

ENHANCED HYBRID RECOMMENDER SYSTEM USING SOCIAL FRIEND NETWORK

Shalmali A. Patil, Reena Pagare
Student, CSE Department, MITCOE, Pune
shalupatil15@gmail.com
Assistant Professor, CSE Department, MITCOE, Pune
reenapwp5@gmail.com

ABSTRACT

Lots of people employ recommender systems to diminish the information overload over the internet. This leads the user in a personalized manner to hit upon interesting or helpful objects in a huge space of possible options. Amongst different techniques, Collaborative filtering recommender system has pulled off great success. But this technique pays no heed towards the social relationship of the users. This problem gave birth to the Social recommender system technology which possesses the capability to recognize user's likings and preferences and their social relationships. In this paper, we present novel method where we combine collaborative filtering recommender system with social friend network to use social relationships. For this, we have made use of data related to users which provides their interests as well as their social relationship. Our method helps to find the friends with dissimilar tastes and determine the close friends amongst direct friends of targeted user which has more similar tastes. This proposed approach resulted in more precise and realistic results than traditional system.

Keywords

Collaborative Filtering; Recommender System; Social Networks.



Council for Innovative Research

Peer Review Research Publishing System

JOURNAL: INTERNATIONAL JOURNAL OF MANAGEMENT & INFORMATION TECHNOLOGY

Vol. 10, No 4

editorsijmit@gmail.com

www.ijmit.com



INTRODUCTION

In last few decades, e-Commerce has expanded in a large degree. So, Recommender Systems turn out to be a great way of harvesting and supplying customers to the business which helps for boosting the sell in the market. Recommender system is the obligatory technique in the field of information retrieval. Thus, websites and commerce sites such as Amazon.com, Yahoo! Music, Ebay.com, and Douban.com begin to use different types of recommender systems together with various recommendation techniques [4]. Amongst those various recommendation techniques, collaborative filtering (CF) technique is the most incredible technique which has seen with a remarkable success. It has been studied, improved and spread in research domain as well as in industry. It estimates the predictions based on the similarity calculation between different users. Initially, it finds out the most similar/nearest users (as neighbors) to the target user based on their earlier preferences. And then, it produces recommendation with the information supplied by selected neighbors.

But, traditional collaborative filtering is not much effective to generate recommendation, because it simply thinks about the similarity of taste between different users available in user-item dataset. These users can be the strangers in real life. Traditional RS has the drawback that it cannot distinguish between strangers or friends. Also, it still suffers from different challenges such as sparsity, scalability and cold start. Hence, it is believed that accuracy of recommender system can be effectively improved if social relationships are taken in to consideration. In real life also, it is observed that people often ask for advices to their friends or acquaintances for different types of decisions. So as to implement this concept in recommender system, Social Recommender Systems came into existence. In this era, social networks are tremendously popular. People tend to spend their lots of time over them for surfing, chatting, giving their opinions on different issues, etc. So it became the large source for collecting information about users. But, on social recommender system, it is believed that the average taste of user's friends is more similar to that user. But, this approach becomes insensible when a user has number of friends with diverse taste.

Hence, in this study, we developed a hybrid approach to provide effective recommendations. So as to achieve this aim, we collected data about preference ratings of users and social network relations from the social network. Then, close friends of a target user are found out from the social network whose preferences are then increased by using mapping function. As well as, top nearest neighbors are found out from traditional collaborative filtering technique. At last, those ratings of close friends and neighbors are combined to predict the recommendation for targeted user. The performance of this approach is evaluated in terms of Mean Absolute Error and Root Mean Square Error by comparing it with traditional and social recommender system.

The rest of the paper is organized as follows. Section II provides an overview about Recommender systems and social networks. Section III describes the related research about different techniques or approached implemented in relevant fields. Section IV investigates our hybrid CF approach. Experimental evaluation is in Section V and conclusion and future work are in Section VI.

1.INTRODUCTION TO RECOMMENDER SYSTEM AND SOCIAL NETWORKS

A. RECOMMENDER SYSTEM

In the earlier 1990's [18], utilization of internet speedily spread. Due to the significant enlargement of online information, it turned out to be a constant challenge to facilitate Internet users to deal with the consequent information burdening. That's why, Recommendation Systems came into existence. It helps users for searching out the interesting information or items amongst a giant dataset based on a query of a user. In addition to that, providing personalized item recommendations is the main purpose of implementing the recommender systems. It takes advantage of the users' likings, interests and hobbies along with their online behavior. By using this information, recommender systems discover the user-item relationship, which helps the customers to come across items, products or information which they may like. For all these progression, users' past interaction history is taken into consideration.

From various users, huge amount of data is gathered over time and then various techniques are applies to generate the recommendations. The way of utilizing this user profile information divides the recommender system into three different categories:

- Content-based (CB): The recommender system will make use of information filtering for generating
 recommendations. For this, user's profile is studied and based on that, the recommender will recommend different
 items whose content is akin to those of the profile.
- Collaborative Filtering (CF): The recommender system will recommend items based on active user's neighbors by considering similarities or connections between them.
- · Hybrid approach: It fuses CB and CF [5].



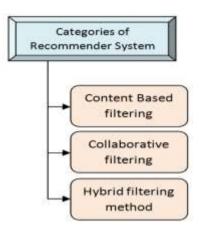


Fig 1. Different Categories of Recommender System

Collaborative filtering is one of the most flourishing technologies in recommender systems. Many well known e-Commerce websites and other developing areas on the internet such as digital library use this technology to improve quality and efficiency of their services. Amongst different existing techniques of collaborative filtering, correlation and supervised learning are found to be most popular techniques [19]. Correlation-based techniques are reasonably accurate. They use Pearson correlation and Cosine similarity to determine similarity for finding out a neighborhood of similar taste for every user and then weighted average of ratings of the neighbors is used to predict the rating for target user for a specific item or product.

Traditional CF can function in two different methods which are user-based and item-based CF. In user-based CF approach, the ratings of users who have like-minded nature or similar taste with the target user are found out and then those ratings are used to generate recommendation for the target user. On the other side, in item-based CF approach, item-item matrix is used to determine the similarity between items based on their purchase behavior and then ratings related to those items are used to generate recommendation for the target user [7].

However, classification of collaborative filtering System can be done in two sub-categories as memory-based CF and model-based CF. In memory-based approach, the whole collection of previous rated items by a user is kept as raw data and processed within memory. These ratings are then used to predict recommendations. In contrast, model-based approaches gain knowledge of a model from user-item database and apply this model for making prediction. This method is highly accurate; but expensive in nature as models need to be created [1].

While studying different challenges in recommender system, cold start, scalability and sparsity are addressed as major challenges. When accurate recommendations are being generated and it is found that the ratings for the target user are very few then appropriate recommendation cannot be generated due to insufficient data to be learnt by the system. This problem is acknowledged as cold start problem. On the other hand, rise in the number of ratings or enormous growth in number of users and items give birth to scalability problem. This increases the complexity of the system. Hence, system becomes unsuccessful to predict suitable results in small amount of time [4].

B. SOCIAL NETWORKS

Reputation of social network such as Facebook, Orkut and MySpace has enormously increased in the recent era. The social networks sites are web-based services. These let users to interact with each other and allow them to start relationship in between them. People generally use the platform of social networks to communicate with the other users whom they have real life connection. People can share text in addition to multimedia content such as images, songs and videos on social network. Nowadays, it is common that a user has a hundreds of connections or relationships with the users who share similar interests and activities. On social network, any kind of individual is called as 'actor' who actively participates in social interactions. Social network is a platform for them to chat, shop, share knowledge and information or grow their business by advertising for endorsements. Based on the similarity of taste, one user responds to another user positively or negatively. Also, it is feasible to allow users to organize his or her connections in groups, which helps them to share any type of content to entire group [9].

However, different applications creates different types of social environment named as Friendly environment, Business environment, General environment and Research environment [1],[8].

2. RELATED RESEARCH

In this section, we analyze quite a few prime techniques for recommender systems, together with traditional recommender systems which are mainly based on collaborative filtering methods and social recommender systems which have gained lots of attention in recent times.

Collaborative filtering recommender systems have studied by applying Different approaches to solve scalability problems. Siavash and Ali in [11] have proposed a hybrid recommender system. This is a composition of user-based collaborative filtering and density-based user clustering by considering demographic information of user. SongJie Gong in [7] has put forward another hybrid approach to combine user clustering and item clustering to make it more scalable and accurate.



Jing Jiang, Guangquan Zhang, Jie Lu and Guodong Long have proposed to implement the Map-reduce framework in collaborative recommender system to enable parallel computations as well as to exploit data locality in [12]. Xingyuan in paper [13] has applied clustering technique over collaborative filtering to separate the user and strived to get better performance of collaborative filtering.

Excluding clustering methods, a diverse hybrid approach is projected by Farman, Ghulam and Sung in [14]. This system uses the temporal and demographic information about the user along with the traditional recommender system. This system has resulted in an accurate, scalable and quick system. Alper Bilge and Huseyin Polat have projected a new system named as content based profiling in [15] to reduce the dimension space for similarity calculation burden by using users purchase and rating behavior. Fatih, Faruk, Mouzhi, and Dietmar in [16] have proposed new scheme named RF-REC to get better efficiency, scalability and ease.

Over these various researches, when a literature analysis is done in [17], it concluded that the hybrid recommender system which takes the demographic and temporal information of user into consideration, gives the good quality of result. Nonetheless, traditional and trust-aware recommender systems face different downsides. No any traditional recommender system can distinguish between friends and neighbors. However, trust relationships have proved that they are different from social relationship. So consideration of trust relationships for finding taste similarity between different people provides imperfect results.

The literature survey has provided new way approach. Hence, we put forward a novel approach to use social friend network information for enhancing collaborative filtering in social recommender system.

3. APPROACH

Our proposed technique is the hybridization of recommender system with social network information about user relationships with social regularization.

A. PROBLEM DESCRIPTION

When scenario of recommendation is scrutinized in the actual world, then it provides the view of social network as shown in the above Fig. 2. However, user-item matrix is shown in Fig. 3.

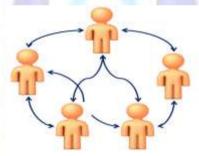


Fig 2. Social Network

User/Item	I_1	I_2	I_3	I_4	I ₅
U_1	3		2	1	-
U_2		4	/		2
U_3		3	1	5	
U_4	1		4	3	
U_5		5	4		3

Fig 3. Social Network

In Fig. 2, 5 different users were shown who share a bidirectional social relationship with each other. The edges show the connection or relationship between them. Bidirectional relationship explicitly states that if ui is uf 's friend then uf must be ui's friend. As demonstrated in Fig. 3, each user has provided his ratings over different items on a 5-point integer scale. The scale is used to derive the degree of liking of that user to that specific item. So it ranges from 1 to 5 for 'hate' to 'most-liked' opinion of user, respectively.

The aim of our study is to predict the missing values to calculate the accuracy of our proposed system by making use of social friend network with recommender system.



B. ALGORITHM

The simple flow of the algorithm is shown in figure 4.

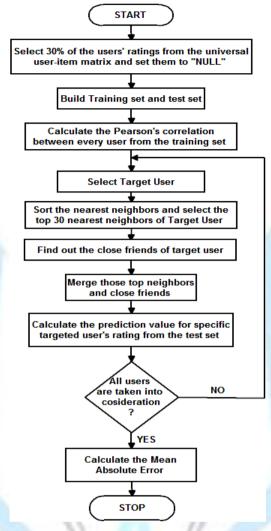


Fig 4. Proposed Algorithm

To conduct the experiments, we have used the two dataset which are co-related to each other. One is Universal user-item matrix database which is similar as traditional user-item matrix database. Another is social friend network database, in which social relationships of different user is provided. At the preprocessing phase, 30% users' ratings are to be selected, copied to another file and nullified from the Universal user-item matrix. So, newly generated set is utilized as training set. And further set, in which 30% selected users' ratings are copied is used as a test set. Now, every user whose rating is nullified in training set is taken into consideration and Pearson's correlation between every other user is calculated. The formula for Pearson's correlation is as follows, in (1).

$$Sim(i,f) = \frac{\sum_{j \in I(i) \cap I(f)} (R_{ij} - \overline{R_i}) \cdot (R_{fj} - \overline{R_f})}{\sqrt{\sum_{j \in I(i) \cap I(f)} (R_{ij} - \overline{R_i})^2} \cdot \sqrt{\sum_{j \in I(i) \cap I(f)} (R_{fj} - \overline{R_f})^2}}$$
(1)

Where, m x n rating matrix R expresses terms as m users providing ratings over n items. The similarity function Sim(i, f) provides similarity between user u_i and user u_f . This calculation provides the neighbors for those users [3].

In the next phase, those neighbors get sorted in descending manner; hence the top 30 nearest neighbors get selected from them. In parallel, for every targeted user, its level-1 friend (direct friend) is found out from social friend network database. But, as per our proposed approach, it is beneficial to find the close friends over the direct friends of the targeted user. So social regularization is applied, this amplifies the preferences of the close friends of target user and lessens the preferences of the remaining friends.

The mapping function is used to bind the similarities into [0, 1]. The formula is shown in (2).

$$f(x) = (x+1)/2 (2)$$



Where, x is similarity $Sim(i,f) \in [0,1]$ $u_i's$ direct friends. Above term is used to increase or decrease the preferences of the users as per their similarity. If similarity value is found to be greater, then the user u_f will put in additional for finding average taste of user u_i . And if similarity value is found to be less, then the user u_f will contribute a smaller amount for finding average taste of user u_i [3].

Using this calculation, close friends whose preferences are amplified are selected for further process. As a final point, in third phase, those top 30 neighbors in addition to friends of target user is merged and by making use of weighted average method, prediction of rating is generated for the targeted user. Following formula for weighted average method in (3) is used to find out the prediction:

$$P_{ij} = \overline{U}_i + \frac{\sum_{a \in Raters} (U_{a,j} - \overline{U}_a) \operatorname{Sim}(a,i)}{\sum_{a \in Raters} |\operatorname{Sim}(a,i)|}$$
(3)

Where, P_{ij} is the prediction score. It is calculated by using user i's average rating score on other items with the weighted average of all the ratings for the item j [2].

This whole process is repeatedly done for each user, until it predicts the ratings for all the 30% nullified places in training set. Those predicted ratings get compared over real ratings from the test set, to calculate average Mean Absolute Error (MAE).

4. EXPERIMENT

A. DATASET

To perform proposed work, we needed the dataset which contains traditional user-item matrix along with the social relationships present between different users in the dataset. So we chose Douban dataset. On 6th March, 2005, this website got launched as a Chinese Web 2.0. This website is

one of the largest online communities. It supplies biggest online dataset for music, movies and books in china. It offers reviews, ratings of users as well as recommendation services in support of music, movies and books. For these, users are capable of providing 1 to 5 ratings. Also, it supplies social network service as Facebook provides. It gives the chance to users to connect with his/her friends on Douban through their Email account. So, Douban has turn out to be ideal dataset for our research work [2].

Here, user can associate with different groups of interest. When we crawled into these different groups, we found that it consists of 129,490 unique users. Also, 58,541 unique movie items are present there. These collectively provide total 16,830,839 number of movie ratings by the different users. We found 1,692,952 claimed social relationships in the social friend network. Douban dataset consists of two files. One is "uir.index" file in which user-item ratings are given. And other is "social.index" file in which user social friend network with different user to user relationships are provided.

Format of "uir.index" file:

User Id Item Id Rating

Format of "social.index" file:

User_Id_1 User_Id_2

Douban only supports friendship mechanism as like Facebook. Hence, "social.index" gives the mutual relationship information. This indirectly says that if relationship of '123 456' is available then relationship of '456 123'(reverse order) should also be available. But due to some crawling issues, some relationships are absent from the given dataset. So if it is found that reverse relationship is absent at any point of time then it can be directly added.

B. EVALUATION METRIC AND VALIDATION

We use Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) is to measure the prediction accuracy of our proposed approach in comparison with other traditional collaborative filtering algorithm considering neighbors and collaborative filtering algorithm considering friends methods. The system with lower MAE or RMSE is known as the system with better accuracy. The metrics MAE is defined in (4):

$$MAE = \frac{\sum_{i=1}^{N} \mid p_{i-}r_i \mid}{N}$$
 (4)

The metrics RMSE is defined in (5):

RMSE =
$$\sqrt{\frac{\sum_{i=1}^{N} (p_{i-}r_{i})^{2}}{N}}$$
 (5)

Where,

N: total number of predictions

p_i and r_i: predicted rating and actual rating, respectively [2].



C. EXPERIMENTAL RESULTS

In this section, in order to demonstrate the performance enhancement of our proposed approach, we evaluate our method with the following approaches:

- [1] Traditional Recommender System: this method is uses traditional collaborative filtering method.
- [2] Social Recommender System: In this is the method, preferences of all direct friends of target user are taken into consideration
- [3] Hybrid Recommender System: this is our proposed method.

To determine the effectiveness of the proposed approach, we carried out the experiment where we nullified 20%, 25% and 30% ratings of respective 200, 300 and 800 users from the dataset. For each user whose ratings are nullified, we ran our experiment using our proposed algorithm, traditional collaborative filtering algorithm considering neighbors and collaborative filtering algorithm considering friends.

Table 1. Performance Comparison (A Smaller MAE or RMSE Value means a High Accuracy)

		200 Users			300 Users			800 Users		
Training Data	Metrics	Tradition al RS	Social RS	Hybrid RS	Tradition al RS	Social RS	Hybrid RS	Tradition al RS	Social RS	Hybrid RS
	MAE	0.86251 1715	0.66849 9597	0.58516 1162	0.88596 942	0.68374 3873	0.58513 3493	0.88711 4845	0.68995 6278	0.57247 6273
80%	RMSE	1.70861 576	1.62162 2852	1.46271 8829	1.70521 0172	1.63953 7131	1.47516 0091	1.65858 6718	1.64740 6464	1.48647 78
	MAE	0.89130 4348	0.68158 1629	0.59277 0237	0.87476 5988	0.68721 2694	0.58135 3417	0.88769 1684	0.69136 7368	0.59196 9275
75%	RMSE	1.76736 775	1.63815 2097	1.47401 0163	1.65477 1656	1.63509 3459	1.48412 4931	1.67634 2555	1.64855 4873	1.47652 965
	MAE	0.76503 8923	0.60227 6366	0.58580 0179	0.77943 3631	0.69332 3867	0.60048 0409	0.71908 2834	0.69994 1343	0.59537 7478
70%	RMSE	1.73198 9649	1.65547 8584	1.36366 6642	1.71194 715	1.64967 1702	1.47516 0091	1.74783 7639	1.65942 4296	1.46208 614

For those experiments, results are shown in Fig. 5 and Fig. 6 when 200 users were taken into consideration. We examined that the prediction accuracy increased.

However, table (1) provides the detailed statistics for the experiments which we have accomplished. These results show that the proposed approach possesses the lower values for both MAE and RMSE matrices. Hence, hybrid recommender system achieves better performance than traditional and social Recommender System.

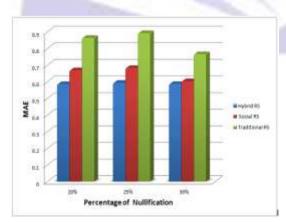


Fig 5. MAE - Performance analysis of each Recommender system with different nullification percentage

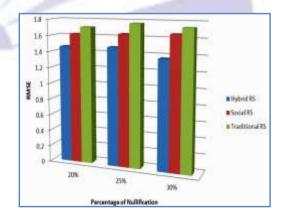


Fig 6. RMSE - Performance analysis of each Recommender system with different nullification percentage

Implementation of the proposed approach on the Douban dataset needed to compare with the recommendation results of other techniques. Hence, we totally implemented 3 different approaches. Initially, we applied traditional recommendation



algorithm considering neighbors on the same dataset. For that, we found average Mean Absolute Error as 0.7694. Then, we applied collaborative filtering algorithm considering friends (level-1/direct friends from social network relationship database), which produced the average Mean Absolute Error as 0.6948. However, when we ran our proposed approach, it outperformed other two experimental results. The average Mean Absolute Error of proposed approach is 0.5894. As the projected method shows the result with the lower MAE, it concludes that the method is more accurate than other two methods.

5. CONCLUSION

We proposed an approach where social recommender system embeds social regularization. In the experimentation, we consider the problem of predicting more accurate recommendations. Results of the experiments proved that the proposed approach raises the accuracy of the predicted values improving the quality of the recommendation system. We proved that, if we consider the close friends of the user from social networking information and along with that, we consider the top nearest neighbors of the same user from user-item matrix, and merge them to generate the recommendations then it offers better results minimizing MAE and RMSE. The average MAE of traditional RS, social RS and hybrid RS is 0.839, 0.677 and 0.587, respectively. Also, average RMSE of traditional RS, social RS and hybrid RS is 1.706, 1.643 and 1.462, respectively. Through the experimental study we can see that our algorithm gives the good performance. And thus, our proposed system achieves 29.9 % and 13.2 % more accuracy over traditional and social recommender system, correspondingly.

ACKNOWLEDGMENT

Author thanks Prof. Reena Pagare for her valuable discussions and guidance in information retrieval and collaborative filtering related research work as well as for reading the numerous revisions of this paper.

REFERENCES

- [1] Reena Pagare, Shalmali A. Patil, "Social Recommender System by embedding Social Regularization" Advance Computing Conference (IACC), February 2014, pp. 471 476.
- [2] F. Liu and H. J. Lee., "Use of social network information to enhance collaborative filtering performance", Expert Syst. Appl., 37(7):4772–4778, 2010.
- [3] Hao Ma, Dengyong Zhou, Chao Liu, Michael R. Lyu, Irwin King, "Recommender Systems with Social Regularization", WSDM'11, February 9–12, 2011, Hong Kong, China.
- [4] Cosimo Birtolo, Davide Ronca, and Roberto Armenise, "Improving accuracy of recommendation system by means of Item-based Fuzzy Clustering Collaborative Filtering", 11th International Conference on Intelligent Systems Design and Applications(ISDA), IEEE 2011, pp. 100 - 106.
- [5] Mustansar Ali Ghazanfar and Adam Prugel-Bennett, "A Scalable, Accurate Hybrid Recommender System", Third International Conference on Knowledge Discovery and Data Mining, IEEE 2010, pp. 94 98.
- [6] Wanyu Deng, Qinghua Zheng and Lin Chen, "Real-Time Collaborative Filtering Using Extreme Learning Machine", International Conference on Web Intelligence and Intelligent Agent Technology – Workshops, IEEE, September 2009,pp. 466 – 473.
- [7] SongJie Gong, "A Collaborative Filtering Recommendation Algorithm Based on User Clustering and Item Clustering", Journal of Software, July 2010, Vol. 5, No. 7
- [8] Fatemeh Khoshnood, Mehregan Mahdavi and Maedeh Kianisarkaleh, "Designing a Recommender System Based on Social Networks and Location Based Services", International Journal of Managing Information Technology (IJMIT) Vol.4, No.4, November, 2012
- [9] Yu Qiao, "Social networks and E-commerce", TKK T-110.5190 Seminar on Internetworking, 2008-04-28/29
- [10] Hao Ma, Irwin King and Michael R. Lyu, "Learning to Recommend with Social Trust Ensemble", international ACM SIGIR conference on Research and development in information retrieval, July 2009, pp. 203-210
- [11] Siavash Ghodsi Moghaddam, Ali Selamat, "A Scalable Collaborative Recommender Algorithm Based on User Density- Based Clustering", 3rd international conference on Data Mining and Intelligent Information Technology Applications (ICMiA), IEEE 2011, pp. 246-249
- [12] Jing Jiang, Jie Lu, Guangquan Zhang, Guodong Long, "Scaling-up Item-based Collaborative Filtering Recommendation Algorithm based on Hadoop", IEEE World Congress on SERVICES, 2011, pp. 490-497
- [13] Xingyuan Li, "Collaborative Filtering Recommendation Algorithm Based on Cluster", International Conference on Business Computing and Global Information (BCGIN), IEEE 2011, pp. 645-648
- [14] Farman Ullah, Ghulam Sarwar, Sung Chang Lee, "Hybrid Recommender System with Temporal Information", (ICOIN), International Conference on Information Networking, IEEE 2012, pp. 421-425
- [15] Alper Bilge and Huseyin Polat, "An Improved Profile-based CF Scheme with Privacy", Fifth IEEE in International Conference on Semantic Computing, IEEE 2011, pp. 133-140

ISSN 2278-5612



- [16] Fatih Gedikli, Faruk Bagdat, Mouzhi Ge and Dietmar Jannach, "RF-REC: Fast and Accurate Computation of Recommendations based on Rating Frequencies", IEEE 13th Conference on Commerce and Enterprise Computing (CEC), 2011, pp. 50-57
- [17] Reena Pagare, Shalmali A. Patil, "Study of Collaborative Filtering Recommendation Algorithm Scalability Issue", International Journal of Computer Applications (0975 8887), Volume 67 No. 25, April 2013
- [18] Joseph A. Konstan · John Riedl, "Recommender systems: from algorithms to user experience", User Modeling and User-Adapted Interaction, 2012, Vol. 22, pp. 101-123
- [19] Wanyu Deng, Qinghua Zheng and Lin Chen, "Real-Time Collaborative Filtering Using Extreme Learning Machine", IEEE/ WIC/ ACM International Conference on Web Intelligence and Intelligent Agent Technologies, Volume 1, pp 466-473, 2009

