Network Analysis (contd ..)

Centrality

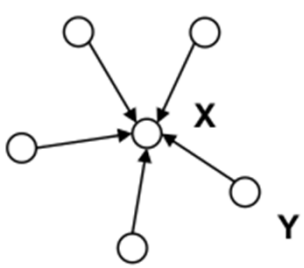
Centrality

- A common goal in SNA (social network analysis) is to identify the "central" nodes of a network.
- What does "central" mean ?

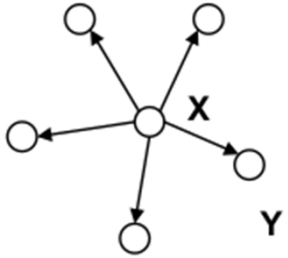
- active?
- important?
- non-redundant?

Definition of 'central' varies by context/purpose.

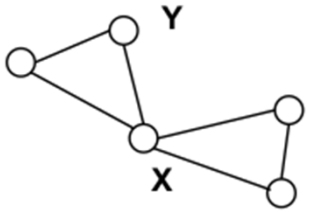
Centrality: Who's Important Based On Their Network Position



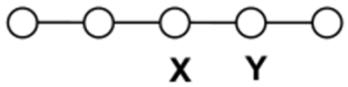
indegree



outdegree



betweenness



closeness

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

Common centrality measures

- We will define and compare three centrality measures:

- **Degree centrality (based on degree)**
- **Betweeness centrality (based on geodesics)**
- **Closeness centrality (based on average distances)**

Degree Centrality

Idea: A central actor is one with many connections.

- Undirected degree centrality (number of edges connecting to the node)
- Indegree centrality (number of edges coming into the node)
- Outdegree centrality (number of edges coming out of the node)

Degree Centrality is a best measure when number of connections is important

Indegree and Outdegree

Which countries have high indegree?

- USA
- Japan
- UK

Indicates more imports to above countries

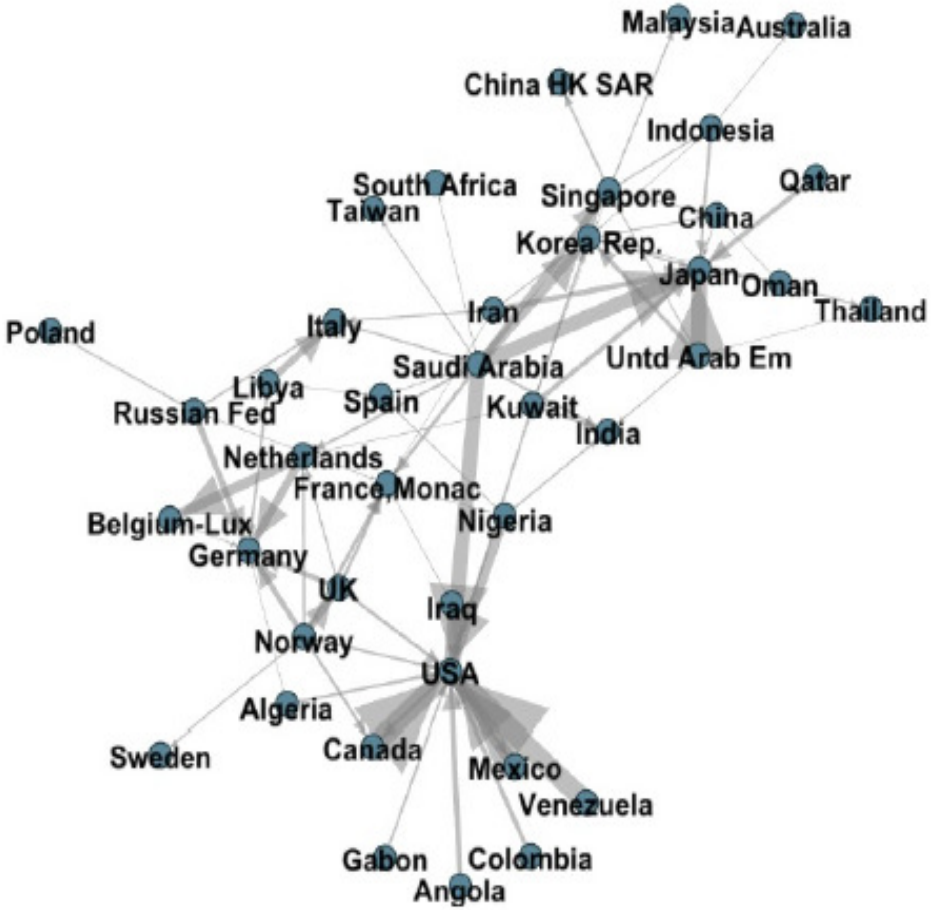
Which country has high outdegree

- Saudi Arabia

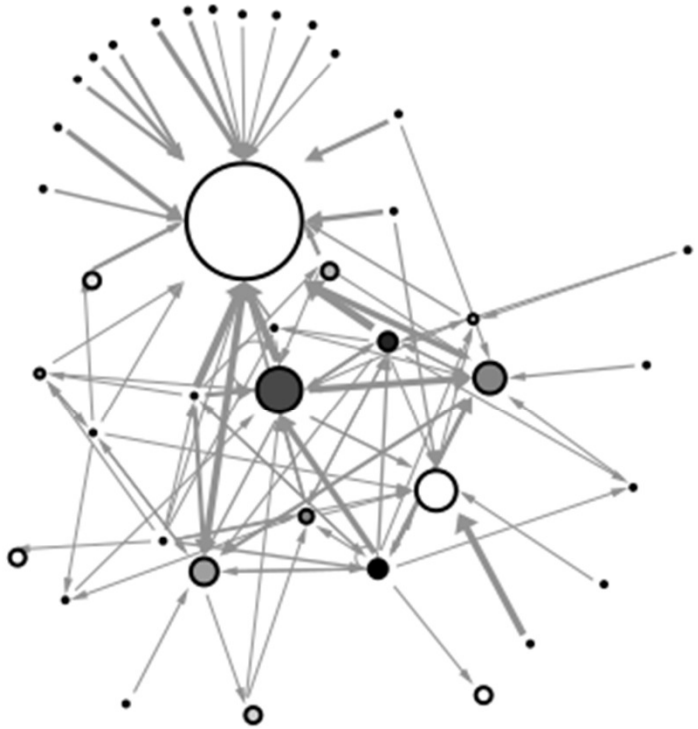
Indicates more exports from Saudi Arabia

Trade in petroleum 1998

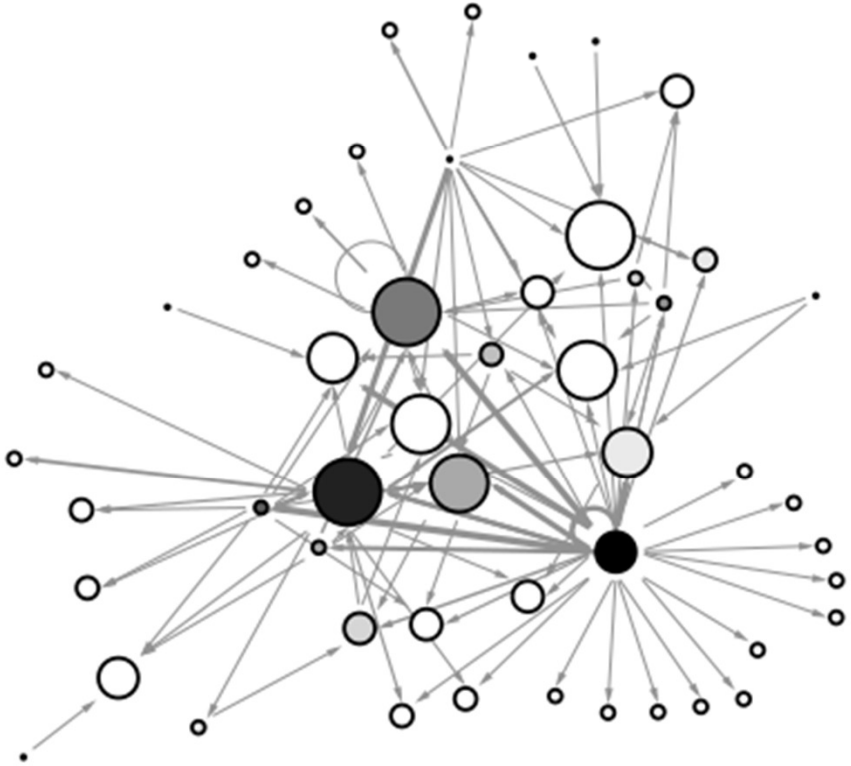
Source :NBER---United Nations Trade Data



example financial trading networks



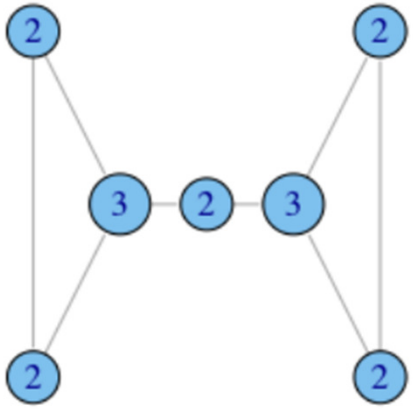
high centralization: one node trading with many others



low centralization: trades are more evenly distributed

In What Contexts May Degree Be Insufficient To Describe Centrality?

- Ability to broker between group
- Likelihood that information originating anywhere in the network reaches you…



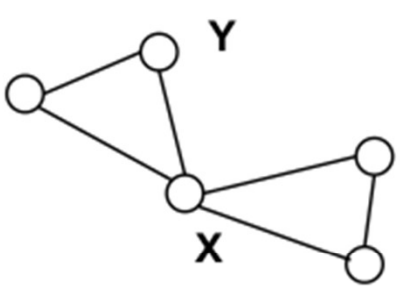
Betweenness Centrality

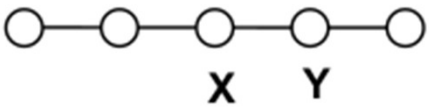
- Betweenness centrality quantifies the number of times a node acts as a bridge along path between two other nodes.
- Betweenness Centrality =
-

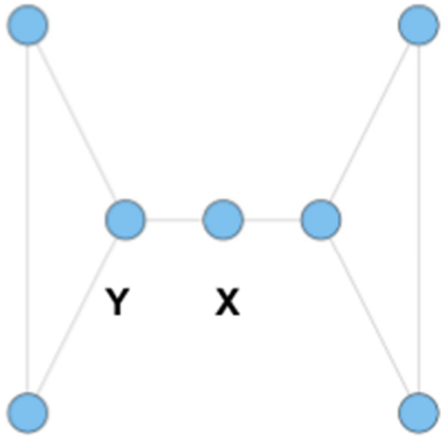
$$C_B(i) = \sum_{j<k} g_{jk}(i) / g_{jk}$$

- Where g_{jk} = the number of geodesics connecting j and k, and $g_{jk}(i)$= the number of geodesics that node i is on.

X has higher betweenness in the following graphs







Closeness is based on the length of the average shortest path between a vertex and all vertices in the graph

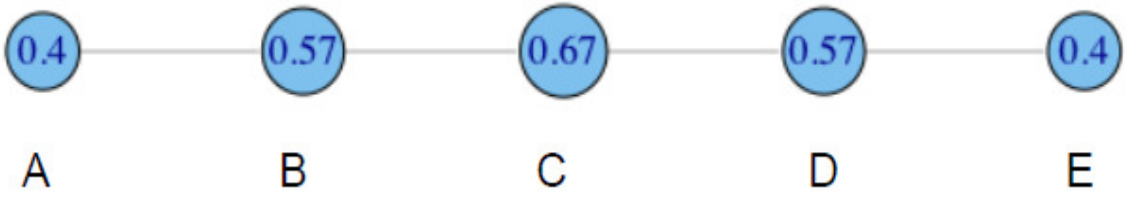
Closeness Centrality

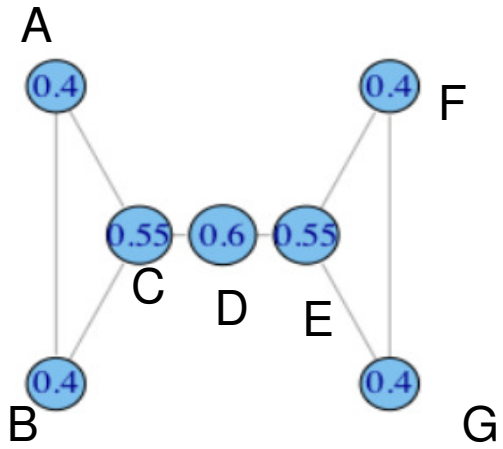
Where d(i,j) = shortest distance from i to j

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

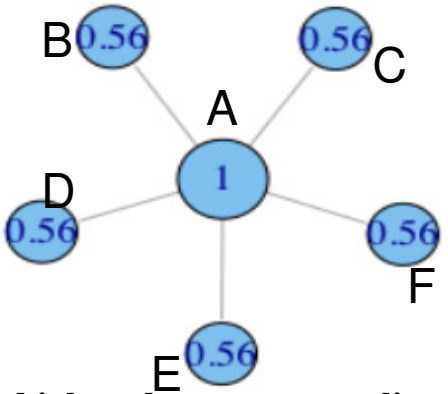
Closeness examples

C has higher closeness centrality and is closer to all nodes compared to others





D has higher closeness centrality and is closer to all nodes compared to others



A has higher closeness centrality and is closer to all nodes compared to others

❖Eigenvector'Centrality

Other centrality measures

❖Katz'Centrality‘

http://cs.brynmawr.edu/Courses/cs380/spring2013/section02/slides/05_Centrality.pdf

http://www2.unb.ca/~ddu/6634/Lecture_notes/Lecture_4_centrality_measure.pdf

References

http://cs224w.stanford.edu