Improving the Accuracy of Tagging Recommender System by Using Classification

Jian Song, Liang He*, Xin Lin

Department of Computer Science and Technology, East China Normal University, Shanghai, China

songjian211@126.com, lhe@cs.ecnu.edu.cn, xlin@cs.ecnu.edu.cn

***Abstract*—Collaborative tagging system has become more and more popular and recently achieved widespread success due to flexibility and conceptual comprehensibility of tagging systems. Recommender system has the access to adopt tagging systems to achieve better performance. In this paper we consider that the items can be categorized into different classifications in which users show different interests. Here we adopt a two-step recommender method called TRSUC (Tagging Recommender Systems by Using Classification) which can be described as Inner-Class Recommender or Global Recommender in which we use tag as the intermediary entity between user and item. The experiment using MovieLens as dataset shows that we acquire better results than the recommender algorithms without classifying the items.**

***Keywords*—Collaborative Tagging; Recommender System; Classification**

I. INTRODUCTION

Social tagging is the process by which users add metadata in the form of keywords to annotate and categorize items such as songs, movies, merchandise, etc. Social tagging is associated to the web2.0 technologies and has already become an important source of information for recommender system. There are many outstanding recommender systems using tags to make recommendation. For example, Last.fm, Amazon, MovieLens, etc., which recommend items to users based on tags who possess common interest with other similar users. Compare to the two common state-of-the-art recommender methods, using tags has a lot of advantages. As a matter of fact, collaborative-filtering system uses pattern in the user ratings to make recommendation, and this method shows barely anything about the items' true content or what they are related to, likewise content-based system also suffers from some limitations such as the information retrieval techniques and features associated with the objects (e.g., graphical images, audio and video) [2]. Because tags are deliberately annotated by users, which possess useful information not only about the users who tagged but also about the items they labelled. Tags can represent concepts meaningful and reflect the comprehension to the items by the users themselves. While different users have different understanding of the items, tags can also indicate the different aspects of one item, that is to say, they can reflect the items full-scale character. Since tags are easily comprehended by users, tags serve as a bridge enabling users to better understand an unknown relationship between an item and themselves [5].

In this paper, we choose to use tag to describe the features of both the users and items. Besides, Classification is always crucial information of items. When users choose items or topics, what they do might use is the classification information to filter out the irrelevant frustration resources which might disgruntle them, distracting their mind and wasting their time. And in our research we find that user also show different interests in different classifications. So in this paper, we purpose a new approach called TRSUC by mapping users' interests to different classification to improve the performance of the tagging recommender systems. The final results we acquire demonstrate that recommendation system with tagging system through classification can outperform that without the step of classification.

The remainder of this paper is organized as follows: the next section provides details on related work, mostly concentrating on research work in related areas, section III describes our method for improve tagging recommender system. In section IV we present the results of our study. The summary and future work of the research is presented in section V

II. RELATED WORK

Before the collaborative tagging system appears, collaborative filtering (CF) algorithms were widely used. GroupLens Usenet recommender employed user-based algorithms: given a particular user, recommend movies which similar users prefer [5]. Similarly, the item-based algorithms predict users' ratings for a specific item based on their ratings for similar item. Beside, paper [12] proposed a Time-context-based Collaborative Filtering Algorithm. In [11], the authors adopted SVD (singular value decomposition) due to its accuracy, efficiency, and ease of implementation. Some hybrid methods were also introduced like [2], combining user-based and item-based CF to improve the performance of their recommender system.

Both the collaborative filtering algorithms and SVD algorithms rely on patterns between user ratings and they lose a lot of information about the items. So some recommender systems like Fab [10] where the researchers eliminated many of the weaknesses web pages by combining CF and content-based algorithms to make recommendation better. But the applications of this method are always limited by the feature of the items and the technologies of the information retrieval.

Collaborative tagging systems have become popular and have achieved widespread success recently due to the flexibility and conceptual comprehensibility of tagging systems. A well-accepted tripartite model of collaborative tagging systems has already been theorized as figure 1 [9].

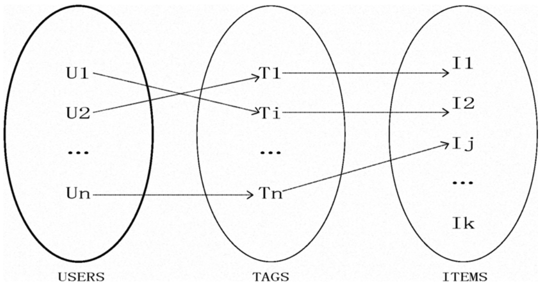

Figure 1. Tripartite graph structure of a tagging system.

There are three main entities that compose any tagging systems: the users of the systems (people who actually do the tagging), the items being tagged, and the tags themselves. The tags provide the link between the users and the items they tagged and also reflect both the preference of the users and the content of the items. Each of these can be seen as forming separate spaces consisting of sets of nodes, which are linked together by edges. The first space, the user space consists of the set of all users of the tagging system where each node is a user. The second space is the tag space, the set of all tags, where a tag corresponds to a term or phrase in natural language. The third space is the item space, the set of all tags. A tagging instance can be seen as the two edges that links together which means a user to a tag and that tag to a given item.

For now, many approaches have been adopted to make recommendation by using tags. In [8], the author used tags to model user profile and took advantage of WordNet to calculate the similarity of the users profile to find users' neighbours. In GroupLens's recent research [5], the researchers proposed that users' preferences for tags based on their interactions with tags and movies, while users' preference for movies based on their preferences for tags.

In this research, we have a similar view with GroupLens's research. In our opinion, tags can represent both the users' and the items' feature. Generally speaking, tags are annotated to items and can be portrayed as the description of items. Here, we also adopt tags which users have assigned to represent his interests. In this way, tags play a role of intermediator to help users find his potential items. Besides, classification is really considerable information to describe items. In our work, we find that users have different preference to each classification, that is to say, we can use items' classification information to divide users' interests. Using this method we can map users' interests to each classification with different weights. Taking into account of all the above points, we propose a new method to improve the accuracy of tagging recommender systems by using classification and the details of the method will be discussed in next section.

III. RECOMMENDER METHOD

A. Tag-based recommendation

In this section, we first focus on the method using tags to connect users and items. As we discussed above that tags can both represent users' interests and items' features like [3], we adopt the terms: tag relevance and tag preference, however we have a different method to measure tag relevance and tag preference.

1) ***Tag relevance:*** We use the term tag relevance to describe the relationship between a tag and an item. In the field of information retrieval and text mining, the TF-IDF weight is always used to evaluate how important a single word to a document, but it offset by the frequency of the word in the corpus. The TF is the measure of the importance of a word to a document, while the IDF is used to filter out some very common words (e.g., the, is). In [7], the authors classified tags as generally factual, subjective, or personal. In addition, most tags falls into the factual class, followed by subjective class, that is to say, most tags are strongly represent what the item is about. Paper [1] explored a significant conclusion, the combined tags of many users' tagging give rise to a stable pattern in which the proportions of each tag are nearly fixed. According to the aforementioned conclusions, we argue that tags applied by many users is probably more relevant to the given item, so the TF weight can be used as the measure of the relevance between tags and items and those tags can be used to describe the feature of a certain item. Consequently, we give the formal definition:

$$\text{tag\_rel}(i,t)=\frac{n_{t,i}}{\sum_{k\in T_i} n_{k,i}} \qquad (1)$$

2) ***Tag preference:*** We define tag preference as the user's sentiment towards a tag. This is the process to select the weighted tags to represent a specific user. The key design choices concern how to compute the value for a given user and tag. In this research, we infer users' tag preference base on their ratings. If a user tends to give high ratings to items that have a particular tag, and the system could infer that the user should have a positive preference towards the aforementioned tag. To estimate a user's preference for a single tag, we compute a weighted average point of a user's ratings of movies with it. In addition, tag relevance is another factor we consider. For example, in our dataset, users have tagged "Reservoir Dogs" with "violence" 7 times, while "Planet of the Apes" has only been tagged once. Evidently, Reservoir Dogs should be assigned a higher weight when computing a user's preference for “violence”. We now formally define the measure we use to estimate tag preference, if I_u is user u’s rating for item i:

$$\text{tag}\_pref(u,t) = \frac{\sum_{i \in I_u} r_{u,i} \cdot tag\_rel(t,i)}{\sum_{i \in I_u} tag\_rel(t,i)} \qquad (2)$$

Here, is the set of items tagged by user u with tag t. The sum in both the numerator and denominator ignore items the user has not rated, and then we normalize tag_pref (u, t). So that the weights for each tags user prefer sum to 1.0.

3) *Tag-based recommendation:* In the previous sections we evaluate methods for inferring users’ preferences for tags based on interest in tags and items. We now shift our focus to how to use those inferred tag relevance and tag preference to make recommendations. The tag-based recommendation is inspired by algorithms from information retrieval that calculate the similarity between a user’s profile and a document’s term vector [5]. In information retrieval, the columns in each vector correspond to words. In the tag-based recommendation, the columns correspond to tags. To generate a prediction for an item I, the tag-based recommendation calculates the dot product between users’ preferences for item I’s tags and the tag relevance between tag t and item I. Then user u’s predicted score for item i is:

$$\begin{aligned} &predict\_score(u,i) \\ &= \sum_{t \in T_i} tag\_pref(u,t) \cdot tag\_rel(i,t) \end{aligned} \qquad (3)$$

B. Tag-based recommendation using classification

1) *Class preference*: Items in the real world normally can be classified into different categories. And there are two main methods to classify items: classification algorithm and clustering algorithm. Classification algorithms classify items by using item categories predefined; while clustering algorithms usually auto generate categories by analysing characteristic (often tags of content). KMeans algorithm is a common clustering algorithm. In this paper, we use classification algorithm to classify items.

In this section, we propose a new conception: class preference. When a user is browsing a website or choosing an item, classification information is always the first step to filter out items they do not concern, for example, in the movie fields, some users may like “Action” films a lot, so they will not pay much attention to movies fall into other classifications like “Romance”, while, some other users behave just the opposite. This means that user may show different interests to each classification. In this way, we can divide users’ interest into several pieces according to the classification information of the items. So, we can map users’ interests to each classification with different weighting which we called class preference and it can be illustrated as figure 2:

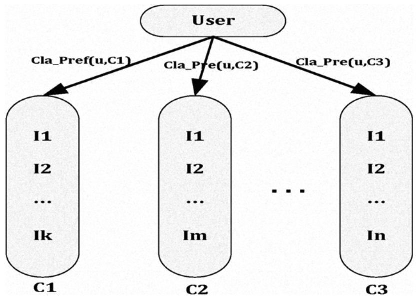

Figure 2. The mapping of a user to the classifications.

With the class preference coefficient we can help users to accomplish the classification choosing step.

Now, what we care is how to estimate a user’s preference to each classification. In this research, we choose ratings to be the measure of class preference. If a user tend to give the items belong to a specifically classification high scores, we conclude that the user may like the classification much. So we go through the space of items of each classification and the whole item space, compare the average ratings of each classification and the average ratings of all the items the user assigned which we define as follows:

$$w(u, c_k) = \frac{\sum i \in c_k \cap i \in i_u \, r_{u,i}}{\sum i \in i_u \, r_{u,i}} \qquad (4)$$

We then adopt exponential function to make the weight smoother and closer to 1.0, and the modified weight could be:

$$class\_pref(u, c_k) = e^{(w(u,c_k)-1)} \qquad (5)$$

2) *Tag Recommendation using Classification*: As we discussed above, users often first filter out some information they do not care much and then choose the interest items in the specifically classification. Here our method to make recommendation is just reverse this process. Our recommendation is divided into two steps; figure 3 is the model of our recommender system:

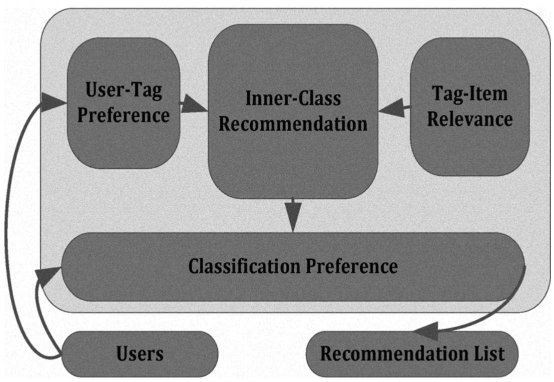

Figure 3. The model of our recommender system.

The Inner-Class Recommender Process: This process makes recommendation in each classification respectively. Here, we adopt the method we talked above: using tag-preference and tag-relevance to describe the users and items respectively. Then we make recommendation using the tag-based recommendation approach and generate a recommendation list in each classification. So the Inner-Class Recommender can be expressed as follow:

$$inner\_score(u,i) = \sum_{t \in T_i} tag\_pref(u,t) \cdot tag\_rel(i,t) \quad (6)$$

The Global Recommender Process: The Global Recommender Process is the way we simulate users' behavior for helping them to find the items users more interested in. The final recommendation list is generated by calculating Item_Rank (u, i):

$$Item\_Rank(u,i) = \max\{class\_pref(u,c_k) \cdot Inner\_Score(u,i)\} \quad (7)$$

Here, some items may fall into several classifications, so we put the items with highest predict score.

IV. EXPERIMENTS

We demonstrate the working of our approach on the dataset from MovieLens, which contains 10000054 ratings and 95580 tags applied to 10681 movies by 71567 users of the online movie recommender service. In the dataset movies are classified into 18 classifications which we use for our analyses: Action, Adventure, Animation, Children's, Comedy, Crime, Documentary, Drama, Fantasy, Film-Nior, Horror, Musical, Mystery, Romance, Sci-Fi, Thriller, War, and Western. We performed two experiments to compare the performance of TRSUC with the tagging recommender systems without using classification information.

In the first experiment, we will show the class preference calculated in our method. Three users are randomly chosen from our dataset, as we illustrate in figure 4, uses show different interest in each classification. User 69388 pays more attention to Animation, Film_Noir, IMAX and Western, but gives much lower score to Horror movies. So when we make recommendation we should lay a lower weight to Horror movies to user 69388. Other users are also having varying degree of preference to different classifications. This makes us more confidence to use classification information to modify the result of recommendations for users.

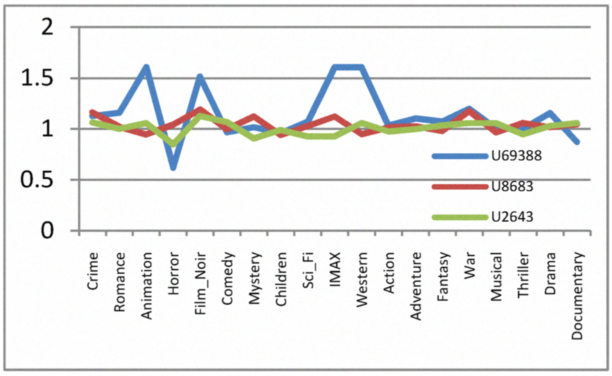

Figure 4. The distribution of users' preferences to different classifications.

In the second experiment, the movies which previously tagged by the user are withheld. We then generate a top-N list of recommendation and record whether each withheld movie should appear or not. For each active user u, I represents the set of movies that u has tagged. Let R_{ui} be the set of movies in the recommendation list. The average precision is thus defined as:

$$top\_n\_precision = \left(\sum_{u \in U} \frac{|R_{ui}|}{|I_{ui}|}\right) \Big/ |U| \quad (8)$$

We go through the whole user space and do the same experiment on each user. Then the statistical average is calculated for all the users. This experiment measures the algorithm's ability to find the relevance movies for active users. High top-n-precision corresponds to greater accuracy in the algorithm's recommendations. In our experiment, for each active user, we search the top-5, top-10 and top-15 recommendation list each time. The result of TRSUC and the tagging recommender system without classification are showed in figure 5.

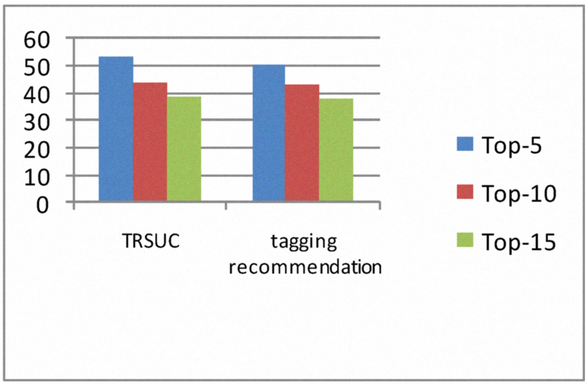

Figure 5. Compare precision between TRSUC and tagging recommendation without classification.

Evidently, our approach makes a higher precision, especially in top-5.

In the end, we present a detailed analysis of the time complexity of the proposed approach TRSUC, and the comparison between TRSUC and traditionally tagging recommendation is also discussed. Although the calculation of the class preference may eat into a lot of CPU time, but as we all know, user's preference to each classification will change not very quickly, so we can accomplish the calculation of the class preference offline, and repeat this process periodically to update this coefficient. The inner-class recommender process goes through the items in each classification, compare to going through the whole item space of the traditionally tagging recommendation, our method scatter the items in various classifications and generate the recommender lists in each class. So, our method will do contributions to real-time recommendations online. And as the clustering of user preference have a low frequency, so we argue that this will not be the bottle neck or weight to the system.

V. CONCLUSIONS

In our work, we use tags represent both users' and items' feature, and find the relationship between users and items using tag relevance and tag preference. Besides that we also found that users often filter out some items they do not pay much attentions by using the classification information. So, we reverse the process users find items, before generate the final recommendation list we assigned different weights for each classification to modify recommendation list in different classification. The experiment results demonstrate that our method improved the accuracy of tagging recommender systems. For the future work, we want to find a way to make the classification preference coefficient more accuracy, besides ratings, we will consider users' tagging behaviors. And we need to find a better way to describe users' and items' feature using tags to make them close contact with each other.

ACKNOWLEDGMENT

The authors would like to thank the support from the Natural Science Foundation of China(60903169), the National Science and Technology Ministry of China Foundation (2007BAH09B04), and the Significant Scientific and Technological Project of Shanghai Science and Technology Commission foundation (08DZ15001(10)).

REFERENCES

[1] Scott A. Golder and Bernardo A. Huberman. "The Structure of Collaborative Tagging Systems," HP Labs Technical Report. 2006. http://www.hpl.hp.com/research/idl/papers/tags/

[2] G.Adomavicius, A.Tuzhilin. "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and possible Extensions," IEEE transactions on Knowledge and Data Engineering, Vol. 17, No. 6. pp. 734-749, 2005.

[3] J,Vig, S.Sen, and J.Riedl. "Tagsplanations: Explaining Recommendations Using Tags," Proceedings of the International Conference on Intelligent User Interfaces, 2009, pp 47-56.

[4] S.Sen, J.Vig, and J.Riedl. "Learning to Recognize Quality Tags," Proceedings of IUI, 2009 , pp 87-96.

[5] S.Sen, J.Vig, and J.Riedl. "Tagommenderers: Connecting Users to Items Through Tags," International World Wide Web Conference, April 2009, pp 671-680.

[6] S.Sen, F.M.Harper, A.LaPitz, and J.Riedl. "The Quest for Quality Tags," Proceedings of the 2007 international ACM conference on Supporting group work, 2007, pages 361-370.

[7] S.Sen, S.K.Lam, A.M. Rashid, D.Cosley, D.Frankowski, J.Osterhouse, F.M. Harper, and J.Riedl. "tagging, communities, vocabulary, evolution," Proceedings of the ACM 2006 Conference on CSCW, 2006, pp: 181-190.

[8] Shiwan Zhao, Nan Du, Andreas Nauerz, Xiatian Zhang, Quan Yuan, Rongyao Fu. "Improved Recommendation based on Collaborative Tagging Behaviors," Proceedings of the 13th international conference on Intelligent user interfaces, 2008, pp: 413-416.

[9] H.Haplin, V.Robu, H.Shepherd. "The complex Dynamics of Collaborative Tagging," Proceedings of the 16th international conference on World Wide Web, 2007, pp:211-220.

[10] M. Balabanovi , Y. Shoham, "Fab: content-based, collaborative recommendation," Communications of the ACM. 1997.

[11] S. Funk. "Netflix Update: Try This At Home," Sifter. Org/simon/journal/20061211.html, 2006.

[12] He Liang,Wu Faqing, "A Time-context-based Collaborative Filtering Algorithm," The 2009 IEEE International Conference on Granular Computing, 2009,p209-213.