2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology

Improving Collaborative Filtering in Social Tagging Systems for the Recommendation of Scientific Articles

Denis Parra-Santander, Peter Brusilovsky
School of Information Sciences
University of Pittsburgh
Pittsburgh, PA, USA
dap89@pitt.edu, peterb@pitt.edu

Abstract—**Social tagging systems pose new challenges to developers of recommender systems. As observed by recent research, traditional implementations of classic recommender approaches, such as collaborative filtering, are not working well in this new context. To address these challenges, a number of research groups worldwide work on adapting these approaches to the specific nature of social tagging systems. In joining this stream of research, we have developed and evaluated two enhancements of user-based collaborative filtering algorithms to provide recommendations of articles on CiteULike, a social tagging service for scientific articles. The result obtained after two phases of evaluation suggests that both enhancements are beneficial. Incorporating the number of raters into the algorithms, as we do in our NwCF approach, leads to an improvement of precision, while tag-based BM25 similarity measure, an alternative to Pearson correlation for calculating the similarity between users and their neighbors, increases the coverage of the recommendation process.**

Keywords-social tagging; recommender systems; information retrieval; collaborative filtering

I. INTRODUCTION

The new generation of Web 2.0 systems – also known as Social Web systems – presented another challenge to researchers and practitioners in the area of recommender systems. An already broad stream of new information created by owners and developers of Web sites and information systems was joined by another stream of information produced by the users of various kinds of social systems – from user articles in various blogs and wiki sites, to shared bookmarks, pictures, and movies on social bookmarking and tagging sites, to a range of information about users themselves on social linking sites. The need for proactive recommendation is arguably high in social systems. Not only is the volume of user-contributed information potentially much larger, but traditional information access infrastructure (such as indexes, directories, information maps) is typically less advanced than in traditional systems. While sometimes considered a luxury in classic information systems [2], personalization has become a necessity in social Web systems.

A number of research groups worldwide have already started to explore classic personalization techniques in this new context; however, early results demonstrate that a mere reuse of old techniques is not an efficient way forward as these technologies may not work in the new contexts as efficiently as they worked in traditional information systems [1; 3; 11; 15]. We experienced this problem ourselves while attempting to transfer a specific search and browsing personalization approach known as ASSIST from the traditional to social information access context. While personalization techniques showed promise in association with the ACM Digital Library project [7], this technology failed to deliver the expected value in the context of finding and recommending YouTube videos [3]. We believe that a significant amount of new research is required to produce efficient personalized information access technologies for the new context.

This paper explores the problem of personalization in a specific kind of social systems known as collaborative tagging systems. The systems of this kind assembled a large volume of user-contributed items, such as Web bookmarks in Delicious, pictures in Flickr, and bibliographic references in CiteULike. However, by the nature of these systems, they lack any kind of centrally provided description, metadata or hierarchical categorization as in more traditional Web systems (i.e., online stores, Web directories, library catalogs). Each contributed item may include user-contributed tags and comments instead. Since the user-driven information access technologies in this context are limited to tag-based browsing and search, social tagging system become an attractive platform for the application of recommender approaches of technologies such as content-based [16] and collaborative filtering [17].

The importance of providing personalized recommendations in collaborative tagging systems has been recognized by many researchers. Over the last several years a number of content-based [8; 14], collaborative [1], and hybrid [19; 20; 21] recommenders for various collaborative tagging systems have been developed and evaluated. As shown by the experience of these pioneer works, collaborative tagging systems do require innovative ideas for both major recommendation approaches. Large volumes of

978-0-7695-4191-4/10 $26.00 © 2010 IEEE
DOI 10.1109/WI-IAT.2010.261

content contributed by diverse users challenged both content-based and collaborative approaches. Content-based approaches were challenged by the low volume and consistency of item descriptions, while collaborative approaches had to deal with a much sparser user feedback. Moreover, user ratings (which are important for quality recommendation) were typically not available – only the fact that an item was contributed or bookmarked by the user was present in the majority of systems. To compensate for the loss of quality content and weaker feedback, a number of pioneer projects in this field creatively exploited the presence of tags, which is a distinguishing feature of collaborative tagging systems. Content-based recommenders explored the use of tags as item descriptors, frequently applying various dimensionality-reduction approaches to deal with tag inconsistency [5; 18]. Collaborative recommender systems explored the use of tags to more reliably calculate similarity between users [4; 6; 20; 21].

The work presented in the paper attempts to further expand the research on tags-aware recommender systems. Our primary focus is improving tag-aware collaborative recommendation approaches. Our experience with several collaborative tagging systems convinced us that the large volume of content and opportunistic nature of tagging undermine the assumptions of traditional collaborative filtering Two users who are very similar in their interests frequently have too few common items bookmarked, which prevents a traditional system from recognizing them as neighbors for user-based collaborative filtering. We agree with the authors cited above that, in this context, the similarity of tags applied by users may provide a more reliable evidence of their interest similarity. Matching user tags may help a collaborative filtering system to identify more neighbors for each user and as a result to generate more precise and more complete recommendation as was attempted in the cited papers. However, we also believe that non-centralized user-driven nature of tag assignment makes tag-based similarity inherently noisier than classic item-based similarity. i.e., while using tag similarity should increase a chance of a collaborative filtering system to identify more "true" neighbors, it should also produce a good fraction of superficial neighbors, i.e., noise. In this situation, the first (neighbor-finding) stage of a tag-aware collaborative filtering algorithm should apply more sophisticated noise-reducing approaches than classic TF*IDF and vector cosine similarity used in a number of recent tag-aware recommenders [14]. Moreover, we also believe that noise-cancelling approaches should also be applied on the second (prediction) stage to decrease the impact of irrelevant papers associated with superficially matched users.

In our recent work we focused on two specific noise-cancelling approaches – one on the neighbor-finding stage and one on the prediction stage. As an alternative to the traditional item-based Pearson correlation approach, we used a tag-matching approach based on Okapi BM25 [13] – one of the best known keyword-matching technologies in the field of information retrieval. As an alternative to the traditional prediction approach we suggested its modification, Neighbor-weighted Collaborative Filtering (NwCF), which takes into account the number of neighbors recommending each item to reduce the prediction noise. We believe that in the context of collaborative tagging systems, where each bookmarking action requires an investment of user efforts, the number of users who bookmarked an item can serve as an additional indicator of item quality and relevance (similar to the citation count in academic literature or link count in Web search). Surprisingly enough, we found only one paper, which attempted to use this indicator for improving the quality of recommendation [10].

To assess the feasibility of these approaches we performed a small-scale user study comparing traditional collaborative filtering system with two experimental systems, which have either neighbor-finding or prediction stage or both replaced with our alternative approaches [15]. The study was performed with seven users in the context of CiteULike , a well-known social tagging system that provides a service for storing, organizing and sharing research references. The results of this pilot study provided evidence that both experimental systems work better than classic collaborative filtering in CiteULike context. While the nature and the scale of the pilot study made it impossible to reliable compare all combinations of the suggested approaches, it provided encouragement for further work presented in this paper.

The primary goal of the work presented below was to run a reliable comparison of all combinations of the experimental approaches using standard n-fold-based evaluation approach. To run this evaluation we decided to continue our work with CiteULike and performed a large volume of crawling of CiteULike data. The account of this work is presented in the following order. Section 2 describes the characteristics of the dataset used on this study. Section 3 describes the recommender approaches compared in our study: Classic Collaborative Filtering (CCF), Neighbor-weighted Collaborative Filtering (NwCF) and BM25-based similarity (BM25). In section 4 we describe the performed experiment and we present the results. Section 5 discusses the results. Section 6 summarizes the presented work and lists ideas for future work..

II. DATASETS

For our study we used a dataset consisting of crawling CiteULike for 38 days during June and July of 2009. The characteristics of this dataset are presented in Table I.

TABLE I. DESCRIPTION OF THE DATASET

Item	# instances
Users	5,849
Items	574,907
Tags	139,993
Tagging incidents	2,337,571

Items in CiteULike are mainly references to research articles. A tagging incident is the tuple {u,i,t} representing a user, a posted item, and one of the tags used in a post. Following a common evaluation methodology of recommender systems [1; 12], we filtered out the dataset to keep a p-core, with a p of 20 for users and a p of 2 for articles. It means that each retained user appears in at least 20 posts and that each article has been posted for at least 2 different users. We followed the cleaning procedure suggested in [1]. The characteristics of this filtered dataset are presented in Table II.

TABLE II. CHARACTERISTICS OF THE FILTERED DATASET

Item	# instances
Users	784
Items	26,599
Tags	26,009
Posts	71,413
Tagging incidents	218,930
avg # items per user	91
avg # users per item	2.68
avg # tags per user	88.02
avg # users per tag	2.65
avg # tags per item	7.07

III. ALGORITHMS

User-based collaborative filtering process consists of two steps. The first step is finding the neighborhood of the center user, i.e., a set of the most similar users. The second step consists of using this neighborhood to rank the items to be recommended, and recommend the Top N items. These items are taken from the set of items which the neighbors rated positively, and which the center user has not posted on her library. We implemented Classic Collaborative Filtering (CCF) as the baseline method and two enhancements: BM25-based similarity (BM25) as an alternative of the first step, and Neighbor-weighted Collaborative Filtering (NwCF) as an alternative of the second step.

A. Classic Collaborative Filtering (CCF)

This approach is described in detail in [17]. In the classic CF model, the similarity between two users is calculated using the Pearson correlation over the ratings of their common items. The formula for the Pearson correlation, as stated in [17], is:

$$userSim(u,n) = \frac{\sum_{i \subset CR_{u,n}} (r_{ui} - \bar{r}_u)(r_{ni} - \bar{r}_n)}{\sqrt{\sum_{i \subset CR_{u,n}} (r_{ui} - \bar{r}_u)^2} \sqrt{\sum_{i \subset CR_{u,n}} (r_{ni} - \bar{r}_n)^2}} \quad (1)$$

In the formula, *r* stands for rating, u denotes the center user and n a neighbor. r_{xi} represents the rating given by the user *x* to the item *i*, and is the average rating of the user x over all her items. $CR_{u,n}$ denotes the set of co-rated items between *u* and *n*, being *i* an element in that set. Next, we rank the articles of these users to recommend to the center user, using the formula of predicted rating for user *u* with average adjusts described in [17].

$$pred(u,i) = \bar{r}_u + \frac{\sum_{n \subset neighbors(u)} userSim(u,n) \cdot (r_{ni} - \bar{r}_n)}{\sum_{n \subset neighbors(u)} userSim(u,n)} \quad (2)$$

B. BM25-based Similarity (BM25)

BM25, also known as Okapi BM25, is a non-binary probabilistic model used in information retrieval [13]. It calculates, given a search query, the relevance of each document in a collection. As we try to take advantage of the set of tags of each user, we made two analogies, comparing the tags of the center user with a query, and the set of tags of each neighbor with a document. Based on this, we use BM25 to calculate similarity and thus we obtain her neighborhood. Our proposed BM25-based similarity model is taken from the calculation of the Retrieval Status Value of a document (RSV_d) of a collection given a query [13]:

$$RSV_d = \sum_{t \in q} IDF \cdot \frac{(k_1 + 1) tf_{td}}{k_1((1-b) + b \times (L_d / L_{ave})) + tf_{td}} \cdot \frac{(k_3 + 1) tf_{tq}}{k_3 + tf_{tq}} \quad (3)$$

In our model RSV_d represents the similarity score between the center user (the terms of the query *q*) and one neighbor (the terms of the document *d*). This similarity is calculated as a sum over every tag t posted by the center user. The neighbor *d* is represented as her set of tags with their respective frequencies. L_d is the document length, in our case is the sum of the frequencies of each tag of the neighbor *d*. L_{ave} is the average of the L_d of every neighbor. The term tf_{td} is the frequency of the tag *t* into the set of tags of the neighbor *d*. tf_{tq} represents the frequency of the tag *t* into the query, i.e., the set of tags of the center user. Finally, *k1*, *k3* and *b* are parameters that we set to 1.2, 1.2 and 0.8 respectively, values slightly different from those suggested by default in [13], which gave us the best results in our previous study [15]. After calculating this similarity measure, we choose the top N most similar neighbors, and calculate the ranking of the recommended articles using the formula (2) or (4).

C. Neighbor-weighted Collaborative Filtering (NwCF)

This method enhances the ranking step by taking into account the number of raters represented as *nbr(i)* in the formula (4). It is useful to filter out potentially noisy items, which have been rated by only one or at most two users. In this way, we push up in the recommendation list the items rated by a larger number of neighbors. The new predicted rating is given by

$$pred'(u,i) = \log_{10}(1 + nbr(i)) \cdot pred(u,i) \qquad (4)$$

IV. Experiments

To assess both separate and cumulative impact of the suggested approaches we compared four conditions (CCF, NwCF, CCF+BM25, NwCF+BM25). As stated in section 3, the CCF recommendation process has 2 steps: 1) calculation of user-similarity, and 2) raking the items to be recommended. While NwCF is an alternative for the second step, BM25 is an alternative for the first one, so they could be both used separately (retaining the classic CCF implementation of the other step) or combined, replacing both steps of CCF.

Using the dataset described in section 2, we performed the evaluation using an IR perspective, comparing MAP@10, a modified version of Mean Average Precision (MAP) [13] and User Coverage (UCov) after a 10-fold cross-validation evaluation.

As accuracy metric, we initially considered MAP, which would be calculated by averaging over the average precision (AP) of the list of recommendations for every user. However, a recommender system rarely displays the complete list of possible recommended items to the center user, which can be hundreds or thousands. It typically displays the top N items, so we decided to calculate the AP for each user and cutting at retrieval point 10 (AP@10). The formula can be expressed as:

$$AP@10 = \frac{\sum_{r=1}^{10} (P(r) \times rel(r))}{number \text{ of } relevant \text{ items @ retrieval point 10}} \qquad (5)$$

In the formula, r is the rank, rel() a binary function on the relevance of a given rank, and P() precision at a given cut-off rank. MAP@10 corresponds to averaging AP@10 over a set of users. On the other side, UCov quantifies the percentage of users for whom the system can generate recommendations.

The evaluation is done by a 10-fold cross-validation process, which is visualized in figure 1. First, the dataset is divided into 10 folds, where each fold is assigned randomly 10% of the users (A). The process follows by selecting one fold as the testing set and the remaining 9 folds as the training set (B).

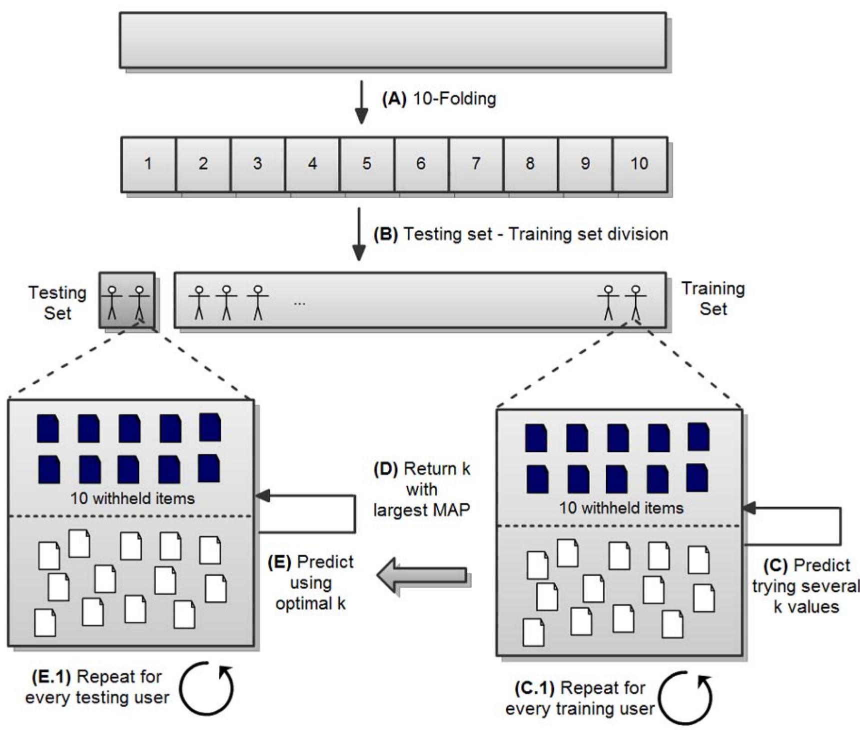

Figure 1. Description of the 10-fold cross-validation process.

The training set is then used to optimize the main parameter in a collaborative filtering algorithm: k, the size of the neighborhood. To calculate the optimal size of the neighborhood, for each user in the training set we withhold 10 articles to be predicted, and we measure the quality of our prediction by calculating its AP@10 (C).

Then, we average over the AP@10 of the whole set of users in the training set obtaining the MAP@10 (D). The neighborhood size (k) with the highest MAP@10 in the training step is used to calculate the MAP@10 of the fold withhold for the testing step (E). We repeat the steps (B), (C), (D) and (E) for every fold, hence, 10 times. Finally, we average over the MAP@10 of each fold to calculate the final MAP@10.

To compare the accuracy of the baseline method CCF against NwCF, BM25, and their combination, we followed the process described in the two previous paragraphs, testing statistical significance of our results with the Wilcoxon paired test of each method compared to the baseline method, CCF. We didn't use a paired t-test since the AP@10 values were not normally distributed.

TABLE III. MAP@10 RESULTS OF THE STUDY

	CCF	NwCF
Pearson (CCF)	0.12875	0.1432*
Tag-based (BM25)	0.0876**	0.1942***

Significance over the baseline:
*p < 0.236, **p < 0.033, ***p < 0.001

As we expected, BM25+NwCF gives the best results, in table III a MAP@10 of 0.1942 compared to a MAP@10 of 0.0876, significant with p<0.001. Now, comparing simple CCF with BM25+CCF, one may argue that the actual improvement is done by NwCF rather than BM25. However, the importance of BM25 stems not from MAP@10 results, but from its coverage (99.23% compared to 81.12% over 784 users), i.e., more users can receive recommendations. It makes BM25 a sound alternative for Pearson correlation to calculate the similarity between users (especially taking into account that many social bookmarking systems receive few user ratings, or they just do not implement rating as functionality). We extend the explanation of this result in the section 6.

TABLE IV. USER COVERAGE RESULTS OF THE STUDY

	CCF	NwCF
Pearson (CCF)	81.12%	81.12%
Tag-based (BM25)	99.23%	99.23%

In the case of NwCF, the results lead us to the same results as in our previous study [15]. Table III shows an improvement of NwCF (MAP@10 0.1432) over CCF (MAP@10 0.12875), with p<0.236. The statistics of table II can give us an explanation. With an average of just 2.68 users per item, it is difficult to assure that the ratings given to one item reflect an objective evaluation. Improving the recommendation ranking by preferring items rated by a larger amount of users decreases the noise of that evaluation. This correlates well with real life practices. For example, in the online store Amazon.com we tend to trust more in the aggregated rating given to one item when it has been evaluated by a larger amount of people.

V. DISCUSSION

The results of this study confirm that creative approaches inspired by the nature of the collaborative tagging systems might perform better in this new context than pragmatic implementation of traditional approaches. In our case, we explored a alternative tag-based approach using a state-of-the art information retrieval model to calculate similarity between users and it paid off: BM25 based similarity, combined with NwCF, performed better than CCF in both phases of the study.

BM25, a similarity measure that demonstrated its potential as an alternative to Pearson correlation, can be especially useful in social bookmarking systems since they do not commonly provide ratings. Though showing worse results than CCF when is not combined with NwCF, its increased user coverage highlights the importance of this method. Furthermore, when combined with NwCF, it provides better results than Pearson combined with NwCF. A natural question is: How the large difference in MAP between BM25+CCF and BM25+NwCF can be explained? We think that, as shown by the user coverage, BM25 is able to produce a larger but also "noisier" neighborhood than CCF. It allows finding more truly relevant users, while also mixing them with superficially relevant users. This additional noise is however, reduced by NwCF. Consequently, the combination of BM25+NwCF has the ability to bring more neighbors and also to re-enforce the signal of the most relevant neighbors. In comparison, CCF, which relies on item co-tagging, carries less noise, so its improvement with NwCF is not as large as combining BM25 with NwCF.

Yet, there is a trade-off on using BM25 between its accuracy and its scalability. Despite the good results of BM25+NwCF, BM25 by itself is more computationally-expensive than CCF. BM25 can be speed up by storing users and tags by a batch process using an indexer such as Lucene or Lemur , so at the moment of the recommendation the information of users is quickly retrieved. However, this indexing process must be executed off-line: running it each time a new user is added or a new article is posted would be too time and resource consuming, at the expense of lacking the most updated information when creating the recommendations. A very active user posting articles frequently, or a popular article being posted often, can produce similar scalability problems on CCF and NwCF.

Our experience with NwCF demonstrates that the inclusion of the amount of raters in the ranking formula is an important contribution. This result hints that the number of raters is a part of the "social knowledge", which can increase the quality of outcome that CCF ignores. NwCF helps to reduce the noise of items rated by too few users, so it can be considered as an important tool given the sparse nature of the dataset: many more items than users, as shown in the tables I and II. This can be also seen as an option to alleviate the cold-start problem, common to new users and items in recommender systems. The new users still have not added enough articles to their library, and new items have not been shared by enough people to be suggested by the collaborative filtering algorithm. However, we cannot claim NwCF as an ultimate solution, since our final evaluation dataset has users sharing at least one item and items posted by at least two users to have a chance of recommendation. One option for those cases is using a content-based approach, or a hybrid approach that combines a collaborative filtering solution with a content-based one, as described in [1; 9], such as computing cosine similarity or similarity based on available metadata.

VI. SUMMARY AND FUTURE WORK

We developed and compared four approaches for item recommendation in the collaborative tagging system CiteULike. The baseline, Classic Collaborative filtering, used Pearson correlation to calculate similarity between users and a classic adjusted ratings formula to rank the recommendations. The second approach, Neighbor-weighted Collaborative Filtering (NwCF), enhanced traditional ranking (prediction) by taking into account the number of raters in the ranking formula of the recommendations. The third approach explored an innovative way to form the user neighborhood based on Okapi BM25 model over users' tags, while keeping the CCF ranking intact. The combined BM25-NwCF approach uses both, BM25 for neighborhood formation and NwCF for ranking. Our results demonstrate that NwCF improves significantly the precision of recommended items over CCF. BM25 alone does not improve the precision results, yet it increases the user coverage, i.e., the amount of users who receive recommendations. The combination of BM25 and NwCF gives the best results, by combining the potentials of each approach: increasing the coverage of users and items, in the case of BM25, and improving the precision, for NwCF.

Some issues explored in this research remain open. One of them is how to beat the cold-start problem when providing recommendations in social tagging systems. Using a content-based or a hybrid-based approach can help to mitigate this problem. Scalability is also an area of concern and we are already working in developing novel spreading activation algorithms to calculate local ranking of users and items around their network.

VII. REFERENCES

[1] T. Bogers, Recommender systems for social bookmarking Tilburg University, 2009.

[2] P. Cotter and B. Smyth, WAP-ing the Web: Content personalization for WAP-enabled devices. in: Proc. Adaptive Hypermedia and Adaptive Web-based systens P. Brusilovsky, O. Stock and C. Strapparava, eds., Lecture Notes in Computer Science 1892, Springer-Verlag, Trento, Italy, 2000, pp.98-108.

[3] M. Coyle, J. Freyne, P. Brusilovsky and B. Smyth, Social Information Access for the Rest of Us: An Exploration of Social YouTube. in: Proc. 5th International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH'2008) W. Nejdl, J. Kay, P. Pu and E. Herder, eds., Lecture Notes in Computer Science 5149, Springer Verlag, Hannover, Germany, 2008, pp.93-102.

[4] A. Dattolo, F. Ferrara and C. Tasso, Supporting Personalized User Concept Spaces and Recommendations for a Publication Sharing System. in: Proc. 17th International Conference on User Modeling, Adaptation, and Personalization (UMAP 2009) G.-J. Houben, G. McCalla, F. Pianesi and M. Zancanaro, eds., Lecture Notes in Computer Science 5535, Springer, Trento, Italy, 2009, pp.325-330.

[5] M. de Gemmis, P. Lops, G. Semeraro and P. Basile, Integrating tags in a semantic content-based recommender. in: Proc. the 2008 ACM conference on Recommender systems, RecSys '08 ACM, Lausanne, Switzerland, 2008, pp.163-170.

[6] H. Drachsler, D. Pecceu, T. Arts, E. Hutten, L. Rutledge, P.v. Rosmalen, H. Hummel and R. Koper, ReMashed – Recommendations for Mash-Up Personal Learning Environments. in: Proc. 4th European Conference on Technology Enhanced Learning (ECTEL 2009) U. Cress, V. Dimitrova and M. Specht, eds., Lecture Notes in Computer Science 5794, Springer-Verlag, Nice, France, 2009, pp.788-793.

[7] R. Farzan, M. Coyle, J. Freyne, P. Brusilovsky and B. Smyth, ASSIST: adaptive social support for information space traversal. in: Proc. 18th conference on Hypertext and hypermedia, HT '07ACM Press, Manchester, UK, 2007, pp.199-208.

[8] C.S. Firan, W. Nejdl and R. Paiu, The benefit of using tag-based profiles. in: Proc. Fifth Latin American Web Congress (LA-Web 2007) V.A.F. Almeida and R.A. Baeza-Yates, eds.,IEEE, Santiago de Chile, 2007, pp.32–41.

[9] J. Gemmell, T. Schimoler, M. Ramezani, L. Christiansen and B. Mobasher, Improving FolkRank With Item-Based Collaborative Filtering. in: Proc. Workshop on Recommender Systems and the Social Web at the 2009 ACM conference on Recommender systems, RecSys '09ACM, New York, NY, 2009.

[10] J. Herlocker, J. Konstan and J. Riedl, An Empirical Analysis of Design Choices in Neighborhood-Based Collaborative Filtering Algorithms. Information Retrieval, 5(4) (2002) 287-310.

[11] A. Hotho, R. Jäschke, C. Schmitz and G. Stumme, Information Retrieval in Folksonomies: Search and Ranking. in: The Semantic Web: Research and Applications, Lecture Notes in Computer Science 4011, 2006, pp.411-426.

[12] R. Jäschke, L. Marinho, A. Hotho, L. Schmidt-Thieme and G. Stumme, Tag Recommendations in Folksonomies. in: Proc. 11th European conference on Principles and Practice of Knowledge Discovery in Databases: PKDD 2007 Lecture Notes in Computer Science 4702, Warsaw, Poland, 2007, pp.506-514.

[13] C. Manning, P. Raghavan and H. Schutze, Introduction to Information Retrieval, Cambridge University Press, Cambridge, 2008.
[14] S. Niwa, T. Doi and S. Honiden, Web Page Recommender System based on Folksonomy Mining for ITNG'06 Submissions. in: Proc. Third International Conference on Information Technology: New Generations, ITNG 2006. , 2006, pp.388-393.
[15] D. Parra and P. Brusilovsky, Collaborative Filtering for Social Tagging Systems: An Experiment with CiteULike. in: Proc. the 2009 ACM conference on Recommender systems, RecSys '09ACM, New York, NY, 2009, pp.237-240.
[16] M.J. Pazzani and D. Billsus, Content-based recommendation systems. in: The Adaptive Web: Methods and Strategies of Web Personalization, P. Brusilovsky, A. Kobsa and W. Neidl, eds., Lecture Notes in Computer Science 4321, Springer-Verlag, Berlin Heidelberg New York, 2007, pp.325-341.
[17] J.B. Schafer, D. Frankowski, J. Herlocker and S. Sen, Collaborative filtering recommender systems. in: The Adaptive Web: Methods and Strategies of Web Personalization, P. Brusilovsky, A. Kobsa and W. Neidl, eds., Lecture Notes in Computer Science 4321, Springer-Verlag, Berlin Heidelberg New York, 2007, pp.291-324.
[18] P. Symeonidis, User Recommendations based on Tensor Dimensionality Reduction. in: Artificial Intelligence Applications and Innovations III, 2009, pp.331-340.
[19] V. Zanardi and L. Capra, Social ranking: uncovering relevant content using tag-based recommender systems. in: Proc. RecSys '08: Proceedings of the 2008 ACM conference on Recommender systemsACM, Lausanne, Switzerland, 2008, pp.51-58.
[20] S. Zhao, N. Du, A. Nauerz, X. Zhang, Q. Yuan and R. Fu, Improved recommendation based on collaborative tagging behaviors. in: Proc. the 13th international conference on Intelligent user interfaces, IUI '08ACM, Gran Canaria, Spain, 2008, pp.413-416.
[21] Y. Zhen, W.-J. Li and D.-Y. Yeung TagiCoFi: Tag Informed Collaborative Filtering ACM Press, New York, NY (2009) 69-76