Collaborative Book Recommendation System using Trust based Social Network and Association Rule Mining

Anand Shanker Tewari
CSE Department
NIT Patna
Patna, India
e-mail: anand@nitp.ac.in

Asim Gopal Barman
ME Department
NIT Patna
Patna, India
e-mail: asim@nitp.ac.in

***Abstract*—Almost all e-commerce websites generates recommendations for their users but most of them are irrelevant. Collaborative filtering is one of the most widely used recommendation generation technique by e-commerce websites. Collaborative filtering generates recommendations for the target user from the collaboration of the other users who have the similar interest derived from their ratings. With the advent of the Web 2.0, web based social networks have become one of the major source of information. Users can now make friends, share thoughts, images etc. on the Internet and express different level of trust on their web friends. Recommendations generated by the trusted friends are more relevant than other users. This paper propose a book recommendation system that generates recommendations from the collaboration of trusted friends of the target user and uses association rule mining to capture current reading trend of users in the network.**

Keywords— Recommendation system; Trust; Collaborative filtering, Association rule mining.

I. INTRODUCTION

With the expansion of the Internet and other related technologies, massive amount of information is available. Recommendation system recommends appropriate information to the user depending on his need. Recommendation system has vital role in increasing the revenue of the E–commerce websites. Collaborative filtering is among the most successful and oldest recommendation technology [1-4]. Amazon has started its book recommendation using collaborative filtering. Collaborative filtering recommends product based on the ratings of the other similar users in the database. With the arrival of Web 2.0 many options are available to the users than just retrieving the data. In web based social networks users can create his profile, make friends and express opinions about anything. Recommender systems are now taking the advantage of this user networking for recommending products to the target user. People prefer to take recommendations from their friends [5]. With the help of trust mechanism users can assign different level of trust on their friends. Information about level of trust of one user on other user plays very important role on generating relevant recommendations. In this paper we have proposed a collaborative Trust based Book Recommendation System (TBRS) that generates the book recommendations to the target user using his friends on the trust network. TBRS uses target user's friends as potential recommenders for collaborative filtering and association rule mining for finding current book trend in the market.

The paper is further organized as follows, section 2 discusses collaborative filtering, section 3 throws light on association rule mining, section 4 defines trust in social network, section 5 explains TBRS and section 6 shows practical implementation and evaluation of the TBRS with other benchmark methods.

II. COLLABORATIVE FILTERING

The term collaborative filtering was first introduced by Goldberg et al.[6], they created a system for filtering email named as Tapestry. Collaborative filtering also called as social filtering is one of the ways to generate recommendations [3]-[4]. Collaborative filtering generates recommendations for the target user in collaboration of the other users in the database. User based Collaborative filtering functions by collecting the ratings of the target user in a given domain and matching it with other users in the same domain and returns useful personalized recommendation for the target user [7]. Similarity between the two users can be calculated using Pearson's correlation formula as shown in equation 1.

$$Sim_{\mathrm{m,n}}=\frac{\sum_{a\in I_{mn}}(R_{m,a}-A_m)(R_{n,a}-A_n)}{\sqrt{\sum_{a\in I_{mn}}(R_{m,a}-A_m)^2\sum_{a\in Imn}(R_{n,a}-A_n)^2}} \quad (1)$$

Where $R_{m,a}$ is the rating of the item a by user m, A_m is the average rating of user m, and I_{mn} is the item set rated by both user m and user n. Next select all the neighbours of the target user whose similarity is above certain threshold as potential recommenders. Then the rating P_{xk} of the target user x to the target item k is calculated using equation 2.

$$P_{xk}= A_x + \frac{\sum_{m=1}^{c}(R_{m,k}-A_m)^* Sim(x,m)}{\sum_{m=1}^{c} Sim(x,m)} \quad (2)$$

Where A_x the average rating of the target user x, R_{mk} is the rating of the neighboring user m to the target item k, A_m is the

978-1-5090-5256-1/16/$31.00 ©2016 IEEE

Authorized licensed use limited to: Seoul National Univ of Science & Tech (SNUT). Downloaded on September 03,2020 at 02:13:14 UTC from IEEE Xplore. Restrictions apply.

average rating of user *m*, *Sim(x,m)* is the similarity between the target user *x* and the neighbouring user *m* and *c* is the total number of the neighbours of the target user.

III. ASSOCITAION RULE MINING

Association rule mining (*ARM*) discovers correlation among large number of items [8]. *ARM* analyses customer shopping cart to guess occurrence of items. Let $I = \{i_1, i_2, i_3 \ldots\ldots, i_m\}$ is a collection of items or itemset. An association rule can be written as $A \rightarrow B$, where $A \subseteq I$, $B \subseteq I$ and $A \cap B = \varnothing$ [9]. Association rule evaluation metrics are as follows.

- *Support:* It shows number of transactions in the database that contains both *A* and *B*.

$$Support(A \cup B) = P(A \cup B) \quad (3)$$

- *Confidence:* It tells how frequently item *B* occurs in a transaction, if *A* has already bought.

$$Confidence\ (A \rightarrow B) = P\ (B \mid A) \quad (4)$$

The main idea behind *ARM* is to find rules based on criteria given below.

a) Generate all item sets having *support* factor bigger than or equal to the domain expert set minimum support.

b) Generate all the rules having the *confidence* factor bigger than or equal to the domain expert set minimum confidence.

IV. TRUST IN SOCIAL NETWORK

Personal trust means one person trusts another person, persons or thing(s) in the specific domain or situation. Interpersonal trust implies that two or more people trust each other in certain domain [10]. Web 2.0 has provided a platform for social networking websites with millions of users, where users can share thoughts, images, make new friends. Some of the social networking websites have offered a facility to the users to express different kind of trust on different friends. Every person prefers to get recommendation from friends rather than strangers. The trust of one person on another varies from person to person.

A. Web of Trust

Web based social networks provides freedom to users to create their own contents, express their likes and dislikes about certain items and/or persons. Many websites allows users to express their degree of trust on another user. *Fig. 1* shows the web of trust where users are expressing their trust on another users in the form of ratings. In this network users

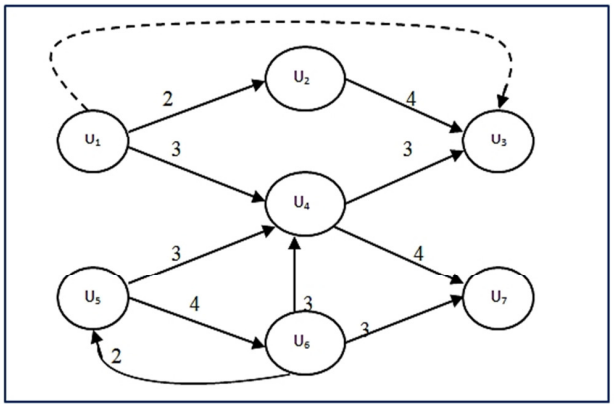

Fig. 1. Trust Network with trust values

are represented by nodes and trust between the two users are shown by edges along with their trust values [11].

B. Trust Propagation

Recommender system usually needs users to recommend products to the target user. In web based social networks user is connected to only few number of users and expresses his trust values on those users. Trust propagation finds the trust between the two users that are not directly connected to each other. As shown in figure 1 trust propagation can calculate the trust of user U_1 on user U_3 represented by the dotted line [5].

C. Local and Global Trust Metrics

Trust metrics are used to compute the trust of the target user on unknown users on the basis of complete social network. Trust metrics are of two types. i.e. local and global [11]. *Local* trust metrics considers personal view of the target user trust on other users and predicts different value of trust for different users. *Global* trust metrics predicts the reputation of the target user in the social network. It does not consider subjective trust of each user on other users but average it across standardized global values. PageRank uses global metrics [12].

D. Tidal Trust

Tidal Trust computes the local trust. It works as modified breadth first search, proposed by J.A. Golbeck [13]. It calculates the inferred trust between the two users that are not directly connected to each other. i.e. U_i and U_s. U_i is called as *source* and U_s is called as *sink* as shown in *Fig. 2*. The inferred trust of U_i on U_s *i.e.* $t_{i,s}$ is a weighted average of source's neighbours' rating of the sink as shown in formula 5 [14]. Where t_{js} represent the trust rating of U_j on U_s. As soon as the first path is found between *source* and *sink* the depth limit of the graph is set and search will continue to find any other path at same depth to the *sink*. Next the trust threshold (*max*) is

Authorized licensed use limited to: Seoul National Univ of Science & Tech (SNUT). Downloaded on September 03,2020 at 02:13:14 UTC from IEEE Xplore. Restrictions apply.

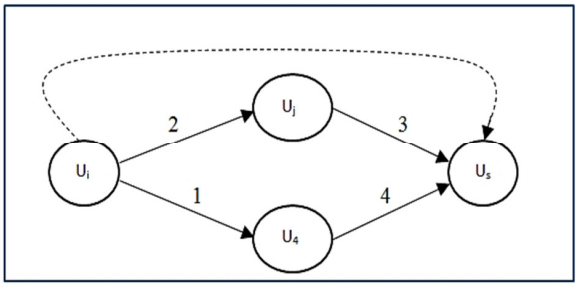

Fig. 2. An illustration of inferred trust inTrust Network

calculated as maximum of all the strengths of all the trust paths leading to the *Sink.* In *Fig. 2* trust network value of *max* is 2.

$$t_{is} = \frac{\sum_{j \in \text{adj(i)},\, t_{ij} \geq max} t_{ij} t_{js}}{\sum_{j \in \text{adj(i)},\, t_{ij} \geq max} t_{ij}} \quad (5)$$

V. TRUST BASED BOOK RECOMMENDATION SYSTEM

The sole purpose of the trust based book recommendation system is to recommend books to the target user that is useful and according to his likes. It uses the combine features of collaborative filtering, association rule mining and user's trust on its neighbours. TBRS uses properties of web based social networks to find target user friends. The system has following steps:

1. The system will create the profile of each user and stores users past bought books and their ratings.

2. Given a trust based social network as shown in *Fig. 3*, *TBRS* finds all the directly connected friends of the target user (U_t) and stores it in the list (L_1). *i.e.* now $L_1=\{U_1, U_2, U_3\}$.

3. Next the system finds the friends of friends of the target user and stores it in the another list *(L_2)*. So L_2 stores *{U_4, U_2, U_5, U_6}* as shown in *Fig. 3*.

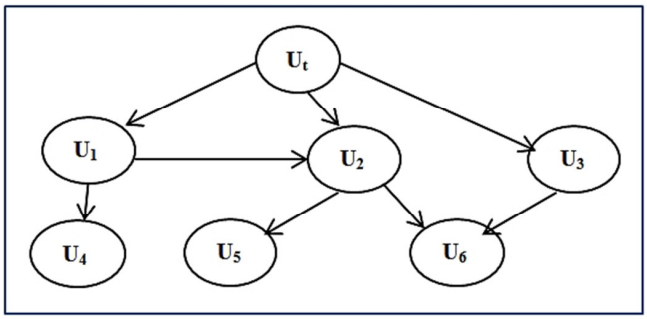

Fig. 3. Trust Network showing target user and his friends

4. Then *TBRS* removes all those users in the second list (L_2) that are already present in the first list (L_1). Now L_2 has only three users. *i.e.*$\{U_4, U_5, U_6\}$.

5. Next it calculates the inferred trust of target user on all the users listed in L_2 using Tidal trust.

6. *TBRS* now creates the pool of friends using L_1 and L_2. Then it removes all those users from the pool whose direct or inferred trust ratings are below threshold, as defined by the domain expert.

7. Now the system has most trusted friends of the target user, those may act as potential recommenders of the target user.

8. Next the system computes the similarity between the target user and his trusted friends found at *step 7* using equation 1.

9. Apply association rule generation algorithm on the books transaction database. Adjust the *support* and *confidence* parameters to generate stronger rules.

10. Afterwards *TBRS* applies generated association rules on the order history of the target user and stores the books that he may buy.

11. Now *TBRS* calculates the collaborative ratings of all the books found at step *10* using equation *2* and similarity calculated among trusted friends at step *8*.

12. Finally *TBRS* generates the recommendations for the target user in decreasing order of books ratings calculated at step 11. The block diagram of *TBRS* is shown in *Fig. 4*.

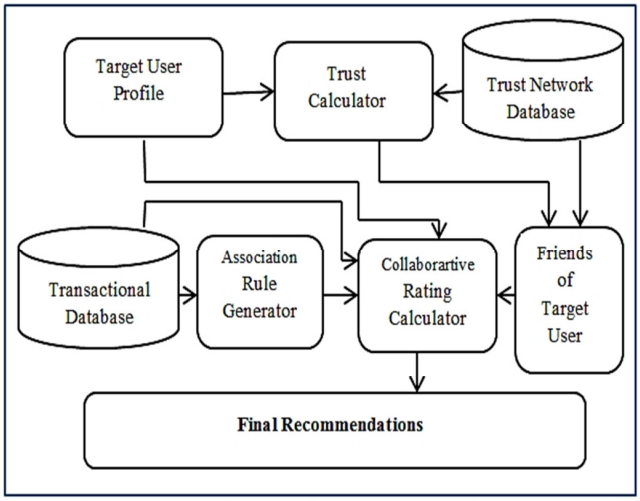

Fig. 4. Block diagram of trust based book recommendation system

Authorized licensed use limited to: Seoul National Univ of Science & Tech (SNUT). Downloaded on September 03,2020 at 02:13:14 UTC from IEEE Xplore. Restrictions apply.

VI. PRCATICAL IMPLEMENTATION AND EVALUATION

TBRS is tested over live data and the results are compared with standard benchmark recommendation methods like collaborative filtering. It uses the trust network of students of NIT Patna, India and generated collaborative ratings using Apache Mahout, which is an open source machine learning library written in Java. The complete TBRS runs on Intel Core i5 machine with 4 GB RAM. Association rules are generated for various combination of support and confidence, but the rules generated at 10% support and 80% confidence are more accurate for given data than other combinations. To evaluate the performance of proposed system precision measuring technique is used. Precision is defined as the ratio of total number of relevant recommendations to the total number of recommendations. During every execution of the experiment TBRS performs better than collaborative filtering algorithm as visible in figure 5.

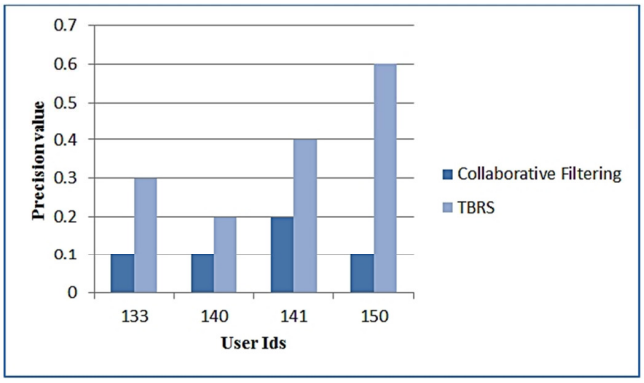

Fig. 5. Comparative precision graph of TBRS with collaborative filtering

VII. CONCLUSION

Almost all people are considering the advice of their friends, whenever they are going to do new purchases. This recommendation system has realized this aspect of the people in the form of trust. Much established recommendation technique collaborative filtering is used to find the similarity between the two users and predicting the ratings of the recommended books. Association rule mining is used to find the current market trends. This recommendations system consults the friends of the target user in the form of trust ratings before making any recommendations. The recommendations generated by this system are more relevant than traditional methods and are of buyer's interest.

REFERENCES

[1] W. Hill, L. Stead, M. Rosenstein and G. Furnas, "Recommending and evaluating choicesin a virtual community of use", Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1995, pp. 194-201.

[2] J. Konstan, B.N. Miller, D. Maltz, , J.L. Herlocker, L.R. Gordon and J. Riedl, "GroupLens: Applying Collaborative Filtering to Usenet News", Communications of the ACM, Vol. 40, No. 3, 1997, pp. 77-87.

[3] P. Resnick, N. Iacovou, M. Suchak, P.Bergstrom, and J.Riedl, "GroupLens: An Open Architecture for Collaborative Filtering of Netnews.", ACM conference on Computer supported cooperative work, 1994, pp. 175-186.

[4] U.Shardanand, and P. Maes, "Social Information Filtering: Algorithms for Automating "word of mouth" ",Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1995, pp. 210-217.

[5] Xu Xiaowei, Wang Fudong, "Trust -Based Collaborative Filtering Algorithm", Fifth International Symposium on Computational Intelligence and Design, 2012, pp. 321-324.

[6] David Goldberg, David, Nichols, Brian M. Oki and Douglas Terry, "Using Collaborative filtering to weave an information tapestry.", Communications of the ACM, Vol. 35, No. 12, 1992, pp. 61-70.

[7] Jonathan L. Herlocker, Joseph A. Konstan, Al Borchers and John Reidl, "An Algoritrhimic framework for performing collaborative filtering", Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 1999, pp. 230-237.

[8] J. Han, M. Kamber, J. Pei, "Data Mining: Concepts and Techniques", The Morgan Kaufmann Series, 2001.

[9] R. Agrawal, T. Imielinski, A. Swami, "Mining association rules between sets of items in large databases", ACM SIGMOD International Conference on Management of Data, Vol. 22, No. 2, 1993, pp. 207-216.

[10] D. H. Knight, N. L. Chervany, " The meaning of trust, Technical Report", WP9604, University of Minnesota, Management Information Systems Research Center, 1996.

[11] P. Massa and P. Avesani, "Trust-aware collaborative filtering for recommender systems", Proceedings of Confederated International Conference, On The Move to Meaningful Internet" CoopIS, DOA, ODBASE. Springer LNCS, Vol. 3290, 2004, pp. 492-508.

[12] L. Page, S. Brin, R. Motwani, T. Winograd, "The PageRank Citation Ranking, Bringing order to the Web", Technical report, Stanford InfoLab, 1999.

[13] Jennifer Ann Golbeck, "Computing and Applying Trust in Web-based Social Networks", Ph.D. Dissertation, University of Maryland, 2005.

[14] J. Golbeck, "Trust and Nuanced Profile Similarity in Online Social Networks", ACM Transactions on the Web (TWEB), Vol. 3, No. 4, 2009, pp. 1-33.

Authorized licensed use limited to: Seoul National Univ of Science & Tech (SNUT). Downloaded on September 03,2020 at 02:13:14 UTC from IEEE Xplore. Restrictions apply.