Social and Information Networks

Module 2 - Centrality Measures

Reference Book:

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

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
  - 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

Group Degree Centrality

Other Measures based on Degree

- Standard statistical summary of the actor degree indices is the
  - Variance of the degrees, $S_D^2 = \left[\sum_{i=1}^{g}(C_D(n_i) - \overline{C}_D)^2\right] / g$
- Mean Degree, $\overline{C}_D = \sum C_D(n_i)/g$
- Standardized Average Degree = Mean Degree/ (g-1)

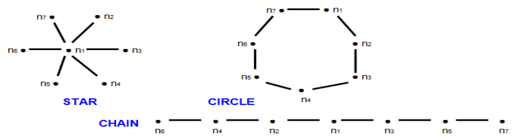
$$\sum C_D(n_i)/g(g-1) = \sum C'_D(n_i)/g = \Delta$$

- Hence, Density is also the Standardized Average Degree

Group Degree Centrality

Other Measures based on Degree

- Density of a graph is the most widely used group-level index
- Densities of the three graphs(given below) are 0.286 (star), 0.333 (circle)and 0.286 (line).



Group Closeness Centralization

Freeman's general group closeness index is based on the standardized actor closeness centralities,

$$\sum_{i=1}^{g}[C'_C(n^*) - C'_C(n_i)]$$

•where $C'_C(n^*)$ is the largest standardized actor closeness in the set of actors.

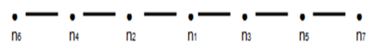
•Freeman shows that the maximum possible value for the numerator is [(g - 2)(g - 1)]/(2g - 3),

•Hence the index of Group Closeness is

$$C_C = \frac{\sum_{i=1}^{g}[C'_C(n^*) - C'_C(n_i)]}{[(g-2)(g-1)]/(2g-3)}$$

Group Closeness Centralization

- Example:
- For the line graph, the index is 0.277



$$C_C = \frac{\sum_{i=1}^{g}[C'_C(n^*) - C'_C(n_i)]}{[(g-2)(g-1)]/(2g-3)}$$

- $C'_c(n_1)$ = (7-1) / (1+ 1+ 2+ 2+3+3) = 0.5
- $C'_c(n_6)$ = $C'_c(n_7)$ = (7-1) / (1+ 2+ 3+4+5+6) = 0.286
- $C'_c(n_4)$ = $C'_c(n_5)$= (7-1) / (1+ 1+ 2+ 3+4+5) = 0.375
- $C'_c(n_2)$ = $C'_c(n_3)$= (7-1) / (1+ 2+ 1+ 2+ 3+4) = 0.462
- C_C =((0.5 -0.5)+(0.5 - 0.286) *2 +(0.5 - 0.375)*2 + (0.5 - 0.462)*2) / (5*6)/11)
  = 0.277

Group Betweenness Centralization

- Group centralization indicies based on betweenness
  - allows a researcher to compare different networks with respect to the heterogeneity of the betweenness of the members of the networks
- Freeman's group betweenness centralization index has numerator

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

  - where $C_B(n^*)$ is the largest realized actor betweenness index for the set of actors
- Freeman shows that the maximum possible value for this sum is (g - 1)(g - 1)(g - 2)/2,

Group Betweenness Centralization

- Hence, the index of Group Betweenness is

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

- With standardized indices, Group Betweenness is

$$C_B = \frac{\sum_{i=1}^{g}[C'_B(n^*) - C'_B(n_i)]}{(g-1)}$$

  - since

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

Centrality - Directional Relations

Degree Centrality:

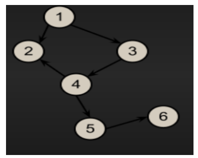
- Centrality indices focus on the choices made, outdegree of each actor is considered rather than the degree
- Actor-level index of degree centralization is

$$C'_D(n_i) = x_{i+}/(g-1)$$

Betweenness Centrality

- Actor Betweenness Centrality measure

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



Centrality - Directional Relations

Closeness Centrality:

Given the g X g distance matrix

- Actor-level centrality indices for closeness are calculated by taking the sum of row i of the distance matrix to obtain the total distance n_i is from all the other actors, and then dividing by g - 1 (the minimum possible total distance).
  - The reciprocal of this ratio gives us an actor-level index for closeness

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

- Actor-level centrality index based on closeness cannot be defined if the digraph is not strongly connected
  - Because, some of the { d(n_i , n_j)} will be infinity
- Hence, consider only those actors that 'I' can reach, ignoring those that are unreachable from 'I'.

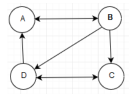
Centrality - Directional Relations

- Actor-level centrality index can be generalized by considering the influence range of n_i as the set of actors who are reachable from n_i
  - This set contains all actors who are reachable from i in a finite number of steps
- Define J_i -> the number of actors in the influence range of actor i (equals the number of actors who are reachable from n_i)
- **Improved actor-level closeness centrality index**
  - considers how proximate n_i is to the actors in its influence range
  - focuses on distances from actor i to the actors in its influence range
  - average distance these actors are from n_i is $\sum d(n_i, n_j)/J_i$
  - a ratio of the fraction of the actors in the group who are reachable to the average distance that these actors are from the actor i

$$C_C^{\bullet}(n_i) = \frac{J_i/(g-1)}{\sum d(n_i, n_j)/J_i}$$

Centrality - Directional Relations

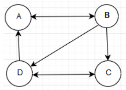
- Example



$$C_C^{\bullet}(n_i) = \frac{J_i/(g-1)}{\sum d(n_i, n_j)/J_i}$$

Centrality - Directional Relations

- Example (cont'd)



$$C_C^{\bullet}(n_i) = \frac{J_i/(g-1)}{\sum d(n_i, n_j)/J_i}$$

Node	C_D	C_C	C_B
A	1	(3/3) / ((1+1+2+2+2+1)/3) = 1/3	2 (C->B,D->B)
B	2	(3/3) / ((1+1+1+3+1+2)/3) = 1/3	2 (A->C,A->D)
C	1	(3/3) / ((2+2+3+1+1+1)/3) = 3/10	0
D	2	(3/3) / ((1+2+2+1+1+1)/3) = 3/8	2 (C->A,C->B)

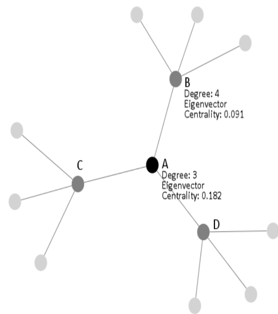
Eigen Vector Centrality Measure

- Degree Centrality depends on the connections a node has
  - But what if these connections are pretty isolated?
- A central node should be one connected to powerful nodes
- Eigenvector centrality (Eigencentrality) is a measure of the influence of a node in a network
- Eigenvector centrality expands upon the notion of the degree of a node, incorporating information about the degree of a node's alters
- Degree for node A in a social network measures how many ties A has
- **Eigenvector centrality of node A is measured based on how many ties A's alters have.**
- Google's PageRank are variants of the eigenvector centrality

Eigen Vector Centrality Measure

Why Eigen Vector Centrality

- Node A has a degree of three, Node B, has a degree of four. Node B is more popular in the network if we only extend our vision out to a distance of 1 from each node.
- But A is connected to nodes that are connected to many other nodes, while B is connected to less-popular nodes



Eigen Vector Centrality Measure

Eigenvalues and eigenvectors of matrices

- Consider *n*-dimensional vectors that are formed as a list of *n* scalars, such as the three-dimensional vectors

$$x = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} \quad \text{and} \quad y = \begin{bmatrix} -20 \\ -60 \\ -80 \end{bmatrix}$$

- These vectors are said to be scalar multiples of each other, if there is a scalar λ such that

  x = λ y , Here λ = −1/20

- Let us consider the linear transformation of n-dimensional vectors defined by an n by n matrix A,

  A v = w

Eigen Vector Centrality Measure

- If *v* and *w* are scalar multiples, A v = w = λ v
  - then v is an eigenvector of the linear transformation A and the scale factor λ is the eigenvalue corresponding to that eigenvector
- Equation A v = w = λ v is the eigenvalue equation for the matrix A
- Equivalently,

  (A − λ I) v = 0

  - where I is the n by n identity matrix

$$Av = w,$$

or

$$\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

where, for each row,

$$w_i = A_{i1}v_1 + A_{i2}v_2 + \cdots + A_{in}v_n = \sum_{j=1}^{n} A_{ij}v_j$$

Eigen Vector Centrality Measure

- (A − λ I) v = 0 , has a non-zero solution v if and only if the determinant of the matrix (A − λI) is zero.
  - Eigenvalues of A are values of λ that satisfy the equation | A − λ I | = 0
  - Obtain the Characteristic equation and solve for λ
  - Substitute λ in Eigen value equation to get the eigen vector

Example

Eigenvalues of a square matrix: If A is an $n \times n$ matrix, the eigenvalues of A are the solutions to the characteristic equation

$$\mathrm{Det}(A - \lambda I_n) = 0$$

where λ is a variable.

Example: Suppose

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$$

Note that

$$\lambda I_2 = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$

So,

$$A - \lambda I_2 = \begin{bmatrix} 1-\lambda & 3 \\ 2 & 2-\lambda \end{bmatrix}$$

Hence,

$$\mathrm{Det}(A - \lambda I_2) = (2-\lambda)(1-\lambda) - 6 = \lambda^2 - 3\lambda - 4$$

Example (cont'd)

So, the characteristic equation for A is given by

$$\lambda^2 - 3\lambda - 4 = 0$$

- The solutions to this equation are: $\lambda = 4$ and $\lambda = -1$.
- These are the **eigenvalues** of the matrix A.
- The largest eigenvalue (in this case, $\lambda = 4$) is called the **principal** eigenvalue.
- For each eigenvalue λ of A, there is a 2×1 matrix (vector) x such that $Ax = \lambda$x. Such a vector is called an **eigenvector** of the eigenvalue λ.
- For the above matrix A, for the principal eigenvalue $\lambda = 4$, an eigenvector x is given by

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Eigen Vector Centrality Measure

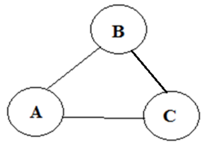
- **Using the adjacency matrix to find Eigenvector centrality**
- For a given graph G := (V , E) with | V | number of vertices let A = ($a_{v,t}$) be the adjacency matrix, i.e. $a_{v,t}$ = 1 is linked to vertex t or 0 otherwise.
  - The relative centrality score of vertex v can be defined as:

$$x_v = \frac{1}{\lambda} \sum_{t \in M(v)} x_t = \frac{1}{\lambda} \sum_{t \in G} a_{v,t} x_t$$

  - where M (v) is a set of the neighbors of v and λ is a constant
  - In vector notation, the eigenvector equation is

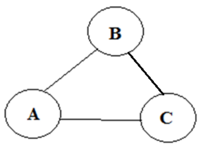
    A x = λ x
- There will be many different eigenvalues λ for which a non-zero eigenvector solution exists.
- Since the entries in the adjacency matrix are non-negative, by the Perron–Frobenius theorem, the greatest eigenvalue results in the desired centrality measure

Eigen Vector Centrality Measure- Example



Eigen Vector Centrality Measure- Example

- Solution



$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix}$$

$$= -\lambda^3 + 3\lambda + 2$$

$$= (\lambda - 2)(\lambda + 1)^2 \qquad \lambda_1 = 2,\ \lambda_{2,3} = -1$$

Example (cont'd)

- Solution (cont'd)

$$\lambda_1 = 2 \qquad \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\eta_1 = \eta_3$$
$$\eta_2 = \eta_3$$

The eigenvector is

$$\vec{\eta} = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix} = \begin{pmatrix} \eta_3 \\ \eta_3 \\ \eta_3 \end{pmatrix}, \qquad \eta_3 \neq 0$$

$$\vec{\eta}^{(1)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad \eta_3 = 1$$

Eigen Vector Centrality Measure- Example