A UNIFIED APPROACH ON RECOMMENDATIONS IN SOCIAL TAGGING SYSTEMS BY LIMITING UNWANTED REQUEST

Manopriya Marimuthu[1], Ms.R.P.Narmadha[2]

[1](M.E Computer Science and Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore) India
[2](Assistant Professor, Department of Computer Science and Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore) India

ABSTRACT

Social Tagging System is the process where user (u) makes their interest by tagging (t) on a particular item (i).These STS are in associated with web 2.0 and has been sourceful information for the users. It provides different types of recommendation in contrast to the current recommendation algorithm, Collaborative Filtering (CF) which apply to two – dimensional data. These data are modeled by a 3-order tensor, on which multiway latent semantic analysis and dimensionality reduction is performed using both the Higher Order Singular Value Decomposition (HOSVD) method and the Kernel-SVD smoothing technique. We provide the 4-order tensor approach, which we named as Tensor Reduction. In particular, the tensor equivalently represents a quadruplet. And also can improve the social tagging efficiency by which unwanted request has been controlled. The results show significant improvements in terms of effectiveness.

KEYWORDS:- *SOCIAL TAGS, RECOMMENDER SYSTEMS, TENSORS, HOSVD.*

1. INTRODUCTION

Social tagging is the process by which many users add metadata in the form of keywords, to annotate and categorize songs, pictures, products, etc. Social tagging is associated to the "web2.0" technologies and has already become an important source of information for recommender systems. For example, music recommender systems such as Last.fm and MyStrands allow users to tag artist, songs, or albums. In e-commerce sites such as Amazon, users tag products to easily discover common interests with other users. Moreover, social media sites, such as Flickr and YouTube use tags for annotating their content. All these systems can further exploit these social tags to improve the search mechanisms and to personalize the user recommendations. Social tags carry useful information not only about the items they label, but also about the users who tagged. Thus, social tags are a powerful mechanism that reveals three-dimensional correlations between users, tags, and items. Several social tagging systems (STSs), e.g., Last.fm, Amazon, YouTube, etc., recommend items to users, based on tags they have in common with other similar users. Traditional recommender systems use techniques such as Collaborative Filtering (CF) [9], which apply to two-dimensional data, i.e., users and items. Thus, such systems do not capture the multimodal use of tags. To alleviate this problem, Tso-Sutter et al. [13] propose a generic method that allows tags to be incorporated to standard CF algorithms, by reducing the three-dimensional correlations to three 2D correlations and then applying a fusion method to reassociate these correlations.

Another type of recommendation in STSs, e.g., Facebook, Amazon, etc., is to recommend tags to users, based on what tags other users have provided for the same items. Tag recommendations can expose different facets of an information item and relieve users from the obnoxious task to come up with a good set of tags. Thus, tag recommendation can reduce the problem of data sparsity in STSs, which results by the unwillingness of users to provide an adequate number of tags. Recently, several algorithms have been proposed for tag recommendation [4], [5], which project the three-dimensional correlations to three 2D correlations. Then, the two-dimensional correlations are used to build conceptual structures similar to hyperlink structures that are used by Web search engines.

A third type of recommendation that can be provided by STSs is to recommend interesting users to a target user, opting in connecting people with common interests and encouraging people to contribute and share more content. With the term interesting users, we mean those users who have similar profile with the target user. If a set of tags is frequently used by many users, then these users spontaneously form a group of users with common interests, even though they may not have any physical or online connections. The tags represent the commonly interested Web contents to this user group. For example, Amazon recommends to a user who used a specific tag, other new users considering them as interesting ones. Amazon ranks them based on how frequent they used the specific tag.

2. An Overview of Related Work

In this section, we briefly present some of the research literature related to Social Tagging. We also present related work in tag, item, and users recommendation algorithms. Finally, we present works that applied HOSVD in various research domains.Social Tagging is the process by which many users add metadata in the form of keywords to share content. So far, the literature has studied the strengths and the weaknesses of STSs. In particular, Golder and Huberman [12] analyzed the structure of collaborative tagging systems as well as their dynamical aspects. Moreover, Halpin et al. [3] produced a generative model of collaborative tagging in order to understand the dynamics behind it. They claimed that there are three main entities in any tagging system: users, items, and tags.

In the area of item recommendations, many recommender systems already use CF to recommend items based on preferences of similar users, by exploiting a two-way relation of users and items [9]. In 2001, Item-based algorithm was proposed, which is based on the items' similarities for a neighborhood generation. However, because of the ternary relational nature of Social Tagging, two-way CF cannot be applied directly, unless the ternary relation is reduced to a lower dimensional space. Jaschke et al. [11], in order to apply CF in Social Tagging, considered for the ternary relation of users, items, and tags two alternative two-dimensional projections. These projections preserve the user information, and lead to log-based like recommender systems based on occurrence or nonoccurrence of items, or tags, respectively, with the users. Another recently proposed state-of-the-art item recommendation algorithm is tag-aware Fusion [13]. They propose a generic method that allows tags to be incorporated to standard CF algorithms, by reducing the three-dimensional correlations to three 2D correlations and then applying a fusion method to reassociate these correlations.

In the area of tag recommendation, there are algorithms which are based on conceptual structures similar to the hyperlink structures used in Search Engines. For example, Collaborative Tag Suggestions algorithm [5], also known as Penalty-Reward algorithm (PR), uses an authority score for each user. The authority score measures how well each user has tagged in the past. This authority score can be computed via an iterative algorithm similar to HITS. Moreover, the PR algorithm "rewards" the high correlation among tags, whereas it "penalizes" the overlap of concepts among the recommended tags to allow high coverage of multiple factes for anitem. Another state-of-the-art tag recommendation algorithms FolkRank[4]. FolkRank exploits the conceptual structures created by people inside the STSs. Their method is inspired by the seminal PageRank [10] algorithm, which reflects the idea that a web page is important, if there are many pages linking to it, and if those pages are important themselves.

FolkRank employs the same underlying principle for Web Search and Ranking in Social Tagging. The key idea of FolkRank algorithm is that an item which is tagged with important tags by important users becomes important itself. The same holds for tags and users: thus, we have a tripartite graph of vertices which mutually reinforcing each other by spreading their weights. FolkRank is like the Personalized PageRank, which is a modification of global PageRank, and was first proposed for Personalized Web searchind[10].Finally, Xu et.al [6] proposed a method that recommends tags by using HOSVD. Hoever, their method does not cover all three type of recommendation in STSs and misses the comparison with the state-of-art algorithms. In contrast,the approach proposes a unified framework for all recommendation types in STSs. We also combine HOSVD with Kernal-SVD to handle data sparsity,attaining significant improvements in accuracy of recommendation in comparison with simple HOSVD,as will be shown experimentally. In area of discovering shared interest in social networks there are two kind of existing approaches [1]. One is user-centric,which focuses on detecting social interest based on the online connection among users; the other is object-centric, which detects common interest based on common objects fetched by users in a social community.

In the user-centric approach, it can be analyzed for user's online connection and interaction to discover users with particular interest of the given user. Different from this of approach, we aim to find the people who share the same interest no matter whether ther are connected by a social graph or not. In the object-centric approach it explores the common interest among the users based on the common items they fetched in peer-to-peer networks. However, they cannot differentiate the various social interests on the same items due to the fact that users may have different interest for an information item and an item may have multiple facets. In contrast, the approach focuses on directly detecting social interest and recommending users by taking advantage of social tagging, by utilizing users tags. Differently from existing approaches, the method develops a unified framework to concurrently model all three dimensions. Usage data are modeled by a 3-order tensor, on which latent semantic analysis is performed using the HOSVD. Moreover, to address the sparseness problem, we propose the combination of kernel-SVD [7], [8] with HOSVD, which substantially improve the accuracy of item and tag recommendation. HOSVD is a generalization of singular value decomposition (SVD) anh has been successfully applied in several areas. In particular, Wang and Ahuja [2] present a novel multilinear algerbra-based approach to reduced dimensionality representation of multidimensional data,such as image ensembles, video sequences, and volume data. Howeever, they transform the initial tensor (through Clique Expansion algorithm) into lower dimensional spaces, so that clustering algorithm (such as k-means) can be applied.

3. The Preliminaries

In this section, we summarize the HOSVD procedure.

SVD: The SVD of a matrix $F_{I1\times I2}$ can be written as a product of three matrices, as shown in equation (1):

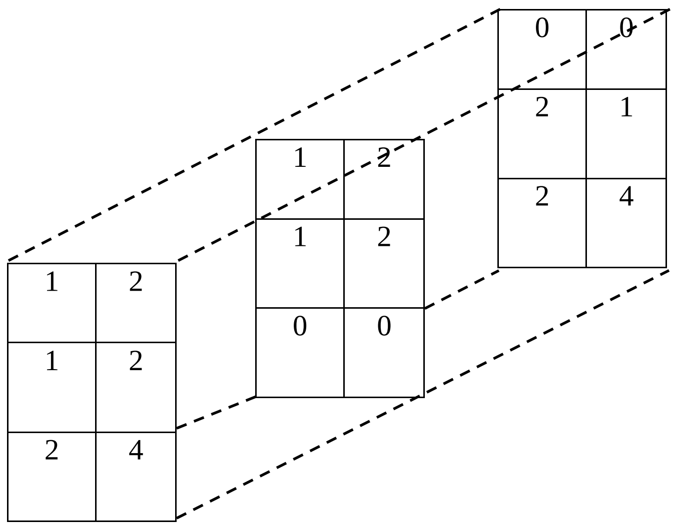
$$FI1\times I2 = U_{I1\times I1} \,.\, S_{I1\times I2} \,.\, V^{T}{}_{I2\times I2} \,; \qquad (1)$$

Where U is the matrix with the left singular vectors of F, V^{T} is the transpose of the matrix V with the right singular vectors of F, and S is the diagonal matrix of (ordered) singular values of F.

Tensors: A tensor is a multidimensional matrix. An N-order tensor A is denoted as $A \in R^{I1\ldots IN}$, with elements $a_{i1,\ldots,iN}$. In this paper, for the purposes of the approach, we only use 3-order tensors.

HOSVD: The high-order singular value decomposition generalizes the SVD computation to multidimensional matrices. To apply HOSVD on a 3-order tensor A, three matrix unfolding operations are defined as follows:

$$A1 \in R^{I1\times I2I3};\ A2 \in R^{I2\times I1I3};\ A3 \in R^{I1I2\times I3},$$



$$\begin{pmatrix} 1 & 2 & 1 & 2 & 0 & 0 \\ 1 & 2 & 1 & 2 & 2 & 4 \\ 2 & 4 & 0 & 0 & 2 & 4 \end{pmatrix}$$

Fig.1. 3 2D matrices

Fig.1: An example tensor A and its 1-mode matrix unfolding A1, where A1, A2, and A3 are called the 1-mode, 2-mode, and 3-mode matrix unfolding of A, respectively. Each A_n ,$1 \le n \le 3$, is called the n-mode matrix unfolding of A and is computed by arranging the corresponding fibers of A as columns of An. The left part of Fig. 1 depicts an example tensor, whereas the right part its 1-mode matrix unfolding $A1 \in R^{I1\times I2I3}$, where the columns (1-mode fibers) of A are being arranged as columns of A1.

4. THE PROPOSED APPROACH

We first provide the outline of the approach, which we name Tensor Reduction, through a motivating example. In this section, we elaborate on how HOSVD is applied on tensors and on how the recommendation of items is performed according to the detected latent associations. Note that a similar approach is followed for the tag and user recommendations.

Three models of Tagging:

After registering a free account, Tagged users can customize their profile page, to which they can post a biography about themselves and their interests, post status updates to inform their friends of their whereabouts and actions, upload photos and albums, and send and receive messages from other users. Model: 1 which represents the user and tag representation in which Fig. 2, represents tagging an item contains a text field, which describes about a particular thing in an item. Item contains photos, videos, etc through which tagging is made.

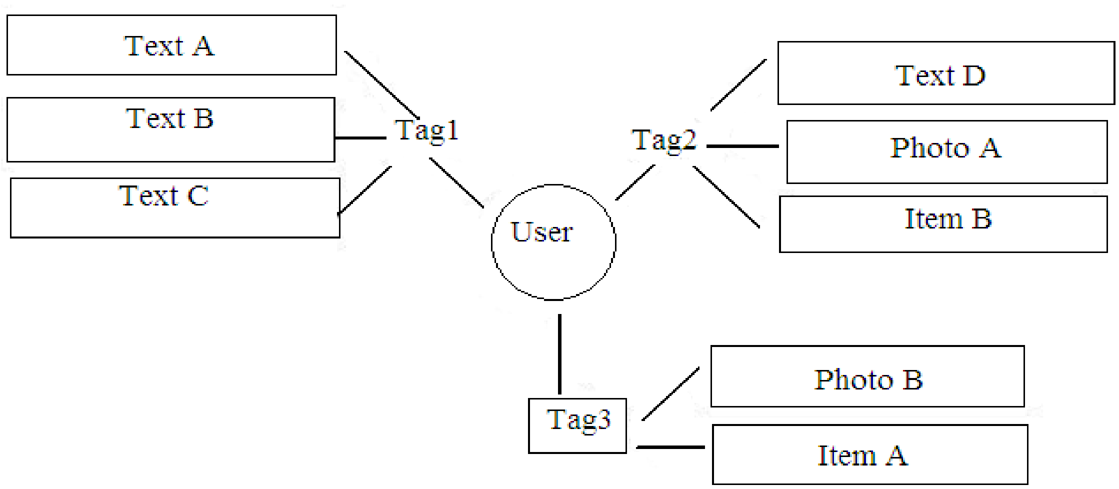


Fig. 2, Model: 1 User – Tag information

The second model fig.3, portrays users that are connected together through their use of tags. This is where real social networking comes in, as users are tagging to relate their concept of information to another user's concept of some piece of information. It may be used inconsistently, by tagging in order that other users see desired information despite the fact that the information is not really classified under their expected concept of that tag.

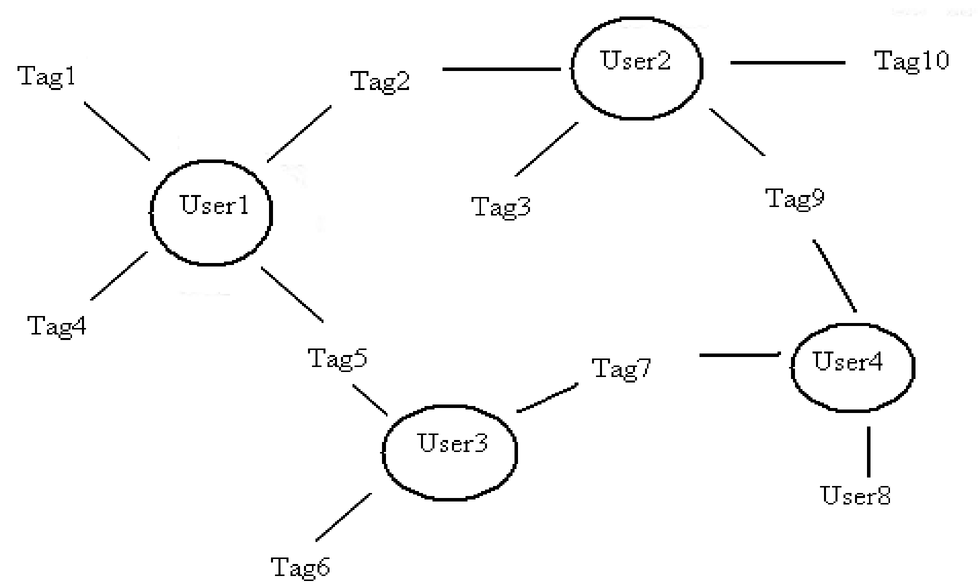


Fig. 3, Model: 2 User - Tag - User

Within the final model, in fig. 4, tags are used to link banks of data (or information) to other information. The tags are acting as metadata to allow search engines to know which information is related to other information. This type of tagging is used greatly within ontologies.

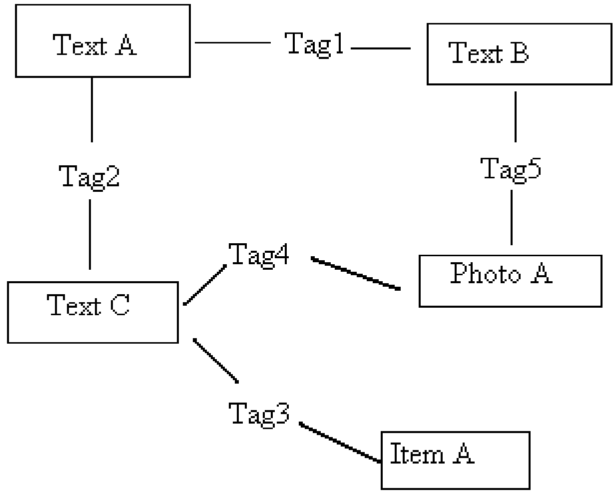


Fig. 4, Model:3 Information - Tag - Information

When using a social tagging systems, to be able to retrieve the information items easily, a user u tags an item i with a tag t. After some time of usage, the tagging system accumulates a collection of usage data, which can be represented by a set of triples {u; i; t}. The Tensor Reduction approach applies HOSVD on the 3-order tensor constructed from these usage data. In accordance with the HOSVD technique introduced, the Tensor Reduction algorithm uses as input the usage data of A and outputs the reconstructed tensor A^. A^ measures the associations among the users, items, and tags. Each element of A^ can be represented by a quadruplet {u; i; t; p}, where p measures the likeliness that user u will tag item i with tag t. Therefore, items can be recommended to u according to their weights associated with {u; t} pair. In this section, in order to illustrate how the approach works, we apply the Tensor Reduction algorithm to a running example. As illustrated in Fig. 5, three users tagged three different items (Web links).

In Fig. 5, the part of an arrow line (sequence of arrows with the same annotation) between a user and an item represents that the user tagged the corresponding item, and the part between an item and a tag indicates that the user tagged this item with the corresponding tag. Thus, the annotated numbers on the arrow lines gives the correspondence between the three types of objects. For example, user U1 tagged item I1 with tag "BMW," denoted as T1. The remaining tags are "Jaguar," denoted as T2, and "CAT," denoted as T3. From Fig. 5, we can see that users U1 and U2 have common interests on cars, while user U3 is interested in cats. A 3-order tensor A 2 $R^{3\times33}$, can be constructed from the usage data.

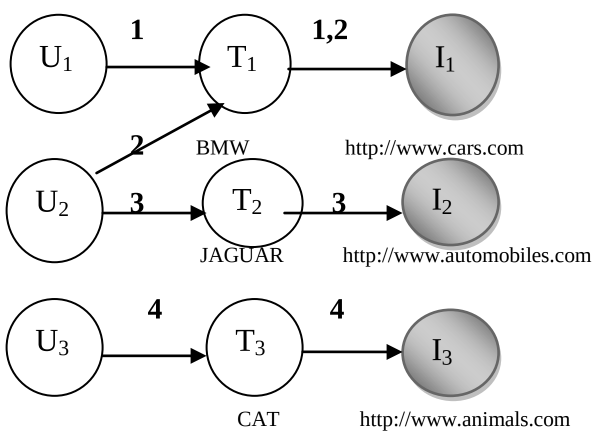


Fig. 5, Usage data of the running example.

We use the co-occurrence frequency (denoted as weight) of each triplet user, item, and tag as the elements of tensor A, which are given in Table 1. Note that all associated weights are initialized to 1. After performing the Tensor Reduction analysis (details of how to do this are given in the following section), we can get reconstructed tensor

TABLE 1
The Elements of the Example Tensor

Arrow Line	User	Item	Tag	Weight
1	U_1	I_1	T_1	1
2	U_2	I_1	T_1	1
3	U_2	I_2	T_2	1
4	U_3	I_3	T_3	1

TABLE 2
The Elements of the Reconstructed Tensor

Arrow Line	User	Item	Tag	Weight
1	U_1	I_1	T_1	0.72
2	U_2	I_1	T_1	1.17
3	U_2	I_2	T_2	0.72
4	U_3	I_3	T_3	1
5	$\boldsymbol{U_1}$	$\boldsymbol{I_2}$	$\boldsymbol{T_2}$	**0.44**

of A^, which is presented in Table 2, whereas Fig. 6 depicts the contents of A^ graphically (the weights are omitted). As shown in Table 2 and Fig. 6, the output of the Tensor Reduction algorithm for the running example is interesting, because a new association among these objects is revealed. The new association is between U1; I2, and T2. This association is represented with the last (bold faced) row in Table 2 and with the dashed arrow line in Fig. 6). If we have to recommend to U1 an item for tag T2, then there is no direct indication for this task in the original tensor A. However, we see that in Table 2 the element recommend the item I2 to user U1, who used tag T2.The resulting recommendation is reasonable, because U1 is interested in

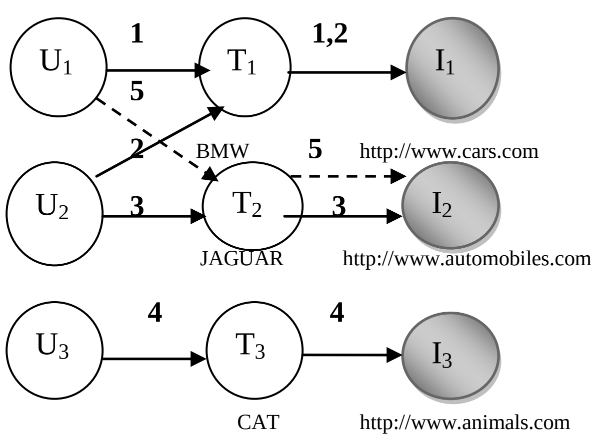


Fig. 6, Illustration of the Tensor Reduction Algorithm output for the running example.

cars rather than cats. That is, the Tensor Reduction approach is able to capture the latent associations among the multitype data objects: user, item, and tags. The associations can then be used to improve the item recommendation procedure, as will be verified by the experimental results. Moreover, for purposes of tag recommendations, we can view the tensor from a different perspective. In particular, the tensor equivalently represents a quadruplet {u, i, t, p} where p is the likeliness that user u will tag item i with tag t. Therefore, tags can be recommended to u according to their weights associated with {u, i} pair. In the running example, if user U1 is about to tag I2, he will be recommended tag T2. Finally, for recommending users, the tensor can be viewed as a quadruplet {t, I, u, p}, where p is the likeliness that tag t will be used to label item i by the user u. Therefore, new users can be recommended for a tag t, according to their total weight, which results by aggregating all items, which are labeled with the same tag by the target user. In the running example, if user U1 tagged item I2 with tag T2, he would receive user U2 as user recommendation.

5. CONCLUSION

Social tagging systems provide recommendations to users based on what tags other users have used on items. In this paper, we developed a unified framework to model the three types of entities that exist in a social tagging system: users, items, and tags. We examined multiway analysis on data modeled as 3-order tensor, to reveal the latent semantic associations between user, items and tags. The multiway latent semantic analysis and dimensionality reduction is performed by combining the HOSVD method with the Kernel-SVD smoothing technique. The approach improves recommendations by capturing user's multimodal perception of item/tag/user. Moreover, we study a problem of how to provide user recommendation which can have significant applications in real systems but which have not been studied in depth so far in related research. We also performed experimental comparison of the proposed method against state-of-the-art recommendations algorithms, with two real data sets.

The results can show significant improvements in terms of effectiveness measured through recall/precision. As future work, we intend to examine different methods for extending SVD to high-order tensors such as the Parallel Factor Analysis. We also indent to apply different weighting methods for the initial construction of a tensor. A different weighting policy for the tensor's initial values could improve the overall performance of the approach and also limiting the requests.

REFERENCES

[1] X. Li, L. Guo, and Y. Zhao, "Tag-Based Social Interest Discovery," Proc. ACM World Wide Web (WWW) Conf., 2008.
[2] H. Wang and N. Ahuja, "A Tensor Approximation Approach to Dimensionality Reduction," Int'l J. Computer Vision, vol. 76, no. 3,pp. 217-229, 2008.
[3] H. Halpin, V. Robu, and H. Shepherd, "The Complex Dynamics of Collaborative Tagging," Proc. 16th Int'l Conf. World Wide Web (WWW '07), pp. 211-220, 2007.
[4] A. Hotho, R. Jaschke, C. Schmitz, and G. Stumme, "Information Retrieval in Folksonomies: Search and Ranking," The Semantic Web: Research and Applications, pp. 411-426. Springer, 2006.
[5] Z. Xu, Y. Fu, J. Mao, and D. Su, "Towards the Semantic Web: Collaborative Tag Suggestions," Proc. Collaborative Web Tagging Workshop at World Wide Web (WWW '06), 2006.
[6] Y. Xu, L. Zhang, and W. Liu, "Cubic Analysis of Social Bookmarking for Personalized Recommendation," Frontiers of WWW Research and Development—APWeb '06, pp. 733-738. Springer, 2006.
[7] T. Chin, K. Schindler, and D. Suter, "Incremental Kernel SVD for Face Recognition with Image Sets," Proc. Int'l Conf. Automatic Face and Gesture Recognition (FGR), pp. 461-466, 2006.
[8] N. Cristianini and J. Shawe-Taylor, Kernel Methods for Pattern Analysis. Cambridge Univ. Press, 2004.
[9] J. Breese, D. Heckerman, and C. Kadie, "Empirical Analysis of Predictive Algorithms for Collaborative Filtering," Proc. Conf. Uncertainty in Artificial Intelligence, pp. 43-52, 1998.
[10] L. Page, S. Brin, R. Motwani, and T. Winograd, "The Pagerank Citation Ranking: Bringing Order to the Web," technical report, 1998.
[11] R. Jaschke, L. Marinho, A. Hotho, L. Schmidt-Thieme, and G.Stumme, "Tag Recommendations in Folksonomies," Proc. Knowledge Discovery in Databases (PKDD '07), pp. 506-514.
[12] S. Golder and B. Huberman, "The Structure of Collaborative Tagging Systems," technical report, 2005.
[13] K Tso-Sutter, B. Marinho, and L.Schmidt- Thieme, "Tag-Aware Recommender Systems by Fusion of Collaborative Filtering Algorithms," Poc. ACM Symp. Applied Computing(SAC) conf, 2008.

Authors

Ms.Manopriya Marimuthu received B.E degree in CSE from Anna University, Chennai and Currently pursuing M.E degree in Computer Science and Engineering in Sri Shakthi Institute of Engineering and Technology, under Anna University of Technology, Coimbatore. Her research interest includes Computer Networks and Data Mining.

Ms.R.P.Narmadha received the B.E degree in Computer Science and Engineering in Kumaraguru College of Engeering and Technology under Bharathiar University, received the M.E degree in IT in Sathyabama University, Chennai and received the M.B.A degree under Bharathiar University. She is currently working as Assistant Professor in Department of CSE in Sri Shakthi Institute of Engineering and Technology, Coimbatore. Her teaching experience is of 11½ year. Her main research interest is Software Engineering, Data Mining and Image Processing.