SOCIAL NETWORK ANALYSIS — CENTRALITY MEASURES

ICS 668 FINAL PROJECT

MAY 2011

By

成玉

Cheng, Yu

Supervisor:

Daniel Suthers, PhD

Contents

List of Figures

1 Introduction

In sociology, we understand the importance of “power”, a fundamental property of social structures. But what is this “power” in social structures? How would we describe and analyze this “power”? There are some main approaches that social network analysis has developed to study “power”, and the concept of centrality analysis is in this category.

In a social network, we often find ourselves ask the question: “Who are the most important actors?” These measures of importance or prominence attempt to describe a property of actor location in a network. From studying the social networks, we learn that important actors typically occupy strategic location in a network. [1]

In this paper we are going to discuss the degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality. Each of these approaches describes the locations of individuals in terms of how close they are to the "center" of the action in a network. Among these approaches the definitions of center differ. Within different analysis, sociologists have studied carefully why the “central” positions tend to be more powerful positions compare to the rest of the network. In the Discussion section we will see how these centrality measures are used in the complex sociotechnical systems.

We will consider measurements for individual actors in this paper, but a group of actors can be aggregated and measured as a unit as well. From the network’s point of view, it can be a network with vertices that are made of multiple actors. For instance, in a contact social network, individuals form a network, and subgraphs of goups of people also forms a network.

2 Social Network Graphs

In the centrality measures, we describe social phenomenons with respect to its corresponding social network, which is a social structure made up of individuals or organizations, which are tied by one or more specific types of interdependency, such as friendship, kinship, common interest, financial exchange, dislike, sexual relationships, and so on. [2]

2.1 Types of Graphs

Graphs representing social networks have different forms depending on the social phenomenon a particular social network is describing. Several examples are given below.

2.1.1 Simple Directed Graph

A simple weighted undirected graph can be used to represent a interpersonal relation social network, such as a network showing how many courses individual students share with others.

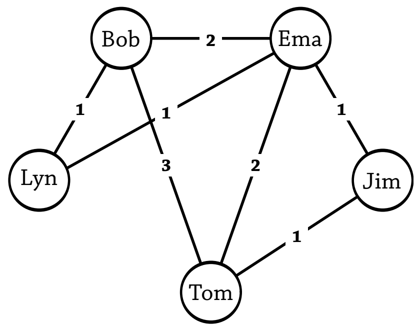

Figure 1: A social network showing course-sharing relations among students

Actor	***Courses Taken***
Bob	Math, Biology, English, History
Ema	Math, Biology, Chemistry
Lyn	Biology
Tom	Math, English, History, Chemistry
Jim	Chemistry

2.1.2 Multigraph or Pseudograph

A multigraph or pseudograph can be used to represent a information flow social network. Two actors could be connected through multiple communication channels, such as emails, phone numbers, and social network sites.

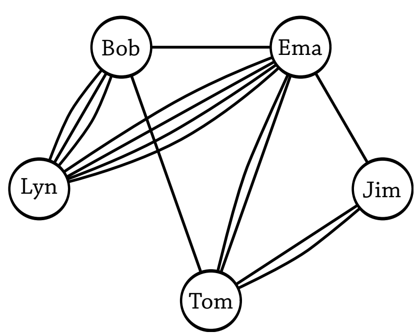

Figure 2: A social network showing the information-flow (bi-direction) relations

Pair of Actors	***Contact Methods***
Bob & Ema	Email
Bob & Lyn	Phone, Email, Skype
Bob & Tom	Mail Addresses
Lyn & Ema	Phone, Email, Skype, Mail Addresses
Ema & Tom	Phone, Mail Addresses
Ema & Jim	Email
Jim & Tom	Phone, Mail Addresses

2.1.3 Weighted Undirected Graph

A simple directed graph can be used to represent a leadership social network. The directed edges indicate the leadership relation between individuals.

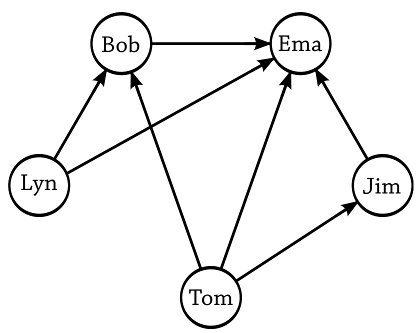

Figure 3: A social network showing the leadership relation

Actor	***Reports To***
Bob	Ema
Ema	N/A
Lyn	Bob, Ema
Tom	Bob, Ema, Jim
Jim	Ema

In the first two examples, the graphs are both undirected and hence representing a symmetric relationship. The centrality measures used to require symmetrical or undirected relations. [6] Researchers later pointed out that the use of symmetrical ties in the individual-level centrality measures may be regarded as a limitation, it is quite possible for actual relationships to be asymmetrical. [7] Directed graphs are now commonly used in centrality analysis as well.

2.2 Social Network Graph Terminology

In order to describe a graph, we will first introduce some graph related terminologies. Graph theory has a large set of specialized vocabulary. Different research areas might use different terms for the same thing. The same term might as well be used with different meanning.

2.2.1 Actor

In a social network graph, we call the vertices of the graph actors. In the social network presented above, we have 5 actors.

$$Bob,\ Ema,\ Jim,\ Tom,\ Lyn$$

2.2.2 Walk

We call a sequence of nodes connected by edges a walk. For example, in Figure 2 the following sequence of actors form a walk

$$Lyn \rightarrow Bob \rightarrow Ema \rightarrow Tom \rightarrow Ema \rightarrow Jim$$

2.2.3 Path

We call a walk with no repeated nodes a path. In the walk shown above, we see that Tom appeared two times, so this walk is not a path. The following walk form a path.

$$Lyn \rightarrow Bob \rightarrow Ema \rightarrow Jim$$

2.2.4 Geodesic

We call the shortest path between two nodes a geodesic. In the the graph presented in Figure 2, let's consider the costs for taking each edge are the same. We see that the path shown above is not the shortest way to pass information from Lyn to Jim. In this path Jim is two hops away from Lyn. The shortest path, geodesic, between Lyn and Jim is the path going through Ema. In other words, if Lyn wants to get a message to Jim fast, she would pass the message to Ema rather than Bob. The following path the geodesic between Lyn and Jim.

$$Lyn \rightarrow Ema \rightarrow Jim$$

All pair-wise shortest paths of a given simple graph can be solved with the well-known Dijkstra's algorithm within polynomial time. [3] Hence, in the rest of this paper, we will consider the geodesics are given for a given static social network.

2.2.5 Distance

We call the length of the shortest path, geodesic, the distance between two actors. Let's consider the cost for each edge to be 1. As we decided previously that the geodesic between Lyn and Jim is the path through Ema, we can compute the distance between Lyn and Jim as the following.

$$Distance_{Lyn \leftrightarrow Jim} = 1 + 1 = 2.$$

2.2.6 Adjacency Matrix

Through out this article we will use simple graphs to model the social networks. In other words, we would not consider multigraph or the reflexivity on actors. We will use adjacency matrices to represent these graphs. An adjacency matrix is a means of representing which vertices of a graph are adjacent to which other vertices. [11] In the paper, the adjacency matrix of a graph on n actors is the $n \times n$ matrix where entry a_{ij} is the edge weight between actor i and j. For the social network that shows sharing-course relations among students, we have the following adjacency matrix.

$$A = \sum_{i,j}^{n,n} x_{ij} = \begin{array}{c} \\ B \\ E \\ J \\ T \\ L \end{array} \begin{array}{c} \begin{array}{ccccc} B & E & J & T & L \end{array} \\ \begin{bmatrix} 0 & 2 & 0 & 3 & 1 \\ 2 & 0 & 1 & 2 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 3 & 2 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

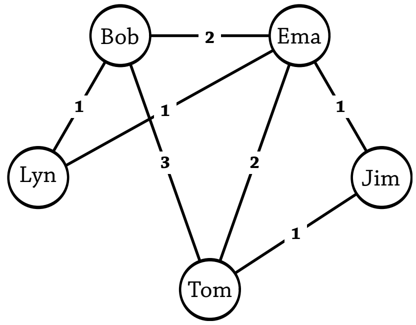

Figure 4: Adjacency matrix for a weighted simple graph

For un-weighted graphs, we will consider the edge weights as 1. In the leadership social network, we have the following adjacency matrix.

$$A = \sum_{i,j}^{n,n} x_{ij} = \begin{array}{c} \\ B \\ E \\ J \\ T \\ L \end{array} \begin{array}{c} \begin{array}{ccccc} B & E & J & T & L \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

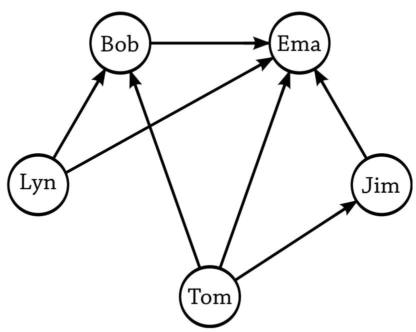

Figure 5: Adjacency matrix for a directed simple graph

As we noticed that the matrix in the first example is symmetric because the relation of the social phenomenon we are describing is symmetric. Bob is related to Ema by sharing 2 courses. Ema is also related to Bob by sharing 2 courses. The matrix in the second example is however not symmetric because the underlying social relationship not symmetric. Bob reports to Ema, so Bob is related to Ema in a way that Ema is the leader of the two, not the other way around.

3 Centrality Measures

Within graph theory and network analysis, there are various measures of the centrality of a vertex within a graph that determine the relative importance of a vertex within the graph. We are going to examine these measure with respect to various social networks.

3.1 Degree Centrality

An actor with high degree centrality is "where the action is" in the network since they are in contact or adjacent to many other actors. Actors with low degree centrality are more peripheral in the network. [1] The degree centrality is defined as the number of ties in which an actor is involved [4].

In a symmetric graph, the degree centrality is measured as the sum of the weights associated to every edge incident to the corresponding actor. It is the sum of the corresponding row or column of the adjacency matrix. It can be expressed with the following equation, where $C_D(a)$ is the degree centrality measure of the actor a, $d(a)$ is a degree of a in the graph.

$$C_D(a) = d(a).$$

For example in the following friendship network. we can compute the degree centrality for every actor.

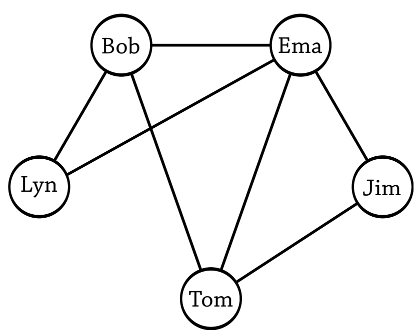

Figure 6: A friendship social network

$$
\begin{array}{ccc} & & \begin{array}{ccccc} B & E & J & T & L \end{array} \\ A & = & \begin{array}{c} B \\ E \\ J \\ T \\ L \end{array} \left[\begin{array}{ccccc} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{array} \right] \end{array}
$$

$$
\Rightarrow \quad C_D(B) = 3,\ C_D(E) = 4,\ C_D(J) = 2,\ C_D(T) = 3,\ C_D(L) = 2
$$

Degree centrality is an appropriate measure for a friendship social network if we are interested in looking for the most popular person in the network. The most popular person should have the highest number of friends. Thus, degree centrality is the most appropriate measure. In the example above, the most popular individual is Ema.

3.2 Betweenness Centrality

The betweenness centrality answers the question: how "between" is an actor to all other actors in the network? It is firstly proposed by Freeman. [5] The idea was around since very early days of sociology. Bavelas suggested in 1949 that a particular person in a group is strategically located on the shortest communication path connecting pairs of others, that person is in a central position. [9] Basically, to reach an action, many actors need to go through this actor. To have a large betweenness centrality, an actor must be between many other actors on their geodesics. Actors that are between other actors may control interactions between these other actors. That is, they could be gatekeepers. [1] There are other variants of betweenness centrality, such as the stress centrality proposed by Shimbel. [10]

The betweenness centrality is a measure of the number of times a vertex occurs on a geodesic. It can be expressed with the following equation, where $C_B(a)$ is the closeness centrality measure of the actor a, g_{ij} is the number of geodesics linking i and j, and $g_{ij}(a)$ is the number of geodesics between i and j that passes through a

$$
C_B(a) = \sum_{i<j} \frac{g_{ij}(a)}{g_{ij}}.
$$

For example in the following friendship network. we can compute the betweenness centrality.

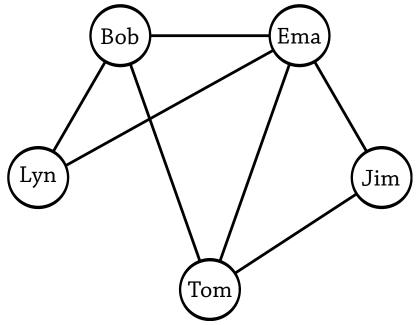

Figure 7: An information flow social network

$$
\begin{aligned}
G_{B\leftrightarrow E} = \{B \leftrightarrow E\} \quad &\Rightarrow \quad g_{BE} = 1 \\
G_{B\leftrightarrow J} = \{B \leftrightarrow E \leftrightarrow J,\ B \leftrightarrow T \leftrightarrow J\} \quad &\Rightarrow \quad g_{BJ} = 2 \\
G_{B\leftrightarrow T} = \{B \leftrightarrow T\} \quad &\Rightarrow \quad g_{BT} = 1 \\
G_{B\leftrightarrow L} = \{B \leftrightarrow L\} \quad &\Rightarrow \quad g_{BL} = 1 \\
G_{E\leftrightarrow J} = \{E \leftrightarrow J\} \quad &\Rightarrow \quad g_{EJ} = 1 \\
G_{E\leftrightarrow T} = \{E \leftrightarrow T\} \quad &\Rightarrow \quad g_{ET} = 1 \\
G_{E\leftrightarrow L} = \{E \leftrightarrow L\} \quad &\Rightarrow \quad g_{EL} = 1 \\
G_{J\leftrightarrow T} = \{J \leftrightarrow T\} \quad &\Rightarrow \quad g_{JT} = 1 \\
G_{J\leftrightarrow L} = \{J \leftrightarrow E \leftrightarrow L\} \quad &\Rightarrow \quad g_{JL} = 1 \\
G_{T\leftrightarrow L} = \{T \leftrightarrow B \leftrightarrow L,\ T \leftrightarrow E \leftrightarrow L\} \quad &\Rightarrow \quad g_{TL} = 2
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow \quad & C_B\left(B\right) = \frac{1}{2},\ C_B\left(E\right) = \frac{1}{2} + \frac{1}{2} = 1 \\
& C_B\left(J\right) = 0,\ C_B\left(T\right) = \frac{1}{2},\ C_B\left(L\right) = 0.
\end{aligned}
$$

Actors with a high betweenness centrality are interesting because they control information flow in a network. These actors may be required to carry more information. And therefore, such actors may be the subject of targeted attack. For example, in the information flow network above, if we are interested in slowing down information propagation, we would start with interfering with the actor Ema. According to the betweenness centrality study, removing Ema from this network would affect the most amount of geodesics. Specifically, this action would affect two geodesics. They are $B \leftrightarrow E \leftrightarrow J$ and $T \leftrightarrow E \leftrightarrow L$.

3.3 Closeness Centrality

The closeness centrality answers the question: how "close" is an actor to all other actors in the network? "Close" actors are able to quickly interact with many other actors. These actors are usually effective in transmitting information. Closeness is equated to "minimum distance" to other actors, geodesics. An actor with good closeness centrality would be able to propagate information to all of the other actors in the network quickly in comparison with the other actors. Therefore, as geodesics increase in length, the closeness centrality decreases.

The closeness measure is weighted inversely of the *farness* between actors. The farness measure of an actor is the sum of all geodesics between this actor with the rest of the network. The closeness measure can be expressed with the following equation, where $C_C(a)$ is the closeness centrality measure of the actor a, $g(a,b)$ is length of the geodesic between actor a and actor b.

$$C_C(a) = \left[\sum_{i \in G-a} g(a,i) \right]^{-1}.$$

For example in the following information flow network. we can compute the betweenness centrality.

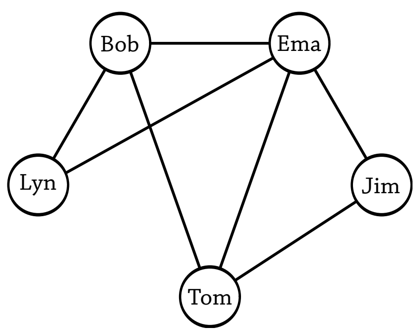

Figure 8: An information flow social network

$$G_{B \leftrightarrow E} = \{B \leftrightarrow E\} \Rightarrow g(B,E) = g(E,B) = 1$$

$$G_{B \leftrightarrow J} = \{B \leftrightarrow E \leftrightarrow J,\ B \leftrightarrow T \leftrightarrow J\} \Rightarrow g(B,J) = g(J,B) = 2$$

$$G_{B \leftrightarrow T} = \{B \leftrightarrow T\} \Rightarrow g(B,T) = g(T,B) = 1$$

$$G_{B \leftrightarrow L} = \{B \leftrightarrow L\} \Rightarrow g(B,L) = g(L,B) = 1$$

$$G_{E \leftrightarrow J} = \{E \leftrightarrow J\} \Rightarrow g(E,J) = g(J,E) = 1$$

$$G_{E \leftrightarrow T} = \{E \leftrightarrow T\} \Rightarrow g(E,T) = g(T,E) = 1$$

$$G_{E \leftrightarrow L} = \{E \leftrightarrow L\} \Rightarrow g(E,L) = g(L,E) = 1$$

$$G_{J \leftrightarrow T} = \{J \leftrightarrow T\} \quad \Rightarrow \quad g(J,T) = g(T,J) = 1$$
$$G_{J \leftrightarrow L} = \{J \leftrightarrow E \leftrightarrow L\} \quad \Rightarrow \quad g(J,L) = g(L,J) = 2$$
$$G_{T \leftrightarrow L} = \{T \leftrightarrow B \leftrightarrow L,\ T \leftrightarrow E \leftrightarrow L\} \quad \Rightarrow \quad g(T,L) = g(L,T) = 2$$

$$\begin{aligned}
\Rightarrow \quad & C_C(B) = [g(B,E) + g(B,J) + g(B,T) + g(B,L)]^{-1} = \frac{1}{5} \\
& C_C(E) = [g(E,B) + g(E,J) + g(E,T) + g(E,L)]^{-1} = \frac{1}{4} \\
& C_C(J) = [g(J,B) + g(J,E) + g(J,T) + g(J,L)]^{-1} = \frac{1}{6} \\
& C_C(T) = [g(T,B) + g(T,E) + g(T,J) + g(T,L)]^{-1} = \frac{1}{5} \\
& C_C(L) = [g(L,B) + g(L,E) + g(L,J) + g(L,T)]^{-1} = \frac{1}{6}.
\end{aligned}$$

Actors with a high closeness centrality are interesting because they propagate information in a social network with the most efficiency in comparison with the rest of the actors. Therefore, if we are looking for someone to quickly spread a message to a social network, we would start with sending the message to the actor with the highest closeness centrality. In this example, we would first send the message to Ema since Ema can reach the rest of the network with the smallest effort. Specifically, Ema has a farness of 4 to the rest of the network, while the other actors have either 5 or 6.

3.4 Eigenvector Centrality

Eigenvector centrality is a measure of the importance of an actor in a network. It assigns relative scores to all actors in the network based on the principle that connections to high-scoring actors contribute more to the score of the actor than equal connections to low-scoring actors. [12] The eigenvector centrality defined in this way accords each actor a centrality that depends both on the number and the quality of its connections: having a large number of connections still counts for something, but a vertex with a smaller number of high-quality contacts may outrank one with a larger number of mediocre contacts.

Using the graph notations, with the ith actor, the centrality score, x_i is proportional to the sum of the scores of all actors which are connected to it. We can express this relation with the following formula, where A_{ij} is the adjacency matrix of the network graph, n is the number of actors in this network, and λ is a constant that denotes the share each actor takes from the sum of its neighbors' centrality scores.

$$x_i = \frac{1}{\lambda} \sum_{j=1}^{n} A_{ij} x_j.$$

Before we go on, we shall look at an example. In the following trust social network. we can use the expression above to demonstrate an individual actor's eigenvector centrality, x_i's.

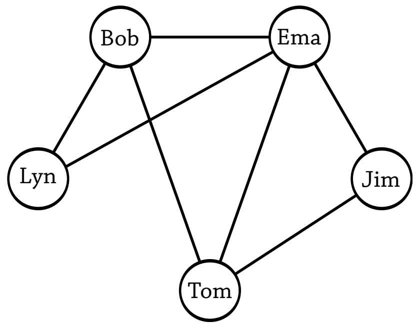

Figure 9: An trust social network

$$
\begin{array}{ccc}
 & & \begin{array}{ccccc} B & E & J & T & L \end{array} \\
A & = & \begin{array}{c} B \\ E \\ J \\ T \\ L \end{array} \left[\begin{array}{ccccc} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{array}\right]
\end{array}
$$

$$
\begin{aligned}
\Rightarrow x_B &= \frac{1}{\lambda}\left(A_{BE}\cdot x_E + A_{BJ}\cdot x_J + A_{BT}\cdot x_T + A_{BL}\cdot x_L\right) \\
&= \frac{1}{\lambda}\left(x_E + x_T + x_L\right) \\
x_E &= \frac{1}{\lambda}\left(A_{EB}\cdot x_B + A_{EJ}\cdot x_J + A_{ET}\cdot x_T + A_{EL}\cdot x_L\right) \\
&= \frac{1}{\lambda}\left(x_B + x_J + x_T + x_L\right) \\
x_J &= \frac{1}{\lambda}\left(A_{JB}\cdot x_B + A_{JE}\cdot x_E + A_{JT}\cdot x_T + A_{JL}\cdot x_L\right) \\
&= \frac{1}{\lambda}\left(x_E + x_T\right) \\
x_T &= \frac{1}{\lambda}\left(A_{TB}\cdot x_B + A_{TE}\cdot x_E + A_{TJ}\cdot x_J + A_{TL}\cdot x_L\right) \\
&= \frac{1}{\lambda}\left(x_B + x_E + x_J\right) \\
x_L &= \frac{1}{\lambda}\left(A_{LB}\cdot x_B + A_{LE}\cdot x_E + A_{LJ}\cdot x_J + A_{LT}\cdot x_T\right) \\
&= \frac{1}{\lambda}\left(x_B + x_E\right).
\end{aligned}
$$

For the formula above, where λ is a constant and A is the adjacency matrix, we can express it in a matrixvector notation. In other words, we rewrite the formula above as the following.

$$x \quad = \quad \frac{1}{\lambda} \cdot A \cdot x.$$

Following the trust social network work example, we have the system of equations below to describe the centrality scores for all actors.

$$\begin{bmatrix} x_B \\ x_E \\ x_J \\ x_T \\ x_L \end{bmatrix} \quad = \quad \frac{1}{\lambda} \cdot \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_B \\ x_E \\ x_J \\ x_T \\ x_L \end{bmatrix}$$

$$\begin{aligned} \Rightarrow \quad x_B &= \frac{1}{\lambda} \left(x_E + x_T + x_L \right) \\ x_E &= \frac{1}{\lambda} \left(x_B + x_J + x_T + x_L \right) \\ x_J &= \frac{1}{\lambda} \left(x_E + x_T \right) \\ x_T &= \frac{1}{\lambda} \left(x_B + x_E + x_J \right) \\ x_L &= \frac{1}{\lambda} \left(x_B + x_E \right). \end{aligned}$$

This computation serves as an verification of the formula conversion. We see that they are expressing the same system of equations in a slightly different manner. Now, we can just move the constant λ to the left side of the formula and obtain the standard form for eigenvalues and eigenvectors, where λ is a eigenvalue and x is a eigenvector.

$$\lambda \cdot x \quad = \quad A \cdot x.$$

At this point, given any social network, we can derive its adjacency matrix and compute its eigenvalues and eigenvectors. In general, though, there will be many than one different eigenvalues for which an eigenvector solution exists. In the social network analysis, however, we have an additional requirement. That is all the entries in the eigenvector need to be positive. [12] This implies that only the greatest eigenvalue results in the desired centrality measure.

The conclusion above directly follows the Perron-Frobenius theorem, which asserts that a real square matrix with positive entries has a unique largest real eigenvalue and that the corresponding eigenvector has strictly positive components. [**?**] This largest eigenvalue is called the *principal eigenvalue* and the corresponding eigenvector is called the *principal eigenvector*.

At this point, following the standard procedure of computing eigenvalues and eigenvectors, we are ready to compute the eigenvector centrality for each actor with a given network graph. We will continue to use the trust social network as our example. We first compute the greatest eigenvalue. Note that for space considerations, the steps of computing the determinate of are omitted as well as the steps of solving the characteristic polynomial.

$$\begin{aligned}
A &= \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix} \\
det\left(\lambda I_5 - A\right) &= det\left(\begin{bmatrix} \lambda & -1 & 0 & -1 & -1 \\ 1 & \lambda & -1 & -1 & -1 \\ 0 & -1 & \lambda & 1 & 0 \\ -1 & -1 & 1 & \lambda & 0 \\ -1 & -1 & 0 & 0 & \lambda \end{bmatrix}\right) \\
&= \lambda^5 - 7\lambda^3 - 6\lambda^2 + 3\lambda + 2 \\
&\Rightarrow \lambda_1 \approx 2.94 \\
&\quad\; \lambda_2 \approx -1.62 \\
&\quad\; \lambda_3 \approx -1.47 \\
&\quad\; \lambda_4 \approx 0.62 \\
&\quad\; \lambda_5 \approx -0.46 \\
&\Rightarrow \lambda \approx 2.94.
\end{aligned}$$

Plugging in the greatest eigenvalue, we can compute its corresponding eigenvectors, which is the centrality score solution for this network graph. The details of the computation is again omitted for space considerations.

$$\begin{aligned}
\lambda \cdot x &= A \cdot x \\
2.94 \cdot \begin{bmatrix} x_B \\ x_E \\ x_J \\ x_T \\ x_L \end{bmatrix} &= \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_B \\ x_E \\ x_J \\ x_T \\ x_L \end{bmatrix}
\end{aligned}$$

$$\Rightarrow \begin{bmatrix} x_B \\ x_E \\ x_J \\ x_T \\ x_L \end{bmatrix} \approx \begin{bmatrix} 1.34 \\ 1.59 \\ 1.00 \\ 1.34 \\ 1.00 \end{bmatrix}$$

$$\begin{aligned} \Rightarrow \quad & C_B = 1.34,\; C_E = 1.59,\; C_J = 1.00 \\ & C_T = 1.34,\; C_L = 1.00. \end{aligned}$$

Note that the eigenvector solution is not unique. Basically, we've found a set of eigenvectors in the following form, where α is a constant but the ratios are the same, which is what we care about.

$$\alpha \cdot \begin{bmatrix} 1.34 \\ 1.59 \\ 1.00 \\ 1.34 \\ 1.00 \end{bmatrix}.$$

It is suggested that the eigenvector of the largest eigenvalue of an adjacency matrix could make a good network centrality measure. Unlike degree centrality, which weights every contact equally, the eigenvector weights contacts according to their own centrality. Eigenvector centrality can also be seen as a weighted sum of not only direct connections but indirect connections of every length. Thus it takes into account the entire pattern in the network. [8] For example, in a trust social network, if a person that is directly or indirectly trusted by other highly-trusted individuals, then this person is tend to be trustworthy.

4 Discussions

These social network measures provide sociologists powerful tools to study complex sociotechnical systems (STS). We live in a world that itself is a complex STS. Sociologists, policy and organizational analysts have identified different levels of complex systems consisting individuals, groups, organizations, and societies. [15] The interaction between people and technology in workplaces has long been recognized as a complex system as well. [16] To study STSs with SNA tools was, however, developed more recently. These tools help sociologists systematically analyze, represent, and model STSs. [15]

4.1 SNA Analysis in STS

STS in organizational development is an approach to complex organizational work design that recognizes the interaction between people and technology in workplaces. The term also refers to the interaction between society's complex infrastructures and human behavior. In this sense, society itself, and most of its substructures, are complex sociotechnical systems.[16] Socio-technical networks encode connections between people, connections between technical artifacts and connections between people and artifacts. [17]

There are a various of ways that our tools, such as classical social network techniques, e.g., centrality measures, can help us understand social networks. For example, with the centrality measures we can drill down into a social network and identify locations of critical individuals, groups, and technologies.

We can be given any network, such as a friendship network, or information-flow network. The nodes can be individuals, groups, computers, etc. As discussed previously, sub-graphs can be applied in the centrality measures just as well. The centrality measures can be used to locate critical nodes. These nodes could be critical as an employee in a company, a week point of an monetary-flow network for a profiting organization, etc.

4.2 Example Centrality Measure in STS

For example, with the betweenness centrality, we can identify that the removal of an individual or group would alter the network configuration significantly. The information flow among the network could be inhibited. The network could become less able to adapt. The performance of the network could be reduced.

Let us continue with the example above with a specific STS scenario. In the <u>*process improvement*</u> practice of an organization, a series of actions is taken to identify, analyze, and improve existing processes within this organization so that the entity can meet new goals and objectives. These actions often follow a specific methodology or strategy to create successful results. [19] Modifications of one, or a few, links in the process could significantly alter the performance of fitting the new goal. This link could be an individual, a sub-organization, or even a software component. It is likely that this component acts as a bridge between multiple otherwise disconnected groups of components. For instance, this link could be a software component that acts as a point of contact or an interface in an organization. This link would also be an individual with high cognitive load, and this individual is likely to be emergent leaders for a variety of reasons including they are most likely to tell others to do things and most likely to be in a position of power in terms of what and who they know.

With SNA tools providing us systematical analytical tools, sociologists are able to obtain a comprehensive picture of the network including both the human factors and the technology factors, and hence better apply their knowledge in STS practices.

5 Conclusions

We have studies both the theories and the social phenomenons of different centrality measures in the context of the social network analysis. In general we consider centrality measures the importance of an actor in a social network. Based on different goals of interacting with a social network, different actors might take the roles of being the most important actor. In other words, we have different types of centrality measures in varying application scenarios. Centrality measures are used widely in the STS studies. They help sociologists understand these complex systems.

The degree centrality measures the connectivity of actors in the network. For example, in a friendship social network, the most popular individual would have the most number of friends and hence would have the highest degree centrality. If we are interested in making friends with this social group, we would start with making friends with the individual with the highest degree centrality.

The betweenness centrality measures how important an actor is in guarding the pathways with in a social network. For example, in an information flow social network, the actor that is been used on the most number of geodesics has the highest betweenness centrality. If we are interested in jeopardizing the information flow in a social network, we would start with attacking the actor with the highest betweenness centrality.

The closeness centrality measures how close an actor is to the rest of social network. For example, in an information flow social network, the actor that has the smallest sum of pair-wise distances with the other actors is considered to be most important with respect to the closeness centrality. If we are interested in propagating a message as fast as possible throughout the network, we would first pass this message to the individual with the highest closeness centrality.

The eigenvector centrality measures the importance of actors differently from the three methods above. The eigenvector centrality is a weighted sum of not only direct connections but indirect connections of every length. Thus it is the only measure that takes into account the entire pattern in the network. For example, in a trust social network, the individual with highest average score of the trust values from its direct and indirect neighbors is considered to be most important with respect to the eigenvector centrality. If we are interested in having a large amount of money transactions, we would start with contacting the individual that is most trustworthy, and hence with the highest eigenvector centrality.

References

[1] Wasserman, S. and Faust, K., (1994) Social Network Analysis Methods and Applications. Cambridge University Press

[2] http://en.wikipedia.org/wiki/Social_network

[3] http://en.wikipedia.org/wiki/Dijkstra's_algorithm

[4] Freeman, L. C. (1979). Centrality in Social Networks: Conceptual clarification. Social Networks,1, 215-239.

[5] Freeman, L. C. (1977). A set of measures of centrality based on betweenness. Sociometry, 40: p. 35-41.

[6] Stephenson, K., & Zelen, M. (1989). Rethinking centrality. Social Networks, 11: 1-37.

[7] Ibarra, H., & Andrews, S. B. (1993). Power, social influence, and sense making: Effects of network centrality and proximity on employee perceptions. Administrative Science Quarterly, 38, 277-303.

[8] Bonacich, P. (1972). Factoring and weighting approaches to clique identification. Journal of Mathematical Sociology, 2, 113-120.

[9] Bavelas, A., (1948) A mathematical model for group structure. Human Organizations, 7: p. 16-30.

[10] Shimbel, A., (1953) Structural parameters of communication networks. Bulletin of Mathematical Biophysics, . 15: p. 501-507.

[11] http://en.wikipedia.org/wiki/Adjacency_matrix

[12] http://en.wikipedia.org/wiki/Centrality#Eigenvector_centrality

[13] Newman, M. E. J. (2008) "Mathematics of networks," in The New Palgrave Encyclopedia of Economics.

[14] http://en.wikipedia.org/wiki/Perron%E2%80%93Frobenius_theorem

[15] Carley, Kathleen M. (2000). "Information Security: The Human Perspective." CASOS Working Paper, Carnegie Mellon University

[16] http://en.wikipedia.org/wiki/Sociotechnical_systems#Sociotechnical_system

[17] C. Bird, N. Nagappan, P. Devanbu, H. Gall, and B. Murphy. (2009) Putting it all together: Using socio-technical networks to predict failures. In ISSRE 09, Bengaluru-Mysuru, India.

[18] Carley, Kathleen M. Ju-Sung Lee and David Krackhardt, (2001), Destabilizing Networks, Connections 24(3):31-34.

[19] http://en.wikipedia.org/wiki/Sociotechnical_systems#Process_improvement