NETWORK MEASURES

A. Centrality Measures

Closeness centrality is a centrality measure which examines the distance between the network nodes and ranks them according to their average distance(Freeman, 1978). The average distance is given by $\boldsymbol{l_i = \frac{1}{n}\sum_j d_{ij}}$, where n is the number of nodes and $\boldsymbol{d_{ij}}$ is the shortest distance between nodes i and j. A small average distance indicates that the node can quickly access other nodes and, thus, the specific node is a central one. The formula of closeness centrality is the inverse of the average short distances, $\boldsymbol{C_i = \frac{1}{l_i} = \frac{n}{\sum_j d_{ij}}}$, so that it is consistent with the rest of the measures, where a higher value indicates a higher centrality.

However, some issues occur especially when there are clusters formed in the network. Two nodes that belong to a different cluster would result to an infinite distance and consequently to zero closeness centrality. To bypass this, we could calculate closeness centrality only for nodes in the same cluster. Again, this would result to nodes that belong to small clusters having a closeness centrality higher compared to nodes in larger clusters which is contradictory. To overcome the adversities, the modification of closeness centrality formula has been proposed to $\boldsymbol{C'_i = \frac{1}{n-1}\frac{1}{\sum_{j(\neq i)} d_{ij}}}$, which solves the issue with nodes belonging in different clusters and nodes that are closer to each other obtain a higher score.

Betweenness centrality, is a centrality measure that examines whether a node is used as a bridge to connect other nodes (Freeman, 1977). More specifically, it measures the shortest paths that pass through each node. A node that has high betweenness centrality is an influential one, meaning that a high amount of information disseminates through it. The formula that describes betweenness centrality is $\boldsymbol{x_i = \sum_{st} n_{st}^i}$, where $\boldsymbol{n_{st}^i}$ is the number of shortest paths between nodes s and t that pass through node i. This equation can be transformed to $\boldsymbol{x_i = \sum_{st}\frac{n_{st}^i}{g_{st}}}$, where $\boldsymbol{g_{st}}$ is the number of shortest paths between nodes s and t, so that we take into consideration the total number of shortest paths between two nodes. Finally, a useful transformation would be a normalization, which can be achieved by dividing betweenness centrality with the total number of node pairs, $\boldsymbol{n^2}$.

Eigenvector centrality measures the influence of a node in a network and ranks the nodes according to the number of links that their neighbors form (Bonacich, 2015). The mathematical equation that we use to calculate eigenvector centrality is $\boldsymbol{x_i = \frac{1}{\lambda}\sum_j x_j}$, where λ is a constant that describes the proportionality between the eigenvector centrality of node i and the centrality of its neighbors, and $\boldsymbol{\sum_j x_j}$ is the sum of eigenvector centralities of node’s i neighbors.

B. Integration Measures

Characteristic path length is the average path length between a pair of nodes. It is used to measure the efficiency of information dissemination on a network (Albert &Barabási, 2002). The shorter the average path the quicker it is for the information to travel among the network nodes.

It is given by: $\boldsymbol{l_G} = \frac{\mathbf{1}}{\boldsymbol{n(n-1)}} \sum_{\boldsymbol{i \neq j}} \boldsymbol{d(x_i, x_j)}$, where n is the number of nodes and $\boldsymbol{d(x_i, x_j)}$ the shortest path length between two nodes.

Eccentricity is the largest shortest path between a node and any other node (Hage&Harary, 1995). It is an indicator of the distance between a node and the node that is most distant from it.

Diameter is the maximum of the network's shortest paths (maximum eccentricity).

Radiusis the minimum of the maximum shortest paths between a node to all other nodes (minimum eccentricity).

Global efficiency is used to measure information transfer within a network (Latora&Marchiori, 2001). It is defined as the inverse of the average characteristic path length between the network's nodes, and it can be calculated as $\boldsymbol{E} = \frac{\mathbf{1}}{\boldsymbol{N(N-1)}} \sum_{\boldsymbol{i \neq j}} \frac{\mathbf{1}}{\boldsymbol{d_{ij}}}$.

C. Segregation Measures

Clustering coefficient is used to measure the tendency of nodes to form clusters (Watts &Strogatz, 1998). Local clustering coefficient focuses on each node separately and its clustering ability and is given by $C_i = \frac{\lambda_G(i)}{\tau_G(i)}$, where for node i $\lambda_G(i)$ is the number of triangles that it participates and $\tau_G(i)$ is the number of possible triangles (open or closed triplets). A value of C_i close to zero indicates that the neighbors of node i are not connected to each other, while on the contrary, values close to 1 show that most of the neighbors are also connected to each other. In order to measure the clustering coefficient of the entire network we may average the local clustering coefficients. In addition, there is global clustering coefficient or **transitivity** for the entire network, which is defined as $C = \frac{3 \times number\ \ of\ triangles}{number\ \ of\ open\ and\ closed\ triplets}$. Their difference lies on the fact that the latter favors higher degree nodes, while the first lower ones.

Modularity is used to measure the degree to which the division of a network into communities is a good or bad one (Newman, 2006). A good division would suggest that there are many links within a community and very few among the communities. Modularity is given by $Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta\left(c_i c_j\right)$, where m is the number of links, A is network's adjacency matrix, k is the degree of a node and $\delta\left(c_i c_j\right)$ is 1 if nodes i and j are in the same community and 0 otherwise.

D. Resilience Measures

Local efficiency measures the efficiency of information spreading within local subgraphs or neighborhoods and reveals how efficiently is the information transferred between the direct neighbors of node i when it is removed from the network (Latora&Marchiori,2001). The local efficiency is given by $E_{local} = \frac{1}{N_{Gi}(N_{Gi}-1)} \sum_{jk} \frac{1}{\boldsymbol{d_{jk}}}$, where N_{Gi} is the number of nodes in subgraph/neighbor G_i, and $\boldsymbol{d_{jk}}$ is the shortest average path length of all neighbors of node i.

E. Similarity Measures

Assortativity or assortative mixing applies to networks with nodes of various types or properties, and it shows the preference of nodes of the same type to connect to each other (Newman, 2003). One such property could be the degree value. The assortativity coefficient is given by $r = \frac{Tr\boldsymbol{e} - ||\boldsymbol{e}^2||}{1 - ||\boldsymbol{e}^2||}$, where **e** is the type mixing matrix with element e_{st} such that e_{st} is the number of links that connects nodes of types s and t, and ||x|| represents the sum of all elements of matrix x (cardinality). Values close to 1 suggest that the network is perfect assortative, value equal to 0 suggests that the network is non-assortative and values close to -1 suggest that network is completely disassortative.

References

Albert, R., & Barabási, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, *74*(1), 47. https://doi.org/10.1103/RevModPhys.74.47

Bonacich, P. (2015). Power and Centrality: A Family of Measures. *Https://Doi.Org/10.1086/228631*, *92*(5), 1170–1182. https://doi.org/10.1086/228631

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215–239. https://doi.org/10.1016/0378-8733(78)90021-7

Freeman, L. C. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry*, *40*(1), 35. https://doi.org/10.2307/3033543

Hage, P., & Harary, F. (1995). Eccentricity and centrality in networks. *Social Networks*, *17*(1), 57–63. https://doi.org/10.1016/0378-8733(94)00248-9

Latora, V., & Marchiori, M. (2001). Efficient Behavior of Small-World Networks. *Physical Review Letters*, *87*(19), 198701. https://doi.org/10.1103/PhysRevLett.87.198701

Newman, M. E. J. (2003). Mixing patterns in networks. *Physical Review E*, *67*(2), 026126. https://doi.org/10.1103/PhysRevE.67.026126

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*(23), 8577–8582. https://doi.org/10.1073/PNAS.0601602103

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of “small-world” networks. *Nature*, *393*(6684), 440–442. https://doi.org/10.1038/30918