International Journal of Advanced Trends in Computer Applications
www.ijatca.com

Hybrid Recommender System for Research Papers

Atul Kumar[1], SumitKaur[2], Gurbinder Kaur[3], HarpreetKaur Virk[4]

[1]Atul Kumar
PG Student
Department of Computer Science &Engg.
Chandigarh University, Gharuan, Mohali(Punjab)
thakur.atul225@gmail.com

[2]SumitKaur
Assistant Professor
Department of Computer Science &Engg.
Chandigarh University, Gharuan, Mohali(Punjab)
sumit.bhangu87@gmail.com

[3]GurbinderKaur
PG Student
Department of Computer Science &Engg.
Chandigarh University, Gharuan, Mohali(Punjab)
gudduaujla@gmail.com

[4]Harpreet KaurVirk
PG Student
Department of Computer Science &Engg.
Chandigarh University, Gharuan, Mohali(Punjab)
Kaurhp.17@gmail.com

Abstract: *Recommender systems these days have changed the way people search for new items, information, and even other person on social networking. Recommender systems look for pattern of behavior of users to know what that particular user will prefer from catalogue of items he has never experienced before. Evolution of the technology behind recommender systems has taken place from the past 20 years. This paper proposed a hybrid recommender system for research papers, which can be used as the alternative to the current systems. This system is hybridization of content filtering and collaborative filtering techniques. This system based on bottom up approach to classify research papers i.e. users can tag their papers with their own keywords. This system uses the content and collaborative filtering techniques along with author analysis, date analysis and title analysis. This system also optimized the retrieved results by using collaborative approaches like user to item techniques and item to item technique. Cold start problem for the new items also eliminated in proposed system.*

Keywords: *Recommender system, collaborative tagging, content filtering, collaborative filtering, and cold start problem.*

I. Introduction

Recommendations are the suggestions given to us by others in our daily life. Weall are actually depend on suggestions from others people to do our routine work for example before watching a movie we ask our friends to review that movie. And the recommender systems which provide the related recommendations to the users are the best recommender systems for example if kids are watching some cartoon video then ads or recommendations should be according to the kids not to the adults. Recommender Systems (RSs) are software tools and techniques providing suggestions for items to be of use to a user and the suggestions relate to various decision-making processes, such as what items to buy, what music to listen to, or what online news to read [1].These days recommender systems play an important

role in highly rated e-commerce and Internet sites like Amazon.com, Flipkart.com, Myntra.com, Jabong.com, YouTube, Yahoo, Netflix. Similarly in a year more and more research papers are publishing exponentially, therefore it is hard to find the related research paper for research work. The inefficiency of the current recommender systems makes it more difficult to find related research articles to the researchers.

This paper proposed a hybrid recommender system for research papers which is based on content filtering and collaborative filtering techniques like User to Item and Item to Item approach. In content based recommender system recommendations are given to users on basis of description of items and the user profile, in content based recommender systems, keywords are used to describe the items. The content of each item are the set of terms, they are usually the words that occur frequently in the document. Recommendations are made by matching user profile with the same terms which occurs in the research paper. There are many problems occurs in implementation of a content based recommender system. First, terms can either be assigned automatically or manually. When terms are assigned automatically a method has to be chosen that can extract these terms from items. Second, the terms have to be represented such that both the user profile and the items can be compared in a meaningful way. Third, a learning algorithm has to be chosen that is able to learn the user profile based on seen items and can make recommendations based on this user profile [2]. Content filtering approach is commonly used with the text documents.

The proposed system is using the collaborative tagging technique that follow the bottom up approach for the recommender system which simply means that tagging a research articles is not done by the system administrator only, user or author of the research articles have authority to tag their own research paper[14]. Collaborative tagging is a powerful approach that allows user to index or tag their items.
All users can participate and contribute their own personal vocabularies (keywords) to generate a collaboratively built 'bottom-up' system which more accurately reflects users' thinking of the world around them [3]. Also implementing the collaborative tagging approach is very economical. The content filtering technique in the proposed system based on tag assigned by the user itself.

The proposed system uses basic collaborative approaches like Memory based collaborative filtering and Model based collaborative filtering. Further the proposed system can be optimized using meta-heuristic approach like Ant Colony Optimization (ACO), PSO.

Cold start problem is very common in recommender systems. The proposed system also provides solution to the cold start problem for the new items i.e. research papers by combining both content and collaborative filtering techniques. Cold start problem occur when new item or new user is added i.e. we do not know how to recommend that new item or what to recommend to that new user using the system.

The paper itself is structured as follows. Section 2 provides the related work of content and collaborative based recommender systems. Section 3 provides proposed methodology used by system to recommend the research paper. Section 3 provides discussion and future work. Finally, Section 5 summarizes the paper and points out future work.

II. Related Work

This section of the paper has highlighted some previous work done in the field of content and collaborative recommender systems. However in this section complete list of related work is not presented.In the field of recommender systems for research paper there is not much work is done. The use content and collaborative filtering for recommendations are suggested by different authors. The strength of currently used academic search engines lies in finding documents containing specific keywords. Due to synonyms and unclear nomenclatures, this approach delivers in practice, often unsatisfying results [4]. In text based analysis there are so many problems occurs like unclear nomenclatures, synonyms or ambiguity in the meanings of words. Therefore if different words or terms are used then it is difficult to relate similar research papers.

The cold start problem occurs when we using collaborative filtering for the recommending research paper because it is highly rely on the rating from different users, but if there will be no user or user do not rate any item then there will cold start problem for sure.Collaborative filtering used in the field of recommender system is criticized for various reasons. Some authors claim that collaborative filtering would be ineffective in domains where more items than users exist. Others believe that users would not be spending time for explicitly rating research papers [4].

Several authors via research papers documented the benefits of use of collaborative tagging. Vander Wal [8] and Mathes [9] have discussed the potential benefits of tagging for personal information management. Vander Wal [8] has observed that in tagging systems there exists a powerful tool, allowing users to index their

information resources with their own keywords [3].Collaborative tagging technique follow the bottom up approach for the recommender system which simply means that tagging a research articles is not done by the system administrator only, user or author of the research articles have authority to tag their own research articles.

Nan Zheng, Qiudan Li [5] proposed the recommender system based on tags and time information. They show empirically using data from a real-world dataset that tag and time information can well express user's taste and we also show that better performances can be achieved if such information is integrated into CF [5].

Another Collaborative filtering recommender system based on tag information is proposed by author [6]. This paper explores the utilization of tagging information to provide the related recommendations. This is based on the distinctive three dimensional relationships among users, tags and items, a new user profiling and similarity measure method is proposed. Author experiments suggest that the approach proposed is better than the traditional collaborative filtering recommender systems using only rating data [6].

Several author also proposed the incorporation of different filtering techniques with the meta-heuristic techniques to provide accurate and precise results. PunamBedi, Ravish Sharma [7] has proposed the trust based recommender system using ant colony optimization. They incorporated the collaborative filtering technique with the ant colony optimization.

From the literature survey we concluded that there are only few recommender systems which are exist for research paper recommendation. Those who are exists currently using either content or collaborative filtering techniques and provide insufficient results. Both the techniques have their own limitations. Also from the literature survey we found that the accuracy and precision of the results can be improved by the use collaborative tagging and Meta heuristic techniques.

III. PROPOSED WORK

This paper proposed the hybrid recommender system for research papers based on content filtering and collaborative filtering techniques. The methodology to recommend research paper by proposed hybrid recommender system follows steps is mentioned and explains as follows.

(A) This System allows user to upload research papers.

(B) This System facilitates user to tag their own papers i.e. with their own keywords

(C) Once the research paper is uploaded to the system. User can search research papers by entering keywords then the recommendations will be provided on the basis tag matching i.e. content filtering.

(D) Recommendations to the user provided according to the user profile and its past search history.

(E) This system also uses collaborative filtering i.e. when user will also get recommendations according to liking of the similar users

(F) Two approaches of collaborative filtering will be proposed in this paper which further can be compare and best approach can be integrated into the system

i) **User–User Collaborative Filtering**: User–user collaborative filtering, also known as k-NN collaborative filtering. User–user collaborative filtering find other users whose past rating behavior is similar to that of the current user and use their ratings on other items to predict what the current user will like[14]

ii) **Item–Item Collaborative Filtering:** To predict preferences this approach uses similarities between user's rating behaviors, item–item collaborative filtering uses similarities between the rating patterns of items. If two items tend to have the same users like and dislike them, then they are similar and users are expected to have similar preferences for similar items[14].

iii) To find similarity between users and items correlation coefficients are used. The correlation coefficient of two variables, is the covariance of the two variables divided by the product of their individual standard deviations. It is a normalized measurement of how the two variables are linearly related [14].

(G) Cold start problem: Cold start problem is very common problem in recommender system where we do not know how to recommend that new item or what to recommend to that new user

added to the system. This system provides equal chance to all new research articles by using both content and collaborative filtering

(H) Performance measurement- Performance of the system [3] [10] can be measure using two parameters.

- **RECALL** is the ratio of the number of relevant records retrieved to the total number of relevant records in the database. It is usually expressed as a percentage.

$$\text{Recall} = \frac{\text{No. of target research articles retrieved}}{\text{No. of target research articles}}$$

- **PRECISION** is the ratio of the number of relevant records retrieved to the total number of irrelevant and relevant records retrieved. It is usually expressed as a percentage.

$$\text{Precision} = \frac{\text{No. of target research articles retrieved}}{\text{No. of research articles retrieved}}$$

X= No. of relevant records retrieved
Y= No. of relevant records not retrieved
Z= No. of irrelevant records retrieved

$$\text{RECALL} = \frac{X}{X+Y} * 100$$

$$\text{PRECISION} = \frac{X}{X+Z} * 100$$

Figure 3.3: Precision and Recall

IV. Discussion and Future Work

Finding similar or related research paper is most difficult task in current scenario where there are large amount of data is stored in the databases. Recommender systems are solution to this problem and attracting researcher's to explore this area in past few years. This paper tries to solve this problem for recommending research articles by combining both content and collaborative filtering. In this paper we propose the bottom–up approach i.e. collaborative tagging technique to tag research articles and provide recommendations on the basis of similar tags, also two approach user to item and item to item is proposed for collaborative filtering . This system also proposed auto-tagging of the research paper by extracting information from the uploaded document. But the wrong tag by users can affect the efficiency of the system. This can be easily integrated into the current systems. The results can optimize by using the Meta heuristic approaches like Ant Colony Optimization. This paper also solve the cold start problem for new items where we do not know how to recommend that new item or what to recommend to that new user added to the system by using both content and collaborative filtering. But it is also not eliminated completely; more techniques can be integrated to the system to solve this problem.Also this system can be further improved by combining collaborative and content filtering techniques with collaborative tagging.

V. Conclusion

This paper proposed the hybrid system to recommend research papers to the user based on both collaborative and content filtering. This paper also purposes the bottom up approach to tag the research paper. This paper also proposes the two collaborative approaches i.e. user to item approach and item to item approach which further can be compared and the best can be integrated to the system to provide accurate, precise and optimized results. This paper proposed the technique which can be easily integrated to the current systems. Also this system can be further improved by using other collaborative and content filtering techniques.

REFERENCES

[1] F Ricci, L Rokach, B Shapira, "Introduction to recommender systems handbook," Springer US, 2011.
[2] Content based filtering http://recommender-systems.org/ content-based-filtering/
[3] Bhatia, Kapil. "Collaborative Tagging for Software Reuse."Computer Science\ & Engineering Department, Thapar Institute of Engineering\ & Technology, Deemed University (2006).
[4] Gipp, Bela, JöranBeel, and Christian Hentschel. "Scienstein: A research paper recommender system." In International Conference on Emerging Trends in Computing, pp. 309-315. 2009.
[5] Zheng, Nan, and Qiudan Li. "A recommender system based on tag and time information for social tagging systems." Expert Systems with Applications 38, no. 4 (2011): 4575-4587.
[6] Liang, Huizhi, YueXu, Yuefeng Li, and RichiNayak. "Collaborative filtering recommender systems using tag information." In Web Intelligence and Intelligent Agent Technology, 2008. WI-IAT'08. IEEE/WIC/ACM International Conference on, vol. 3, pp. 59-62. IEEE, 2008.
[7] Bedi, Punam, and Ravish Sharma. "Trust based recommender system using ant colony for trust computation." Expert Systems with Applications 39, no. 1 (2012): 1183-1190.
[8] Vander Wal, T. , "Explaining and Showing Broad and NarrowFolksonomies",Vanderwal.net,http://www.vanderwal .net/random/ entrysel.php?bl og=1635,2005
[9] Mathes, A. , "Folksonomies – Cooperative Classification and CommunicationThrough Shared Metadata", Adam Mathes.com,USAhttp://adammathes.com/

academic/computermediatedcommunication /folksonomies.pdf, 2004
[10]Basics of Search guide available at http://www.frame.org.uk/ Searching for Information/Basics.htm
[11] Bobadilla, Jesus, Fernando Ortega, Antonio Hernando, and Javier Alcalá. "Improving collaborative filtering recommender system results and performance using genetic algorithms." Knowledge-based systems 24, no. 8 (2011): 1310-1316
[12] Beel, Joeran, Marcel Genzmehr, Stefan Langer, Andreas Nürnberger, and Bela Gipp. "A comparative analysis of offline and online evaluations and discussion of research paper recommender system evaluation."In Proceedings of the International Workshop on Reproducibility and Replication in Recommender Systems Evaluation, pp. 7-14.ACM, 2013.
[13] Mobasher, Bamshad, Robin Burke, RunaBhaumik, and Chad Williams. "Toward trustworthy recommender systems: An analysis of attack models and algorithm robustness." ACM Transactions on Internet Technology (TOIT) 7, no. 4 (2007): 23.
[14] Ekstrand, Michael D., John T. Riedl, and Joseph A. Konstan. "Collaborative filtering recommender systems." Foundations and Trends in Human-Computer Interaction 4.2 (2011): 81-173.