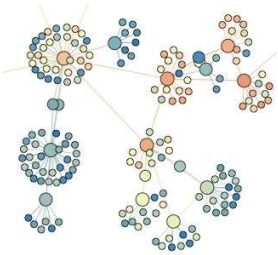
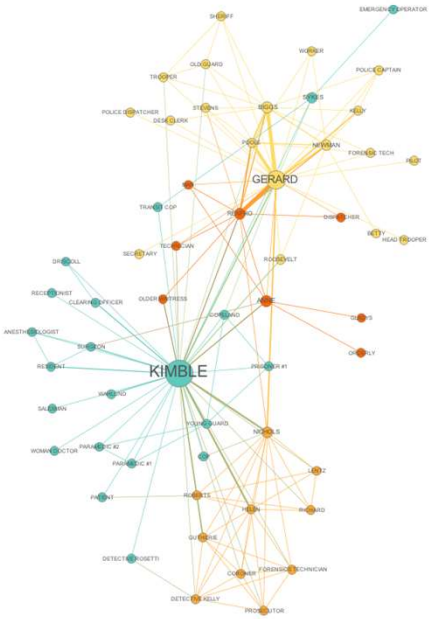


SNA 3A: Centrality

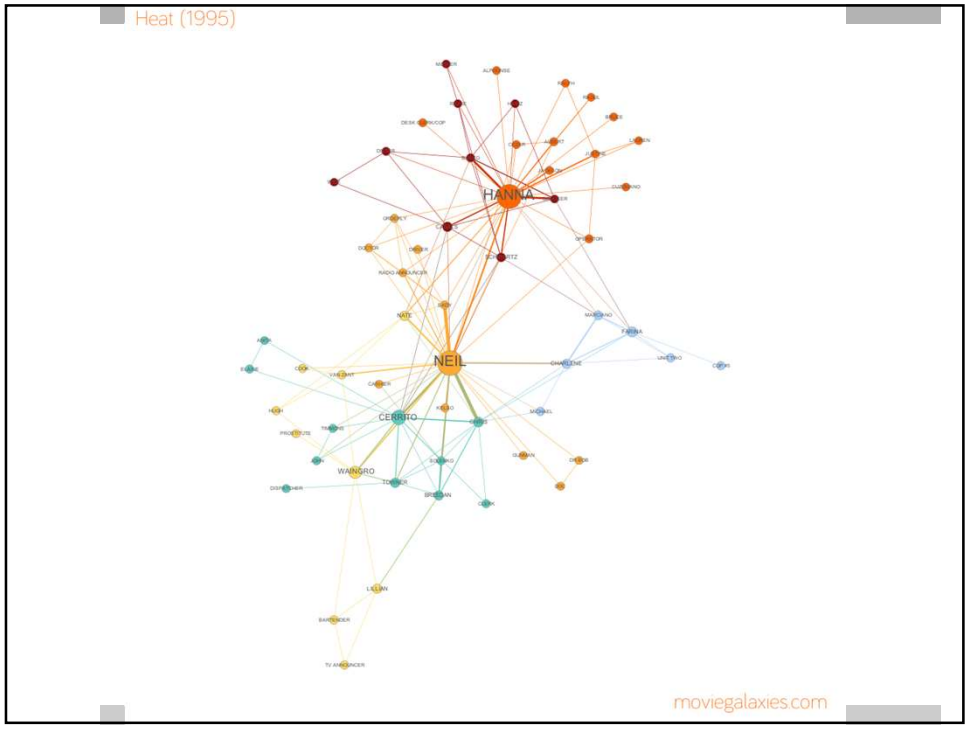
Lada Adamic

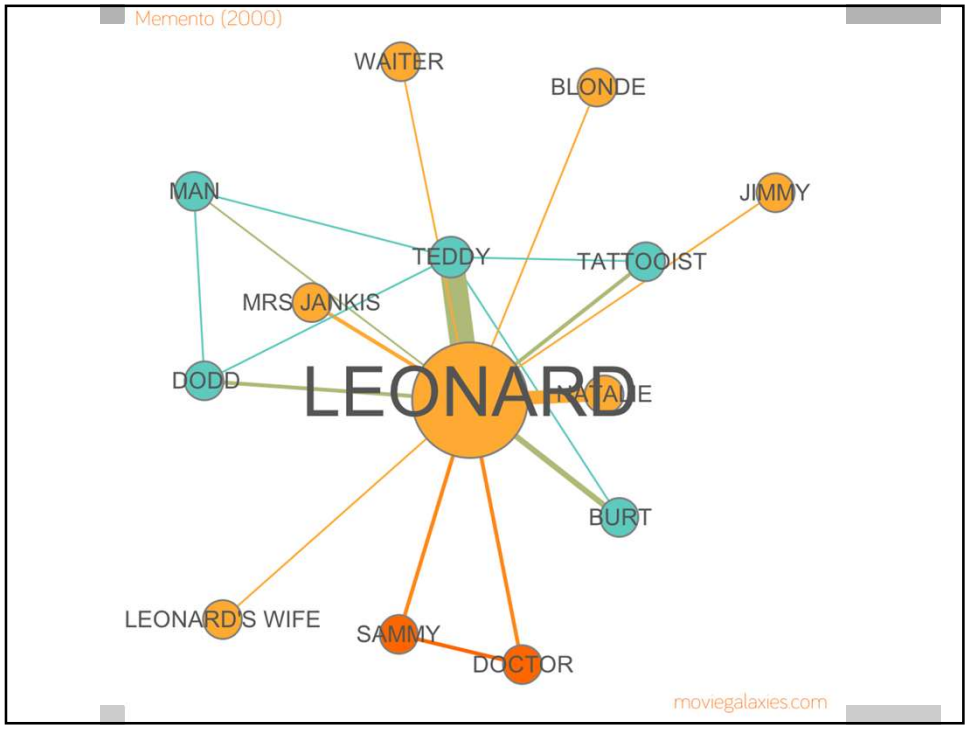
school of information

The Fugitive (1993)

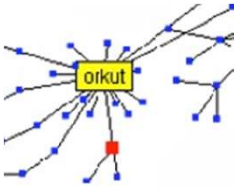


moviegalaxies.com

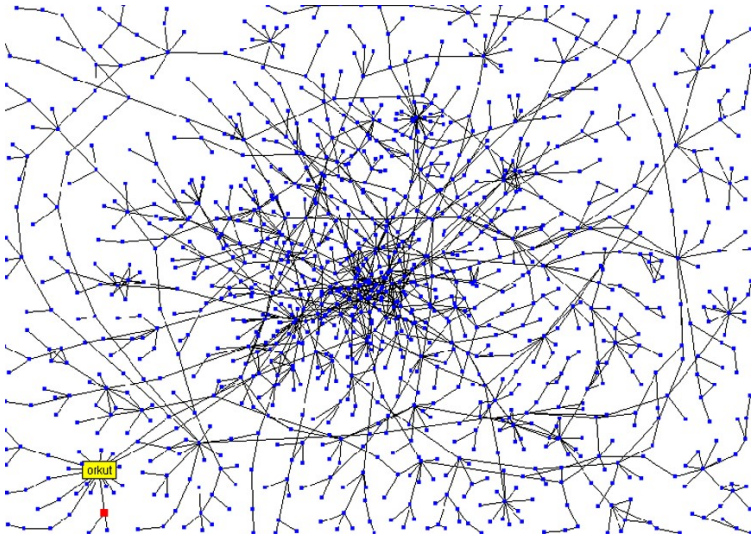

Heat (1995)
HANNA
NEIL
CERRITO
WAINGRO
moviegalaxies.com


Memento (2000)
WAITER
BLONDE
MAN
JIMMY
TEDDY
TATTOOIST
MRS JANKIS
DODD
LEONARD
NATALIE
BURT
LEONARD'S WIFE
SAMMY
DOCTOR
moviegalaxies.com

is counting the edges enough?



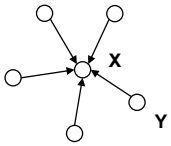
Stanford Social Web (ca. 1999)



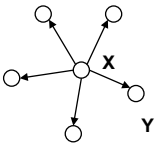
network of personal homepages at Stanford

different notions of centrality

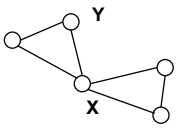
In each of the following networks, X has higher centrality than Y according to a particular measure



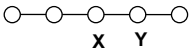
indegree



outdegree

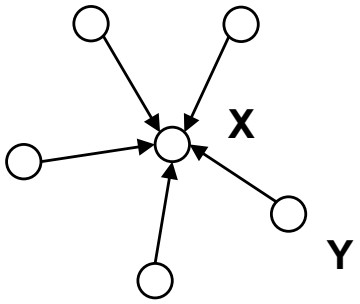


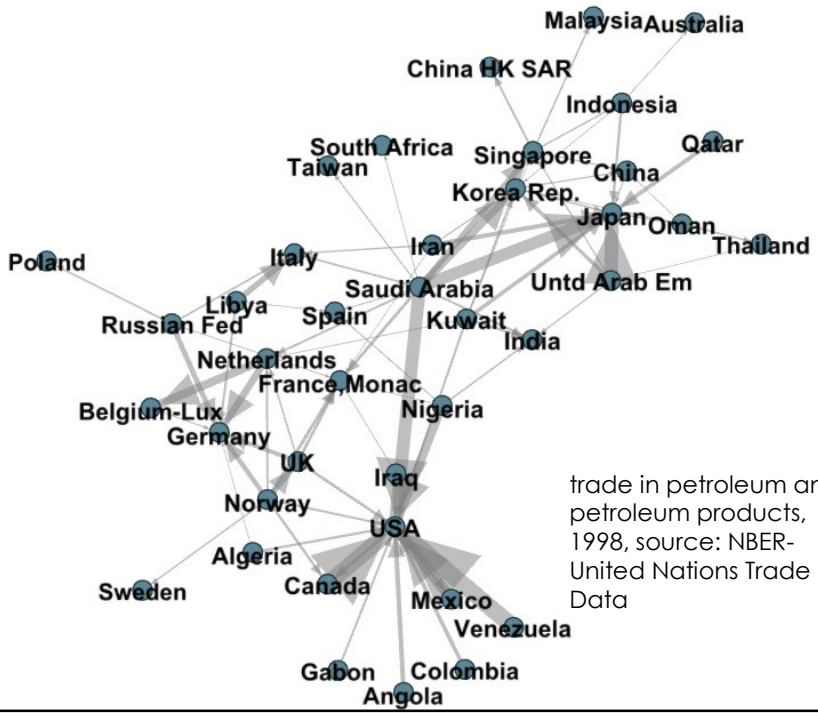
betweenness



closeness

review: indegree



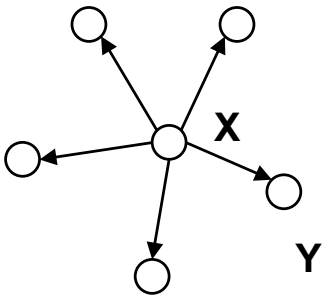


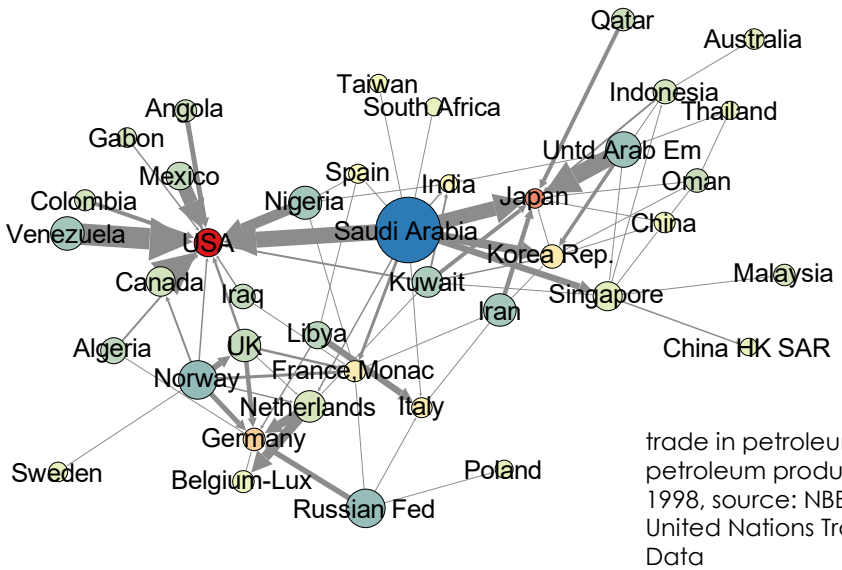
trade in petroleum and petroleum products, 1998, source: NBER-United Nations Trade Data

Quiz Q:

- Which countries have high indegree (import petroleum and petroleum products from many others)
  - Saudi Arabia
  - Japan
  - Iraq
  - USA
  - Venezuela

review: outdegree

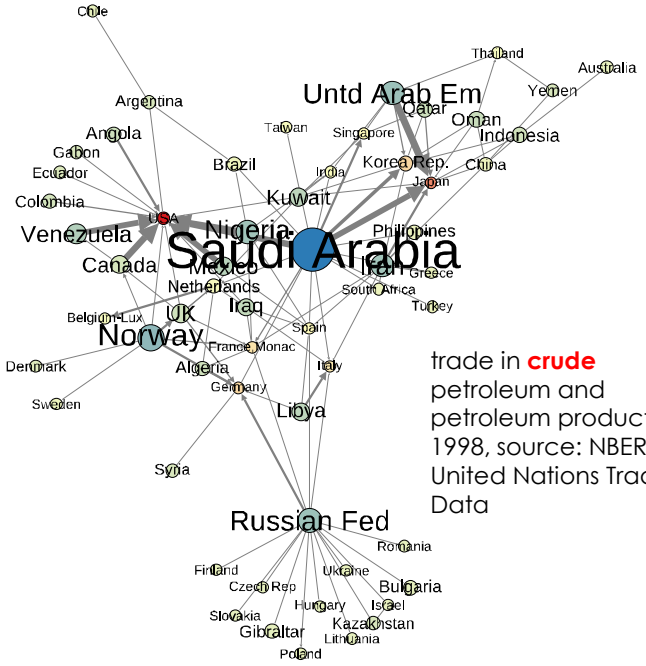




trade in petroleum and petroleum products, 1998, source: NBER-United Nations Trade Data

Quiz Q:

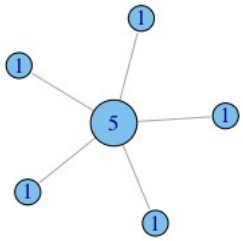
- Which country has low outdegree but exports a significant quantity (thickness of the edges represents $$ value of export) of petroleum products
  - Saudi Arabia
  - Japan
  - Iraq
  - USA
  - Venezuela



trade in **crude** petroleum and petroleum products, 1998, source: NBER-United Nations Trade Data

putting numbers to it

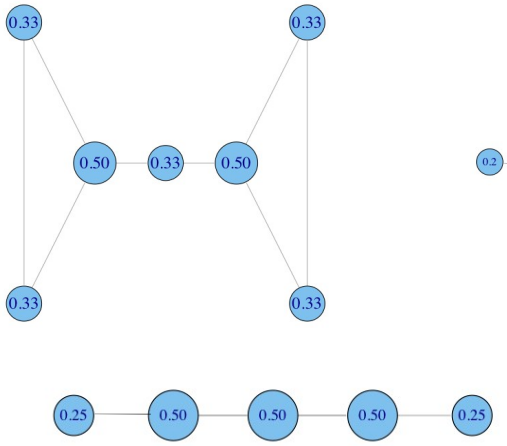
Undirected degree, e.g. nodes with more friends are more central.



Assumption: the connections that your friend has don't matter, it is what they can do directly that does (e.g. go have a beer with you, help you build a deck...)

normalization

divide degree by the max. possible, i.e. (N-1)



centralization: skew in distribution

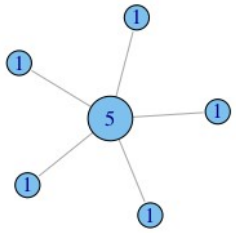
How much variation is there in the centrality scores among the nodes?

Freeman's general formula for centralization (can use other metrics, e.g. gini coefficient or standard deviation):

maximum value in the network

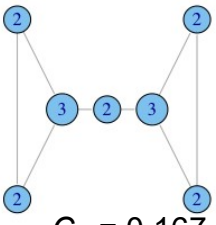
$$C_D = \frac{\sum_{i=1}^{g} [C_D(n^*) - C_D(i)]}{[(N-1)(N-2)]}$$

degree centralization examples



C_D = 1.0

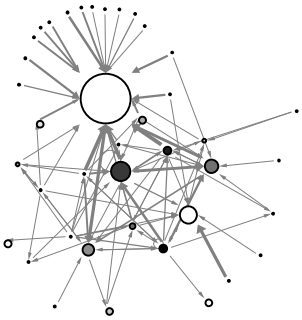
C_D = 0.167



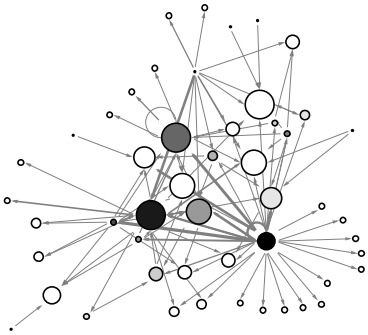
C_D = 0.167

real-world examples

example financial trading networks



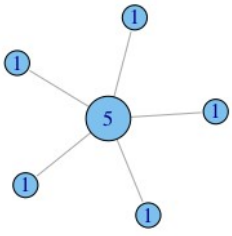
high in-centralization:
one node buying from
many others

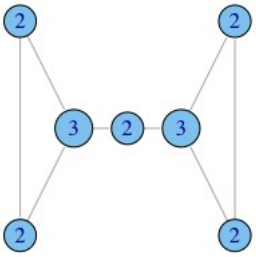


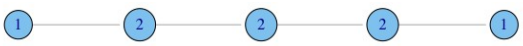
low in-centralization:
buying is more evenly
distributed

what does degree not capture?

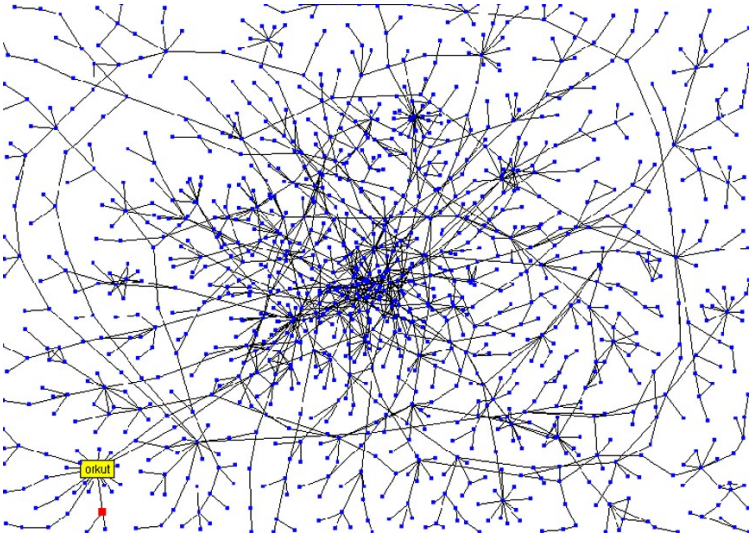
In what ways does degree fail to capture centrality in the following graphs?





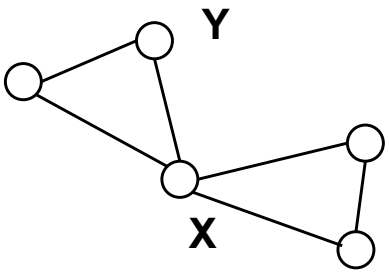


Stanford Social Web (ca. 1999)

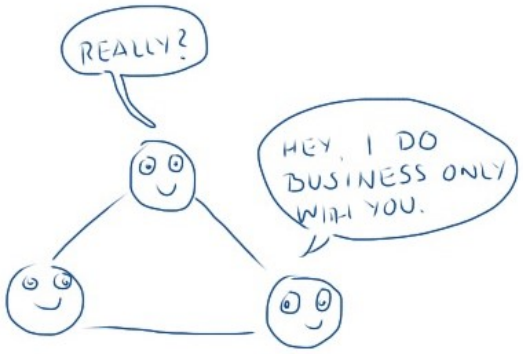


network of personal homepages at Stanford

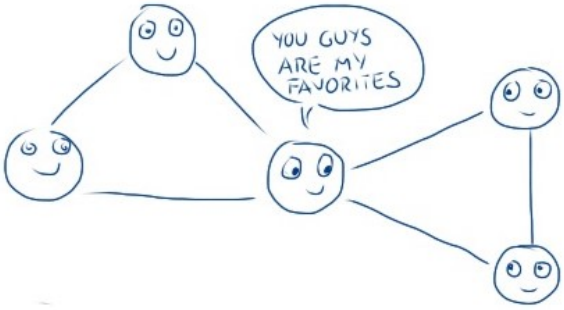
Brokerage not captured by degree



constraint

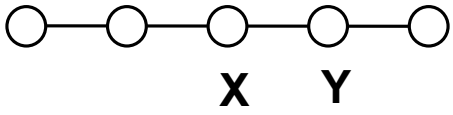


constraint



betweenness: capturing brokerage

- intuition: how many pairs of individuals would have to go through you in order to reach one another in the minimum number of hops?



betweenness: definition

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

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

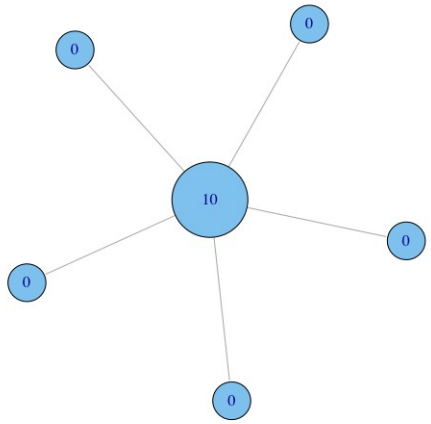
Usually normalized by:

$$C_B'(i) = C_B(i) / [(n-1)(n-2)/2]$$

number of pairs of vertices excluding the vertex itself

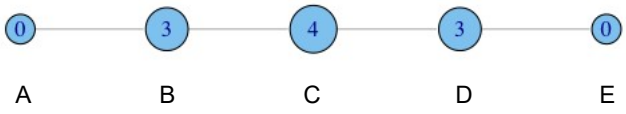
betweenness on toy networks

- non-normalized version:



betweenness on toy networks

- non-normalized version:

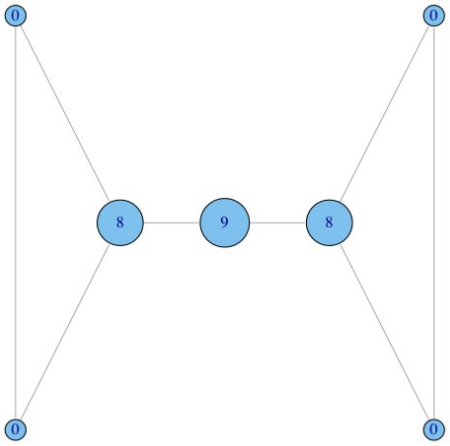


- A lies between no two other vertices
- B lies between A and 3 other vertices: C, D, and E
- C lies between 4 pairs of vertices (A,D),(A,E),(B,D),(B,E)

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

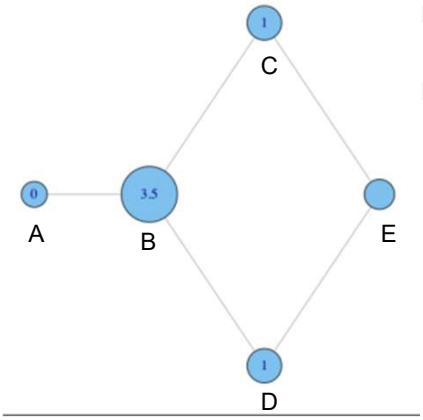
betweenness on toy networks

- non-normalized version:



betweenness on toy networks

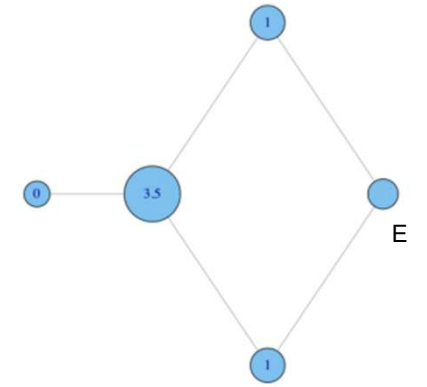
- non-normalized version:



- why do C and D each have betweenness 1?
- They are both on shortest paths for pairs (A,E), and (B,E), and so must share credit:
  - ½+½ = 1

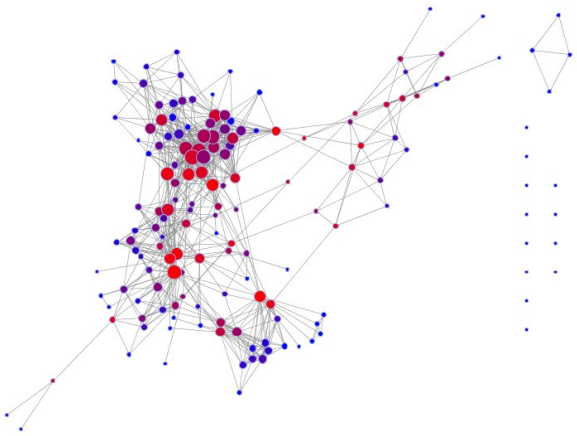
Quiz Question

- What is the betweenness of node E?



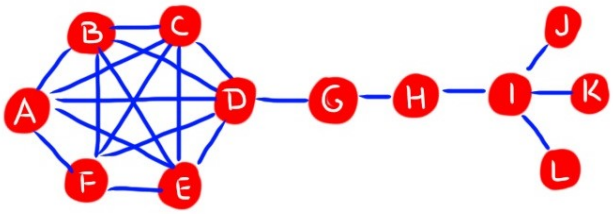
betweenness: example

Lada's old Facebook network: nodes are sized by degree, and colored by betweenness.



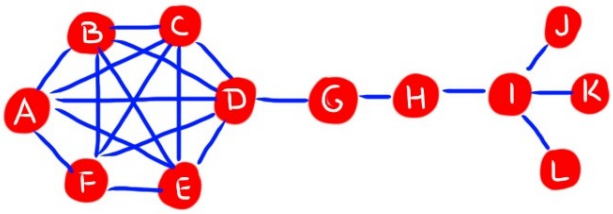
Quiz Q:

- Find a node that has high betweenness but low degree



Quiz Q:

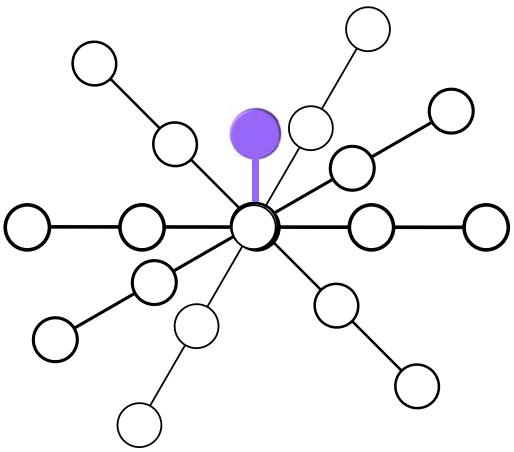
- Find a node that has low betweenness but high degree



closeness

- What if it's not so important to have many direct friends?
- Or be "between" others
- But one still wants to be in the "middle" of things, not too far from the center

need not be in a brokerage position



closeness: definition

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

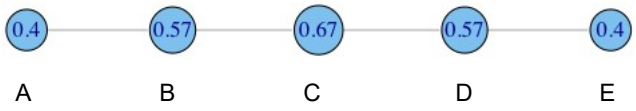
Closeness Centrality:

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

Normalized Closeness Centrality

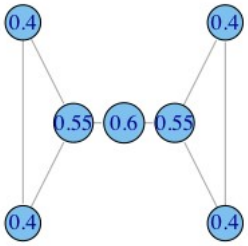
$$C'_C(i) = (C_C(i))/(N-1)$$

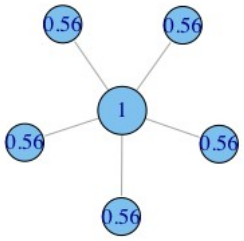
closeness: toy example



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

closeness: more toy examples





Quiz Q:

Which node has relatively high degree but low closeness?