes this contradictory result may lie in the characteristics of datasets. In the Delicious dataset, social relationships between users are much more sparse than that of the Lastfm dataset(see Table 3); hence, a user may be affected more by those indirectly connected users of like-mind than by his/her straightly connected friends. In contrast, a user in the Lastfm would be more likely to consult his/her friends when searching resources, consequently adding social relations brings positive effects. For the MovieLens dataset, the attributes information of the movies does not necessarily explain the similarities between the movies, for example, some films may belong to different genres (categories), even if they share an actor or a director; therefore, adding such information cannot increase the probability of success in finding similar movies.

We secondly examine the performance of recommendation in diversity metric. For all the datasets, the recommendation diversity is leveraged on augmenting the basic bipartite graphs, except the Delicious dataset, where the total number of distinct resources recommended to all users

Table 4 Recommendation performance of the DaPPR with basic bigraphs and augmented bigraphs, where NDR@20 is short for the total number of distinct resources recommended to all users

	Accuracy				Diversity			Novelty	
	P@1	P@5	P@10	P@20	HD@10	HD@20	NDR@20	Nov@10	Nov@20
Delicious($d = 0.15$)									
PPR(UR, $b = 0$)	0.0564	0.0501	0.0506	0.0613	0.9976	0.9973	9619	0.5549	−0.051
PPR(AUR, $b = 0$)	0.0363	0.0346	0.0351	0.0401	0.9978	0.9975	9086	1.0057	0.4038
DaPPR(UR, $b = 1$)	***0.2857***	***0.2413***	***0.1965***	***0.1405***	0.9975	0.9962	***10132***	0.7243	0.2332
DaPPR(AUR, $b = 1$)	0.1817	0.1446	0.1135	0.0759	***0.9984***	***0.9982***	9838	***1.1953***	***0.7610***
Lastfm ($d = 0.5$)									
PPR(UR, $b = 0$)	0.0670	0.0486	0.0402	0.0334	0.8319	0.7997	970	−0.3140	−0.9070
PPR(AUR, $b = 0$)	*0.0810*	*0.0497*	0.0399	0.0318	0.8937	0.8589	1396	−0.0580	−0.6687
DaPPR(UR, $b = 0.06$)	0.0726	0.0503	***0.0441***	***0.0344***	***0.9454***	***0.9328***	***1735***	−0.1895	−0.7857
DaPPR(AUR, $b = 0.06$)	***0.0838***	***0.0525***	0.0394	0.0324	0.9322	0.9161	1695	***−0.0002***	***−0.5994***
MovieLens($d = 0.5$)									
PPR(RT, $b = 0$)	***0.0331***	***0.0264***	***0.0199***	***0.0175***	0.9800	0.9641	1355	−0.1621	−0.8296
PPR(ART, $b = 0$)	0.0265	0.0159	0.0146	0.0136	0.9920	0.9849	1688	0.4058	−0.3628
DaPPR(RT, $b = 0.1$)	***0.0331***	0.0252	0.0192	0.0172	0.9871	0.9770	1456	−0.1160	−0.7638
DaPPR(ART, $b = 0.1$)	0.0265	0.0146	0.0146	0.0139	***0.9921***	***0.9853***	***1698***	***0.4093***	***−0.3552***

Also, the best results are marked in bold-italic fonts

(NDR@20) decreases as adding social relations to the UR bigraph.

In the next, as far as the novelty of recommendation concerned, the novelty metrics are significantly improved by using augmented bigraphs for both the PPR and the DaPPR. It means that returned resources are quite different from the resources, which have been previously seen by target users. The result also implies that exploiting augmented bipartite graphs is useful to yield surprising recommendations.

Based on these observations, we can conclude that augmenting bipartite graphs helps to improve the diversity and the novelty of recommendations, however, not necessarily enhance the recommendation accuracy. For a recommender system, the primary objective is to provide accurate recommendations to users. Consequently, the use of basic folksonomy schema may be sufficient in most cases. To generate diverse and novel recommendations, we can adjust only the balance factor of the DaPPR.

5 Related works

As for the diversity of recommender systems, it can be further classified as individual or aggregate diversity. Individual diversity means dissimilarity among items suggested to an individual user. Aggregate diversity means dissimilarity among items being recommended across all users under considerations [24].

There have been many works addressing the individual diversity as well as keep the accuracy. Bradley & Smyth [4] proposed three algorithms for improving individual diversity, which concerned content-based recommendation techniques. According to experiments, their Bounded Greedy Selection algorithm has greatly reduced the retrieval cost and caused minimal loss of similarity among target query and recommendations. Ziegler et al. [34] proposed topic diversification, a novel method designed to balance and diversify personalized recommendation lists to reflect the user's complete spectrum of interests. The key technology is using classification systems to quantify the similarity between two product sets, forms an essential part of topic diversification. The authors also propose intra-list similarity, a new metric which is well suited to capture the diversity using the proposed algorithm. Their experimental results are demonstrated that users preferred the altered diversified list even some loss of accuracy occurred, than the unaltered accurate list. Following Ziegler et al, many works seek to find out the best diverse subset of items to be recommended with minimal loss of accuracy. Hurley & Zhang [15] proposed such an approach, in which resultant list's similarity to target query and diversity within list, is taken as a binary optimization problem. A recent evaluation metric, item novelty, is also proposed measuring how much an item is different from existing items in the user profile. By adjusting the novelty value, the tolerance in accuracy loss is balanced. Wang & Zhu [30] introduced an optimization method to increase intra-diversity among items grounding on the well-known Portfolio Theory. In addition, Tong et al. [28] built an optimization method for diversified ranking in graph based on the personalized PageRank. A greedy-based model is proposed to increase dissimilarity among items ranked to a single unique user. Beyond those methods which diversify the subset of items to be recommended to users by considering the content dissimilarity, some works specially use other contextual information to diversify recommendations. Servajean et al.[25] firstly investigated profile diversity to address the problem of returning highly popular but too-focused documents. They adopt Fagin's threshold-based algorithms to return the most relevant and most popular documents that satisfy content and profile diversities. Preliminary experiments showed that exploiting profile diversity is promising to improve individual diversity. Lathia et al. [19] evaluated three CF algorithms from the point of view of the diversity in the sequence of recommendation lists they produce over time and examined how a number of characteristics of user rating patterns (including profile size and time between rating) affect diversity. Experiments have proved that temporal diversity is an another important facet of recommender systems.

Few works are presented with respect to aggregate diversity. Zhou et al. [33] developed an approach that combines Probabilistic Spreading algorithm (ProbS) and Heat Spreading algorithm (HeatS). The ProbS and the HeatS, respectively, simulate mass diffusion and heat diffusion in the physics. As the ProbS is accuracy focused while the HeatS is diversity focused, the hybrid spreading model combines them together can yield best Top-N results balancing both accuracy and diversity, without relying on any semantic- or context-specific information. The principles behind the ProbS and the HeatS are similar to the PageRank, since both of them derive from the random walk model. However, the model of ProbS+HeatS is not equal to our DaPPR model according to above-mentioned experiments. Moreover, both the ProbS and the HeatS are specifically designed to explore the singleton relation of bipartite graph; in contrast, the PPR can be easily extended to take advantages of rich semantic information. Adomavicius & Kwon [1] designed some ranking-based techniques that can improve aggregate diversity of recommendation lists across all users. To balance the aggregate accuracy and diversity, they formed a parameterized function for item-based popularity ranking approach, through which the level of accuracy and diversity to be maintained is controlled. They conducted online experiments by different rating prediction techniques in combination with the proposed seven ranking-based techniques. The analysis shows that the popularity of items can be a good tool to enhance the diversity without or with limited losses of the accuracy. Gan & Jiang [9] proposed a method called PLUS to adjust user similarities for collaborative filtering models using a power function. According to large-scale validation experiments, PLUS achieves a reasonable trade-off between recommendation accuracy and diversity and is robust to similarity measures. Also, Niemann & Wolpers [23] proposed a new collaborative filtering approach that is based on the items' usage contexts. The way increases the rating predictions for niche items with fewer usage data available and improves the aggregate diversity of the recommendations. Compared with our method, most of the works presented here are based on the collaborative filtering model, which performs well on rating-based recommender systems with more dense data and hence are fundamentally different with us. In addition, FolkRank [17] that achieves outstanding performance on recommending tags can be seen as a special case of the DaPPR where b always takes 1. However, for many application systems, a compromise between diversity and accuracy as the balance factor takes 1 is not meaningful, since the recommendations may be far from the preferences of users.

The individual diversity essentially differs from the aggregate diversity, since balancing of individual recommendation list does not mean a balance fulfilled in the aggregate view [1]. Also, the optimization method has been proposed to find out the best diverse subset of an individual recommendation list are usually time-consuming; therefore, any attempt to enhance the aggregate diversity using optimization in the individual view would be a terrible thing for a large-scale recommender system. For balancing the aggregate accuracy and diversity, a simple yet efficient algorithm would always be a right choice. In this sense, our works provide a promising solution for solving the accuracy–diversity dilemma of recommender system.

6 Conclusions

We have proposed the DaPPR for resources recommendation in folksonomy-based social systems. The method is simple yet effective, since it allows not merely a compromise between the two imperatives but also allows us to simultaneously increase both accuracy and diversity of recommendations. Compared with existing works, the DaPPR can also yield better results balancing both accuracy and diversity (or novelty). The DaPPR can potentially enhance recommendation effectiveness in combination with more rich contextual information, particularly, provide diverse and surprising recommendation results to users. However, due to the complexity of the application systems, how to improve the ability of recommender systems, to achieve a better balance between the accuracy and the diversity (novelty) of recommendations, is still a difficult task. Other methods (e.g., collaborative filtering and rank aggregation [31]) need to be taken into account to form advanced hybrid models; accordingly, the DaPPR can be taken as an excellent component to play to its strengths. Moreover, since the technical challenges facing recommender systems involve similar paradoxes [33], the DaPPR can be used to not only folksonomy-based systems but also other recommender systems. For example, in cloud computing environment [20], the context associated with cloud services can be abstracted as a graph similar to folksonomy schema; our approach yet can be used to find diverse services to users.

Acknowledgments This work was supported by the Applied Basic Research Project of Yunnan Province (No.2013FB009) and the Scientific Research Project of Yunnan University (No.2010YB024), Special Funds for "Middle-aged and Young Core Instructor Training Program" of Yunnan University, the National Natural Science Foundation of China (No.61070013, No.U1135005), and "Hundred Talents Recruitment Program" of Global Experts of Hubei. This work of Jun He was supported by the National Natural Science Foundation of China (No.61203273). We are grateful to anonymous reviewers for their useful comments and suggestions which contribute to substantially improving this paper.

References

1. Adomavicius G, Kwon Y (2012) Improving aggregate recommendation diversity using ranking-based techniques. IEEE Trans Knowl Data Eng 24(5):896–911
2. Bellogin A, Cantador I, Castells P (2013) A comparative study of heterogeneous item recommendations in social systems. Inf Sci 221:142–169
3. Boldi P, Santini M, Vigna S (2009) Pagerank: functional dependencies. ACM Trans Inf Syst 27:19
4. Bradley K, Smyth B (2001) Improving recommendation diversity. In: Proceedings of the twelfth national conference in artificial intelligence and cognitive science, 2001. Springer, pp 75–84
5. Brin S, Page L (1998) The anatomy of a large-scale hypertextual web search engine. Comput Netw 30:107–117
6. Cantador I, Brusilovsky I, Kuflik T (2011) 2nd workshop on information heterogeneity and fusion in recommender systems (hetrec 2011). In: Proceedings of the fifth ACM conference on recommender systems, ACM, pp 387–388
7. Fleder D, Hosanagar K (2009) Blockbuster cultures next rise or fall: the impact of recommender systems on sales diversity. Manag Sci 55:697–712
8. Fouss F, Pirotte A, Renders JM, Saerens M (2007) Random-walk computation of similarities between nodes of a graph with application to collaborative recommendation. IEEE Trans Knowl Data Eng 19:355–369
9. Gan M, Jiang R (2013) Improving accuracy and diversity of personalized recommendation through power law adjustments of user similarities. Decis Support Syst 55:811–821

10. Golder S, Huberman B (2006) Usage patterns of collaborative tagging systems. J Inf Sci 32:198–208
11. Gupta M, Li M, Yin Z, Han J (2010) Survey on social tagging techniques. ACM SIGKDD Explor 12:58–72
12. Haveliwala T (2003) Topic-sensitive pagerank: a context-sensitive ranking algorithm for web search. IEEE Trans Knowl Data Eng 15:784–796
13. Herlocker J, Konstan J, Terveen L, Riedl J (2004) Evaluating collaborative filtering recommender systems. ACM Trans Inf Syst 22:5–53
14. Huang Z, City J, Zeng D, Chen H (2007) A comparison of collaborative-filtering recommendation algorithms for e-commerce. IEEE Intell Syst 22:68–78
15. Hurley N, Zhang M (2011) Novelty and diversity in top-n recommendation analysis and evaluation. ACM Trans Int Technol 10:14
16. Jannach D, Lerche L, Gedikli F, et al (2013) What recommenders recommend-An analysis of accuracy, popularity, and sales diversity effects. In: Proceedings of the 21st conference on user modeling, adaptation and personalization, Springer, pp 25–37
17. Jaschke R, Marinho L, Hotho A, Schmidt-Thieme L, Stumme G (2008) Tag recommendations in social bookmarking systems. AI Commun 21:231–247
18. Konstas I, Stathopoulos I, Jose J (2009) On social networks and collaborative recommendation. In: Proceedings of the 32nd international ACM SIGIR conference on research and development in information retrieval, ACM, pp 195–202
19. Lathia N, Hailes S, Capra L, Amatriain X (2010) Temporal diversity in recommender systems. In: Proceedings of the 33rd international ACM SIGIR conference on research and development in information retrieval, ACM, pp 210–217
20. Liu J, Zhou J, Wang JF, Chen X, Zhou HN (2013) Cloud service: automatic construction and evolution of software process problem-solving resource space. J Supercomput 64(3):1108–1132
21. Luo XF, Xu Z, Yu J, Chen X (2011) Building association link network for semantic link on web resources. IEEE Trans Autom Sci Eng 8(3):482–494
22. McGinty L, Smyth B (2003) On the role of diversity in conversational recommender systems. In: Proceedings of the 5th international conference on case-based reasoning research and development, Springer, pp 276–290.
23. Niemann K, Wolpers M (2013) A new collaborative filtering approach for increasing the aggregate diversity of recommender systems. In: Proceedings of the 19th ACM SIGKDD international conference on knowledge discovery and data mining, ACM, pp 955–963
24. Patil CB, Wagh RB (2013) Recommendation diversity for web personalization: a survey. Int Mag Adv Comput Sci Telecommun 4:15–18
25. Servajean M, Pacitti E, Amer-Yahia S, et al (2013) Profile diversity in search and recommendation. In: Proceedings of the 22nd international conference on World Wide Web companion, ACM, pp 973–980
26. Smyth B, McClave P (2003) Similarity vs. diversity. In: Proceedings of the 3th International conference on case-based reasoning research and development, Springer, pp 347–361
27. Sun YC, Yan HL, Lu C, Bie RF, Zhou ZB (2013) Constructing the web of events from raw data in the web of things. J Mob Inf Syst. doi:10.3233/MIS-130173
28. Tong HH, He JR, Wen Z, Konuru R, Lin C (2011) Diversified ranking on large graphs: an optimization viewpoint. In: Proceedings of the 17th ACM SIGKDD international conference on knowledge discovery and data mining, ACM, pp 1028–1036
29. Vargas S, Castells P (2011) Rank and relevance in novelty and diversity metrics for recommender systems. In: Proceedings of the fifth ACM conference on recommender systems, 2011. ACM, pp 109–116
30. Wang J, Zhu J (2009) Portfolio theory of information retrieval. In: Proceedings of the 32nd international ACM SIGIR conference on research and development in information retrieval, ACM, pp 115–122
31. Wu H, Yu H, Li B, Pei YJ (2013) Personalized recommendation via rank aggregation in social tagging systems. In: Proceedings of the 10th international conference on fuzzy systems and knowledge discovery, IEEE, pp 912–916
32. Zhang ZK, Zhou T, Zhang YC (2011) Tag-aware recommender systems: a state-of-the-art survery. J Comput Sci Technol 26:767–777
33. Zhou T, Kuscsik Z, Liu JG, Medo M, Wakeling J, Zhang YC (2010) Solving the apparent diversity-accuracy dilemma of recommender systems. Proc Nat Acad Sci 107:4511–4515
34. Ziegler CN, McNee SM, Konstan JA, Lausen G (2005) Improving recommendation lists through topic diversification. In: Proceedings of the 14th international conference on World Wide Web, ACM, pp 22–32