Indonesian Journal of Electrical Engineering and Computer Science
Vol. 29, No. 3, March 2023, pp. 1642~1647
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v29.i3.pp1642-1647

Complex networks analysis: centrality measures

Ali Ali Saber, Noor Kaylan Hamid
Department of Computer Engineering Techniques, College of Engineering Technology, AL-Kitab University, Kirkuk, Iraq

Article Info

Article history:

Received Aug 2, 2022
Revised Oct 28, 2022
Accepted Nov 2, 2022

Keywords:

Betweenness centrality
Centrality
Closeness centrality
Complex networks
Degree centrality

ABSTRACT

The centrality of an edge in a graph is proposed to be the degree of sensitivity of a graph distance function to the weight of the edge under consideration. Many centrality metrics are available in network analysis and are effectively used in the investigation of social network properties. Node position is one of them. In this paper, we propose a novel importance of nodes showing how to locate the most essential nodes in a network and to construct a centrality measure for each node in the network, sort the nodes by centralities, and focus on the top ranked nodes, which are the most relevant in terms of this centrality measure. Our research aims to explain how to identify the most important nodes in networks. A centrality metric should be established for each node in the network, and then the nodes based on their centralities, focusing on the top-ranked nodes, which in light of this importance, might be regarded as the most pertinent measure.

Corresponding Author:

Ali Ali Saber
Department of Computer Engineering Techniques, College of Engineering Technology
AL-Kitab University
Altun Kpori, Kirkuk, Iraq
Email: ali.a.saber@uoalkitab.edu.iq

1. INTRODUCTION

Recently, researchers are becoming increasingly interested in social networks. With the growing popularity of complex social networks, a large variety of tools for investigating and analyzing these networks have been created. The difficulty of extracting the most significant, central members of the network is one of the most essential issues in complex network analysis. Several measures have been developed for this purpose, like betweenness centrality, closeness centrality and degree centrality [1]. Complex networks, such as the world wide web, the internet, social friendship networks, and collaboration networks, are regularly met in our daily lives [2]. Each node in these types of networks has certain distinct properties that are utilized to determine its relevance based on the application environment. We need to find extremely influential or essential nodes in a variety of real-world applications [3].

The value of a node varies depending on the application context. A high-degree node can have a big impact on how information spreads in the area [4], [5]. A high betweenness node will be necessary for spreading information between communities, but it may not be prominent locally due to its local connections [4], [6]. However, the purpose of this research is to give further proof of algorithms's ability to cope with relational issues, not to develop a quicker technique to compute these measurements [7]. In this study, we will analyze three well-known vertex centralities:

Degree algorithm, it simply calculates the number of neighbors a vertex is connected. In betweenness algorithm determines the number of shortest routes that pass through the specified vertex. High betweenness vertices are more critical for the network's cohesiveness, and their removal may cause the graph to become disconnected. Closeness centrality it's a distance-based centrality that calculates the average distance between a particular vertex and all other vertices that may be reached [8]. The connectivity of a graph portion (site, or potentially edge) through direct edge connections to other sites determines its centrality. The frequency with which a graph portion occurs on meaningful (perhaps shortest) pathways between all pairs of other sites determines a site's centrality, and the degree to which a graph portion is near to the rest of the graph determines its centrality (measured by an average distance to the sites of the graph). Also, the purpose of this paper is to focus on and explore the following topics: Node count, degree, and shortest pathways are all basic network statistics. Similarity detection and node importance – determining how similar two social network users are. Node detection is the process of identifying important nodes that often interact with another nodes.

Previous work has been done for finding the most weighed nodes in complex network used Katz centrality. They estimated the related effect each nodes in a network by including its both its close neighbors and its non-neighboring nodes, which are linked by immediately adjacent nodes. For the first filtering constraint, the centrality is calculated using the user-defined values, and the user may then chose or choose the values of α and β. The user is in charge of the first level of filtering chains [9]. The user controls the first level of filtering chains, which is what const means [10]. Filtering away those nodes with higher degrees with a combination of neighbors and the most heavily weighted value with much more higher and much less lower centralities is the second or next step of filtering. While the user can change or vary the first filtering restriction, the second filtering constraint varies from network to network and is set for a particular network. The findings of this work demonstrate that the number of nodes in the search space has lowered by more than 75%. They conducted various values of α and evaluated the link between the number of search space nodes and the number of search space nodes [11]. The α values and network characteristics like degree ranking or distribution. So, it helps the user in choosing α value for running the algorithm. They also showed that there is a significant difference in rankings produced in both the approaches and this enables the users to capture two different kind of perspectives then studying the properties of weighted nodes. Using the Kats centrality measures gives only the important nodes in the given data. And there are more different other factors that affect the nodes importance in the complex network, such as their level of activity and the many kinds of activities they engage in their work [12]. Next, we go over our primary issues that have drawn researchers to this field. The primary goal of defining centrality metrics is to rank nodes. It's still a work in progress to determine a node's global rank using local data based on various centrality measurements. To determine the relevance of nodes based on their local and global connectedness, freeman developed three major centrality metrics in [13]. Undirected and unweighted networks were also covered by the suggested definitions. However, these unweighted networks are insufficient to communicate all of the system's information. These extensions will be discussed in our work.

2. THE PROPOSED METHOD

The goal of this research is finding the most important nodes In complex network which used in this work on real data in Minnesota.mat This dataset contains over 1,039 nodes and 6,234 edges, next we try to discuss the algorithms and the ideas for finding most weighted nodes or important nodes [14]. One of the most necessary has been needed is centrality metrics or significant ways for this goal, with a strong interest in discovering the most important nodes, nodes with high weight. In this paper, we will employ measures of centrality such as betweenness, closeness and degree to find the most weighted nodes in a complex network. Individuals who have a significant impact on a social network are known as important nodes [15], [16]. Finding and estimating key nodes in social networks is extremely useful for social network structure study.

3. METHOD

Firstly we propose the betweenness centrality algorithm which evaluates The frequency of each node acts or performs as a connecter or a bridge over the shortest route connecting two more nodes [17]. Being between implies that a vertex can affect the flow of information between several network nodes. This was presented by Linton Freeman in accordance with his theory [18]. The maximum betweenness is found in vertices with a large likelihood of being on a randomly determined shortest path or top between two randomly selected vertices. The betweenness of the formula results in the node v as in (1).

$$BC(v) = \sum uv \in V \frac{\propto uw\,(v)}{\propto uw} \tag{1}$$

$\propto$ uw = Total number of shortest paths between node u and node w.
$\propto$ uw (v)= Total number of shortest paths between node u and node w that pass through v.

Next algorithm which is used in this work closeness centrality, which state the average length of the closeness between two nodes or the shortest path between one node and all other nodes is used in this approach to calculate the closeness centrality of each node in the network [19]. It may be thought as a measure of how long it takes for information to propagate or disseminate throughout a network from a specific vertex or top point [20]. It may be used to identify nodes that can swiftly disseminate a rumor or false information throughout the network. The equation for it is as follows:

$$C_c(v_i) = \sum_{j=1}^{N} \frac{1}{d(v_i, v_j)} \quad (2)$$

Where Cc is the closeness centrality of a node vi.

Finally, we will use degree centrality algorithm, there are two types of degree centrality metrics in a directed network: in-degree and out-degree. By these two types of degree centrality we can find important nodes soThe number of ties that are directed to a node in the graph is measured by its in-degree, and the number of links that each node in the graph points or directs to other nodes is measured by its out-degree [21], [22]. The equations for them are as follows:

$$c_{in}(v_i) = d_{in}(v_i) \quad (3)$$

$$c_{out}(v_j) = d_{out}(v_j) \quad (4)$$

where $d_{in}(v_i)$ in (3) and $d_{out}(v_j)$ in (4) are the corresponding in-degree and out- degree centralities of nodes v_i and v_j

4. RESULTS AND DISCUSSION

In this part we are showing the results finding the most weighted nodes in a given network. The following approaches might be used to calculate the relevance of nodes. This technique investigates the number of nodes' neighbors as well as the impact of node significance. The notion of centrality in network analysis is the process of identifying the most significant node inside a network [13], [23]. This unique method of measurement may be used to identify the most important people in a social network as well as the Internet's main infrastructure nodes [24], [25]. Now let's look at some prominent network centrality metrics like closeness, degree and betweenness.

a) Loading the data minnesota.mat, which comprises a graph illustrating Minnesota's nodes as shown in Figure 1. As can be seen, all nodes have the same color before to applying centrality measurements to the data. We use the Euclidean distance between the xy coordinates of the end nodes of each edge to determine edge weights that roughly match to the length of the roadways.

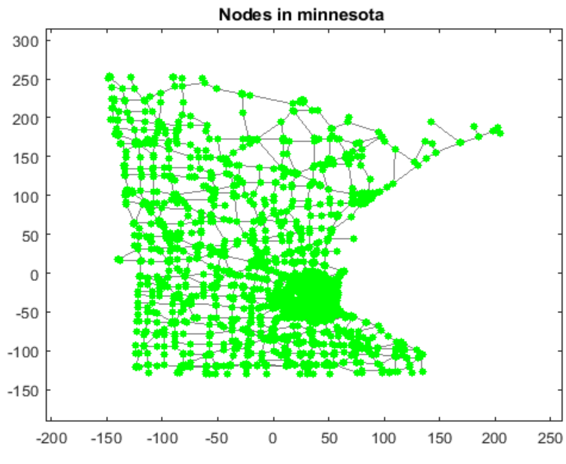

Figure 1. Nodes in minnesota

b) Then we compute each node's closeness centrality. Node's color should be proportionate to the centrality score. We calculate the weighted closeness centrality score as well, using the edge weights to represent the cost of traversing each edge. Because distances are now assessed as the sum of the lengths of all traveled edges, rather than the number of edges traveled, using road lengths as edge weights enhances the score quality as shown in Figure 2.

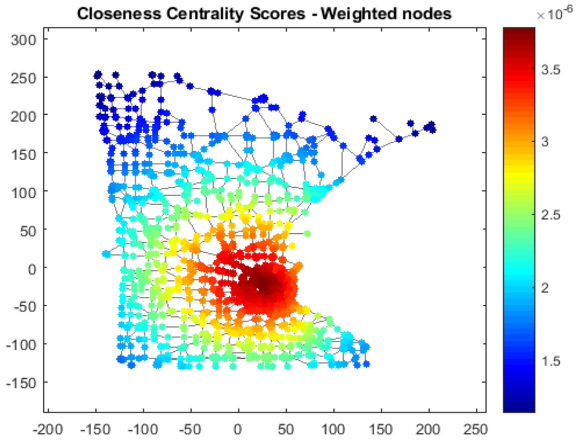

Figure 2. Closeness scores

c) Calculate the graph's weighted betweenness centrality scores to find the roads that are most frequently located on the shortest path between two nodes. Normalize the centrality ratings by factoring in the chance that a traveler on the shortest path between two random nodes will pass through a specific node as seen in Figure 3.

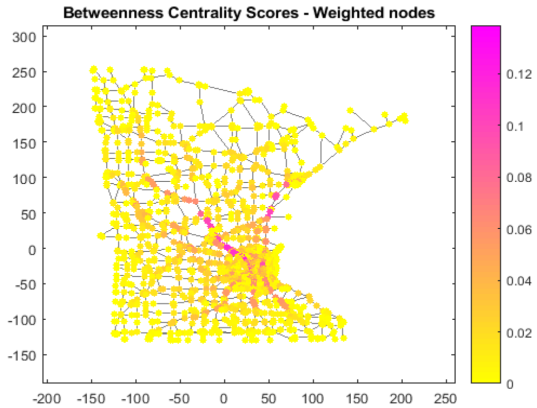

Figure 3. Betweeness scores

- The dark-colored nodes (important nodes) in the diagram above are the most weighted (values ranging from 0.08 to 0.12) or most essential nodes are listed in Table 1.
- Because they have in-degree by important nodes, weighted nodes are important.
- A high weight was given to an important web page, whereas a lesser weight was given to a less significant web page.
- As shown in the figure, the essential nodes are connected by a short path, it is likely that they are also connected directly.
- Important nodes are nodes that have a greater impact on the network's structure and operation than other nodes. Understanding the architecture of a network structure requires assessing the importance of nodes in a complex network.
- The degree is the quantity of nodes that are directly connected to each other (i.e., neighbors whose distance is equal to 1) and share a direct edge. The number of edges a node has with other nodes is measured by its out-degree, whereas its in-degree measures its edges with other nodes.
- The graph shows that the size of the nodes in the buddy network correlates the centrality in closeness. The proximity value is normalized by the total sum of the least feasible distances, because the entire sum of distances is dependent on the number of nodes.

– Standardize the network betweenness the total number of all-pairs shortest paths via a vertex determines a node's centrality. The few nodes that this measure has highlighted are those that are related to important nodes.

Table 1. Position of nodes in the network

Values	Nodes Position
0.1 - 0.12	Betweeness
0.6 - 0.8	Betweeness
0 - 0.04	Closeness

5. CONCLUSION

One indicator of a node's importance in the social network its position, which has been examined in this work. The most linked nodes which are found using degree centrality. Closeness centrality identifies the nodes in the network's middle and betweenness centrality is focused on nodes that act as bridges between different regions of the network. We've also explored how centralities must be adjusted for various network types. By taking the directionality of the link and their weights. In summary, centrality combines a vast number of criteria and methodologies for analyzing the importance of nodes in networks and identifying the most significant. In our future work we intend to look at how flow-based centrality and random-walk betweenness centrality may benefit from a similar approach. We also consider a variant of this technique that may be run on local sub graphs, making it more suitable for bigger graphs.

ACKNOWLEDGEMENTS

Thanks to computer technical engineering colleges in Al-Kitab University staff for providing research facilities to excute the work.

REFERENCES

[1] T. Challenge and I. Topology, "The challenge of interaction topology," *Los Alamos Sci.*, no. 29, pp. 94–109, 2005.
[2] D. J. Klein, "Centrality measure in graphs," *J. Math. Chem.*, vol. 47, no. 4, pp. 1209–1223, 2010, doi: 10.1007/s10910-009-9635-0.
[3] R. Lambiotte and M. T. Schaub, "Modularity and dynamics on complex networks," *Modul. Dyn. Complex Networks*, 2022, doi: 10.1017/9781108774116.
[4] H. Xu, J. Zhang, J. Yang, and L. Lun, "Identifying important nodes in complex networks based on multiattribute evaluation," *Math. Probl. Eng.*, vol. 2018, 2018, doi: 10.1155/2018/8268436.
[5] Y. Yang, L. Yu, Z. Zhou, Y. Chen, and T. Kou, "Node importance ranking in complex networks based on multicriteria decision making," *Math. Probl. Eng.*, vol. 2019, no. Dc, 2019, doi: 10.1155/2019/9728742.
[6] K. Li and Y. He, "The complex network reliability and influential nodes," *AIP Conf. Proc.*, vol. 1864, no. August 2017, 2017, doi: 10.1063/1.4992961.
[7] D. Sarkar, C. Andris, C. A. Chapman, and R. Sengupta, "Metrics for characterizing network structure and node importance in spatial social networks," *Int. J. Geogr. Inf. Sci.*, vol. 33, no. 5, pp. 1017–1039, 2019, doi: 10.1080/13658816.2019.1567736.
[8] J. Matta, G. Ercal, and K. Sinha, "Comparing the speed and accuracy of approaches to betweenness centrality approximation," *Comput. Soc. Networks*, vol. 6, no. 1, pp. 1–30, 2019, doi: 10.1186/s40649-019-0062-5.
[9] Y. Yu, B. Zhou, L. Chen, T. Gao, and J. Liu, "Identifying important nodes in complex networks based on node propagation entropy," *Entropy*, vol. 24, no. 2, pp. 1–17, 2022, doi: 10.3390/e24020275.
[10] J. Zhou, X. Yu, and J. A. Lu, "Node importance in controlled complex networks," *IEEE Trans. Circuits Syst. II Express Briefs*, vol. 66, no. 3, pp. 437–441, 2019, doi: 10.1109/TCSII.2018.2845940.
[11] F. Jian and S. Dandan, "Complex network theory and its application research on P2P networks," *Appl. Math. Nonlinear Sci.*, vol. 1, no. 1, pp. 45–52, 2016, doi: 10.21042/amns.2016.1.00004.
[12] S. Kaur, A. Gupta, and R. Saxena, "Identifying central nodes in directed and weighted networks," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 8, pp. 905–914, 2021, doi: 10.14569/IJACSA.2021.01208100.
[13] A. Mester, A. Pop, B. E. M. Mursa, H. Greblă, L. Dioşan, and C. Chira, "Network analysis based on important node selection and community detection," *Mathematics*, vol. 9, no. 18, 2021, doi: 10.3390/math9182294.
[14] H. Aznaoui, A. Ullah, S. Raghay, L. Aziz, and M. H. Khan, "An efficient GAF routing protocol using an optimized weighted sum model in WSN," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 22, no. 1, pp. 396–406, 2021, doi: 10.11591/ijeecs.v22.i1.pp396-406.
[15] J. Zhang, X. K. Xu, P. Li, K. Zhang, and M. Small, "Node importance for dynamical process on networks: A multiscale characterization," *Chaos*, vol. 21, no. 1, 2011, doi: 10.1063/1.3553644.
[16] M. O. Adebiyi, E. E. Adeka, F. O. Oladeji, R. O. Ogundokun, M. O. Arowolo, and A. A. Adebiyi, "Evaluation of load balancing algorithms on overlappiing wireless accesspoints," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 21, no. 2, pp. 895–902, 2020, doi: 10.11591/ijeecs.v21.i2.pp895-902.
[17] A. H. Alsaeedi, A. H. Aljanabi, M. E. Manna, and A. L. Albukhnefis, "A proactive metaheuristic model for optimizing weights of artificial neural network," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 20, no. 2, pp. 976–984, 2020, doi: 10.11591/ijeecs.v20.i2.pp976-984.
[18] C. Maier and D. Simovici, "Saturated betweenness centrality sets of vertices in graphs," *J. Adv. Inf. Technol.*, vol. 12, no. 4, pp. 287–295, 2021, doi: 10.12720/jait.12.4.287-295.
[19] L. Y. Ann, P. Ehkan, M. Y. Mashor, and S. M. Sharun, "FPGA-based architecture of hybrid multilayered perceptron neural network," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 14, no. 2, pp. 949–956, 2019, doi: 10.11591/ijeecs.v14.i2.pp949-956.
[20] P. Brodka, K. Musial, and P. Kazienko, "A performance of centrality calculation in social networks," *CASoN 2009 - Int. Conf.*

Comput. Asp. Soc. Networks, no. January 2014, pp. 24–31, 2009, doi: 10.1109/CASoN.2009.20.
[21] J. Zhang and Y. Luo, "Degree centrality, betweenness centrality, and closeness centrality in social network," *In 2017 2nd international conference on modelling, simulation and applied mathematics (MSAM2017)*, vol. 132, no. Msam, pp. 300–303, 2017, doi: 10.2991/msam-17.2017.68.
[22] E. Estrada, "Graph and network theory," *Dev. Water Sci.*, vol. 32, no. C, pp. 317–339, 1988, doi: 10.1016/S0167-5648(08)70928-6.
[23] Y. Y. Liu, J. J. Slotine, and A. L. Barabási, "Controllability of complex networks," *Nature*, vol. 473, no. 7346, pp. 167–173, 2011, doi: 10.1038/nature10011.
[24] Z. Ghalmane, C. Cherifi, H. Cherifi, and M. El Hassouni, "Centrality in complex networks with overlapping community structure," *Sci. Rep.*, vol. 9, no. 1, pp. 1–29, 2019, doi: 10.1038/s41598-019-46507-y.
[25] C. Chekuri, A. Ene, and A. Vakilian, "Node-weighted network design in planar and minor-closed families of graphs," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7391 LNCS, no. PART 1, pp. 206–217, 2012, doi: 10.1007/978-3-642-31594-7_18.

BIOGRAPHIES OF AUTHORS

Ali Ali Saber received the B.Sc. degree in Computer Science from Kirkuk University, Iraq, in 2018 and the M.S. degrees in computer software engineering from Tehran University, Tehran, Iran, more than three years of experience as a lecturer in the field of computer engineering including experience in software programs, over eight years of experience working in HR section, A key team member with strong leadership and ability to work under pressure. Trilingual with fluent verbal and written skills in Arabic, Turkish, English, Persian and Kurdish languages. Experienced in dealing with different cultures and nationality. Communicate effectively, thrives on responsibility and challenge. He can be contacted at email: ali.a.saber@uoalkitab.edu.iq.

Noor Kaylan Hamid received the B.Sc. degree in Computer Science from Kirkuk University Iraq in 2007-2008, and M.Sc. Degree in Computer science and information technology Accurate specialty (Mobile Network) Salahaddin University Iraq in 2011-2014. More than two years of experience as a lecturer in the field of computer engineering including experience in software programs, written skills in Arabic, English and Kurdish languages. Experienced in dealing with different cultures and nationality. She can be contacted at email: noor.g.h@uoalkitab.edu.iq.