Recommendation in the Social Web

Robin Burke, Jonathan Gemmell,
Andreas Hotho, and Robert Jäschke

■ *Recommender systems are a means of personalizing the presentation of information to ensure that users see the items most relevant to them. The social web has added new dimensions to the way people interact on the Internet, placing the emphasis on user-generated content. Users in social networks create photos, videos, and other artifacts, collaborate with other users, socialize with their friends, and share their opinions online. This outpouring of material has brought increased attention to recommender systems as a means of managing this vast universe of content. At the same time, the diversity and complexity of the data has meant new challenges for researchers in recommendation. This article describes the nature of recommendation research in social web applications and provides some illustrative examples of current research directions and techniques.*

It is difficult to overstate the impact of the social web. This new breed of social applications is reshaping nearly every human activity from the way people watch movies to how they overthrow governments. Facebook allows its members to maintain friendships whether they live next door or on another continent. With Twitter, users from celebrities to ordinary folks can launch their 140 character messages out to a diverse horde of "followers." Flickr and YouTube users upload their personal media to share with the world, while Wikipedia editors collaborate on the world's largest encyclopedia.

These varied applications all share some fundamental characteristics. First, the focus is often partly (if not entirely) on the users themselves, and especially on their assocations and interactions with others. Second, many of these applications facilitate the organization and creation of online content. Finally, social applications often allow users to respond to posted content in a variety of ways, sometimes in real time, embedding each item in a rapidly evolving social network.

The complex information space generated by the social web offers a rich and dynamic environment for users to share information, discover new content, and meet new people. However, the success of the social web has made some of these benefits difficult to realize, due to the vast amount of information available. Twitter has 50 million tweets per day; which should you read? Every minute more than 24 hours of video is uploaded to YouTube; what should you watch? Recommender systems have the potential to filter the oceans of data that make up the social web and provide a personalized view for each user.

The types of recommender systems most commonly studied by researchers and discussed elsewhere in this special issue are not always a great fit for the task of recommendation in the social web. Data used to recommend products in an e-commerce setting, for example, represent a user-item relationship: a user might give a movie three stars, or a database might show which products a user has purchased. Much recommender systems

  ISSN 0738-4602

research has been focused on algorithms appropriate to data in this form.

The data associated with the social web is radically multidimensional in comparison. To see why this is so, consider a social tagging system, such as the music sharing site Last.fm or the citation sharing site BibSonomy. In such sites, users associate items with tags, short text labels in an unconstrained vocabulary, and share these annotations with others. We can think of a tag as a kind of rating, a valuation that the user associates with an item. However, in contrast to numeric ratings, which allow direct comparison on a scale, tags are textual and often idiosyncratic in their form. In addition, a user often applies multiple tags when annotating an item. So, in contrast to scalar ratings, a tagging system has valuations that are effectively vectors of tags, and the tag space may contain hundreds of thousands of unique terms.

Another way to view tags would be analogous to item descriptors such as might be found in a content-based recommender system. However, unlike an item catalog, the descriptors in tagging systems are not assigned by a centralized entity using a standard vocabulary. They are created by a host of individuals for a variety of purposes, so a tag must be interpreted relative to an individual. For example, for user *A*, who is a car fancier, the tag "jaguar" might be similar to user *B*'s tag "xj-series," whereas user *C*'s "jaguar" might be a label used only for jungle cats. Instead of a single one-dimensional description representing the content of an item, we have a family of them, each offering a different user's perspective.

This article describes the landscape of the social web, sampling some of the components common across its many applications, and discusses some of the recommendation tasks that naturally arise from this landscape. Taking social tagging systems as the most well-studied application in this area, we sample a few of the current techniques being employed to address its challenges.

Landscape of the Social Web

The social web can be considered as the aggregation of online interactions among users. These interactions can take myriad forms, some of which are shown in figure 1. Some interactions are between individuals while others revolve around online content or how that content is valued. In this section, we examine some of the core entities found in the social web and their characteristics with respect to recommendation.

People

Research on the web often models that system as an information graph, with web pages linking to each other in complex patterns. In this model, the authors do not participate and indeed, anonymity and difficulty in establishing authorship are two hallmarks of the web as we know it. The social web, on the other hand, lends itself to graphs in which users are connected to each other, either directly through social links of various types or indirectly through connections to content.

The interactions between a user and his or her online social contacts can vary as dramatically as they do in real-world environments. Interactions can be directional, when they are formed through subscription arrangements, as in Twitter, where the user being followed has no control over who chooses to be follower. They can also be bidirectional—as in the now famous "friend" relationship on Facebook. Interactions may involve groups of different sizes. Sometimes a user is posting content to a closed group of friends; sometimes, to the whole world. These actions all require different interpretations.

Certain aspects of social links have been studied by researchers interested in applying concepts of social trust in generating recommendations (Victor, DeCock, and Cornelis 2011). Some sites, such as epinions.com, allow users explicitly to create links to others as individuals whose valuations are reliable. There has been extensive research on the integration of these social network considerations into conventional rating-based recommendation, and many studies show that having the right connections can be a very powerful aid in producing accurate recommendations.

Helping users build social connections is a basic recommendation operation in the social web. Friend recommendation, for example, is a common operation that can directly leverage the social graph, as in Facebook's built-in friend finder that locates individuals with overlapping social circles. Still, because of the nature of the social environment, user recommendation has multiple interpretations. A user seeking a recommendation might be looking for a friend or colleague, but he or she might also be seeking a potential client, a romantic prospect, or a like-minded hobbyist, who could be more distant in the social network. In some applications, especially corporate intranets, the recommendation of individuals with particular skills or expertise may be an important task. Such an expert is unlikely to be found in the user's immediate social circle.

Items

The social web presents a vast array of items that can be recommended. One reason that people come to social websites is to hear about new items of the type that recommender systems have typically presented: movies, music, restaurants. However, in the social web, there will often be multiple routes into a given item. A music track might be

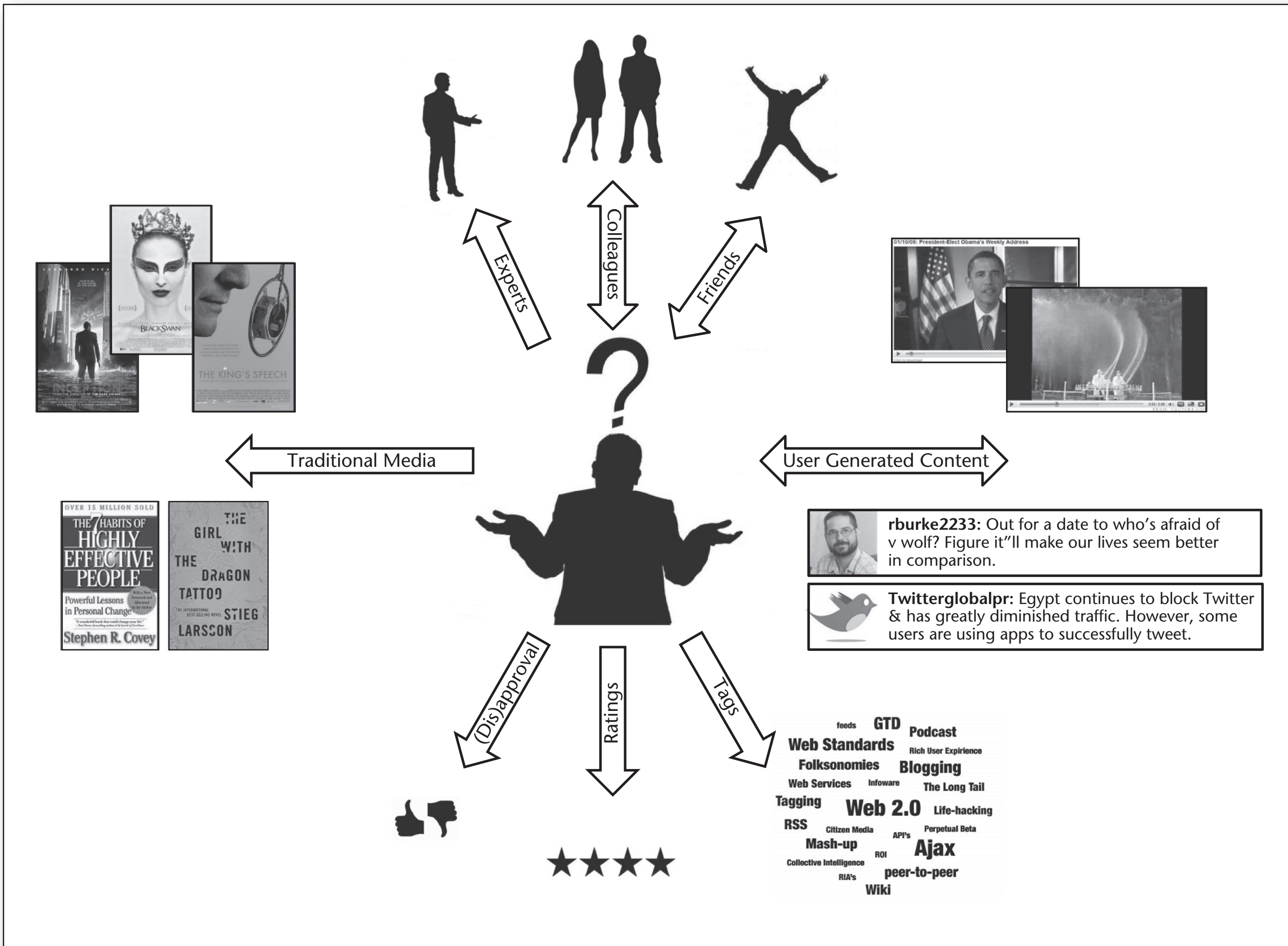

Figure 1. Landscape of the Social Web.

A user has many options on how to with connect other users, what types of content to consume, and how to respond to this content.

accessed through the artist, through a tag, through an album, through a friend's "favorites" list, and others. The multitude of links and interactions surrounding an item provides additional clues about its meaning and significance, but also poses a challenge for the recommender.

In addition to exogenous media items, a social website will also contain user-generated content, which is also fair game for recommendation. User-generated items, such as blog entries, status updates, and other data, have the benefit of being closely embedded in the networks of interaction that a social website hosts. Often, however, they are not intended to stand alone and can sometimes only be interpreted in context. For example, the Twitter post about "Who's Afraid of Virginia Woolf?" in figure 1 makes much more sense if you know that the poster is a college professor, like the male lead in the play.

Recommendations of user-generated content must also take context into account. Consider a user who just viewed photographs annotated with "wedding" and "New York." Would such a user be interested in seeing all photos with these tags? Most likely, the user interest is related to the particular individuals getting married, and pictures of strangers' New York weddings would not be welcome recommendations. On the other hand, if we know that the user is a professional photographer, that context would give rise to an entirely different set of recommendations.

Valuations

Users in social websites can express their opinions about contents of the site in a variety of ways. They can tag, post, tweet, rate, or write reviews. These reactions can be read, responded to, or forwarded to others. Substantial valuations like reviews may become items of recommendation in their own right. Many users prefer a peer review to a professionally authored one, especially if the author has some commonality with themselves.

In social annotation systems, a common application of recommendation is to assist users in labeling items through the recommendation of tags. If properly implemented, tag recommendation reduces the cognitive load of tagging, encouraging users to do more of it, and also reduces the amount of noise in the tag vocabulary by reducing redundancies like "NewYork," "New-York," "New_York," and others.[1]

Recommendation Tasks

The complexity of recommendation in the social web is not solely a function of the variety of possible items to be recommended. Its multidimensional nature means that a wide variety of recommendation tasks and modalities can be supported.

We can distinguish recommendation tasks by their input and output, and by the semantics of the recommendation operation. The wide variety of outputs has already been noted. The essential input to any recommender system is the profile of the user for whom the recommendation is sought. However, in the social web, user profile information may be augmented in various ways. For example, tag recommendation requires information about the resource to which the tag will be applied.

Because of the multidimensional nature of the supporting data, the social web supports a wide variety of recommendation semantics. For example, in recommending items, the system can present items personalized in terms of past viewing behavior, in terms of the behavior of friends, or in terms of a broader group of peers with similar interests. These would be presented to and interpreted by the user in different ways.

Adomavicius and Tuzhilin (2005) defined the recommendation problem as that of predicting items with the highest utility for a given user. This can be achieved by a function that computes the utility for a user-item pair $u(c, i)$ where c is a user and i is an item. In the social web, we extend this formulation in three ways. First, by augmenting the input to include requirements related the recommendation. For example, when recommending tags in a social tagging system, we are interested in the utility of a tag relative to a user-resource combination: "What tag would this user want to assign to this resource?" So, the utility needs to be a function of the whole triple: user, tag, and item. Second, we note that in the social web, recommenders are not restricted to a user/item dichotomy, but may recommend users, tags, reviews, and many other things, so the object of recommendation must be broadened to include anything in the system. Finally, we see that in the social web a variety of types of utility may be relevant, as in the case of user recommendation, which can include friends, business contacts, or professional experts. So, a recommender system for a social web application may need a family of functions $u_x(c, r, o)$ where r is some set of requirements on top of the user profile, o can be any object contained in the application, and x stands for different types of utility that the recommendations need to satisfy.

Techniques for Recommendation

Given the diversity of avenues and needs for recommendation in the social web, it is not surprising that many different techniques have been employed. The kind of recommendation, the motivation of the user, the type of underlying item, the form of valuation, and the mode of user interaction all affect the quality of the data as well as the relative performance of the recommender. Research has shown that, in spite of the complexity, sparsity, and inherent noisiness of many of the dimensions of social web data, recommenders generally do better when they take advantages of more of these dimensions rather than fewer. This is in contrast to some results with rating-based data sets that have found little benefit to adding integrating content and collaborative data in some cases (Pilászy and Tikk 2009).

The preceding discussion will have given some sense of the breadth of research in social web recommendation and the variety of problems addressed. In the remainder of this article, we will look in more detail at the problem of recommendation in social tagging systems. This is the most developed area of research in social web recommendation, and we will look at three different approaches to the problem of integrating the multiple dimensions (users, tags, items) of the data.

Graphs

Since the multidimensionality of the social web is reflected in the graph it forms, one can apply graph-based algorithms for recommendation to exploit directly the relationships between the entities, for example, between users, items, and valuations (figure 2). The breakthrough in web search at the end of the 1990s was founded on a graph-based method: the PageRank algorithm (Brin and Page 1998). PageRank reflects the idea that a web page is important if there are many pages linking to it, and if those pages are important themselves. It is natural to apply similar methods for recommendation in the social web. The key idea of the FolkRank algorithm (Benz et al. 2010) is that an item that is tagged with important tags by important users becomes important itself. The same holds, symmetrically, for tags and users. We have thus a graph

The ECML PKDD Discovery Challenge 2009

In 2009, the Knowledge and Data Engineering group at the University of Kassel organized the second ECML PKDD Discovery Challenge. It focused on the task of recommending tags in the social bookmark and publication sharing system BibSonomy.[2] Three tag recommendation tasks were formulated, based on the experience gained in the 2008 challenge.

The first task was a cold-start task that included users previously unknown to the system, new items, and sometimes new items and new users together.

The second task concentrated on a more conventional scenario where something is known about the user and item.

The third challenge was online. Participants were invited to connect their recommendation engines to BibSonomy for live provision of results to its users. This challenge introduced the issue of response time in addition to recommendation quality.

To enable the competition, the tag recommendation feature of BibSonomy was adapted to use a multiplexer, which distributed requests to all connected recommendation engines. When a user edited a post, all recommenders got the same requests and had to provide an answer within one second. The request and all answers were stored in a database for evaluation purposes. One recommender was randomly selected and the result was presented to the user.

The system not only stored the post and the chosen tags but also allowed for monitoring the click behavior of the user. If the user was not satisfied with the recommended tags he or she could ask for another set.

The results of the challenge were very encouraging. Most of the recommenders delivered their recommendations in time even if they were connected through the Internet. But some systems that produced high-quality recommendations in offline challenges struggled to meet the real-time conditions of the online setting.

In all, 21 research groups participated in the challenge and more than 150 users downloaded the data sets. For the online task, 10 participants from seven countries came up with working solutions. Thirteen recommendation components were fielded for a five-week period. The online competition was a unique and exiting experience for the participants and for BibSonomy users, and the winning recommendation solution is still in place providing recommendations in the system. For more information see Eisterlehner et al. (2009) and Jäschke (2011).

of vertices that are mutually reinforcing by spreading their weights.

Because of the different nature of folksonomies compared to the web graph (undirected triadic hyperedges instead of directed binary edges), PageRank cannot be applied directly on folksonomies. This problem is overcome in two steps. First, we transform the hypergraph into an undirected graph. Then we apply a differential ranking approach that deals with the skewed structure of the network and the undirectedness of folksonomies and that allows for topic-specific rankings.

Folksonomy-Adapted PageRank. First we convert the folksonomy hypergraph $\mathbb{F} = (U, T, I, Y)$ into an *un*directed tri-partite graph $G_{\mathbb{F}} = (V, E)$. The set V of nodes of the graph consists of the disjoint union of the sets of tags, users, and items (that is, $V = U \cup T \cup I$). All cooccurrences of tags and users, users and items, tags and items become edges between the respective nodes. That is, each triple (u, t, i) in Y gives rise to the three undirected edges $\{u, t\}$, $\{u, i\}$, and $\{t, i\}$ in E.

Like PageRank, we employ the random surfer model, that is based on the idea that an idealized random web surfer normally follows links (for example, from an item page to a tag or a user page), but from time to time jumps to a new node without following a link (Brin and Page 1998). This results in the following definition.

The rank of the vertices of the graph is computed (like in PageRank) with the weight spreading computation

$$\mathbf{w}_{\tau+1} \leftarrow dA^{\mathrm{T}}\mathbf{w}_{\tau} + (1-d)\mathbf{p} \tag{1}$$

where $\mathbf{w}$ is a weight vector with one entry for each node in V, A is the row-stochastic version of the adjacency matrix[3] of the graph $G_{\mathbb{F}}$ defined above, $\mathbf{p}$ is the random surfer vector—which we use as preference vector in our setting, and $d \in [0, 1]$ is determining the strength of the influence of $\mathbf{p}$. By normalization of the vector $\mathbf{p}$, we enforce the equality

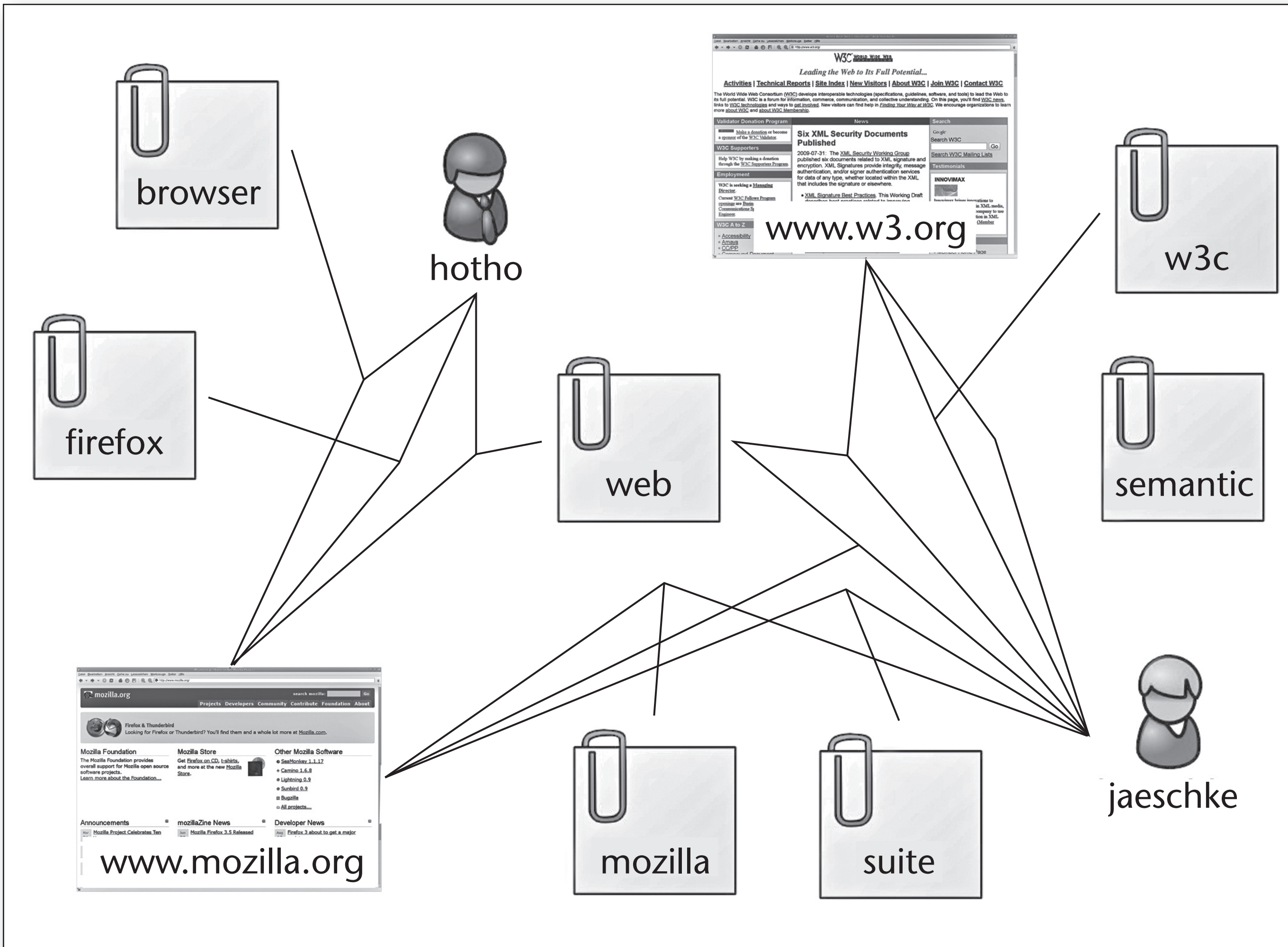

Figure 2: Excerpt of a Folksonomy Hypergraph from a Social Bookmarking System.

Each assignment of a tag to an item that a user performs is represented by a hyperedge in this tripartite graph of users, tags, and items. The example shows three posts made by two users.

$||\mathbf{w}||_1 = ||\mathbf{p}||_1$. This (together with the condition that there are no rank sinks — which holds trivially in the undirected graph $G_{\mathbb{F}}$ ensures that the weight in the system will remain constant. The rank of each node is its value in the limit $\mathbf{w} := \lim_{\tau\to\infty} \mathbf{w}_\tau$ of the iteration process.

For a global ranking, one will choose $\mathbf{p} = \mathbf{1}$, that is, the vector composed by 1s. In order to generate recommendations, however, $\mathbf{p}$ can be tuned by giving a higher weight to certain nodes. For a tag recommendation we increase the weight for the user node *u* and the item node *i*. The recommended tags are then the top tag nodes in the ranking.

As the graph $G_{\mathbb{F}}$ is undirected, most of the weight that went through an edge at step τ will flow back at $\tau + 1$. The results are thus rather similar (but not identical, due to the random surfer) to a ranking that is simply based on edge degrees. The experiments we performed showed that the topic-specific rankings are biased by the global graph structure. As a consequence, we developed the following differential approach.

FolkRank—Topic-Specific Ranking. The undirectedness of the graph $G_{\mathbb{F}}$ makes it very difficult for other nodes than those with high edge degree to become highly ranked, no matter what the preference vector is. This problem is solved by the *differential* approach in FolkRank, which computes a topic-specific ranking of the elements in a folksonomy. In our case, the topic is determined by the user/item pair (*u, i*) for which we intend to compute the tag recommendation.

1. Let $\mathbf{w}^{(0)}$ be the fixed point from equation 1 with $\mathbf{p} = \mathbf{1}$.
2. Let $\mathbf{w}^{(1)}$ be the fixed point from equation 1 with $\mathbf{p} = \mathbf{1}$, but $\mathbf{p}[u] = 1 + |U|$ and $\mathbf{p}[i] = 1 + |I|$.
3. $\mathbf{w} := \mathbf{w}^{(1)} - \mathbf{w}^{(0)}$ is the final weight vector.

Thus, the difference between the two fixed points is determined. Much of the network is unaffected by the extra emphasis placed on the nodes for *u* and i, and so the differential approach removes these neutral nodes and allows the algorithm to concentrate on those that receive weight by virtue of their association with u and i. We call the resulting weight $\mathbf{w}[x]$ of an element x of the folksonomy the FolkRank of x.

Multimode Recommendations For generating tag recommendations for a given user/item pair (u, i), we compute the ranking as described and then restrict the result set to the top tag nodes. Similarly, one can compute recommendations for users (or items) by giving preference to certain users (or items). Since FolkRank computes a ranking on all three dimensions of the folksonomy, this produces the most relevant tags, users, and items for the given users (or items).

Remarks on Complexity One iteration of the adapted PageRank requires the computation of $dA\mathbf{w} + (d-1)\mathbf{p}$, with $A \in \mathbb{R}^{s \times s}$ where $s := |U| + |T| + |R|$. If τ marks the number of iterations, the complexity would therefore be $(s^2 + s)\tau \in \mathcal{O}(s^2\tau)$. However, since A is sparse, it is more efficient to go linearly over all tag assignments in Y to compute the product $A\mathbf{w}$. After rank computation we have to sort the weights of the tags to collect the top tags.

Results. We evaluated the performance of FolkRank against other baseline methods on a data set from Delicious (Jäschke et al. 2008). The precision-recall plot in figure 3 shows how the recall increases, when more tags of the recommendation are used. Simultaneously, the precision drops. The plot reveals the quality of the recommendations given by FolkRank compared to other baseline approaches. The top 10 tags given by FolkRank contained on average 80 percent of the tags the users decided to attach to the selected item. For its top recommendations, FolkRank reaches a precision of 58.7 percent.

Factorization

The multidimensional data of the social web can also be represented in the form of a tensor that can be factorized to exploit the latent semantic structure between users, items, and valuations. Rendle et al. (2009) introduced a model to learn the factorization of the folksonomy tensor for tag recommendation. This new model focuses on a specific error function that is better suitable to predict tags in a personalized way.

Factorization Model. Given a subset S of Y that represents the training data, the goal is to learn a predictor $\hat{Y}$ that predicts the tag assignments from Y that are not known in S. Therefore, the tensor Y is estimated by the three matrices

$$\hat{U} \in \mathbb{R}^{|U| \times k_U}, \hat{I} \in \mathbb{R}^{|I| \times k_I}, \hat{T} \in \mathbb{R}^{|I| \times k_T}$$

and the core tensor

$$\hat{C} \in \mathbb{R}^{k_U \times k_I \times k_T}$$

The factorization of the predictor $\hat{Y}$ can then be expressed as follows:

$$\hat{Y} := \hat{C} \times_u \hat{U} \times_i \hat{I} \times_t \hat{T} \tag{2}$$

The low-rank feature matrices represent the corresponding users, items, and tags, respectively, by a small number of latent dimensions k_U, k_I, and k_T. The core tensor $\hat{C}$ contains the connections between the latent factors. After the parameters $\hat{C}$, $\hat{U}$, $\hat{I}$, $\hat{T}$ have been learned,

$$\hat{y}_{u,i,t} = \sum_{\tilde{u}} \sum_{\tilde{i}} \sum_{\tilde{t}} \tilde{c}_{\tilde{u},\tilde{i},\tilde{t}} \cdot \hat{u}_{u,\tilde{u}} \cdot \hat{i}_{i,\tilde{i}} \cdot t_{t,\tilde{t}} \tag{3}$$

predicts how well the tag t fits for the given user/item pair u, i.

Learning the Model. The model parameters $\hat{C}$, $\hat{U}$, $\hat{I}$, $\hat{T}$ can be learned by optimizing some quality criterion that compares $\hat{Y}$ with Y. Symeonidis, Nanopoulos, and Manolopoulos (2008) proposed to factorize Y through minimizing the elementwise loss on the elements of $\hat{Y}$ by optimizing the square loss, that is,

$$\underset{\hat{C},\hat{U},\hat{I},\hat{T}}{\arg\min} \sum_{(u,i,t) \in U \times I \times T} \left(\hat{y}_{u,i,t} - y_{u,i,t}\right)^2$$

This resembles higher order SVD (HOSVD), the multidimensional analog of singular value decomposition (SVD) for tensors. See Kolda and Bader (2009) for a recent survey.

Rendle et al. (2009), on the other hand, propose *ranking with tensor factorization*, a method for learning an optimal factorization of Y for the specific problem of tag recommendations. Therefore, the observed tag assignments for a post $(u, i) \in P_S$ are divided into positive ($T^+_{u,i}$), negative ($T^-_{u,i}$), and missing values. Only the positive and negative values are used in the optimization:

$$\underset{\hat{C},\hat{U},\hat{I},\hat{T}}{\arg\max} \sum_{(u,i) \in P_s} \frac{1}{\left|T^+_{u,i}\right|\left|T^-_{u,i}\right|} \sum_{t^+ \in T^+_{u,i}} \sum_{t^- \in T^-_{u,i}} \frac{1}{H\left(u,i,t^+,t^-\right)}$$

with

$$H\left(u,i,t^+,t^-\right) = 1 + e^{\hat{y}_{u,i,t^-} - \hat{y}_{u,i,t^+}}$$

The optimization is performed using gradient descent with a stochastic update approach.

Results. Figure 4 shows published results for the ranking with tensor factorization (RTF) technique on a data set from BibSonomy. As the graph shows, as more latent factors are included in the RTF model (RTF 8-128), F1 (the harmonic mean of recall and precision) increases, and at 128 factors, the performance is comparable to FolkRank on this tag recommendation task. The importance of optimizing only over the positive and negative observations is demonstrated by the performance of the HOSVD method, which is significantly poorer. The advantage of Rendle's approach is that one can control the trade-off between speed and quality by selecting an appropriate number of dimensions.

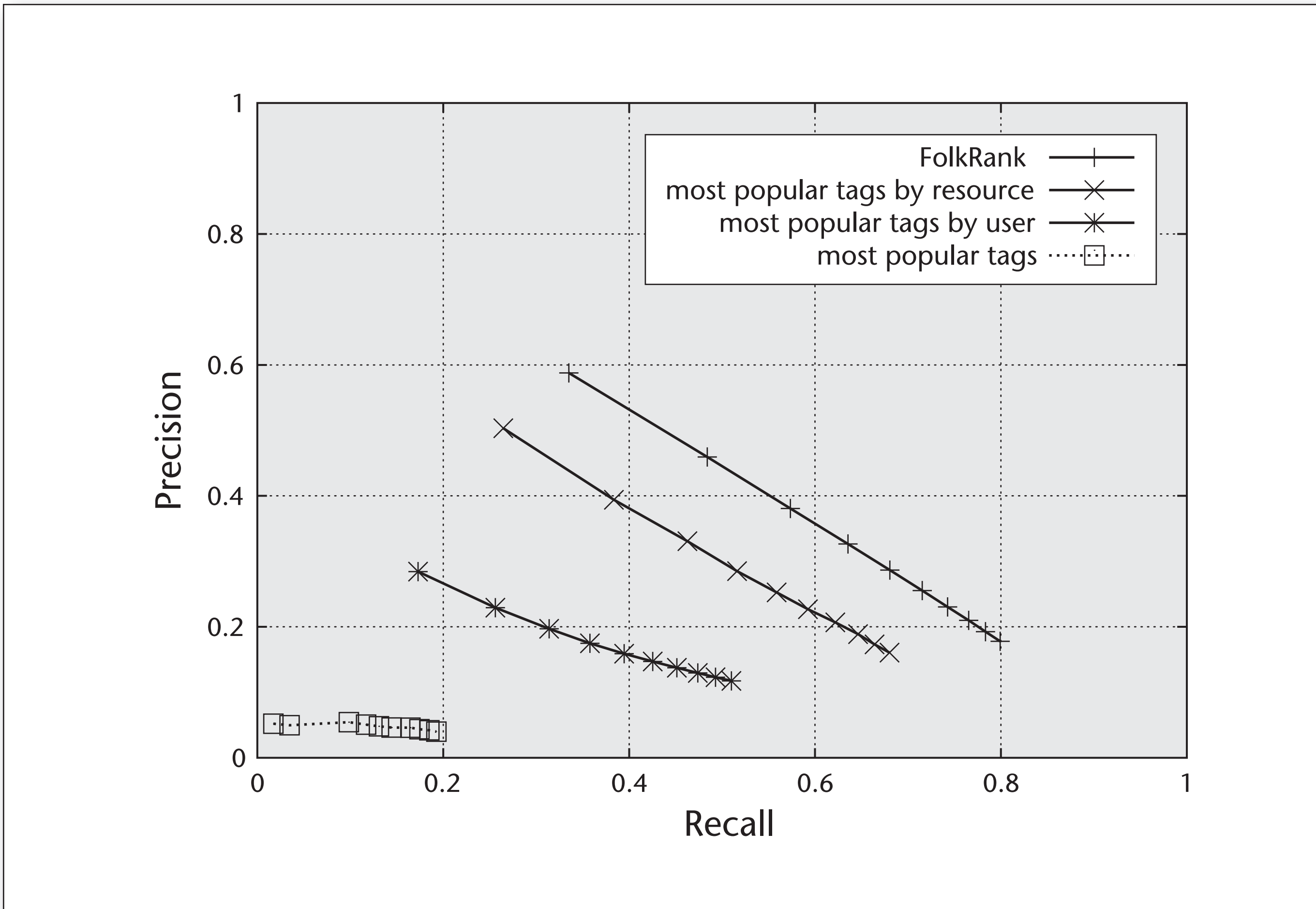

Figure 3. Recall and Precision for an Increasing Number of Recommended Tags on a 2005 Delicious Postcore at Level 10.

Hybrids

Hybrid recommenders integrate the results of several component recommenders into a single recommendation set (Burke 2002). They have been shown effective in traditional e-commerce settings, with similar performance to other integrative solutions. In the context of the social web, hybrids have been applied most notably to social annotation systems, especially tag and item recommendation (Gemmell et al. 2010a and 2010b).

Algorithm. As one example, consider the linear weighted hybrid described in Gemmell et al. (2010b). A schematic of this design is shown in figure 5. As shown, a number of different recommendation components work together to assign a recommendation score combining their individual values with a linear weight or α. The weights are learned through random-restart hill climbing.

The recommendation task in this case is to find interesting items within a social tagging system, knowing only the user and his or her prior tagging behavior. The aim of the hybrid is to achieve good performance from components that are individually simple but leverage different aspects of the data. For example, we can build a two-dimensional recommendation algorithm by ignoring the tag dimension of the data and looking just for users who have tagged similar items. Or we can ignore the items and look only for neighbors with similar tag usage. Even though each of these algorithms independently might not be terribly effective, results show that combined in a linear hybrid they can achieve performance comparable to more complex algorithms.

Results. Figure 6 presents an example of results (Gemmell et al. 2010b) using data from CiteULike,[4] a social tagging and publication sharing site, similar to BibSonomy. In this experiment, the hybrid combines six constituent components: *Pop* ignores the user and merely returns a score based on the popularity of the item. *TagSim* treats both the user and item as a vector of tags and computes the cosine similarity between the two. *KNNur* and *KNNut* employ user-based collaborative filtering modeling users as either items or as tags. *KNNru*

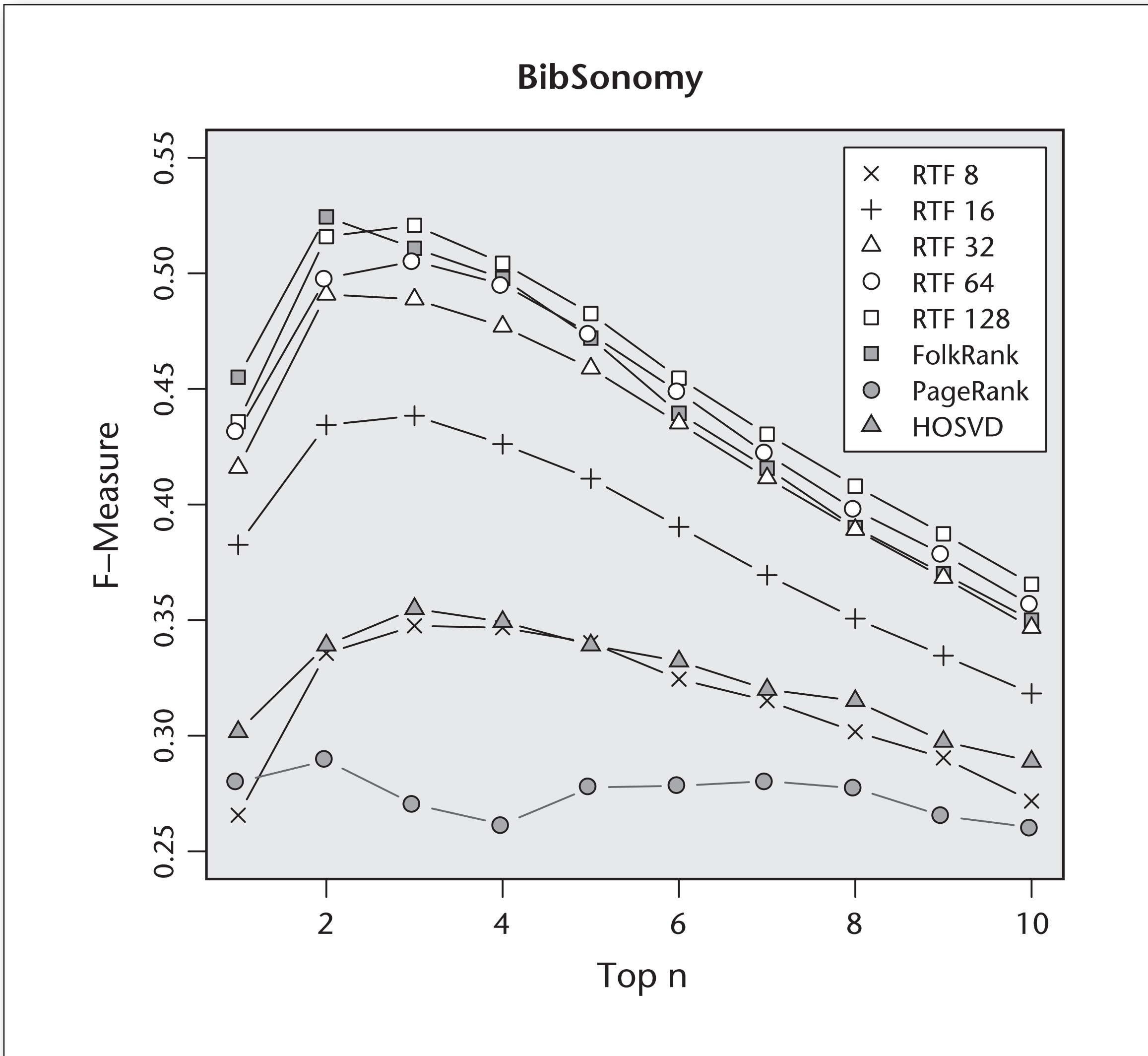

Figure 4. The F1-Measure for an Increasing Number of Recommended Tags on a BibSonomy Data Set.

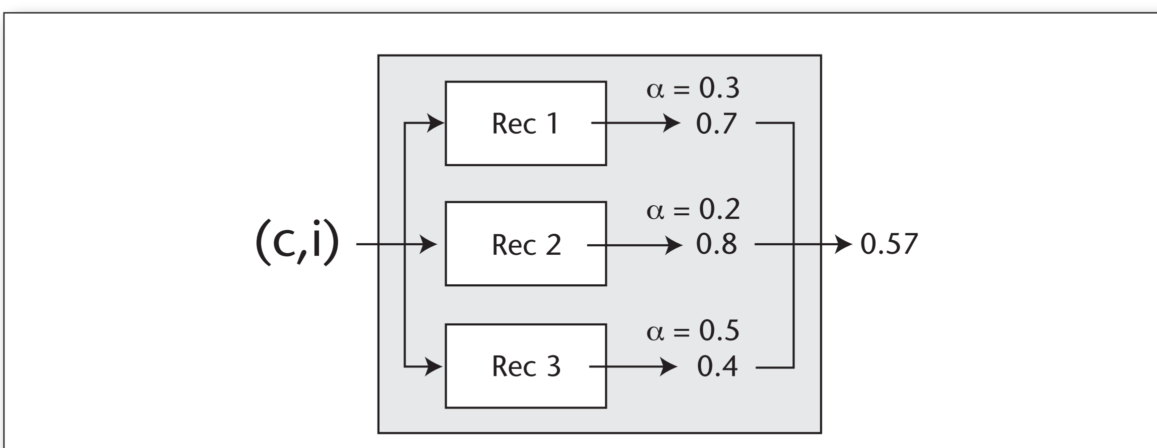

Figure 5. The Linear Weighted Hybrid.

The linear weighted hybrid takes a user and item as input and passes it along to each component. The components individually produce relevance scores, which the hybrid aggregates into a final result based on the α values.

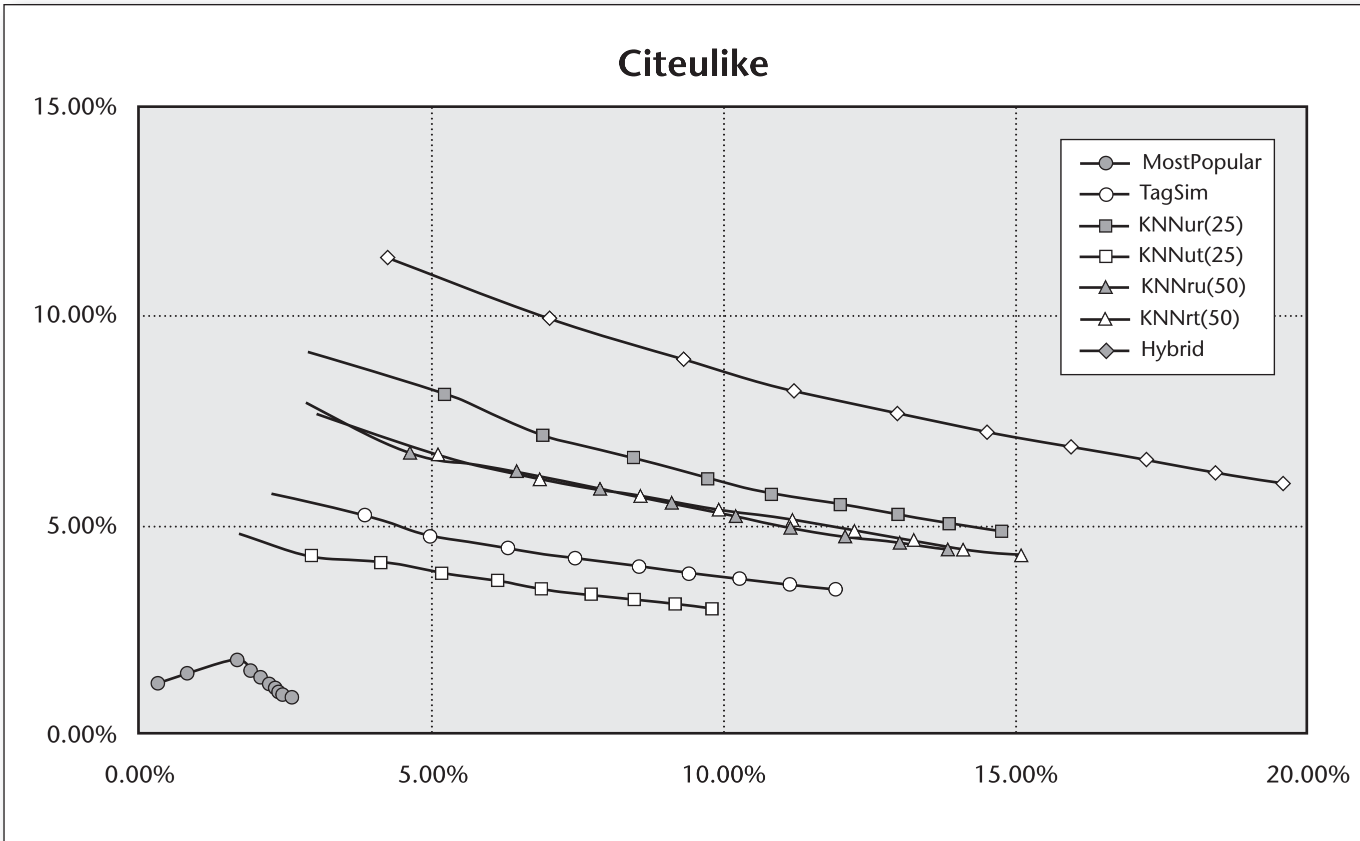

Figure 6. The Recall (x-axis) and Precision (y-axis) Plotted for Item Recommendations Sets of Size 1 through 10.

Pop	*TagSim*	*KNNur*	*KNNut*	*KNNru*	*KNNrt*
0.217	0.184	0.270	0.025	0.162	0.142

Table 1. Contribution of the Individual Components.

and *KNNrt* model items either as users or as tags in item-based collaborative filtering components.

The success of the hybrid here was achieved by leveraging all of the components, not just the strongest performing individuals, into an integrative model that exploits complementary dimensions of the data. This can be seen by examining the weights that were learned in the course of building the hybrid, shown in table 1.

All of the components contributed to the hybrid, regardless of how well they performed alone. *KNNur* for example had an α of 0.270, the largest individual contribution. *KNNut* in contrast had an α of 0.025, perhaps because it offered little information that the other user-based collaborative filtering method did not already contribute. The remaining recommenders offered significant contributions, underscoring the need to leverage multiple dimensions of the data. Another interesting case is that of the popularity-based recommender. Alone it was the worst performing recommender. As a component it had a relatively large α of 0.217, meaning it accounted for more than 20 percent of the hybrid's final score.

Discussion. The weighted hybrid is a simple, extensible, and flexible approach to the problem of integrating the multiple dimensions of social web data, and its accuracy in tag and item recommendation is comparable to other state-of-the-art algorithms. The convergence time for the alpha weights is dependent on the number of algorithms being combined, so it is essential to assemble a diverse, but not too large, collection of components. As seen in our example, diversity may be more important than raw accuracy since a poorly performing algorithm can become a strong contributor.

Conclusion

The social web is an important emerging area in recommender systems research. Compared to recommendations problems typically found in e-commerce settings, the social web presents unique challenges—in particular, the diversity of the items to which recommendation may be applied, even within a single application; the wide range of recommendation semantics that prevail depending on the nature of the application; and the complexity and multidimensional nature of the social web data to which recommendation algorithms are applied.

For these reasons, the choice of recommendation approach is likely to be highly dependent on the specific recommendation task and domain. Social tagging is the best-studied social web recommendation application, and in this article, we have illustrated three different approaches to this problem.

Notes

1. Interestingly, tag recommendation is somewhat controversial among those concentrating on the ontological aspects of tagging, since an unbiased consensus requires that valuations be made independently between users and the recommender defeats user independence.

2. See www.bibsonomy.org.

3. $a_{ij} := \frac{1}{\text{degree}(i)}$

if $\{i, j\} \in E$ and 0 else

4. See www.citeulike.org.

References

Adomavicius, G., and Tuzhilin, A. 2005. Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering* 17(6): 734–749.

Benz, D.; Hotho, A.; Jäschke, R.; Krause, B.; Mitzlaff, F.; Schmitz, C.; and Stumme, G. 2010. The Social Bookmark and Publication Management System Bibsonomy. *The VLDB Journal* 19(6): 849–875.

Brin, S., and Page, L. 1998. The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer Networks and ISDN Systems* 30(1–7): 107–117.

Burke, R. 2002. Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction* 12(4): 331–370.

Eisterlehner, F.; Hotho, A.; and Jäschke, R., eds. 2009. *ECML PKDD Discovery Challenge 2009 (DC09):* Proceedings of the International Workshop at the European Conference on Machine Learning and Printiples and Practice of Knowledge Discovery in Databases, Ceur Workshop Proceedings, volume 497. Aachen, Germany: RWTHAachen University.

Gemmell, J.; Schimoler, T.; Mobasher, B.; and Burke, R. 2010a. Hybrid Tag Recommendation for Social Annotation Systems. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*. New York: Association for Computing Machinery.

Gemmell, J.; Schimoler, T.; Mobasher, B.; and Burke, R. 2010b. Resource Recommendation in Collaborative Tagging Applications. In *Proceedings of the 11th International Conference on E-Commerce and Web Technologies*, 1–10. Berlin: Springer Verlag.

Jäschke, R. 2011. *Formal Concept Analysis and Tag Recommendations in Collaborative Tagging Systems.* Amsterdam: IOS Press.

Jäschke, R.; Marinho, L.; Hotho, A.; Schmidt-Thieme, L.; and Stumme, G. 2008. Tag Recommendations in Social Bookmarking Systems. *AI Communications* 21(4): 231–247.

Kolda, T. G., and Bader, B. W. 2009. Tensor Decompositions and Applications. *SIAM Review* 51(3): 455–500.

Pilászy, I., and Tikk, D. 2009. Recommending New Movies: Even a Few Ratings Are More Valuable Than Metadata. In *Proceedings of the 3rd ACM Conference on Recommender Systems*, 93–100. New York: Association for Computing Machinery.

Rendle, S.; Marinho, L. B.; Nanopoulos, A.; and Schmidt-Thieme, L. 2009. Learning Optimal Ranking with Tensor Factorization for Tag Recommendation. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 727–736. New York: Association for Computing Machinery.

Symeonidis, P.; Nanopoulos, A.; and Manolopoulos, Y. 2008. Tag Recommendations Based on Tensor Dimensionality Reduction. In *Proceedings of the 2008 ACM Conference on Recommender Systems,* 43–50. New York: Association for Computing Machinery.

Victor, P.; DeCock, M.; and Cornelis, C. 2011. Trust and Recommendations. In *Recommender Systems Handbook,* ed. F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, 645–675. Berlin: Springer.

Robin Burke is a professor in the College of Computing and Digital Media at DePaul University. His research interests are in recommender systems, including knowledge-based recommendation, hybrid recommendation, and recommendation applications in social web systems. He has published more than 140 articles and conference papers in these and related fields and is a member of the steering committee of the ACM Recommender Systems conference.

Jonathan Gemmell is a PhD student at DePaul University's School of Computing. Gemmell's research focuses on designing and evaluating recommendation algorithms for the social web, particularly social annotation systems. He is a senior researcher at the Web Intelligence Laboratory and a lecturer at DePaul University, teaching diverse courses ranging from data analysis and bioinformatics to the history of games.

Andreas Hotho is a professor at the University of Würzburg. He holds a Ph.D. from the University of Karlsruhe, where he worked from 1999 to 2004 at the Institute of Applied Informatics and Formal Description Methods (AIFB) in the areas of text, data, and web mining; semantic web; and information retrieval. From 2004 to 2009 he was a senior researcher at the University of Kassel. Since 2005 he has led the development of the social bookmark and publication sharing tool BibSonomy. His research focus is on the combination of machine learning/data mining and semantic web—called semantic web mining—and on the analysis of emerging structures in the web 2.0, especially in tagging systems.

Robert Jäschke studied mathematics at the Technische Universität Dresden (Germany). Since 2005 he has worked at the Knowledge and Data Engineering Group at the University of Kassel. As leading developer of the social bookmark and publication sharing system BibSonomy, his research activities are centered on the application of data-mining methods on folksonomy data and their integration into actual tagging systems. In 2010 he finished his PhD on formal concept analysis and tag recommendations in collaborative tagging systems.