Tag-Aware Recommender Systems: A State-of-the-art Survey

Zi-Ke Zhang[1,2], Tao Zhou[1,3], Yi-Cheng Zhang[1,2,*]

1 Web Sciences Center, University of Electronic Science and Technology of China - Chengdu 610054, PRC

2 Department of Physics, University of Fribourg - Chemin du Musée 1700 Fribourg, Switzerland

3 Department of Modern Physics, University of Science and Technology of China - Hefei 230026, PRC
E-mail: {zhangzike@gmail.com; zhutou@ustc.edu; yi-cheng.zhang@unifr.ch}

Abstract   In the past decade, Social Tagging Systems have attracted increasing attention from both physical and computer science communities. Besides the underlying structure and dynamics of tagging systems, many efforts have been addressed to unify tagging information to reveal user behaviors and preferences, extract the latent semantic relations among items, make recommendations, and so on. Specifically, this article summarizes recent progress about tag-aware recommender systems, emphasizing on the contributions from three mainstream perspectives and approaches: network-based methods, tensor-based methods, and the topic-based methods. Finally, we outline some other tag-related works and future challenges of tag-aware recommendation algorithms.

Keywords   social tagging systems, tag-aware recommendation, network-based, tensor-based, topic-based methods

1   Introduction

The last few years have witnessed an explosion of information that the exponential growth of the Internet [104] and World Wide Web [27] confronts us with an information overload: there are too much data and sources to be able to be found out those most relevant for us. Indeed, we have to make choices from thousands of movies, millions of books, billions of web pages, and so on. Evaluating all these alternatives by ourselves is not feasible at all. As a consequence, an urgent problem is how to automatically find out the relevant items for us. Internet search engine [10], with the help of keyword-based queries, is an essential tool in getting what we want from the web. However, the search engine does not take into account personalization and returns the same results for people with far different habits. In addition, not all needs or tastes can be easily presented by keywords. Comparatively, *recommender system* [78], which adopts knowledge discovery techniques to provide personalized recommendations, is now considered to be the most promising way to efficiently filter out the overload information. Thus far, recommender systems have successfully found applications in e-commerce [80], such as book recommendations in *Amazon.com* [59], movie recommendation-

s in *Netflix.com* [5], video recommendations in *TiVo.com* [3], and so on.

A recommender system is able to automatically provide personalized recommendations based on the historical record of users' activities. These activities are usually represented by the connections in a user-item bipartite graph [42, 112]. So far, collaborative filtering (CF) is the most successful technique in the design of recommender systems [36], where a user will be recommended items that people with similar tastes and preferences liked in the past. Despite its success, the performance of CF is strongly limited by the sparsity of data resulted from: (i) the huge number of items is far beyond user's ability to evaluate even a small fraction of them; (ii) users do not incentively wish to rate the purchased/viewed items [77]. Besides the fundamental user-item relations, some accessorial information can be exploited to improve the algorithmic accuracy [1]. User profiles, usually including age, sex, nationality, job, etc., can be treated as prior known information to filter out possibly irrelevant recommendations [46], however, the applications are mostly forbidden or strongly restricted to respect personal privacy. Attribute-aware method [93] takes into account item attributes, which are defined by domain experts. Yet it is limited to the attribute vocabulary, and, on the other hand, attributes describe global properties of items which are essentially not helpful to generate personalized recommendations. In addition, content-based algorithms can provide very accurate recommendations [70], however, they are only effective if the items contain rich content information that can be automatically extracted out, for example, these methods are suitable for recommending books, articles and bookmarks, but not for videos, tracks or pictures.

Recently, the *network theory* provides us a powerful and versatile tool to recognize and analyze such relation-based complex systems where nodes represent individuals, and links denote the relations among them. Therefore, many social, biological and technological and information systems can be represented as complex networks. In addition, a vast amount of efforts has been addressed in understanding the structure, evolution and dynamics of complex networks [2, 25, 68, 8, 20]. However, the advent of *Web 2.0* and its affiliated applications bring a new form of user-centric paradigm which can not be fully described by pre-existing models on neither unipartite or bipartite networks. One such example is the user-driven emerging phenomenon, *folksonomy* [64], which not only allows users to upload resources (bookmarks, photos, movies, publications, etc.) but also freely assign them with user-defined words, so-called *tags*. Folksonomy requires no specific skills for user to participate in, broadens the semantic relations among users and resources, and eventually achieves its immediate success in a few years. Presently, a large number of such applications can be found online, such as *Del.icio.us* (with tags of bookmarks by users), *MovieLens* (with ratings of movies by users), *CiteULike* (with tags of publications by users), *BibSonomy* (with tags of bookmarks and references by users), *Flickr* (with tags of images by users), *Last.fm* (with tags of music by users) etc. From the view of physics, all these online systems have performed similar statistical properties, e.g. Zipf's law like rank-frequency distribution [15] and Heaps' laws growth phenomenon [14], between which the in-depth understanding are studied

---

http://del.icio.us/
http://www.movielens.org/
http://www.citeulike.com/
http://www.bibsonomy.org/
http://www.flickr.com/
http://www.last.fm/

in recent works [109, 62]. With the help of those platforms, users can not only store their own resources and manage them with social tags, but also look into other users' collections to find what they might be interested in by simply keeping track of the baskets with tags. Unlike traditional information management methods where words (or indices) are normally pre-defined by experts or administrators, e.g. the library classification systems. A tagging system allows users to create arbitrary tags that even do not exist in dictionaries. Therefore, those user-defined tags can reflect user behaviors and preferences with which users can easily make acquaintance, collaborate and eventually form communities with others who have similar interests [82].

2 Overview of Tag-based Recommender Systems

Nowadays, people are confronting huge amount of information and making much effort in searching relevant or interesting items. However, as discussed in previous section, it is impossible for individuals to filter metadata from various structures and massive number of sources, especially in a user-generated information era [69]. The motivation of users' contribution is straightforward: they build their own data based on which they become further involved in web-based communications. Social tagging is becoming one of most popular tools in playing important rules among various social activities. Ding *et al.* [24] provided good overviews of social tagging systems with emphasis on both its social impact and ontology modeling.

As a consequence, social tags can be naturally considered as kind of additional yet useful resource for designing effective recommendation algorithms. Firstly, tags are freely associated by users, which can reflect their personalized preferences. Secondly, tags express

the semantic relations among items, which can help evaluating the underlying item qualities. Thirdly, the co-occurrence properties of tags can be employed to build both user communities and item clusters, which be further made use of to find relevant yet interesting items for targeted individuals. Therefore, tags provide us a promising way to solve some stubborn problems in recommender systems, e.g. the *cold-start* problem [108].

Up to date, a remarkable amount of researches have discussed how to apply tags in the domain of recommender systems. Hotho *et al.* [39] proposed a modified PageRank [10] algorithm, namely *FolkRank*, to rank tags in folksonomies with the assumption that important tags are given by important users, which is akin to HITS [48] algorithm in internet networks. The *FolkRank* is then be adopted to recommend tags [43]. In addition, due to the user-generated property, tags are considered to have high personalized information, hence can be used to design methods for both personalized searching [102] and recommendation. A good overview of social bookmarking and its applications in recommender systems can be found in a recent Ph.D. thesis [9]. However, although tags are especially useful for both organizing and searching resources, there are also many studies arguing that not all tags can benefit recommendation [6] because of the various limitations of tags, such as *polysemy, synonymy, ambiguity* [64, 99, 89, 55], etc. These flaws are also the side effects of the uncontrolled vocabulary, thus it remains some open issues in tagging systems: (i) singularity vs. plurality: e.g. the words *cat* and *cats* somehow have very similar meanings, however, refer to two different words in tagging systems; (ii) polysemy vs. synonymy: e.g. the word *apple* may refer to a kind of fruit, while it can also indicate the well-known computer company, *Apple Inc.*, as well as its products; on the other hand, the words *mac, macintosh,* and *apple* all point to the

products of *Apple Inc.*, however, it fails again to uncover their underlying relations in tagging systems; (iii) different online tagging systems allow users to give different formats of tags, e.g. *Del.icio.us* only allows words to be assigned, which subsequently results in compound words with various symbols (e.g. underline, dashline, colon, etc.), leading to an unlimited formats of metadata. Such freestyle tags additionally exemplify the explosion of observed datasets, hence interfere in the analyses of the structure and user behaviors in tagging systems. Recently, researches have devoted much effort to solve those issues. Firstly, clustering-based methods [87, 13] are proposed to alleviate the word reduction problem. Secondly, semantic methods are discussed to use ontology-based to organize tags and reveal the semantic relations among them [67, 47]. Thirdly, dimension reduction and topic-based methods are put forward to discover the latent topics [91, 98], and graph-based methods are proposed [110, 61] to solve the sparsity problem in large-scale datasets.

In the following, we firstly give the evaluation metrics measured in this survey. Secondly we summarize some of the most recent and prominent tag-aware recommendation algorithms, showing and discussing how they make use of the aforementioned representations to address the some unresolved issues in recommender systems. Basically, there are three kinds of recommendations in social tagging systems: (i) predicting friends to users; (ii) recommending items to users; (iii) pushing interesting topics (tags) to users. However, as mentioned above, the most urgent problem in information era is to filter irrelevant items for individuals, therefore, in this survey, we mainly discuss the second case, and introduce some related methods discussing (i) or (iii) if necessary. Finally, we conclude with comparison of the surveyed methods and outline some future challenges of tag-aware recommendation algorithms.

3   Tag-Aware Recommendation Models

Formally, a social tagging network consists of three different kind of communities: users, items and tags, which subsequently form an entry set of personalized folksonomy, *personomy* [60], each follows the form $\mathbb{F}=\{$user, item, tag_1, tag_2, $\cdots$, $tag_t\}$, where t is the number of tags assigned to this item by the very user. Correspondingly, in a recommender system, a full folksonomy can be considered in two ways: (i) to be consisted of three sets, respectively of users $U = \{U_1, U_2, \cdots, U_n\}$, items $I = \{I_1, I_2, \cdots, I_m\}$, and tags $T = \{T_1, T_2, \cdots, T_r\}$. Consequently, each binary relation can be described by a adjacent matrix, A, A' and A'' for user-item, item-tag and user-tag relations, respectively. If U_i has collected I_j, we set $a_{ij} = 1$, otherwise $a_{ij} = 0$. Analogously, we set $a'_{jk} = 1$ if I_j has been assigned by the tag T_k, and $a'_{jk} = 0$ otherwise. Furthermore, the users' preferences on tags can be represented by a adjacent matrix A'', where $a''_{ik} = 1$ if U_i has adopted T_k and $a''_{ik} = 0$ otherwise; (ii) a ternary [91, 75] or hypergraph [31, 114, 106] based structure: only complete ternary relation is taken into account to be existed as a real link. That is to say, each relation of (u, i, t), represented as an existing component $Y = 1$, if it exists in a folksonomy $\mathbb{F}$, and $Y = 0$ otherwise.

3.1   Evaluation Metrics

For a traditional recommender system, each data set, E, is randomly divided into two parts to test the performance of proposed algorithms: the training set, E^P, is treated as known information, while the testing set, E^T, is used for testing. In this survey, the training set always contains 90% of entries, and the remaining 10% of entries, constitute the testing set. In addition, each division should guarantee $E^T \bigcap E^P = \varnothing$ and $E^T \bigcup E^P = E$ in order to make sure no redundant information is used. Furthermore, To give solid and compre-

hensive evaluation of the proposed algorithm, we consider metrics of both accuracy [33] and diversity [111] to characterize the performance of recommendations.

3.1.1 Metrics of Accuracy

1. *Ranking Score* (*RS*) [112].— In the present case, for each entry in the testing set (i.e. a user-item pair), *RS* is defined as the rank of the item, divided by the number of all uncollected items for the corresponding user. Apparently, the less the *RS*, the more accuracy the algorithm is. $\langle RS \rangle$ is given by averaging over all entries in the testing set.

2. *The area under the ROC curve* [34, 19].— In the present case, the area under the ROC curve, abbreviated by AUC, for a particular user is the probability that a randomly selected removed item for this user (i.e., an item in the testing set and being collected by this user) is given a higher score by our algorithm than a randomly selected uncollected item (i.e, an item irrelevant to this user in neither the training set nor the testing set). The AUC for the whole system is the average over all users. If all the scores are generated from an independent and identical distribution, $AUC \approx 0.5$. Therefore, the degree to which the AUC exceeds 0.5 indicates how much better the algorithm performs than pure chance.

3. *Recall* [36].— Note that, the AUC takes into account the order of all uncollected items, however, in the real applications, user might only care about the recommended items, that is, the items with highest scores. Therefore, as a complementary measure, recall is employed to quantify the accuracy of recommended items, which is defined as:

$$Recall = \frac{1}{n} \sum_{i=1}^{n} N_r^i / N_p^i, \qquad (1)$$

where N_p^i is the number of items collected by U_i in the testing set, and N_r^i is the number of recovered items in the recommendations for U_i. We use the averaged recall instead of simply counting N_r/N_p with $N_r = \sum_i N_r^i$ and $N_p = \sum_i N_p^i$ since it is fair to give the same weight on every user in the algorithm evaluation. Assuming the length of recommendation list, L, is fixed for every user, recall is very sensitive to L and a larger L generally gives a higher recall.

3.1.2 Metrics of Diversity

1. *Inter Diversity* (*InterD*) [113, 112].— *InterD* measures the differences of different users' recommendation lists, thus can be understood as the inter-user diversity. Denote I_R^i the set of recommended items for user U_i, then

$$InterD = \frac{2}{n(n-1)} \sum_{i \neq j} \left(1 - \frac{|I_R^i \cap I_R^j|}{L} \right), \qquad (2)$$

where $L = |I_R^i|$ is the length of recommendation list. In average, greater or less *InterD* mean respectively greater or less personalization of users' recommendation lists.

2. *Inner Diversity* (*InnerD*) [113].— *InnerD* measures the differences of items within a user's recommendation list, thus can be considered as the inner-user diversity. It reads,

$$InnerD = 1 - \frac{2}{nL(L-1)} \sum_{i=1}^{n} \sum_{j \neq l, j, l \in I_R^i} S_{jl}, \qquad (3)$$

where $S_{jl} = \frac{|\Gamma_{I_j} \cap \Gamma_{I_l}|}{\sqrt{|\Gamma_{I_j}| \times |\Gamma_{I_l}|}}$ is the cosine similarity between items I_j and I_l, where Γ_{I_j} denotes the set of users having collected object I_j. In average, greater or less $InnerD$ suggests respectively greater or less topic diversification of users' recommendation lists.

3.2   Network-based Models

Recently, there are a variety of attempts utilizing tagging information for recommendation from a perspective of graph theory, Generally, a tag-based network can be viewed as a tripartite graph which consists of three integrated bipartite graphs [112] or a hypergraph. Therefore, network-based methods are widely used to describe the tag-based graph. Up to date, bipartite graph has been largely applied to depict massive number of online applications. For example, users rate movies, customers comment books, individuals participate in online games, etc. In a typical bipartite graph, there are two mutually connected communities, which contrastively have no link within each community, shown in Fig. 1.
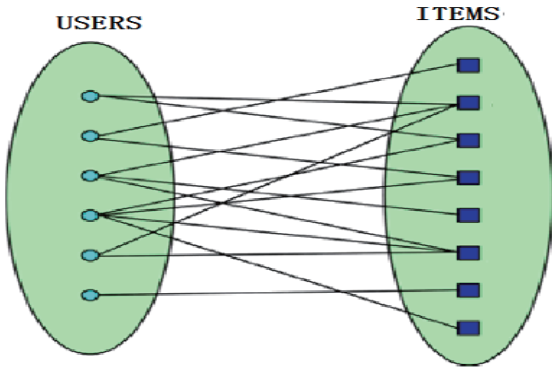

Fig. 1. (Color online) Illustration of a user-item bipartite network [83] composed by 6 users and 8 items, in which only inter-community links are allowed.

Inspired by this elegant structure, two underlying network-based methods: Probability Spreading (ProbS) [112, 111] and Heat Spreading (HeatS) [105, 111], were proposed as a starting point to apply network theory in recommender systems.

ProbS is also known as random walk (RW) in computer science and mass diffusion (MD) in physics. Given a target user U_i, ProbS will generate final score of each item, f_j, for her/him according to following rules:

Suppose that a kind of resource is initially located on items. Each item averagely distributes its resource to all neighboring users, and then each user redistributes the received resource to all his/her collected items. The final resource vector for the target user U_i, $\vec{f^p}$, after the two-step mass diffusion is:

$$f_j^p = \sum_{l=1}^{n} \sum_{s=1}^{m} \frac{a_{lj} a_{ls} a_{is}}{k(U_l) k(I_s)}, \; j = 1, 2, \cdots, m, \; (4)$$

where $k(U_l) = \sum_{j=1}^{m} a_{lj}$ is the number of collected items for user U_l, and $k(I_s) = \sum_{i=1}^{n} a_{is}$ is the number of neighboring users for item I_s.

Comparatively, HeatS works based on the reverse rules of ProbS. At each step, each target will receive resources according to how active or popular it is, while ProbS distributes resources based on its own activity or popularity. Thus, the final resource vector for the target user U_i, $\vec{f^h}$, after the two-step heat spreading is:

$$f_j^h = \frac{1}{k(I_j)} \sum_{l=1}^{n} \sum_{s=1}^{m} \frac{a_{lj} a_{ls} a_{is}}{k(U_l)}, \; j = 1, 2, \cdots, m,$$
$$(5)$$

Therefore, HeatS will depress the score of popular items and is inclined to recommend the relatively cold ones, while ProbS will enhance the scoring ability of popular items.

Based on the aforementioned methods, a variety of algorithms have been proposed to add tags in order to generate better recommendation performance. Zhang *et al.* [110] firstly proposed a tag-aware diffusion-based method, considering tags as additional information, which
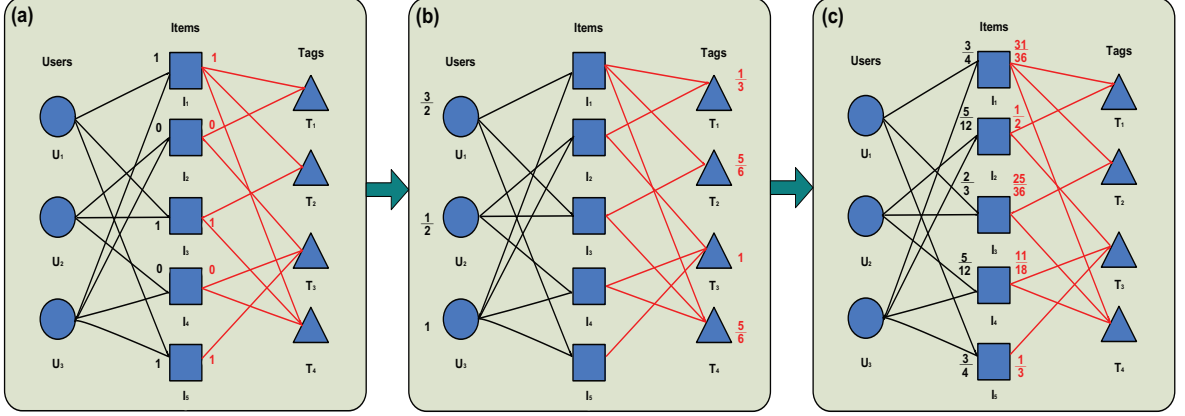
Fig. 2. (Color online) Illustration of a user-item-tag tripartite graph consists of 3 users, 5 items and 4 tags, as well as the recommendation process described in [110]. The tripartite graph is decomposed to user-item (black links) and item-tag (red links) bipartite graphs connected by items. For the target user U_1, the scoring process works as: (a) firstly, highlight the items, I_1, I_3, I_5, collected by the target user U_1 and mark them with unit resource. That is to say: $f_{I_1} = f_{I_3} = f_{I_5} = 1$, and $f_{I_2} = f_{I_4} = 0$. (b) secondly, distribute the resources from items to their corresponding users and tags, respectively; e.g. $f_{U_3} = f_{I_1} * \frac{1}{2} + f_{I_2} * \frac{1}{2} + f_{I_5} * \frac{1}{2} = 1 * \frac{1}{2} + 0 + 1 * \frac{1}{2} = 1$ and $f_{T_4} = f_{I_1} * \frac{1}{3} + f_{I_3} * \frac{1}{2} + f_{I_4} * \frac{1}{2} = 1 * \frac{1}{3} + 1 * \frac{1}{2} + 0 = \frac{5}{6}$; (c) finally, redistribute the resources from users and tags to their neighboring items. e.g. $f_{I_4}^p = f_{U_2} * \frac{1}{3} + f_{U_3} * \frac{1}{4} = \frac{1}{2} * \frac{1}{3} + 1 * \frac{1}{4} = \frac{5}{12}$ and $f_{I_4}^{pt} = f_{T_3} * \frac{1}{3} + f_{T_4} * \frac{1}{3} = 1 * \frac{1}{3} + \frac{5}{6} * \frac{1}{3} = \frac{11}{18}$.

extended the resulting paradigm as reduced bipartite graphs, known as *tripartite graph*. In such a graph, one kind of nodes (users, items or tags) plays as a centric role to bridge the remaining two. Fig. 2 shows an example of item-centric model. In such a graph, each item of a target user will respectively distribute to its neighboring users and tags, and then all the items in database will receive their resources from their neighboring nodes. Hence, the final resource for the target user U_i, $\vec{f^t}$, after two-step mass diffusion (see Fig. 2), will be integrated in a linear way:

$$f_j^t = \lambda f_j^p + (1 - \lambda) f_j^{pt}, \qquad (6)$$

where $f_j^{pt} = \sum_{l=1}^{r} \sum_{s=1}^{m} \frac{a_{is} a'_{ls} a'_{lj}}{k'(I_s)k(T_l)}$ is the resource of item j received from item-tag bipartite graph, $k(T_l) = \sum_{j=1}^{m} a'_{jl}$ is the number of neighboring items for tag T_l, $k'(I_s) = \sum_{l=1}^{r} a'_{sl}$ is the number of neighboring tags for item I_s, and $\lambda \in [0, 1]$ is a tunable parameter to obtain the optimal performance. Table 1 shows the corresponding AUC results for three datasets: *Del.icio.us*, *MovieLens* and *BibSonomy*, in which the AUC values are significantly improved by considering item-tag bipartite relation. In addition, [110] also experimentally demonstrated that the incorporation of tags can enhance the Recall results for various ranges of recommendation length. Besides the accuracy, [110] extensively showed that tags could also promote the recommendation diversification, hence enlarge the selection vision for users.

Recently, a variety of researchers have designed tag-aware algorithms by modifying the above model. Shang *et al.* [85] proposed a user-centric diffusion-based similarity, which considered users as the communication hubs to measure the coincidence among users, and found it could obtain more accurate recommendations. In addition, the tag usage frequency were measured as edge weight in user-item bipartite

networks. Shang and Zhang [84] directly regarded the frequency as weight and applied diffusion method to improve the recommendation accuracy. Wu and Zhang [100] viewed the tag usage patten in a document vocabulary manner and applied the inverse document frequency (TF-IDF) model [79] to calculate the weight for user-item relations. They found this weighting method could enhance the recommendation diversity. Furthermore, Zhang *el al.* [108] took such tag usage frequency into account on the user-tag and then spread the tag-based preferences to all the corresponding tags' neighboring items. The numeric results showed it could significantly enhance the algorithmic accuracy for relatively inactive or new users, and it also found that different tag usage patterns might result in different algorithmic diversity: the more diverse topic of tags users like, the more diverse results the algorithm would generate. Consequently, two fundamental roles of tags [106, 107], describing and retrieving items, were firstly found applications in recommender systems. Up to date, Liang *et al.* [58] have noticed that the above methods decomposed the user-item-tag relationships into two bipartite graphs and made recommendations, which, to some extents, ignored the remaining one binary relation (e.g. user-tag for [110], user-item for [108]). As a result, by further eliminating the noise of tags, they used the semantic meaning of tags to represent topic preferences of users and combined it with item preferences of users to measure user-based similarity. Subsequently, the hybrid similarity was used in a standard collaborative filtering framework to obtain better *Recall* results in two datasets: *Amazon.com* and *CiteULike.com.* Similar measurements of user-based and item-based similarities were also widely applied by various researches [92, 94].

3.3 Tensor-based Models

Recently, the tensor factorization (TF) [49] based method has attracted increasing at-

tention to be applied in designing recommendation algorithms in social tagging systems [103, 91, 90, 44, 75]. Generally, by using tensor, a ternary relation, $\mathbb{A} = \{u, i, t\}$, can be represented as [90]

$$a_{(u,i,t)} = \begin{cases} 1, if (u, i, t) \subseteq Y, \\ 0, otherwise. \end{cases} \tag{7}$$

There are also other researches that define the missing values for empty triples in which the items have never been tagged, while the negative values are set for the triples in which the items are tagged in other tensors [75]. Fig. 3 shows the illustration of the above two definitions.
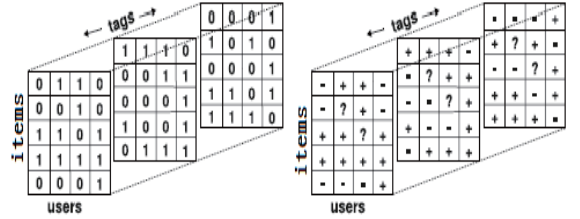

Fig. 3. Illustration of tensor-based tag assignment ternary relation. Left panel: a visible tag assignment, $a_{(u,i,t)}$, is set 1, and 0 otherwise [90]. Right panel: $a_{(u,i,t)}$ is set negative as the triple of which the item is tagged in other existing triples rather than $a_{(u,i,t)}$. The missing values are given to other empty triples [75].

For the purpose of recommendation, Y can be represented by three low-rank *feature metrics*, $\hat{U}$, $\hat{I}$, $\hat{T}$ and one *core tensor*, $\hat{C}$, shown as

$$\hat{Y} = \hat{C}_{\times u}\hat{U}_{\times i}\hat{I}_{\times t}\hat{T}, \tag{8}$$

where the core tensor $\hat{C}$ and the feature matrices $\hat{U}$, $\hat{I}$ and $\hat{T}$ are the parameters to be learned and $\times x$ is the tensor x-mode dimension multiplication factor between a tensor and a matrix [75]. In addition, the size of feature

Table 1. Comparison of algorithmic accuracy, measured by the AUC. *Pure U-I* and *Pure I-T* denote the pure diffusions on user-item bipartite graphs and item-tag bipartite graphs, respectively corresponding to $\lambda=1$ and $\lambda=0$. The optimal values of λ as well as the corresponding optima of AUC are presented for comparison.

Data set	Pure U-I	Pure I-T	Optimum	λ_{opt}
Del.icio.us	0.8098	0.8486	0.8588	0.32
MovieLens	0.8065	0.8163	0.8233	0.44
BibSonomy	0.7374	0.7600	0.7852	0.44

matrices are:

$$\hat{C} \subseteq \mathbb{R}^{k_U \times k_I \times k_T}, \hat{U} \subseteq \mathbb{R}^{|U| \times k_U},$$
$$\hat{I} \subseteq \mathbb{R}^{|I| \times k_I}, \hat{T} \subseteq \mathbb{R}^{|T| \times k_T}, \tag{9}$$

where k_U, k_I, k_T are the latent dimensions of the low-rank approximations for users, items and tags, respectively. Then, recommendations can be generated as

$$\hat{y}_{(u,i,t)} = \sum_{\hat{u}} \sum_{\hat{i}} \sum_{\hat{t}} \hat{c}_{(\hat{u},\hat{i},\hat{t})} \cdot \hat{u}_{(u,\hat{u})} \cdot \hat{i}_{(i,\hat{i})} \cdot \hat{t}_{(t,\hat{t})}, \tag{10}$$

where the tilde denotes the feature dimensions and the hat indicates the elements of the feature matrices. Finally, the personalized recommendations list of items or tags will be displayed to the target user in a descending order.

The tensor factorization is based on Singular Value Decomposition (SVD) [96], with which the ternary relation can be reduced to low dimensions, hence easier to be proceeded for recommendation. [91] used it corresponding to a TF model optimized for square-loss where all not observed values are learned as 0s. In further, [90] developed a unified framework to model the three types of entities. Then, the three-order tensor dimension decomposition was performed by combining Higher Order Singular Value Decomposition (HOSVD) [21] method and the Kernel-SVD [86, 17] smoothing technique on two real-world datasets: *Last.fm* and *BibSonomy*. The results showed improvements in Recall and Precision. [75] proposed a better learning approach for TF models, which optimized the model parameters for the AUC values. The optimization of this model is related to Bayesian personalized ranking (BPR) proposed in [76]. They both tried to optimize over pairs of ranking constraints, where the former focused on AUC optimization, and the latter optimized for pair classification. [74] discussed the relationship between them in details.

3.4 Topic-based Models

Generally, the *core* challenge of recommender systems is to estimate the likelihood between users and items. In the last two decades, many efforts have been devoted to build various models to measure such probabilities in information retrieval. Deerwester *et al.* [22] proposed Latent Semantic Analysis (LSA) to use a term-document matrix describing the occurrences of terms in documents. Normally, each element in the matrix is weighted by TF-IDF [79] revealing the importance of the very term in its corresponding documents. In addition, Hofmann [38] introduced the Probability Latent Semantic Analysis (PLSA) to improve recommendation quality for various settings by assuming a latent lower dimensional topic model as origin of observed co-occurrence distributions. Comparing with the standard LSA, PLSA is based on a mixture decomposition derived from a latent topic model which

would statistically result in a more principled approach having a solid foundation. Eq. 11 gives a formula way of PLSA

$$
\begin{aligned}
P(w,d) &= \sum_z P(z)P(d|z)P(w|z) \\
&= P(d)\sum_z P(z|d)P(w|z),
\end{aligned}
\tag{11}
$$

where word w and document d are both generated from the latent topic z, which is chosen conditionally to the document according to $P(z|d)$, and a word is then generated from that topic according to $P(w|z)$. However, PLSA does not allocate the topic distribution for each document, which might potentially lose information of documents with multiple subjects. Therefore, recently, a more widely used model, Latent Dirichlet Allocation (LDA) [7], was proposed to overcome this issue by allowing multiple latent topics with a priori Dirichlet distribution, a conjugate prior of multinomial distribution, assigned to each single document. Besides, LDA assumes that the documents are represented as random mixtures over the latent topics, each of which is given by a distribution over words. For each document d in collection D, LDA works as (see Fig. 4):
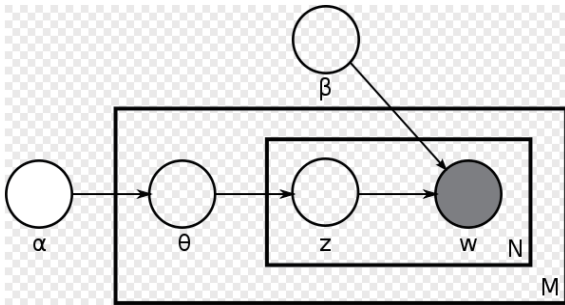

Fig. 4. (Color online) Illustration of generative process for LDA model (from wikipedia.org), where α, β, θ are parameters to be learned, z is the latent topic variable, w is observed variable of words, and the direction of arrows indicates the process flow.

(i) Choose θ_i from $Dir(\alpha)$, where i runs over the document collection; (ii) For each

word w_{ij} in document d_i, choose a latent topic $z_{ij} \sim Multinomial(\theta_i)$ and then choose a word $w_{ij} \sim Multinomial(\beta_{z_{ij}})$. Finally, after learning the parameters by Gibbs sampling [29] or expectation-maximization (EM) algorithm [23], the probability of the document collection can be given as

$$
\begin{aligned}
P(D|\alpha,\beta) = {} & \prod_i \int p(\theta_i|\alpha)* \\
& \left(\prod_j \sum_{z_{ij}} p(z_{ij}|\theta_i)p(w_{ij}|z_{ij},\beta) \right) d\theta_i
\end{aligned}
\tag{12}
$$

Recently, those topic-based models are applied in social tagging systems for both tag and item recommendations. In [98, 95], they proposed a PLSA-based hybrid approach unifying user-item and item-tag co-occurrence to provide better item recommendations. In these two works, they measured the co-occurrence probabilities of user-item and item-tag by summing over the latent topic variables, and then maximized the likelihood of fused scenarios.

Comparatively, LDA is more widely used for tag recommendation. Xi *et al.* [88] employed LDA for eliciting topics from the words in documents, as well as the co-occurrence tags, where words and tags form independent vocabulary spaces, and then recommended tags for target documents. Krestel *et al.* [53, 52], on the other hand, used LDA to extract hidden topics from the available tags of items and then recommended tags from these latent topics. Bundschus *et al.* [12] integrated both user information and tag information into LDA algorithm. Its generative process extracted user specific latent topics using a Topic-Tag Model adding tags and User-Topic-Tag Model adding the user layer. It assumed that users had a multinomial distribution over topics, hence, the users' interests could be modeled by each tag assignment. Finally, they used two-step latent topic realizations (user-item based and tag-based topics) to provide personalized tag rec-

ommendations. In addition, Bundschus *et al.* [11] summarized different topic modeling approaches with respect to their ability to model annotations. Different from applying Bayesian rule to decompose the joint probability of item-tag and user-tag co-occupance, Harvey *et al.* [35] introduced a similar LDA-based approach for tag recommendation by decomposing the joint probability of latent topics given the tag assignments. Furthermore, Li *et al.* [57] combined LDA and GN community detection algorithm [32, 56] to observe the topic distributions of communities, as well as community evolving over time in social tagging systems. On this basis, they found that users in the same community tended to be interested in similar topics, which would shed some lights on recommendation for groups.

4 Conclusions and Outlook

In this survey, we summarized the progress of studies on tag-aware recommender systems (RS), emphasizing on the recent contributions by both statistical physicists and computer scientists in three aspects: (i) network-based methods; (ii) tensor-based methods; (iii) topic-based methods. Generally, there is no single method that can full address all the problems existing in RS. Network-based and tensor-based methods can overcome the sparsity of large-scale data, hence can be used for designing efficient algorithms. However, they only focus on the network structure, while lack considerations of relations among tags. Comparatively, topic-based methods can distinguish tags into different topics, hence can produce more meaningful and understandable recommendations. However, since most of topic-based methods use machine learning to iteratively refine the results, they require high-efficient hardwares for computation, and thus consume more computation time. Similar problem lies in tensor-based methods for dimension reduc-

tion process. Therefore, a unified model might be considered to fully make use of their advantages and provide a more promising method in tag-aware recommender systems.

Nowadays, RS is not a new problem in information science, the advent of new Web2.0 paradigms bring versatile tools and information to help build better recommendation models by integrating traditional methods. Recently, the studies of complex networks would benefit tag-based algorithms, because the in-depth understanding of network structure, user behaviors and network dynamics can be used to design advanced tag-aware recommendation algorithms (e.g., making use of the information about hypergraph [106, 97] and tripartite graph [54, 16] of social tagging networks to better predict underlying interests). On the other hand, tag-based algorithms can also help the trend detection [40] over time.

Although the studies of tag-aware recommender systems have achieved fruitful goals, there are still challenges, as well as some new directions remained to be solved (discovered) in future: (i) the complete hypergraph [114, 106] should be well considered to fully address the integrity of tagging networks without decomposing any information and thus is a promising way to provide recommendations with better performance; (ii) most of current related researches emphasize on recommending single type of nodes, however, predicting the joint node pairs (e.g. item-tag pair [71]) comparatively lacks of study. The joint pair recommendation would provide more personalized preference, hence be a new application of tag-aware recommender systems; (ii) since the tags are freely assigned by users, which consequently results in much noise of added tags. Tag clustering [87, 30, 73] methods and anti-spam [51] technique would be both promising ways to reduce the noise and help provide high-quality recommendations; (iv) the probability-based models are mainly used to provide tag recom-

mendations in most recent researches, while how to well use them to benefit item recommendations is still an open challenge. In addition, those models would also help to prevent rumor spreading [72, 45] and trend detection [50]; (v) the multi-layered network [65] consists of user social interactions, tag co-occurrence relations and user-item-tag ternary infomation can be considered to describe the hierarchical structure of social tagging systems, and thus the Social Network Analysis (SNA) [81] and social influence [66, 41, 18] based techniques can be used to provide more substantial recommendations, and social predictions [37, 63] as well; (vi) most tagging platforms are dynamical systems and evolve over time [26], thus the study of human dynamics [4] in analyzing the temporal behaviors and interests can provide real-time recommendations [28, 101].

5 Acknowledgements

References

[1] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Trans. Knowl. Data Eng.*, 17:734, 2005.

[2] R. Albert and A.-L. Barabási. Statistical Mechanics of Complex Networks. *Rev. Mod. Phys.*, 74:47, 2002.

[3] K. Ali and W. van Stam. TiVo: making show recommendations using a distributed collaborative filtering architecture. In *Proc. 10th ACM SIGKDD Intl. Conf. Knowl. Disc. Data Min.*, page 394, 2008.

[4] A.-L. Barabási. The origin of bursts and heavy tails in human dynamics. *Nature*, 435:207–211, 2005.

[5] J. Bennett and S. Lanning. The netflix prize. In *Proc. KDD Cup Workshop*, 2007.

[6] K. Bischoff, C. Firan, W. Nejdl, and R. Paiu. Can all tags be used for search? In *Proc. 17th ACM Conf. Info. Knowl. Manag.*, pages 193–202, 2008.

[7] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3:993–1022, 2003.

[8] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Huang. Complex Networks: Structure and Dynamics. *Phys. Rep.*, 424:175, 2006.

[9] T. Bogers. *Recommender Systems for Social Bookmarking.* PhD thesis, Tilburg University, December 2009.

[10] S. Brin and L. Page. The anatomy of a large-scale hypertextual Web search engine. *Comput. Netw. ISDN Syst.*, 30:107–117, 1998.

[11] M. Bundschus, V. Tresp, and H. P. Kriegel. Topic models for semantically annotated document collections. In *NIPS workshop: Applications for Topic Models: Text and Beyond*, 2009.

[12] M. Bundschus, S. Yu, V. Tresp, A. Rettinger, M. Dejori, and H. Kriegel. Hierarchical bayesian models for collaborative

tagging systems. In *2009 Ninth IEEE International Conference on Data Mining*, pages 728–733. IEEE, 2009.

[13] A. Capocci and G. Caldarelli. Folksonomies and clustering in the collaborative system CiteULike. *J. Phys. A*, 41:224016, 2008.

[14] C. Cattuto, A. Barrat, A. Baldassarri, G. Schehr, and V. Loreto. Collective dynamics of social annotation. *Proc. Natl. Acad. Sci. U.S.A.*, 106:10511, 2009.

[15] C. Cattuto, V. Loreto, and L. Pietronero. Semiotic dynamics and collaborative tagging. *Proc. Natl. Acad. Sci. U.S.A.*, 104:1461, 2007.

[16] C. Cattuto, C. Schmitz, A. Baldassarri, V. D. P. Servedio, V. Loreto, A. Hotho, M. Grahl, and G. Stumme. Network properties of folksonomies. *AI Commun.*, 4:245, 2007.

[17] T. J. Chin, K. Schindler, and D. Suter. Incremental kernel SVD for face recognition with image sets. In *Proc. Inll Conf. Automatic Face Gesture Recognition (F-GR)*, pages 461–466, 2006.

[18] G. Cimini, M. Medo, T. Zhou, D. Wei, and Y. Zhang. Heterogeneity, quality, and reputation in an adaptive recommendation model. *Eur. Phys. J. B*, DOI: 10.1140/epjb/e2010-10716-5, 2011.

[19] A. Clauset, C. Moore, and M. E. J. Newman. Finding community structure in very large networks. *Nature*, 453:98, 2008.

[20] L. d. F. Costa, F. A. Rodrigues, G. Traviesor, and P. R. U. Boas. Characterization of complex networks: A survey of measurements. *Adv. Phys.*, 56:167–242, 2007.

[21] L. De Lathauwer, B. De Moor, and J. Vandewalle. A multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl.*, 21:1253–1278, 2000.

[22] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman. Indexing by latent semantic analysis. *J. Am. Soc. Info. Sci.*, 41:391–407, 1990.

[23] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. B*, 39:1–38, 1977.

[24] Y. Ding, E. Jacob, Z. Zhang, S. Foo, E. Yan, N. L. George, and L. Guo. Perspectives on Social Tagging. *J. Am. Soc. Inf. Sci. Technol*, 60:2388–2401, 2009.

[25] S. N. Dorogovtsev and J. F. F. Mendes. Evolution of networks. *Adv. Phys.*, 51:1079, 2002.

[26] M. Dubinko, R. Kumar, J. Magnani, J. Novak, P. Raghavan, and A. Tomkins. Visualizing tags over time. 1:7, 2007.

[27] D. Fensel, J. Hendler, and H. Lieberman. *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential.* Cambridge, US, 2003.

[28] Z. Gantner, S. Rendle, and L. Schmidt-Thieme. Factorization models for context-/time-aware movie recommendations. In *Proce. Workshop on Context-Aware Movie Recommendation*, pages 14–19, 2010.

[29] A. E. Gelfand and A. F. M. Smith. Sampling-based approaches to calculating marginal densities. *J. Am. Stat. Assoc.*, 85:398–409, 1990.

[30] J. Gemmell, A. Shepitsen, B. Mobasher, and R. Burke. Collaborative filtering

in social tagging systems based on joint item-tag recommendations. *Intelligent Techniques for Web Personalization & Recommender Systems*, 2008.

[31] G. Ghoshal, V. Zlatić, G. Caldarelli, and M. E. J. Newman. Random hypergraphs and their applications. *Phys. Rev. E*, 79:066118, 2009.

[32] M. Girvan and M. E. J. Newman. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. U.S.A.*, 99:7821, 2002.

[33] A. Gunawardana and G. Shani. A survey of accuracy evaluation metrics of recommendation tasks. *J. Mach. Learn. Res.*, 10:2935–2962, 2009.

[34] J. A. Hanley and B. J. McNeil. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143:29, 1982.

[35] M. Harvey, M. Baillie, I. Ruthven, and M. Carman. Tripartite hidden topic models for personalised tag suggestion. *Adv. Infor. Retr.*, pages 432–443, 2010.

[36] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst.*, 22:5, 2004.

[37] P. Heymann, D. Ramage, and H. Garcia-Molina. Social tag prediction. In *Proc. 31st Annual Intl. ACM SIGIR Conf. Research Development Information Retrieval*, pages 531–538. ACM, 2008.

[38] T. Hofmann. Probabilistic latent semantic indexing. In *Proc. 22nd Ann. Intl. ACM SIGIR Conf. Research and Development Information Retrieval*, pages 50–57. ACM, 1999.

[39] A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. Information retrieval in folksonomies: Search and ranking. *Lect. Notes Comput. Sci.*, 4011:411–426, 2006.

[40] A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. Trend detection in folksonomies. *Semantic Multimedia*, pages 56–70, 2006.

[41] J. Huang, X. Q. Cheng, J. Guo, H. W. Shen, and K. Yang. Social Recommendation with Interpersonal Influence. In *Proc. 2010 Conf. ECAI 2010: 19th Euro. Conf. on Artif. Int.*, pages 601–606, 2010.

[42] Z. Huang, H. Chen, and D. Zeng. Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. *ACM Trans. Inf. Syst.*, 22:116, 2004.

[43] R. Jäschke, L. Marinho, A. Hotho, L. Schmidt-Thieme, and G. Stumme. Tag recommendations in folksonomies. *Knowledge Discovery in Databases: PKDD'07*, pages 506–514, 2007.

[44] P. B. Kantor, F. Ricci, L. Rokach, and B. Shapira. *Recommender Systems Handbook*. Springer, 2010.

[45] R. Karp, C. Schindelhauer, S. Shenker, and B. Vocking. Randomized Rumor Spreading. In *Proc. 41st IEEE FOCS*, pages 565–574, 2000.

[46] P. Kazienko and M. Adamski. AdROSA–Adaptive personalization of web advertising. *Info. Sci.*, 177:2269, 2007.

[47] H. Kim, S. Scerri, J. Breslin, S. Decker, and H. Kim. The state of the art in tag ontologies: a semantic model for tagging and folksonomies. In *Proc. 2008 Intl. Conf. Dublin Core Metadata Appl.*, pages 128–137, 2008.

[48] J. M. Kleinberg. Authoritative sources in a hyperlinked environment. *J. ACM*, 46:604–632, 1999.

[49] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Rev.*, 51(3):455–500, 2009.

[50] A. Kontostathis, L. Galitsky, W. M. Pottenger, S. Roy, and D. J. Phelps. A survey of emerging trend detection in textual data mining. *Survey of Text Mining: Clustering, Classification, and Retrieval*, 2003.

[51] G. Koutrika, F. A. Effendi, Z. Gyöngyi, P. Heymann, and G.-M. Hector. Combating Spam in Tagging Systems: An Evaluation. *ACM Trans. Web*, 2:22, 2008.

[52] R. Krestel and P. Fankhauser. Tag Recommendation using Probabilistic Topic Models. *ECML PKDD Discovery Challenge 2009 (DC09)*, page 131.

[53] R. Krestel, P. Fankhauser, and W. Nejdl. Latent dirichlet allocation for tag recommendation. In *Proc. 3rd ACM Conf. Recomm. Syst.*, pages 61–68. ACM, 2009.

[54] R. Lambiotte and M. Ausllos. Collaborative tagging as a tripartite network. *Lect. Notes Comput. Sci.*, 3993:1114, 2006.

[55] S. Lee and H. Yong. Component based approach to handle synonym and polysemy in folksonomy. In *Computer and Information Technology*, pages 200–205, 2007.

[56] J. Leskovec, K. J. Lang, and M. Mahoney. Empirical comparison of algorithms for network community detection. In *Proc. 19th Intl. Conf. WWW*, pages 631–640. ACM, 2010.

[57] D. Li, B. He, Y. Ding, J. Tang, C. Sugimoto, Z. Qin, E. Yan, J. Li, and T. Dong. Community-based Topic Modeling for Social Tagging. In *CIKM'10: Proc. 9th ACM Intl. Conf. Infor. Knowl. Manag.*, pages 54–63. ACM, 2010.

[58] H. Liang, Y. Xu, Y. Li, R. Nayak, and X. Tao. Connecting Users and Items with Weighted Tags for Personalized Item Recommendations. In *HT '10: Proc. 21st ACM Conf. Hypertext and Hypermedia*, pages 51–60. ACM, 2010.

[59] G. Linden, B. Smith, and J. York. Amazon. com recommendations: Item-to-item collaborative filtering. *IEEE Internet Comput.*, 7:76, 2003.

[60] M. Lipczak. Tag recommendation for folksonomies oriented towards individual users. In *ECML PKDD Discovery Challenge (RSDC08)*, pages 84–95, 2008.

[61] Z. Liu, C. Shi, and M. Sun. FolkDiffusion: A Graph-Based Tag Suggestion Method for Folksonomies. *Infor. Retr. Technol.*, pages 231–240, 2010.

[62] L. Lü, Z.-K. Zhang, and T. Zhou. Zipf's Law Leads to Heaps' Law: Analyzing Their Relation in Finite-Size Systems. *PLoS ONE*, 5:e14139, 2010.

[63] L. Lü and T. Zhou. Link prediction in complex networks: A survey. *Physica A*, 390:1150–1170, 2011.

[64] A. Mathes. Folksonomies-cooperative classification and communication through shared metadata. *Computer Mediated Communication*, LIS590CMC, 2004.

[65] D. Medhi and D. Tipper. Multi-Layered Network Survivability–Models, Analysis,

Architecture, Framework and Implementation: An Overview. In *Proc. 2000 DARPA Information Survivability Conference & Exposition*, pages 173 – 186, 2000.

[66] M. Medo, Y.-C. Zhang, and T. Zhou. Adaptive model for recommendation of news. *EPL*, 88:38005, 2009.

[67] P. Mika. Ontologies are us: A unified model of social networks and semantics. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5:5–15, 2007.

[68] M. E. J. Newman. The Sturcture and Function of Complex Networks. *SIAM Rev.*, 45:167, 2003.

[69] O. Nov. What motivates wikipedians? *Commun. ACM*, 50:60–64, 2007.

[70] M. J. Pazzani and D. Billsus. Content-based recommendation systems. *Lect. Notes Comput. Sci.*, 4321:325, 2007.

[71] J. Peng, D. D. Zeng, H. Zhao, and F. Wang. Collaborative filtering in social tagging systems based on joint item-tag recommendations. In *Proc. 9th ACM Intl. Conf. Infor. Knowl. Manag.*, pages 809–818. ACM, 2010.

[72] B. Pittel. On spreading a rumor. *SIAM J. Appl. Math.*, 47:213–223, 1987.

[73] D. Ramage, P. Heymann, C. D. Manning, and H. Garcia-Molina. Clustering the tagged web. In *Proc. 2nd ACM Intl. Conf. Web Search and Data Mining*, pages 54–63. ACM, 2009.

[74] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme. BPR: Bayesian personalized ranking from implicit feedback. In *Proce. 25th Conf. Uncertainty Arti. Intel.*, pages 452–461, 2009.

[75] S. Rendle, L. B. Marinho, A. Nanopoulos, and L. S. Thieme. Learning optimal ranking with tensor factorization for tag recommendation. In *Proc. 15th ACM SIGKDD Intl Conf. Knowl. Discovery Data Mining*, pages 727–736. ACM, 2009.

[76] S. Rendle and L. Schmidt-Thieme. Pairwise interaction tensor factorization for personalized tag recommendation. In *Proc. 3rd ACM Intl. Conf. Web Search Data Mining*, pages 81–90. ACM, 2010.

[77] P. Resnick, K. Kuwabara, R. Zeckhauser, and E. Friedman. Reputation systems. *Commun. ACM*, 43:45, 2000.

[78] P. Resnick and H. R. Varian. Recommender systems. *Commun. ACM*, 40:56, 1997.

[79] G. Salton and M. J. McGill. *Introduction to Model Information Retrieval*. MuGraw-Hill, New York, 1983.

[80] J. B. Schafer, J. Konstan, and J. Riedl. E-commerce recommendation applications. *IEEE Internet Comput.*, 5:115, 2001.

[81] J. Scott. Social Network Analysis. *ACM Trans. Web*, 2:109, 1988.

[82] S. Sen, S. K. Lam, A. M. Rashid, D. Cosley, D. Frankowski, J. Osterhouse, F. M. Harper, and J. Riedl. Tagging, communities, vocabulary, evolution. In *Proc. 20th Anniversary Conf. Computer Supported Cooperative Work*, page 190. ACM, 2006.

[83] M.-S. Shang, L. Lü, Y.-C. Zhang, and T. Zhou. Empirical analysis of web-based user-object bipartite networks. *EPL*, 90:48006, 2010.

[84] M.-S. Shang and Z.-K. Zhang. Diffusion-Based Recommendation in collaborative Tagging Systems. *Chin. Phys. Lett.*, 26:11, 2009.

[85] M.-S. Shang, Z.-K. Zhang, T. Zhou, and Y.-C. Zhang. Collaborative filtering with diffusion-based similarity on tripartite graphs. *Physica A*, 389:1259, 2010.

[86] J. Shawe-Taylor and N. Cristianini. *Kernel methods for pattern analysis*. Cambridge Univ Press, 2004.

[87] A. Shepitsen, J. Gemmell, B. Mobasher, and R. Burke. Personalized recommendation in social tagging systems using hierarchical clustering. In *Proc. the 2008 ACM Conf. Recomm. Syst.*, pages 259–266, 2008.

[88] X. Si and M. Sun. Tag-LDA for scalable real-time tag recommendation. *J. Comput. Infor. Syst.*, 6:23–31, 2009.

[89] L. Spiteri. Structure and form of folksonomy tags: The road to the public library catalogue. *Webology*, 4, 2007.

[90] P. Symeonidis. User recommendations based on tensor dimensionality reduction. *Artificial Intelligence Applications and Innovations III*, pages 331–340, 2009.

[91] P. Symeonidis, A. Nanopoulos, and Y. Manolopoulos. Tag recommendations based on tensor dimensionality reduction. In *Proc. 2008 ACM Conf. Recomme. Syst.*, pages 43–50. New York, USA, ACM, 2008.

[92] M. Szomszor, C. Cattuto, H. Alani, K. O-Hara, A. Baldassarri, V. Loreto, and V. D. P. Servedio. Folksonomies, the semantic web, and movie recommendation. In *Proc. the 4th Euro. Semantic Web Conf.*, pages 71–84, 2007.

[93] K. Tso and L. Schmidt-Thieme. Attribute-aware collaborative filtering. In *From Data and Information Analysis to Knowledge Engineering*, pages 614–621, 2006.

[94] K. H. L. Tso-Sutter, L. B. Marinho, and L. Schmidt-Thieme. Tag-aware recommender systems by fusion of collaborative filtering algorithms. In *Proc. the 2008 ACM Symposium on Applied Computing*, pages 1995–1999. ACM, 2008.

[95] A. S. Umbrath, W. Wetzker R.and Umbrath, and L. Hennig. A hybrid PLSA approach for warmer cold start in folksonomy recommendation. *Recommender Systems & the Social Web*, pages 87–90, 2009.

[96] M. Wall, A. Rechtsteiner, and L. Rocha. Singular value decomposition and principal component analysis. *A practical approach to microarray data analysis*, pages 91–109, 2003.

[97] J. W. Wang, L. L. Rong, Q. H. Deng, and J. Y. Zhang. Evolving hypernetwork model. *Eur. Phys. J. B*, 77:493–498, 2010.

[98] R. Wetzker, W. Umbrath, and A. Said. A hybrid approach to item recommendation in folksonomies. In *Proc. of the WSDM'09 Workshop Exploiting Semantic Annotation Information Retrieval*, pages 25–29. ACM, 2009.

[99] H. Wu, M. Zubair, and K. Maly. Harvesting social knowledge from folksonomies. In *Proc. 7th Conf. Hypertext hypermedia*, pages 111–114, 2006.

[100] P. Wu and Z.-K. Zhang. Enhancing personalized recommendation in weighted social tagging networks. *Physical Procdia*, 3:1877–1885, 2010.

[101] L. Xiang, Q. Yuan, S. Zhao, L. Chen, X. Zhang, Q. Yang, and J. Sun. Temporal recommendation on graphs via long-and short-term preference fusion. In *Proc. 16th ACM SIGKDD Intl. Conf. Knowl. Disc. Data Min.*, pages 723–732, 2010.

[102] S. Xu, S. Bao, B. Fei, Z. Su, and Y. Yu. Exploring folksonomy for personalized search. In *Proc. the 31st Annual Intl. ACM SIGIR Conf. Res. Dev. Info. Retr.*, pages 155–162, 2008.

[103] Y. Xu, L. Zhang, and W. Liu. Cubic analysis of social bookmarking for personalized recommendation. *Frontiers of WWW Research and Development-APWeb 2006*, pages 733–738, 2006.

[104] G.-Q. Zhang, G.-Q. Zhang, Q.-F. Yang, S.-Q. Cheng, and T. Zhou. Evolution of the Internet and its cores. *New J. Phys.*, 10:123027, 2008.

[105] Y.-C. Zhang, M. Blattner, and Y.-K. Yu. Heat conduction process on community networks as a recommendation model. *Phys. Rev. Lett.*, 99:154301, 2007.

[106] Z.-K. Zhang and C. Liu. A hypergraph model of social tagging networks. *J. Stat. Mech.*, page P10005, 2010.

[107] Z.-K. Zhang and C. Liu. Identifying the Role of Social Tags and its Application in Recommender Systems. *Int. J.Complex Systems in Science*, 1:10, 2011.

[108] Z.-K. Zhang, C. Liu, Y.-C. Zhang, and T. Zhou. Solving the Cold-Start Problem in Recommender Systems with Social Tags. *EPL*, 92:28002, 2010.

[109] Z.-K. Zhang, L. Lü, J.-G. Liu, and T. Zhou. Empirical analysis on a keyword-based semantic system. *Eur. Phys. J. B*, 66:557, 2008.

[110] Z.-K. Zhang, T. Zhou, and Y.-C. Zhang. Personalized recommendation via integrated diffusion on user-item-tag tripartite graphs. *Physica A*, 389:179, 2010.

[111] T. Zhou, Z. Kuscsik, J.-G. Liu, M. Medo, J. R. Wakeling, and Y.-C. Zhang. Solving the apparent diversity-accuracy dilemma of recommender systems. *Proc. Natl. Acad. Sci. U.S.A.*, 107:4511, 2010.

[112] T. Zhou, J. Ren, M. Medo, and Y.-C. Zhang. Bipartite network projection and personal recommendation. *Phys. Rev. E*, 76:046115, 2007.

[113] T. Zhou, R.-Q. Su, R.-R. Liu, L.-L. Jiang, B.-H. Wang, and Y.-C. Zhang. Effect of initial configuration on network-based recommendation. *New J. Phys.*, 11:123008, 2009.

[114] V. Zlatić, G. Ghoshal, and G. Caldarelli. Hypergraph topological quantities for tagged social networks. *Phys. Rev. E*, 80:036118, 2009.