International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 1, pp: 480-483 www.ijaem.net ISSN: 2395-5252

Community Detection with Influential and Follower Nodes

Avani Kesarwani[a] | Dr. Ashutosh Singh[b]
[a] Department of Electronics Engineering, Harcourt Butler Technical University, Kanpur-208002, UP, india
[b] Department of Electronics Engineering, Harcourt Butler Technical University, Kanpur-208002, UP, india

Date of Submission: 16-06-2020 Date of Acceptance: 02-07-2020

ABSTRACT
In present era most of the services are converted into online format, for example- supply of essential goods, e-governance data etc. Human have habits to connect with the individuals having common comportment and interests. So the network community is helpful for the business man, if a business man recognize a group of people, who are interested about some specific product, then he show adds only to them instead of showing to complete network.
In this paper we proposed a method of community detection based on finding influential nodes with its follower nodes. In this method firstly we search the influential node in the network and then we have to add follower nodes with the influenced node, which form a community. Then we have to disconnect all outer connections of influential node and follower node and add any outlying vertices of network with that community and remove that community from the network. Repeat same process until network is converted into communities.
KEYWORDS: Community Detection, Katz Centrality, Follower Similarity.

I. INTRODUCTION

Networks are crowd of similar types of objects with their relationships. In which objects are the assembly of individuals, computers, various similar devices etc. These objects are grouped at one place with respect to their common features. These groups are called communities inside a network [1]. There are various algorithms are developed for finding communities inside a network such as partition method, hierarchical method, clustering methods etc. These algorithms require previous knowledge of size and number of communities. Some more algorithms are developed on the basis of modularity, near vertices etc. These algorithm works for some accurate results of community structure. The previously developed algorithms are only working to find community inside a network but if we want community with its leader vertex, then these algorithms were failed. Such as if we want to use these algorithm for finding communities of terrorist network such as network of 9/11 terrorists, then this algorithm was failed to find its leader vertex. Because if we know about the leader of community then it's easy to destroy terrorist group easily by targeting the leader of community which work as catalyst for a group. There are various models are developed for detecting communities along with the catalysts vertex. Here we propose a new method of finding community with influential vertex.

II. PROPOSED ALGORITHMS

In this proposed method we develop a new algorithm for finding communities inside a network with their influential node [4]. In every network there exists a person, who provides new ideas and more active in their group. These persons are treated as catalysts for other members of community. So the some members of network follow their thoughts and join the group of that person. These members are recognized as follower of that influential node. Therefore the name of this algorithm is given as Influential Follower Community Detection (IFCD) Algorithm. In which we firstly detect the influential node then form its community by its follower.

2.1 Influential node selection

The basic phenomenon of proposed work to find a community inside the network is that, first we have to find a node which is more dominant and spreading lots of fresh and ingenious thoughts [2]. These types of nodes treated as influential nodes. These influential nodes are the center of their respective communities. Here we use katz centrality method to find central node of a community. This centrality is developed in 1953 by Leo Katz, which is works to find relative degree of vertices that means the total distance between two pairs of vertices.
Katz centrality is depends upon the distance between two node pair and according to that distance the matrix of graph is multiplied by an

attenuation factor named as α. The value of attenuation factor is less than 1, in this case attenuation factor is taken as 0.5. For calculation of katz centrality between two node pair i and j, we have to apply certain expression-

$$C_{katz}(i) = \sum_{k=1}^{\infty} \sum_{j=1}^{n} \alpha^k (A^k)_{ij}$$

Which node i have highest katz centrality that node is treated as more influential node and that node is more dominant in the network. At a time only one vertex is treated as a influential vertices. After the selection of influential vertex we have to find follower vertices of that dominant node.

2.2 Follower similarity calculations

There are various methods are there which presents the similarity between two node pairs[5]. But here we use a new similarity method by analyzing some existing method.

The concept behind this similarity method is that there are various other members exist in a network, which follow the thoughts of dominating person or influential person. So this new similarity method is defined as the ratio of their common neighbors between two node pair to the cubic root of multiplication of their squire's degree. Followers of influential node is calculated as

$$\text{Follower Similarity} = \frac{|\Gamma(x) \cap \Gamma(y)|}{\sqrt[3]{\{k^2x * k^2y\}}}$$

Where, x is the influential vertex and y is the vertex of network with which we measure similarity of dominant node. K_x and K_y is the degree of x and y respectively.

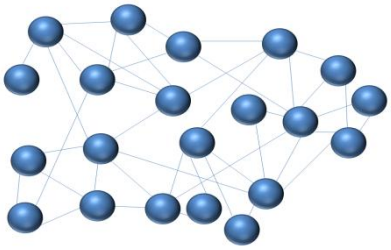

Figure 1 Initial graph

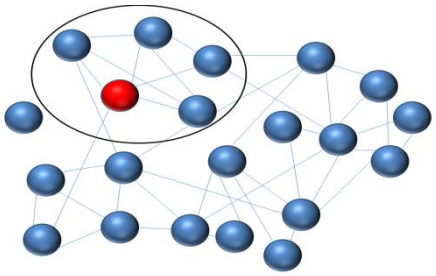

Figure 2 graph after similarity measure

This figure 1 shows a initial graph of a toy network and figure 2 shows the network structure after dominant node selection and follower similarity measure.

2.3 Node Addition

After removing all outer connections of influential as well as follower vertex, if there exists any vertex, which is completely outlaying from the network then we have to add that vertex with the selected community. This concept is shown clearly in the figures 3 with outlaying vertex and figure 4 with added outlying vertex in selected community.

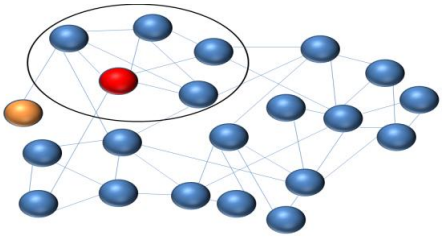

Figure 3 Network with outlying vertex

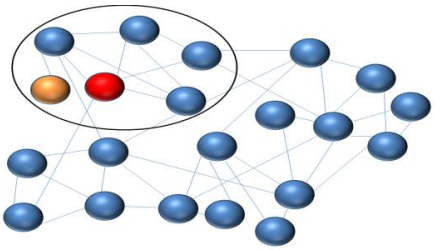

Figure 4 Addition of vertex

after finding a community we have to remove that community from the network and repeat same process until complet network is divided into community structure.

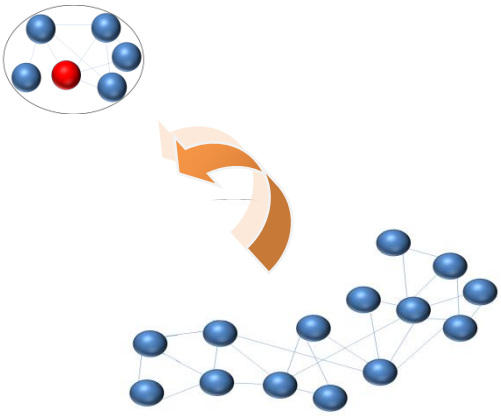

Figure 5 Removed community and remaing network

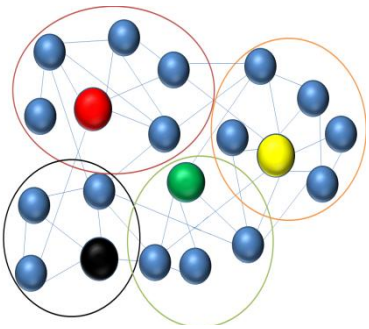

Figure 6 Network with community

In the Figure 5 and figure 6, the iteration of proposed algorithm is properly illustrated.

III. RESULT AND DISCUSSION

For the testing of our proposed algorithm, here we implement this algorithm into two datasets one is Zachary karate club and another one is the American college football datasets. The figure 7 represents the real existing community structure of Zachary karate club and figure 8 shows the community structure detected by **Influence Follower Community Detection (IFCD)** algorithm.

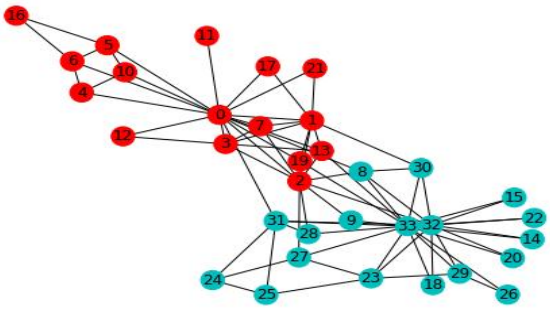

Figure 7 Ground truth community of Karate Club

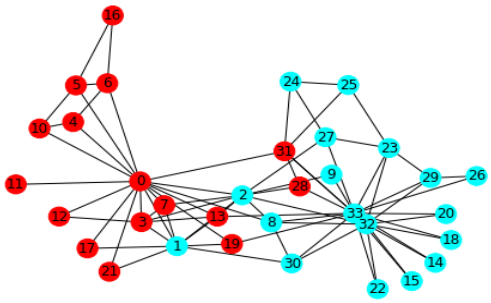

Figure 8 Resultant community of Karate Club

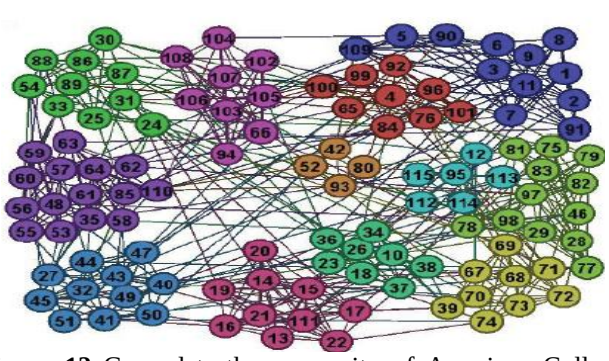

Figure 12 Ground truth community of American College Football [13]

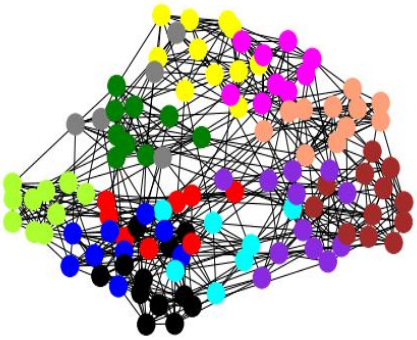

Figure 13 Resultant community of American College Football

the accuracy of proposed IFCD algorithm is detected by modularity, NMI and ARI measures.

Modularity: It defines the dense and sparse connection of a detected community.
NMI (Normalized Mutual Information): It compares the resultant community with original community. It is defined as ratio of mutual information to the addition of corresponding entropy[3].
ARI (Adjacent Rand Index): It is also b a parameter of testing the resultant community structure.

Table 1 Parameters values for proposed (IFCD) algorithm

Datasets	Modularity	NMI	ARI
Zachary Karate Club	0.231755424063115	0.3984265305315489	0.4828441189372699
American College Football	0.530674430301589	0.8796707560287794	0.8124488720027632

The above Table 1 shows the values of these three parameters for proposed (IFCD) algorithm.

IV. CONCLUSION

In this study we can see that the values of parameters are provides good results in American College Football datasets as compare to the Zachary Karate Club. So we can conclude that the result of proposed algorithm is good for large number of vertices and large communities.

REFERENCES

[1]. Wang, Y., Di, Z., Fan, Y., Identifying and characterizing nodes important to community structure using the spectrum of the graph. PLoS One 6 (11), p.e27418, (2018).

[2]. Gao, S., Ma, J., Chen, Z., Wang, G., Xing, C., Ranking the spreading ability of nodes in complex networks based on local structure. Phys. Stat. Mech. Appl.403, 130–147, (2014).

[3]. Shah, D., Zaman, T., Community detection in networks: the leader-follower algorithm. Phys. Stat. (November), ArXiv10110774, (2010).

[4]. Ahajjam, S., Badir, H., El Haddad, M., Detection of leader's nodes in complex networks. Glob. J. Eng. Sci. Res., 40–44, vol SESA, (2015).

[5]. Sara Ahajjam, Mohamed El Haddad, Hassan Badir, A new scalable leader-community detection approach for community detection in social networks, Laboratory of Information and Communication Technologies, National School of Applied Sciences, ENSA, Tangier, Moroccoa, (2018).

Avani Kesarwani, et. al. "Community Detection with Influential and Follower Nodes." International Journal of Advances in Engineering and Management (IJAEM), 2(1), 2020, pp. 480-483.

International Journal of Advances in Engineering and Management
ISSN: 2395-5252

IJAEM

Volume: 02 Issue: 01 DOI: 10.35629/5252

www.ijaem.net
Email id: ijaem.paper@gmail.com