An Hybrid Approach for an Improved Recommendation System by Combining the Concepts of Fuzzy Clustering and Voting Theory Techniques

Pankaj Agrawal*, Prashant Agnihotri**, Prof. Imran Ali Khan*** and Prof. Damodar Tiwari****
*MTech Research Scholar Dept. of CSE , Bansal Institute of Research and Technology , Bhopal , Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal M.P., India
-**Bansal Institute of Research and Technology, Bhopal

Abstract: On the world wide web, there's need to filter, prioritize and effectively deliver pertinent data so as to alleviate the problem of information overload, which has made a possible problem to many Internet users. Recommender systems solve this issue by searching through large volume of dynamically generated information to provide users with personalized content and solutions. Recommender systems are information filtering systems which manage the problem of information overload by filtering vital info fragment out of substantial amount of dynamically generated information according to user's preferences, attention, or observed behaviour about item. Recommender system has the power to predict if a specific user would prefer an item or not based on the consumer's profile. The current function combines the methods for Fuzzy Clustering and Voting theory approaches in order to design an efficient Recommender System.

Keywords: Recommender Systems, Fuzzy Clustering, Fuzzy C-Means Algorithm, Voting Theory.

Introduction

Ever increasing quantity of information about the internet has generated the need for automatic filtering, refinement and personalized presentation of information for users to help decision making. There have been efforts to design data filtering systems which filter and present information in line with the tastes of the individual user. By way of instance, the simple act of a user buying or browsing an item might be seen as an endorsement for this item. Such kinds of feedback are generally used by online merchants like Amazon.com, and the group of the sort of data is totally effortless concerning the work required of a customer. The simple idea of recommender systems would be to use these many sources of information to infer customer interests. The thing to which the recommendation is supplied is referred to as the consumer, and the merchandise being recommended is also called a product. Consequently, recommendation evaluation is often dependent on the previous interaction between users and items, because beyond interests and proclivities are often good indicators of future choices [5]. Various types of recommender systems are as given below:

Content-based filtering

Content-based recommender systems operate with profiles of users that are created at the start. A profile includes information about an individual and his preference. Taste is dependent on how the user rated items. Generally, when creating a profile, recommender systems create a survey, to find initial information about a user in order to avoid the new-user problem [6].

Collaborative Filtering

Collaborative filtering became among the most researched techniques of recommender systems since this approach was mentioned and described by Paul Resnick and Hal Varian in 1997. [7] The idea of collaborative filtering is in locating users in a neighborhoodwhich discuss appreciations [8]. If two users possess the same or almost same graded things in common, then they have similar tastes. Such users construct a set or even a so called neighbourhood. A user receives recommendations to those items that he/she hasn't rated earlier, but that were already favourably rated by users in his/her neighbourhood.

Related Work

In the procedure DBSCAN clustering algorithm is used for clustering the consumers, and then implement voting algorithms to recommend items to the consumer depending on the audience to which it belongs. The idea is to partition the consumers of the RS using clustering algorithm and use the Recommendation Algorithm separately to each partition. Our system recommends item to a user in a certain cluster only using the rating numbers of the other users of the cluster. This helps us to reduce the running time of the algorithm as we avoid computations over the whole data. Our objective is to enhance the running time in addition to maintain an acceptable recommendation quality. An another method first uses similarity

calculation method to reduce the sparseness. Then fuzzy clustering method is used to look for nearest neighbors for goal user. Then matrix is utilized after clustering to create recommendations [2]. Additionally, recommender system was defined from the view of E-commerce for a tool that helps users search through records of information that's related to users' attention and preference [4]. Recommender system was described as a way of assisting and augmenting the social process of using recommendations of others to make choices when there is no adequate personal knowledge or experience of the options [9]. Recommender systems handle the issue of data overload that consumers normally encounter by providing them with personalized, exclusive articles and service recommendations. Lately, various approaches for construction recommendation systems have been developed, which may utilize collaborative filtering, content-based filtering or hybrid filtering [10-12]. Collaborative filtering method is the most mature and the most commonly implemented. Collaborative filtering urges items by identifying different users with similar taste; it uses their opinion to recommend items to the busy user. Collaborative recommender systems are implemented in different application areas. GroupLens is a news-based architecture which employed collaborative methods in assisting users to find articles from enormous news database [13]. Ringo is an internet social information filtering program which uses collaborative filtering to create users profile based on their evaluations on audio albums [14]. The system employs collaborative filtering method to overcome scalability issue by creating a table of comparable things offline through the usage of item-to-item matrix. The machine then recommends other products which are similar online according to the users' purchase history. On the other hand, content-based techniques match content tools to user characteristics. Content-based filtering techniques normally base their forecasts on consumer's information, plus they dismiss contributions from other users as with the case of collaborative techniques [16, 17].

Problem Identification

The current work focuses on addressing two major challenges namely sparsity and scalability.

1. **Scalability:** In many of the environments in which these systems make recommendations, there are millions of users and products. Thus, a large amount of computation power is often necessary to calculate recommendations.
2. **Sparsity:** The number of items sold on major e-commerce sites is extremely large. The most active users will only have rated a small subset of the overall database. Thus, even the most popular items have very few ratings.

Proposed Methodology

The proposed methodology combines techniques of fuzzy clustering and voting theory in order to produce a hybrid approach. In the first phase fuzzy clustering technique is applied in order to cluster similar objects so that the recommendations are generated for target user as per neighbourhood users in efficient manner. In the second phase voting theory is applied in order to select cluster with maximum ratings or reviews.

The basic steps involved in the algorithm are as follows:

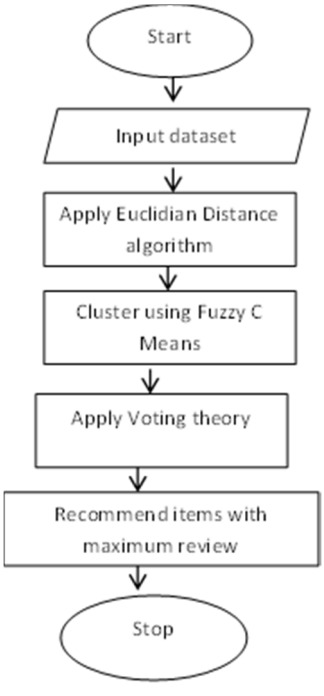

Figure 1: Recommender System using Hybrid Approach

1. Collect the review/rating of each user for the given item for the recommending the item for the target user.
2. Build user-review matrix, but these matrixes cause sparseness i.e. most of the element in the given matrix contains zero elements.
3. Build a user-user matrix also known as relational matrix, by using this matrix sparness is removed, these is done by using Euclidean distance algorithm. The distance formulae are as follows:

$$d_{x,y} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \ldots\ldots (1)$$

4. Clustering is performed on resultant matrix by using the fuzzy c means clustering method.

5. Voting theory technique is applied in order to remove the scalability of clustering techniques.

6. The cluster with maximum rating is obtained after applying Voting theory for displaying result in front of user.

Fuzzy C Means Clustering Algorithm Pseudo code

1. Set the number of clusters, the fuzzy parameter (a constant > 1), and the stopping condition
2. Initialize the fuzzy partition matrix
3. Set the loop counter k = 0
4. Calculate the cluster centroids, calculate the objective value j
5. For each pixel, for each cluster, compute the membership values in the matrix
6. If the value of j between consecutive iterations is less than the stopping condition, then stop; otherwise, set k=k+1 and go to step 4
7. Defuzzification and segmentation

Expected Outcome

The proposed work as following advantages:

1. The system stores user preferences in the form of user provided weights of different attributes. The system also stores the relative importance of each of these attributes as specified by the users.
2. The system recommends a few most probably likeable items to the users. Principles of voting theory are applied to recommend items on the basis of the cluster to which the user belongs.

The future outcome of the proposed work will be and Hybrid Approach to develop an efficient Recommendation System by combining Fuzzy clustering and Voting theory techniques. Some major outcomes are:-

1. The proposed system will be scalable as the computational requirements will be reduced as only the scalability will be reduced because of selection of data on the basis of voting algorithm. Only top ranked records will be selected for further processing.
2. The sparsity of the data will be reduced as a relational matrix is formed on the basis of relation distance algorithm.
3. An analysis of the improved approach will be displayed.

Conclusion

With the ever increasing e-commerce data, storing and querying massive data is becoming a big challenge for recommender systems. With increase in number of users and item the recommendation system needs to be more accurate and is becoming more and more difficult to meet real-time demands. The proposed work address two major problems in existing recommended systems namely sparsity and scalability by creating a hybrid approach.The proposedapproach first solves the sparsity problem by reducing the zero elements in the given sparse matrix and converts it into non-zero elements by converting it into a dense matrix and thenclustering technique is applied to the given matrix to produce clusters which holds the similar objects/reviews. And finally, after obtaining the clusters votingtheory technique is applied to select the maximum value of the given cluster thus reducing the scalability problem.

References

[1] Honey chatting: a novel instant messaging systemRobust to eavesdropping over communication,Joo-imkim and ji won yoon,IEEE, 2016 ,center for information security technologies (CIST)
[2] Korea university, seoul, republic of korea
[3] Honey encryption beyond message recovery security,Joseph jaeger,Thomasristenpart&Qiang tang,Eurocrypt,2016.
[4] Schafer JB, Konstan J, Riedl J. Recommender system in ecommerce. In: Proceedings of the 1st ACM conference on electronic commerce; 1999. p. 158–66.

[5] Resnick P, Varian HR. Recommender system's. Commun ACM 1997;40(3):56–8. http://dx.doi.org/10.1145/245108.24512.[5] An Introduction to Recommender Systems, C.C. Aggarwal, Springer International Publishing Switzerland 2016.
[6] L. Candillier, F. Meyer, F. Fessant, and K. Jack, "State-of-the-art recommender systems," 2009.
[7] P. Resnick and H. R. Varian, "Recommender systems," Commun. ACM, vol. 40, pp. 56–58, March 1997. [Online]. Available: http://doi.acm.org/10.1145/245108.245121
[8] Z. Sun and N. Luo, "A new user-based collaborative filtering algorithm combining data-distribution," in Proceedings of the 2010 International Conference of Information Science and Management Engineering - Volume 02, ser. ISME '10. Washington, DC, USA: IEEE Computer Society, 2010, pp. 19–23. [Online]. Available: http://dx.doi.org/10.1109/ISME.2010.48
[9] Acilar AM, Arslan A. A collaborative filtering method based on Artificial Immune Network. ExpSystAppl 2009;36(4):8324–32.
[10] Chen LS, Hsu FH, Chen MC, Hsu YC. Developing recommender systems with the consideration of product profitability for sellers.Int J Inform Sci 2008;178(4):1032–48.
[11] Jalali M, Mustapha N, Sulaiman M, Mamay A. WEBPUM: a web-based recommendation system to predict user future movement. ExpSystApplicat 2010;37(9):6201–12.
[12] Adomavicius G, Tuzhilin A. Toward the next generation of recommender system. A survey of the state-of-the-art and possible extensions. IEEE Trans Knowl Data Eng 2005;17(6):734–49.
[13] Ziegler CN, McNee SM, Konstan JA, Lausen G. Improving recommendation lists through topic diversification. In: Proceedings of the 14th international conference on World Wide Web; 2005. p. 22–32.
[14] Min SH, Han I. Detection of the customer time-variant pattern for improving recommender system. ExpSystApplicat 2010;37(4):2911–22.
[15] Celma O, Serra X. FOAFing the Music: bridging the semantic gap in music recommendation. Web Semant: SciServ Agents World Wide Web 2008;16(4):250–6.
[16] Lieberman H. Letizia: an agent that assists web browsing. In: Proceedings of the 1995 international joint conference on artificial intelligence. Montreal, Canada; 1995. p. 924–9. [17] Pazzani MJ. A framework for collaborative, content-based and demographic filtering.ArtificIntell Rev 1999;13:393–408, No. 5(6)