

Alleviating the Cold Start Problem in Recommender Systems Based on Modularity Maximization Community Detection Algorithm

S. Vairachilai¹, M. K. Kavithadevi², M. Raja¹

¹Department of CSE, Faculty of Science & Technology, IFHE Hyderabad, Telangana, India

²Department of CSE, Thiagarajar College of Engineering, Madurai, Tamil Nadu, India

Email: vairachilai2676@gmail.com

Received 22 March 2016; accepted 20 April 2016; published 7 June 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Recommender system (RS) has become a very important factor in many eCommerce sites. In our daily life, we rely on the recommendation from other persons either by word of mouth, recommendation letters, movie, item and book reviews printed in newspapers, etc. The typical Recommender Systems are software tools and techniques that provide support to people by identifying interesting products and services in online store. It also provides a recommendation for certain users who search for the recommendations. The most important open challenge in Collaborative filtering recommender system is the cold start problem. If the adequate or sufficient information is not available for a new item or users, the recommender system runs into the cold start problem. To increase the usefulness of collaborative recommender systems, it could be desirable to eliminate the challenge such as cold start problem. Revealing the community structures is crucial to understand and more important with the increasing popularity of online social networks. The community detection is a key issue in social network analysis in which nodes of the communities are tightly connected each other and loosely connected between other communities. Many algorithms like Givan-Newman algorithm, modularity maximization, leading eigenvector, walk trap, etc., are used to detect the communities in the networks. To test the community division is meaningful we define a quality function called modularity. Modularity is that the links within a community are higher than the expected links in those communities. In this paper, we try to give a solution to the cold-start problem based on community detection algorithm that extracts the community from the social networks and identifies the similar users on that network. Hence, within the proposed work several intrinsic details are taken as a rule of thumb to boost the results higher. Moreover, the simulation experiment was taken to solve the cold start problem.

How to cite this paper: Vairachilai, S., Kavithadevi, M.K. and Raja, M. (2016) Alleviating the Cold Start Problem in Recommender Systems Based on Modularity Maximization Community Detection Algorithm. *Circuits and Systems*, 7, 1268-1279.

<http://dx.doi.org/10.4236/cs.2016.78111>

Keywords

Collaborative Recommender Systems, Cold Start Problem, Community Detection, Pearson Correlation Coefficient

1. Introduction

The goal of a Recommender System [1] is to generate meaningful recommendations to a collection of users for items and services that might interest them which is based on the Information Filtering (IF) Mechanism. In general, the users are always interested in searching information, items like Films, Audio files, Video files, News, Images, Music and Web records, etc. Typically, a recommender system equates few reference features from the user's profile and looking forward to the rating that a user would give to an item they had not yet considered. These features may be from the information item (Content-Based filtering approach) or the candidate's social environment (Collaborative filtering approach). Content Based Filtering is making the recommendations by using information about the items. It will recommend a few items to the user if the recommended items are as same as the content of items that the user liked earlier. In particular, different user items are compared with items in which formerly rated by the user and the finest identical items are recommended. The collaborative filtering approach is based on collecting and analyzing the information in huge quantity on users' behaviors, actions or favorites and calculating what user's desire based on their comparison with other candidates. The importance of the collaborative filtering method is that it does not depend on system analyzable content. The term collaborative filtering was introduced in the first commercial recommender system; called Tapestry [2]. It was developed at the Xerox Palo Alto Research Centre and is defined in a 1992 issue of the Communications of the ACM. The Tapestry RS was used to increase the use of electronic mail, which resulted in users being inundated by a huge stream of incoming documents. The main idea of Collaborative recommendation approaches [3] is to exploit information about the previous activities or opinions of an existing user community for predicting which items the current user of the system will most probably like or be interested in. Because of this collaborative filtering refers to people-to-people correlation [4]. Collaborative filtering (CF) is considered to be the most popular and widely implemented technique in RS. Though they are domain independent, requires minimal information about user and item features, it can still achieve an accurate prediction rate [5]. In the recent research, the system has been demonstrated as a Hybrid recommender approach which is the combination of various inputs and/or combination of a different mechanism.

Collaborative filtering (CF) has some challenges while collecting and analyzing the information from the database. The challenges are: 1) Sparsity: There are many items to be recommended, even if there are many users, the user/rating matrix is sparse, and it is very difficult to find users that have rated the same items; 2) Cold Start: Difficult to make recommendations to the new users that have not rated any items yet and difficult to deal with items that have not been rated or bought yet. Weike Pan, Manos Papagelis *et al.* [6] [7] has explained that collaborative filtering is restricted owing to the sparsity problem in the recommender systems. Using trust inferences, Manos Papagelis proposes a method for improving sparsity. Trust implications are transitive relations between candidates in the environment of an underlying social network and are reasonable sources of supplementary data that help e-commerce with the sparsity. When existing statistics are inadequate for finding out the similar users, the sparsity problem occurs and in general, it is a major problem that restricts the quality of recommendations of collaborative filtering [8].

Asela Gunawardana *et al.* [9] describes an algorithm for training the model and provides experimental outputs evaluating the model in two cold start recommendation systems. In order to improve the cold-start problem, system bonds the parameters of the Boltzmann machine through restrictions in terms of the content information associated with the items. Boltzmann machines are very weak to cold start problems and do not make use of content information. To produce synthetic files for training recommender systems, this work attempts have been made to use a various heuristic method based on the content information. Sahebi S., Cohen W. *et al.* proposed a solution to the cold- start problem for the user based on community [10]. M. Claypool, P. Melville *et al.* proposed a solution for the cold start problem by asking them a series of questions about their tastes or some survey about the items in order to get rating [11] [12].

In social networks, each entity is represented as a node in the network, and the communication between two

nodes has been represented as an edge. The analysis of social networks is important in many fields such as recommender systems [13], sociology [14], email communication [15], etc. The important property that is found in many social networks is community structure. The communication within the community is more than the other communities. **Figure 1** displays the community structure, with four groups of nodes tightly connected each other and loosely connected between other communities. The communities in these networks are beneficial in many applications such as marketing and advertising [16], finding out a similar customer in e-commerce site like shopclues®, flipkart®, jabong®, etc., that enhances the business opportunities for retailer [17], statistical physics and applied mathematics. The algorithms for finding out communities in networks have been proposed and improved in the recent years. The well-known algorithm is divisive hierarchical clustering algorithm proposed by Girvan and Newman [18], called GN algorithm. This algorithm defines the betweenness of an edge is the number of geodesic distances between pairs of nodes that run through that edge of the network. The modularity maximization community detection algorithm (MMCDA) works well in simulation test case and real world scenario [19]-[21] among the community detection algorithm. The most widely used and accepted algorithm for detecting the communities in the social network was MMCDA. The proposed cold start community-based algorithm exhibits the dynamic behavior in the network. The dynamic behavior of the network used in various fields of applications: e-mail communication [22], mobile communication in cellular networks [23], co-authorship pattern [24] and Recommender System [1]. The rest of the paper is organized as follows: Section 2 explains the modularity maximization community detection algorithm. Section 3 explains the cold start problem in the recommender system. Section 4 explains the problem statement. Section 5 explains the proposed algorithm. Section 6 explains the result and discussion. Conclusion and Future work are illustrated in Section 7.

2. Modularity Maximization Community Detection Algorithm

In order to evaluate the community division, we define the quality function called modularity [18]. Modularity means that the links within a community are higher than the expected links in that community. The modularity can be either positive or negative, the positive values represent the possible presence of community structure and the negative values represent the non-existence of community structure. If the modularity values between 0.3 and 0.7 on the real world network, then it has the strongest community structure. Equation (1), where A_{ij} is the link between node i and j , the value of A_{ij} is 1 if there exists a link between i to j , otherwise zero, $k_i k_j / 2m$ is the expected number of links between i and j , $\delta(C_i C_j)$ indicates in which community the node i belongs to and m are the total numbers of links in the community. Equation (2), where e_{ii} is the actual number of links in the community and a_i is the expected number of links for that community. Equation (3), where e_{ij} is the percentage of edges between communities i & j . Therefore, the modularity (Q) functions are defined as follows:

$$Q = \frac{1}{2m} \sum_{ij} \left[A_{ij} - \frac{K_i K_j}{2m} \right] \delta(C_i C_j) \tag{1}$$

$$Q = \sum_i e_{ii} - (a_i)^2 \tag{2}$$

$$a_i = \sum_i e_{ij} \tag{3}$$

To explain the calculation of the modularity of a community in **Figure 2** shows a small example. For a

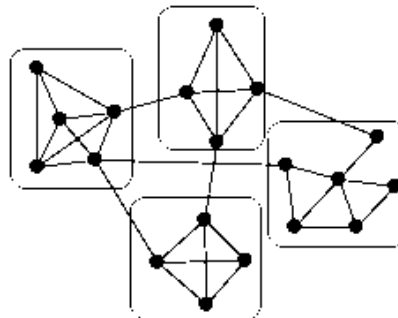
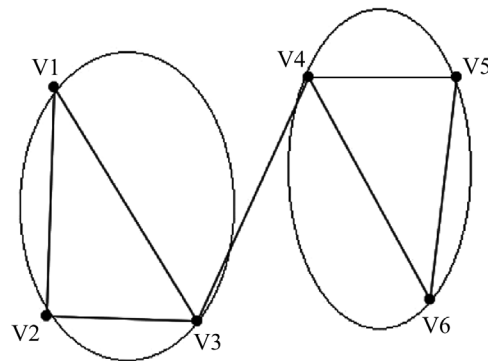


Figure 1. A simple graph with four communities in a social network.



	C1			C2		
	V1	V2	V3	V4	V5	V6
V1	0	1	1	0	0	0
V2	1	0	1	0	0	0
V3	1	1	0	1	0	0
V4	0	0	1	0	1	1
V5	0	0	0	1	0	1
V6	0	0	0	1	1	0

Figure 2. Modularity computation for a small network with six vertices with two communities.

community division of the example graph with 6 vertices into the clusters $C_1 = \{v_1, v_2, v_3\}$, $C_2 = \{v_4, v_5, v_6\}$ the values of e_{ij} are the total sums of the matrix elements belong to a pair of C_{1i} and C_2 divided by the total sum of all matrix elements: $Q(C_1) = \left(\frac{6}{14}\right) - \left(\frac{7}{14}\right)^2 = 0.17857$, $Q(C_2) = \left(\frac{6}{14}\right) - \left(\frac{7}{14}\right)^2 = 0.17857$. The modularity of the partition of the example graph into the two communities is $Q = 0.357$. The overall flow for the modularity maximization algorithm shown in **Figure 3**.

Algorithm 1 Pseudo Code of Modularity Maximization Method

Input: Undirected network G

Output: Community is assigned to each node with maximum modularity (Q)

Repeat

- Initially, each node belongs to its own communities.
- Compute the actual number of links in the community one.
- Divide the actual number of links with the total number of links in the network.
- Compute the expected number of links to the same community.
- Subtract the actual number of links in the community from the expected number of links to the same community.
- Calculate the increase and decrease of modularity measure Q for all possible communities.
- Merge the pairs with the greatest increase (or smallest decrease) in Q.
- Repeat the process from the step 2 for the next community.
- Update the modularity value by adding the modularity value for the two communities.
- Until return community with high modularity.

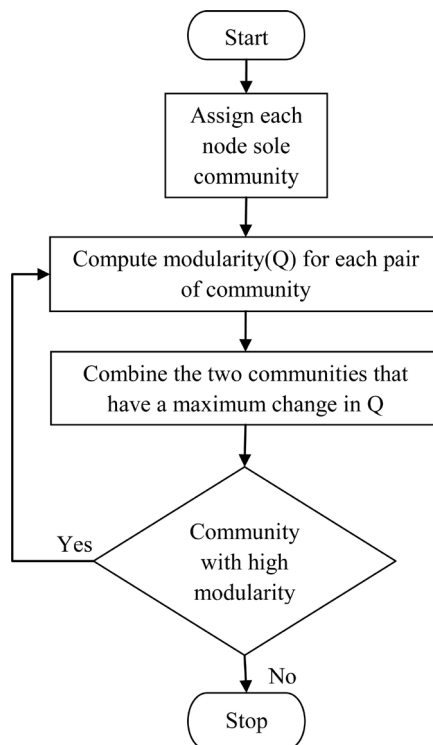


Figure 3. Flowchart for modularity maximization method.

3. Cold Start Problem in Recommender System

The cold-start problem can be observed as one of the special problems in the recommender system. The cold-start occurs in Recommender Systems (RS) due to the lack of information about the users and the items [25] [26]. The User-Cold Start Problem occurs, new users arrive frequently, or may seem new when they don't log into RS. In particular; users that never rated an item do not connect with any other user in the recommender system. Therefore, they do not belong to any relevant community in the system, so the preference of the user is not known to the system. Without any information about the user, the system is not able to guess the user's preferences and generate recommendations. Therefore, the new user cannot get any recommendation until she/he rates a few items. The collaborative filtering requires the historic rating of the user to calculate the similarities for the determination of the neighborhood. The Item-Cold Start Problem occurs when a new item enters into the system; the new items appear frequently an e-commerce site. Because it is a new product, it is not known to the users of the system. Since the new item is rated by a substantial number of users; the recommender system would not be able to recommend it to the users. These problems lead to significant degradation of recommending quality. In traditional collaborative filtering, the cosine similarity is used to measure the similarity between items. The basic cosine similarity measure does not take the difference in the average rating behavior of the user; this was solved by adjusted cosine similarity measure [2]. To make a prediction for an item not rated by the active user done in two ways: 1) User-based nearest neighbor's recommendation; 2) Item-based nearest neighbor's recommendation. Equation (4), where U be the set of users that rated both items i and j . Equation (5), where p is the product and u is the user of that product. The Equation (4) used to find the similarity between the items and the Equation (5) used to calculate the predicted for product p of the user.

$$sim(i, j) = \frac{\sum_{u \in U} (r_{u,i} - \bar{r}_u)(r_{u,j} - \bar{r}_u)}{\sqrt{\sum_{u \in U} (r_{u,i} - \bar{r}_u)^2} \sqrt{\sum_{u \in U} (r_{u,j} - \bar{r}_u)^2}} \quad (4)$$

$$pred(u, p) = \frac{\sum_{i \in ratedItems(u)} sim(i, p) * r_{u,i}}{\sum_{i \in ratedItems(a)} sim(i, p)} \quad (5)$$

In traditional collaborative recommender system, one of the most complicated tasks in the recommender system is to determine the similarity among the users. The correlation coefficient is used to measure the similarity among the users. A correlation coefficient is a statistical measure of the degree to which changes in the value of one variable predict change in the value of another. The Pearson correlation coefficient named after Karl Pearson (1857-1936), it is a measure of the linear correlation between active users and all other users who rated the target item, gives a value between +1 and -1 inclusive. The value 1 is the total positive correlation, 0 is no correlation, and -1 is a negative correlation. The assumption of this method is if the user had similar taste in the past will have the similar taste in the future and also a user preference to remain constant and steady over time. Pearson correlation coefficient can be easily implemented and can achieve high accuracy [27]. The input for the CF recommender system is the user- movie rating matrix. The ratings are stored in the database. There are several other similarity measures can be used, including Spearman rank correlation, Kendall's correlation, mean squared differences, and entropy. Equation (6), where a is an active user, u is the neighbor to compare, P is the set of items, $r_{a,p}$ is the rating of the active user a to the product p , $r_{u,p}$ is the rating of the neighbor u to the product p , n is the total number of items, \bar{r}_a is the average rating of the active user, and \bar{r}_u is the average rating of the neighbor. The Equation (6) used to find the similarity between the users and the Equation (8) used to calculate the predicted for the product p of the user.

$$sim(a, u) = \frac{\sum_{p \in P} (r_{a,p} - \bar{r}_a)(r_{u,p} - \bar{r}_u)}{\sqrt{\sum_{p \in P} (r_{a,p} - \bar{r}_a)^2} \sqrt{\sum_{p \in P} (r_{u,p} - \bar{r}_u)^2}} \quad (6)$$

$$\bar{r}_a = \frac{\sum_{a=1}^n r_a}{n}; \quad \bar{r}_u = \frac{\sum_{u=1}^n r_u}{n} \quad (7)$$

$$pred(a, p) = \bar{r}_a + \frac{\sum_{u \in N} sim(a, u) * (r_{u,p} - \bar{r}_u)}{\sum_{u \in N} sim(a, u)} \quad (8)$$

4. Problem Statement

By applying community detection techniques in online e-commerce network, we can get 1) group of users similar to the target user; 2) predict rating for the active user based on similar community user; 3) recommend a product/service to the new user; 4) recommend a new product such as new movie to the users in RS. For the Item-Cold Start problem, detect the community based on the user-movie matrix, then creates a short description of the new movie. This description can be used as the seed to grow around the communities and its description as the recommendation for the new movie. For the User Cold-Start problem, the user is new to the recommender system (X), but his/her profile is available in the recommender systems (Y). The modularity maximization algorithm detects the community based on the user-movie matrix available in recommender systems (Y). We can use the external community profile available in recommender system Y such as amazon[®] to recommend the item for the new user in a recommender system in X such as ebay[®]. Community detection is used to find out a similar group of user to new target user in the system. We can assume that the two users are similar in terms of book ratings, to have the similar taste in the movie also; or if two of them can be friends. We can make use of this information to predict that the new user's interest in another dimension can be used as a solution to the user cold-start problem. **Table 1** shows an example rating matrix for five users and five movies on a 1-5 scale, with 5 the highest rating in a movie recommender system; cells marked '?' indicate unknown values (the user has not rated that movie). The input data for the recommendation algorithms based on the recommender system types shown in **Table 2**. The community with high based on the Cold Start Community-Based Algorithm shown in **Figure 4**.

5. Proposed Algorithm—Cold Start Community-Based Algorithm (CSCBA)

Input: User-Movie Rating Database

Output: Community is Extracted

- Initially, extract the user-movie rating information from the Movie rating database.
- Creates a matrix from the given set of values.

Table 1. Sample rating database for collaborative recommendation systems of movies on a 1-5 scale.

Movies/users	Movie 1	Movie 2	Movie 3	Movie 4	Movie 5
Active User	5	3	1	2	?
User1	4	3	1	5	3
User2	?	?	4	4	5
User3	5	3	4	1	4
User5	1	3	?	4	?

Table 2. Input data for the recommendation algorithms based on the recommender system types.

Recommender System	User Profile	Contextual parameters	Community Data	Product features	Knowledge Model
Collaborative	Yes	Yes	Yes	No	No
Content-Based	Yes	Yes	No	Yes	No
Knowledge-based	Yes	Yes	No	Yes	Yes

Source from [3]: Dietmar Jannach, Markus Zanker, Alexander Felfernig, Gerhard Friedrich (2010), An Introduction Recommender System (Page number 128).

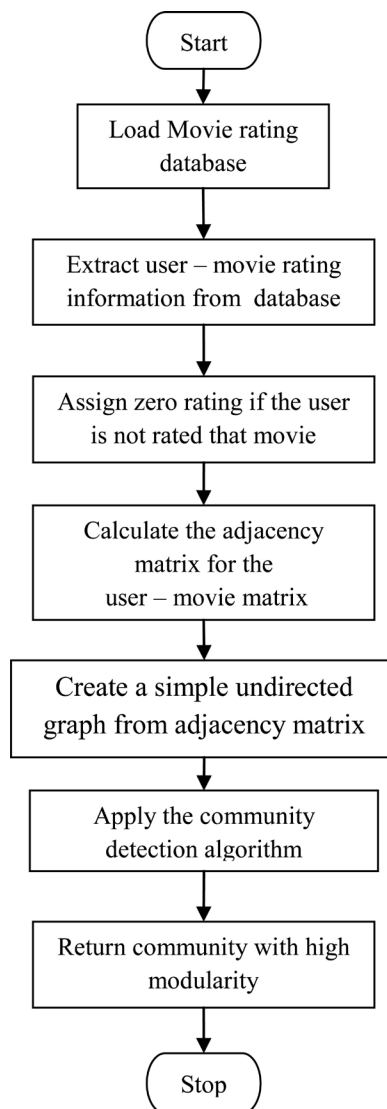


Figure 4. Flowchart for cold start community-based algorithm.

- If the user is not rated for any movie assign zero rating for that movie.
- Find the adjacency matrix for the user-movie matrix.
- Create graphs from adjacency matrix.
- Find the simple graphs that do not contain loops and multiple edges.
- The undirected graph is extracted from the simple graph.
- Apply the community detection algorithm.
- Return community with high modularity.

6. Result and Discussion

The Movie Lense rating data set was collected from the website movielens.umn.edu. The data set contains about 100,000 ratings (1-5) from 943 users on 1664 movies. To simulate the User Cold-Start problem, we removed all the other information from the rating data set like the description about the movie, time, etc. We performed the community detection in the user-movie rating matrix. After detecting the community, we can find out the relevant movies to the users. This can be used for the new user in another e-commerce site to recommend the movies. Similarly for Item Cold-Start problem, detecting the community based on the user-movie matrix, creates a short description of the new movie. This description used as a seed to grow around the communities and its description as a recommendation for the new movie. Based on this observation, in the prediction formula to predict the unknown rating for the active user can be calculated by using the community in the network. The developed algorithm is applied in the network that was extracted from the movie database. The output of the community detected is shown in [Figure 5](#). The short description of the new item is fed into the each community shown in [Table 3](#). The relevant movies per community based on Cold Start Community-Based Algorithm shown in [Table 3](#). The social networks are dynamic graphs. Dynamic graph properties such as the number of graph vertices, graph edges, and connections between them are determined in a random way. The cold start community-based algorithm shows the dynamic behavior while implementing in the movie-user database. In such cases, the edges may be rapidly added to and deleted from the network, as a result of which the community member of the network may vary significantly at any point of time. The Cold Start Community-Based Algorithm gives the current community structure. The dynamic nature of such structure needs to be studied over a dynamic network. The

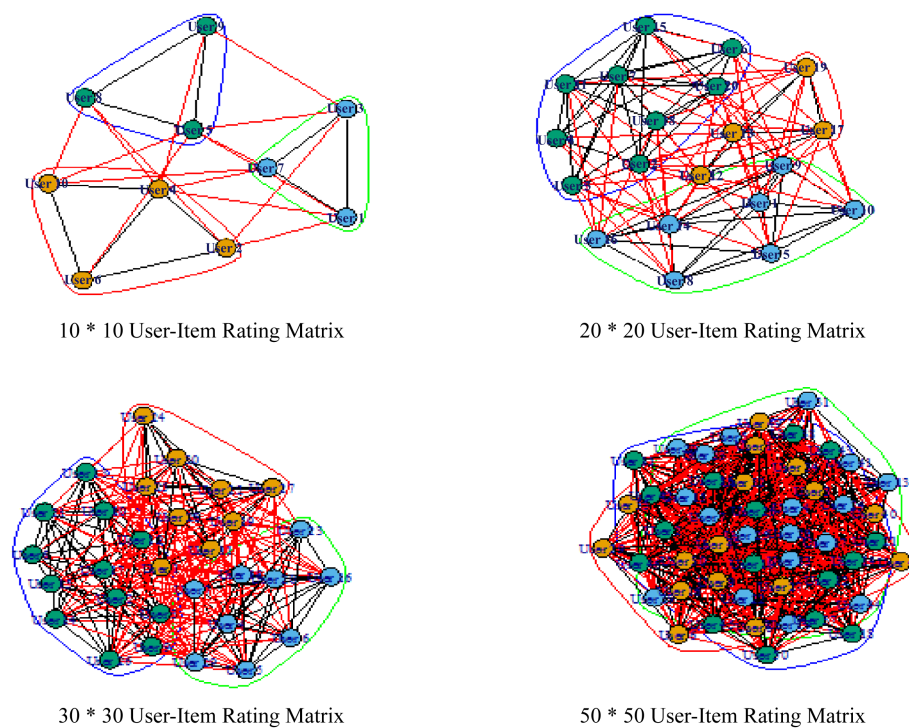


Figure 5. Community detected based on modularity maximization community detection algorithm. Each community shown in different colors.

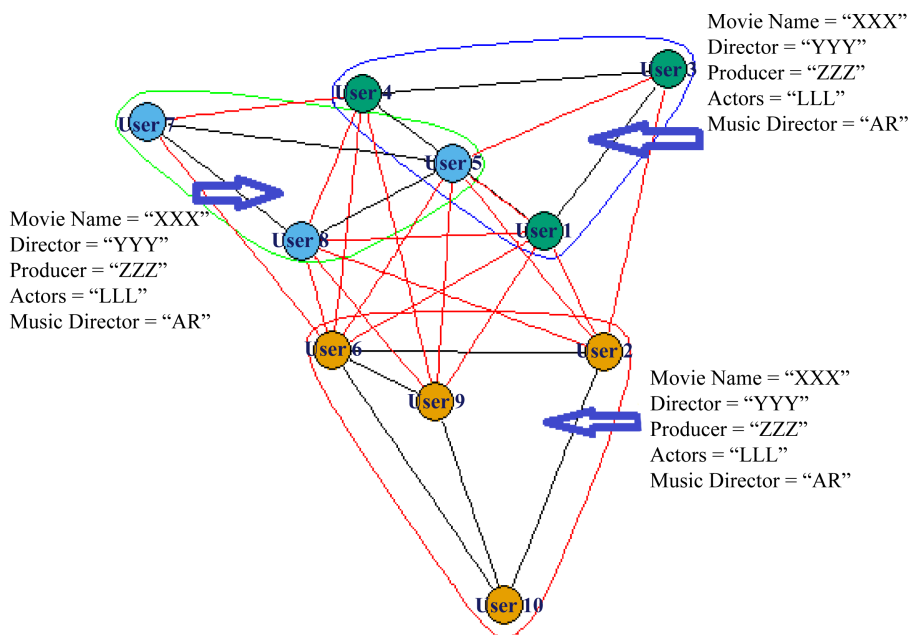


Figure 6. Community detected based on modularity maximization community detection algorithm. Short descriptions about each item to feed into the community.

Table 3. Relevant movies per community based on cold start community-based algorithm.

Community Id	Name of the Movie
1	Butch Cassidy and the Sundance Kid, Glory, Lawrence of Arabia, Casablanca, Cool Hand Luke, Harold and Maude
2	Return of the Jedi, Silence of the Lambs, Raising Arizona, Lawrence of Arabia, Princess Bride, Raiders of the Lost Ark.
3	Fargo, Star Wars, Titanic, Silence of the Lambs, In the Company of Men, Leaving Las Vegas
4	Fargo, Return of the Jedi, Godfather, Shawshank Redemption, Leaving Las Vegas, Empire Strikes Back
5	This Is Spinal Tap, Lone Star, Postino, Cinema Paradiso, Three Colors: Red, Fifth Element.
6	Schindler’s List, Babe, Princess Bride, Killing Fields, Godfather: Part II, This Is Spinal Tap, Cool Hand Luke
7	Shawshank Redemption, Return of the Jed, Graduate, Empire Strikes Back, Fargo, Godfather: Part II

following is the formal definition of the dynamic network. Given a dynamic network $\langle G_1, \dots, G_T \rangle$ consisting of simple graphs $G_T = \langle G_t, E_t \rangle$ over a uniquely labeled set of vertices $\{V_t\}$ and a set of edges E_t . This is much more appropriate when the relevant movie for any community will vary over the period of time. Using them, we can easily find out the similar user in the network. The proposed algorithm is tested with the other community detection algorithm. The comparison result is shown in **Figure 7** and **Figure 8**. The total number of community and the single community structure in that community shown in **Table 4**. The Equation (9) is derived from the Equation (8). Equation (9), where r_a is the average rating of the active user, $\sum_{u \in \text{community}(a)} \text{sim}(a, u)$ is the total positive correlation value based on community detected by the Cold Start Community-Based Algorithm, $r_{u,p}$ is the rating of the neighbor u to the product p and r_u is the average rating of the neighbor [3]. A formula for computing a prediction for the rating of the user a for movie p based on the community.

$$\text{pred}(a, p) = \bar{r}_a + \sum_{u \in \text{community}(a)} \text{sim}(a, u) * (r_{u,p} - \bar{r}_u) \tag{9}$$

7. Conclusion and Future Work

In this paper, we have analyzed the cold-start problem in the recommender system. One of the most important properties of complex networks is community structure. The nodes of the communities are tightly connected within the community and loosely connected among other communities. The community detection algorithm

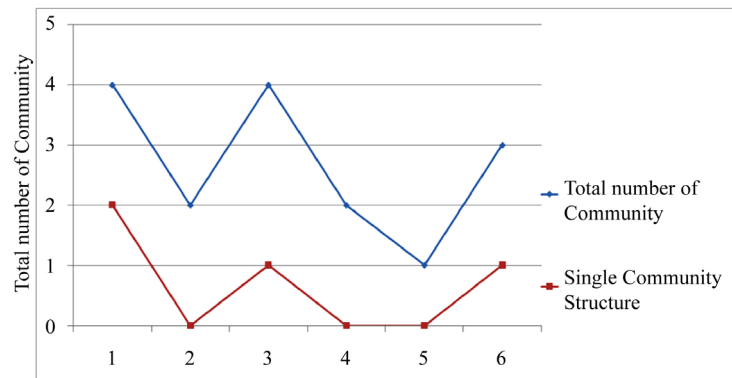


Figure 7. Comparison chart based on the community detected by community detection algorithms.

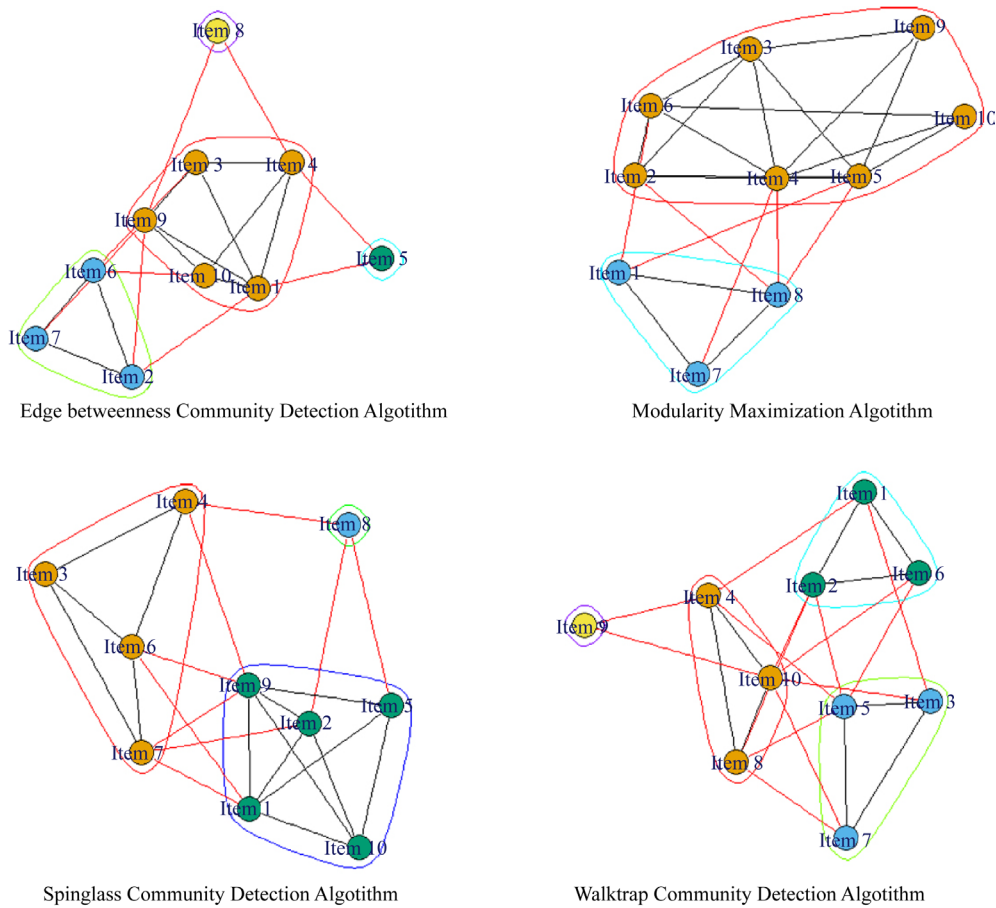


Figure 8. Comparison of community detection algorithms based on the community detected.

Table 4. Comparison of community detection algorithms based on the community detected.

S.No	Algorithm	Total number of Community	Single Community Structure
1	Edge betweenness	4	2
2	Modularity Maximization	2	0
3	Walktrap	4	1
4	Leading eigenvector	2	0
5	label. propagation	1	0
6	Spin glass	3	1

used to detect the community in the complex network. The cold start community-based algorithm alleviates the cold start problem in collaborative filtering recommender system. Among all other community detection algorithms, cold-start community-based algorithm gives a better result. This algorithm extracts the community from the network and finds out the community member within the community. It extracts the community based on current community structure. It shows the dynamic behavior that some user added or deleted in the network. The real-time networks such as social networks are dynamic in nature, as the nodes and edges keep on getting added and deleted from the network. This method is also used to find out the similarity value and predict the rating for the unrated item also. In future, the present work may be extended to study the dynamic nature of community structure by analyzing social media network as a dynamic network and also to find out a solution to the sparsity problem in collaborative filtering based on community.

References

- [1] Melville, P. and Sindhwan, V. (2010) Recommender Systems. IBM T.J. Watson Research Centre, Yorktown Heights, NY. <http://vikas.sindhwan.org/recommender.pdf>
- [2] Goldberg, D., Nichols, D., Oki, B.M. and Terry, D. (1992) Using Collaborative Filtering to Weave an Information Tapestry. *Communications of the ACM*, **35**, 61-70. https://www.ischool.utexas.edu/~i385d/readings/Goldberg_UsingCollaborative_92.pdf <http://dx.doi.org/10.1145/138859.138867>
- [3] Jannach, D., Zanker, M., Felfernig, A. and Friedrich, G. (2010) Recommender System: An Introduction.
- [4] Schafer, J.B., Konstan, J.A. and Riedl, J. (2001) E-Commerce Recommendation Applications. *Data Mining and Knowledge Discovery*, **5**, 115-153. <http://dx.doi.org/10.1023/A:1009804230409>
- [5] Pilaszy, I. and Tikk, D. (2009) Recommending New Movies: Even a Few Ratings Are More Valuable than Metadata. *Third ACM Conference on Recommender Systems (RecSys'09)*, New York, 23-25 October 2009, 93-100. <http://dx.doi.org/10.1145/1639714.1639731>
- [6] Papagelis, M. (2005) Alleviating the Sparsity Problem of Collaborative Filtering Using Trust Inferences. *Proceedings of the Third International Conference on Trust Management*, **3477**, 224-239. http://dx.doi.org/10.1007/11429760_16
- [7] Pan, W.K. (2010) Transfer Learning in Collaborative Filtering for Sparsity Reduction. *Proceedings of the 24th AAAI (Association for the Advancement of Artificial Intelligence) Conference on Artificial Intelligence (AAAI-10)*.
- [8] Huang, Z., Chen, H. and Zeng, D. (2004) Applying Associative Retrieval Techniques to Alleviate the Sparsity Problem in Collaborating Filtering. *Transaction on Information Systems*, No. 1, 116-142.
- [9] Gunawardana, A. (2008) Tied Boltzmann Machines for Cold-Start Recommendations. *Proceedings of the 2008 ACM Conference on Recommender Systems*, Lausanne, 23-25 October 2008, 8 p. http://www.msr-waypoint.net/pubs/69521/gunawardana08_tied_boltz_machin_cold_start_recom.pdf
- [10] Sahebi, S. and Cohen, W. (2011) Community-Based Recommendations: A Solution to the Cold-Start Problem. *Proceedings of the Workshop on Recommender Systems and the Social*, El Paso. <http://www.dcs.warwick.ac.uk/~ssanand/RSWeb11/7Sahebi.pdf>
- [11] Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Nets, D. and Sartin, M. (1999) Combining Content-Based and Collaborative Filters on an Online Newspaper. *Proceedings of ACM SIGIR Workshop on Recommender Systems*, Berkeley, 19 August 1999, 1-11. <http://web.cs.wpi.edu/~claypool/papers/content-collab/content-collab.pdf>
- [12] Melville, P., Mooney, R.J. and Nagarajan, R. (2002) Content-Boosted Collaborative Filtering for Improved Recommendations. *Proceedings of the National Conference on Artificial Intelligence*, AAAI Press, MIT Press, Menlo Park, CA, Cambridge, MA, London, 187-192. <http://www.cs.utexas.edu/~ml/papers/cbcf-aaai-02.pdf>
- [13] Palau, J., Montaner, M., Lopez, B. and de la Rosa, J.L. (2004) Collaboration Analysis of Recommender Systems Using Social Networks. *Proceedings of the 8th International Workshop on Cooperative Information Agents CIA'04*, **3191**, 137-151. http://dx.doi.org/10.1007/978-3-540-30104-2_11
- [14] Wasserman, S. and Faust, K. (1994) Social Network Analysis: Methods and Applications. Cambridge University Press. <http://www.icpsr.umich.edu/summerprog/biblio/2012/Social%20Network%20Analysis%20An%20Introduction.pdf> <http://dx.doi.org/10.1017/CBO9780511815478>
- [15] Tyler, J.R., Wilkinson, D.M. and Huberman, B.A. (2003) Email as Spectroscopy: Automated Discovery of a Community Structure within Organizations. *Communities and Technologies*, 81-96. http://dx.doi.org/10.1007/978-94-017-0115-0_5
- [16] Kempe, D., Kleinberg, J. and Tardos, E. (2003) Maximizing the Spread of Influence through a Social Network, *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '03, 24-

- 27 August 2003, Washington DC, 137-146. <https://www.cs.cornell.edu/home/kleinber/kdd03-inf.pdf>
- [17] Reddy, P.K., Kitsuregawa, M., Sreekanth, P. and Rao, S.S. (2002) A Graph Based Approach to Extract a Neighborhood Customer Community for Collaborative Filtering. *Proceedings of the 2nd International Workshop on Databases in Networked Information Systems*, DNIS '02, Aizu, 16-18, December 2002, 188-200. http://dx.doi.org/10.1007/3-540-36233-9_15
- [18] Girvan M. and Newman, M.E.J. (2002) Community Structure in Social and Biological Networks. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 7821-7826. <http://dx.doi.org/10.1073/pnas.122653799>
- [19] Newman, M.E. (2004) Fast Algorithm for Detecting Community Structure in Networks. *Physical Review E*, **69**, Article ID: 066133. <http://dx.doi.org/10.1103/physreve.69.066133>
- [20] Newman, M.E.J. and Girvan, M. (2003) Finding and Evaluating Community Structure in Networks. Preprint Cond-mat/0308217, *Physical Review E*, **69**, Article ID: 026113. <http://arxiv.org/pdf/cond-mat/0308217.pdf>
- [21] Aaron Clauset, M.E.J. Newman, and Christopher Moore (2004) Finding Community Structure in Very Large Networks. *Physical Review E*, **70**, Article ID: 066111. <http://ece-research.unm.edu/ifis/papers/community-moore.pdf>
- [22] Tyler, J.R., Wilkinson, D.M. and Huberman, B.A. (2003) Email as Spectroscopy: Automated Discovery of Community Structure within Organizations. *Proceedings of the First International Conference on Communities and Technologies*, January 2003, 81-96. http://dx.doi.org/10.1007/978-94-017-0115-0_5
- [23] Liu, J., Li, T., Cheng, G., Yu, H. and Lei, Z.M. (2013) Mining and Modeling the Dynamic Patterns of Service Providers in Cellular Data Network Based on Big Data Analysis. *China Communications*, **10**, 25-36. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6723876>
<http://dx.doi.org/10.1109/CC.2013.6723876>
- [24] Rodriguez, M.A. and Pepe, A. (2008) On the Relationship between the Structural and Socioacademic Communities of a Co-Authorship Network. *Journal of Informetrics*, **2**, 195-201. <http://arxiv.org/pdf/0801.2345v3.pdf>
<http://dx.doi.org/10.1016/j.joi.2008.04.002>
- [25] Schein, I., Popescu, A., Popescu, L.H.R., Ungar, L.H. and Pennock, D.M. (2002) Methods and Metrics for Cold-Start Recommendations. *Proceeding of ACM SIGIR Conference on Research and Development in Information Retrieval*, ACM Press, Tampere, 11-15 August 2002, 253-260. http://repository.upenn.edu/cgi/viewcontent.cgi?article=1141&context=cis_papers
- [26] Adomavicius, G. and Tuzhilin, A. (2005) 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**, 734-749. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1423975>
<http://dx.doi.org/10.1109/TKDE.2005.99>
- [27] Breese, J.S., Heckerman, D. and Kadie, C. (1998) Empirical Analysis of Predictive Algorithms for Collaborative Filtering. *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, Madison, May 1998, 43-52. <http://research.microsoft.com/pubs/69656/tr-98-12.pdf>