International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2017 IJSRCSEIT | Volume 2 | Issue 5 | ISSN : 2456-3307

Collaborative Filtering Based Product Recommendation System for Online Social Networks

Sanjeev Dhawan[1], Kulvinder Singh[2] and Naveen Kumar[3]

[1, 2]Faculty of Computer Science and Engineering , Department of Computer Science and Engineering, University Institute of Engineering and Technology, Kurukshetra University, Kurukshetra-136119, Haryana, India.

[3]M.Tech. Computer Engineering, University Institute of Engineering and Technology, Kurukshetra University, Kurukshetra-136119, Haryana, India.

ABSTRACT

In recent years the demand of social networking websites has been increased due to frequent uses of these sites by people. Social Networks play important role in the product's recommendation because nowadays peoples are connected with each other through Facebook, Twitter, Google+ etc., so it is easy to recommend a product to friends by social websites that whether to go for this product or not. Even companies are marketing their products on these social media's. In this paper, an attempt has been made to discuss product recommendation system and its related techniques. This paper is divided into four sections. In section I, product recommendation system is discussed. In section II, works related to different papers has been discussed. Section III includes different existing techniques with their benefits and drawbacks. Section IV presents problem statement and in section V different performance metrics has been presented.

Keywords: Social Networks (SNs), Products, Recommendation and Collaborative Filtering (CF).

I. INTRODUCTION

In the era of information technology and the Internet, people are getting very much confused with the huge amount of information, which leads to information overload. In this environment, whether it is information of consumers or product are experiencing a great challenge. As a consumer, it is very difficult to find information of interest from a lot of information available on any social network and also much harder to find the product of user's interest from enormous amount of products. Recommended system is the solution to this problem and makes intelligent agent system, and related techniques have been widely employed in some large famous commercial systems, such as Amazon, Ebay. With the enlargement of e-commerce, the number of users and goods are increasing exponentially, but traditional collaborative filtering technology performance is getting worse. Recommendation systems are based on the rankings and ratings given to product by the users [1]. So far, the collaborative filtering algorithm is still one of the most successful e-commerce recommendation systems [2].

A. Collaborative Filtering

Collaborative Filtering is the technique of automatic filtering or analysing items by gathering the opinions of other people. The terminology collaborative filtering (CF) has only been alive for a little more than a decade, CF takes its geneses from something humans have been doing for centuries sharing opinions and interests with others. For years, people have followed the outdated

tradition of discussing only those products which they have purchased, restaurants they have tried, and music they have listened and then used these handfuls of information to form opinions. Computers and the web technology allow us to visualize beyond the simple word-of-mouth. Instead of limiting ourselves to tens or hundreds of individuals, the Internet allows us to consider the opinions of thousands and even more. The speed of computer allows us to exercise these opinions in practical time and determine not only what a much larger community thinks about an item, but also expand a truly personalized view of that item using the most appropriate opinions for a given user or group of users.

The term user refers to any individual who issues ratings to a system. Most often, we utilize this term to refer those people who use a system to receive information (e.g., recommendations) although it also refers to those who provided the data (ratings) used in triggering this information. A collaborative filtering system generates forecast or suggestions for a given user for one or more items. Items can be anything such as art, books, CDs, restaurants, journal articles, or vacation destinations to which a human can imagine and provide a rating. Ratings in a Collaborative filtering system can be used in a variety of forms.

1) Item-Based Collaborative Filtering

It is most predictable and different method in the area of filtering algorithms that was proposed recently and is based on item relations and not on user relations, as in typical Collaborative Filtering. In Item-based Collaborative Filtering process, we took the group of items, that the dynamic users have rated, then analyse how similar are those item to the goal item and then choose the k most similar items {a1, a2, ak}, based on their parallel similarities {sim1, sim2, ..., simk}. The calculations can then be computed by taking weighted average scores given by the dynamic users on these associated items. The first step in this new approach is the 'Representation'. Its solution is related as with the classic Collaborative Filtering procedure: represent the data in an ordered manner [3].

2) Content-Based Collaborative Filtering

The basic thinking behind Content-Based Collaborative Filtering is to use a content-based analyser to enhance current user data which is communicated via the user-item matrix 'R' and then deliver personalized suggestions through collaborative filtering. The content-based analyst is practical on each row from the first user-item matrix, corresponding to every individual user, and gradually creates a pseudo user-item matrix, 'PR'. At the end, each row 'i', of the pseudo user-item matrix 'PR' consists of the scores provided by user 'ui', when available and those grades analysed by the content-based predictor [4].

3) Link Analysis Algorithm

New recommendation algorithm which we developed lately is based on the thoughts from link analysis research [5]. Association analysis procedures found to have necessary application in Web page ranking and social network study.

B. Sparsity Problem in Recommendation System

Sparsity is a big problem for recommendation system. A huge number of products are sold on various e-commerce sites but only small subset of products is rated by frequent active users. Thus, it leads to a very big problem because the most popular items have very few ratings due to which users are not able to purchase good products. To resolve this problem an intelligent recommendation system is required which is capable of finding the ratings of those products which are not rated by the active users.

II. RELATED WORK

Recommendation system is to solve this problem and make intelligent agent system, and related techniques have been widely employed in some large famous

commercial systems, such as Amazon, Ebay etc. Here, in this section literature review on different existing techniques proposed by various researchers has been presented.

Yu et al. [6] developed a probabilistic framework for memory-based CF (PMCF). This demonstrates their proposed PMCF framework allowed an accurate and efficient prediction of user's preferences. Similarly, Gediminas Adomavicius and Alexander Tuzhilin [7] presented an overview of the field of recommender systems and described the current generation of recommendation methods. A new heuristic similarity measure called PIP for collaborative filtering has been presented by Hyung Jun Ahn [8] that is widely used for automated product recommendation in Internet stores. Liu et al. [9] proposed a hybrid method that considers customer's purchase sequences over time and their purchase data for the current period. In particular, they show that how the input variables can be coded into GA chromosomes in various modes. The use of a recommender system was introduced by Afsarmanesh et al. [10] to assist designer's sub-products and reusing the existing specifications, as well as recommending business services that can enhance defined sub-products. Hasan et al. [11] presented a new similarity along with the concept of super similarity, average similarity and super dissimilarity. Similarly, Yang et al. [12] proposed a latent social trust network model to improve the recommendation performance. Whereas, a novel two-layer neighbour selection scheme was proposed by Zhang et al. [13] that enhances the quality of neighbour selections by selecting the most influential and trustworthy neighbors. Hooda et al. [14] presented framework, in which the combination of multiple similar metrics derived from both implicit and explicit social networks were used to obtain ration of items. The framework uses keywords of products that are searched by customers and based on these keywords recommend the best products for them. Dhawan et al. [15] presented enhanced collaborative filtering mechanisms in which rating of products were assigned by user's recent preference and interest. Dhawan et al. [16] surveyed emotion mining techniques in social networks. Sometimes most of the active and vibrant users have a great impact on influencing others views and thoughts, such an impact can be negative as well as positive depending upon the conveyer and receptor's ability to think and understand. Also, these factors affect a person's reaction to a particular situation and there by influence others.

Although the approaches mentioned above can improve the recommendation performance to some extent, the problems referring to sparse ratings and it affect the recommendation quality badly.

III. EXISTING TECHNIQUES

A. Collaborative

The recommendation is done on the basis of user access history and rating. The advantage of collaborative techniques is that there is no requirement of manual data maintenance but at the same time drawbacks of this technique are faces scalability, cold start and sparsity problem.

B. Content-Based

Its work is based on similar item attributes. Benefit of this technique is high recommendation accuracy for smaller groups, whereas manual maintenance of data causes error in process.

C. Knowledge Based

On the basis of Implicit and Explicit data recommendation is done in this technique. It uses implicit as well as explicit data to predict items that the users may like. This technique is highly complex.

D. Demographic

In this technique recommendation relies upon the user's behaviour and interest. This technique finds the relationship between liked items and users. Drawback

of this technique is the delay in user profile updation when the user's interest changes.

IV. PROBLEM STATEMENT

In latent social network, products are recommended based on three conditions [17]. The first condition was coupling trust, in this condition two friends recommended one common neighbour .The second condition was co-citation trust, in this condition the users are trusted by numbers of neighbour and the last condition was common trust relationship between users based on similar interests. The drawback of this mechanism was that by checking all these condition creates lots of overhead. Another drawback was recommendation is totally dependent on neighbour so if a neighbour behaves maliciously and recommend a wrong product then it may create problem.

So to overcome these kinds of problem a novel mechanism is proposed. In this technique products are recommended on the basis of calculated common interest using trust mechanism. In this technique data set is formulated into user-item matrices which are based on user's ratings.

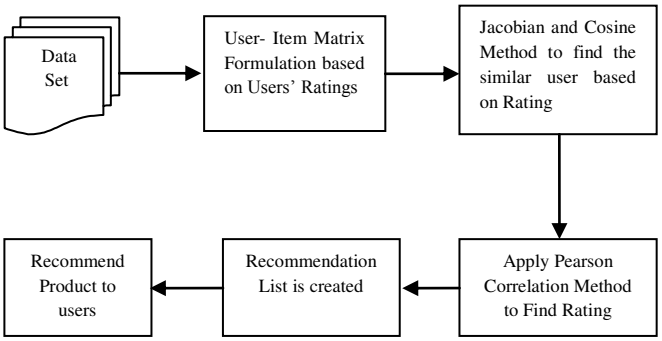

Figure 1. Proposed Product Recommendation System.

After that Jaccard and Cosine Method are used to find the similarities between the user's interests. Interest refers to both liking and disliking any product by users. Using these methods, users of similar interest are clustered or grouped together. This group of users is called the neighborhood of the user to whom we want to make recommendations. In these groups, users of similar interest as the given user are selected and are represented using User-Item Matrix. On this basis a Recommender System can be built which can find those products that are liked by lots of users in the group. Then it able to find the rating of the product on behalf of the other users rating prediction using Pearson Correlation method and then recommends those products to other users on the basis of rating using some social networking sites.

This proposed method helps to overcome the sparsity problem. This method uses the Similarities between the users and then finds the rating of the products even if products are rated by fewer users.

V. METRICS USED IN RECOMMENDER SYSTEMS

The metrics used in recommendation system are classified as below:

Mean Absolute Error (MAE) it measures to what extent a system can predict ratings of users. Classification accuracy metrics measure how well a system is able to classify items correctly such as Precision, Recall, F1-measure and ROC Sensitivity.

Coverage metrics measure the percentage of items for which the system can make recommendations.

Confidence metrics measure how certain the system is of the accuracy of the recommendations.

A. Mean Absolute Error (MAE)

Compute the average of the absolute difference between the predictions and true ratings.

$$MAE=\frac{\sum_{\{i,j\}}|p_{i,j}-r_{i,j}}{n}$$

Here n is the total no. of ratings over all users, pi,j is the predicted rating for user i on item j and ri,j is the actual rating. When MAE value goes low then predication rate increases.

B. Normalized Mean Absolute Error (NMAE)

Normalizes MAE is used to express errors as percentages of full scale.

$$NMAE = \frac{MAE}{r_{max} - r_{min}}$$

Here, rmax and rmin are the upper and lower bounds of the ratings.

C. Root Mean Squared Error (RMSE)

It is one of the most popular performance measurements metric for recommender systems. It is expressed as

$$RMSE = \sqrt{\frac{1}{n}\sum_{\{i,j\}}(p_{i,j} - r_{i,j})^2}$$

RMSE amplifies the contributions of the absolute errors between the predictions and true values.

D. Accuracy

Accuracy simply measures how often the classifier makes the correct prediction.

$$Accuracy = \frac{Numbers\ of\ Correct\ Predictions}{Numbers\ of\ \ Total\ Data\ Points}$$

E. Precision and Recall

Precision and Recall metrics are usually used together. Precision means out of the items that the ranker/classifier predicted to be relevant, how many are truly relevant. On the other hand, recall measures out of all the items that are truly relevant, how many are found by the ranker/classifier.

$$Precision = \frac{Numbers\ of\ Correct\ Answers}{Numbers\ of\ \ Total\ Items\ Returned\ By\ Ranker}$$

$$Recall = \frac{Numbers\ of\ Correct\ Answers}{Numbers\ of\ \ Total\ Items\ Relevant\ Items}.$$

F. F-score

It is the harmonic mean between Precision and Recall.

$$F\text{-measure} = \frac{2\ Precision\ Recall}{Precision + Recall}$$

VI. CONCLUSION

In this paper, the main focus on the sparsity problem and enhance the recommendation accuracy in collaborative filtering systems. A new collaborative filtering algorithm is proposed to increase the recommendation accuracy. This approach is effective in finding similarities between user's interests which is used to find the ratings of products using some techniques like Pearson Correlation. The proposed solution may help to handle the sparsity problem and achieved significantly better recommendation quality than standard collaborative filtering approach.

VII. REFERENCES

[1] White, T.Chu, W.Salehi-Abari, A Media Monitoring Using Social Networks , IEEE Second International Conference on Social Computing, pp. 661-668, Aug. 2010.

[2] WANG Guoxia, LIU Heping. Survey of Personalized Recommendation System, Computer Engineering and Application, pp. 66-76, Nov. 2013.

[3] Ziyang Zhang, Yuhong Liu, Zhigang Jin and Rui Zhang, Selecting Influential and Trustworthy Neighbors for Collaborative Filtering Recommender Systems, IEEE 7th Annual computing and Communication Workshop and Conference, pp. 1-7, January 2017.

[4] Aggarwal, C. C., Wolf, J. L., Wu K., and Yu, P. S., Horting Hatches an Egg: A New Graph-theoretic Approach to Collaborative Filtering, The 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA. pp. 201-212, August 1999.

[5] F. Shao, F. Long, J. Liang, et al. Personalized recommendation based on trust and preference, Applied Mechanics and Materials, Volume 713-715, pp. 2288-2291, 2015.

[6] Kai Yu, Anton Schwaighofer, Volker Tresp, Xiaowei Xu, and Hans-Peter Kriegel, Probabilistic Memory-Based Collaborative Filtering, IEEE Transactions on Knowledge and Data Engineering, Volume 16, Issue 1, pp.56-69, Feb. 2004.

[7] Gediminas Adomavicius and Alexander 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, Volume 17, Issue 6, pp. 734-749, April 2005.

[8] Hyung Jun Ahn, A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem, Elsevier, Volume 178, Issue 1, pp. 37-51 January 2008.

[9] Duen-Ren Liu, Chin-Hui Lai and Wang-Jung Lee, A Hybrid of Sequential Rules and Collaborative Filtering for Product Recommendation, The 4th IEEE International Conference on Enterprise Computing, E-Commerce and E-Services (CEC-EEE), pp. 211-220, 2007.

[10] Simon Fong, Yvonne Ho and Yang Hang, Using Genetic Algorithm for Hybrid Modes of Collaborative Filtering in Online Recommenders, Eighth International Conference on Hybrid Intelligent Systems, pp. 174-179, 2008.

[11] Hamideh Afsarmanesh, Mohammad Shafahi and Mahdi Sargolzaei, On service-enhanced product recommendation\ Guiding users through complex product specification, IEEE International Conference on Computing and Communications Technologies (ICCCT'15), pp.43-48, 2015.

[12] Mahamudul Hasan, Shibbir Ahmed, Md. Ariful Islam Malik and Shabbir Ahmed, A Comprehensive Approach towards User-Based Collaborative Filtering Recommender System, IEEE International Workshop on Computational Intelligence (IWCI), pp.159-164., 2016.

[13] Xian Yang, Jiangning Wu, Yanzhong Dang and Lili Rong, A Product Recommendation Approach Based on the Latent Social Trust Network Model for Collaborative Filtering, IEEE International Conference on Software Quality, Reliability and Security Companion pp. 178-185, 2016.

[14] Rahul Hooda, Kulvinder Singh and Sanjeev Dhawan, Social Commerce Hybrid Product Recommender, International Journal of Computer Applications, Volume 100– No.12, pp. 43-49, August 2014.

[15] Sanjeev Dhawan, Kulvinder Singh and Jyoti, High Rating Recent Preferences Based Recommendation System, 4thInternational Conference on Eco-friendly Computing and Communication Systems, pp. 259-264, 2015.

[16] Sanjeev Dhawan, Kulvinder Singh and Deepika Sehrawat, Emotion Mining Techniques in Social Networking Sites, International Journal of Information & Computation Technology, Volume 4, No. 12 , pp. 1145-1153, 2014.

[17] Xian Yang, Jiangning Wu, Yanzhong Dang and Lili Rong, A Product Recommendation Approach Based on the Latent Social Trust Network Model for Collaborative Filtering, IEEE International Conference on Software Quality, Reliability and Security Companion, pp. 178-185, 2016.