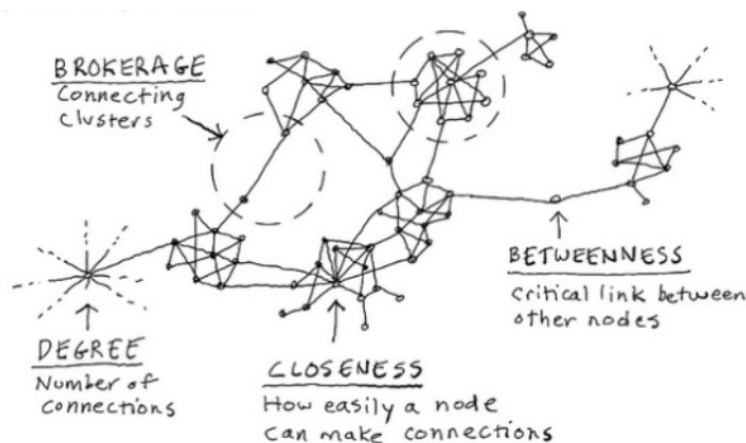


Node ranking by position & centrality

Node centrality measures and *link analysis algorithms* are techniques used to evaluate the importance, position, or structural role of nodes within a network. A variety of approaches exist, including clustering coefficients, spectral analysis methods, geodesic-based measures, fragment-based metrics, and others.



In the following, we assume a simple undirected graph, although most measures can be straightforwardly generalized to directed graphs, multigraphs, or weighted graphs.

Closeness centrality

Closeness centrality quantifies a node's importance based on its proximity to other nodes, meaning that important nodes tend to occupy "central" positions within the network ([Newman, 2018](#)).

It is defined either as

$$CC_i = \frac{1}{n-1} \sum_{j \neq i} \frac{1}{d_{ij}}$$

or

$$CC_i = \frac{n-1}{\sum_{j \neq i} d_{ij}},$$

and can be computed in time $\mathcal{O}(nm)$ by performing a breadth-first search from each node.

Closeness centrality is particularly useful in logistics applications (e.g., transportation networks) but may be less meaningful in social networks. Moreover, in small-world networks with $d \sim \log n$, the measure typically spans a very narrow range.

Betweenness centrality

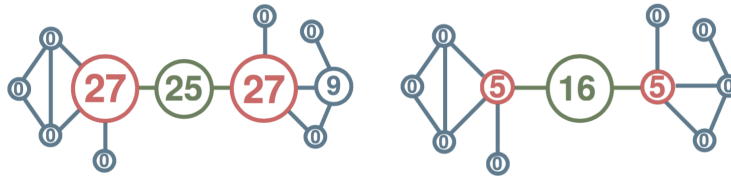
Betweenness centrality quantifies the importance of nodes as bridges connecting different "parts" of a network ([Freeman, 1977](#)).

Let g_{st} denote the number of shortest paths between nodes s and t , and let g_{st}^i be the number of these paths that pass through node i . Then,

$$BC_i = \frac{1}{n^2} \sum_{st} \frac{g_{st}^i}{g_{st}},$$

which can be computed in time $\mathcal{O}(nm)$ using Brandes' algorithm ([Brandes, 2001](#)).

Betweenness centrality is particularly useful in communication networks and other contexts where bridging nodes are critical. Note that the measure considers only shortest paths and may mix local centers with global bridges.



Degree centrality

Degree centrality quantifies node importance based on the number of neighbors.

$$DC_i = d_i = \frac{1}{n-1} \sum_{j \neq i} A_{ij} = \frac{k_i}{n-1}$$

For directed networks (e.g., information networks), the in-degree centrality $d_i^{in} = \frac{k_i^{in}}{n-1}$ is typically more informative of node importance than the out-degree centrality. Note that in scale-free networks with a power-law degree distribution $p_k \sim k^{-\gamma}$, the measure spans a very wide range.

Eigenvector centrality

Eigenvector centrality builds upon the degree centrality by considering not only the number of neighbors but also their importance, so that a node is more central if it is connected to

other central nodes ([Bonacich, 1987](#)).

Therefore, eigenvector centrality is defined as

$$EC_i = e_i = \frac{1}{\lambda} \sum_j A_{ij} e_j,$$

where λ is the leading eigenvalue of A (i.e. $\lambda e = Ae$). The measure can be computed in time $\mathcal{O}(n)$ using the power iteration method.

Note that in directed networks, $e_i = 0$ for nodes with zero in-degree $k_i^{in} = 0$, and this effect propagates, so that only nodes belonging to large strongly connected components acquire non-zero importance.

Katz status

This issue is addressed by *Katz status*, which assigns each node a small baseline importance $\beta > 0$ for free ([Katz, 1953](#)).

$$KS_i = s_i = \alpha \sum_j A_{ij} s_j + \beta_i$$

For convenience, we may set $\beta_i = 1$, whereas $\alpha < \frac{1}{\lambda}$ is required for the power iteration method to converge.

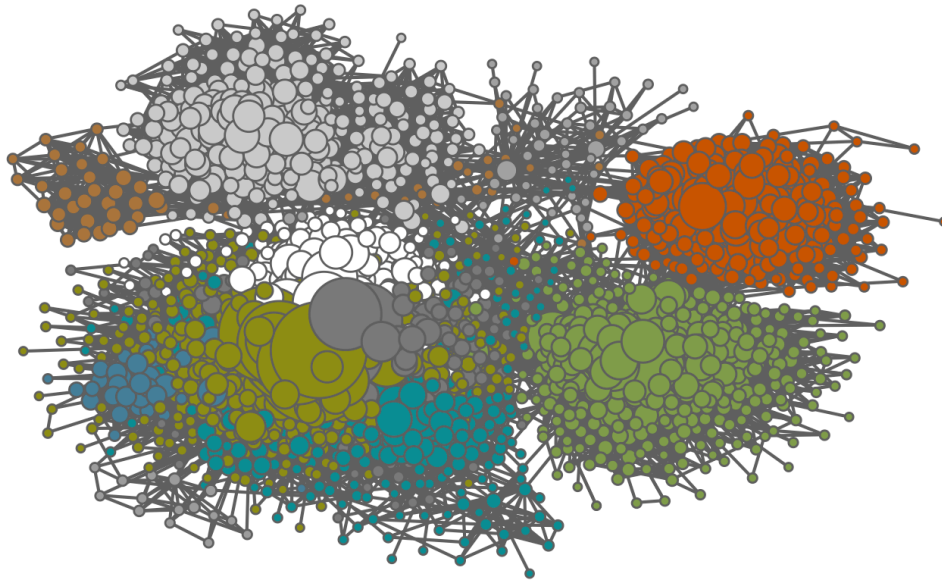
PageRank score

Finally, the well-known *PageRank score* builds upon Katz status, with nodes distributing their importance equally among their neighbors ([Brin & Page, 1998](#)).

The PageRank formula is defined as

$$PR_i = p_i = \alpha \sum_j A_{ij} \frac{p_j}{k_j} + \frac{1 - \alpha}{n},$$

where α is called the damping factor, which is traditionally set to 0.85. The PageRank algorithm runs in time $\mathcal{O}(n)$ and typically requires between 10 and 100 iterations of the power iteration method to converge.



Notice that the PageRank score can be interpreted as the probability p_i that a random walker on the graph is located at node i . At each step, the walker moves to a random neighbor with probability α , or teleports (i.e. jumps) to a randomly chosen node with probability $1 - \alpha$.

Personalized PageRank

The PageRank formulation can be adapted not only to measure node importance but also to quantify the similarity of all nodes to a designated teleport node t . In this personalized version, the random walker always teleports (i.e. restarts) to the same node t .

Random walks with restarts ([Tong et al., 2006](#)) or *personalized PageRank* is defined as

$$p_i^{(t)} = \alpha \sum_j A_{ij} \frac{p_j^{(t)}}{k_j} + (1 - \alpha) \delta_{it},$$

where δ_{it} is the Kronecker delta, ensuring that the walker restarts at node t . The method can also be extended to multiple teleport nodes, allowing personalization to a set of nodes rather than a single node.

1 IA														18 VIIIA																						
1	DC													2 EC																						
	Degree Centrality													Eigenvector centrality																						
2	BC		CC												13 IIIA		14 IVA		15 VA		16 VIA		17 VIIA													
	Betweenness Centrality		Closeness Centrality												SC		C _{COEF}		C _{COEF} ⁻¹		MNC		EC _{COEF}		PR											
															Subgraph Centrality		Clustering Coefficient		inverse C _{COEF}		max. neigh. comp.		edge clustering coefficient		PageRank											
3	RL		IC												91		14		15		16		17		18											
	RangeLimited Betweenness		Information Centrality												CC _{XX}		SC ₀		LAC		DMNC		SEC _{COEF}		LR											
															odd Subgraph Centrality		loc. avg. Connectivity		dens. max. neigh. comp.		sum of EC _{COEF}		LeaderRank													
4	BN		RC		IG		DC _{XX}		BC _{XX}		CC _{XX}		ECC _{XX}		KS _{XX}		PR _{XX}		IG _{XX}		RC _{XX}		DC _{XX}		BC _{XX}		SC _E		KL		COC _{COEF}		PEC _{COEF}		KS	
	Bottleneck Centrality		Radiality Centrality		Integration																				even Subgraph Centrality		Clique Level		coopr. weight C _{COEF}		PCC+EC _{COEF}		KatzStatus			
5	RWBC		RWCC		CC _{2,3,4}		ECC _{XX}		PR _{XX}		KS _{XX}		CCC _{COEF}		RC _{XX}		IG _{XX}		DCBC		BCCC		CCKS		KSPR		DCPR		β		SC ₂		NC		EC ₂	
	RandomWalk Betweenness		RandomWalk Closeness		2,3,4-localized-CC																						Bipartivity		2-localized-SC		Neighborhood Centrality		2-localized-EC			
6	σ		ECC		WDC		DCECC		CCECC		BCECC		KSECC		PRECC		IGECC		DCCC		BCKS		CCPR		KSIG		DCIG		DCC _{COEF}		SC ₃		LI		EC ₃	
	stress Centrality		Eccentricity		Weighted Degree																						2-localized-SC		Lobby Index		3-localized-EC					
7	BC _{2,3,4}		ECC ⁻¹		SDC		DCRC		CCRC		BCRC		KSRC		PRRC		IGRC		DCKS		BCPR		CCIG		DCPR		BCIG		ECCRC		BCC _{COEF}		SC ₄		EC ₄	
	2,3,4-localized-BC		Inverse Eccentricity		Sphere Degree Centrality																						4-localized-SC		4-localized-EC							

z	mass
C	Hybrid
Name	

22	2	23	2	24	2	25	1	26	1	27	1	117	1
FC	FD	US	DIS	ASS	DAM	UC							
Functional Centrality	Functional Diversity	UniScore	Pairwise Dis-connectivity	Assortative Mixing	Damage	United compl. Centrality							

28	1	29	1	30	1	31	1	33	1	116	0	118	1
EI	CM	NαC	MC	HGI	HYP	HC							
Essentiality Index	Complexity Measure	Normalized α Centrality	Moduland Centrality	Harary Graph Information	Hyperbolic Index	Harmonic Centrality							

- Betweenness-based
- Distance-based
- Linear Combinations
- Subgraph-based
- Clustering Coefficient-based
- Edge Clustering Coefficient-based
- Spectral-based
- Miscellaneous

©David Schoch (University of Konstanz)