CIFF
Centro Internacional
de Formación Financiera

Social Network Analysis

Lecture 2: Centrality in Networks

David Martín-Corral Calvo
dmartincc@smartvel.com

Enero / Marzo

MASTER EN BA & BD

Recap:

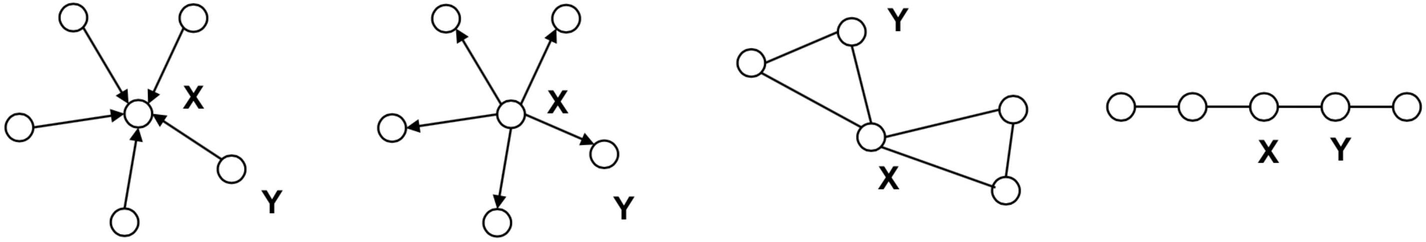
- Networks can be represented as matrices
- Useful metrics:
  - Degree
  - Connected components

Outline:

- Centrality in networks
  - Betweenness
  - Closeness
  - Eigenvector centrality
  - PageRank

Notions of centrality

In each of the following networks, X has higher centrality than Y according to a particular measure



indegree outdegree betweenness closeness

Betweenness

$$C_B(i) = \sum_{j<k} g_{jk}(i) / g_{jk}$$

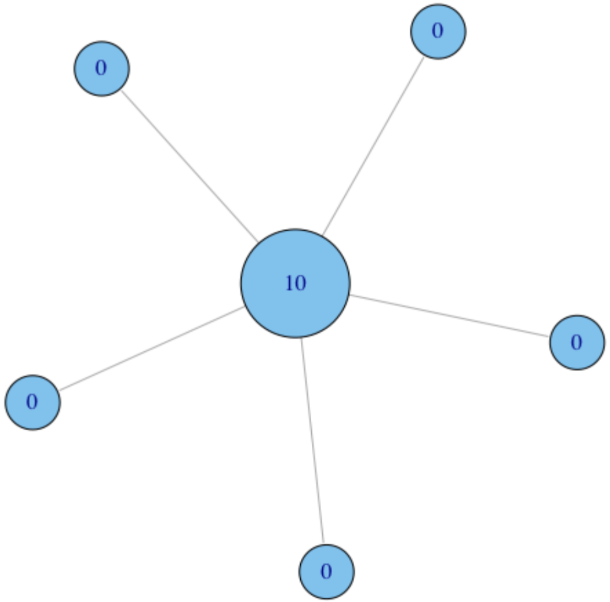
Where g_{jk} = the number of shortest paths connecting *jk*
$g_{jk}(i)$ = the number that actor *i* is on.

Usually normalized by:

$$C'_B(i) = C_B(i) / \boxed{[(n-1)(n-2)/2]}$$

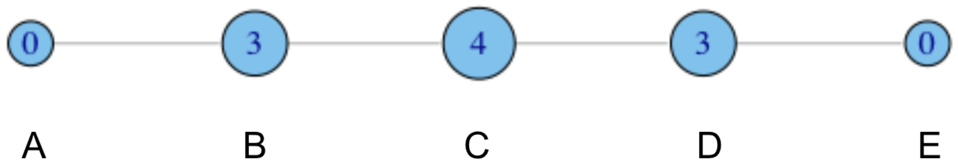
number of pairs of vertices
excluding the vertex itself

Betweenness: example



Betweeness: example

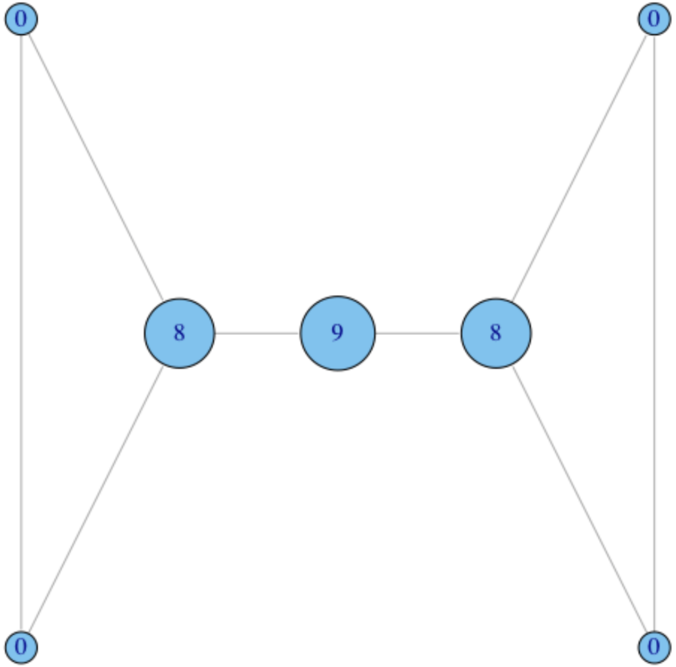
- non-normalized version:



- A lies between no two other vertices
- B lies between A and 3 other vertices: C, D, and E
- C lies between 4 pairs of vertices (A,D),(A,E),(B,D),(B,E)

- note that there are no alternate paths for these pairs to take, so C gets full credit

Closeness: example



Closeness

Closeness is based on the length of the average shortest path between a node and all other nodes in the network

Closeness Centrality:

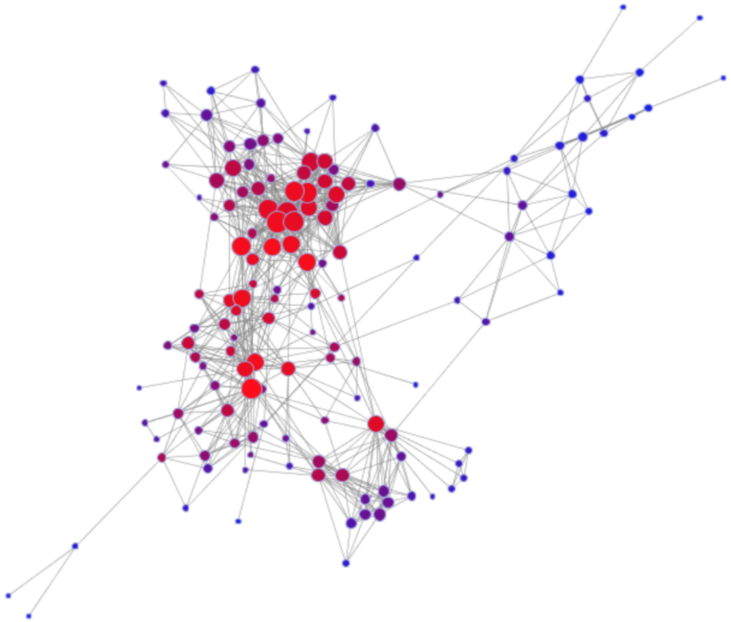
$$C_c(i) = \left[\sum_{j=1}^{N} d(i,j) \right]^{-1}$$

Normalized Closeness Centrality

$$C_C^{'}(i) = (C_C(i))/(N-1)$$

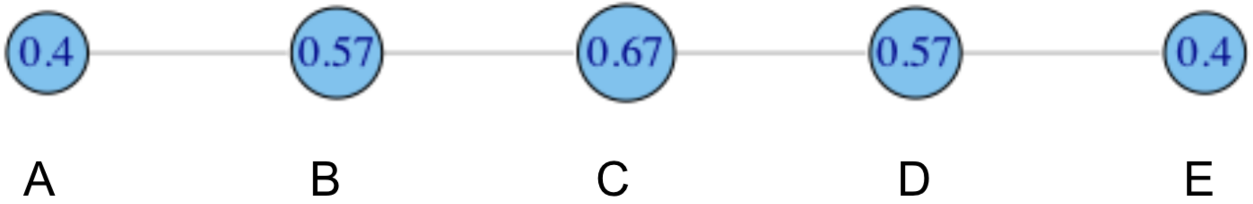
Centro Internacional de Formación Financiera

Closeness



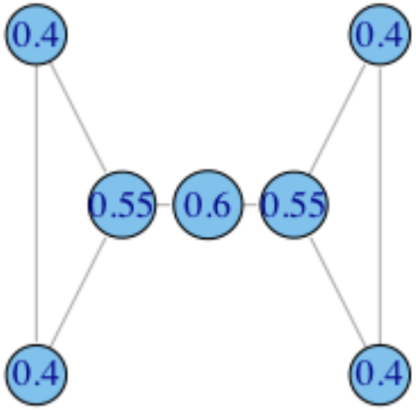
Degree is the size
Color is the closeness

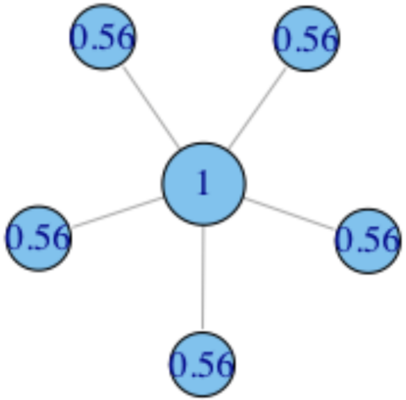
Closeness: example



$$C_c'(A) = \left[\frac{\sum_{j=1}^{N} d(A,j)}{N-1} \right]^{-1} = \left[\frac{1+2+3+4}{4} \right]^{-1} = \left[\frac{10}{4} \right]^{-1} = 0.4$$

Closeness: example





Eigenvector centrality

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

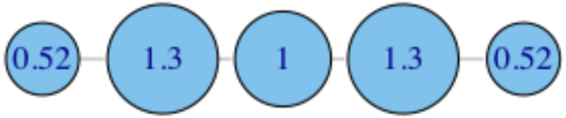
$$c(\beta) = \alpha (I - \beta A)^{-1} A \mathbf{1}$$

- α is a normalization constant
- β determines how important the centrality of your neighbors is
- $\mathbf{A}$ is the adjacency matrix (can be weighted)
- $\mathbf{I}$ is the identity matrix (1s down the diagonal, 0 off-diagonal)
- $\mathbf{1}$ is a matrix of all ones.

Eigenvector centrality: example

β=.25

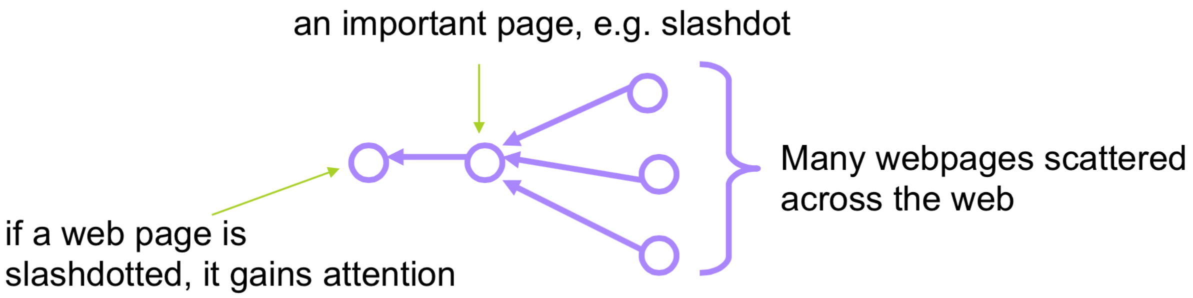
β=-.25



Why does the middle node have lower centrality than its neighbors when β is negative?

Page Rank

- PageRank brings order to the Web:
  - it's not just the pages that point to you, but how many pages point to those pages, etc.
  - more difficult to artificially inflate centrality with a recursive definition



Recap:

Many measures:

degree, betweenness, closeness, eigenvector

In indirected networks:

indegree, outdegree, page rank

Activity: Facebook network analysis

We are going to analyse facebook's user networks (with Gephi), Les Miserables and Football Transfers.

1. Layout Force Atlas 2, set gravity under 50
2. Select Giant Component
3. Calculate Degree
4. Calculate Network diameter
5. Calculate Page Rank
6. Eigenvector centrality

Important Links:

Strong and Weak Ties (Chapter 1 and 2)

The PageRank Citation Ranking: Bringing Order to the Web

Suggested readings:

Tweet the Debates

CIFF

Centro Internacional
de Formación Financiera