A new Approach of Community Detection Based on Seed Node

A thesis in partial fulfilment for the degree of

MASTER OF COMPUTER APPLICATIONS

Submitted By
Bheem Kumar
(15/CA/639)

Under the supervision of

Dr. Prasenjit Choudhury
(Assistant Professor)
NIT Durgapur

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY, DURGAPUR

MAY 2018

NATIONAL INSTITUTE OF TECHNOLOGY DURGAPUR

MAHATMA GANDHI AVENUE

DURGAPUR, INDIA - 713209

DECLARATION BY THE CANDIDATE

I certify that

1. The work contained in this thesis is original and has been done by me under the guidance of my supervision.

2. The work has not been submitted to any other Institute for any degree or diploma.

3. I have followed the guidelines provided by the Institute in preparing the thesis.

4. I have conformed to the norms and guidelines given in the Ethical Code of Conduct of the Institute.

5. Whenever I have used materials (data, theoretical analysis, figures and text) from other sources, I have given due credit to them by citing them in the text of the thesis and giving their details in the references.

Date: may 2018 Bheem kumar

Palace: NIT Durgapur 15/CA/639

NATIONAL INSTITUTE OF TECHNOLOGY DURGAPUR

MAHATMA GANDHI AVENUE

DURGAPUR, INDIA - 713209

CERTIFICATE OF RECOMMENDATION

This is to certify that the thesis entitled "**A new Approach of Community Detection Based on Seed Node**", submitted by **Bheem kumar** for the partial fulfilment of the requirements for the award of the degree of Master of Computer Applications, is a bonafide research work under the guidance of Dr. Prasenjit Choudhury. The results embodied in this thesis have not been submitted to any other University or Institute for the award of any degree or diploma. In our opinion, this thesis is of the standard required for the partial fulfilment of the requirement for the award of the degree of **Master of Computer Applications**.

____

Internal Guide

Dr. Prasenjit Choudhury
Asst. Professor, Dept. of CSE

National Institute of Technology

Durgapur - 713209, INDIA

____

(Counter Signed by)

Dr. Goutam Sanyal

Professor & Head, Dept. of CSE

National Institute of Technology

Durgapur - 713209, INDIA

NATIONAL INSTITUTE OF TECHNOLOGY DURGAPUR

MAHATMA GANDHI AVENUE

DURGAPUR, INDIA - 713209

CERTIFICATE OF APPROVAL

The following thesis is hereby approved as a creditable study of Computer Applications subject carried out and presented in a manner satisfactory to warrant its acceptance as aprere quisite to obtain the degree for which it has been submitted. It is understood that by this approval the undersigned don't necessarily endorse or approve any statement made, opinion expressed or conclusion drawn there in but approve the thesis only for the purpose for which it has been submitted.

BOARDS OF THESIS EXAMINERS

1. ______________________ **4.** ______________________

2. ______________________ **5.** ______________________

3. ______________________ **6.** ______________________

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my project guide **Dr. Prasenjit Choudhury**, Assistant Professor, Department of Computer Science and Engineering, NIT Durgapur for their constant guidance, encouragement and support. Their dedicated involvement in every step gave me the confidence to successfully take this project to completion.

I wish to acknowledge Dr. Goutam Sanyal , HOD of CSE, for providing supporting environment for my work. I would also like to show gratitude and respect to other faculty members of our department for their encouragement. Last but not the least,

I would like to thank my family, my parents and my friends for their unconditional support.

ABSTRACT

From Earlier days Community detection is one of the most popular topic and is an important structure in social network. Community detection is required to define that who is belong to which community. Further numerous approaches are already proposed for finding community over a social network. In this paper we will propose a new approach of community detection based on seed centric approach in which we will discuss that how we are finding the seed node. We will also discuss about how seed set is expanding from remaining nodes which do not belongs to any community directly. The basic idea underlying in this approach is to identifying special nodes in the target network, called seed so that we can detect good community. Different algorithms adopt different approaches of seed selection definitions and seed set expansion definitions for communities construction. We will apply our algorithm in three classical data sets and will compare to other algorithms.

Keywords—Social networks, seed selection, Community Detection;

Table of Content

1. INTRODUCTION

There are varieties of the social networks we are using these days such as Facebook, Twitter, Amazon, LinkedIn, etc. We are growing with the information age. Now a day most of the people connected with this type of social network. With the development of the smart devices, more and more people login into their social networks through their smartphones, computer, laptop etc, and share the informations with their friends online. The social network is usually modeled as graphs. Large-scale social networks have become the complex networks, like www networks [1], metabolic networks [2-4], epidemiology networks [5,6], ecological webs [7], etc. A social network consists of a set of nodes along with edges connecting the nodes. The nodes represent the object in social networks, such as people, commodities, etc. The edges represent the relationships between objects. Basically, communities are set of nodes with higher edge density. In recent year, finding seed nodes and how to use the seed nodes for community detection has become a popular topic in the complex network analysis and field of data mining. A single vertex is put only in one community, but in some cases vertices are present in more than one community. In social network analysis, finding seed nodes have a wide range of applications. E.g., if we can dig out the most influential customers in marketing, then through the community structured by the seed nodes, a product brand can be rapidly promoted. +In this paper, we are using clustering coefficient and degree of the node as the basic theory to discover seed nodes, and apply seed nodes for community detection in social networks. The modularity is used as the measure standard to evaluate the accuracy and efficiency of the algorithm. Lastly we will visualize all three classical dataset and will compare from another community detection algorithm.

2. RELATED WORK

Community detection is important in the field of social network analysis. In 2002, Girvan and Newman [18] published an article which explored the community structure of social networks and biological networks. This paper started a period of rapid development in finding communities. In 2004, Newman and Girvan [19] proposed the concept of modularity, and introduced the null model into the analysis of the community structure. They also proposed the concept of betweenness, from different aspects, such as the perspective of shortest path, random walk, current. They adopted a divisive hierarchical clustering algorithm for community discovery. In the following years, many community detection algorithms were proposed based on modularity, such as the greedy algorithm [20,21], the simulated annealing algorithm [22] and the ultimate optimization algorithm [23]. These algorithms all use the null model as the reference.

In 2007 Fortunato and Barthelemy published an article in PANS [24], pointing out that the algorithms based on modularity have resolution limits, which make the accuracy of community detection by these algorithms suspect. The modularity and null model proposed by Newman and Girvan needs to be further improved, but it emphasized the idea of using the null model as the reference when compared with real local networks. Using random walk to find community structure was adopted in early work such as [25-30]. For example, in 2005, Pascal Pons and Matthieu Latapy [30] designed an algorithm based on random walk for community detection. As for the concept of order statistic, it was proposed by Radicchi F, Lancichinetti A, Ramasco [31, 32]. We combine the order statistic theory and the seed nodes to find communities structure in social networks. The algorithm can be extended to detect overlapping community.

In recent years, many new algorithms have been proposed for community detection, such as, in 2012, a cooperative and heuristic community detection algorithm proposed by Ruixin Ma [33]. In the same year, Shuzhi Li [34] proposed a novel community structure partition algorithm based on multi-gene family. Apart from that, much attention was focused on the analysis of the community structure. For example, Gyeong-Mi Park [35] gave a full description about the structure analysis of social network from the perspective of characters in literature Texts.

This paper is consisted of 6 sections. In the first section, we talk about the Introduction of community detection in social networks. In the second section, the preprocess procedure using random walk is introduced. In the third section, we discuss about the modularity and how to find seed nodes, how to detect the community in social networks are talked about. In forth part, the algorithm is tested in classical datasets and the experiments results also visualized. Fifth section contain the conclusion and the last section will be references.

2.1 Modularity

Modularity is one of the measure structure of networks or graphs. It was designed to measure the strength of division of a network into modules (also called groups, clusters or communities). Networks with high modularity have dense connections between the nodes within modules but sparse connections between nodes in different modules. Modularity is often used in optimization methods for detecting community structure in networks. The value of the modularity lies in the range [-0.5,1]. It is positive if the number of edges within groups exceeds the number expected on the basis of chance. For a given division of the network's vertices into some modules, modularity reflects the concentration of edges within modules compared with random distribution of links between all nodes regardless of modules.

2.2 Seed-Centric Algorithms

2.2.1 General Description

Algorithm 1 presents the general outlines of a typical seed-centric community detection algorithm. We recognize three principal steps:

1. Seed computation.
2. Seed local community computation
3. Community computation out from the set of local communities computed in step 2.

#*<u>Notation</u>*
G - Graph
V - Vertices
E - Edges
Cs - Community
C - Communities
S - Seeds
s = seed
Cs = Local Community

Algorithm 1 General seed-centric community detection algorithm

Require: G =< V,E > a connected graph,
1: C ← ∅
2: S ← compute seeds(G)
3: for s ∈ S do
4: Cs ← compute local com(s,G)
5: C ← C + Cs
6: end for
7: return compute community(C)

Each of the above mentioned steps can be implemented applying different techniques. A quick survey of existing approaches allows to identify five main criteria that can be used to classify seed-centric approaches.

2.2.2. Classification Criteria

Five criteria has been identified for classifying seed-centric approaches. The first three criteria are relative to the step of seed computation. The two last ones are related to the two last steps : seed local community computation and community computation steps. The five identified criteria are the following:

1. **Seed nature:** Form a pure topological point of view, a single seed can be: a single node a set of nodes (not necessarily connected), or a set of nodes composing a subgraph that is densely connected. Concerning the role that a seed plays in a network, most existing algorithms search for seeds that are likely to be at the core of communities to be detected. This is mainly the case of leader-based approaches [11-13] and set-seeds based approaches. One exception is the Yasca algorithm that search for seeds playing various roles in the target network.

2. **Seed number:** The number of seed nodes can be given as an input to the algorithm. This is much a classical approach similar to the use of the standard k-means data clustering algorithm [15]. In most situation, it is hard to know the

number of communities to discover. Some heuristics have been proposed in order to automatically compute the set of all possible seeds. For example, in leader-based algorithms (where seeds are supposed to be community leaders), nodes that have a higher centrality than their direct neighbors are considered as leaders (i.e. seeds) [11,13].

3. **Seed selection policy:** We can distinguish here between two main approaches for seed selection: random election and informed selection. Applying a random selection consists on selecting randomly a set of eligible seeds. For example in [12] authors propose to select K-top central nodes as leaders. Algorithms applying a random selection strategy often apply an iterative process, where communities discovered by the end of a first iteration are used to select new set of leaders and the algorithm iterates, and thus until convergence. Algorithms defining seed as a group of nodes or as a subgraph apply mainly an informed selection policy such as the one cited above (selecting nodes with higher centrality that direct neighbors).

4. **Seed local community computation:** One first classical approach for seed local community computation consists on applying an expansion approach. Local (or ego-centered) community detection algorithms can simply be applied for that purpose. One major drawback of expansion strategy is that it does not ensure covering all nodes of the network in the set of detected communities. To overcome this problem, one can add outliers node to the most near community as proposed in [16]. Another approach, is to apply an agglomerative approach where each node in the network search to join the community of the nearest detected seed [11,12].

5. **Community computation:** Once all local communities of all seeds have been identified, a global decomposition of the network into, eventually overlapped, communities should be computed. In most existing approaches The final result is simply taken to be the set of seed communities. This is natural when seeds are selected as nodes at the core of theirs communities. In [17], authors propose an alternative approach where an ensemble clustering approach is applied.

3. OUR ALGORITHM BASED ON SEED NODES FOR COMMUNITY DETECTION

In this section we will discuss about our algorithms that how algorithm is selecting the seed node and how seed set expansion is happening.

3.1 Preliminaries for find seed nodes

CC - Here CC means clustering coefficient of any particular node. The clustering coefficient means the relation between the neighbors of that particular node is how much dense. The clustering coefficient of a node is between the range of [0,1]. CC is the list of clustering coefficient of all the nodes available in the graph G.

Deg - Here Deg represents the degree of a particular node means how many nodes are connected to that node. Deg is the list of degree of all the nodes available in the graph G.

CCMD - Here CCMD is actually the product of the CC and Deg, that is we are multiplying the clustering coefficient of a node with the degree of the same node. CCMD is the list of product of corresponding nodes of CC and Deg.

LC - Here LC stands for Local Community. The LC is created when seed node is identified in the networks and we had to mark all the neighbors of that particular seed node so we created the a local variable LC to store the seed node and their neighbors. This LC is created for local use not for long time use, and this Local Community is not a complete community it’s because of seed set expansion is not yet done.

LOC - After creating Local community we have to store that community. For this purpose we created a new list that is LOC that means List Of Community. There is more chances that multiple community can be detected in a network. So for storing all the community we used LOC variable. The use of LOC variable is in seed set expansion where remaining nodes that are not belongs to any community directly. Next we will explore the RN(remaining nodes)

RN - The RN stands for Remaining Nodes which is note belongs to any community during community detection then it stored in RN variable. For seed set expansion.

a - a is simple a node id.

3.2 The Proposed Algorithm Explanation

First we are taking the input as graph G(V,E), after then all the process will run under while loop and the loop will continue until some criteria is not fulfilled. Under the loop find out the degree of all the nodes available in graph in variable *Deg* and also find out the clustering coefficient of all the nodes in *CC* variable. Then we multiplied the clustering coefficient with Degree of nodes with corresponding node id and stored in *CCMD***.** After that we find the maximum *CCMD* from the *CCMD* list and if the degree of that *CCMD* is greater than threshold 1 and the clustering coefficient of *CCMD* is greater than threshold 2 then it becomes selected as seed node otherwise loop becomes break. Here we took threshold 1 is equal to 3 it's because there may be chances that it can create a community of three node so some community will become very big and some community will become very small like 3 nodes community t. And for same purpose threshold 2 is taken. In threshold 2 we are conditioning that if clustering coefficient of seed node is less than the half of max clustering coefficient of list CCMD then the loop will become break. We calculated the maximum clustering coefficient in before starting the loop. after selecting the seed node its time to detect the community for that we copy the seed and neighbors of seed node in local community *(LC)* and then local community is stored permanent in the list of community that is (LOC) and if the nodes does not belongs to any community directly then it stored in remaining node *(RN)* variable, and all these nodes are added into communities using seed set expansion.

In seed set expansion for all nodes in RN we first checked that in which community have maximum number of neighbors of remaining node then i put the node into community which have highest number of neighbors of remaining node.

<u>#Notation</u>

CC - *Clustering Coefficient of all the node available in the graph G.*
Deg - *Degree of all the nodes available in the graph G*
CCMD - *Clustering Coefficient Multiplied by Degree of corresponding node*
LC - *Local Community*
LOC - *List of Community*
RN - *Remaining Nodes*
a - *node id*

ALGORITHM 1 Our algorithm to Finding seed node and detecting communities

```
input = G
output = seed node and communities
RN = 0
while(True)
        CC = Clustering Coefficient of all the node in graph G
        Deg = degree of all the nodes in graph G
        CCMD = CC *  Deg
        seed = max  CCMD in CCMD list
        Deg = degree of seed in  Deg list
        if(neighbors of seed < threshold1 or CCMD[seed] < threshold2)  then
                RN = remaining nodes in graph G
                break the loop
        else
                LC = seed + neighbors of seed
                LOC = LOC + LC
                G = G – LC
        end  if
        if( G is empty ) then
                Break the loop
        end  if
end  while
```

After finding seed node and detecting the community there is chances that some nodes do not belongs to any community directly so we stored that type of nodes in remaining nodes (RN) list for expansion. Below is seed seed set expansion algorithm by which we can included that remaining nodes into the created community according to their clustering coefficient and degree of the remaining nodes.

ALGORITHM 2 Our algorithm to Seed set expansion

Input *= RN*
Output *= communities*
while*(length of RN > 0)*
    for *each node* ***a*** *∈ RN* ***do***
        find the neighbors of ***a*** *in each community(LOC)*
        put the node ***a*** *in that community, which have maximum number of neighbors of* ***a***
    end for
    RN = RN **-** *a*
end while

4. EXPERIMENTAL RESULT

In this section we evaluate a selection of seed-centric approaches on a set of classical benchmark networks for which a ground-truth decomposition into communities era known. First we described the used datasets. Then we explicit the applied evaluation criteria and lastly we compare results of selected seed-centric approaches along with some classical community detection algorithms mainly based on the principle of modularity optimisation.

4.1 Datasets

Network name	#node	#edge	References
Karate Club	34	78	[8]
Dolphins network	62	159	[9]
Political books	105	441	[10]

4.1.1 Zachary's karate club

This network is a social network of friendships between 34 members of a karate club at a US university in 1970 [8]. Following a dispute the network was divided into 2 groups between the club's Seed-Centric Approaches for Community Detection 203 administrator and the club's instructor. The dispute ended in the instructor creating his own club and taking about half of the initial club with him. The network can hence be divided into two main communities.

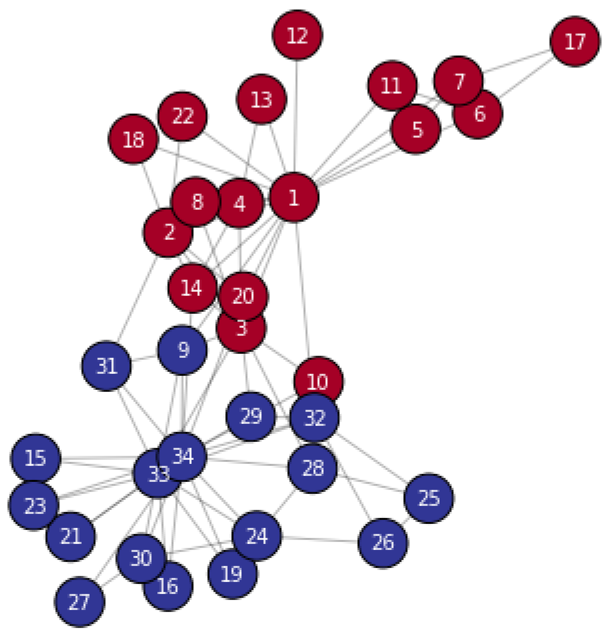

Fig. 1 After partition of Zachary's karate club

4.1.2 Dolphins social network

This network is an undirected social network resulting from observations of a community of 62 dolphins over a period of 7 years [9]. Nodes represent dolphins and edges represent frequent associations between dolphin pairs occurring more often than expected by chance. Analysis of the data revealed two main groups.

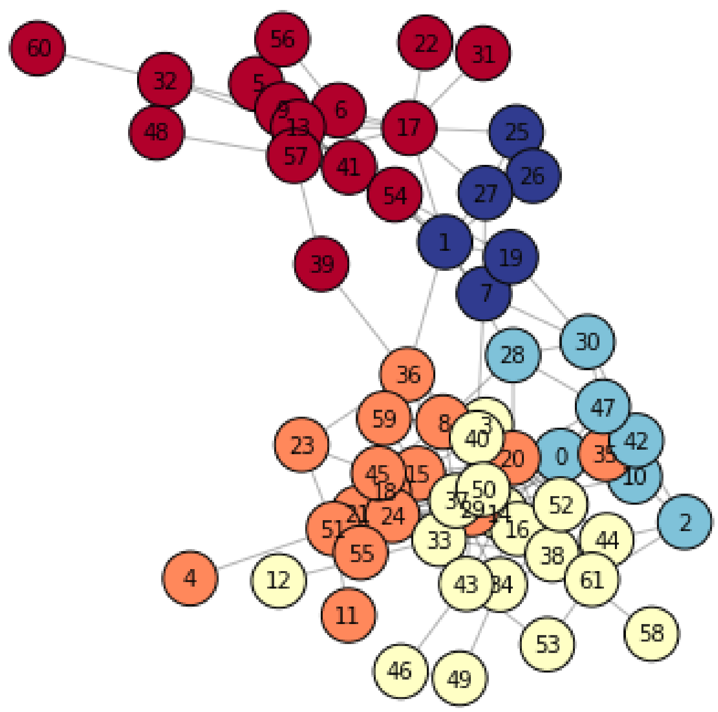

Fig. 2 After partition of Dolphins Network

4.1.3 American political books

This is a political books co-purchasing network [10]. Nodes represent books about US politics sold by the online bookseller Amazon.com. Edges represent frequent co-purchasing of books by the same buyers, as indicated by the ”customers who bought this book also bought these other books” feature on Amazon. Books are classified into three disjoint classes: liberal, neutral or conservative. The classification was made separately by Mark Newman based on a reading of the descriptions and reviews of the books posted on Amazon.

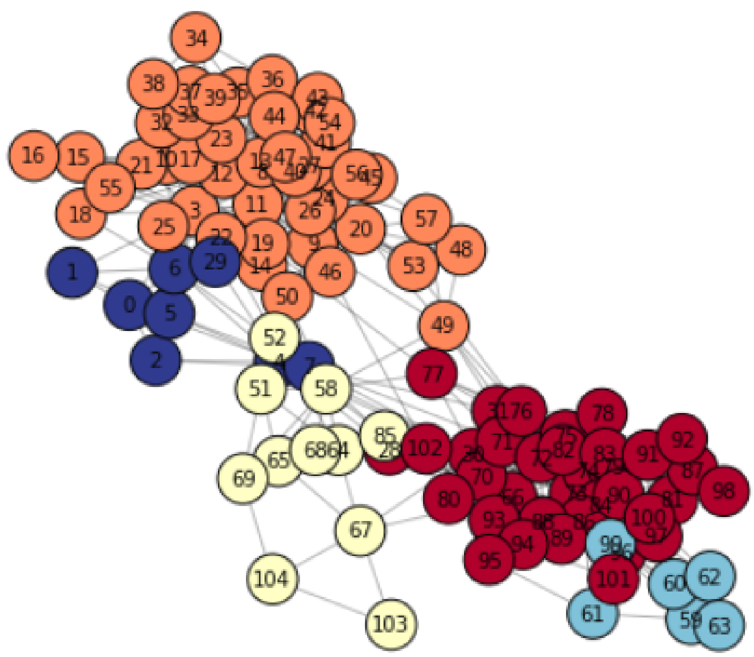

Fig. 3 After partition of American political books Network

4.2 Size Distribution vs Modularity Class

Here we are visualizing the relation between the communities (modularity class) and the number of nodes in each community. The horizontal line is showing the modularity class and the vertical line showing the number of nodes in each node modularity class.

1. **Zachary's karate club**

Fig. 4 karate club

2. **Dolphins Network**

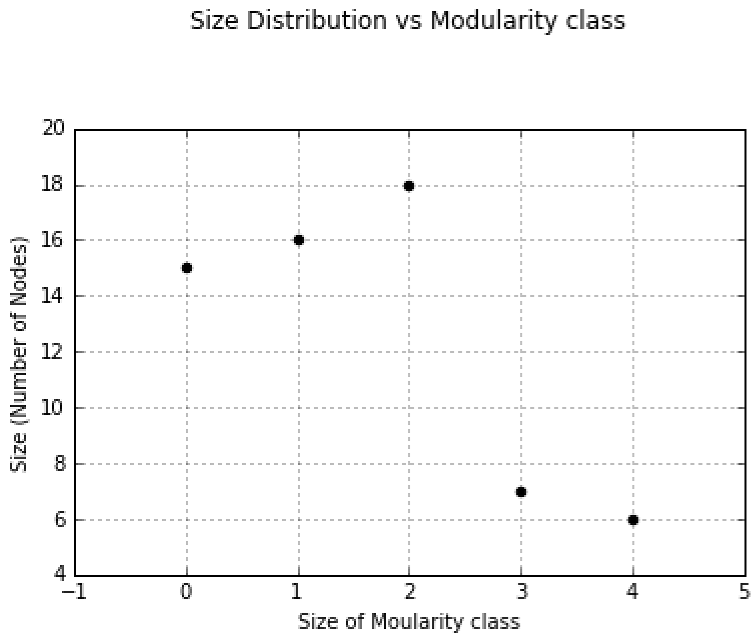

Fig. 5 Dolphins Network

3. **US Political Books**

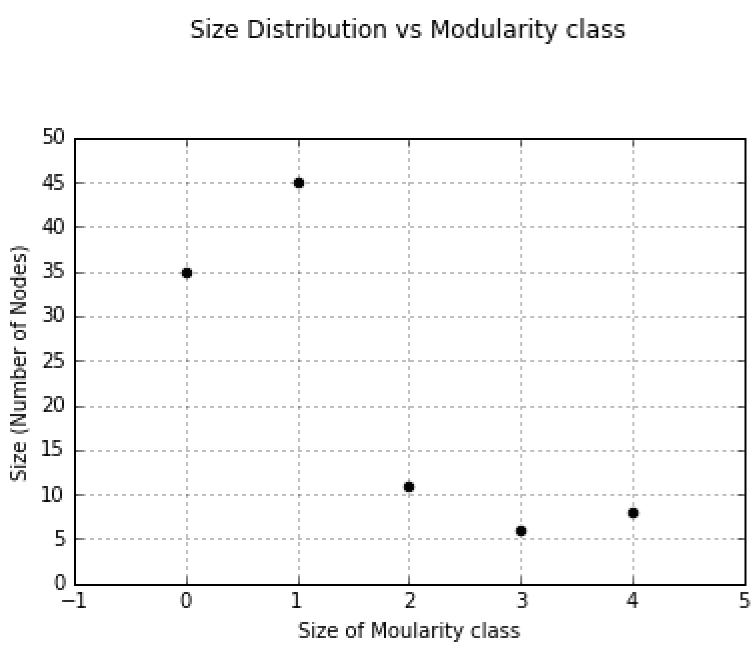

Fig. 6 US Political Books

4.3 Experiments Results

After experiments of all the dataset in our algorithm we stored the number of communities found in each communities of all the networks and we also calculated the modularity of all the networks, stored in Table 1. With corresponding network.

Table 1

Dataset	**Nodes**	**Edges**	**Communities**	**Modularity**
Karate Club	34	78	2	0.372
Dolphins	62	159	5	0.505
Political books	105	441	4	0.524

Table 1. Result of Experiments

4.4 Experiments Results and comparison

In this table we are comparing our algorithm with existed algorithms with Communities and Modularity

Table 2

Dataset	Algorithm	Communities	Modularity
Karate	Newman	5	0.40
	Lovain	4	0.41
	Walktrap	5	0.35
	Licod	3	0.24
	Yasca	2	0.34
	Our Algorithm	**2**	**0.37**
Political Books	Newman	5	0.51
	Lovain	4	0.52
	Walktrap	4	0.50
	Licod	6	0.42
	Yasca	3	0.24
	Our Algorithm	**4**	**0.52**
Dolphins	Newman	5	0.51
	Lovain	5	0.51
	Walktrap	4	0.48
	Licod	2	0.35
	Yasca	3	0.53
	Our Algorithm	**5**	**0.50**

4.5 Visualization of Modularity Comparison

This bar graph is showing the comparison between our proposed algorithm and existed algorithms on the basis of modularity of networks. By this bar graph we see that how well our algorithm is created. Here we used modularity for algorithm comparison, because modularity shows that how good our network is divided.

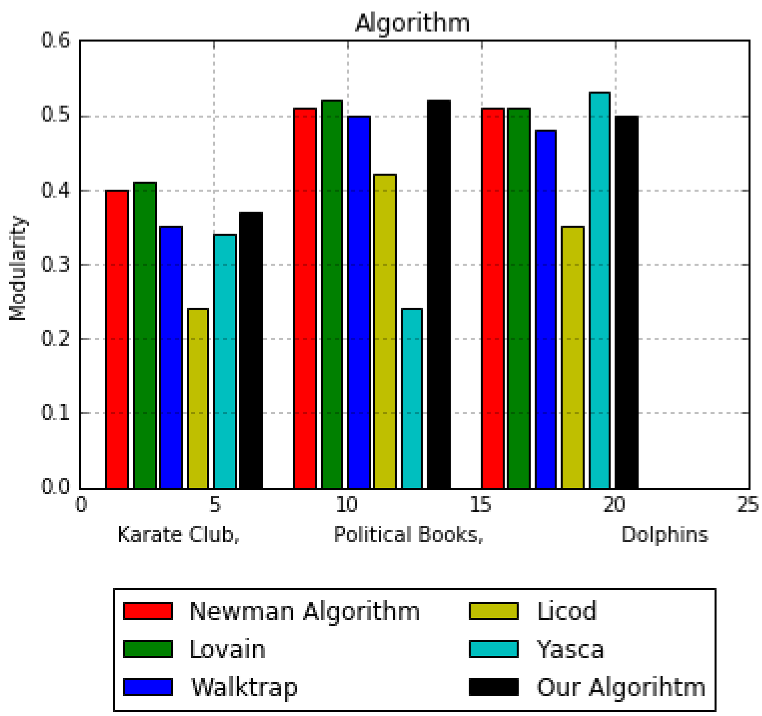

Fig. 7: Comparison modularity

5. CONCLUSION

A new algorithm for community detection is proposed based on seed node. We used three classical dataset for experiments. The experiments i did using small networks as existed algorithm. The performance of this algorithm is moderate and still this algorithm can be improved especially in validation of finding seed nodes. In above Figure (fig. 7) at least in political books our algorithm is equal the Lovain algorithm which is better than remaining algorithm.

6. REFERENCES

[1] Faloutsos M, Faloutsos P, Faloutsos C. On power-law relationships of the internet topology[C]//ACM SIGCOMM Computer Communication Review. ACM, 1999, 29(4): 251-262.

[2] Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N.& Barabási, A.-L. (2000) Nature 407, 651–654.

[3] Wagner A, Fell D A. The small world inside large metabolic networks[J]. Proceedings of the Royal Society of London. Series B: Biological Sciences, 2001, 268(1478): 1803-1810.

[4] Ravasz E, Somera A L, Mongru D A, et al. Hierarchical organization of modularity in metabolic networks[J]. science, 2002, 297(5586): 1551- 1555.

[5] Moore, C. & Newman, M. E. J. (2000) Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top. 61, 5678–5682.

[6] Pastor-Satorras R, Vespignani A. Epidemic spreading in scale-free networks[J]. Physical review letters, 2001, 86(14): 3200.

[7] Williams, R. J. & Martinez, N. D. (2000) Nature (London) 404, 180–183

[8] Zachary, W.W.: An information flow model for conflict and fission in small groups.Journal of Anthropological Research 33, 452–473 (1977)

[9] Lusseau, D., Schneider, K., Boisseau, O.J., Haase, P., Slooten, E., Dawson, S.M.:The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations. Behavioral Ecology and Sociobiology 54, 396–405 (2003)

[10] Krebs, V.: Political books network, http://www.orgnet.com/

[11] Kanawati, R.: Licod: Leaders identification for community detection in complex networks. In: SocialCom/PASSAT, pp. 577–582 (2011)

[12] Khorasgani, R.R., Chen, J., Zaiane, O.R.: Top leaders community detection approach in information networks. In: 4th SNA-KDD Workshop on Social Network

[13] Shah, D., Zaman, T.: Community Detection in Networks: The Leader-Follower Algorithm. In: Workshop on Networks Across Disciplines in Theory and Applications, NIPS (2010)

[14] Whang, J.J., Gleich, D.F., Dhillon, I.S.: Overlapping community detection using seed set expansion. In: He, Q., Iyengar, A., Nejdl, W., Pei, J., Rastogi, R. (eds.) CIKM, pp. 2099–2108. ACM (2013)

[15] Aggarwal, C.C., Reddy, C.K. (eds.): Data Clustering: Algorithms and Applications. CRC Press (2014)

[16] Whang, J.J., Gleich, D.F., Dhillon, I.S.: Overlapping community detection using seed set expansion. In: He, Q., Iyengar, A., Nejdl, W., Pei, J., Rastogi, R. (eds.) CIKM, pp. 2099–2108. ACM (2013)

[17] Kanawati, R.: Yasca: A collective intelligence approach for community detection in complex networks. CoRR abs/1401.4472 (2014)

[18] Girvan M, Newman M E J. Community structure in social and biological networks[J]. Proceedings of the National Academy of Sciences, 2002, 99(12): 7821-7826.

[19] Newman M E J, Girvan M. Finding and evaluating community structure in networks[J]. Physical review E, 2004, 69(2): 026113.

[20] Newman M E J. Fast algorithm for detecting community structure in networks[J]. Physical review E, 2004, 69(6): 066133.

[21] Blondel V D, Guillaume J L, Lambiotte R, et al. Fast unfolding of communities in large networks[J]. Journal of Statistical Mechanics: Theory and Experiment, 2008, 2008(10): P10008.

[22] Guimera R, Amaral L A N. Functional cartography of complex metabolic networks[J]. Nature, 2005, 433(7028): 895-900.

[23] Duch J, Arenas A. Community detection in complex networks using extremal optimization[J]. Physical review E, 2005, 72(2): 027104.

[24] Fortunato S, Barthelemy M. Resolution limit in community detection[J]. Proceedings of the National Academy of Sciences, 2007, 104(1): 36-41

[25] Lambiotte R, Delvenne J C, Barahona M. Laplacian dynamics and multiscale modular structure in networks[J]. arXiv preprint arXiv:0812.1770, 2008.

[26] Evans T S, Lambiotte R. Line graphs, link partitions, and overlapping communities[J]. Physical Review E, 2009, 80(1): 016105.

[27] Zhou H. Network landscape from a Brownian particle's perspective[J]. Physical Review E, 2003, 67(4): 041908..

[28] Zhou H. Distance, dissimilarity index, and network community structure[J]. Physical review e, 2003, 67(6): 061901.

[29] D. Lai, H. Lu, and C. Nardini, Physica A: Statistical Mechanics and its Applications 389, 2443 (2010).

[30] Pons P, Latapy M. Computing communities in large networks using random walks[M]//Computer and Information Sciences-ISCIS 2005. Springer Berlin Heidelberg, 2005: 284-293.

[31] Lancichinetti A, Radicchi F, Ramasco J J. Statistical significance of communities in networks[J]. Physical Review E, 2010, 81(4): 046110.

[32] Radicchi F, Lancichinetti A, Ramasco J J. Combinatorial approach to modularity[J]. Physical Review E, 2010, 82(2): 026102.

[33] Ma R, Deng G, Wang X. A Cooperative and Heuristic Community Detecting Algorithm[J]. Journal of Computers, 2012, 7(1).

[34] Li S, Wang X. Complex Networks Community Structure Division Algorithm Based on Multi-gene Families Encoding[J]. Journal of Computers, 2013, 8(12).

[35] Park G M, Kim S H, Cho H G. Structural Analysis on Social Network Constructed from Characters in Literature Texts[J]. Journal of Computers, 2013, 8(9).