# Centrality Metrics in Graph Node Networks

By: Gargie Tambe, Asish Bharadwaj, Chetan Vellanki

June 2023

Introduction

Centrality measures play a critical role in the field of graph analytics as they provide valuable insights into the importance and influence of individual nodes within a network. By quantifying the centrality or "centrality" of nodes, researchers can identify key entities that have a significant impact on network dynamics, control the flow of information, or act as vital connectors within the network structure.
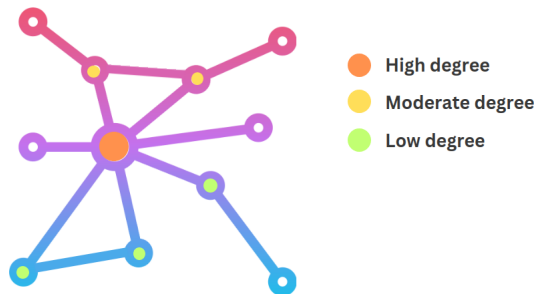
SOCIAL NETWORKING DEGREE CENTRALITY

The concept of centrality encompasses multiple perspectives, allowing researchers to evaluate the significance of nodes from various angles. This versatility is particularly valuable, as it enables the analysis of complex systems such as human brain networks or social networks, where understanding the role and importance of individual nodes is crucial for comprehending the overall system dynamics.

This review aims to explore and evaluate the key characteristics and limitations of commonly used centrality measures, focusing on Degree Centrality, Betweenness Centrality, and Eigenvector Centrality. Additionally, we will demonstrate how centrality measures can be effectively combined with linear algebraic concepts to gain a deeper understanding of the operation and behavior of complex systems.

Furthermore, we will delve into real-world scenarios and practical applications of centrality analysis, showcasing how these measures have been successfully utilized in diverse fields such as social network analysis, disease spread modeling, recommendation systems, and urban planning.

Types of Centrality Metrics

- **Degree Centrality**

  Degree centrality is a fundamental measure in network analysis that quantifies the importance of a node within a graph based on its degree, which represents the number of connections it has to other nodes. It serves as a simple and intuitive indicator of node prominence, with nodes having a higher degree of centrality considered more central or influential in the network. Degree centrality is widely used in various domains, including social network analysis, transportation networks, and biological networks, providing valuable insights into key nodes and network structure. By considering the immediate connections of a node, degree centrality offers a valuable perspective on node importance without implying any negative connotations.

  In a given graph $G(V, E)$, for a given vertex v, the degree centrality is defined as the following -

  $$C_D(v) = deg(v)$$

  Time complexity: Calculating the degree centrality for all nodes in a graph, it takes $\Theta(V^2)$ time in a dense adjacency matrix and for edges, $\Theta(E)$ time in a sparse matrix representation.

  Let $X(Y, Z)$ be the $|Y|$ node connected graph that maximizes the following quantity (where $y*$ is node with highest degree centrality in $|X|$:

  $$H = \sum_{j=1}^{|Y|} [C_D(v*) - C_D(y_j)]$$

  So, the degree centrality of the graph G will be:

  $$C_D(G) = \frac{\sum_{i=1}^{|V|} [C_D(v*) - C_D(V_i)]}{H}$$

  The value of H is maximized when graph X contains one central node to which all other nodes are connected (star topology). In that case,

  $$H = (n-1) \times ((n-1)-1) = n^2 - 3n + 2$$

By utilising this metric, it is possible to identify influential individuals in a network, including those with the most extensive contact networks, those transitioning rapidly between peers, and those with access to the most essential information.

- **Betweenness Centrality**

  Betweenness centrality is a metric that assesses the significance of a node in a network based on its frequency in the geodesic distance or shortest path between all pairs of nodes. It identifies nodes that serve as connectors between distinct groups within the network, playing a pivotal role in controlling the flow of interactions or information. Although originating from social network analysis, betweenness centrality finds applicability in various network types, including social, transportation, and computer networks.

  The betweenness centrality of a vertex v in a graph can be calculated using the following formula, which involves linear algebra:

  $$BC(v) = \sum_{u,v \ \epsilon \ V} \frac{\sigma_{uw}(v)}{\sigma_{uw}}$$

  where V represents the set of all nodes in the graph, σ_{uw} denotes the total number of shortest paths between node u and v, and $\sigma_{uw}(v)$ is the total number of shortest paths between nodes u and w that pass through vertex v.

  Following steps are followed to calculate the betweenness centrality of a vertex v in a graph -

  1. Calculate unique pairs of nodes (unique edges): Start by considering all possible pairs of nodes in the graph. These pairs represent the edges or connections between nodes.

  2. Determine the shortest paths: Find the shortest paths between all pairs of nodes in the graph. This can be done using algorithms such as Dijkstra's algorithm or Floyd-Warshall algorithm. In case of a weighted graph, the paths should be minimal weighted paths.

  3. Calculate the sum of all fractions of shortest paths through vertex v: For each pair of nodes (u, w) calculate the number of shortest paths between them (σ_{uw}) and the number of paths that pass through the vertex v $(\sigma_{uw}(v))$. Then, sum up the fractions $\frac{\sigma_{uw}}{\sigma_{uw}(v)}$ for all pairs of nodes in which vertex v lies on the shortest path.

  4. Normalize the betweenness centrality (BC): If you are comparing networks of different sizes, it is common practice to normalize the betweenness centrality scores. This normalization takes into account the fact that larger networks tend to have higher centrality scores. There are various normalization techniques, such as dividing each

node's betweenness centrality by the maximum possible betweenness centrality value.

The betweenness centrality metric provides a unique perspective on the significance of nodes compared to other centrality measures. It focuses on individuals who have control over the flow of data throughout the network, including brokers, controllers, intermediaries, gatekeepers, and similar roles. Removing nodes with high betweenness centrality from the network would have the most disruptive impact on the interactions within the network.

- **Eigenvector Centrality**

  Eigen Vector Centrality is a metric that assesses the importance of nodes in a graph based on their connections to highly important nodes. It assigns scores to nodes, considering that a node contributes more to its score if it is connected to other nodes with high scores. This measure is calculated using an iterative algorithm that converges to stable scores. Eigenvector Centrality is used in various domains, such as social network analysis, biology, and recommendation systems, to identify influential nodes within a network.

  The Eigenvector Centrality (EVC) of a node in a network can be computed using the adjacency matrix of the network.

  For a graph $G(V, E)$ where $|V|$ is the number of vertices of the graph network. $A = (a_{v,t})$ is the adjacency matrix of the graph. $a_{v,t} = 1$ if there exists an edge between vertex v and vertex t and $a_{v,t} = 0$, otherwise.

  The formula for calculating EVC is as follows:

  $$x_v = \frac{1}{\lambda} \sum_{t \epsilon M(v)} x_t = \frac{1}{\lambda} \sum_{t \epsilon G} a_{v,t} x_t$$

  where $M(v)$ is the set of neighbours of v and λ is the eigenvalue.

  Multiplying both LHS and RHS by λ on both sides, we see that $\sum_{t \epsilon G} a_{v,t} x_t$ is the product of the adjacency matrix A with the vector v i.e. Ax. The LHS part of the product of vector x, i.e. λx.

  Rearranging, we get, $Ax = \lambda x$.

Applications and Related Works

- Degree centrality analysis in social networking plays a pivotal role in understanding the structure and dynamics of social relationships. By measuring the number of connections a node has in a network, degree centrality provides insights into individual popularity, influence, and information dissemination potential. It helps identify influential hubs, brokers, and opinion leaders who serve as bridges between different communities,

facilitating the spread of ideas and social influence. Conversely, it highlights individuals who may be isolated or peripheral, offering insights into barriers to information diffusion and social integration. Overall, degree centrality analysis informs marketing strategies, community engagement initiatives, and interventions aimed at understanding and leveraging social network dynamics.

- Betweenness centrality has diverse applications, and one notable use case is in the analysis of criminal or terrorist networks. Nodes with high betweenness centrality in such networks often indicate crucial or highly involved actors. These nodes exert control over the flow of data among different groups, serving as bridges for information dissemination. Nodes with high betweenness centrality hold favorable positions in the network, facilitating greater control over information propagation. As a consequence, they become potential single points of failure. Disrupting or removing these nodes can effectively hinder communication and disrupt the network's operations.

- The PageRank algorithm incorporates Eigen Vector Centrality as a fundamental element and is extensively employed by major companies, including Google, for website ranking purposes. By considering the World Wide Web as a directed graph, the algorithm evaluates the significance of web pages based on the quantity and quality of incoming links. Eigen Vector Centrality, leveraging eigenvectors and eigenvalues, iteratively computes centrality scores for web pages, ultimately establishing a ranking system that assigns priority to influential pages. This algorithmic application has brought about a paradigm shift in web search, enabling the provision of more accurate and pertinent search results to users, thereby enhancing their search experience.

References

- Centrality Metrics for Graphs - Works Blog
- Graph Centrality Measures: Types and Explanation - Turing
- Centrality - Wikipedia
- Eigenvector Centrality (Centrality Measure) - GeeksforGeeks
- Degree Centrality (Centrality Measure) - GeeksforGeeks
- Betweenness Centrality (Centrality Measure) - GeeksforGeeks

Future Agenda

In our final report, we aim to analyze and understand the centrality of nodes within our network by applying the formulas provided. To achieve this, we will

leverage the power of the Python library Networkx. Networkx offers a robust set of functionalities for graph creation, manipulation, and analysis. By utilizing Networkx, we will be able to plot our graphs and calculate various centrality metrics with ease. This approach will provide us with valuable insights into the importance and influence of nodes within our network.

In summary, our approach of utilizing Networkx in our final report will facilitate a comprehensive analysis of network centrality. Through graph visualization and the calculation of centrality metrics, we will gain a deeper understanding of the network's structure and identify significant nodes. These insights will contribute to a more informed analysis of the network's dynamics and provide valuable implications for various applications, such as social network analysis, transportation networks, or information dissemination systems.

Work Contribution

- Gargie Tambe: Worked on Degree Centrality and designed the mid-report.

- Asish Bharadwaj: Worked on Betweenness Centrality

- Chetan Vellanki: Worked on Eigen-Vector Centrality