Centrality Measures

Leonid E. Zhukov

School of Data Analysis and Artificial Intelligence
Department of Computer Science
National Research University Higher School of Economics

Structural Analysis and Visualization of Networks

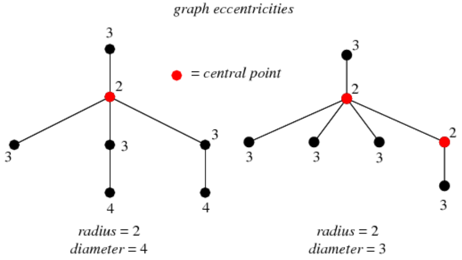
NATIONAL RESEARCH
UNIVERSITY

Which vertices are important?

M.Grandjean, 2014

Graph-theoretic measures

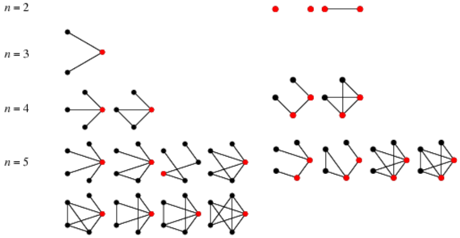
- The **eccentricity** $\epsilon(v)$ of a vertex v is the maximum distance between v and any other vertex u of the graph $\epsilon(v) = \max_{u \in V} d(u, v)$
- Graph **diameter** is the maximum eccentricity $d = \max_{v \in V} \epsilon(v)$
- Graph **radius** is the minimum eccentricity $r = \min_{v \in V} \epsilon(v)$.
- A point v is a **central point** of a graph if the eccentricity of the point equals the graph radius $\epsilon(v) = r$



from Eric Weisstein *MathWorld*

Graph-theoretic measures

- Graph **center** is a set of of vertices with graph eccentricity equal to the graph radius $\epsilon(v) = r$ - set of central points
- Graph **periphery** is a set of vertices that have graph eccentricities equal to the graph diameter $\epsilon(v) = d$



graphs with 1 center node *graphs with 2 center nodes*

from Eric Weisstein *MathWorld*

Centrality Measures

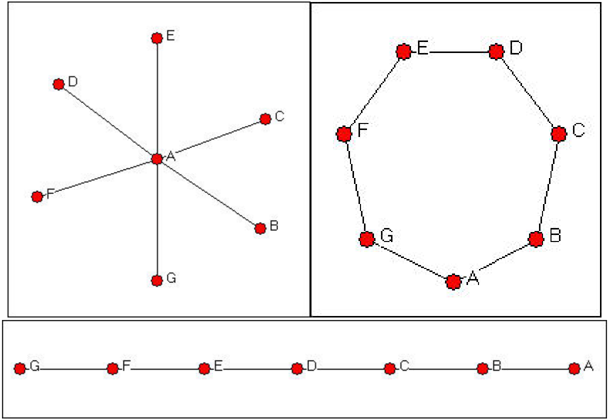
Sociology.
Most "important"actors: actor location in the social network

- Actor centrality - involvement with other actors, many ties, source or recipient. Undirected network.
- Actor prestige - recipient (object) of many ties, ties directed to an actor. Directed network.

In this lecture: undirected graphs, symmetric matrix $A_{ij} = A_{ji}$, $\mathbf{A} = \mathbf{A^T}$

Linton Freeman, 1979

Three graphs



Star graph Circle graph Line Graph

Degree centrality

Degree centrality: number of nearest neighbours

$$C_D(i) = k(i) = \sum_j A_{ij} = \sum_j A_{ji}$$

Normalized degree centrality

$$C_D^*(i) = \frac{1}{n-1} C_D(i)$$

High centrality degree -direct contact with many other actors
Low degree - not active, peripheral actor

Closeness centrality

Closeness centrality: how close an actor to all the other actors in network

$$C_C(i) = \frac{1}{\sum_j d(i,j)}$$

Normalized closeness centrality

$$C_C^*(i) = (n-1)C_C(i)$$

Actor in the center can quickly interact with all others, short communication path to others, minimal number of steps to reach others

[*** Harmonic centrality $= \sum_j \frac{1}{d(i,j)}$ ***]

Alex Bavelas, 1948

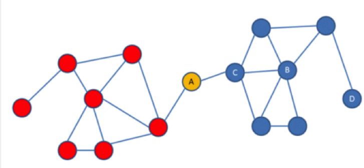
Betweenness centrality

Betweenness centrality: number of shortest paths going through the actor $\sigma_{st}(i)$

$$C_B(i) = \sum_{s \neq t \neq i} \frac{\sigma_{st}(i)}{\sigma_{st}}$$

Normalized betweenness centrality

$$C_B^*(i) = \frac{2}{(n-1)(n-2)} C_B(i)$$



Probability that a communication from s to t will go through i (geodesics)

Linton Freeman, 1977

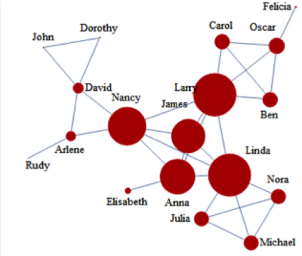
Eigenvector centrality

Importance of a node depends on the importance of its neighbors (recursive definition)

$$v_i \leftarrow \sum_j A_{ij} v_j$$

$$v_i = \frac{1}{\lambda} \sum_j A_{ij} v_j$$

$$\mathbf{A}\mathbf{v} = \lambda \mathbf{v}$$

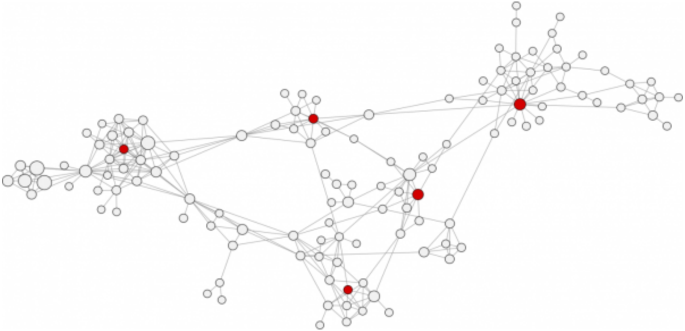


Select an eigenvector associated with largest eigenvalue $\lambda = \lambda_1$, $\mathbf{v} = \mathbf{v}_1$

Phillip Bonacich, 1972.

Centrality examples

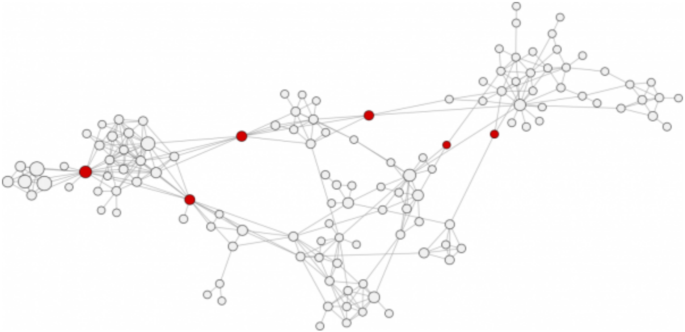
Closeness centrality



from www.activenetworks.net

Centrality examples

Betweenness centrality



from www.activenetworks.net

Centrality examples

Eigenvector centrality

from www.activenetworks.net

Katz status index

Weighted count of all paths coming to the node: the weight of path of length n is counted with attenuation factor β^n, $\beta < \frac{1}{\lambda_1}$

$$k_i = \beta \sum_j A_{ij} + \beta^2 \sum_j A_{ij}^2 + \beta^3 \sum_j A_{ij}^3 + \ldots$$

$$\mathbf{k} = (\beta \mathbf{A} + \beta^2 \mathbf{A}^2 + \beta^3 \mathbf{A}^3 + \ldots)\mathbf{e} = \sum_{n=1}^{\infty} (\beta^n \mathbf{A}^n)\mathbf{e} = (\sum_{n=0}^{\infty} (\beta \mathbf{A})^n - \mathbf{I})\mathbf{e}$$

$$\sum_{n=0}^{\infty} (\beta \mathbf{A})^n = (\mathbf{I} - \beta \mathbf{A})^{-1}$$

$$\mathbf{k} = ((\mathbf{I} - \beta \mathbf{A})^{-1} - \mathbf{I})\mathbf{e}$$

$$(\mathbf{I} - \beta \mathbf{A})\mathbf{k} = \beta \mathbf{A}\mathbf{e}$$

$$\mathbf{k} = \beta (\mathbf{I} - \beta \mathbf{A})^{-1} \mathbf{A}\mathbf{e}$$

Leo Katz, 1953

Bonacich centrality

Two-parametric centrality measure $c(\alpha, \beta)$
β - degree to which an individual status is a function of the statuses of those to whom he is connected (can be positive if connected to powerful and negative, if connected to powerless)
α - normalization parameter

$$c_i(\alpha, \beta) = \sum_j (\alpha + \beta c_j) A_{ij}$$

$$\mathbf{c} = \alpha \mathbf{A} \mathbf{e} + \beta \mathbf{A} \mathbf{c}$$

$$(\mathbf{I} - \beta \mathbf{A}) \mathbf{c} = \alpha \mathbf{A} \mathbf{e}$$

$$\mathbf{c} = \alpha (\mathbf{I} - \beta \mathbf{A})^{-1} \mathbf{A} \mathbf{e}$$

α - found from normalization $||\mathbf{c}||_2 = \sum c_i^2 = 1$

Phillip Bonacich, 1987

Modified versions

- Katz centrality (Newman):

$$x_i = \alpha \sum_j A_{ij} x_j + \beta_i$$

$$\mathbf{x} = \alpha \mathbf{A} \mathbf{x} + \boldsymbol{\beta}$$

$$\mathbf{x} = (\mathbf{I} - \alpha \mathbf{A})^{-1} \boldsymbol{\beta}$$

- Alpha-centrality (Bonacich):

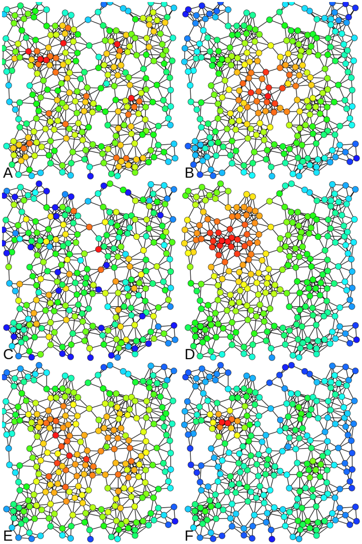
$$x_i = \alpha \sum_j A_{ij} x_j + 1$$

$$\mathbf{x} = \alpha \mathbf{A} \mathbf{x} + \mathbf{e}$$

$$\mathbf{x} = (\mathbf{I} - \alpha \mathbf{A})^{-1} \mathbf{e}$$

Bonacich, 2001,Newman, 2010

Centrality examples



from Claudio Rocchini

- A) Degree centrality
- B) Closeness centrality
- C) Betweenness centrality
- D) Eigenvector centrality
- E) Katz centrality
- F) Alpha centrality

Centralization

Centralization (network measure) - how central the most central node in the network in relation to all other nodes.

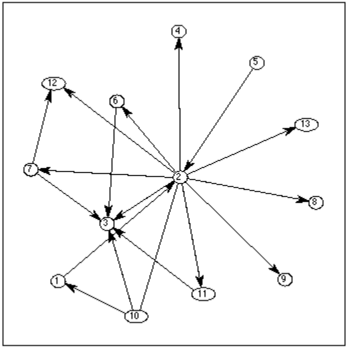
$$C_x = \frac{\sum_i^N [C_x(p_*) - C_x(p_i)]}{\max \sum_i^N [C_x(p_*) - C_x(p_i)]}$$

C_x - one of the centrality measures
p_* - node with the largest centrality value
max - is taken over all graphs with the same number of nodes (for degree, closeness and betweenness the most centralized structure is the star graph)

Linton Freeman, 1979

Prestige

Prestige - measure of node importance in directed graphs



- Degree prestige $k_{in}(i)$
- Proximity prestige (closeness)
- Status or Rank prestige (Katz, Bonacich)

Metrics comparison

- **Pearson correlation** coefficient

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

  Shows linear dependence between variables, $-1 \leq r \leq 1$
  (perfect when related by linear function)

- **Spearman rank correlation** coefficient (Sperman's rho):
  Convert raw scores to ranks - sort by score: $X_i \rightarrow x_i$, $Y_i \rightarrow y_i$

$$\rho = 1 - \frac{6\sum_{i=1}^{n}(x_i - y_i)^2}{n(n^2 - 1)}$$

  Shows strength of monotonic association
  (perfect for monotone increasing/decreasing relationship)

Ranking comparison

- The Kendall tau rank distance is a metric that counts the number of pairwise disagreements between two ranking lists
- Kendall rank correlation coefficient, commonly referred to as Kendall's tau coefficient

$$\tau = \frac{n_c - n_d}{n(n-1)/2}$$

  n_c - number of concordant pairs, n_d - number of discordant pairs
- $-1 \le \tau \le 1$, perfect agreement $\tau = 1$, reversed $\tau = -1$
- Example

Rank 1	A	B	C	D	E
Rank 2	C	D	A	B	E

$$\tau = \frac{6-4}{5(5-1)/2} = 0.2$$

References

- Centrality in Social Networks. Conceptual Clarification, Linton C. Freeman, Social Networks, 1, 215-239, 1979
- Power and Centrality: A Family of Measures, Phillip Bonacich, The American Journal of Sociology, Vol. 92, No. 5, 1170-1182, 1987
- A new status index derived from sociometric analysis, L. Katz, Psychometrika, 19, 39-43, 1953.
- Eigenvector-like measures of centrality for asymmetric relations, Phillip Bonacich, Paulette Lloyd, Social Networks 23, 191?201, 2001