# Network Centrality
# Part 1

Saptarshi Ghosh
Department of CSE, IIT Kharagpur
Social Computing course, CS60017

Includes material borrowed from various online sources, including slides by Lada Adamic

# Node centrality

- Relative importance of a node in a network

- Importance varies according to application
  - How influential a person is within a social network
  - How important a webpage is in the Web
  - Which persons to vaccinate when a disease is spreading

- There is an analogous concept of edge centrality, but we will focus on node centrality

# Node centrality measures

- Many proposed centrality measures
  - Network structure based
  - Activity based (e.g., number of times a user is mentioned on Twitter or Facebook)
  - Temporal (e.g., Test-of-Time awards to research papers)
  - Hybrid
  - … and more

- We will focus on the first two types of measures

# Degree centrality

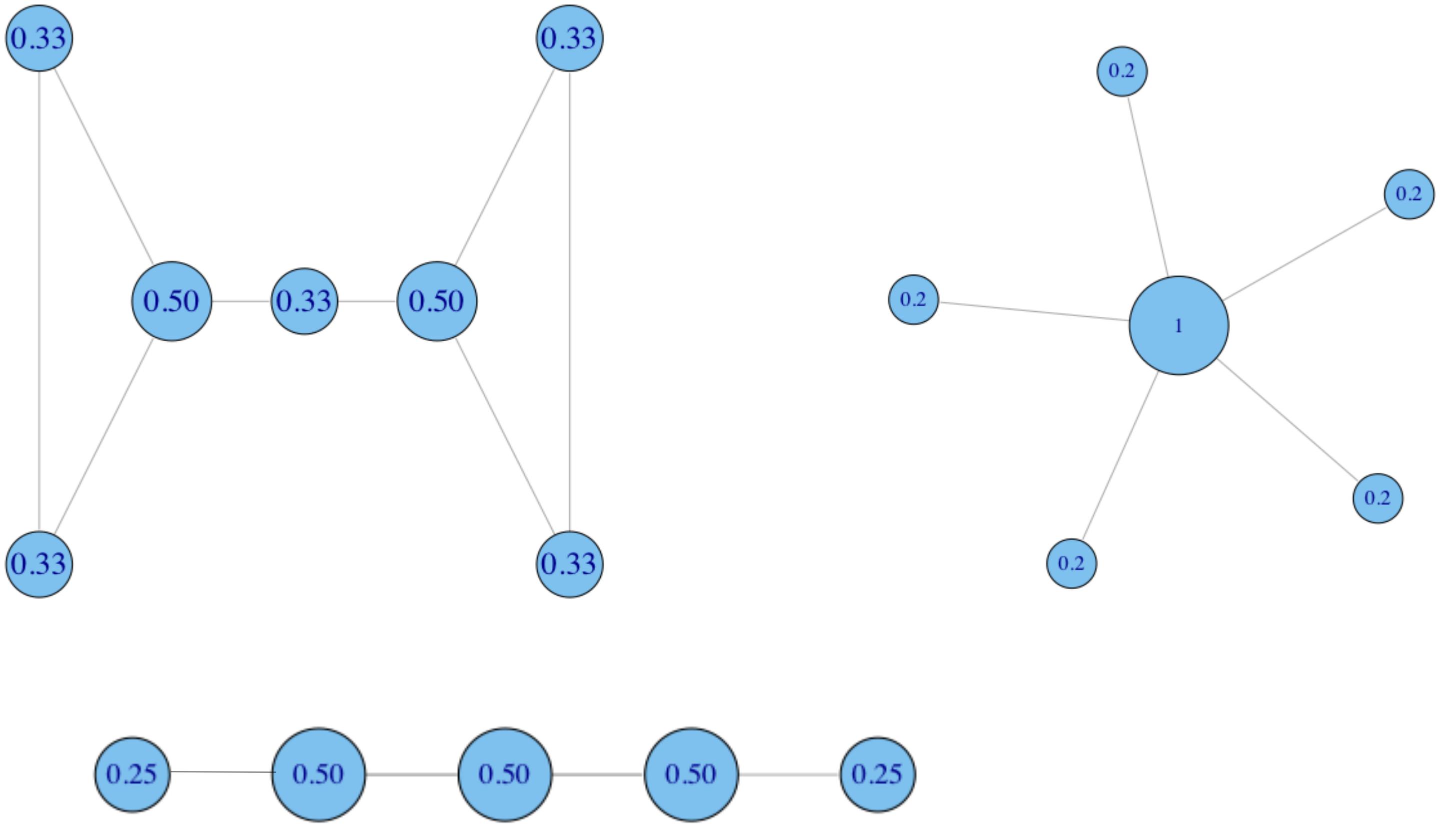
- Simply, centrality measured by degree of a node
  - A node of higher degree is more important

- Undirected graphs
  - Number of friends of a user in Facebook
  - Important stations in railway networks

- Directed graphs: usually indegree of node
  - Number of pages linking to a given page in the Web
  - Number of followers of a user in Twitter

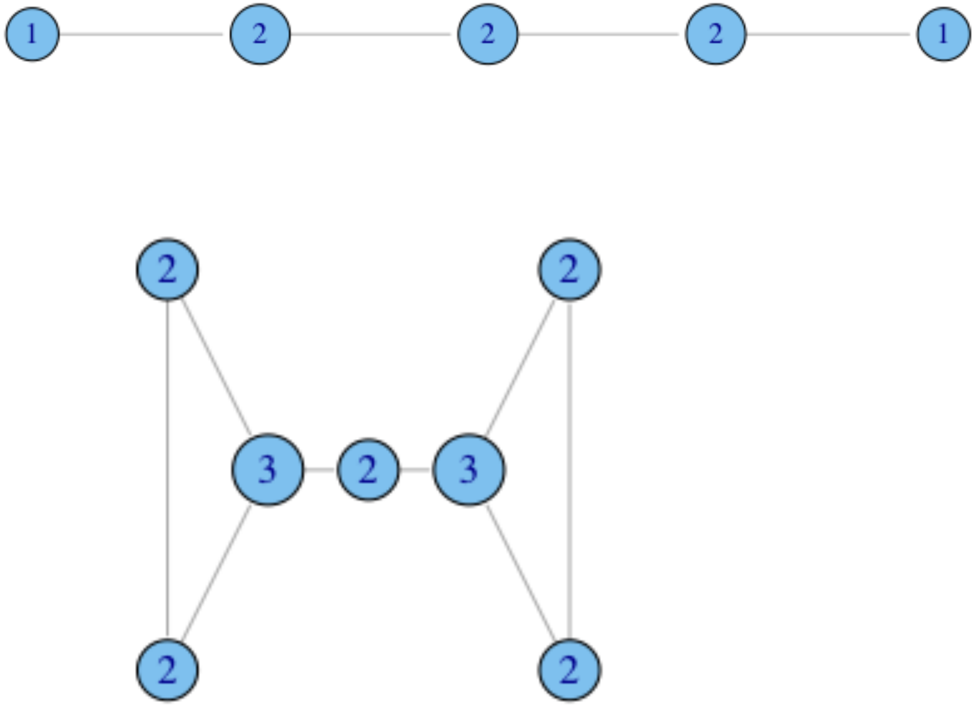
# Normalized degree centrality

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



# When degree isn't everything

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




- ability to broker between groups
- likelihood that information originating anywhere in the network reaches you…

# Closeness centrality

- Intuition
  - Farness of node $s$: sum of its shortest distances to all other nodes
  - Closeness of node $s$: inverse of farness

# Closeness centrality

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

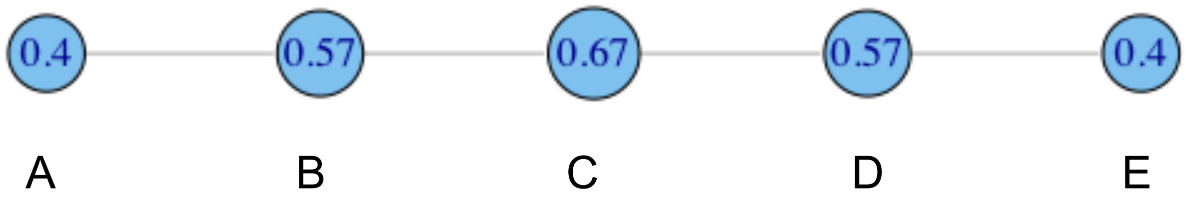
Closeness Centrality:

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

Normalized Closeness Centrality

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

# Closeness centrality: 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 centrality

- Higher the closeness centrality of *s*, the lower is its total distance to all other nodes

- Applications
  - Where to set up a hospital in a town?
  - How fast can information spread from *s* to all other nodes?

# Betweenness centrality

- 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 of node s:
  - For each pair of vertices (u, v), find the shortest paths between them (u or v is not s itself)
  - Compute the fraction of these shortest paths which pass through node s
  - Sum this fraction for all pairs of nodes (u, v)

# Betweenness centrality: definition

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

Where $g_{jk}$ = the number of geodesics connecting *jk*, and
$g_{jk}$ (i) = the number of these geodesics that actor *i* is on.

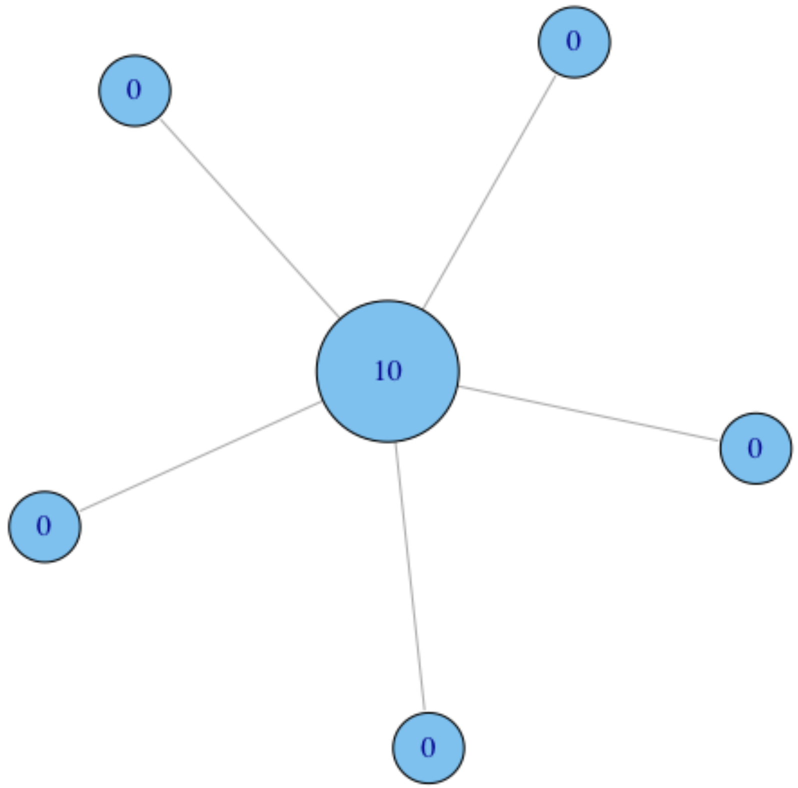
Can be normalized by:

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

number of pairs of vertices
excluding the vertex itself

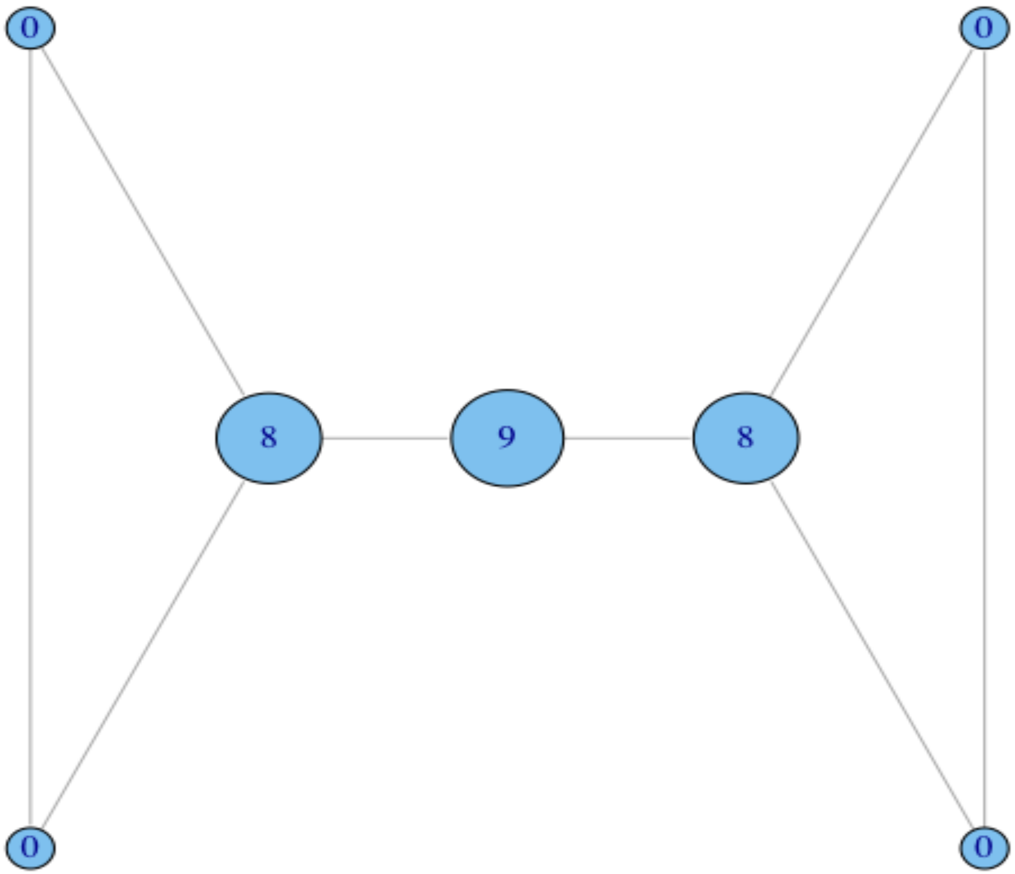
# Betweenness on toy networks

- non-normalized version:

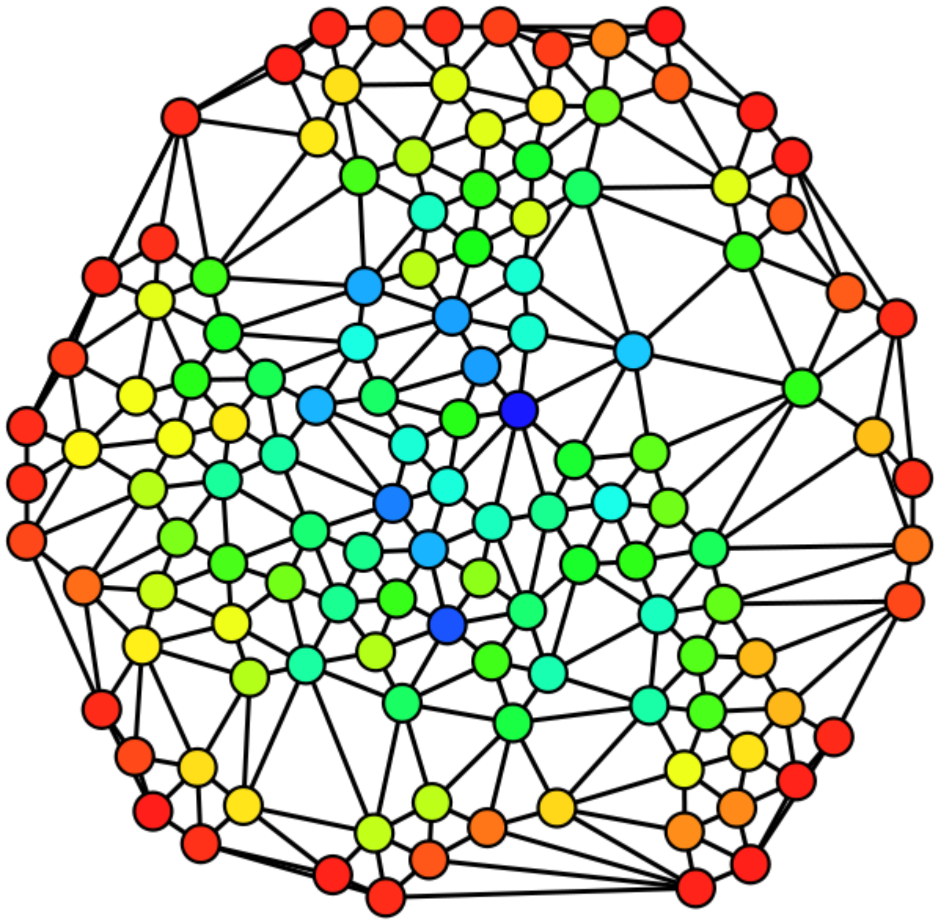


# Betweenness on toy networks

- non-normalized version:



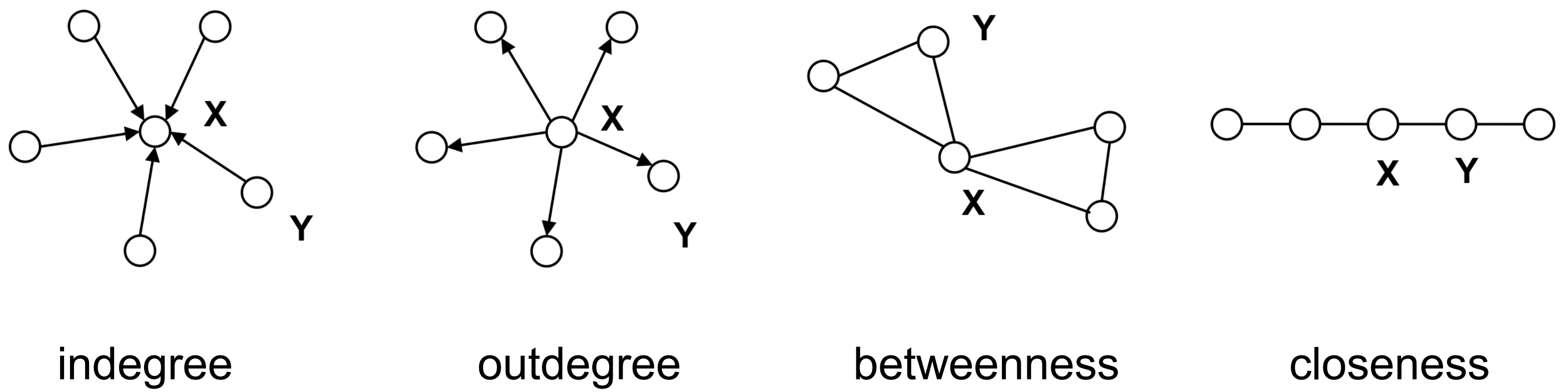
# Example of betweenness centrality



Betweenness centrality coded by color

Red: 0 betweenness
Blue: maximum betweenness

# Centrality measures - visual comparison




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