A Fast Effective Multi-Channeled Tag Recommender

Jonathan Gemmell, Maryam Ramezani, Thomas Schimoler,
Laura Christiansen, and Bamshad Mobasher

Center for Web Intelligence
School of Computing, DePaul University
Chicago, Illinois, USA
{jgemmell,mramezani,tschimoler,lchris,mobasher}@cti.depaul.edu

Abstract. Collaborative tagging applications allow users to annotate online resources, resulting in a complex three dimensional network of interrelated users, resources and tags often called a folksonom A pivotal challenge of these systems remains the inclusion of the varied information channels introduced by the multi-dimensional folksonomy into recommendation techniques. In this paper we propose a composite tag recommender based upon popularity and collaborative filtering. These recommenders were chosen based on their speed, memory requirements and ability to cover complimentary channels of the folksonomy. Alone these recommenders perform poorly; together they achieve a synergy which proves to be as effective as state of the art tag recommenders.

Key words: Folksonomies, Tag Recommenders, Hybrid Recommenders

1 Introduction

Collaborative tagging has emerged as a popular method for organizing and sharing online content with user-defined keywords. Delicious[1], Flickr[2] and Last.fm[3] are among the most popular destinations on the Web allowing users to annotate bookmarks, digital photographs and music. Other less popular tagging applications serve niche communities enabling users to tag blogs, business documents or scholarly articles.

At the heart of collaborative tagging is the post; a user describes a resource with a set of tags. A collection of posts results in a complex network of interrelated users, resources and tags commonly referred to as a folksonomy [10].

The rich tapestry of a folksonomy presents an enticing target for data mining techniques such as recommenders. Recommenders reduce a burdensome number of items to a manageable size correlated to the user's interests. Recommendation in folksonomies can include resources, tags or even other users. In this work

[1] delicious.com
[2] www.flickr.com
[3] www.last.fm

we focus on tag recommendation, the suggestion of tags during the annotation process.

Tag recommendation reduces the cognitive effort from generation to recognition. Users are therefore encouraged to tag more frequently, apply more tags to a resource, reuse common tags and perhaps use tags the user had not previously considered. User error is reduced by eliminating capitalization inconsistencies, punctuation errors, misspellings and other discrepancies. The final result is a cleaner denser dataset that is useful in its own right or for further data mining techniques.

Despite the richness folksonomies offer, they present unique challenges for tag recommenders. Traditional recommendation strategies, often developed to work with two dimensional data, must be adapted to work with the three dimensional nature of folksonomies. Otherwise they risk disregarding potentially useful information. To date the most successful tag recommenders are graph-based models, which exploit the links between users, resources and tags. However, this approach is computationally intense and ill suited for large scale implementation.

In this work we propose a composite tag recommender incorporating several distinct recommendation strategies. These recommenders are combined to generate a new hybrid. As such no single recommender is required to fully exploit the data structure of the folksonomy. Instead the recommenders may specialize in a single channel. The aggregation of these recommenders, none of which performs well on its own, produce a synergy allowing the composite recommender to outperform its constituent parts.

Our hybrid includes popularity models and item-based collaborative filtering techniques. Popularity based approaches include information garnered from the crowd with little computational cost. Item-based collaborative filtering focuses more closely on the user's profile incorporating a degree of personalization.

We provide a through evaluation of the composite recommender and its constituent parts. Our experiments reveal that the composite model produces results far superior to the capabilities of their individual components. We further include a comparison with the highly effective but computationally inefficient graph-based approach. We show that a low cost alternative can be constructed from less time consuming recommenders and perform nearly as well as the state or the art graph based approaches.

The rest of the paper is organized as follows. In Section 2 we discuss related work. In Section 3 we offer a model of folksonomies and describe tag recommendation. We further describe four recommendation algorithms. Informational channels in folksonomies are discussed in Section 4. We design a hybrid recommender in Section 5. Our experimental evaluation is presented in Section 6 including a discussion of the dataset, methodology and results. Finally we end the paper with a discussion of our conclusions and directions for future work in Section 7.

2 Related Work

As collaborative tagging applications have gained in popularity researchers have begun to explore and characterize the tagging phenomenon. In [9] and [4] the authors studied the information dynamics of Delicious, one of the most popular folksonomies. The authors discussed how tags have been used by individual users over time and how tags for an individual resource stabilize over time. They also explored two semantic difficulties: tag redundancy, when multiple tags have the same meaning, and tag ambiguity, when a single tag has multiple meanings. In [9] the authors provide an overview of the phenomenon and explore reasons why both folksonomies and ontologies will have a place in the future of information access.

There have been several recent research investigations into recommendation within folksonomies. Unlike traditional recommender systems which have a two-dimensional relation between users and items, tagging systems have a three dimensional relation between users, tags and resources. Recommender systems can be used to recommend each of the dimensions based on one or two of the other dimensions. In [17] the authors apply user-based and item-based collaborative filtering to recommend resources in a tagging system and uses tags as an extension to the user-item matrices. Tags are used as context information to recommend resources in [13] and [12].

Other researchers have studied tag recommendation in folksonomies. In [7] user-based collaborative filtering is compared to a graph-based recommender based on the Pagerank algorithm for tag recommendation. The authors in [5] use association rules to recommend tags and introduce an entropy-based metric to define how predictable a tag is. In [8] the title of a resource, the posts of a resource and the user's vocabulary are used to recommend tags.

General criteria for a good tagging system including high coverage of multiple channels, high popularity and least-effort are presented in [18]. They categorize tags as content-based tags, context-based tags, attribute tags, subjective tags, and organizational tags and use a probabilistic method to recommend tags. In [2] the authors propose a classification algorithm for tag recommendation. The authors in [15] use a co-occurrence-based technique to recommend tags for photos in Flickr. The assumption is that the user has already assigned a set of tags to a photo and the recommender uses those tags to recommend more tags. Semantic tag recommendation systems in the context of a semantic desktop are explored in [1]. Clustering to make real-time tag recommendation is developed in [16].

3 Tag Recommendation

Here we first provide a model of folksonomies, then review several common recommendation techniques which we employ in our evaluation. A folksonomy can be described as a four-tuple $D = \langle U, R, T, A \rangle$, where, U is a set of users; R is a set of resources; T is a set of tags; and A is a set of annotations, represented as user-tag-resource triples: $A \subseteq \{\langle u, r, t\rangle : u \in U, r \in R, t \in T\}$. A folksonomy can, therefore, be viewed as a tripartite hyper-graph [11] with users, tags, and resources represented as nodes and the annotations represented as hyper-edges connecting a user, a tag and a resource.

Aggregate projections of the data can be constructed, reducing the dimensionality but sacrificing information [14]. The relation between resources and tags, RT, can be formulated such that each entry, $RT(r,t)$, is the weight associated with the resource, r, and the tag, t. This weight may be binary, merely showing that one or more users have applied that tag to the resource. In this work we assume $RT(r,t)$ to be the number of users that have applied t to the r: $RT_{tf}(r,t) = |\{a = \langle u, r, t\rangle \in A : u \in U\}|$. Analogous two-dimensional projections can be constructed for UT in which the weights correspond to users and tags, and UR in which the weights correspond to users and resources.

Many authors have attempted to exploit the data model for recommendation in folksonomies. In traditional recommendation algorithms the input is often a user, u, and the output is a set of items, I. Tag recommendation differs in that the input is both a user and a resource. The output remains a set of items, in this case a set of recommended tags, T_r. Given a user-resource pair, the recommendation set is constructed by calculating a weight for each tag, $w(u,r,t)$, and recommending the top n tags.

3.1 Popularity Based Approaches

We consider two popularity based models which rely on the frequency a tag is used. **PopRes** ignores the user and relies on the popularity of a tag within the context of a particular resource. We define the resource based popularity measure as:

$$w(u,r,t) = \frac{|\{a = \langle u, r, t\rangle \in A : u \in U\}|}{|\{a = \langle u, r, t\rangle \in A : u \in U, t \in T\}|} \tag{1}$$

PopUser, on the other hand, ignores the resource and focuses on the frequency of a tag within the user profile. We define the user based popularity measure as:

$$w(u,r,t) = \frac{|\{a = \langle u, r, t\rangle \in A : r \in R\}|}{|\{a = \langle u, r, t\rangle \in A : r \in R, t \in T\}|} \tag{2}$$

Popularity based recommenders require little online computation. Models are built offline and can be incrementally updated. However both these models focus on a single channel of the folksonomy and may not incorporate otherwise relevant information into the recommendation.

3.2 Item-Based Collaborative Filtering

KNN_RT models resources as a vector over the tag space. As before the weights of the vectors may be calculated through a variety of means. Given a resource and a tag, we define the weight as the entry of the two dimensional projection, $RT(r, t)$, the number of times r has been tagged with t.

When a user selects a resource to annotate, the similarity between it and every resource in the user profile is calculated. A neighborhood of the k most similar resources, S, is thus constructed. We then define the item-based collaborative filtering measure as:

$$w(u, r, t) = \frac{\sum_s^S sim(s, r) * d(u, s, t)}{k} \tag{3}$$

where $d(u, s, t)$ is 1 if the user has applied t to s and 0 otherwise. Like *popUser*, this recommender focuses strongly on the user's tagging practice. However this recommender includes an additional informational channel, identifying resources in the user profile that are similar to the query resource. This technique therefore includes resource-to-resource information.

If the system waits to compute the similarity between resources until query time, this recommender will also scale well to larger datasets so long as user profiles remain small. Alternatively similarities between resources can be computed offline. Consequently the computation at query time is dramatically reduced and the algorithm becomes viable for large collaborative tagging implementations.

3.3 Folkrank

Folkrank was proposed in [6]. It computes a Pagerank vector from the tripartite graph of the folksonomy. This graph is generated by regarding $U \cup R \cup T$ as the set of vertices. Edges are defined by the three two-dimensional projections of the hypergraph, RT, UR and UT.

If we regard the adjacency matrix of this graph, W, (normalized to be column-stochastic), a damping factor, d, and a preference vector, p, then we iteratively compute the Pagerank vector, w, in the usual manner: $w = dAw+(1-d)p$.

However due to the symmetry inherent in the graph, this basic Pagerank may focus too heavily on the most popular elements. The Folkrank vector is taken as a difference between two computations of Pagerank: one with and one without a preference vector. Tag recommendations are generated by biasing the preference vector towards the query user and resource [7]. These elements are given a substantial weight while all other elements have uniformly small weights.

We include this method as a benchmark as it has demonstrated to be an effective method of generating tag recommendations. However, it imposes steep computational costs.

4 Informational Channels of Folksonomies

The model of a folksonomy suggests several informational channels which may be exploited by data mining applications such as tag recommenders. The relation between users, resources and tags generate a complex network of interrelated items as shown in Figure 1.

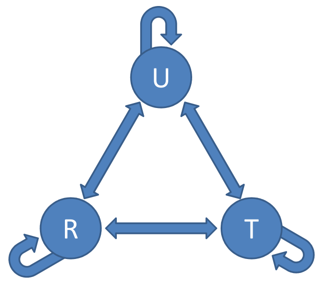

Fig. 1. Informational channels of a folksonomy.

The channel between resources and tags reveals a highly descriptive model of the resources. The accumulation of many users' opinions (often numbered in the thousands or millions) results in a richness which taxonomies are unable to approximate. Conversely the tags themselves are characterized by the resources to which they have been assigned.

As users annotate resource with tags they define their interests in as much as they describe a resource. The user-tag channel therefore reveals the users interests and provides opportunities for data mining algorithms to offer a high degree of personalization. Likewise a user may be defined by the resources which he has annotated as in the user-resource channel.

These primary channels can be used to produce secondary informational channels. The user-user channel can be constructed by modeling users as a vector of tags or as a vector of resources and applying a similarity measure such as cosine similarity. Many variations exist. However the result reveals a network of users that can be explored directly or incorporated into further data mining approaches. The resource-resource and tag-tag channels provide similar utility, presenting navigational opportunities for users to explore similar resources or tags.

5 A Multi-Channeled Tag Recommender

The most successful tag recommenders to date have included multiple informational channels. Folkrank explicitly includes the user-resource, user-tag and resource-tag channels in the graph model. Moreover since the algorithm calculates the Pagerank vector of the graph it implicitly includes the secondary channels of the folksonomy. The success Folkrank has achieved is due to its ability to incorporate multiple informational channels into a single tag recommender.

However the success it has achieved is blunted by the computational effort required to produce a recommendation; a new Pagerank vector is computed for each query.

Here we construct a hybrid recommender. The constituent parts by themselves perform poorly when compared to Folkrank. However, when aggregated into a single recommender they achieve a synergy which exploits several channels of the folksonomy while retaining their modest computational needs.

Our model includes *PopRes*, *PopUser* and *KNN_RT*. We employ a weighted approach to combine the recommenders. First in order to ensure that weight assignments are on the same scale for each recommendation approach, we normalize the weights given to the tags by $w(u, r, t)$ to 1 producing $w'(u, r, t)$. We then combine the weights in a linear combination:

$$w(u, r, t) = \alpha w'_{PopRes}(u, r, t) + \beta w'_{PopUser}(u, r, t) + \gamma w'_{KNN\_RT}(u, r, t) \tag{4}$$

such that weights $\alpha + \beta + \gamma = 1$ and all values are positive. If α is set near 1 then hybrid would rely mostly on *PopRes*.

Tags promoted by *PopRes* will have a strong relevance to the resource, while tags promoted by *PopUser* will include tags in the user's profile. *PopRes* alone will ignore personal tags that the user often users. *PopUser*, on the other hand, will ignore tags related to the context of the query resource. Together these recommenders can include both aspects in the recommendation set. Moreover by including *KNN_RT* tags which the user has applied to resources similar to the query resource are promoted.

PopRes explicitly includes the resource-tag information. *PopUser*, on the other hand, includes user-tag information. Both these models are based on popularity and are single-minded in their approach ignoring all data except the informational channel to which they are employed. We use *KNN_RT* to introduce more subtlety into the hybrid. It focuses heavily on the user-tag channel, but gives more weight to tags that have been applied to similar resources. Hence it also includes resource-tag information. Moreover by focusing exclusively on resources in the user profile it includes the user-resource channel. Finally, *KNN_RT* includes resource-resource information when it calculates the neighborhood of similar resources.

This hybrid does not include user-user information or tag-tag information. Additional recommenders could be included to cover these informational channels. However, we have built this hybrid with the goals of speed and simplicity. The two popularity based approaches are among the fastest and simplest recommendation algorithms. The item-based collaborative filtering recommender is used to tie together these approaches incorporating similarities among resources into the model while retaining its speed.

	Complete	PostCore(2)
Users	3,617	253,615
URLs	235,328	41,268
BibTeXs	143,050	22,852
Tags	93,756	1,185
Tag Assignments	1,401,104	14,443
Bookmark Posts	263,004	7,946
BibTeX Posts	158,924	13,276

Table 1. Bibsonomy datasets.

6 Experimental Evaluation

In this section we describe the dataset used for experimentation. We then describe our experimental methodology and metrics. Finally we discuss the results of our experiments.

6.1 Data Set

The dataset was provided by Bibsonomy[4] for the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD) 2009 Challenge. BibSonomy was originally launched as a collaborative tagging application allowing users to organize and share scholarly references. It has since expanded its scope allowing users to annotate URLs.

The data includes all public bookmarks and publication posts of BibSonomy until 2009-01-01. The data was cleaned by removing all characters which are neither numbers nor letters from tags. Additionally the system tags *imported*, *public*, *systemimported*, *nn* and *systemunfiled* where removed.

Task 1 for the 2009 Challenge utilizes the complete dataset. Task 2 however focuses on the post-core at level 2 geared toward graph based approaches. For the post-core all users, tags, and resources which appear in only one post were removed. This process was repeated until convergence and produced a core in which each user, tag, and resource occurs in at least two posts. Reducing a dataset to its core was first proposed in [3]. In [6] it was adapted for folksonomies. The experiments for this work rely on post-core at level 2.

6.2 Experimental Methodologies

We employ the leave one post out methodology as described in [7]. One post from each user was placed in the testing set consisting of a user, u, a resource, r, and all the tags the user has applied to that resource. These tags, T_h, are analogous to the holdout set commonly used in Information Retrieval evaluation. The remaining posts are used to generate the recommendation models.

[4] www.bibsonomy.org

The tag recommendation algorithms accepts the user-resource pair and returns an ordered set of recommended tags, T_r. From the holdout set and recommendation set utility metrics were calculated. For each metric the average value was calculated across all test cases.

6.3 Experimental Metrics

Recall is a common metric of recommendation algorithms that measures coverage. It measures the percentage of items in the holdout set, T_h, that appear in the recommendation set T_r. It is defined as:

$$r = (|T_h \cap T_r|)/|T_h| \tag{5}$$

Precision is another common metric that measures specificity and is defined as:

$$p = (|T_h \cap T_r|)/|T_r| \tag{6}$$

In order to conform to the evaluation methods of the ECML-PKDD 2009 Challenge, we use the F1-Measure common in Information Retrieval to evaluate the recommendations. We compute for each post the recall and precision for a recommendation set of five tags. Then we average precision and recall over all posts in the test data and use the resulting precision and recall to compute the F1-Measure as:

$$f1 = (2 * p * r)/(p + r) \tag{7}$$

6.4 Experimental Results

Our approach required that several variables be tuned. For *KNN_RT*, after extensive experimentation of k in increments of 1 we set k equal to 15. We observed that as k increased from 0 to 15 recall and precision both increased rapidly until it suffers from diminishing returns.

We evaluated the weights α, β and γ in .05 increments attempting every possible combination. Best results were found when $\alpha = 0.35$, $\beta = 0.15$ and$\gamma = 0.50$. As such *KNN_RT* accounts for 50% of the model, *PopRes* acounts for 35% and *PopUser* acounts for 15%.

KNN_RT identifies resources in the user profile most similar to the query resource and promotes the tags applied to these resources. This approach is most effective when the user has generated a large user profile. Since users often employ tags as an organizational tool they often reuse tags. Hence the success of *KNN_RT* stems from its ability to identify which previously used tags are most appropiate given the context of the query resource.

PopRes, on the other hand, ignores the user profile and concentrates on the popularity of a tag given the query resource. When the tags provided by *KNN_RT* are insufficient, perhaps because the user has yet to build a deep user profile or

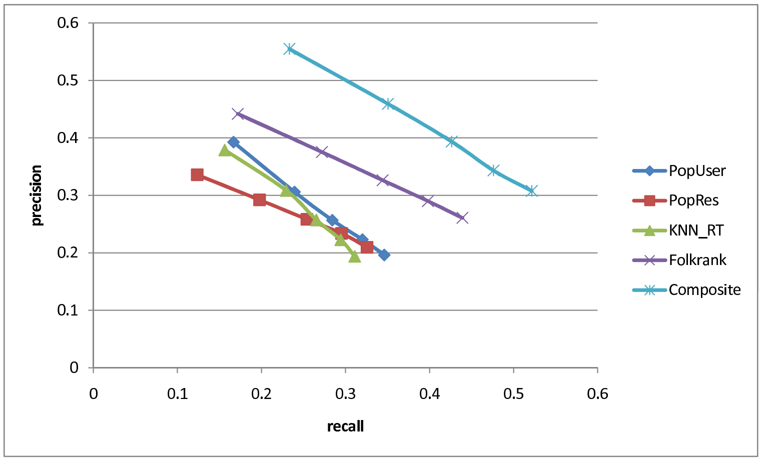

Fig. 2. Evaluation of recommendation techniques: recall vs. precision.

is tagging a resource dissimilar to items in the profile, *PopRes* is able to provide relevant suggestions.

Finally *PopUser* promotes tags in the user profile regardless of the similarity to the query resource. It may promote idiosyncratic, subjective or organizational tags that do not necessarily relate to the context of the query resource but are often applied by the user.

Our evaluation of the composite recommenders in Figures 2 and 3 reveals that *PopRes*, *PopUser* and *KNN_RT* achieve only modest success when used alone. However when combined together as a hybrid recommender the three are able to cover multiple informational channels and produce a synergy allowing the hybrid to produce superior results.

Not only is the hybrid recommender able to outperform the baseline recommenders it is also able to outperform Folkrank, a highly effective tag recommender. Moreover the hybrid retains the computationally efficiency of its parts making it suitable for deployment in large real work collaborative filtering applications.

7 Conclusions and Future Work

In this paper we have introduced the idea of informational channels in folksonomies and have proposed a fast yet effective tag recommender composed of three separate algorithms. The constituent recommenders were chosen for their speed and simplicity as well as their ability to cover complimentary informational channels. We have demonstrated that these recommenders while performing poorly alone, create a synergy when combined in a linear combination. The hybrid recommender is able to surpass the effective graph based approaches while

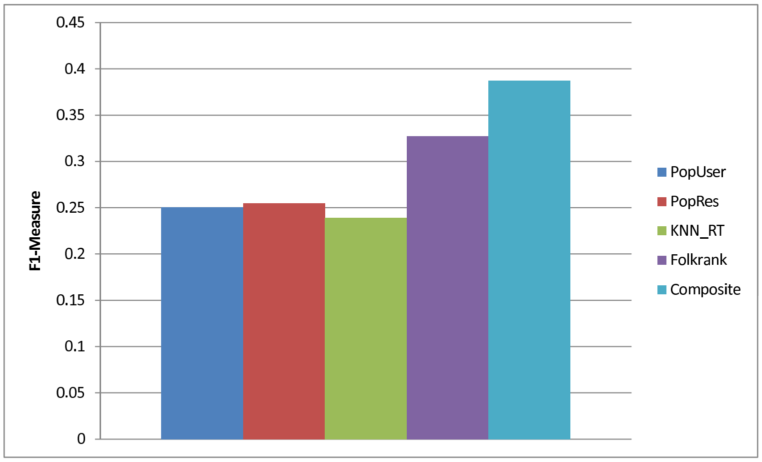

Fig. 3. Evaluation of recommendation techniques: F1-measure.

retaining the efficiency of its parts. Future work will include an examination of alternative hybrid recommenders and present work on other datasets.

8 Acknowledgments

This work was supported in part by the National Science Foundation Cyber Trust program under Grant IIS-0430303 and a grant from the Department of Education, Graduate Assistance in the Area of National Need, P200A070536.

References

1. B. Adrian, L. Sauermann, and T. Roth-Berghofer. Contag: A semantic tag recommendation system. In T. Pellegrini and S. Schaffert, editors, *Proceedings of I-Semantics' 07*, pages pp. 297–304. JUCS, 2007.
2. P. Basile, D. Gendarmi, F. Lanubile, and G. Semeraro. Recommending smart tags in a social bookmarking system. In *Bridging the Gep between Semantic Web and Web 2.0 (SemNet 2007)*, pages 22–29, 2007.
3. V. Batagelj and M. Zaveršnik. Generalized cores. *Arxiv preprint cs/0202039*, 2002.
4. S. A. Golder and B. A. Huberman. Usage patterns of collaborative tagging systems. *Journal of Information Science*, 32(2):198, 2006.
5. P. Heymann, D. Ramage, and H. Garcia-Molina. Social tag prediction. In *SIGIR '08: Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 531–538, New York, NY, USA, 2008. ACM.
6. A. Hotho, R. Jaschke, C. Schmitz, and G. Stumme. Information retrieval in folksonomies: Search and ranking. *Lecture Notes in Computer Science*, 4011:411, 2006.
7. R. Jaschke, L. Marinho, A. Hotho, L. Schmidt-Thieme, and G. Stumme. Tag Recommendations in Folksonomies. *LECTURE NOTES IN COMPUTER SCIENCE*, 4702:506, 2007.

8. M. Lipczak. Tag recommendation for folksonomies oriented towards individual users. In *Proceedings of the ECML/PKDD 2008 Discovery Challenge Workshop, part of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, 2008.
9. G. Macgregor and E. McCulloch. Collaborative tagging as a knowledge organisation and resource discovery tool. *Library Review*, 55(5):291–300, 2006.
10. A. Mathes. Folksonomies-Cooperative Classification and Communication Through Shared Metadata. *Computer Mediated Communication, (Doctoral Seminar), Graduate School of Library and Information Science, University of Illinois Urbana-Champaign, December*, 2004.
11. P. Mika. Ontologies are us: A unified model of social networks and semantics. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(1):5–15, 2007.
12. R. Y. Nakamoto, S. Nakajima, J. Miyazaki, S. Uemura, and H. Kato. Investigation of the effectiveness of tag-based contextual collaborative filtering in website recommendation. In *Advances in Communication Systems and Electrical Engineering*, pages 309–318. Springerlink, 2008.
13. R. Y. Nakamoto, S. Nakajima, J. Miyazaki, S. Uemura, H. Kato, and Y. Inagaki. Reasonable tag-based collaborative filtering for social tagging systems. In *WICOW '08: Proceeding of the 2nd ACM workshop on Information credibility on the web*, pages 11–18, New York, NY, USA, 2008. ACM.
14. C. Schmitz, A. Hotho, R. Jaschke, and G. Stumme. Mining association rules in folksonomies. In *Proc. IFCS 2006 Conference*, pages 261–270. Springer, 2006.
15. B. Sigurbjörnsson and R. van Zwol. Flickr tag recommendation based on collective knowledge. In *WWW '08: Proceeding of the 17th international conference on World Wide Web*, pages 327–336, New York, NY, USA, 2008. ACM.
16. Y. Song, Z. Zhuang, H. Li, Q. Zhao, J. Li, W.-C. Lee, and C. L. Giles. Real-time automatic tag recommendation. In *SIGIR '08: Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 515–522, New York, NY, USA, 2008. ACM.
17. K. H. L. Tso-Sutter, L. B. Marinho, and L. Schmidt-Thieme. Tag-aware recommender systems by fusion of collaborative filtering algorithms. In *SAC '08: Proceedings of the 2008 ACM symposium on Applied computing*, pages 1995–1999, New York, NY, USA, 2008. ACM.
18. Z. Xu, Y. Fu, J. Mao, and D. Su. Towards the semantic web: Collaborative tag suggestions. *Collaborative Web Tagging Workshop at WWW2006, Edinburgh, Scotland, May*, 2006.