International Journal of Soft Computing 10 (1): 83-86, 2015
ISSN: 1816-9503


# Weight Based Approach for Resource Recommendation

P. Pabitha, G. Amirthavalli and M. Rajaram
Department of Computer Technology, MIT, Anna University, Chennai, India

**Abstract:** Recommendation systems refines enormous amount of data and provides suggestions on items or ratings of items which is appealing to the user. Although, the recommender systems depends on content based approach or collaborative filtering technique to make recommendations, these methods suffers from cold start and data sparsity problem. To overcome the limitations of the aforementioned problems, a weight based approach is proposed for better performance. The main criterion for building a recommender system is to exploit useful content and provide better recommendations with minimal processing time. The proposed system is a web based client side application which uses user profiles to form neighbourhood and calculates ratings using weights. For newcomers a profile is constructed based on past history and current preference. The resources which might be of interest to the user are predicted from calculated ratings.

**Key words:** Weight based approach, resources, neighbourhood, recommender system, ratings

## INTRODUCTION

Recommender systems (wikipedia.org) are now an integral part of online sites. They are very useful in recommending items or products to user according to their interests. They are active information filtering systems that attempt to predict the items which the user might be interested. They work by namely user's opinion, ratings and preferences to make recommendations accurately.

Recommender systems are of many kinds namely: content-based-recommend items similar to the items users preferred in the past, collaborative-recommend items that people with similar preferences liked in the past, hybrid based-combine content and collaborative-based methods, preference-based recommendations-recommend items according to relative preference for the user instead of individual ratings.

The resources in the recommender systems are items in the web page. An item may be a document, internet bookmark and image, audio or video file. The tags are the words, phrase or terms that is used to represent the item. This kind of metadata helps describe an item and allows it to be found again by searching or browsing. They are informal data created based on user's choice.

**Literature review:** Researchers describes a content based recommender system that utilizes the properties of Folksonomies (tags) (Semeraro *et al*., 2012). The system analyses the features of art works previously rated by a visitor and build a visitor model or profile. The user profile is then matched against the attributes of new items in order to provide personalized suggestions. Training examples for a user who has access to the system for the first time is not available.

The system designed uses quadratic concepts to provide personal recommendations on tags, users and resources (Jelassi *et al*., 2013). It requires user's profile to make recommendations. All the tags corresponding to a particular resource are returned rather than meaningful tags. All the resources which are viewed by other users are returned rather than those which are apt for the new user.

In an agent assisted system the information retrieval is done in two phases. The first phase is FolkRank algorithm where the prediction is done on the Folksonomy structure constructed and then applying the adapted version of Page Rank algorithm. The second phase is Tag Querying which uses Fuzzy formal concept analysis is used to model two kinds of profiles namely tag and resource profile (Fenza *et al*., 2011).

Tags are recommended using collaborative filtering approach based on user profile in the system devised by Xu *et al*. (2012). A weighted network of co-occurrence of tags is constructed for selecting the top pairs of tags. The method to determine the weight for ranking the tags is not known. The pairs of tags that should be shown to the users for extracting preferences are not specified.

In item based collaborative filtering the tags (items) are used to compute the similarity between the current user and each user. It filters out users with different interests but it is not clear whether the same can be used for recommending tags (Hui *et al*., 2011).

**Corresponding Author:** P. Pabitha, Department of Computer Technology, MIT, Anna University, Chennai, India

Tag based smoothing technique for recommending items uses quantitative values for smoothing (Peng and Zeng, 2010). Tags ranked at the front (left) of a bookmark are more important than those tags ranked at the end (right). It makes use of tags for item recommendation without discarding the user-item interaction information. The two co-efficient defined in the study are saving confidence and saving likelihood.

The neighbourhood formation in the system projected by the researcher uses clustering techniques (Sarwar *et al*., 2002). Clustering techniques work by identifying groups of users who appear to have similar preferences. Once the neighbourhood is obtained, Classical Collaborative Filtering algorithm is used to generate prediction from that. Clustering techniques usually produce less-personal recommendations than other methods and most often lead to worse accuracy than nearest neighbour algorithms.

The hybrid approach is used for solving the problem of finding the ratings of unrated items in a user-item ranking matrix through a weighted combination of user-based and item-based collaborative filtering (Badaro *et al*., 2013). The technique provides improvements in addressing two major challenges of recommender systems: accuracy of recommender systems and sparsity of data by simultaneously incorporating users' correlations and items ones.

Some of the identified are Coldstart-information is not available for new users, data sparsity-less amount information.

## MATERIALS AND METHODS

The proposed approach aims to recommend resources to the user by considering the user's preferences and previous browsing history. This approach combines collaborative and content based filtering techniques to enhance the quality of suggestions. The query from the user is received as input based on which a user profile containing three fields namely ID, age and preference is created. The newly built user profile is then updated in the database for further processing.

The group of users whose preference matches with the current user goal is identified by considering the users profile. The profile of the current user is then compared with different groups of users that were classified previously and the most appropriate neighborhood is identified. The two sets of weights $\alpha$ and $\beta$ are used to boost up the performance and normalize the ratings to avoid data sparsity problem. Since, the rating scale is from 1-5 two weights are used. The $\alpha$ is the maximal appropriate of ratings ranging from 4-5 while $\beta$ is the minimal appropriate of ratings ranging from 1-3. The value of $\alpha$ and $\beta$ are set based on the count of the ratings done by other users:

$$\alpha+\beta = 1$$

$$\alpha \leq 1 \text{ and } \beta \leq 1$$

- Case 1: only high ratings (5, 4) ($\alpha = 1, \beta = 0$)
- Case 2: only low ratings (3, 2, 1) ($\alpha = 0, \beta = 1$)
- Case 3: count of high and low ratings are same ($\alpha = 0.5, \beta = 0.5$)
- Case 4: count of higher ratings>lower ratings ($\alpha = 0.7, \beta = 0.3$)
- Case 5: count of higher ratings<lower ratings ($\alpha = 0.3, \beta = 0.7$)

Ratings are calculated by using the following equation:

$$\alpha(\text{avg of ratings} \geq 4)+\beta(\text{avg of ratings}<4)$$

The architecture for recommending the resources by using the weight based technique is given in Fig. 1.

**Implementation:** The system designed is used for recommending books to the user. It is implemented using Net Beans (wikipedia.org) and My SQL database (wikipedia.org). The dataset used contains about 150 user profiles, 150 books, >3500 ratings. The user profile consist of three fields namely id, age, preference.

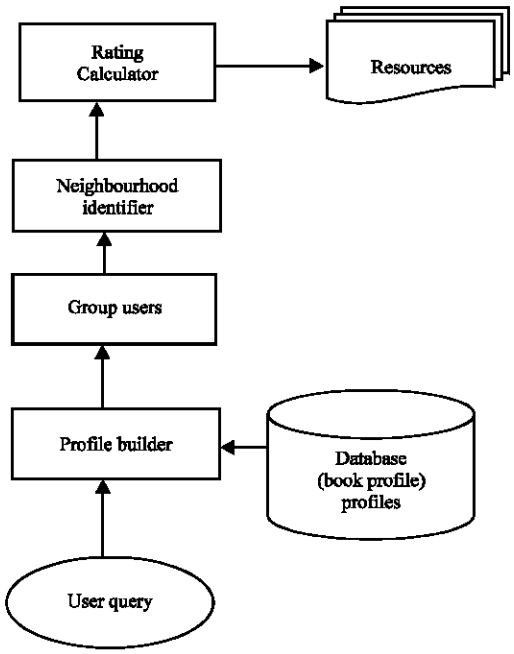


Fig. 1: Architecture of weight based recommender

Table 1: Results obtained for the input

| Input | Profile built | Neighbourhood (user id) | Books identified | Ratings calculate | Output |
|---|---|---|---|---|---|
| Mammoth, Thriller | (15, 18-24, Thriller) | 31 32 33 34 35 36 40 41 44 45 50 51 52 53 54 55 56 57 58 64 65 66 67 68 69 70 71 73 144 146 | The Spy who came in from the cold, whiteout, the silence of the lambs, shutter island, coma. The surgeon, the andromed a strain, one shot, shadow divers | 75 = 4.18, 69 = 4.01, 74 = 4.01, 68 = 4.0, 70 = 3.64, 71 = 3.64, 77 = 3.6, 78 = 3.57 67 = 3.41 | The Androme da strain, eye of the needle, the surgeon, whiteout, the silence of the lambs |

The book profile contains book id, name and genre. The age group is classified as follows under 10, 10-17, 18-24, 25-39, 40-55 and above 55. The genres of books are Science fiction, fairy tales, romance and humour. The steps for recommending books to the user are explained as.

**Step 1:** Get the user input, build user profile based on his preferences and group the users according to their profile.

**Step 2:** Identify the neighbourhood in which the active user is present.

**Step 3:** Find the required resources from the neighbourhood.

**Step 4:** Recommend the books from the calculated ratings.

The actual process which happens for the input as "mammoth" and type as "thriller" is tabulated in Table 1. Based on the past browsing record a user profile is constructed. The users are grouped based on age. The neighbourhood of the active user is found to be the users in between the ages 18-24. The books which are more popular among the users in the neighbourhood are retrieved. To filter large amount of books a threshold value of 7 is chosen and used during retrieval process.

The ratings are calculated for the books identified from the neighbourhood and stored in hash map. A Hash map is represented by <key, value> pair. This is mapped to the system as <Bookid, ratings>. Finally the hash map is sorted and the top books are recommended. For the input the output obtained is The Andromeda Strain, Eye of the Needle, The surgeon, Whiteout, The Silence of the Lambs and Shutter island.

## RESULTS AND DISCUSSION

To improve the efficiency of the system the best value of weights are chosen by plotting the average ratings obtained by applying various sets of $\alpha$ and $\beta$ on different genres of books.

In the Fig. 2 the straight line corresponds to the average ratings obtained for different genres of books the line with square marker corresponds to the value of $\alpha = 0.4$, $\beta = 0.6$, the dashed line corresponds to the value of $\alpha = 0.5$, $\beta = 0.5$ and the dotted line corresponds to the value of $\alpha = 0.3$, $\beta = 0.7$. The following inferences on comparing with the ratings calculated over average are made:

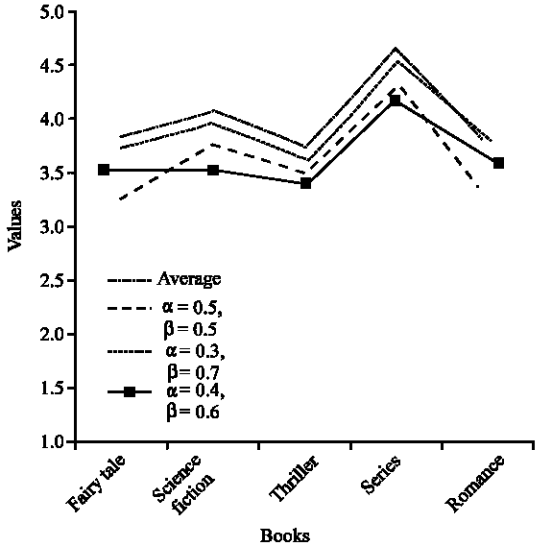


Fig. 2: Ratings obtained for different values of $\alpha$ and $\beta$

- For $\alpha = 0.5$, $\beta = 0.5$ the ratings calculated are very much lower than those obtained by calculating the average thereby decreasing the accuracy
- For $\alpha = 0.4$, $\beta = 0.6$ the ratings obtained increase the accuracy than $\alpha = 0.5$, $\beta = 0.5$ but it either increases or decreases in irregular way
- For $\alpha = 0.3$, $\beta = 0.7$ the predicted ratings appear to be normalized to overcome the varying high range and low range ratings

From Fig. 2, it is evident that using the values 0.3 and 0.7 as weights for calculating the ratings has provided a boost for predictions. The remaining values of weights have either increased or decreased the ratings in an abnormal manner.

## CONCLUSION

Recommender systems suggest ideas to the users which will help them to take better decisions and learn about their interests. The technique used to implement system uses weights to predict the ratings in an effective manner. The usage of weights has been analyzed by experimenting with different values, genres of books and has been found to improve the quality of predictions by calculating ratings which provide normalization, when less amount of ratings are available. Thus, the books which the user might prefer is suggested by creating a new profile and predicting the ratings based on weights. The proposed system exploits the popularity of a resource which is indicated by the ratings. With no ratings available there occurs a scenario where the resources cannot be predicted even it being better than other resources. To improve the quality of recommendations the user-resource similarity can be considered. The grouping of the user profiles is a simpler way of identifying neighbourhood which can be enhanced further by forming clusters using Machine Learning algorithms that will be helpful in predicting resources effectively. For the most popular resource tags can be recommended using similarity Learning algorithms.

## REFERENCES

Badaro, G., H. Hajj, W. El-Hajj and L. Nachman, 2013. A hybrid approach with collaborative filtering for recommender systems. Proceedings of the 9th International Wireless Communications and Mobile Computing Conference, July 1-5, 2013, Sardinia, Italy, pp: 349-354.

Fenza, G., V. Loia and S. Senatore, 2011. Agent-assisted tagging aimed at folkonomy-based information retrieval. Proceedings of the IEEE Symposium on Intelligent Agent, April 11-15, 2011, Paris, France, pp: 1-8.

Hui, S., L. Pengyu and Z. Kai, 2011. Improving item-based collaborative filtering recommendation system with tag. Proceedings of the 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce, August 8-10, 2011, Deng Leng, China, pp: 2142-2145.

Jelassi, M.N., S. Ben Yahia and E.M. Nguifo, 2013. A personalized recommender system based on users' information in folksonomies. Proceedings of the 22nd International Conference on World Wide Web Companion, May 13-17, 2013, Rio de Janeiro, Brazil, pp: 1215-1224.

Peng, J. and D. Zeng, 2010. Tag-based smoothing for item recommendation. Proceedings of the IEEE International Conference on Service Operations and Logistics and Informatics, July 15-17, 2010, Qingdao, Shandong, pp: 452-456.

Sarwar, B.M., G. Karypis, J.A. Konstan and J. Riedl, 2002. Recommender systems for large-scale e-Commerce: Scalable neighborhood formation using clustering. Proceedings of the 5th International Conference on Computer and Information Technology, December 27-28, Dhaka, Bangladesh, pp: 1-6.

Semeraro, G., P. Lops, M. De Gemmis, C. Musto and F. Narducci, 2012. A folksonomy-based recommender system for personalized access to digital artworks. J. Comput. Cult. Heritage, Vol. 5, 10.1145/2362402.2362405.

Xu, D., Z. Wang, Y. Zhang and P. Zong, 2012. A collaborative tag recommendation based on user profile. Proceedings of the 4th International Conference on Intelligent Human-Machine Systems and Cybernetics, Volume 1, August 26-27, 2012, Nanchang, Jiangxi, pp: 331-334.