World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:6, 2011

Folksonomy-based Recommender Systems with User's Recent Preferences

Cheng-Lung Huang, Han-Yu Chien, Michael Conyette

***Abstract*—**Social bookmarking is an environment in which the user gradually changes interests over time so that the tag data associated with the current temporal period is usually more important than tag data temporally far from the current period. This implies that in the social tagging system, the newly tagged items by the user are more relevant than older items. This study proposes a novel recommender system that considers the users' recent tag preferences. The proposed system includes the following stages: grouping similar users into clusters using an E-M clustering algorithm, finding similar resources based on the user's bookmarks, and recommending the top-N items to the target user. The study examines the system's information retrieval performance using a dataset from del.icio.us, which is a famous social bookmarking web site. Experimental results show that the proposed system is better and more effective than traditional approaches.

***Keywords*—**Recommender systems, Social bookmarking, Tag

I. Introduction

FOLKSONOMY is known as collaborative tagging or social tagging, which allows users to collaboratively create and manage tags to classify and categorize contents or users' collections. Collaborative tagging in Web 2.0 is becoming widely used as an important tool to classify dynamic content for searching and sharing [1]. Recently, researchers have shown that social tagging can be used to classify blogs [2], to enhance information retrieval [2][3] and to improve recommender systems [1].

The tags collected by the user represent part of this user's preference or interests in the social bookmarking website. That is, the recent tags represent the user's current preference or interests. For the recommender systems in the world wide web, however, user interest changes with time, and thus learning user's interest categories in a dynamic environment like the web is challenging [4][5][6]. In an environment in which the user gradually changes interests, the tag data close to the current temporal period is usually more important than that temporally far from the current period. This is called the time decay of user's interests in the recommender systems [7]. The phenomenon of time decay suggests that in the social tagging system, newly tagged items by the user are more relevant for this user currently. That is, based on the tagging information, users are usually interested in items they tagged more recently. For example, a certain user may have been interested in a PDA 6 months ago. He is interested in an iPad now, and so the tag of iPad is currently used. Because his/her current interest is an iPad it is more proper to recommend the iPad than a PDA.

The question thus becomes how best to handle the problem of time decay for tag-based recommender systems? This triggered our research to introduce a new TF-IDF model with the notion of recency to provide higher importance on the tags with more recent time periods. The new model will improve the recommendation quality based on the tag-based user profile.

The rest of this paper outlines the development of this model and is organized as follows. Section 2 introduces the related works. Section 3 describes the framework of the proposed methodology. Section 4 demonstrates the empirical experiment. Section 5 provides conclusions.

C.-L. Huang is with the Department of Information Management, National Kaohsiung First University of Science and Technology, Kaohsiung, Taiwan, ROC. (phone: +886-7-6011000 Ext 4127; fax: +886-7-6011042; e-mail: clhuang@nkfust.edu.tw).

H.-Y. Chien was with the Department of Information Management, National Kaohsiung First University of Science and Technology, Kaohsiung, Taiwan, ROC.

M.Conyette teaches for the Okanagan School of Business, 7000 College Way, Vernon, BC, V1B 2N5, Canada; e-mail:mconyette@okanagan.bc.ca.

II. Related Works

Social resource sharing systems are web-based systems that allow users to upload their resources, and to label them with arbitrary tags [8]. Users, resource items, and tags are three important roles in this kind of systems. Users label the resource item using social tags. These systems can be categorized according to what kinds of resources are supported, such as bookmarks, bibliographic references, photos, merchandise, or video. "Delicious" (del.icio.us) is a social bookmarking web service for storing, sharing, and discovering web bookmarks. Delicious uses a non-hierarchical classification system in which users can tag each of their bookmarks with freely chosen keyword terms.

People tag resources for future retrieval and sharing [9]. Tags can convey information about the content and creation of a resource [10]. Tags identify what the resource is about and the characteristics of a resource [11]. The tags accumulated by the user represent part of this user's preference or interest in the social bookmarking website. The current study models the user's preference by using the tag-based information.

Kim et al. [1] used a tag-based user profile in collaborative filtering recommender systems to alleviate limitations of the cold-start and sparsity problems. Unlike previous researchers like this, the current study constructs a two-stage recommender approach that hybridizes the collaborative filtering and content-based filtering.

III. Proposed Methodology

The proposed system includes two parts: collaborative filtering and content-based filtering stages.

A. Collaborative filtering stage

(1) Tag frequency–inverse user frequency (TF-IUF) for users

The user tagging profile and resource item tagging profile are collected and stored in a database. The user's tag information includes the tag name and its corresponding number of items collected by this user. The resource item's tag information includes the tag name and its corresponding frequency tagged by these users.

The tag-based user profile can be transformed into a vector of TF-IUF (tag frequency–inverse user frequency) which is modified from the TF-IDF (term frequency–inverse document frequency) for representing the document description. The weight of tag *j* in user *u*'s tag collection is defined as:

$$UserTagW_{u,j} = UserTagTF_{u,j} \times UserTagIUF_j \quad (1)$$

$UserTagTF_{u,j}$ is the local weight of tag frequency and it is defined as:

$$UserTagTF_{u,j} = \frac{freq_{u,j}}{\max_u freq_{u,j}} \quad (2)$$

where $freq_{u,j}$ is the number of occurrences of tag *j* in user *u*'s tag collection; $\max_u freq_{u,j}$ is the maximum number of occurrences in user *u*'s tag collection. This equation normalizes or scales the tag occurrences.

The global weight, $UserTagIUF_j$, which represents the relative importance among user *u*'s tag collection, is defined as:

$$UserTagIUF_j = \log(\#Users / \#UserCollectingTag_j) \quad (3)$$

where *#Users* is the total number of users in the training set; $\#UserCollectingTag_j$ is the total number of users who collect tag *j* (in the training set).

This model incorporates local and global information. The $UserTagTF$ accounts for local information and $UserTagIUF$ is the inverse tag importance which represents the global probability of a certain tag for a user.

(2) User clustering and similarity between users

The purpose of this step is to cluster users based on their TF-IDF tag profile. Clustering is an unsupervised data segmentation technique for grouping a set of data objects into classes of similar data objects. Some popular clustering methods can be adopted such as partitioning methods (e.g., k-means clustering), hierarchical methods, grid-based methods, model-based methods and density-based methods [12]. This study used the E-M clustering approach, which can easily perform clustering based on the cosine similarity matrix among users.

The cosine similarity between the user *u* and user *v* can be defined as the inner product of the two users' tag weights:

$$UserSim_{u,v} = \frac{\sum_{j=1}^{N} UserTagW_{u,j} \cdot UserTagW_{u,j}}{\sqrt{\sum_{j=1}^{N} (UserTagW_{u,j})^2 \cdot \sum_{i=1}^{N} (UserTagW_{u,j})^2}} \quad (4)$$

where *N* is the number of common tags that is collected by user *u* and *v*.

B. Content-based filtering for resource items

Users in the same cluster have similar preferences. The content-based filtering based on the resource item's tag information is applied in each cluster. The purpose of this step is to find the similar resource items which may interest the user and then recommend these similar resource items to the target user.

(1) Tag frequency for the resource item

The user defines a resource item using tags. The tag *i*'s normalized frequency for item *q* represented as $ItemTagTF_{q,i}$ is defined below.

$$ItemTagTF_{q,i} = \frac{freq_{q,i}}{\max_q freq_{q,i}} \quad (5)$$

where $freq_{q,i}$ is the number of occurrences of tag *i* that defines item *q*; $\max_q freq_{q,i}$ is the maximum number of occurrences of tags that define item *q*.

(2) Inverse item frequency for the resource item

The tag *i*'s relative importance among collected tags in a cluster represented as $ItemTagIIF$ (inverse item frequency) is defined as:

$$ItemTagIIF_i = \log(\#Items / \#ItemsDefinedByTag_i) \quad (6)$$

where *#Items* is the total number of items in a cluster; $\#ItemsDefinedByTag_i$ is the total number of items (in a cluster) defined by tag *i*.

(3) Tag weight for the resource item

The weight of tag *i* for resource item *q* is defined as:

$$ItemTagW_{q,i} = ItemTagTF_{q,i} \times ItemTagIIF_i \quad (7)$$

(4) Cosine similarity between resource items

The tag-based cosine similarity between resource item *q* and item *r* is calculated as the inner product of the item tag weights:

$$ItemSim_{q,r} = \frac{\sum_{i=1}^{M} ItemTagW_{q,i} \cdot ItemTagW_{r,i}}{\sqrt{\sum_{i=1}^{M} (ItemTagW_{q,i})^2 \cdot \sum_{i=1}^{M} (ItemTagW_{r,i})^2}} \quad (8)$$

where *M* is the number of common tags which label both resource item *q* and *r*.

C. Personalized Resource item recommendation

To recommend items to the target user, content-based filtering is applied. In the social tagging system, newly tagged items by the user are more important for this user. That is, based on the tagging information, users are usually interested in items they tagged more recently. The effect of recency gives the weight to tags or items according to their tagging or bookmarking time. For the recency preference of an item, we consider the collection time for the item and the tagging time for the tags that defined this item. The bookmark time of an item is recorded when a user tags the bookmark or item. The recency of a tag is defined as the collection time of the most recent item. This study defines the relative recency of a tag owned by a certain user as the following.

$$UserTagRecency = 1 - \frac{userLast - tagLast}{userLast - userFirst} \quad (9)$$

where $userFirst$ represents user's first bookmarking time, $userLast$ represents user's last bookmarking time, and $tagLast$ represents the last bookmarking time of the tag by the user.

The recommendation procedure is:

(1) For each item *q* for user *u*, calculate the similarity between item q and other user's item *r*, $ItemSim_{q,r}$ using Eq. (8).

(2) For each item *q* for user *u*, calculate the recency preference of item *q* as follows.

$$TagRecency_q = Max(UserTagRecency_{q,i}),\ i = 1,2,...,m \quad (10)$$

where *m* is the number of tags that define item *q*, and $UserTagRecency_{q,i}$ is recency of tag *i* for item *q* defined in Eq.(9).

(3) Recommend top-N items to user *u* according to the score of the other user's items. The item score is calculated as the following:

$$ItemScore_{q,r} = \alpha \times ItemSim_{q,r} + (1-\alpha) \times TagRecency_q \quad (11)$$

where α is a prespecified factor that determines the impact scale of the user's preference and the item similarity.

IV. Experiments and Results

The experimental dataset was collected from del.icio.us, which is a popular web site that helps users share links to their favorite information items. We crawled through del.icio.us to randomly collect newly active users. The resulting dataset contains 163,274 bookmarked items tagged using 123,784 tags by 397 active users. The user profile of a specific user includes the collected resource items, tags, and the number of items contained in each tag. The item profile for a specific user's collection includes the tags (name) of an items used by this user, and the tags (name and frequency) that is collected (tagged) by all users. These comprise five tables in the database, including user, item, tag, user-tags, user-items, and item-tags.

A. Experimental design

(1)Training and test set

For each user's collection, the collected items were sorted according to the time when it was bookmarked. And then each user's collections were divided into two parts, 80% of items for the training set and the other 20% of items for test purposes based on items' bookmark time. Each user is rotated as the target user. The number of items hit is counted if the recommended items correspond to the taget user's collected items in the test set. That is, a hit is determined by the similarity between the user's collected items and the recommended items predicted. The average performance for target users is computed to evaluate the reliability of the proposed recommender system.

(2) User clustering

We used the E-M clustering approach to cluster users. The default number of clusters is set to 25. We removed clusters whose number of members was less than five to improve the recommendation quality.

(3) Evaluation measures

The test set consist of items bookmarked in the test period by the target users. Three information retrieval (IR) performance measures are considered in evaluating the effectiveness of the recommender system: recall, precision, and F1 which incorporates the first two measures [13]. Precision is the ratio of a target user's hit items retrieved from the recommended top-N items. Recall, which gauges quality, is the number of items of a target user's hit items divided by the total number of items bookmarked in the target user's test period. In this study, the F1-Measure combines recall and precision with equal weights as seen in Eq. (10).

$$F1 = \frac{recall \times precision}{(recall + precision)/2} \quad (10)$$

In the current study hit rate is defined as the number of items that the target user has actually bookmarked within the recommendation list; this measures precision. Thus, a full keywords analysis is performed to calculate the similarity among items. For each item, its HTML tag, cascading style sheet, JavaScript and stop words were removed in advance. Thirty keywords that are most frequent were extracted from each item. If the similarity between the user's items and their prediction items are greater than 70%, it is defined as a hit. Other hit thresholds are shown in the following charts as well.

B. Experimental Results

For model comparison, two types of data are analyzed: one is the TF-IDF with user recency preference, and the other is just the TF-IDF which did not incorporate the user recency preference. The proposed hybrid model with recency preference is represented as "Hybrid with Recency." This model is compared with the model without using the recency preference which is represented as "Hybrid without Recency" to show the relative effectiveness of using the recency component.

In order to make a comparison among other models, three types of benchmark models (see Table 1) were also performed as follows:

(1)The traditional collaborative filtering (TCF), is constructed in order to know its relative performance. Based on the clustering result, TCF recommends top-N frequent (popular) items in the cluster to the user in the same cluster.

(2) The top-N popular recommendation represented as POP, recommends the top-N frequent (popular) items of all users to the target user, without considering the clustering results.

(3) To identify the relative performance improvement of our

proposed system against that of the random recommendation (no-model approach), we conducted a performance comparison with random recommendation, which is represented as "Random." The random recommendation arbitrarily recommends items from a cluster to the users from the same cluster.

The performance measures of precision, recall, and F1-Measure are shown in Fig 1, 2 and 3. Various threshold values of hit were used to compare the performance of these measures. Hit thresholds were set to 0.5, 0.6, 0.7, and 0.9. The number of items given to users to choose from, that is, the recommendation size, was set to 70. Experimental results show the rank of these performance measures is as follows: Hybrid with Recency > Hybrid without Recency > Traditional CF > Popular Top-N > Random.

We summarize this result as follows:

(1)The consideration of recency preferences for the target users is a key necessary element in this study.

(2)Our hybrid recommendation is superior to the traditional collaborative recommendation approach.

(3)Using top-N popular recommendation and random recommendations yield comparatively poor performance. Our proposed model is better than the two recommender systems since it yields higher precision, recall and F1 values.

The proper recommendation size could be an appropriate topic for future studies when one assesses precision and recall measures, and especially when one considers different subject domains such as books, documents, or audio for instance. In this study, the recommendation size of 30, 40, 50, 60, 70, and 80 were conducted for a particular cluster. The precision, recall and F1-measures are illustrated in Figure 4. Here we see that the F1 and recall measures may be slightly improved by increasing the recommendation size, however, it is best not to recommend too many items to users in order to avoid overloading them with information.

TABLE I
THREE BENCHMARK MODELS

	Clustering	Recommending top-N popular items
Traditional CF	yes	yes
Popular Top-N	no	yes
Random	no	no

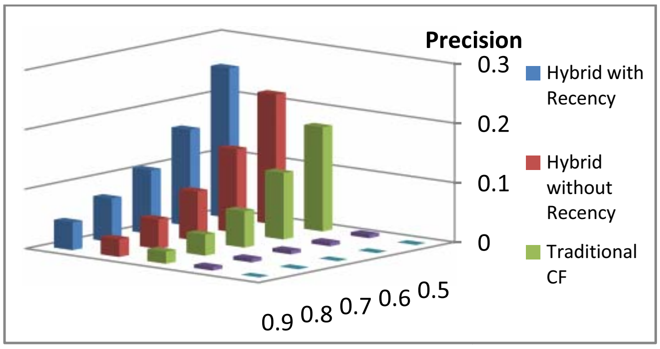

Fig. 1 Precision under various hit thresholds

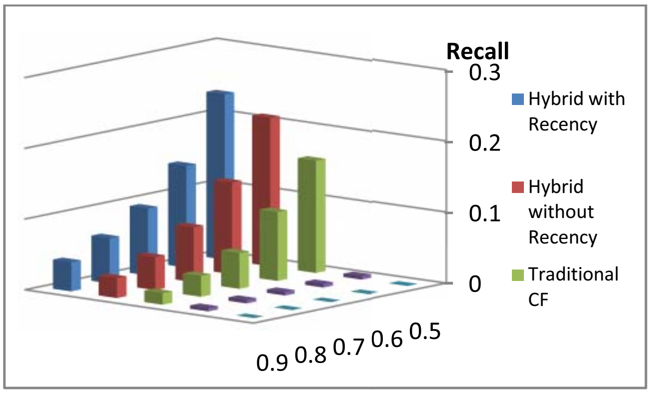

Fig. 2 Recall under various hit thresholds

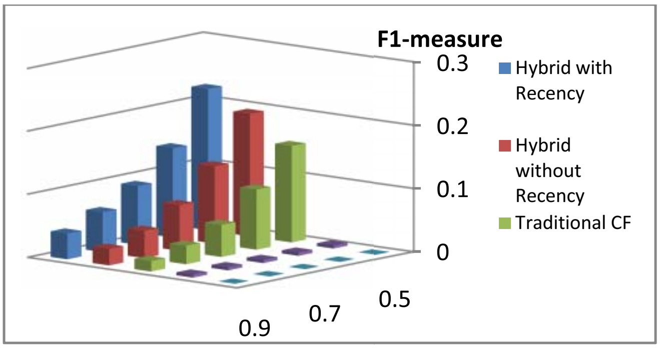

Fig. 3 F1-Measure under various hit thresholds

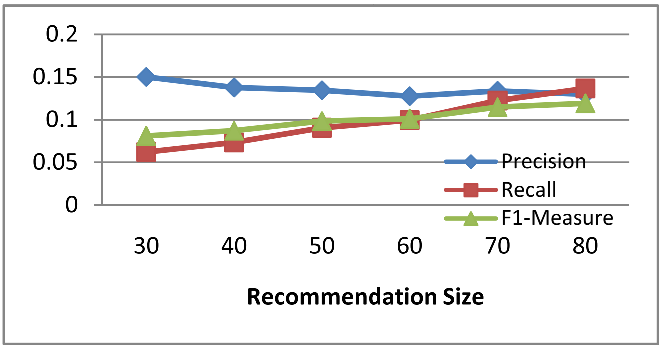

Fig. 4 Performances of the recommendation size

V. CONCLUSIONS AND FUTURE WORKS

In the social tagging system, the newly tagged items by the user are more relevant and important for the user. This study used a two-stage recommendation approach. First, collaborative filtering found similar user groups through a clustering algorithm employing the tag-based user preference profile. And then content-based filtering recommended resource items to target users by analyzing the tag-based content of the target user's collected items.

From the experimental results of the dataset available through the del.icio.us website, we found the proposed hybrid recommender system provided a higher measure of precision and recall which is important since they influence the effectiveness and quality of recommendations. Improved recommendations, more intelligent online tools and decision aids have a promising future on the world wide web and they are expected to be the subject of ongoing research in the years ahead[14].

The current study has demonstrated that tag information can be used to represent users' preferences in a social bookmarking website. Furthermore, the effect of recency plays an important role in the proposed recommender system by making it better than traditional collaborative recommendation systems. Finally, the proposed model can be adapted in many application areas where tagging is suitable, such as books, articles, documents, pictures, audio and video.

REFERENCES

[1] H.-N. Kim, A.-T. Ji, I. Ha and G.-S. Jo, "Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation," Electronic Commerce Research and Applications, vol. 9, Issue 1, pp. 73-83, January-February 2010.
[2] C. H. Brooks and N. Montanez, "An analysis of the effectiveness of tagging in blogs," in: Proceedings of the 2005 AAAI Spring Symposium on Computational Approaches to Analyzing Weblogs. Stanford, CA. March 2006.
[3] P. J. Morrison, "Tagging and searching: Search retrieval effectiveness of folksonomies on the World Wide Web," Information Processing & Management, vol. 44, Issue 4, pp. 1562-1579, July 2008.
[4] D. H. Widyantoro, T. R. Ioerger, J. Yen, "Learning User Interest dynamics with a three-descriptor representation," Journal of the American Society for Information Science and Technology, Vo. 52, Issue 3, pp.212 – 225, February 2001.
[5] D. H. Widyantoro, T. R. Ioerger, J. Yen, "An adaptive algorithm for learning changes in user interests," in: Proceedings of the Eight ACM International Conference on Information and Knowledge Management, pp.405-412, 1999.
[6] M. Pazzani, D. Billsus, S. Michalski and J. Wnek, "Learning and revising user profiles: the identification of interesting web sites," Machine Learning, Vol. 27, No. 3, pp.313-331, 1997.
[7] C.-L. Huang and W.-L. Huang, "Handling sequential pattern decay: Developing a two-stage collaborative recommender system," Electronic Commerce Research and Applications, vol. 8, Issue 3, pp.117-129, May-June 2009.
[8] R. Jäschke, A. Hotho, C. Schmitz, B. Ganter and G. Stumme, "Discovering shared conceptualizations in folksonomies," Web Semantics: Science, Services and Agents on the World Wide Web, Vol. 6, No. 1, pp.38-53, February 2008.
[9] C. Marlow, M. Naaman, D. Boyd and M. Davis, "HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, ToRead," in: Proceedings of Hypertext, New York: ACM Press, 2006.
[10] M. Memmel, M. Kockler and R. Schirru, "Providing multi source tag recommendations in a social resource sharing platform," Journal of Universal Computer Science, vol. 15, no. 3, pp.678-691, 2009.
[11] S.A. Golder and B. A. Huberman, "The structure of collaborative tagging systems," Journal of Information Science, vol. 32, no. 2, pp. 198-208, 2006.
[12] J. Han and M. Kamber, Data Mining Concepts and Techniques. Morgan Kaufmann Publishers, San Francisco, USA, 2006.
[13] G. Kowalski, Information Retrieval Systems: Theory and Implementation. Kluwer Academic Publishers, Norwell, MA, 1997.
[14] M. Conyette, "Determinants of Online Leisure Travel Planning Decision Processes: A Segmented Approach" (Unpublished doctoral dissertation, University of Newcastle, Newcastle), 2010.