Centralities

Ref: https://cambridge-intelligence.com/keylines-faqs-social-network-analysis/

Degree Centrality

Definition: Degree centrality assigns an importance score based purely on the number of links held by each node.

What it tells us: How many direct, 'one hop' connections each node has to other nodes within the network.

When to use it: For finding very connected individuals, popular individuals, individuals who are likely to hold most information or individuals who can quickly connect with the wider network.

A bit more detail: Degree centrality is the simplest measure of node connectivity. Sometimes it's useful to look at in-degree (number of inbound links) and out-degree (number of outbound links) as distinct measures, for example when looking at transactional data or account activity.

Betweenness centrality

Definition: Betweenness centrality measures the number of times a node lies on the shortest path between other nodes.

What it tells us: This measure shows which nodes act as 'bridges' between nodes in a network. It does this by identifying all the shortest paths and then counting how many times each node falls on one.

When to use it: For finding the individuals who influence the flow around a system.

A bit more detail: Betweenness is useful for analyzing communication dynamics, but should be used with care. A high betweenness count could indicate someone holds authority over, or controls collaboration between, disparate clusters in a network; or indicate they are on the periphery of both clusters.

Closeness centrality

Definition: This measure scores each node based on their 'closeness' to all other nodes within the network.

What it tells us: This measure calculates the shortest paths between all nodes, then assigns each node a score based on its sum of shortest paths.

When to use it: For finding the individuals who are best placed to influence the entire network most quickly.

A bit more detail: Closeness centrality can help find good ‘broadcasters’, but in a highly connected network you will often find all nodes have a similar score. What may be more useful is using Closeness to find influencers within a single cluster.

EigenCentrality

Definition: Like degree centrality, EigenCentrality measures a node’s influence based on the number of links it has to other nodes within the network. EigenCentrality then goes a step further by also taking into account how well connected a node is, and how many links their connections have, and so on through the network.

What it tells us: By calculating the extended connections of a node, EigenCentrality can identify nodes with influence over the whole network, not just those directly connected to it.

When to use it: EigenCentrality is a good ‘all-round’ SNA score, handy for understanding human social networks, but also for understanding networks like malware propagation.

A bit more detail: KeyLines calculates each node’s EigenCentrality by converging on an eigenvector using the power iteration method.

PageRank

Definition: PageRank is a variant of EigenCentrality, also assigning nodes a score based on their connections, and their connections’ connections. The difference is that PageRank also takes link direction and weight into account – so links can only pass influence in one direction, and pass different amounts of influence.

What it tells us: This measure uncovers nodes whose influence extends beyond their direct connections into the wider network.

When to use it: Because it factors in directionality and connection weight, PageRank can be helpful for understanding citations and authority.

A bit more detail: PageRank is famously one of the ranking algorithms behind the original Google search engine (the ‘Page’ part of its name comes from creator and Google founder, Sergei Brin).