International Journal of Engineering & Technology, 7 (3.34) (2018) 913-917

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Advanced Overlap Community Detection by Associative Rule Mining and Multi-View Ant Clustering

[1]T.P.Latchoumi, [2]Latha Parthiba , [3]T.P.Ezhilarasi, [4]Golda Dilip, [5]Hema Bhargavi

[1,5]Department of Computer Science and Engineering,VFSTR Deemed to be University, Vadlamudi, Andhra Pradesh, India
[2]Computer Science and Engineering, Community College, Pondicherry University, Pondicherry, India
[3]Computer Science and Engineering, RAAK College of Engineering, Pondicherry, India
[4]Department of Computer Science and Engineering, Saveetha School of Engineering, Chennai, India
**Corresponding author E-mail: [1]lax.senthil@gmail.com , [2]lathaparthiban@yahoo.com , [3]ezhil.9944@gmail.com , [4]goldadilip.sse@saveetha.com , [5]khemabhargavi@gmail.com*

Abstract

Recommendation systems mean based on customers interest techniques and tools are to generate the new products and services. The main issue in this recommendation system is the number of users are more and giving preference to their items takes more time. And also processing the date takes more time. Hence, clustering techniques are used for users into overlapping groups helps in the information sparsely issue and improve recommendation range. Next essential factor in this system is dynamic attention on users in which their importance varies. This work mainly concentrates on using the ant clustering technique to improvethe multi-view clustering method.

Keywords: *Clustering, Association Rule, Data Mining, Recommendation system.*

1. Introduction

Recommendation system (RS) is a widespread class for creating applications on the web that includes calculating the user's responses with options. It is found in several recent web applications that expose the user's huge item collections. The amount of data and things got very massive, in leading the overload of information. It turned into a big issue to discover what actually the user looks. Moreover the issues are solved by search engines partially; however, personalization of data was not specified. So the recommender systems provide the solution for this above-mentioned issues. Filtering and sorting the information is done by RS tools. And also, many researchers studied several facts of RS like semantics, scalability etc.
This paper organized as follows: Sect. 2 gives the background of the Recommender system. In sect. 3, gives the proposed architecture in multiway ant cluster approach in mining. In sect. 4, implementation of proposed work is described. Finally, in sect. 5, a conclusion is made.

2. Background

2.1 Recommendations System

It improves the accuracies of their recommendations by using several methods like context-aware, semantic and other approaches.

2.1.1. Content-based Recommender System

This work created the users profiles at the beginning. The information of a profile contains about a user and his taste. Rated items is defined on user's taste [11]. Normally, creation of users profile in this systems is makes by survey, and to avoid the new user issues it acquire the information about the users initially itself.

2.1.2. Collaborative Filtering Recommender System

This system is one of the greatest studied methods of recommender systems and described by Paul Resnick and Hal Varian [2]. The objective of this filtering is to find users in a community that share obligations.

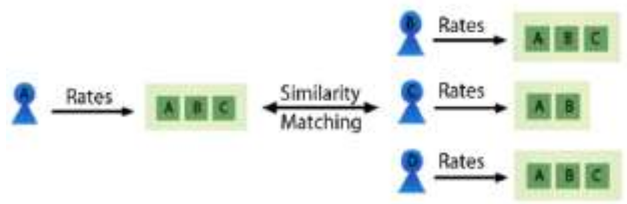


Fig. 1: User Based filtering

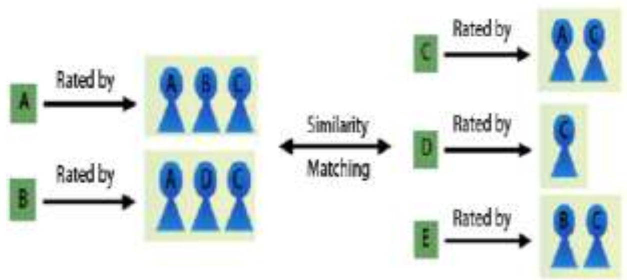


Fig. 2: Item-based filtering

2.1.3. Hybrid Recommender System

This approach is the combination of collaborative [1] and content based system to give the better results. This approach is used to avoid some restrictions and issues of pure recommender systems, like the cold-start issue.

2.1.4. Context-Aware Methods

This will give the information about the user'senvironment and situation details of he/she is in. Such information may perform additional important role in this recommendations of rating the items, and this information itself not providing complete information about which positions the ratings were specified by users. Few recommendations are appropriate to a user in evening and it doesn't match the same user preferences in the morning. And he/she would like to do when it's cold and entirely another choice when it's hot.

2.1.5. Semantic Based Method

Recommender system [3] or a computer couldn't understand the lexical and syntactical analysis but that should be understood by the user. To overcome these issues we are developing the new text mining method using semantic-based analysis. Recommendation systems with text mining methods are called as semantic based recommender systems [8].

3. Proposed Architecture

3.1 Multi-View Ant Cluster Method with Adaptive Association Rule Mining

Proposed system details has been explained here. A new approach called as multi-view clustering is proposed here. Quality is improved in this personalized recommendation system to the users [6], communities have generated the rules through adaptive association rule mining [12]. The input is the movie lens data set which consists of movie details, user details and the ratings given for the various movies by various users.

3.2 Community Generation Using Multi-View Ant Cluster Approach

Many model-based methods are developed for community generation in social networks [5]. Among them, matrix factorization is the most popular method which is widely used to address the issues in training the dataset.

3.2.1 Multi-View Ant Clustering

Ant clustering algorithm has been applied for forming user communities [11-13]. The algorithm describes the steps in ant cluster algorithm.

Algorithm 1
Step 1: Randomly select the one agent.
Step 2: Randomly identified direction on the grid was performed by the agent in step by step process.
Step 3: The agent (probabilistically) chooses whether to drop $p^*_{drop^{(i)}}$ (eq. 3.2) its data item in its present situation or in the instant neighborhood.
Step 4: Then it examines for a new data item immediately to pick up $p^*_{pick^{(i)}}$ (eq. 3.1).
Step 5: Step 3 and step 4 continues until it becomes successful picking process occurs.

For the picking and dropping decisions the following threshold formulae are used

$$p^*_{pick^{(i)}} = \begin{cases} 1.0 & \text{if } f*(i) \le 1.0 \\ \frac{1}{f*(i)^2} & \text{else} \end{cases} \quad 3.1$$

$$p^*_{drop^{(i)}} = \begin{cases} 1.0 & \text{if } f*(i) \ge 1.0 \\ f*(i)^4 & \text{else,} \end{cases} \quad 3.2$$

Where $f*(i)^n$ is a modified version of Lumer and Faieta's neighborhood function?

$$f*(i) = \begin{cases} \frac{1}{\sigma^2}\sum\left(1-\frac{d(i,j)}{\alpha}\right) & \text{if } (f*(i) > 0 \wedge \forall j \left(1-\frac{d(i,j)}{\alpha}\right) > 0) \\ 0 & \text{otherwise} \end{cases} \quad 3.3$$

Where $\frac{1}{\sigma^2}$is a neighborhood scaling parameter, α is a parameter scaling the dissimilarities within the neighborhood function$f*(i)$, d (i, j) a dissimilarity function.

Pseudo code
Input: dataset objects
Output: set of clusters

Randomly scatter o_i object on the grid file
for each agent a_j do
random_select_object (o_i)
pick_up_object o_i
place_agent a_j at randomly selected empty grid location
end for
{*Main loop*}
for t = 1 to tmax do
random_select_agent (a_j)
move_agent a_j to the new location
i = carried object (agent a_j)
Compute f*(oi) and p*drop (o_i)
if drop = True then
while pick = False do
i = random_select_object o
Compute f*(oi) and p*pick (o_i)
Pick_up_object o_i
end while
end if
end for
end

3.2.2. Overlapping Community Detection and Combination

After community formation using the ant cluster algorithm, the overlapping communities can be detected [7]. The overlapping proposition measure δ_{pq} is

$$\delta_{pq} = \beta * \frac{|c_p \cap c_q|}{|c_p \cup c_q|} + (1-\beta) * \frac{|NC_p \cap NC_q|}{|NC_p \cup NC_q|} \quad 3.4$$

For smaller communityδ_{pq}. is calculated using the formula 3.5.

$$\delta_{pq} = \frac{|c_p \cap c_q|}{\min(|c_p| \cup |c_q|)} \quad 3.5$$

3.2.3 Partition User based on Time

For partitioning the user based on time the algorithm is given below:

Algorithm 2: Time Based Partition
Step 1: Select a Community Cn.
Step 2: Analyse the different time period t1……tn available in the community Cn.
Step 3: Detect the overlapping time period and combine together.
Step 4: Output the finalized time frames t1 ….tn.

3.3Adaptive Association Rule Mining

3.3.1 Adaptive Frequent Itemset Mining

The minimum support [9] value plays an important role in affecting the algorithm's performance in this search process. Without increasing the computation time it must be self-adaptive to provide enough rules for high-quality rule mining. Then Sort frequent items in decreasing order based on their support value. The minimum support value is adjusted for each community in order to generate better rules for the recommendation.

3.3.2 Association Rule Mining

Support is used to indicate how frequently the items appear in the database. The confidence indicates the number of times the statements have been found to be true.

Pseudo code for Association Rule

Association rules mining procedure on a certain community.

Input: frequent item set Fsetk and dataset Dk for the kth community.

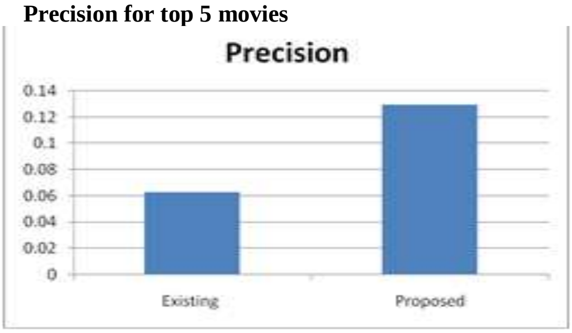
1. for each FsetK do
2. Forms a rule. A-> B
3. find the count of items A in dataset
4. Find the count of {A, B} in frequent item set.
5. A ⇒ B [s, c] support s: denotes the frequency of the rule within transactions.
6. Compute support (A ⇒ B [s, c]) = p(A ∪ B)
7. Confidence c: denotes the percentage of transactions containing A which contain also B. It is an estimation of conditioned probability.
8. compute confidence (A ⇒ B [s, c]) = p(B|A) = support (A,B)/support (A)
9. End for
10. End

4. Implementation

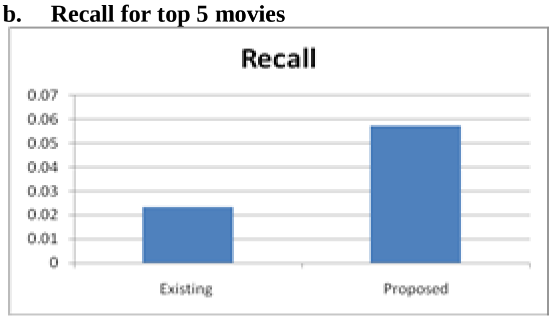
The experiments are conducted with the movie-lens dataset to assess the performance of the existing and proposed system with different values of N. where N is the Top N recommended items to the i^{th} user. The experimental result for the top 5 and top 10 items that are recommended to the users with the data's in the movie –lens dataset give the following values for precision, recall and f1 measure. The average Precision, Recall and F1 measure values of proposed framework for Movie lens dataset for 50 users are as follows

4.1 Evaluation Results for Top 5 Recommended Movies

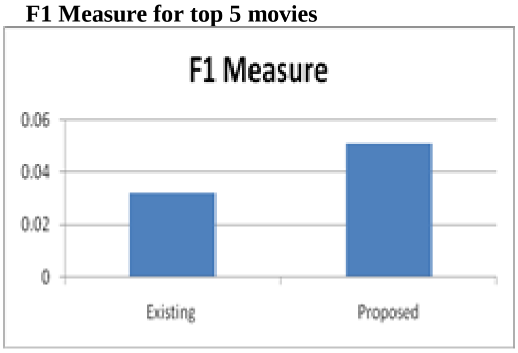
a. Precision for top 5 movies



b. Recall for top 5 movies



c. F1 Measure for top 5 movies



As shown in the graph, proposed algorithm performs better than the TOTAR at different N's (N=5) for the Movie Lens dataset, which indicates improvement in temporal overlapping community structure generation and improves the quality of recommendation.

5. Conclusion

In this paper, Multi-view Ant colony based clustering technique is used to detect the communities by using multiple dimensions such as rating similarity and trust between users. Hence using additional parameters can improve the quality of the communities. Another key factor is overlapping communities in which the user may appear in more than one community or the communities can have more similar nodes. In this case, it is necessary to combine the communities that overlap with each other in order to avoid redundancy of data. Another parameter time adaptive filtering is also included in this paper in an efficient way. This will address the issues present in the existing system and reduce the processing time and give most desirable result in recommendation system.

References

[1] Vassillios statho, Ioannis Konstas poulos and Joemon M.Jose, "On social networks and collaborative recommendation", SIGIR '09 Proceedings of the 32^{nd} international ACM SIGIR conference, USA ©2009.

[2] P. Resnick and H. R. Varian, "Recommender systems and its methods",Communications of the ACM, vol. 40, pp. 56–58, March 1997.

[3] Bahram Amini, Roliana Ibrahim and Mohd Shahizan Othman, "Discovering the impact of knowledge in recommender systems", International Journal of Computer Science & Engineering Survey (IJCSES) Vol.2, No.3, August 2011.

[4] Giancarlo Ruffo and Rossano Schifanella, "Evaluating Peer-to-Peer Recommender Systems that Exploit Spontaneous Affinities", ACM symposium on Applied computing, Pages 1574-1578, USA, 2007.

[5] Shaghayegh Sahebi and William W. Cohen, "Community-Based Recommendations: a Solution to the Cold Start Problem", workshop on Recommender systems and the social Web(RSWEB),ACM RecSys'11,Chicago,23 October,2011.

[6] Xiaofang Ding, Zhixiao Wang*, Shaoda Chen and Ying Huang,"Community-based Collaborative Filtering Recommendation Algorithm",International Journal of Hybrid Information Technology, Vol.8, pp.149-158, No.2 (2015).

[7] Brian Lee Yung Rowe ,Zato Novo, "Using social network graph

analysis for interest Detection", International journal of Data Analytics, School of Professional Studies, CUNY,October 2, 2014.

[8] Maryam Fatemi, Laurissa Tokarchuk, "A Community Based Social recommender System for Individuals & Groups", 978-0-7695-5137-1/2013 IEEE,DOI 10.1109/SocialCom.2013.

[9] Isabel M. Kloumann,"Community membership identification from small seed sets",20thACM international conference on Knowledge discovery and data mining, Pages 1366-1375, New York, USA, 2014.

[10] Badrul M. Sarwar, George Karypis, Joseph Konstan, and John Riedl,"Recommender Systems for Large scale E-Commerce: Scalable Neighborhood Formation Using Clustering", GroupLens Research Group, Minneapolis, MN 55455, USA, 2004.

[11] K. Vijayakumar and C. Arun, "Continuous Security Assessment of Applications in Cloud Environment", International Journal of Control Theory and Applications, ISSN: 0974-5645 volume No. 9(36), Sep 2016, Page No. 533-541

[12] K. Vijayakumar C. Arun, Continuous security assessment of cloud based applications using distributed hashing algorithm in SDLC, Cluster Computing DOI 10.1007/s10586-017-1176-x, Sept 2017.

[13] Purushottam Sharma and Kanak Saxena, "Temporal Weighted Association Rule Mining for Classification", International Journal of Computer Theory and Engineering, Vol. 4, No. 5, October 2012.

[14] C.-H. Lee, Y.-H. Kim, P.-K. Rhee,Web personalization expert with combining collaborative filtering and association rule mining technique", ACM Conference onExpert Systems, Volume 21, Issue 3, October 2001, Pages 131–137.

[15] Urszula Boryczka, "Ant Clustering Algorithm", Intelligent Information Systems, ISBN 978-83-60434-44-4, pages 377–386, 2008.

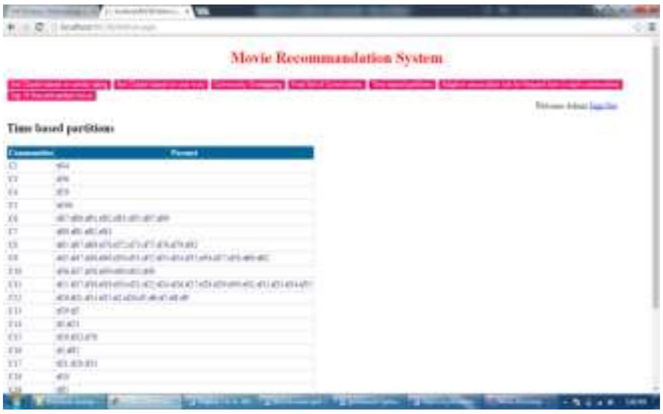
APPENDIX

Screen 1: Ant Clustering Based on Similar Ratings

Screen 2: Ant Clustering based on user Trust

Screen 3: Overlapping communities

Screen 4: Final set of communities

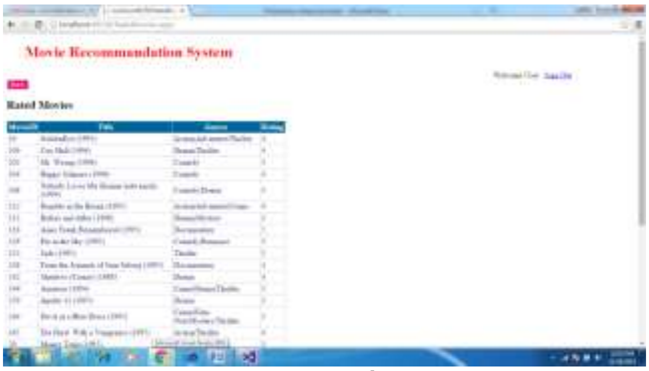


Screen 5: Time Based Partition

Screen 6: Frequent movies in each community

Screen 7: Top 10 recommend movies in each community

Screen 8: Personalized recommendation



Screen 9: User rated movies