Betweenness centrality: shortest paths vs.random walks

Network analysis is the method of representing a complex problem using a graph $G(V, E)$, where V is the vertices of the graph and E is the edges of the graph, and then analysing the graph as a network. The centrality of a network can be found using various measures (Borgatti et al., 2005) such as betweenness centrality, first introduced by Freeman (1977), which we will focus on. The betweenness of a node can, in short, be defined as the number of times a vertex i is needed for the path a source vertex s takes to reach it's target vertex t. However there are many variants to consider, most importantly which path will be taken to reach vertex t. Therefore this project aims to compare two different betweenness measures for various types of networks.

The simplest way to calculate betweenness centrality is by using the geodesic paths only and is called shortest path betweenness centrality. If g_{ij} is the geodesic or shortest path from any vertex i to vertex j and g_{ikj} is the geodesic path from vertex i to j which passes through a vertex k, the shortest path betweenness centrality of k can be found by,

$$c_k^{BET} = \sum_i \sum_j \frac{g_{ikj}}{g_{ij}}. \tag{1}$$

Equation (1) shows that shortest path betweenness centrality can be defined as the sum of the fraction of the geodesic between pairs of vertices that pass through k. However, realistically, not only the set of all shortest paths could be considered, for example only shortest paths up to a length k may be considered for some problem which gives rise to k-shortest path betweenness as explored by White and Smyth (2003) and Koschtzki et al. (2005).

Now consider that the walk taken from the vertex s to t is by random walks as there is no necessity that only ideal paths, such as geodesic paths, may be used (Newman, 2005). Take a communication network as an example, if the information takes a random walk then it does not know where it is going and will simply go from vertex to vertex at random until it gets to vertex t and will then stop.

Calculating the random walk betweenness of a network uses the same equations necessary to calculate the current flow betweenness (Newman, 2005). Define $G = (V, E)$ to be an electrical network, with a conductance function $c : V \to \mathbb{R}$, where an external current satisfies Kirchoff's current law and defined by the s-t-supply, $b_{st} : V \to \mathbb{R}$, enters the network at vertex s and leaves at vertex t. Also, let $x : \overrightarrow{E} \to \mathbb{R}$ be a current in the network, where $\overrightarrow{E}$ is the oriented edge set with oriented edges $\overrightarrow{e}$. Then the current-flow betweenness centrality $c_{CB} : V \to \mathbb{R}$ of an electrical network is given by,

$$c_{CB} = \frac{1}{(n-1)(n-2)} \sum_{s,t \in V} \frac{1}{2} \Big(-|b_{st}(v)| + \sum_{e \ni v} |x(\overrightarrow{e})| \Big) \qquad \forall v \in V. \tag{2}$$

Now let $\mathbf{M}$ be the matrix which is the probability that vertex j will send a message to vertex i, where $M_{ij} = \frac{a_{ij}}{d_j}$ if $j \neq t$, and let $\mathbf{D}$ be the degree matrix of the graph so that $\mathbf{D}^{-1}$ has inverted vertex degrees on its diagonal. The probability, $\mathbf{v}^{st}$, to find the information at i when it is on a random walk from vertex s to t is given by $\mathbf{v}^{st} = D_t^{-1}.(I - M_t)^{-1}.\mathbf{s}$, where $\mathbf{s}$ is a vector which is 1 at vertex s and 0 everywhere else. This leads to show that random

walk betweenness centrality is the same as the current-flow betweenness centrality shown in equation (2), hence $c_{RWB}(v) = c_{CB}(v)$.

References

Freeman, L. C. (1977) A set of measures of centrality based on Betweenness, Sociometry, 40(1), p. 35. doi: 10.2307/3033543.

Borgatti, Stephen P.; Everett, Martin G. (2005). "A Graph-Theoretic Perspective on Centrality". Social Networks (Elsevier) 28: 446484. doi:10.1016/j.socnet.2005.11.005.

White, S. and Smyth, P. (2003) Algorithms for estimating relative importance in networks, Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD 03, . doi: 10.1145/956750.956782.

Koschtzki, Dirk; Katharina A. Lehmann, Leon Peeters, Stefan Richter, Dagmar Tenfelde-Podehl, Oliver Zlotowski (2005). Centrality Indices. In Ulrik Brandes, Thomas Erlebach.Network Analysis Methodological Foundations. LNCS 3418. Springer Verlag, Heidelberg, Germany. pp. 1660. ISBN 978-3-540-24979-5

Newman, M. E. J. (2005) A measure of betweenness centrality based on random walks, Social Networks, 27(1), pp. 3954. doi: 10.1016/j.socnet.2004.11.009.

999