

Review of Social Collaborative Filtering Recommender System's Methods

¹Pratibha Yadav, ²Kumari Seema Rani, ³Sonia Kumari

^{1,2,3}Shyama Prasad Mukherji College (For Women), University Of Delhi, India

¹pratibhamcadu2011@gmail.com, ²raniseemabca@gmail.com, ³Soniakumari.ducs@gmail.com

ABSTRACT

Recommender Systems plays a vital role in e-commerce. The goal of recommender system is to present the user with the personalized information that matches with the user's interest. Now a days, user's interest is leaning towards social networks. Social Networking Sites provide users a platform to connect and share their information with other users who share similar interests with user. The popularity of social networking sites is increasing day by day. Recommender systems are now using the social information for their analysis and prediction process. Collaborative filtering approach is assumed to be the broadly approved technique of recommender system. Collaborative filtering method recommends an item to a user based on the preferences of other users who share analogous interest with the active user. In this paper, we have presented a study of collaborative filtering based social recommender system.

Keywords

Recommender System, Collaborative Filtering, Social Collaborative Filtering

1. INTRODUCTION

World Wide Web has accelerated the rate of information exchange. Users can easily access the information from internet. With the advancement of technology, the information is also expanding at exponential rate. This in turn has become a problem for users as it is not viable for them to explore the whole pool of information. Users are thus facing problem in finding the individualised information. This in turn arises the need for recommender systems. **Recommender System (RS)** is a software which provides users with customized information they are interested in [1][6][25]. RS has been widely implemented by many applications which includes movies, books, music, jokes, news articles, search queries and many more. RSs have emerged as an excellent recommending tool for e-

commerce sites like Amazon [8], E-bay, CDNOW and many more [4]. RS has become a boon for these sites.

RS is mainly classified in three approaches namely Content Based, Collaborative Filtering and Hybrid. Content based technique analyses users past behaviour and preferences for making recommendations [7][23]. Collaborative Filtering approach uses user's profile for finding neighbours of the active user and then uses the information of these neighbours for prediction. Hybrid approach is the integration of content based and collaborative filtering approaches [5][24]. In our paper, we have focused on Collaborative Filtering approach which is described in the following section.

2. Collaborative Filtering

Collaborative Filtering technique is the most accepted techniques of the RS. **Collaborating Filtering** (CF) makes prediction of the items for a user on the analysis of preferences of other users who share similar profile with the active [2][4][9][11]. In CF RS, data is represented in form of a rating matrix of form $user * item$. Let I be the set of items and U be the set of users in a rating matrix. Let $utility(u, i)$ be the utility function which computes the worth of item $i \in I$ for user $u \in U$. In CF RS, $utility(u, i)$ is based on $utility(u_n, i)$ where, $i \in I$ and $u_n \in U$ is the set of neighbours of active user $u \in U$ which have similar preferences as the active user has. CF is further divided into three approaches which are described as follows:

- **Memory based CF:** Memory based CF uses the whole dataset for analysis and prediction process [12]. It uses various measures like Cosine based, Adjusted cosine based, Pearson correlation, Adjusted Pearson correlation, k-Nearest neighbours and many more. Pearson correlation and cosine based are the commonly implemented memory based techniques.
- **Model based CF:** Model based CF first uses the dataset to learn a model by analysing the dataset information and then uses the learned model for prediction [13]. There are various model learning techniques which includes Probabilistic models, Bayesian classifiers, CF using dimensionality reduction techniques, clustering techniques, and many more. Probabilistic models and Bayesian classifiers are the commonly used model based techniques.
- **Hybrid CF:** Hybrid CF is the combination of model and memory based CF techniques. It alleviates the limitations of model based and memory based approaches.

3. Social Collaborative Filtering based Recommender Technique

Social network has revolutionised the communication process. Now, Users are preferring social sites for information exchange. Majorly active and widely used social networking sites are Google+, Facebook, LinkedIn, Twitter and more. The popularity of the social sites is increasing drastically. With the increase in popularity of social sites, RSs have also embraced social information as an input to the analysis and prediction process. The embracement of social information by RS has given rise to performance enhancement. A social RS has outperformed traditional RS by taking into account social interest and trust between users connected via social network [19][20]. Social trust between two users u and v may be established on the basis of feedback [14][15][16]. Feedback can be classified in two types: explicit and implicit. Explicit feedback includes explicitly asking user u for the feedback related to user v for instance by voting or rating. Implicit feedback may be inferred by mapping user behaviour into user preferences for instance how frequently user u interacts or visits user v . Mostly in literature, explicit feedback has been used as input by algorithms. Social CF is classified into two categories namely Matrix factorization based and Neighbourhood based social CF approaches. These approaches are described in the following section.

3.1 Matrix Factorization based social CF

Matrix factorization is a model based approach which focuses on learning a model by analysing data and then using the trained model for prediction [10][14]. Various social matrix factorization approaches have been proposed in literature for instance social trust ensemble model, social matrix factorization model, social recommender model, similarity based social regularisation and circle based recommendation approaches.

3.1.1 Social Matrix Factorization

The Social Matrix Factorization (SMF) model considers the transitivity of trust in social networks [15]. The behaviour of user u can be influenced by

the neighbours of u . Let S be the social network matrix. Let S_n be another matrix derived from S whose each row is normalised to 1. For prediction, model is used. Model includes two matrices: $P \in \mathbb{R}^{i_0 \times r_0}$ which represents latent features of item and $Q \in \mathbb{R}^{u_0 \times r_0}$ represents latent features of user and r_0 is the rank, u_0 denotes number of users and i_0 denotes number of items. Also, $r_0 \ll u_0, i_0$. Now, Predicted rating matrix can be modelled as follows:

$$P_m = r_m + QP^T$$

where, $r_m \in \mathbb{R}$ is a global offset value. The objective function can be optimised by minimising Root Mean Square Error (RMSE) as follows:

$$\frac{1}{2} \sum_{(u,i)obs} (R_{u,i} - P_{m_{u,i}})^2 + \frac{\gamma}{2} (\|P\|_F^2 + \|Q\|_F^2)$$

where, $R_{u,i}$ is the actual rating of item i given by user u and $P_{m_{u,i}}$ is the predicted rating of item i for user u and obs means observed.

3.1.2 Social Recommender Model

In Social Recommender Model (SRM) [27], social matrix S is modified as follows:

$$S_{u,v}' = S_{u,v} \sqrt{\frac{d_v^-}{d_u^+ + d_v^-}}$$

where, d_u^+ is the numbers of users u trusts and d_v^- is the numbers of users who trusts user v . Predicted rating matrix can be modelled as:

$$P_m = r_m + QP^T$$

where, $r_m \in \mathbb{R}$ is a global offset value. Also, social information is also used for model learning. Social matrix is predicted as follows:

$$S_r = s_m + QZ^T$$

where, $Z \in \mathbb{R}^{u_0 \times r_0}$

The objective function can be trained by minimising RMSE as:

$$\sum_{(u,i)obs} (R_{u,i} - P_{m_{u,i}})^2 + \sum_{(u,i)obs} (S'_{u,i} - S_{r_{u,i}})^2 + \frac{\gamma}{2} (\|P\|_F^2 + \|Q\|_F^2 + \|Z\|_F^2)$$

3.1.3 Social Trust Ensemble Model

Social Trust Ensemble Model (STEM) is the combination of MF and social network based approach [28]. Predicted rating matrix can be modelled as follows:

$$P_m = r_m + S_\beta QP^T$$

where, $S_\beta = \beta I + (1 - \beta)S$ and I is the identity matrix.

The objective function can be optimised by minimising RMSE as follows:

$$\sum_{(u,i)obs} (R_{u,i} - P_{m_{u,i}})^2 + \frac{\gamma}{2} (\|P\|_F^2 + \|Q\|_F^2)$$

3.1.4 Similarity based Social regularisation

Similarity between users can be computed using Pearson correlation or cosine based measures. The prediction quality of RS can be improved by taking into consideration the latent features of users during similarity computation. Similarity is computed as follows:

$$\sum_{u \in U} \sum_{v \in F_u} Sim(u, v) \|Q_u - Q_v\|^2$$

where, F_u denotes the set of direct friends of user u and $Sim(u, v)$ denotes the similarity between user u and v computed using Pearson correlation or cosine based approach.

3.1.5 Circle based recommendation

Circle-based Recommendation models are an extension of the SMF model to social networks with inferred friend circles [21]. The basic concept is that a user may not trust a friend in all categories. So while making predictions, only a subset of friends circle is considered. Predicted rating matrix for each category c can be described as follows:

$$\hat{R}_{u,i}^c = r_m^c + Q_u^c P_i^{c,T}$$

where, $Z \in \mathbb{R}^{u_0 \times r_0}$

The objective function can be trained by minimising RMSE as:

$$\begin{aligned} \frac{1}{2} \sum_{(u,i)_{obs}} (R_{u,i}^c - \hat{R}_{u,i}^c)^2 \\ + \frac{\beta}{2} \sum_u \left\| Q_u^c - \sum_v S_{u,v}^c Q_v^c \right\|^2 \\ + \frac{\gamma}{2} (\|P^c\|_F^2 + \|Q^c\|_F^2) \end{aligned}$$

3.2 Neighbourhood Based Social CF Approach

Neighbourhood Based approaches are also known as memory based approaches. It mainly includes two approaches: Nearest Neighbourhood and Social Network Traversal based approach.

3.2.1 Social Network Traversal based CF

Social Network Traversal (SNT) approach traverse the neighbourhood of a user and analyses the ratings of direct/indirect friends of the user. SNT approach is further classified into various techniques. In our paper, we have discussed in brief two SNT approaches. **MoleTrust** approach which computes trust between two indirectly connected users and considers only users within maximum depth for prediction [26]. **Bayesian Inference based Prediction** approach uses conditional probability distribution for similarity computation of two friends connected in social network[18].

3.2.2 Nearest Neighbourhood based CF

Nearest Neighbourhood (NN) based CF approach is an extension of traditional memory based approach. NN incorporates social information with the memory based approach [22]. There are various NN approaches which are illustrated in following section.

3.2.2.1 Trust CF

In Trust CF (TCF), Breadth first traversal (BFS) is used to find the neighbours of the active user u [16]. This neighbourhood is known as trusted neighbourhood. Another neighbourhood known as traditional CF neighbourhood is also computed using Pearson correlation approach. The weight for user u in the trusted neighbourhood is set to $1/d_v$, where d_v is the depth of user v from the source user u in the social trust network. The weight for a user v in the CF neighbourhood is the Pearson Correlation coefficient between v and the source user u . TCF predicts the item on the basis of weighted average of the ratings of the two neighbourhoods.

3.2.2.2 CF-ULF (User Latent Feature) Approach

CF-ULF uses MF to find user latent features. Once latent features are found, Pearson correlation is used for user clustering. k-nearest neighbours of source user u are identified. For deduction of top k-recommended items, voting scheme is used. Voting for user u relating to item i is computed as follows:

$$Vote_{u,i} = \sum_{v \in N_u} \sum_i sim(u,v) \delta_{i \in R_v}$$

where, N_u is the k-nearest neighbours of user u , R_v is the set of relevant items of user v , δ is the Kronecker delta, $sim(u,v)$ is similarity computed using Pearson correlation between user u and v .

3.2.2.3 Pure Trust Approach

Pure Trust (PT) approach uses BFS to find k-trusted users of source user u . Voting for user u relating to item i is in PT approach is computed as follows:

$$Vote_{u,i} = \sum_{v \in T_u} \sum_i wt(u,v) \delta_{i \in R_v}$$

where, T_u is the k-trusted users of user u , R_v is the set of relevant items of user v , δ is the Kronecker delta, $wt(u,v)$ is voting weight from user u which is equal to $1/d_v$.

4. CONCLUSION

Social Networking sites has been gaining popularity rapidly day by day. Recommender systems has been incorporated by e-commerce sites for the expansion of their business. With the introduction of social network, recommender system has also included social information as input. In our paper, we have presented study of various social based collaborative filtering

techniques. Social based recommender system has outperformed traditional recommender system. With the inclusion of social information in recommender systems, security and privacy has become a major concern. In future, privacy and security issues related work can be done in recommender systems.

References

1. G. Adomavicius and 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 17 (6) 734–749, 2005
2. Su, Xiaoyuan, and Taghi M. Khoshgoftaar, "A survey of collaborative filtering techniques", Advances in Artificial Intelligence Vol 3, pp 1-20, 2009
3. J.B. Schafer, J.A. Konstan, and J. Riedl, "E-Commerce Recommendation Applications," Data Mining and Knowledge Discovery, vol. 5, nos. 1/2, pp. 115-153, 2001.
4. Breese, J.S., Heckerman, D., Kadie and C.M., "Empirical analysis of predictive algorithms for collaborative filtering", in: 14th Conference on Uncertainty in Artificial Intelligence, pp. 43–52, 1998
5. Robin Burke, "Hybrid Recommender Systems: Survey and Experiments" User Modeling and User-Adapted Interaction, Vol 12, Issue 4, pp 331-370, November 2002
6. L Lü, M Medo, CH Yeung, YC Zhang, ZK Zhang and Tao Zhou, "Recommender systems", Physics Reports, Vol. 519, Issue 1, pp. 1–49, October 2012
7. R.J. Mooney and L. Roy, "Content-Based Book Recommending Using Learning for Text Categorization", Proc. Fifth ACM Conf. Digital Libraries, pp. 195-204, 2000
8. Linden, G. , Smith, B. and York, J. , "Amazon.com recommendations: item-to-item collaborative filtering" Internet Computing, IEEE Volume 7 Issue 1, pp. 76 – 80, 2003
9. U. Shardanand and P. Maes, "Social Information Filtering: Algorithms for Automating 'Word of Mouth'," Proc. Conf. Human Factors in Computing Systems, 1995
10. Y. Koren, R. Bell and C. Volinsky. "Matrix Factorization Techniques for Recommendation Systems", IEEE computer, 2009.
11. P. Resnick, N. Iakovou, M. Sushak, P. Bergstrom, and J. Riedl, "GroupLens: An Open Architecture for Collaborative Filtering of Netnews," Proc. 1994 Computer Supported Cooperative Work Conf., 1994
12. K. Yu, A. Schwaighofer, V. Tresp, X. Xu and H. P. Kriegel. "Probabilistic Memory-based Collaborative Filtering", IEEE Transactions on Knowledge and Data Engineering, vol. 16, no. 1, pp.56-69, Jan. 2004.
13. Y. Koren, R. Bell and C. Volinsky. "Matrix Factorization Techniques for Recommendation Systems", IEEE computer, 2009.
14. Y. Xin and H. Steck., "Multi-Value Probabilistic Matrix Factorization for IP-TV Recommendations", In ACM Conference on Recommender Systems (RecSys), 2011.
15. C.-N. Ziegler and G. Lausen. "Analyzing Correlation Between Trust and User Similarity In Online Communities", Proceedings of Second International

- Conference on Trust Management, pp. 251-265, 2004.
16. P. Massa and P. Avesani. "Trust-aware Recommender Systems", Proceedings of the 2007 ACM conference on Recommender systems, 2007.
 17. J. O'Donovan and B. Smyth. "Trust in recommender systems", Proceedings of the 10th international conference on Intelligent user interfaces, 2005.
 18. X. Yang, Y. Guo and Y. Liu. "Bayesian-inference based recommendation in online social networks". in IEEE Transactions on Parallel and Distributed Systems (TPDS) pp. 642-651, 2013.
 19. F. Liu and H. J. Lee., "Use of social network information to enhance collaborative filtering performance", Expert Syst. Appl., 37(7) pp. 4772-4778, 2010.
 20. H. Ma, D. Zhou, C. Liu, M. R. Lyu, and I. King., "Recommender Systems with Social Regularization", In ACM International Conference on Web Search and Data Mining (WSDM), 2011.
 21. X. Yang, H. Steck and Y. Liu, "Circle-based Recommendation in Online Social Networks", Proc. of 2012 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'12)
 22. X. Yang, H. Steck, Y. Guo and Y. Liu, "On Top-k Recommendation using Social Networks", In ACM Conference on Recommender Systems (RecSys'12), 2012.
 23. Pasquale Lops, Marco de Gemmis and Giovanni Semeraro, "Content-based Recommender Systems: State of the Art and Trends" pp. 73-105, 2011
 24. M. Balabanovic and Y. Shoham, "Fab: Content-Based, Collaborative Recommendation," Comm. ACM, vol. 40, no. 3, pp. 66-72, 1997
 25. J. Bobadilla , F. Ortega , A. Hernando and A. Gutiérrez," Recommender systems survey", Knowledge-Based Systems, Vol. 46, pp.109-132, July, 2013
 26. P. Massa and P. Avesani. "Controversial users demand local trust metrics: an experimental study on epinions.com community", Proceedings of the 25th American Association for Artificial Intelligence Conference, 2005
 27. H. Ma, H. Yang, M. R. Lyu, and I. King., "Sorec: Social recommendation using probabilistic matrix factorization" In International Conference on Information and Knowledge Management (CIKM), 2008
 28. H. Ma, I. King, and M. R. Lyu., "Learning to recommend with social trust ensemble" In ACM conference on Research and development in information retrieval (SIGIR), 2009