ISSN No:-2456-2165

A Survey on Centrality Measures of a Network

Seema Varghese[1]
[1] Department of Mathematics, Government Engineering College, Thrissur
APJ Abdul Kalam Technological University, Kerala, India

Abstract:- A node's centrality score is an index that identifies its importance inside a network. Depending on the goal of the study, different writers define different centrality measures. Numerous fields, including sociology, neurobiology, communication networks, electrical networks, etc., find use for these measurements. The goal of this work is to provide a thorough overview of degree, closeness, eigen vector, and betweenness centrality.

Keywords:- *Centrality Measure; Degree Centrality; Closeness Centrality; Eigen Vector Centrality; Betweenness Centrality***.**

I. INTRODUCTION

Bavelas first proposed the idea of centrality in relation to human communication in 1948. He put up a link between structural centrality and influence in group dynamics since he was particularly interested in how people interacted in small groups. Effective group problem-solving and perceptions of leadership were linked to centrality. The idea of centrality, however, has been applied in situations other than only experimental studies of group problem-solving. In an effort to understand political unification in light of the diversity of Indian social life, Cohn and Marriott also used the centrality idea in 1958. They essentially questioned the viability of any form of government in a nation the size and variety of India. They concluded that network centers bound and twisted different strands into a planned framework, which connected every facet of Indian social life..

The idea of centrality is still applicable today and is being applied in more and more situations. Everyone seems to concur that centrality is an essential structural element of social networks. But there isn't much agreement on how to measure centrality effectively, and there isn't agreement on what centrality actually is or what its conceptual foundations are either. Over the years, numerous alternative centrality metrics have been proposed. The establishment of metrics should aid in idea clarification by describing a concept's components and how they connect to one another. In the instance of centrality, it appears that the opposite outcome was achieved. The complexity of many of the measurements makes it difficult or impossible to tell what, if anything, they are measuring.

Networks are pervasive in our modern world, encompassing a wide spectrum of applications, from social interactions on social media platforms to the structure of transportation systems and the internet. Understanding the structure and dynamics of these networks is crucial for various fields, which include social science, biology, economics, and computer science. Centrality measures play a fundamental role in network analysis, providing insights into the relative importance and influence of nodes within a network. In this survey, we will delve into centrality measures, their types, applications, and significance in the study of networks. Centrality measures aim to quantify the importance of nodes within a network. These measures are based on the fact that not all nodes are equal; some nodes have a more significant impact on the network's structure and functionality. Centrality measures provide a way to identify and analyze these influential nodes.

Centrality measures have a wide range of applications across various domains. In social networks, centrality measures help identify key influencers, opinion leaders, and individuals who play pivotal roles in the spread of information or behaviors. Centrality measures assist in identifying critical transportation hubs and optimizing routes for efficient transportation systems. In biological networks such as protein-protein interaction networks, centrality measures can identify essential proteins for understanding disease mechanisms and drug target identification. In web search algorithms, centrality measures like PageRank help rank web pages, improving the relevance of search results. In business and organizational networks, centrality measures can identify key employees or departments responsible for information flow and decision-making.

Centrality measures are essential tools in network analysis for several reasons. They help to pinpoint nodes that, if removed or targeted, can have a significant impact on the network's structure and functionality. They provide insights into how information, influence, or resources flow within a network, aiding in the prediction of network behavior and vulnerabilities. Centrality measures assist in designing interventions or strategies that focus on influential nodes to maximize desired outcomes. In network design, centrality measures help in optimizing the placement of important nodes or resources for efficiency and robustness.

In this paper, four types of centrality measures are discussed viz. Degree Centrality, Closeness Centrality, Betweenness Centrality and Eigen vector Centrality. Degree Centrality, Closeness Centrality, Betweenness Centrality and Eigen vector Centrality. The advantages and disadvantages of these centrality measures are surveyed in detail. For all basic graph theoretic terminologies, we refer to [17].

II. DEGREE CENTRALITY

Some of the numerous nodes that make up a network are essential in mediating a huge number of network connections. These nodes play a crucial role in network structure and are frequently recognised by numbers referred to as centrality metrics. The concept of degree centrality was put forward by Bavelas in 1948 to make clear how the structural position of an individual within a social network determines his or her influence in group social activities.

In 1979, Freeman pointed out that there is a common theme across different centrality metrics. It is that the fact that they give the same results when applied to a star graphs. Three fundamental properties can be ascribed to the central vertex. "It has the maximum possible degree. It falls on the shortest possible topological path between all pairs of vertices. it is located at the shortest topological distance from all other vertices" [18].

The first step in network analysis is to comprehend how connection fluctuates between nodes. The number of connections each vertex has to the other vertices in the network, known as vertex degree, is may be the simplest metric we can compute in this context. The distribution of degree values across vertices is frequently heterogeneous in real-world networks; we frequently see that many vertices have few linkages while a smaller number of important vertices receive the majority of connectivity, designating them as putative hubs that allow integration throughout the network.

The Degree Centrality measure of a node *i* is denoted as

$C_D(i) = \sum A_{ij}$

where A is the adjacency matrix. This definition assumes that vertices with many connections exert more influence over network functions than others.

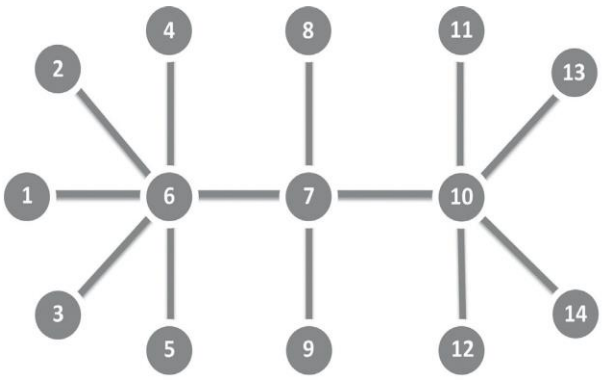

Fig :1 Vertex 6 has highest degree centrality

The limitations of this centrality measure are that all connections are treated alike. It counts quantity but does not consider quality. For example, in Fig :1, node 10 has more degree than node 7. But a closer look will reveal that four of the five neighbours of node 10 are of degree one and hence have less influence in the network. Whereas, node 7 has only four neighbours and two among them are high degree vertices which can dominate the whole graph.

III. CLOSENESS CENTRALITY

A useful metric called closeness centrality predicts how quickly data would go from one vertex to another. How short the geodesics are from vertex *i* to all other vertices is a measure of closeness centrality [18]. Two vertices are topologically close if they are connected by a short path. More generally, a vertex has high closeness centrality if it is connected, on average, by short paths to a large number of other vertices in the network. Vertex efficiency and closeness are equivalent measures.

Closeness centrality is typically stated as "the inverse of the normalised sum of the topological distances in the graph". The farness between the vertices is another name for this parameter [18]. Sometimes closeness centrality is just written as the inverse of distance.

The Closeness Centrality measure of a node *i* is denoted as

$C_C(i) = N\text{-}1 / \sum d_{ij}$

where N is the number of nodes and d_{ij} is the shortest distance between the ith and jth node.

Closeness centrality is a concept in network analysis and graph theory that measures the centrality or importance of a vertex within a network based on its proximity to other vertices. It quantifies how close a vertex is to all other vertices in the network. In essence, it identifies vertices that can quickly reach other vertices in the network with the shortest path.

Vertices with high closeness centrality are considered to be more central within the network because they can reach other vertices more quickly. They are often important for information flow and communication within the network. The closeness centrality calculation is based on the concept of the shortest path, which is the minimum number of edges or steps required to reach one vertex from another. Vertices with lower cumulative shortest path distances tend to have higher closeness centrality. Closeness centrality values are typically normalized by dividing them by (N-1), where N is the total number of vertices in the network. This normalization ensures that the centrality values are comparable across networks of different sizes. A vertex with a closeness centrality value close to 1 is considered highly central, while a vertex with a value closer to 0 is less central. Closeness centrality is one of several centrality measures used in network analysis. It helps identify important vertices in various types of networks, including social networks, transportation networks, and communication networks.

ISSN No:-2456-2165

IV. BETWENNESS CENTRALITY

Betweenness centrality is a crucial component of the analysis of social networks [1, 2], computer networks [3], and numerous other kinds of network data models [4–9]. A unit's closeness to other units in a communication network is not its sole important property. It is more important which units are situated on the geodesics, or shortest routes, between pairs of other units. These organizations have control over the network's information flow. As a measure of a vertex's ability to influence communication, betweenness centrality can be useful.

By measuring how much a vertex is situated on the geodesics joining pairs of other vertices, betweenness centrality [10–14] measures the betweenness of a vertex in a network. In many situations in the actual world, it is quite important. When there is only a single geodesic between any pair of vertices determining betweenness is simple and straightforward and the internal vertices of the geodesic have complete control over communication between pairs of others. When there are several geodesics connecting a pair of vertices, the situation worsens and the control of the internal vertices gets fragmented.

Bavelas initially established the idea of betweenness centrality in 1948 [15]. The potential of a vertex to regulate information flow in the network is the significance of the vertex centrality idea. The degree to which they stand out from others and have the potential to assist, obstruct, or influence the flow of signals makes positions considered to be structurally central. Freeman divided betweenness centrality into three categories in his papers [5, 16]. Two vertex centrality indices—one based on counts and one based on proportions—as well as a measure of network or graph centralization overall are included in the three measurements.

“Betweenness centrality, C_B(V) for a vertex i is defined as

$$C_B\,(i) = \Sigma\, \sigma_{st}\,(i) \,/\, \sigma_{st}$$

where σ_{st} is the number of shortest paths with vertices s and t as their end vertices, while $\sigma_{st}(i)$ is the number of those shortest paths that include vertex i” [16]. High centrality value is an indication that a node lies on a large fraction of geodesics connecting pairs of nodes [19]. Each pair of nodes in a connected network provides a value between 0 and 1 to the betweenness centrality of all other nodes. If there is only a single shortest path joining a particular pair of vertices, then that pair provides a betweenness centrality 1 to each of its internal nodes and zero to all other nodes. For example, in a path graph, a pair of nodes gives a betweenness centrality value 1 to each of its internal nodes and zero to the external nodes. If there are k geodesics of length 2 joining a pair of nodes, then that pair of nodes provides a betweenness centrality $1/k$ to each of the internal nodes.

As the universal vertex of a star is located on the geodesic (which is unique) connecting every pair of other vertices, Freeman [16] shown that the central vertex is the only one in a star to obtain the maximum value taken by C_B(i). The number of these geodesics, which is ${}^{n-1}C_2$ determines the central vertex's betweenness centrality in a star S_n (Fig:2) with n vertices. Since no pendant vertex is located between any geodesic, their betweenness centrality is zero. Once more, it is clear that in a complete network K_n, the betweenness centrality of any vertex is 0 because there are no vertex-geodesic intersections because each geodesic has a length of 1.

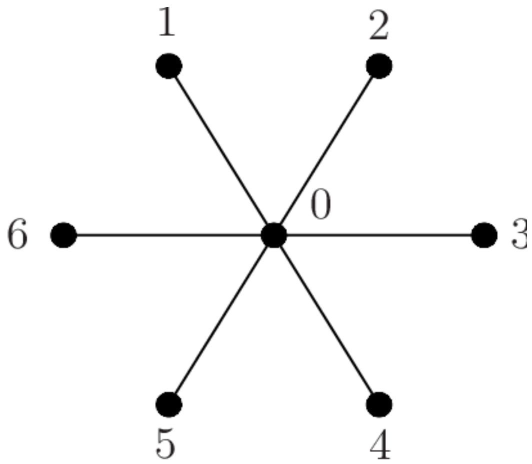

Fig: 2 The Star Graph S_6

V. EIGEN VECTOR CENTRALITY

Eigenvector centrality is a concept in network analysis and graph theory that measures the centrality or importance of a vertex within a network based on the idea that the importance of a vertex depends on the importance of its neighbors. In other words, a vertex is considered central if it is connected to other central vertices.

“Mathematically, eigenvector centrality [18] is calculated using the following formula:
$x(i) = 1 / \lambda * \Sigma(j)\ A(i, j) * x(j)$

where: x(i) is the eigenvector centrality of node i, λ is the largest eigenvalue of the adjacency matrix of the network. A(i, j) is the element in the adjacency matrix that represents the connection between node i and node j. Σ(j) represents the sum over all neighbouring nodes j of node i.”

Eigenvector centrality considers not just the number of connections a vertex has (as in degree centrality) but also the importance of the vertices to which it is connected. In other words, it values connections to vertices that are themselves central. Eigenvector centrality is calculated iteratively. The centrality of a vertex depends on the centrality of its neighbours, and this process continues until it converges to a stable solution.

The largest eigenvalue (λ) of the adjacency matrix plays a crucial role in the calculation. The centrality values are scaled by this eigenvalue, which ensures that all centrality values are positive and allows for comparisons among nodes. Nodes with higher eigenvector centrality values are considered more central or influential within the network. A node with a high eigenvector centrality value is not only

ISSN No:-2456-2165

connected to many other nodes but is also connected to nodes that are themselves central.

Eigenvector centrality is commonly used in various fields, including social network analysis, web page ranking (PageRank is a variant of eigenvector centrality), and the study of information flow in networks. It helps identify key nodes that can exert significant influence or control over the network due to their connections to other influential nodes.

VI. CONCLUSION

For the analysis and comprehension of complex networks in many different disciplines, centrality measurements are essential tools. They enable us to discover the underlying structure, pinpoint important participants, and decide on network architecture, optimisation, and interventions with knowledge. The study of centrality metrics continues to be essential to understanding networks and how they affect our lives as our world gets more interconnected.

REFERENCES

[1]. E. Otte and R. Rousseau, "Social network analysis: a powerful strategy, also for the information sciences," *Journal of Information Science*, vol. 28, no. 6, pp. 441–453, 2002.

[2]. K. Park and A. Yilmaz, "A social network an analysis approach to analyze road networks," in Proceedings of the ASPRS Annual *Conference*, San Diego, Calif, USA, 2010.

[3]. V. Latora and M. Marchiori, "A measure of centrality based on network efficiency," *New Journal of Physics*, vol. 9, article 188, 2007.

[4]. E. Estrada, "Virtual identification of essential proteins within the protein interaction network of yeast," *Proteomics*, vol. 6, no.1, pp. 35–40, 2006.

[5]. L. C. Freeman, "Centrality in social networks conceptual clarification," *Social Networks*, vol. 1, no. 3, pp. 215–239, 1979.

[6]. D. Kosch¨utzki and F. Schreiber, "Centrality analysis methods for biological networks and their application to gene regulatory networks," Gene Regulation and Systems Biology, vol. 2008, no.2, pp. 193–201, 2008.

[7]. A. M. M. Gonz´alez, B. Dalsgaard, and J. M. Olesen, "Centrality measures and the importance of generalist species in pollinationnetworks," Ecological Complexity, vol. 7, no. 1, pp. 36–43, 2010

[8]. M. E. J.Newman, "Scientific collaboration networks. II. sho rtest paths, weighted networks, and centrality," Physical Review E,vol. 64, no. 1, pp. 1–7, 2001

[9]. M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: uses and interpretations," NeuroImage, vol.52, no. 3, pp. 1059–1069, 2010.

[10]. P. Bonacich, "Power and centrality: a family of measures," American Journal of Sociology, vol. 92, no. 5, pp. 1170–1182, 1987.M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[11]. S. P. Borgatti, "Centrality and network flow," Social Networks, vol.27, no.1, pp.55-71, 2005.

[12]. S. P. Borgatti and M. G. Everett, "A graph-theoretic perspective on centrality," *Social Networks*, vol. 28, no. 4, pp. 466–484, 2006.

[13]. E. Estrada, D. J. Higham, and N. Hatano, "Communicability betweenness in complex networks," *Physica A: Statistical Mechanics and its Applications*, vol. 388, no. 5, pp. 764–774,2009

[14]. S. Gago, J. C, Hurajova, and T. Madaras, "On betweenness-uniform graphs," *Czechoslovak Mathematical Journal*, vol. 63, no. 3, pp. 629–642, 2013.

[15]. A. Bavelas, "A mathematical model for group structure. Human organization," *Applied Anthropology*, vol. 7, no. 3, pp. 16–30, 1948.

[16]. L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, pp. 35–41, 1977.

[17]. F. Harary, *Graph Theory*, Addison-Wesley, USA, 1969

[18]. A. Fornito, A. Zalesky, E. T. Bullmore, *Fundamentals of Brain Network Analysis*, Academic Press, UK,2016

[19]. S. Kumar, R. Unnithan, B. Kannan, M. Jathavedan," Betweenness Centrality in some classes of graphs", International Journal of Combinatorics,,Volume 2014, Article ID 241723, 12 pages http://dx.doi.org/10.1155/2014/241723