Social and Information Networks

Module 2 - CentralityMeasures

Reference Book:

Wasserman Stanley, and Katherine Faust. (2009). Social Network Analysis: Methods and Applications, Structural Analysis in the Social Sciences.

Introduction

- In a social network , we may need to identify
  - the most influential person(s)
  - key infrastructure nodes in the Internet or urban networks
  - super-spreaders of disease
  - prominent funding agencies
  - important "actors"

Introduction

- In many social settings, people with more connections tend to have more power and more visible
- Degree is often a highly effective measure of the influence or importance of a node
- In an undirected graph
  - Nodes with higher degree are more central
- In a directed graph
  - Node with higher outdegree is more **central** (choices made)
  - Node with higher indegree is more **prestigious** (choices received)

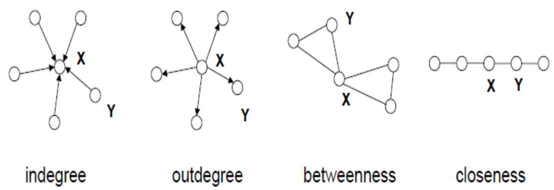
Introduction

- He or she who has many friends => most important
- Centrality: **A measure of who is important** (based on their network position)
- When is the number of connections the best centrality measure?
  - People who will do favors for you
  - People you can talk to / have coffee with

Example: Star Graph

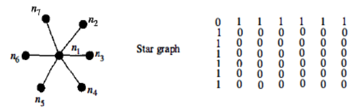
Centrality - Introduction

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



Centrality - Introduction

- An actor is prominent if the ties of the actor makes the actor visible to the other actors in the network.
- An actor's **prominence** reflects its greater visibility to the other network actors (an audience).
- An actor's prominent location takes account of the direct sociometric choices made and choices received (outdegrees and indegrees), as well as the indirect ties with other actors.
- Prominence is based on the pattern of (g-1) (or 2*(g-1)) possible ties in the sociomatrix defining the location of actor i



Centrality - Introduction

The two basic prominence classes are:

Centrality: Actor has high involvement in many relations, regardless of send/receive directionality (volume of activity)

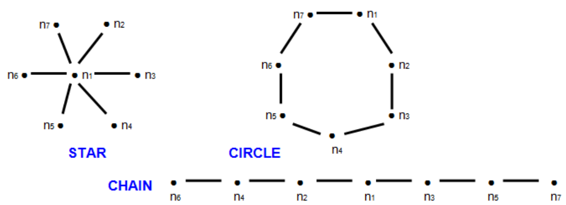
Prestige: Actor receives many directed ties, but initiates few relations (popularity > extensivity)

Centrality - Introduction

- An actor with high degree centrality maintains numerous contacts with other network actors
- A central actor occupies a structural position (network location) that serves as a source or conduit for larger volumes of information exchange and other resource transactions with other actors
- Central actors are located at or near the center in a network
- A peripheral actor maintains few or no relations and thus is located spatially at the margins of a network diagram

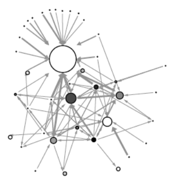
Centrality - Introduction

- Examples:

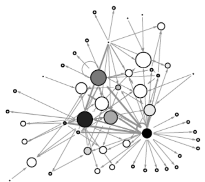


Centrality - Introduction

- Example: Financial trading networks



high centralization: one node trading with many others



low centralization: trades are more evenly distributed

Centrality - Introduction

- Three most widely used centrality measures are
  - Degree
  - Closeness
  - Betweenness

Degree Centrality

- **Actor-level degree centrality** is simply each actor's number of degrees in an undirected graph

  $C_D(n_i) = d(n_i)$ (or)

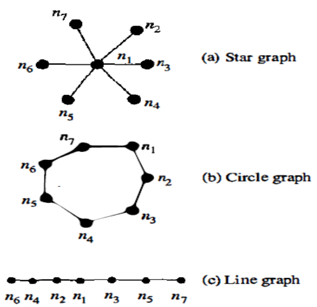
  $$C_D(n_i) = d(n_i) = x_{i+} = \sum_j x_{ij} = \sum_j x_{ji}$$

- To standardize or normalize the degree centrality index, divide it by the maximum possible degree

  $C'_D(n_i) = d(n_i)/(g-1)$

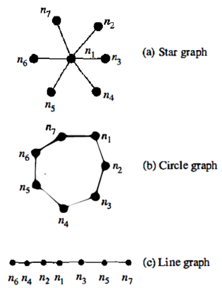
  - Networks of different sizes (g) may be compared

Degree Centrality - Examples



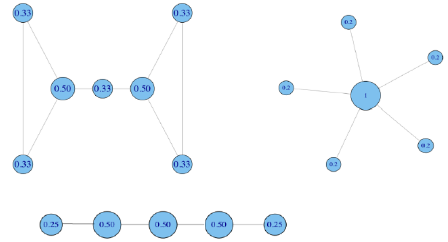
Degree Centrality - Examples

- In the star graph(a), the most central actor (n_1) has degree centrality = 6 but the six peripheral actors each have degree centrality = 1; their standardized values are 1.00 and 0.167, respectively.
- All seven circle graph(b) actors have identical degree centrality (=2), so no central actor exists; their standardized values are each 0.333.
- In the line graph(c), the two end actors have smaller degree centralities (degrees = 1) than those in the middle (=2); the respective standardized scores are 0.167 and 0.333.



Degree Centrality - Examples

Normalized Degree Centrality - Examples

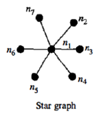


Degree Centrality

Group degree centrality :

- Quantifies the dispersion or variation among individual centralities (Freeman Group Degree Centrality)

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



  - $C_D(n^*)$ is the largest observed value across g actors
  - Reaches its maximum value of unity when one actor "chooses" all other g - 1 actors (that is, has geodesics of length 1 to all the other actors), and the other actors have geodesics of length 2 to the remaining (g - 2) actors
- Determines how centralized the degree of the set of actors is
- A measure of the dispersion or range of the actor indices, since it compares each actor index to the maximum attained value
- Examples: Star Graph, Circle Graph

Closeness Centrality

- 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

Closeness Centrality

- Closeness measure focuses on how close an actor is to all the other actors in the set of actors
- An actor is central if the actor can quickly interact with all others
- A central ego actor has minimum path distances from the g-1 alters
- An actor that is close to many others can quickly interact and communicate with them without going through many intermediaries
- A measure of how long it will take to spread information from actor 'v' to all other nodes sequentially
- **Fairness/peripherality** of a node 'v' is defined as the sum of its distances to all other nodes
- **Actor Closeness is the inverse of fairness**

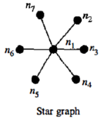
Closeness Centrality

- Actor Closeness is the inverse of fairness = 1/ total distance that actor i is from all other actors

$$C_C(n_i) = \left[\sum_{j=1}^{g} d(n_i, n_j)\right]^{-1}$$

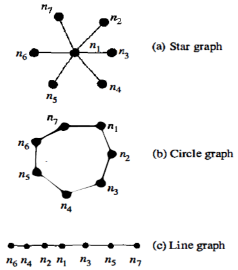
  - Inverse of the sum of geodesic distances from actor i to the g-1 other actors
- Standardizing the indices so that the maximum value equals unity ,

$$C'_C(n_i) = \frac{g-1}{\left[\sum_{j=1}^{g} d(n_i,n_j)\right]} = (g-1)C_C(n_i)$$



Closeness Centrality - Examples

- In the star graph(7 nodes), actor n_1 has closeness = 1.0 while the six peripheral actors, closeness= 0.545
- All circle graph (7 nodes) actors have the same closeness (0.50)
- In the chain (line) graph (7 nodes), the two end actors are less close (0.286) than those in the middle (0.375, 0.4615, 0.50)



$$C'_C(n_i) = \frac{g-1}{\left[\sum_{j=1}^{g} d(n_i,n_j)\right]} = (g-1)C_C(n_i)$$

Closeness Centrality – Examples (cont'd)

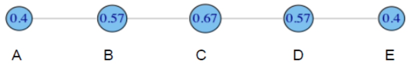
- In the star graph(7 nodes),
  - $C_C'(n_1) = 6/(1+1+1+1+1+1)=1.0$
  - $C_C'(n_2)=C_C'(n_3)=C_C'(n_4)=C_C'(n_5)=C_C'(n_6)=C_C'(n_7) = 6/(1+2+2+2+2+2)=0.545$
- All circle graph (7 nodes) actors have the same closeness (6/(1+2+3+1+2+3) = 0.50)
- In the chain (line) graph (7 nodes),
  - the two end actors are less close (0.286)
    - 6/(1+2+3+4+5+6) =0.286
  - than those in the middle (0.375, 0.461, 0.50)
    - 6/(1+1+2+3+4+5) = 0.375
    - 6/(1+2+1+2+3+4) = 0.4615
    - 6/(1+2+3+1+2+3) = 0.5

$$C'_C(n_i) = \frac{g-1}{\left[\sum_{j=1}^{g} d(n_i,n_j)\right]} = (g-1)C_C(n_i)$$

Closeness Centrality

Example:

Closeness Centrality: Toy Example

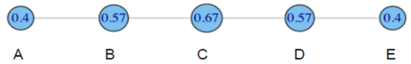


$$C'_c(A) = \left[\frac{\sum_{j=1}^{N} d(A,j)}{N-1}\right]^{-1} = \left[\frac{1+2+3+4}{4}\right]^{-1} = \left[\frac{10}{4}\right]^{-1} = 0.4$$

Closeness Centrality

Example:

Closeness Centrality: Toy Example

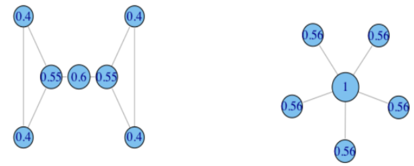


$$C'_C(n_i) = \frac{g-1}{\left[\sum_{j=1}^{g} d(n_i,n_j)\right]} = (g-1)C_C(n_i)$$

$$= \left[\frac{1+2+3+4}{4}\right]^{-1} = \left[\frac{10}{4}\right]^{-1} = 0.4$$

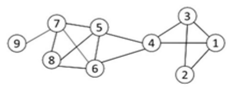
Closeness Centrality

- Examples:



Closeness Centrality

Closeness Centrality Example



Pairwise geodesic distance

Node	1	2	3	4	5	6	7	8	9
1	0	1	1	1	2	2	3	3	4
2	1	0	1	2	3	3	4	4	5
3	1	1	0	1	2	2	3	3	4
4	1	2	1	0	1	1	2	2	3
5	2	3	2	1	0	1	1	1	2
6	2	3	2	1	1	0	1	1	2
7	3	4	3	2	1	1	0	1	1
8	3	4	3	2	1	1	1	0	2
9	4	5	4	3	2	2	1	2	0

$$C_C(3) = \frac{9-1}{1+1+1+2+2+3+3+4} = 8/17 = 0.47$$

$$C_C(4) = \frac{9-1}{1+2+1+1+1+2+2+3} = 8/13 = 0.62$$

Node 4 is more central than node 3

Betweenness Centrality

- A central actor occupies a "between" position on the geodesics connecting many pairs of other actors in the network
- As a cutpoint in the shortest path connecting two other nodes, a between actor might control the flow of information or the exchange of resources
- If more than one geodesic links a pair of actors, assume that each of these shortest paths has an equal probability of being used
- Let g_{jk} be the number of geodesics linking the two actors
- Actor betweenness index for n, is simply the sum of these estimated probabilities over all pairs of actors not including the i^{th} actor:

$$C_B(n_i) = \sum_{j<k} \frac{g_{jk}(n_i)}{g_{jk}}$$

Betweenness Centrality

- Actor betweenness is maximum when the i^{th} actor falls on all geodesics [When maximum number of pairs of actors (not including n_i) is clearly (g - 1)(g - 2)/2]
- The index reaches the maximum when the i^{th} actor falls on all geodesics
- Hence

$$C'_B(n_i) = C_B(n_i)/[(g-1)(g-2)/2]$$

Betweenness Centrality

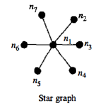
- In the star graph, actor n_1 has betweenness = 1.0 while the six peripheral actors = 0.0.
- All circle graph actors have the same betweenness (0.2).
- In the chain graph, the two end actors have no betweenness (0.0), the exactly middle actor n_1 has the highest betweenness (0.60), while the two adjacent to it are only slightly less central (0.53), (last but one node goes with 0.33)

Betweenness Centrality

- In the star graph(7 nodes),
  - $C_B'(n_1)$ = ((1/1+1/1+1/1+1/1+1/1)+(1/1+1/1+1/1+1/1)+(1/1+1/1+1/1)+(1/1+1/1)+(1/1))/15=1.0
  - $C_B'(n_2)=C_B'(n_3)=C_B'(n_4)=C_B'(n_5)=C_B'(n_6)=C_B'(n_7)$ = 0
- All circle graph (7 nodes) actors have the same betweenness(1/1+1/1+ 1/1)/(6*5/2) = 0.2)
- In the chain (line) graph (7 nodes),
  - the two end actors have no betweenness (0.0)
  - the exactly middle actor n_1 has the highest betweenness (0.60), while the two adjacent to it are only slightly less central (0.53), (last but one node goes with 0.33)
  - ((1/1+1/1+1/1)+(1/1+1/1)+(1/1+1/1)+(1/1)+(1/1)) / 15 = 0.6
  - ((1/1+1/1)+(1/1+1/1)+(1/1+1/1)+(1/1)+(1/1)) / 15 = 0.53
  - ((1/1)+(1/1)+(1/1)+(1/1)+(1/1)) / 15 = 0.33

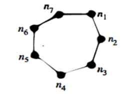
Betweenness Centrality

- In the star graph(7 nodes),
  - $C_B'(n_1)$ = ((1/1+1/1+1/1+1/1+1/1)+(1/1+1/1+1/1+1/1)+(1/1+1/1+1/1)+(1/1+1/1)+(1/1))/15=1.0
  - Description (SP -> Shortest Path)
  - SP from n_2 to (n_3, n_4, n_5 , n_6, n_7) passes thru n1 -> (1/1+1/1+1/1+1/1+1/1)
  - SP from n_3 to (n_4, n_5 , n_6, n_7) passes thru n1 -> (1/1+1/1+1/1+1/1)
  - SP from n_4 to (n_5 , n_6, n_7) passes thru n1 -> (1/1+1/1+1/1+1)
  - SP from n_5 to (n_6, n_7) passes thru n1 -> (1/1+1/1)
  - SP from n_6 to (n_7) passes thru n1 -> (1/1)
  - $C_B'(n_2)=C_B'(n_3)=C_B'(n_4)=C_B'(n_5)=C_B'(n_6)=C_B'(n_7)$ = 0 (Example: None of the SPs from n_1 to other five nodes passes thru n_2 . Hence scope is 0/1)



Star graph

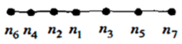
- All circle graph (7 nodes) actors have the same betweenness((1/1+1/1+ 1/1)/(6*5/2) = 0.2)
  - Description (SP -> Shortest Path)
  - SP from n_2 to (n_6, n_7) passes thru n1 -> (1/1+1/1)
  - SP from n_3 to (n_7) passes thru n1 -> (1/1)



Circle graph

Betweenness Centrality

- In the chain (line) graph (7 nodes),
  - the two end actors have no betweenness (0.0)
  - the exactly middle actor n_1 has the highest betweenness (0.60), while the two adjacent to it are only slightly less central (0.53), (last but one node goes with 0.33)
  - ((1/1+1/1+1/1)+(1/1+1/1)+(1/1+1/1)+(1/1)+(1/1)) / 15 = 0.6
    SP from n_2 to (n_3, n_5 , n_7) passes thru n_1 -> (1/1+1/1+1/1)
    SP from n_3 to (n_4, n_6) passes thru n_1 -> (1/1+1/1)
    SP from n_4 to (n_5 , n_7) passes thru n_1 -> (1/1+1/1)
    SP from n_5 to (n_6) passes thru n_1 -> (1/1)
    SP from n_6 to (n_7) passes thru n_1 -> (1/1)
  - ((1/1+1/1)+(1/1+1/1)+(1/1+1/1)+(1/1)+(1/1)) / 15 = 0.53
    SP from n_1 to (n_4, n_6) passes thru n_2 -> (1/1+1/1)
    SP from n_3 to (n_4, n_6) passes thru n_2 -> (1/1+1/1)
    SP from n_4 to (n_5 , n_7) passes thru n_2 -> (1/1+1/1)
    SP from n_5 to (n_6) passes thru n_2 -> (1/1)
    SP from n_6 to (n_7) passes thru n_2 -> (1/1)
  - ((1/1)+(1/1)+(1/1)+(1/1)+(1/1)) / 15 = 0.33
    SP from n_1 to (n_6) passes thru n_4 -> (1/1)
    SP from n_2 to (n_6) passes thru n_4 -> (1/1)
    SP from n_3 to (n_6) passes thru n_4 -> (1/1)
    SP from n_5 to (n_6) passes thru n_4 -> (1/1)
    SP from n_6 to (n_7) passes thru n_4 -> (1/1)



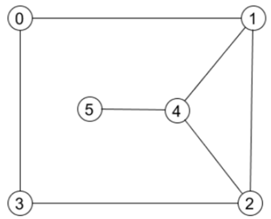
Line graph

$$C_B(n_i) = \sum_{j<k} \frac{g_{jk}(n_i)}{g_{jk}}$$

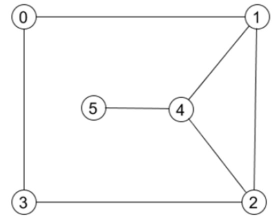
$$C'_B(n_i) = C_B(n_i)/[(g-1)(g-2)/2]$$

Example – Betweenness Centrality

- Example



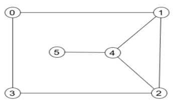
Example – Betweenness Centrality NonStandardized



v	$C_B(v)$
0	$\frac{1}{2}$
1	$\frac{1}{2}+1+1=\frac{5}{2}$
2	$\frac{1}{2}+1+1=\frac{5}{2}$
3	$\frac{1}{2}$
4	$1+1+1+1$
5	0

Table: Betweenness Centrality

Example – Betweenness Centrality NonStandardized (cont'd)

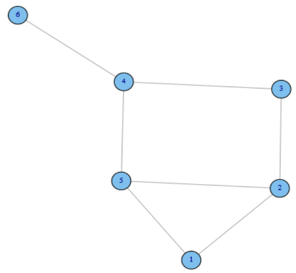


v	$C_B(v)$
0	$\frac{1}{2}$
1	$\frac{1}{2}+1+1=\frac{5}{2}$
2	$\frac{1}{2}+1+1=\frac{5}{2}$
3	$\frac{1}{2}$
4	$1+1+1+1$
5	0

Table: Betweenness Centrality

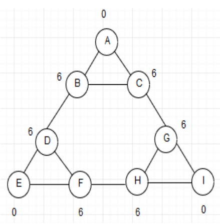
- Description
  - $C_B(0)$ = 0.5
    One SP out of two SPs from node 1 to 3 passes thru 0 -> (1/2)
  - $C_B(1)$ = (1/2+1+1 = 2.5)
    One SP out of two SPs from node 0 to 2 passes thru 1 -> (1/2)
    SP from node 0 to 4 passes thru 1 -> (1/1)
    SP from node 0 to 5 passes thru 1 -> (1/1)
  - $C_B(2)$ = 2.5
    One SP out of two SPs from node 1 to 3 passes thru 2 -> (1/2)
    SP from node 3 to (4, 5) passes thru 2 -> (1/1 + 1/1)
  - $C_B(3)$ = 0.5
    One SP out of two SPs from node 0 to 2 passes thru 3 -> (1/2)
  - $C_B(4)$ = 4
    SP from nodes (0,1,2,3) to (5) passes thru 4 -> (1/1 + 1/1 +1/1+1/1)
  - $C_B(5)$ = 0

Example – Betweenness Centrality NonStandardized



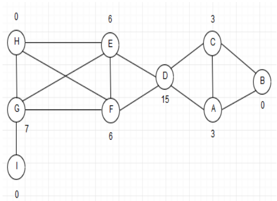
Node	Betwenness
1	0
2	1.5
3	1
4	4
5	3
6	0

Example – Betweenness Centrality NonStandardized



- $C_B(A) = C_B(E) = C_B(I) = 0$
- $C_B(B) = 6$
  SP from node A to {D,E,F} passes thru B
  -> (1/1 + 1/1 +1/1)
  SP from node C to {D.E} passes thru B
  -> (1/1 + 1/1)
  One SP from node C to F passes thru B
  -> 1/2
  One SP from node D to G passes thru B
  -> 1/2
- $C_B(D) = 6$
- $C_B(C) = 6$

Example – Betweenness Centrality NonStandardized



- $C_B(B) = C_B(H) = C_B(I) = 0$
- $C_B(D) = 15$

  SP from node A to {E,F,G,H,I} passes thru D

  -> 5

  SP from node B to {E,F,G,H,I} passes thru D

  -> 5

  SP from node C to {E,F,G,H,I} passes thru D

  -> 5
- $C_B(C) = C_B(A) = 3$

  SP from node B to {D,E,F,G,H,I} passes thru C

  -> 6 * 1 / 2 = 3
- $C_B(E) = C_B(F) = 6$

  SP from nodes {A,B,C,D} to {G,H,I} passes thru E

  -> 4 * 3 * 1/2 = 6
- $C_B(G) = 7$

  SP from nodes {A,B,C,D,E,F,H} to I passes thru G -> 7