Modeling And Analysis Of (MUEJ) Mehran University Research Journal Using Complex Network Approach

Saria Abbasi, Altaf Hussain Abro, Abdul Waheed Mahesar, Khalil-ur-Rehman Khoumbati

Abstract: Many complex systems in this world have been modeled and analyzed as complex networks. In this research paper, we have modeled specific dataset of MUEJ journal as two-mode network to find its structural properties. Our finding suggests that, this network is evolving with few prominent nodes which are playing the role of connector in this network. This trend can be the reason of highly clustered and more robust behavior in terms of specific nodes in the network.

Index Terms: Betweenness, Complex networks, Closeness , Collaborative networks, Centrality measures, Degree, Two-mode network.

1 INTRODUCTION

A network is a set of vertices or nodes and a set of lines which shows relation or flow between the nodes[1]. There are many ways of formally defining the complex network but the most accepted definition of complex network is “a system consisting of may interacting units whose collective behaviors can not be explained from the behavior of the individual units alone”[2].Complex networks have significant role in many diverse fields of life such as technological network (internet, World Wide Web, phone network, electric power grid network),biological (neuron, protein-protein interaction) network, ecological system, finance and various other domain which can be analyzed and modeled as a network[3][4]. All of the above networks are involve a large number of nodes and links. In the field of complex networks, research is focused on the small world effect and scale-free phenomenon which shows that the degree is size independent and it follows the power-law distribution. In [5] author proposed collaboration network and knowledge-based model based on several approaches and theoretical contribution, data were collected from the web of services and journal citation report, the record around (16,351).They compared complex network properties and found knowledge network is more cohesive and extensive. In[6] author focused on collaboration network among Canadian researchers, the data extracted from 1996 to 2010 and provide work for multiple regression models to estimate the impact on network structure variable. In several articles, such as[7][8][9], they demonstrated collaboration network in various fields and applied complex network methodologies on different case studies. The authors has studied the impact of citation of paper and it depend on some factors like author related factors, journal related factors, and paper related factors, total 2087 papers were extracted from WOS, Scopus, and PubMed.

This paper present Robocop multidisciplinary group from scientific fields and built a weighted adjacency matrix based on published paper conferring with standard (e.g ISBN, ISSN), graph partition approach was used to divide the graph into the cluster and calculate the data through Matlab script[10]. Based on the above research trends in collaborative networks, our research objectives focus on collaboration among the scientists of MURJ where they have written two or more papers together. In this study, we find the connectivity patterns of different authors with prominent nodes which makes this network random or scale-free by applying the core network analysis metrics of degree, betweenness, and closeness centrality. The structure of the remaining paper is as follows. In section 2 , we discuss few network analysis metrics. In section 3, we discuss the modeling. In section 4, we discuss the findings of ur research on particular dataset with network visualizations. Finally in section 5, we conclude the paper.

2 NETWORK ANALYSIS METRICS

2.1 Node Degree

Node degree is a simple and basic indicator of node centrality measures in the network. Node degree count the number of links directly connected to another node in the network[11].Node degree ki of the node i can be computed as given in equation (1)

$$K_i = \sum_{j=1}^{N} aij \qquad (1)$$

Here k_i indicates the degree of node i and a is adjacency the entry of a_{ij} shows the connectivity if the value is 1, otherwise it is 0 and N is the total number of nodes in the network.

2.2 Betweenness Centrality

Betweenness centrality measures the number of shortest paths from a node to node in the network or in simple word it counts the number of time a node acts as a bridge among two other nodes in the network. Betweenness [12] can be computed as given in equation (2)

$$C_B(k) = \sum_j^N \sum_i^N \frac{hji(k)}{hji}, \; i \neq j \qquad (2)$$

- *Saria Abbasi (corresponding author) is pursuing masters degree program in Computer Science University of Sindh , Jamshoro Pakistan. E-mail: saria.abbasi@gmail.com*
- *Altaf Hussain Abro is currently working as Assistant Professor in University of Sindh , Jamshoro Pakistan E-mail: altaf.abro@usindh.edu.pk*
- *Abdul Waheed Mahesar is currently working as Assistant Professor in University of Sindh , Jamshoro Pakistan E-mail: waheed.mahessar@usindh.edu.pk*
- *Khalil-ur-Rehman Khoumbati is currently working as Professor in University of Sindh , Jamshoro Pakistan E-mail: khalil@usindh.edu.pk*

Where $C_B(i)$ denotes the betweenness centrality of node i,g_{jk} is the total number of shortest path in between two nodes, and $g_{jk}(i)$ is the total number of routes that pass through node i.

2.3 Closeness Centrality

Closeness Centrality can be defined as how near an actor is to other actors in the network. It is the inverse of the sum of the shortest distance between each node to other nodes. Closeness can be formalized as in given equation (3) [13]

$$Cc(i)=\left[\sum_j^N d(i,j)\right]^{-1s} \quad (3)$$

Here $C_c(i)$ is the closeness of node i, and d(i,j) is the distance between node i and node j.

3 METHODOLOGY

This research study is based on quantitative approach. The data set has been extracted from Mehran university research journal since 2011 to 2016. After the primary analysis of data set we have used two- mode complex network by organizing data set into the matrix where rows and column indicate authors and papers. We have modeled two mode network if the authors have written two or more papers together or they have collaborated with each other. The dataset of Mehran university research journal has two-mode or (bipartite graph) structure, here the top set of nodes (T) shows the author and bottom set of nodes (⊥) represent the papers as shown in fig1. Usually, two mode (bipartite) network are converted into one mode because it is hard to analyze in their original form as there are very rare network analysis metrics are available.

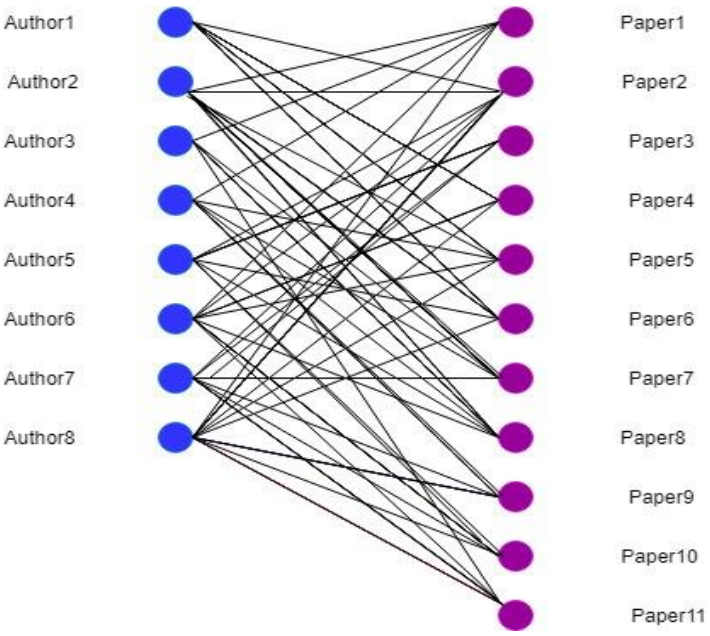

Fig.1*. The bipartite graph of MURJ*

4 FINDINGS

For this study, we applied local measure degree centrality, and global measures betweenness, and closeness centrality to identify the most central scientists in co-authorship network. The data has been extracted from Mehran University Research Journal (MURJ) since 2011 to 2016. The results of network analysis have been achieved in R-Project version 3.0.0 as a simulation tool which is free open source programming and statistical computing language. The power of R is its best graphics, which can be used for better network visualization. Also through additional packages the dynamic and interactive graphics can be obtained. The table 1 shows the top ten authors based on degree centrality. From the degree centrality, we noticed that author 54 has a high degree in terms of the number of connection that is 15 with different authors in the network. The id 276 is a second most connected author in the network with 14 degree. The author 2, and 35 have a degree of 13. Similarly, the authors 79 and 298 has 10 number of connection with other authors in the network and (50, 61, 85, and 87) have same score of degree 9 where they have written less number of papers with other authors. The fig 2(a) illustrate the network visualization of degree centrality. The large size nodes show high degree scientists who have worked with more co-authors than all others in the network and they are strongly connected to each other. Closeness Centrality shows the top ten authors based on closeness scores. It represents all those ten authors are more closely connected with each other's through collaborations. The author 35 has high closeness score indicated that they collaborated widely and receive information more quickly than other in the network and 298 author has the second highest score and well-connected author among others. In fig 2(b) closeness is based on color scale where purple color referred as high closeness centrality and blue node has middle level centrality as compared to rest of them have low level closeness centrality. Fig 2(c) Betweenness centrality show top ten authors with highest betweenness score .The author 298 has top score in term of betweenness centrality and it controls the largest number of communication among those who have written fewer papers together. The author id 2 has second high score in a co-authorship network.

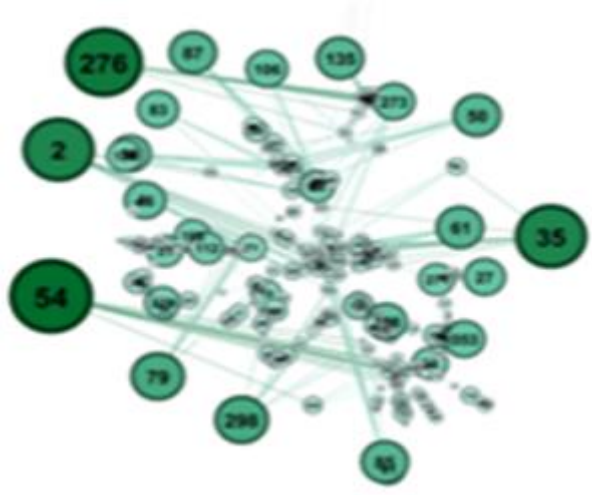

Fig.2 *(a)Visualizations of Degree*

Table 1*.The degree, closeness, and betweenness centrality result.*

Degree Centrality		Closeness Centrality			Betweenness Centrality	
Author_id	Degree	Author_id	Closeness	N.Closeness	Author_id	Betweenness
54	15	35	0.001182	4.75E-06	298	12593.93
276	14	298	0.001144	4.60E-06	2	10096.45
2	13	85	0.001139	4.57E-06	35	9179.368
35	13	2	0.001138	4.57E-06	54	8792.905
79	10	81	0.001026	4.12E06	106	4219.554
298	10	54	0.001016	4.08E-06	112	4177.142
50	9	210	0.001014	4.07E-06	85	3563.411
61	9	442	0.000998	4.01E-06	50	3408.861
85	9	71	0.000994	3.99E-06	21	3277.981
87	9	61	0.000992	3.98E-06	71	3108.303

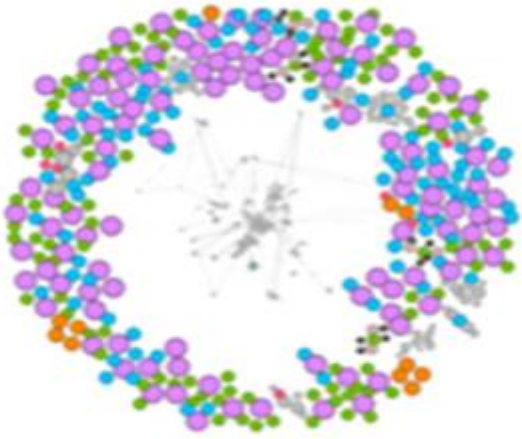

Fig.2 *(b) Closeness*

Fig.2 *(c) Betweenness .*

Fig.2*. (a)Visualizations of Degree, (b) Closeness, and (c) Betweenness Centrality.*

5 Conclusion

In this study, we have analyzed and model the networks of scientists in which nodes are authors of scientific papers and a tie between two or more authors who have written the same paper. The initial findings based on three network analysis metrics i.e, degree, betweenness, and closeness shows that the collaboration is not homogenous as few authors have contributed a lot as compared to others. This trend will surely self organize itself into clustered topology. In future research, we will analyze this dataset from different perspective of clustering and will find the exact topology to see the evolving nature of publications contributed by few prominent authors.

References

[1] M. E. J. Newman, “Scientific collaboration networks. I. Network construction and fundamental results,” Phys. Rev. E, vol. 64, no. 1, p. 16131, 2001.

[2] M. C. González, Contact Networks of Mobile Agents and Spreading Dynamics. 2006.

[3] R. Cazabet, P. Borgnat, and P. Jensen, “Enhancing Space-Aware Community Detection Using Degree Constrained Spatial Null Model,” Springer, pp. 47--55, 2017.

[4] D. Goss-Souza, L. W. Mendes, C. D. Borges, D. Baretta, S. M. Tsai, and J. L. M. Rodrigues, “Soil microbial community dynamics and assembly under long-term land use change,” FEMS Microbiol. Ecol., vol. 93, no. 10, 2017.

[5] J. Guan, Y. Yan, and J. J. Zhang, “The impact of collaboration and knowledge networks on citations,” J. Informetr., vol. 11, no. 2, pp. 407–422, 2017.

[6] A. Ebadi and A. Schiffauerova, “How to become an important player in scientific collaboration networks?,” J. Informetr., vol. 9, no. 4, pp. 809–825, 2015.

[7] M. Bordons, J. Aparicio, B. González-Albo, and A. A.

Díaz-Faes, “The relationship between the research performance of scientists and their position in co-authorship networks in three fields,” J. Informetr., vol. 9, no. 1, pp. 135–144, 2015.

[8] M. E. J. Newman, “Coauthorship networks and patterns of scientific collaboration,” Proc. Natl. Acad. Sci., vol. 101, no. Supplement 1, pp. 5200–5205, 2004.

[9] M. S. Couceiro, F. M. Clemente, and F. M. L. Martins, “Towards the Evaluation of Research Groups based on Scientific Co-authorship Networks: The RoboCorp Case Study,” Arab Gulf J. Sci. Res., vol. 31, no. 1, pp. 36–52, 2013.

[10] I. Tahamtan, A. Safipour Afshar, and K. Ahamdzadeh, “Factors affecting number of citations: a comprehensive review of the literature,” Scientometrics, vol. 107, no. 3, pp. 1195–1225, 2016.

[11] M. E. J. Newman, “Who is the best connected scientist? A study of scientific coauthorship networks,” vol. 64, pp. 1–7, 2000.

[12] L. C. Freeman, S. P. Borgatti, and D. R. White, “Centrality in valued graphs: A measure of betweenness based on network flow,” Soc. Networks, vol. 13, no. 2, pp. 141–154, 1991.

[13] L. C. Freeman, “Centrality in social networks conceptual clarification,” Soc. Networks, vol. 1, no. 3, pp. 215–239, 1978.