STM JOURNALS

International Journal of
Data Structure Studies

Volume 1, Issue 2, 2023
July—December
DOI (Journal): 10.37591/IJDSS

https://journals.stmjournals.com/ijdss

Review IJDSS

Centrality in Social Network Analysis: A Comprehensive Review for Students

Rajesh Yadav*

Abstract

Centrality, a fundamental concept in social network analysis (SNA), plays a pivotal role in understanding the structural dynamics and information flow within networks. This study offers an extensive examination of centrality metrics and their utilization in diverse fields. We begin by defining centrality and exploring its significance in characterizing the importance of nodes within a network. Subsequently, we delve into the literature review by different authors and most employed centrality metrics, including degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality, elucidating their underlying mathematical foundations and practical implications. Furthermore, the study discusses the strengths and limitations of each centrality measure, offering insights into their suitability for different network structures and research objectives. Additionally, we examine real-world applications of centrality in fields such as social sciences, epidemiology, organizational behavior, and information retrieval, highlighting the diverse contexts in which centrality analysis proves invaluable. This review serves as a comprehensive resource for researchers, practitioners, and enthusiasts in the field of social network analysis, offering a nuanced understanding of centrality measures and their broad-ranging applications across disciplines. Major focus is to make students in undergraduate degree and post-graduation aware about centrality measures in social network analysis.

Keywords: Degree centrality, betweenness centrality, SNA, centrality metrics

INTRODUCTION

In the realm of Social Network Analysis (SNA), the concept of centrality stands as a cornerstone, providing essential insights into the pivotal nodes and their influence within a network. Centrality metrics play a crucial role in understanding the dynamics of information flow, identifying key players, and characterizing structural significance. This comprehensive review aims to delve deeply into the diverse facets of centrality measures, elucidating their mathematical underpinnings, applications, and evolving extensions. Centrality, at its core, quantifies the prominence of nodes within a network, offering a quantitative measure of their importance. The review will commence by establishing a fundamental comprehension of centrality metrics, encompassing concepts such as degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality. Each of these metrics offers a unique perspective on node importance, considering factors such as direct connections, intermediary role, proximity to others, and influence within the network.

Furthermore, we will explore the practical implications and real-world applications of centrality measures. As SNA continues to evolve, so does the concept of centrality. Recent developments have led to the emergence of dynamic centrality measures, capable of capturing temporal changes in network structures, as well as multiplex centrality, which addresses the complexities of networks with multiple layers of relationships.

***Author for Correspondence**
Rajesh Yadav
E-mail: ry280888@gmail.com

Assistant Professor, Department of Computer Science, SIES College of Arts, Science and Commerce (Autonomous), Mumbai, Maharashtra, India

Received Date: September 20, 2023
Accepted Date: October 03, 2023
Published Date: October 25, 2023

Citation: Rajesh Yadav. Centrality in Social Network Analysis: A Comprehensive Review for Students. International Journal of Data Structure Studies. 2023; 1(2): 1–8p.

REVIEW OF LITERATURE BY DIFFERENT AUTHORS

1. *Author:* Freeman
   - *Title:* "Centrality in Social Networks: Conceptual Clarification"
   - *Year:* 1978
   - *Summary:* Freeman's seminal work provides a comprehensive review of centrality measures in social network analysis. It provides conceptual clarity and explores various forms of centrality, which encompass degree centrality, closeness centrality, and betweenness centrality [1].
2. *Author:* Borgatti and Everett
   - *Title:* "Models of Core/Periphery Structures"
   - *Year:* 2000
   - *Summary:* This literature review by Borgatti and Everett focuses on core/periphery structures within networks. It explores centrality measures and their role in identifying central and peripheral nodes in social networks [2].
3. *Author:* Opsahl *et al*.
   - *Title:* "Node Centrality in Weighted Networks: Generalizing Degree and Shortest Paths"
   - *Year:* 2010
   - *Summary:* This review delves into centrality measures in weighted networks, emphasizing the challenges and adaptations needed for such networks. It discusses methods for generalizing degree and shortest path centrality to weighted networks [3].
4. *Author:* Brass
   - *Title:* "Being in the Right Place: A Structural Analysis of Individual Influence in an Organization"
   - *Year:* 1984
   - *Summary:* Brass' review focuses on structural aspects of centrality in organizational networks. It explores how individual positionality within the network structure influences their level of influence and access to resources [4].
5. *Author:* Sabidussi
   - *Title:* "The Centrality Index of a Graph"
   - *Year:* 1966
   - *Summary:* Sabidussi's work is a foundational paper that introduces the concept of centrality index in graph theory. It provides mathematical formulations for measuring centrality and discusses its implications in network analysis [5].

DEFINING CENTRALITY AND CHARACTERIZING ITS SIGNIFICANCE

In essence, centrality addresses the question: "How central or critical is a node in the overall structure of the network?" It helps us identify which nodes play pivotal roles in facilitating information flow, mediating interactions, or connecting disparate parts of the network.

Understanding centrality is crucial because it allows us to pinpoint key individuals, organizations, or entities that have a significant impact on the network's structure and dynamics. These central nodes can act as information hubs, opinion leaders, or critical intermediaries, influencing the flow of information, resources, and behaviors within the network. By characterizing the importance of nodes, centrality analysis provides valuable insights into various aspects of network behavior, such as the spread of information, the flow of resources, and the dynamics of influence and control. It forms the foundation for numerous applications in fields ranging from sociology and epidemiology to organizational behavior and information retrieval [6].

CENTRALITY METRICS: EXPLORING MATHEMATICAL FOUNDATIONS AND PRACTICAL IMPLICATIONS

Centrality metrics play a central role in the field of Social Network Analysis, offering valuable perspectives on the significance and impact of nodes within a network. This section will delve into four key centrality measures: degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality, unravelling their mathematical underpinnings and shedding light on their practical significance [7].

Degree Centrality

- *Mathematical foundation:* Degree centrality measures the count of direct connections a node maintains within a network. In the context of a directed network, it is further divided into in-degree (representing incoming links) and out-degree (representing outgoing links), as shown in Figure 1.
- *Practical implications:* Nodes with high degree centrality are considered hubs, acting as focal points for information flow. These nodes are essential for the efficient distribution of information and are likely to exert a substantial influence on other nodes as shown in Table 1.

Betweenness Centrality

- *Mathematical foundation:* Betweenness centrality identifies nodes that act as critical bridges or intermediaries in a network. It is calculated by considering the count of shortest paths that traverse through a node as shown in Figure 2.
- *Practical implications:* Nodes with high betweenness centrality serve as connectors between disparate clusters or groups. Their removal could disrupt communication and information flow, making them strategically important in maintaining network cohesion [8].

High betweenness centrality nodes act like ‘bridges’ in the network and have control over the flow passing between others as shown in Table 2.

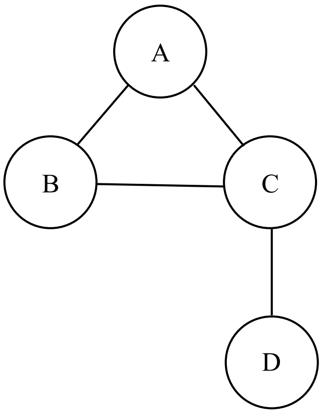


Figure 1. An undirected graph with four vertices.

Table 1. Data summary of degree centrality.

	Adjacency matrix				Degree (# of edges)	N-1 (# of nodes–1)	Degree centrality: degree/(N–1)
	A	***B***	***C***	***D***			
A	0	1	1	0	2	3	2/3
B	1	0	1	0	2		2/3
C	1	1	0	1	3		1
D	0	0	1	0	1		1/3

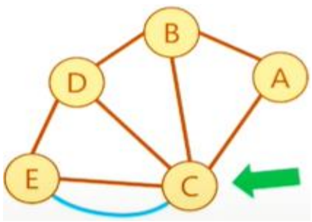


Figure 2. Betweenness centrality graph.

Table 2. Data summary of betweenness centrality.

Node	**Betweenness centrality**	**Degree**	**Avg. shortest distance between node pairs**
C	0.36	5	1.67
B	0.08	3	1.33
D	0.05	3	1.33
A	0	2	1.17
E	0	3	1.17

Closeness Centrality

- *Mathematical foundation:* Closeness centrality evaluates how rapidly a node can connect with all other nodes within the network. This calculation is based on the average length of the shortest paths from a node to all other nodes as shown in Figure 3.
- *Practical implications:* Nodes with high closeness centrality are efficient disseminators of information, as they can quickly reach a large portion of the network. They are crucial for rapid information diffusion or resource allocation [9].

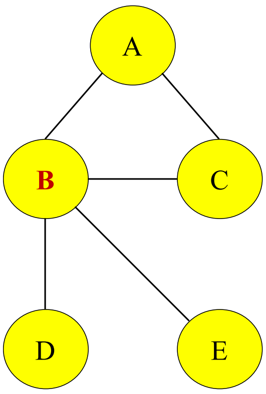
Node B is the most centrally closed node according to these parameters.

Eigenvector Centrality

- *Mathematical foundation:* Eigenvector centrality considers both the direct connections of a node and the centrality of its neighbors. This calculation relies on the concept of eigenvector centrality in the field of linear algebra.
- *Practical consequences:* Nodes with high eigenvector centrality are effectively connected to other nodes that hold significant centrality themselves. This signifies a type of indirect influence, as these nodes are linked to influential nodes in the network as shown in Figure 4 [10].

$$CC(i) = \frac{N-1}{\sum_j d(i,j)}$$

where
$i \neq j$, d_{ij} is the length of the shortest path between nods i and j in the network,
N is the number of nodes.



	A	*B*	*C*	*D*	*E*	farness $\sum_{j=1}^{n} d(i,j)$	$CC(i) = \frac{N-1}{\sum_j d(i,j)}$
A	0	1	1	2	2	6	(5-1)/6 = 0.67
B	1	0	1	1	1	4	1.00
C	1	1	0	2	2	6	0.67
D	2	1	2	0	2	7	0.57
E	2	1	2	2	0	7	0.57

N = 5 (# of nodes)

Figure 3. Closeness centrality.

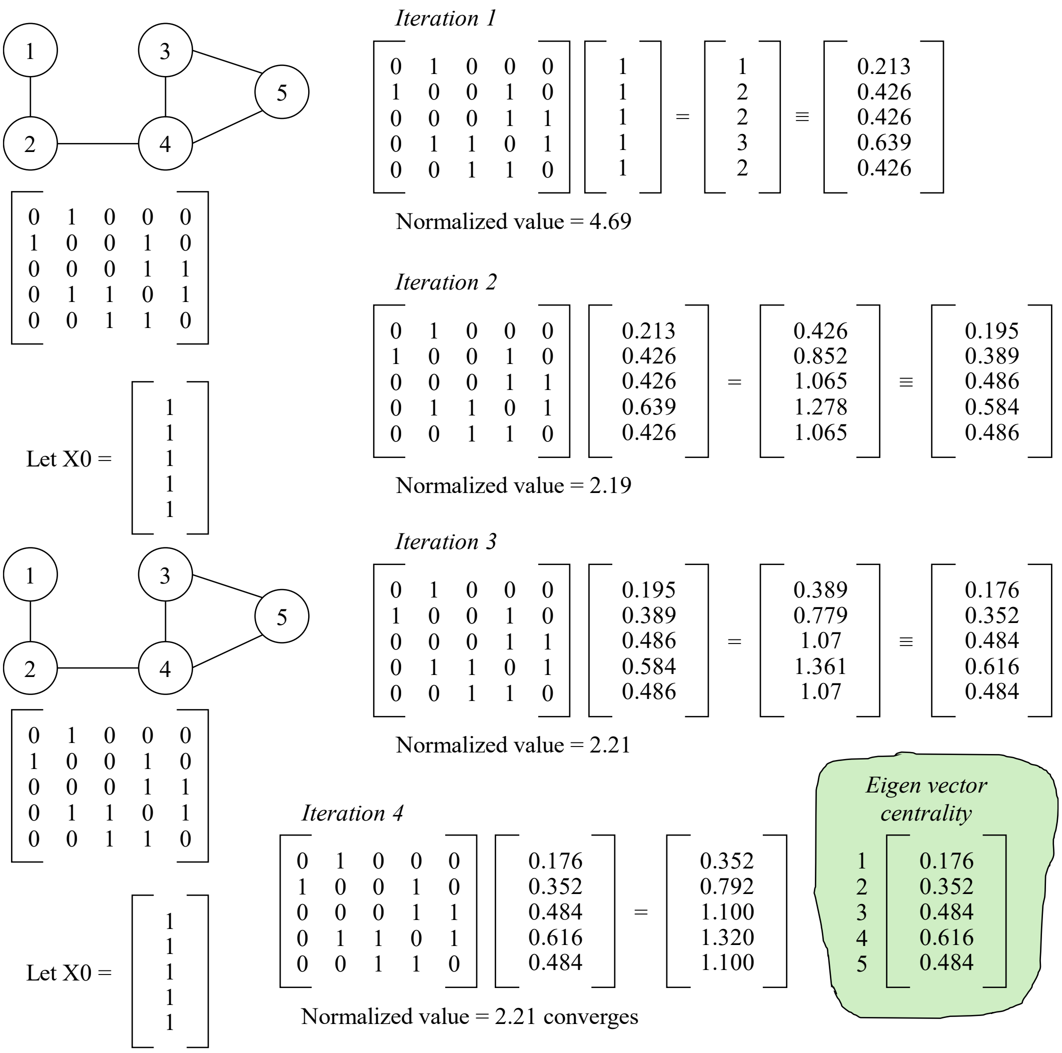


Figure 4. Eigenvector centrality.

STRENGTHS AND LIMITATIONS OF EACH CENTRALITY MEASURE

Degree Centrality

- *Strengths*
  - *Simplicity:* Degree centrality is easy to compute and interpret, making it a straightforward measure of node importance.
  - *Identifying hubs:* It effectively identifies nodes with a high number of direct connections, which often serve as important hubs for information flow.
- *Limitations*
  - *Ignores indirect influence:* It does not account for the influence of nodes that are indirectly connected but have a high centrality themselves.
  - *Not sensitive to edge weights:* It treats all connections as equal, which may not reflect the true strength of relationships in weighted networks.

Betweenness Centrality

- *Strengths:*
  - *Identifying bridge nodes:* It effectively identifies nodes that act as critical bridges or intermediaries in a network, facilitating communication between otherwise disconnected clusters.
  - *Robustness testing:* It helps identify nodes whose removal may have a significant impact on network connectivity.

- *Limitations:*
  - *Computationally intensive:* Calculating betweenness centrality can be computationally expensive, especially in large networks.
  - *Sensitivity to network structure:* It may overemphasize nodes in networks with many parallel paths, potentially leading to inaccurate assessments of centrality.

Closeness Centrality

- *Strengths:*
  - *Efficiency in information flow:* It identifies nodes that can efficiently disseminate information to the entire network, making them crucial for rapid information diffusion.
  - *Closeness to central locations:* It measures proximity to central areas in the network, which can be important for certain applications.
- *Limitations:*
  - *Sensitivity to isolated nodes:* In networks with isolated nodes or disconnected components, closeness centrality may yield misleading results.
  - *Dependence on network structure:* It may not accurately reflect centrality in networks with irregular or non-standard topologies.

Eigenvector Centrality

- *Strengths:*
  - *Accounts for indirect influence:* It considers both direct connections and the centrality of connected nodes, capturing a form of indirect influence.
  - *Identifying influential individuals:* It identifies nodes that are connected to other highly central nodes, potentially representing individuals with significant influence.
- *Limitations:*
  - *Computationally Intensive:* Calculating eigenvector centrality can be computationally demanding, especially in large networks.
  - *Sensitivity to Network Density:* It may overestimate the centrality of nodes in sparse networks and underestimate it in dense networks.

Each centrality metric possesses its unique advantages and constraints, rendering them more suitable for particular network configurations and research goals. Grasping these subtleties is essential when choosing the most fitting centrality metric for a specific analysis.

REAL-WORLD APPLICATIONS OF CENTRALITY ANALYSIS IN VARIOUS FIELDS

Social Sciences

- *Identifying influential individuals:* In social networks, centrality analysis helps identify individuals who wield significant influence over their peers. This is crucial for understanding social dynamics, opinion formation, and the spread of behaviors or ideas.
- *Community detection:* Centrality metrics aid in the detection and characterization of communities or groups within social networks. This is valuable for studying group dynamics, identifying subcultures, or understanding patterns of interaction.
- *Studying information diffusion:* Centrality analysis is used to study how information or behaviors propagate through social networks. It helps identify key nodes that play a central role in the spread of information.

Epidemiology

- *Identifying disease spread:* In epidemiology, centrality metrics are used to identify individuals or locations that are most likely to facilitate the spread of infectious diseases. This data is essential for precise interventions and preventive strategies.
- *Contact tracing:* Centrality analysis aids in contact tracing efforts during disease outbreaks.

- *Optimizing vaccination strategies:* Centrality measures help identify key nodes for vaccination campaigns, ensuring that the most influential individuals are targeted to maximize the impact of vaccination efforts.

Organizational Behavior

- *Identifying key influencers:* In organizational networks, centrality analysis helps identify employees who serve as key influencers or information brokers. This information is useful for understanding communication patterns and decision-making dynamics.
- *Optimizing communication channels:* Centrality analysis assists in identifying the most effective communication channels within an organization. This helps improve information flow and collaboration.
- *Leadership identification:* Centrality metrics are used to identify potential leaders within an organization based on their network centrality. This can aid in succession planning and leadership development.

Information Retrieval

- *Search engine ranking:* In web networks, centrality metrics are used by search engines to rank the importance of webpages. Pages with greater centrality are more likely to be ranked higher in search results.
- *Recommender systems:* Centrality analysis is used in recommendation algorithms to identify influential users or items in social networks or e-commerce platforms. This contributes to enhancing the quality of personalized suggestions.
- *Identifying hubs in citation networks:* In academic citation networks, centrality analysis helps identify highly influential papers or authors, providing insights into the impact of research.

CONCLUSION

In conclusion, this comprehensive review has provided a thorough exploration of centrality in social network analysis, illuminating its fundamental role in understanding the structure, dynamics, and influence within networks. Through an exploration of crucial centrality metrics such as degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality, we have revealed their mathematical underpinnings and real-world significance across various scenarios.

Degree centrality, as a measure of direct connections, serves as a foundational metric for identifying hubs and pivotal nodes within a network. Betweenness centrality illuminates critical bridge nodes that mediate communication pathways, offering insights into network resilience and efficiency. Closeness centrality, by evaluating node proximity, identifies efficient disseminators of information and central players in the network. Eigenvector centrality, accounting for both direct and indirect influence, reveals nodes with a form of indirect sway over the network.

In summary, centrality analysis is a cornerstone of social network analysis, providing invaluable insights into the dynamics of relationships and influence within networks. Through comprehension of the advantages and drawbacks associated with various centrality metrics, researchers and professionals can make informed choices and derive valuable insights from intricate network data. This comprehensive review serves as a foundational resource for those seeking a nuanced understanding of centrality in the realm of social network analysis.

REFERENCES

1. Freeman Linton C. Centrality in social networks conceptual clarification. Soc Networks. 1978; 1(3): 215–239. Available at doi: https://doi.org/10.1016/0378–8733(78)90021–7
2. Borgatti Stephen P, Everett Martin G. Models of core/periphery structures. Soc Networks. 2000; 21(4): 375–395. Available at doi: https://doi.org/10.1016/S0378–8733(99)00019–2.

3. Opsahl Tore, Agneessens Filip, Skvoretz John. Node Centrality in Weighted Networks: Generalizing Degree and Shortest Paths. Soc Networks. 2010; 32(3): 245–251. 10.1016/j.socnet.2010.03.006.
4. Brass Daniel J. Being in the Right Place: A Structural Analysis of Individual Influence in an Organization. Adm Sci Q. 1984; 29(4): 518–39. JSTOR, https://doi.org/10.2307/2392937.
5. Sabidussi Gert. The centrality index of a graph. Psychometrika. 1966; 31(4): 581–603.
6. Frost H. (2022). Eigenvector centrality for multilayer networks with dependent node importance. arXiv:2205.01478 [physics.soc-ph. 2022 May 3.
7. Das Kousik, Sovan Samanta, Madhumangal Pal. Study on centrality measures in social networks: a survey. Soc Netw Anal Min. 2018; 8(1): 13.
8. Landherr Andrea, Bettina Friedl, Julia Heidemann. A critical review of centrality measures in social networks. Wirtschaftsinformatik. (Bus Inf Syst Eng). 2010; 2(6): 371–385.
9. Riquelme F, Gonzalez-Cantergiani P, Molinero X, Serna M. Centrality measure in social networks based on linear threshold model. Knowledge-based Systems. Jan 2018; 140: 92–102.
10. Valente TW, Coronges K, Lakon C, Costenbader E. How correlated are network centrality measures? Connect (Tor). 2008 Jan 1; 28(1): 16–26.

