INDIAN INSTITUTE OF MANAGEMENT CALCUTTA

WORKING PAPER SERIES

WPS No. 693/ January 2012

Sensitivity of Centrality and Centralization Measures to the Level of Decentralization in the Network Structure

by

Ram Babu Roy
Doctoral student, IIM Calcutta, Joka, Diamond Harbour Road, Kolkata 700104, India

&

Uttam Kumar Sarkar
Professor, Indian Institute of Management Calcutta, Joka, Kolkata 700104

Sensitivity of Centrality and Centralization Measures to the Level of Decentralization in the Network Structure

Ram Babu Roy
Cell Phone: +91-99038 83394 Email: rbroy07@iimcal.ac.in
and
Uttam Kumar Sarkar
Cell Phone: +91-94330 18562 Email: uttam@iimcal.ac.in

Indian Institute of Management Calcutta, Kolkata, India 700104

Abstract

Analytic expressions for node centrality and network centralization measures based on degree, betweenness, closeness, and eigenvector centrality are derived for network topologies with varying levels of decentralization. A solution to known difficulties related to normalization in computing graph centralization using eigenvector centrality is proposed. The correlations between these centrality measures are analyzed. Findings reveal that these measures are highly sensitive to the degree of decentralization in the network. The analytical expressions of these measures would be useful to get deeper insights into the sensitivity of these measures to the underlying structure and in approximating their values for resembling empirical network.

Keywords: Centrality, centralization, eigenvector, network topology, correlation, decentralized network

1. Introduction

Complex systems theory helps in understanding emergent and collective phenomena in economic, organizational, and social systems that may warrant preemptive actions to minimize possible negative impact due to its emergent behavior in the society. Different types of network structures are used to represent the interaction and information flow among the agents in a complex adaptive system such as a social network. Network models are being used in studying organizational structures and assessing their effectiveness in delivering the desired outcome. The organizational structure should follow strategy to align the team members with the strategy and often the strategy needs to be charted out considering the existing organizational structure. Inter-firm strategic relationships can potentially be shaped by industry events and managerial actions leading to either structure re-inforcing or structure-loosening impact on the relationship structure (Madhavan *et al.* 1998). Advances in information and communication technology (ICT) had a huge impact on the corporate strategy and organizational structure. Different types of managerial command and control structures are adopted to utilize knowledge and expertise in the organization and hence the communication network structure needs to be designed to suit the specific organizational goal. The organizational structures keep evolving to environmental pressures, changing corporate strategies, and to make country-specific adjustments in response to global economic developments and cultural differences between countries. Recent trend in globalization attempts to exploit decentralized human capital and knowledge base by collaborative works while retaining the administrative control by a central unit. In response to such opportunities, the organizations are tending to follow a more decentralized structure. A reliable measure of the level of decentralization in such a network would play a key role in assessing the impact of decentralization of work centers, governance, transportation and communication infrastructure on economic development, effectiveness of social policy implementations, supply chain management, project management, traffic management, and knowledge transfer in a social network (Bardhan 2002).

The social structure in different subgroups of large product development teams may be different due to regional influence or organizational culture. This often leads to intergroup coordination and communication problems that may cause hindrances to product delivery or meeting deadlines. Therefore, organizations have to do a trade-off between the level of centralization and cost. (Gupta & Govindarajan 1991) emphasizes that the knowledge flow patterns among different subsidiaries within the same MNC located in different locations can differ as reflected in the way a mix of formal and informal administrative control mechanisms are used to shape the actions of various subsidiaries and for each type of transactions, different subsidiaries may occupy central positions. (Inkpen & Tsang 2005) examines the effect of structural, cognitive and relational dimensions of social capital on the transfer of knowledge between the network members. The study proposes that different types of networks require different kinds of facilitating conditions to transfer knowledge among network members. The size of network is also important and one-size-fits-all analyses on network may not be suitable in tackling the complexity involved in the knowledge transfer processes. Centralization of public facilities may be good for better monitoring, control and maintenance but it leads to several problems ranging from accumulation of wastes, increased load on transportation infrastructure to healthcare systems eventually deteriorating the quality of life. Therefore, it is emphasized that decentralization of industrial clusters is required for sustainable development. Similarly, some government organizations such as police services often need a mix of autonomy and coordination and control. Such structure needs to have some central control and coordinating authority at the same time different units should have some level of operational autonomy to make them responsive and agile in dealing with social problems such as terrorism. Similarly, road network and other infrastructure development having role in national integration needs to have a holistic approach at the higher level as well as detailed local interconnections depending on the local needs.

Identification of most important actors in a social network is of immense interest in many practical situations and requires different types of centrality measures suitable for the purpose. Several centrality measures such as degree, closeness, betweenness, and eigenvector centrality have been proposed in the literature to quantify the importance of an actor in various processes within the network (Freeman 1979; Bonacich 1987; Borgatti & Everett 2006). There are individual node level centrality measures as well as at the entire network level known as 'network centralization' to characterize the extent to which a network is centralized around a particular node in the network. Betweenness centrality dictates the importance of a node in information flow, structural connectivity, and reduction of transportation cost as removal of a node with high betweenness centrality may lead to structural breakup of the graph. (Ercsey-

Ravasz & Toroczkai 2010) emphasizes the need for approximation in computation of betweenness centrality as the computation of betweenness centrality is very difficult and time consuming especially for large networks with millions to billions of nodes. Such large networks may be approximated with known number of hubs and leaf nodes connected at those hubs and a general formula would help to compute the approximate centrality values of resembling network of any size and structure. Obtaining analytic expression for node centrality and the network centralization applicable to any arbitrary network is infeasible because it depends on the network topology and there are too many topologies to consider. Despite this difficulty, the analytic expressions for the network centralization corresponding to some standard network topologies should be worth exploring. Little research has been done to compare the behavior of these centrality measures in different types of network topology. Also, there is no unified measure of the level of decentralization in the network. Results on how the network centralization changes as the network gets decentralized into multiple interconnected subgroups are not known.

The network centralization based on degree, betweenness, and closeness has been defined in (Freeman 1979), but the network centralization based on eigenvector centrality is not defined in the literature. There is no literature on systematic comparison of the four important centrality measures on some standard network configurations that may help understand the interrelationships between the centrality measures and relative centralization of various network configurations. Further, there is no benchmark network centralization value available in the literature for different network configurations except for the star configuration to measure and visualize the level of decentralization in the governance or organizational structure represented in the form of a network. There is a lack of closed form solutions for benchmark network configurations except the star configuration to quantify the level of decentralization in the network structures. Therefore the centralization values obtained using small sample network may not be scalable to large network and computation of centralization values is difficult for large networks as well as it becomes difficult to approximate real world network into simpler configurations. This may give erroneous research findings about the impact of decentralization in a large networked system on its performance and dynamics. This work attempts to fill this void in the literature by providing general formulae to compute the network centralization based on all the four centrality measures for some standard types of network configurations with comparable number of nodes as in the problem under investigation to compare and visualize the level of decentralization in the network structure. These formulae are scalable to any number of nodes in the network within some constraints put to achieve structural symmetry for ease in computation. The comparative analysis of these standard network configurations with any empirical network would help in getting insights into the level of decentralization in large empirical networks.

This paper provides theoretical contribution towards analysis of the sensitivity of network centrality measures on the level of decentralization in a network. We analytically derive the formulae for the network centralization for some of the standard topologies of decentralized network and compare their centralization with the most centralized star topology. We suggest possible solutions to extend the concept of network centralization based on eigenvector centrality. Our research attempts to answer some of the following questions. How does the network centralization value based on different types on centrality measures change with the level of decentralization in the network and with the size of the network? How sensitive are these measures to the introduction of a central node to otherwise decentralized network with completely interconnected hubs? At what level of decentralization do the hub nodes become more important than the central node for any particular centrality measure? How do the centrality values scale with the size of the network? Do these centralization measures correlate with each other as the level of decentralization in the network increases? The variance in node centrality corresponding to the four centrality measures for all the types of network has also been compared.

Section 2 looks into the related research in this area. Section 3 analyses the network centralization measures based on the level of decentralization in the network. The results and discussions on the sensitivity of the centralization measures are presented in Section 4. Section 5 draws concluding remarks.

2. Related Research

Centrality and network centralization measures have been used extensively in social network analysis literature to study the influence of important actors and analyze network structures arising in several practical scenarios. The basic idea behind the social network analysis is that the observed network structure of actors is the outcome of actions of various actors and social institutions in the society. Opinion and behavior within a group are more homogeneous compared to that between different groups. People connected across group may have alternating ways of behaving and such people attain an advantageous brokerage position in establishing links across structural holes between distant groups in the network. People with large number of contacts (degree) often play a role of opinion leaders. Opinion leaders help in propagating information across the social boundaries between groups in a network (Burt 1999). Such people can play the role of broker in bridging different groups across the structural holes and bring novel ideas into the groups thus creating a social capital (Burt 2004). Social capital of an individual refers to the benefits he can derive through his social network. The focus of social capital research is on the features of the network that contribute to the individual, whereas with key player research the emphasis is on which individuals are important for the network (Borgatti 2006). (Friedkin 1993) has examined the relationship between the interpersonal power and interpersonal influence in resolving issues related to organizations. His findings suggests that the social structure that gives an individual social power have significant positive effect on the frequency of issue-related communications among the members of the organization that in turn have substantial effect on interpersonal influence. Two competing views on creation of social capital in a network has been presented in (Gargiulo & Benassi 2000; Burt 2001). A cohesive network provides a safety of cooperation while the structural holes provide flexibility in developing new ideas that may lead to innovations in the network. The network measures of social capital has been discussed in (Borgatti *et al.* 1998). It has been shown that the scale free network provides the fastest growth and diffusion of newly innovated knowledge (Lin & Li 2010).

(Renneboog & Zhao 2011) investigated the role of director networks on the top manager's compensation and the pay-setting process in the UK. Literature shows that both formal and informal professional and social network affects the monitoring of economic and financial activities and corporate decision making. CEOs accumulate larger social capital by setting right kind of network amongst top management and directors as different types of network enables different kinds of managerial functions and hence drives the motive for forming right kind of networks in an organization. Indirect network are built for reasons of information gathering and direct networks helps in accumulating more managerial influence. Therefore, a CEO well connected with board directors often derives significantly higher compensation. Assessing the centrality of senior managers is important for any firm in corporate decision making and pay-setting process in order to retain their competitive edge in the market. Network modeling of real world interconnected systems from diverse area is gaining attention in recent years (Uzzi 1997; Doerfel 1998; Newman 2003; Guckenheimer & Ottino 2008; Helbing 2008; Kolaczyk 2009). The topological properties of the network and identification of clusters play crucial role in understanding the internal structure and dynamics of a network (Newman 2008; Mishra *et al.* 2009). Random matrix theory and spectral methods are used to study correlation based networks and identification of clusters in a network (Edelman & Rao 2005; Kim & Jeong 2005; Newman 2006; Heimo *et al.* 2008). The concept of social network analysis has been used for analyzing interdependence structure between stock indices (Roy & Sarkar 2011a) and between the stocks (Roy & Sarkar 2011b) in the global stock market. Trading among actors in stock option market show a distinct social structural patterns that affects the direction and magnitude of price volatility (Baker 1984). Bounded rationality and opportunistic behavior of economic actors gives rise to restrictive micro-networks. Such restrictive micro-networks may create differentiated macro-networks in large markets leading to information asymmetry due to inefficient communication among actors.

The influence of social network on spread and cessation of smoking in a group of socially interconnected people have been investigated by (Christakis & Fowler 2008). They find that generally the whole group of smokers quit smoking together emphasizing the influence of social network on individual behavior to conform with the majority behavior in the group leading to the observed collective behavior of the entire group. Data available from online environment and information and communication systems has enabled vast amount of research in the broader domain of network science and in particular Social network analysis (Rosen *et al.* 2010). In a social network, position of actors defines the role of the actor therefore similarity among relations can be used to determine the group of actors into different classes known as equivalence classes of actors. There are several equivalence classes such as structural, isomorphic, regular etc defined based on the types of ties between the actors. The roles of individuals can in inferred from the pattern of ties that emerge due to the types of role played by the actor in the social network (Wasserman & Faust 1994). Applying modality (different classes of node types) and equivalence concepts may facilitate understanding the social processes and patterns in the complex and voluminous data that is generated through social interactions of various agents (Hanneman & Shelton 2011).

(Friedkin 1991) provides theoretical foundations for three complementary centrality measures based on elementary process model of social influence that explains why ties are formed or are broken in a social network. He classifies the three centrality measures arising due to three complementary effects namely total effect, immediate effect, and mediative effect. (Borgatti 2005) pointed out that the most commonly used centrality measures make implicit assumption about the flow processes in the network and hence the measures derived using a different assumption applied to a flow with different characteristics may lead to wrong interpretation and answers. Therefore the centrality measure used to present a particular phenomenon should match with the appropriate kinds of flow processes. The centrality values for large networks may be estimated by using a sub-sample of the network. But such estimate is likely to suffer from errors due to incomplete set of data in estimating these values. (Costenbader & Valente 2003) have investigated the impact of different levels of sampling of data from the population on the stability of centrality measures and shown that some centrality measures are more stable than others under different levels of sampling. The research highlights the impact of missing data on approximating the values of the centrality for the entire network using a smaller sample. We need a large set of data to estimate the centrality values in order to reduce the impact of missing data that may pose practical problems in many situations. Moreover, research gaps exist in analytical formulations of centrality measures in network topologies of varying degrees of decentralization levels.

3. Impact of Network Configuration and Decentralization on Centrality Measures

We have derived analytical expressions for all the four centrality measures for some standard network configurations with varying levels of decentralization. Such configurations arise in several practical situations (Wasserman & Faust 1994; Rivkin & Siggelkow 2007). In order to compare the sensitivity of the centrality measures, we have computed the node level centrality as well as the entire network level centralization measures based on all the four centrality measures for these network configurations. For those configurations, having complicated analytic expressions for centrality and centralization, we have computed their values numerically using MATLAB. We have also compared the variance of node centrality for these network configurations and the correlation between various centrality measures for some selected network configurations. We now present the derivations of the analytical expressions.

3.1 Network centralization of some network configurations

(Freeman 1979) has defined the generalized measure of graph centralization based on the differences in point centralities in the network and derived the general formula for graph centralization based on degree, betweenness and closeness centrality. The node centrality values are normalized by dividing them by the highest possible centrality value among all the nodes in the network. The general form of network centralization is given by

$$Network\ \ Centralization = \frac{\sum_{x \in U} (C_A^{\max} - C_A(x))}{Max \sum_{x \in U} (C_A^{\max} - C_A(x))} \quad \textbf{(1)}$$

Where $C_A^{\max}$ is the maximum value of normalized node centrality $C_A(x)$ among all the nodes based on a particular type of centrality measure, say A. The denominator denotes the maximum possible value of the numerator among all possible network configurations. This division is required to normalize the centralization value with respect to the most centralized graph. It turns out that the maximum value of the denominator is achieved for star topology that is intuitively the most centralized configuration in case of degree, closeness and betweenness centrality. We observe some counterintuitive behavior of the denominator in case of eigenvector centrality as discussed in the following section. Variance of node centrality in the network is another measure of network centralization (Wasserman & Faust 1994). Intuitively, in a strongly centralized network the variance of node centrality from the centrality of the most central node should be the highest. The closed form expression for this measure is complicated, but the level of centralization using this metric can easily be computed numerically using the expressions derived in this paper for the normalized node centrality value for different types of equivalent nodes in the network.

3.2 Network centralization based on eigenvector centrality of nodes

Eigenvector centrality is useful for analyzing the relative importance of nodes in a network such as social status of actors in a social network and detecting changes in the connectivity patterns in the neural architecture of human brain (Bonacich & Lloyd 2004; Lohmann *et al.* 2010). (Bonacich & Lloyd 2004) emphasizes that a person's social status reduces with a positive link with a notorious person and increases with a negative link with a notorious person. The network centralization based on eigenvector centrality for star configuration can be shown to be equal to $(N-1)-\sqrt{(N-1)}$. The theoretical maximum value for the expression in the denominator can be shown to be equal to N-2 and it is achieved when the network constitutes of only two connected nodes (a dyad) and N-2 independent nodes. The eigenvector centrality of a node in a network is given by the respective component of the eigenvector corresponding to the largest eigenvalue (Bonacich 1987). If we define the eigenvector centrality of a node in a network to be given by the square of respective component, the network centralization based on eigenvector centrality for star configuration can be shown to be equal to $(N-2)$. The later approach of defining the eigenvector centrality is more suitable for computing network centralization as the most central star topology achieves the highest value using this approach. The first approach does not give maximum value for star configuration used as a normalization constant in the denominator of Equation 1. For example, let's consider a network topology constructed using np+1 nodes in which one central node connected to n hub nodes, and each of the hub nodes is connected to (p-1) leaf nodes as shown in Figure 1(L). The sum of the deviation of centrality value from the maximum value in this case becomes

$$\left(n-\sqrt{n+p-1}\right)+(n-1)(p-1)=np-\sqrt{n+p-1}-(p-1) \text{ for } n\geq\lambda .$$

For p =2, and n>5 this expression becomes greater than the sum of deviation for star topology with same number of nodes (i.e. N = np+1) which is equal to $np-\sqrt{np}$.

If we define the eigenvector centrality of a node in a network to be given by the square of respective component, the sum of the deviations of the node centralities from the maximum centrality value based on eigenvector centrality for a network topology in Figure 1(L) becomes

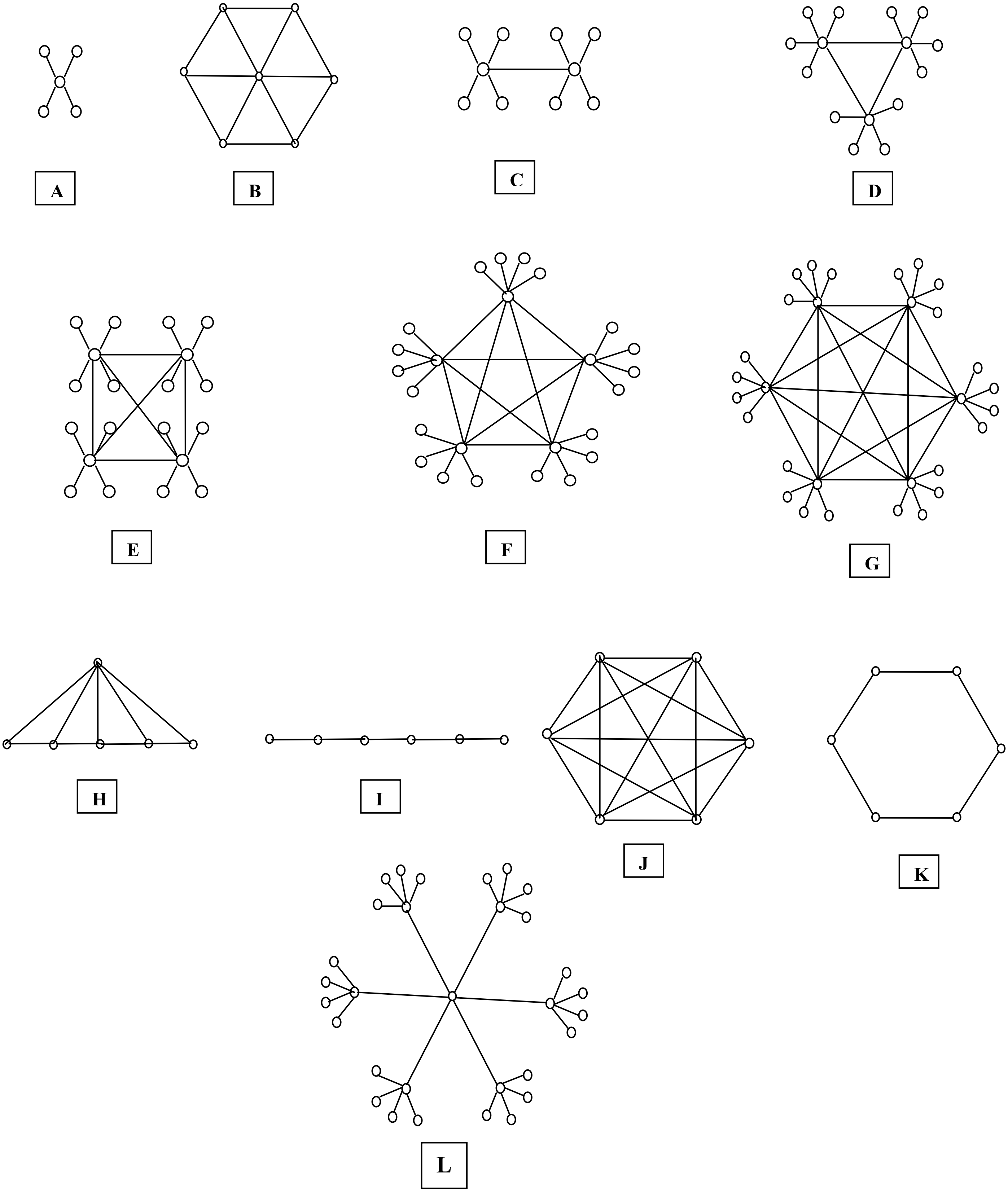

Figure 1: Different types of network topology

$(np-1)-\frac{2(p-1)}{n}$ for $n \geq \lambda$ and $(np-2n+1)$ for $n < \lambda$. This is always less than or equal to the value $(N-2)=(np-1)$ corresponding to the star configuration. Therefore, there is a need for slight modification in the definition of the eigenvector centrality of nodes in order to make the computation of network centralization based on eigenvector centrality using the formula given in equation 1 by (Freeman 1979) meaningful.

We present the computation of the closed form expression based on the centrality measures for various types of networks. The schematic diagram for several network topologies are shown in Figure 1. Apart from the graphs shown in Figure 1, the decentralized network topology with 10, 12, 15, 20 and 30 centers have also been analyzed and the results have been compared.

3.3 Star network: We consider a star network with total N nodes where one node is central and rest N-1 leaf nodes. The computation of network centralization based on degree, betweenness and closeness centrality is illustrated in (Freeman 1979). We have derived the closed form expressions for the variance of normalized centrality of nodes of the network and presented them along with the formulae for the network centralization in the table below.

Degree centrality: The degree of central node = (N-1) and degree of leaf node = 1 each. Normalized degree centrality of nodes = 1 for central and 1/(N-1) for rest others. The difference between the maximum normalized centrality and the actual normalized centrality of nodes = 0 for central and (N-2)/(N-1) for all other. Sum of these differences for all nodes gives the centralization value equal to (N-2).

Node Type	**Degree**	**Normalized Centrality**	**Max Norm Centrality - Norm Centrality**	**(Norm Centrality – Centrality)^2**
Central node	(N-1)	1	0	$\left(1-\frac{2}{N}\right)^2$
(N-1) Leaf nodes	1	$\frac{1}{(N-1)}$	$\frac{(N-2)}{(N-1)}$	$\left(\frac{1}{(N-1)}-\frac{2}{N}\right)^2$
Sum		2	(N-2)	$\frac{(N-2)^2}{N(N-1)}$
Variance of Normalized Centrality				$\frac{(N-2)^2}{N^2(N-1)}$

Betweenness centrality: The central node falls on $\frac{(N-1)(N-2)}{2}$ geodesics (shortest paths) connecting any pair of nodes. The number of geodesics passing through leaf nodes is zero.

Node Type	No. of Geodesics	Normalized Centrality	Max Norm Centrality - Norm Centrality	(Norm Centrality Centrality)^2
Central node	$\frac{(N-1)(N-2)}{2}$	1	0	$\left(1-\frac{1}{N}\right)^2$
(N-1) Leaf nodes	0	0	1	$\left(-\frac{1}{N}\right)^2$
Sum		1	(N-1)	$\frac{(N-1)}{N}$
Variance of Normalized Centrality				$\frac{(N-1)}{N^2}$

Closeness centrality: The central node is at distance 1 from all N-1 nodes and the leaf odes are at distance 1 from the central and 2 from other N-2 nodes. The closeness of a node is defined as (N-1)/sum of distance between the node and all other nodes.

Node Type	Closeness Centrality	Normalized Centrality	Max Norm Centrality - Norm Centrality	(Norm Centrality – Mean Centrality)^2
Central node	$\frac{(N-1)}{(N-1)}$	1	0	$\left(1-\frac{(N^2-2)}{N(2N-3)}\right)^2$
(N-1) Leaf nodes	$\frac{(N-1)}{(2N-3)}$	$\frac{(N-1)}{(2N-3)}$	$\frac{(N-2)}{(2N-3)}$	$\left(\frac{(N-1)}{(2N-3)}-\frac{(N^2-2)}{N(2N-3)}\right)^2$
Sum		$\frac{(N^2-2)}{(2N-3)}$	$\frac{(N-1)(N-2)}{(2N-3)}$	$\frac{(N-1)(N^2-4N+12)}{N(2N-3)^2}$
Variance of Normalized Centrality				$\frac{(N-1)(N^2-4N+12)}{N^2(2N-3)^2}$

Eigenvector centrality: The largest eigenvalue for star network adjacency matrix = $\sqrt{N-1}$. By symmetry the eigenvector centrality of all leaf nodes (say x_2) should be equal and the centrality of the central node (x_1) should be greater than x_2, the centrality of leaf nodes. The eigenvalue equation for a star network with 6 nodes is shown below as an example to illustrate the computation of the eigenvector centrality.

$$\begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_2 \\ x_2 \\ x_2 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ x_2 \\ x_2 \\ x_2 \\ x_2 \end{bmatrix}$$

The normalization condition of the eigenvector puts the following constraints on the values of x_1 and x_2.

$$x_1^2 + (N-1)x_2^2 = 1$$

The eigenvalue equation puts the following constraints on the values of x_1 and x_2.

$$(N-1)x_2 = \lambda x_1 \qquad \text{and} \qquad x_1 = \lambda x_2$$

From these three relations the largest eigenvalue (λ) and the two components x_1 and x_2 can be computed as given below.

$$\lambda = \sqrt{N-1}, \quad x_1 = \frac{1}{\sqrt{2}}, \quad x_2 = \frac{1}{\sqrt{2(N-1)}}$$

Node Type	**Eigenvector Centrality**	**Normalized Centrality**	**Max Norm Centrality – Actual Norm Centrality**	**(Norm centrality – Centrality)^2**
Central node	$\frac{1}{\sqrt{2}}$	1	0	$\left(\frac{(N-1)-\sqrt{(N-1)}}{N}\right)^2$
(N-1) Leaf nodes	$\frac{1}{\sqrt{2(N-1)}}$	$\frac{1}{\sqrt{N-1}}$	$1-\frac{1}{\sqrt{N-1}}$	$\frac{(N-1)(1-\sqrt{(N-1)})^2}{N^2(N-1)}$
Sum		$1+\sqrt{N-1}$	$(N-1)-\sqrt{(N-1)}$	$\frac{(N-2\sqrt{(N-1)})}{N}$
Variance of Normalized Centrality				$\frac{(N-2\sqrt{(N-1)})}{N^2}$

3.4 Decentralized Network without any central node

In order to obtain a generalized network centralization formula for symmetrically decentralized topologies, we consider a network with total N nodes that are equally distributed at n subgraphs with each subgraph having star like configuration and the hubs of all these star subgraphs connected with one another symmetrically.

Degree centrality: Degree of each hubs in the subgraph = $\frac{N}{n}+(n-2)$ and the degree of each leaf node = 1. The general closed form expression for variance of degree centrality is complicated, therefore we omit them. However, the variance of node centrality for any topology can easily be computed numerically using the normalized centrality values given in the Table.

Node Type	Degree	Normalized Centrality	Max Norm Centrality – Actual Norm Centrality
Hub nodes (n)	$\frac{N}{n}+(n-2)$	1	0
(N-n) Leaf nodes	1	$\frac{n}{N+n(n-2)}$	$\frac{N+n(n-3)}{N+n(n-2)}$
Sum			$\frac{(N-n)(N+n(n-3))}{N+n(n-2)}$
Network Centralization			$\frac{(N-n)(N+n(n-3))}{(N+n(n-2))(N-2)}$

Betweenness centrality: The leaf nodes do not lie on any of the geodesics between any other two nodes. Therefore, their betweenness is zero. The hub nodes fall on ${}^{(N-n)/n}C_2$ geodesics between the leaf nodes directly attached to it and another (${}^{(N-n)/n}C_1$* ${}^{N/n}C_1$)(n-1) geodesics between the (n-1) other hubs and the leaf nodes attached to them. Therefore the total number of geodesics on which a hub falls is given by

$$^{\frac{(N-n)}{n}}C_2+(n-1)\,^{\frac{(N-n)}{n}}C_1\,^{\frac{N}{n}}C_1=\frac{(N-n)(2n(N-1)-N)}{2n^2}$$

Node Type	No. of Geodesics	Normalized Centrality	Max Norm Centrality - Norm Centrality	(Norm Centrality – Mean Centrality)^2
Hub nodes (n)	$\frac{(N-n)(2n(N-1)-N)}{2n^2}$	1	0	$\left(1-\frac{n}{N}\right)^2$
(N-n) Leaf node	0	0	1	$\left(-\frac{n}{N}\right)^2$
Sum		n	N-n	$\frac{(N-n)(N-n+n^2)}{N^2}$
Network Centra			$\frac{(N-n)}{(N-1)}$	
Variance of Norm Centrality				$\frac{(N-n)(N-n+n^2)}{N^3}$

Closeness centrality: Each of the hub nodes is at distance 1 from all (N-1)/n leaf nodes directly attached to them and all other (n-1) hub nodes. The hubs are at distance 2 from other (N-n)(n-1)/n leaf nodes attached to the other (n-1) hub nodes. The leaf nodes are at a distance 1 from the hub it is attached to and at distance 2 from (N-2n)/n leaf nodes attached to the same hub as well as at distance 2 from (n-1) other hubs. The distance between a leaf node and the leaf nodes attached to other hubs are 3 units. The closeness centrality of a node is defined as (N-1)/sum of distance between the node and all other nodes. The general closed form expression for variance of need centrality is complicated but can be computed numerically using the normalized centrality values given in the table.

Node Type	Closeness Centrality	Normalized Centrality	Max Norm Centrality – Actual Norm Centrality
Hub nodes (n)	$\frac{(N-1)n}{N(2n-1)-n^2}$	1	0
(N-n) Leaf nodes	$\frac{(N-1)n}{N(3n-1)-n(n+2)}$	$\frac{N(2n-1)-n^2}{N(3n-1)-n(n+2)}$	$\frac{n(N-2)}{N(3n-1)-n(n+2)}$
Sum		$\frac{N^2-2}{2N-3}$	$\frac{n(N-2)(N-n)}{N(3n-1)-n(n+2)}$
Network Closeness Centralization			$\frac{n(2N-3)(N-n)}{(N-1)(N(3n-1)-n(n+2))}$

Eigenvector centrality: The largest eigenvector of the adjacency matrix for a network with total N nodes decentralized into n symmetrical hubs will have n equal components with value x_1 corresponding to the n central nodes of the n hubs and (N-n) equal components with value x_2 corresponding to the (N-n) leaf nodes. The normalization condition of the eigenvector puts the following constraints on the values of x_1 and x_2.

$$nx_1^2 + (N-n)x_2^2 = 1$$

The eigenvalue equation for network with 6 nodes decentralized into 2 hubs is shown below as an example to illustrate the computation of the eigenvector centrality.

$$\begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_2 \\ x_2 \\ x_2 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ x_2 \\ x_2 \\ x_2 \\ x_2 \end{bmatrix}$$

The eigenvalue equation puts the following constraints on the values of x_1 and x_2.

$$(n-1)x_1 + \left(\frac{N}{n} - 1\right)x_2 = \lambda x_1, \qquad x_1 = \lambda x_2$$

From these three relations the largest eigenvalue (λ) and the two components x_1 and x_2 can be computed as given below.

$$\lambda = \frac{(n-1) + \sqrt{(n-1)^2 + \frac{4(N-n)}{n}}}{2}, \quad x_1 = \frac{\lambda}{\sqrt{n\lambda^2 + (N-n)}}, \quad x_2 = \frac{1}{\sqrt{n\lambda^2 + (N-n)}}$$

Node Type	Eigenvector Centrality	Normalized Centrality	Max Norm Centrality – Actual Norm Centrality
Hub nodes (n)	$\frac{\lambda}{\sqrt{n\lambda^2 + (N-n)}}$	1	0
(N-n) Leaf nodes	$\frac{1}{\sqrt{n\lambda^2 + (N-n)}}$	$\frac{1}{\lambda}$	$1-\frac{1}{\lambda}$
Sum		$n+\frac{N-n}{\lambda}$	$\frac{(N-n)(\lambda-1)}{\lambda}$
Network Eigenvector Centralization			$\frac{(N-n)(\lambda-1)}{\lambda((N-1)-\sqrt{N-1})}$

If the eigenvector centrality of a node is defined as equal to the square of the component of the principal eigenvector rather the value of the component, the value of network centralization can be shown to be given by the following equation. Both the equations give the value of network centralization for the star network as 1.

$$C_{EV} = \frac{(N-n)(\lambda^2-1)}{\lambda^2(N-2)}$$

3.5 Decentralized Network with a central node and n hubs

Generalizing the network centralization for star topology to a decentralized topology with one central node connected to n hub nodes where each hub node have a star like configuration with p-1 leaf nodes (i.e. excluding the central and the hub nodes) connected to it. This network will have total np+1 nodes. We have considered both the cases viz. the number of hubs n is less than or equal to p and the number of hubs n is greater than p. The general closed form expression for variance of centrality of nodes is complicated, therefore we omit them. However, the variance of centrality for any topology can easily be computed numerically using the normalized centrality values given in the table.

Degree Centrality: The central node has a degree n and the hub nodes have a degree p each and the leaf node has a degree 1 each.

Node Type	Degree	Normalized Centrality (if n ≤ p)	Max Norm Centrality – Actual Norm Centrality
Central Node(1)	n	$\frac{n}{p}$	$\frac{p-n}{p}$
Hub nodes (n)	p	1	0
n(p-1) Leaf nodes	1	$\frac{1}{p}$	$\frac{p-1}{p}$
Sum		2	$np-2n+1$
Network Centralization			$\frac{n(p-2)+1}{(np-1)}$

For case n>p

Node Type	Degree	Normalized Centrality (if n>p)	Max Norm Centrality – Actual Norm Centrality
Central Node(1)	n	1	0
Hub nodes (n)	p	$\frac{p}{n}$	$\frac{n-p}{n}$
n(p-1) Leaf nodes	1	$\frac{1}{n}$	$\frac{n-1}{n}$
Sum		2	$pn-2p+1$
Network Centralization			$\frac{p(n-2)+1}{(np-1)}$

Betweenness centrality: The central node and the hub nodes fall on the geodesics between various other nodes. The leaf nodes do not lie on any of the geodesics between any other two nodes. Therefore, their betweenness is zero. The central node falls on ${}^{p}C_{1}*{}^{p(n-1)}C_{1}$ geodesics between the p nodes at one hub and the p(n-1) nodes at other hubs. The hub nodes fall on ${}^{(p-1)}C_{1}*({}^{p(n-1)}C_{1}+1)$ geodesics between the (p-1) leaf nodes at one hub and the p(n-1) nodes at other hubs and one central node. The following derivation is valid for n>1, for n =1 the network becomes a star network and hence the value of network centrality becomes 1.

Node Type	No. of Geodesics	Normalized Centrality (n>1)	Max Norm Centrality – Actual Norm Centrality
Central node (1)	$p^2(n-1)$	1	0
Hub nodes (n)	$p^2(n-1)-(np-2p+1)$	$1-\frac{(np-2p+1)}{p^2(n-1)}$	$\frac{(np-2p+1)}{p^2(n-1)}$
n(p-1) Leaf nodes	0	0	1
Sum		n	$\frac{n(np-2p+1)}{p^2(n-1)}+n(p-1)$
Network Centralization	$\frac{(np-2p+1)}{p^3(n-1)}+\frac{(p-1)}{p}$, where n>1, Centralization = 1, for n = 1 (Star Configuration)		

Closeness centrality: The central node is at distance 1 from all n hub nodes at a distance of 2 units from all the n(p-1) leaf nodes directly attached to the hub nodes. A hub node is at distance 1 unit from the central node and all (p-1) leaf nodes directly attached to the hub node. The hub is at a distance of 2 units from all other (n-1) hub nodes and at a distance of 3 units from (n-1)(p-1) leaf nodes attached to these hub nodes. The leaf nodes are at a distance 1 from the hub it is attached to and at distance 2 from (p-2) leaf nodes attached to the same hub as well as at distance 2 from the central node. The distance between a leaf node and other (n-1) hubs and the (n-1)(p-1) leaf nodes attached to the other hubs are 3 units and 4 units respectively. The closeness centrality of a node is defined as (N-1)/sum of distance between the node and all other nodes where N is the total number of nodes in the network.

Total distance from central node = $n + 2n(p-1) = n(2p-1)$
Total distance from hub node = $p + 2(n-1) + 3(p-1)(n-1) = n(2p-1) + (n-2)p + 1$
Total distance from leaf node = $1 + 2(p-1) + 3(n-1) + 4(n-1)(p-1) = 2p(2n-1) - n$

Node Type	**Closeness Centrality**	**Normalized Centrality**	**Max Norm Centrality - Norm Centrality**
Central node (1)	$\frac{p}{2p-1}$	1	0
Hub nodes (n)	$\frac{np}{n(2p-1)+(n-2)p+1}$	$\frac{n(2p-1)}{n(2p-1)+(n-2)p+1}$	$\frac{(n-2)p+1}{n(2p-1)+(n-2)p+1}$
n(p-1) Leaf nodes	$\frac{np}{2p(2n-1)-n}$	$\frac{n(2p-1)}{2p(2n-1)-n}$	$\frac{2(n-1)p}{2p(2n-1)-n}$
Sum	$\frac{np(n-2)+n}{n(2p-1)+(n-2)p+1} + \frac{2np(n-1)(p-1)}{2p(2n-1)-n}$		
Network Closeness Centralization	$\left(\frac{np(n-2)+n}{n(2p-1)+(n-2)p+1} + \frac{2np(n-1)(p-1)}{2p(2n-1)-n}\right) \times \frac{2np-1}{np(np-1)}$, where n>1, Centralization = 1, for n = 1 (Star Configuration)		

Eigenvector centrality: The eigenvector corresponding to the largest eigenvalue of the adjacency matrix for a network with total N = np+1 nodes decentralized into one central node connected to n symmetrical hubs with n(p-1) leaf nodes will have one component with value x1 corresponding to the central node, n equal components with value x2 corresponding to the n hub nodes, and n(p-1) equal components with value x3 corresponding to the n(p-1) equivalent leaf nodes. The eigenvalue equation for network with total 9 nodes with one central node connected to 2 hubs with 3 leaf nodes as each hub (i.e. n = 2 and p =4) is shown below as an example to illustrate the computation of the eigenvector centrality.

$$\begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_2 \\ x_3 \\ x_3 \\ x_3 \\ x_3 \\ x_3 \\ x_3 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ x_2 \\ x_3 \\ x_3 \\ x_3 \\ x_3 \\ x_3 \\ x_3 \end{bmatrix}$$

The normalization condition of the eigenvector puts the following constraints on the values of x1, x2 and x3.

$$x_1^2 + nx_2^2 + n(p-1)x_3^2 = 1$$

The eigenvalue equation puts the following constraints on the values of x1, x2 and x3.

$nx_2 = \lambda x_1$, $\quad x_1 + (p-1)x_3 = \lambda x_2$, $\quad x_2 = \lambda x_3$

From these four relations the largest eigenvalue (λ) and the three components x1, x2 and x3 can be computed as given below.

$$\lambda = \sqrt{n + p - 1}$$

$$x_1 = \frac{n}{\sqrt{n^2 + n\lambda^2 + n(p-1)}}, \quad x_2 = \frac{\lambda}{\sqrt{n^2 + n\lambda^2 + n(p-1)}}, \quad x_3 = \frac{1}{\sqrt{n^2 + n\lambda^2 + n(p-1)}}$$

Node Type	**Eigenvector Centrality**	**Normalized Centrality** (if $\lambda \le n$ $(p-1) \le n(n-1)$)	**Max Norm centrality – Actual Norm Centrality**
Central node (1)	$\frac{n}{\sqrt{n^2 + n\lambda^2 + n(p-1)}}$	1	0
Hub nodes (n)	$\frac{\lambda}{\sqrt{n^2 + n\lambda^2 + n(p-1)}}$	$\frac{\lambda}{n}$	$\frac{n-\lambda}{n}$
n(p-1) Leaf nodes	$\frac{1}{\sqrt{n^2 + n\lambda^2 + n(p-1)}}$	$\frac{1}{n}$	$\frac{n-1}{n}$
Sum			$(n-\lambda) + (n-1)(p-1)$
Network Eigenvector Centralization			$\frac{(n-\lambda) + (n-1)(p-1)}{(np - \sqrt{np})}$

For $\lambda > n$

Node Type	Eigenvector Centrality	Normalized Centrality (if $\lambda > n$ $(p-1) > n(n-1)$)	Max Norm centrality – Actual Norm Centrality
Central node (1)	$\frac{n}{\sqrt{n^2 + n\lambda^2 + n(p-1)}}$	$\frac{n}{\lambda}$	$\frac{\lambda - n}{\lambda}$
Hub nodes (n)	$\frac{\lambda}{\sqrt{n^2 + n\lambda^2 + n(p-1)}}$	1	0
n(p-1) Leaf nodes	$\frac{1}{\sqrt{n^2 + n\lambda^2 + n(p-1)}}$	$\frac{1}{\lambda}$	$\frac{\lambda - 1}{\lambda}$
Sum			$\frac{(\lambda - n) + n(p-1)(\lambda - 1)}{\lambda}$
Network Eigenvector Centralization			$\frac{(\lambda - n) + n(p-1)(\lambda - 1)}{\lambda(np - \sqrt{np})}$

If the eigenvector centrality of a node is defined as the square of the corresponding component of the principal eigenvector, then the network centralization can be derived as given below.

Network Centralization (Eigenvector)	(if $\lambda \le n$ i.e. $(p-1) \le n(n-1)$)	(if $\lambda > n$ i.e. $(p-1) > n(n-1)$)
	$\frac{(n^2 p - n) + 2(p-1)}{n(np-1)}$	$\frac{(np - 2n + 1)}{(np-1)}$

3.6 Wheel Graph: A wheel graph of size N has one central node directly connected to all the (n-1) leaf nodes and the leaf nodes are connected to their neighbors forming a cycle of size (N-1).

Degree centrality: The central node has a degree N-1 and all the leaf nodes have a degree 3 corresponding to connections with two adjacent nodes and one central node.

Node Type	Degree	Normalized Centrality	Max Norm Centrality – Actual Norm Centrality
Hub node	$N-1$	1	0
(N-1) Leaf nodes	3	$\frac{3}{N-1}$	$\frac{N-4}{N-1}$
Sum		4	$N-4$
Network Centralization			$\frac{N-4}{N-2}$

Closeness centrality: The central node is at a distance of one unit from all other nodes. The leaf nodes are at a distance of one unit from the central and two adjacent leaf nodes and a distance of two units from other remaining nodes.

Node Type	Closeness Centrality	Normalized Centrality	Max Norm Centrality – Actual Norm Centrality
Hub node	$N-1$	1	0
(N-1) Leaf nodes	$3+2(N-4)$	$\frac{N-1}{2N-5}$	$\frac{N-4}{2N-5}$
Sum		$\frac{N^2-4}{2N-5}$	$\frac{(N-1)(N-4)}{2N-5}$
Network Centralization			$\frac{(2N-3)(N-4)}{(2N-5)(N-2)}$

Betweenness centrality: Out of the $\frac{(N-1)(N-2)}{2}-(N-1)$ possible geodesics passing through the central node, (N-1) geodesics connecting pair of leaf nodes (one each on either side) adjacent to every leaf node have the same path length equal to two, one passing through the central node and other passing through the intervening leaf node. Therefore, the effective number of geodesics through the central and node is given by

$$\frac{(N-1)(N-2)}{2}-(N-1)-\frac{1}{2}.(N-1)=\frac{(N-2)(N-5)}{2}$$

and that through each of the leaf nodes is given by $\frac{1}{2}$.

Node Type	No. of Geodesics	Normalized Centrality	Max Norm Centrality – Actual Norm Centrality
Hub node	$\frac{(N-2)(N-5)}{2}$	1	0
(N-1) Leaf nodes	$\frac{1}{2}$	$\frac{1}{(N-2)(N-5)}$	$\frac{N^2-7N+9}{(N-2)(N-5)}$
Sum		$\frac{(N-3)^2}{(N-2)(N-5)}$	$\frac{(N-1)(N^2-7N+9)}{(N-2)(N-5)}$
Network Centralization			$\frac{N^2-7N+9}{(N-2)(N-5)}$

Eigenvector centrality: By symmetry the central node will have highest eigenvector component and all other leaf nodes would have equal eigenvector components. The eigenvalue λ and the two components x_1 and x_2 of eigenvector representing the centrality of the central and leaf nodes of the wheel graph can be shown to be given below.

The eigenvalue equation for wheel network with 6 nodes is shown below as an example to illustrate the computation of the eigenvector centrality.

$$\begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_2 \\ x_2 \\ x_2 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ x_2 \\ x_2 \\ x_2 \\ x_2 \end{bmatrix}$$

The normalization condition of the eigenvector puts the following constraints on the values of x1 and x2.

$$x_1^2 + (N-1)x_2^2 = 1$$

The eigenvalue equation puts the following constraints on the values of x1 and x2.

$$(N-1)x_2 = \lambda x_1$$

$$x_1 + 2x_2 = \lambda x_2$$

From these three relations the largest eigenvalue (λ) and the two components x1 and x2 can be computed as given below.

$$\lambda = 1 + \sqrt{N}$$

$$x_1 = \frac{(N-1)x_2}{\lambda} = \sqrt{\frac{(N-1)}{\sqrt{2\sqrt{N}(1+\sqrt{N})}}}$$

$$x_2 = \frac{\lambda}{\sqrt{(N-1)^2 + (N-1)\lambda^2}} = \sqrt{\frac{1+\sqrt{N}}{\sqrt{2\sqrt{N}(N-1)}}}$$

Node Type	**Eigenvector Centrality**	**Normalized Centrality**	**Max Norm Centrality – Actual Norm Centrality**
Hub node	$x_1 = \sqrt{\frac{(N-1)}{\sqrt{2\sqrt{N}(1+\sqrt{N})}}}$	1	0
(N-1) Leaf nodes	$x_2 = \sqrt{\frac{1+\sqrt{N}}{\sqrt{2\sqrt{N}(N-1)}}}$	$\frac{\lambda}{(N-1)}$	$1 - \frac{\lambda}{(N-1)}$
Sum		$1+\lambda$	$(N-1-\lambda) = N-2-\sqrt{N}$
Network Eigenvector Centralization			$\frac{(N-2-\sqrt{N}}{(N-1-\sqrt{N-1}}$

3.7 Triangular Graph: We define a triangular graph as a modification to a line graph with all the nodes in the line graph connected to a central node. In order to define a central node for this configuration, we need to have N > 4. Such graph structure can be seen in organizations having sequential processing department controlled by a centralized office. The communication paths in such organizational structure can be represented by such graph.

Degree centrality: The central node has a degree equal to N-1, the two leaf nodes at the two ends have degree 2, and all the other nodes have degree 3 each.

Node Type	**Degree**	**Normalized Centrality**	**Max Norm Centrality – Actual Norm Centrality**
Central node	$N-1$	1	0
2 Leaf End nodes	2	$\frac{2}{N-1}$	$\frac{N-3}{N-1}$
(N-3) Leaf nodes	3	$\frac{3}{N-1}$	$\frac{N-4}{N-1}$
Sum			$\frac{(N-2)(N-3)}{N-1}$
Network Centralization			$\frac{N-3}{N-1}$

Closeness centrality: The central node is at a distance of one unit from all other nodes. The two leaf nodes at the ends are at a distance of one unit from the central and one adjacent leaf node and at a distance of two units from other remaining nodes. The remaining (N-3) leaf nodes are at a distance of one unit from the central and two adjacent leaf nodes and at a distance of two units from other remaining nodes.

Node Type	**Closeness Centrality**	**Normalized Centrality**	**Max Norm Centrality – Actual Norm Centrality**
Central node	$N-1$	1	0
2 Leaf End nodes	$2+2(N-3)$	$\frac{N-1}{2+2(N-3)}$	$\frac{N-3}{2+2(N-3)}$
(N-3) Leaf nodes	$3+2(N-4)$	$\frac{N-1}{3+2(N-4)}$	$\frac{N-4}{3+2(N-4)}$
Sum			$\frac{(N-1)(N-3)^2}{(N-2)(2N-5)}$
Network Centralization			$\frac{(N-3)^2}{(N-2)(2N-5)}$

Betweenness centrality: Out of the $\frac{(N-1)(N-2)}{2}-(N-1)$ possible geodesics passing through the central node, (N-3) geodesics connecting pair of leaf nodes (one each on either side) adjacent to every leaf node (except at the two ends) have the same path length equal to two, one passing through the central node and other passing through the intervening leaf node. Therefore, the effective number of geodesics through the central and node is given by

$$\frac{(N-1)(N-2)}{2}-(N-1)-\frac{1}{2}.(N-3)=\frac{(N^2-6N+7)}{2}$$

and that through each of the leaf nodes (except at the two ends) is given by $\frac{1}{2}$.

Node Type	**No. of Geodesics**	**Normalized Centrality**	**Max Norm Centrality – Actual Norm Centrality**
Central node	$\frac{(N^2-6N+7)}{2}$	1	0
2 Leaf end nodes	0	0	1
(N-3) Leaf nodes	$\frac{1}{2}$	$\frac{1}{N^2-6N+7}$	$\frac{N^2-6N+6}{N^2-6N+7}$
Sum		$\frac{(N-4)(N-1)}{N^2-6N+7}$	$\frac{(N-2)(N^2-5N+2)}{N^2-6N+7}$
Network Centralization			$\frac{(2N-3)(N^2-5N+2)}{(N-1)(N^2-6N+7)}$

Eigenvector centrality: Deriving closed form expression for eigenvector centrality of this network is complicated because of asymmetry among the leaf nodes. Therefore, it needs to be computed numerically.

3.8 Line graph: Line graph is an open chain with all nodes connected to their adjacent nodes. The nodes at the two ends have only one neighbor each. Such graph structure can be seen in organizations having sequential autonomous processing departments without any centralized control. The coordination between various departments is achieved only through the communication between either or both of the source or sink of the product. The largest eigenvalue for the line graph can be approximated to that of the cyclic graphs for large N as these two graphs are similar except for one missing edge. Degree of all nodes except the two nodes at the ends are 2 and the degree of the nodes at the ends are 1 each. The closeness of the nodes with other nodes can be computed using the general formula for the sum of the shortest distance of ith node from any end to all other nodes = $\frac{N(N-i)}{2}+\frac{i(i-1)}{2}$. The betweenness of a node can be computed using the number of geodesics the ith node from any end falls on. The number of such geodesics = $(i-1)(N-i)$. Eigenvector centrality needs to be computed numerically for this configuration. Deriving closed form solution for all the graph centralization based on the four measures of this network is complicated because of asymmetry among leaf nodes, therefore this needs to be computed numerically using its node centrality values.

3.8 Cyclic graph: A cyclic graph is a graph with all the nodes connected to only two adjacent nodes. Degree of each of the nodes is 2. The closeness of nodes can be computed from the general formula for the distance of nth node from all other nodes = $\frac{N^2}{4} \quad if \quad N \quad is \quad even$

And $\frac{N^2-1}{4} \quad if \quad N \quad is \quad odd$

The betweenness centrality can be computed by observing that the umber of geodesics the ith node falls on = $\frac{N-3}{2}$. The largest eigenvalue for the cyclic networks can be shown to be equal to 2. The components of eigenvector representing the centrality of all the nodes can be shown to be given by $x_1 = \frac{1}{\sqrt{N}}$. Thus we can observe that all the four graph centrality of cycle graph becomes zero as all the nodes in this graph are equivalent having same relative centrality value.

3.10 Complete graph: In a complete graph every node is connected to all other nodes. Degree of all nodes is N-1. Closeness of nodes can be computed by observing that the sum of the shortest distance (geodesics) of any node from all other nodes = N-1. The betweenness centrality of nodes can easily be shown to be zero as the number of geodesics between any pair of nodes on which any other node falls on is equal to 0. The largest eigenvalue for complete network and cyclic networks can be shown to be equal to (N-1). The components of eigenvector representing the centrality of all the nodes can be shown to be given by $x_1 = \frac{1}{\sqrt{N}}$. By symmetry all the nodes in a complete graph are equivalent having equal centrality value in the network and hence all the four network centralization measures are equal to zero.

4. Results and Observations

We have computed the centrality values for two networks, one with total N= 60 nodes and the other with total 5040 nodes decentralized into various configurations with different numbers of hubs. We have compared the results obtained using the formulae derived in this paper and that obtained through numerical computation of the centrality for an example network with 60 nodes arranged in different symmetric permutations. We compare the behavior of these centrality measures on different network configurations with similar size as well as same configuration but with different size.

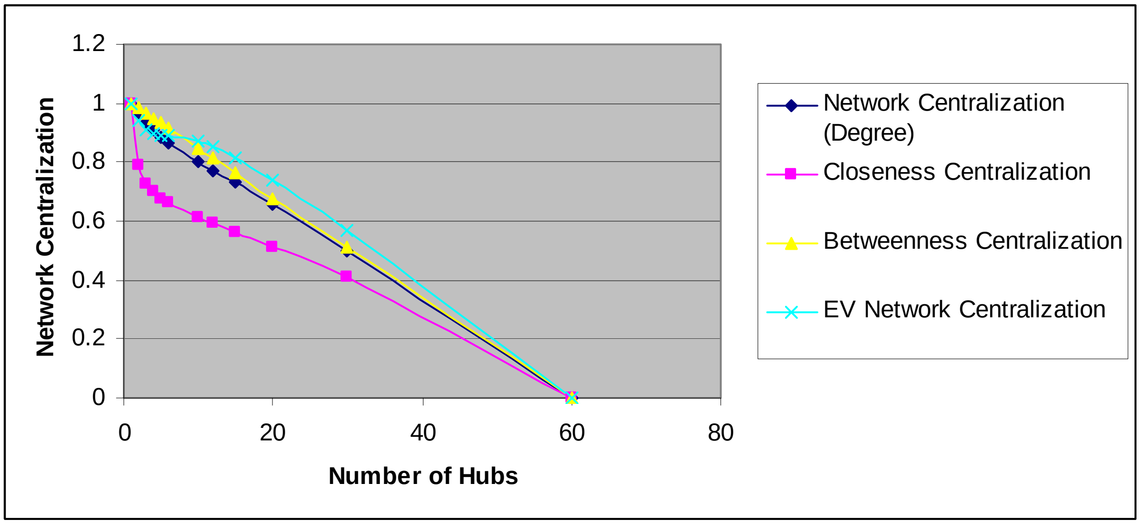

Figure 2: Variation of network centralization with the number of hubs (Total nodes N = 60)

Figure 2 shows the network centralization values for a network with N = 60 nodes. The result with N= 5040 gives similar pattern. The general formula helps us to compute the network centralization value for large network of any size with similar configuration. Figure 3 shows the impact of inserting a central node as shown in Figure 1(L) in otherwise decentralized configuration given in Figure 1(G).

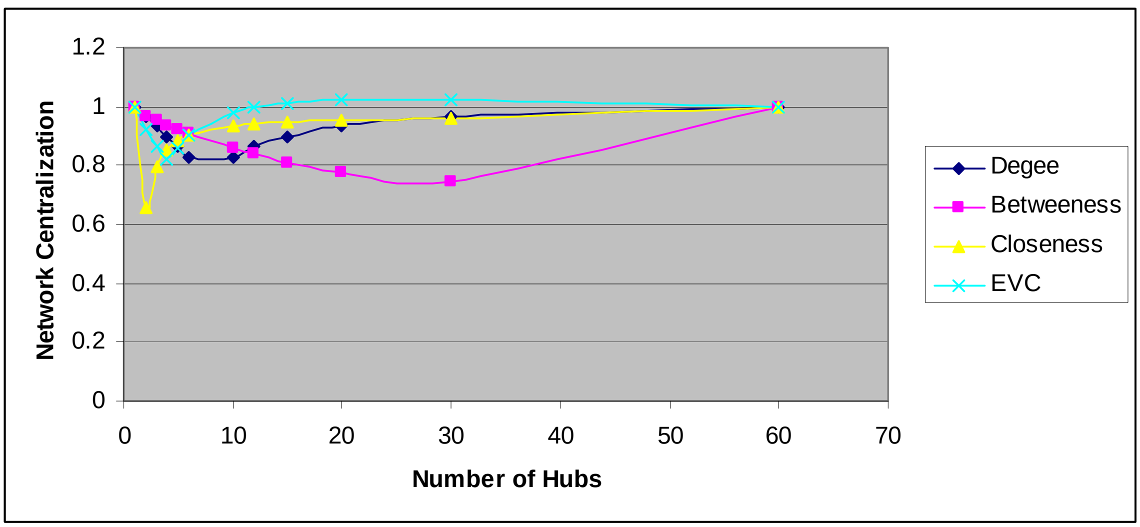

Figure 3: Effect of inserting one central node in otherwise decentralized structure (N=60)

Principal eigenvalue plays a crucial role as a parameter in computing the eigenvector centrality. Computing the eigenvalue numerically for large networks is difficult. Approximation of the largest (principal) eigenvalue of the adjacency matrix of a large network is important due to its crucial role in understanding the dynamical processes in the network (Restrepo *et al.* 2007). We therefore derived analytic expressions for the principal eigenvalue for the two types of decentralized network. Closed form expression derived for the principal eigenvalue helps us compare the principal eigenvalue of some of possible network configurations with any number of nodes. Figure 4 compares the variation of the principal eigenvalue with the level of decentralization in the two types of network configurations (G and L in Figure 1) with size N equal to 60 and 61 respectively. It clearly shows that the principal eigenvalue is strongly dependent on the configuration of the network. We have shown that beyond certain level of decentralization, the principal eigenvalue of the network with a central node is less sensitive (scales as the square root of the size of the network) to the level of decentralization compared to that without a central node (scales as the size of the network). We have seen similar pattern for principal eigenvalues of the two types of network configuration with size equal to 5040 and 5041 respectively.

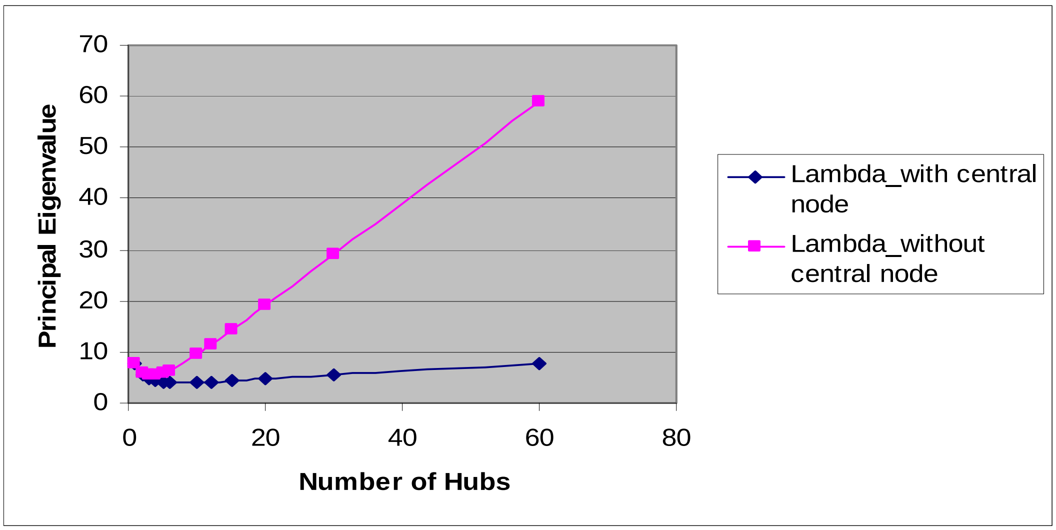

Figure 4: Principal eigenvalue vs number of clusters (N=60)

As discussed earlier, the network centralization based on eigenvector centrality for star configuration can be shown to be equal to $(N-1)-\sqrt{(N-1)}$. This approach does not give maximum value for the normalization constant (i.e. denominator) used in equation 1. Therefore using it as a normalization constant in equation 1 may give erroneous result as shown in Figure 5 where the graph centralization becomes greater than 1 for cases other than a star configuration. We find that the maximum value of the normalization constant for the case of eigenvector centrality is achieved when only two nodes of the network are connected and rest others are independent. Though this configuration is intuitively not a highly centralized configuration, it gives the maximum value equal to $(N-2)$ for the normalization constant to be used as the denominator of equation 1. We can now be assured to have the centralization value less or equal to 1, but this lacks intuitive justification for the case when the value is 1. The eigenvector centrality of a node is defined as the respective component of the principal eigenvector. If we define the eigenvector centrality of a node in a network to be given by the square of respective component, the relative order of centrality would still be maintained. With this definition, the network centralization based on eigenvector centrality for star configuration can be shown to be equal to $(N-2)$. This approach of defining the eigenvector centrality is more suitable for computing network centralization as the most central star topology achieves the highest value using this approach. Since the components of the principal eigenvector are positive and less than one, there is a monotonous relationship between the components and their squares. Therefore, defining EVC in this manner won't affect the relative ranking of the nodes in the network. The network centralization based on the centrality defined as the square of the components of the principal eigenvector gives the maximum network centrality value for star configuration as shown in Figure 5.

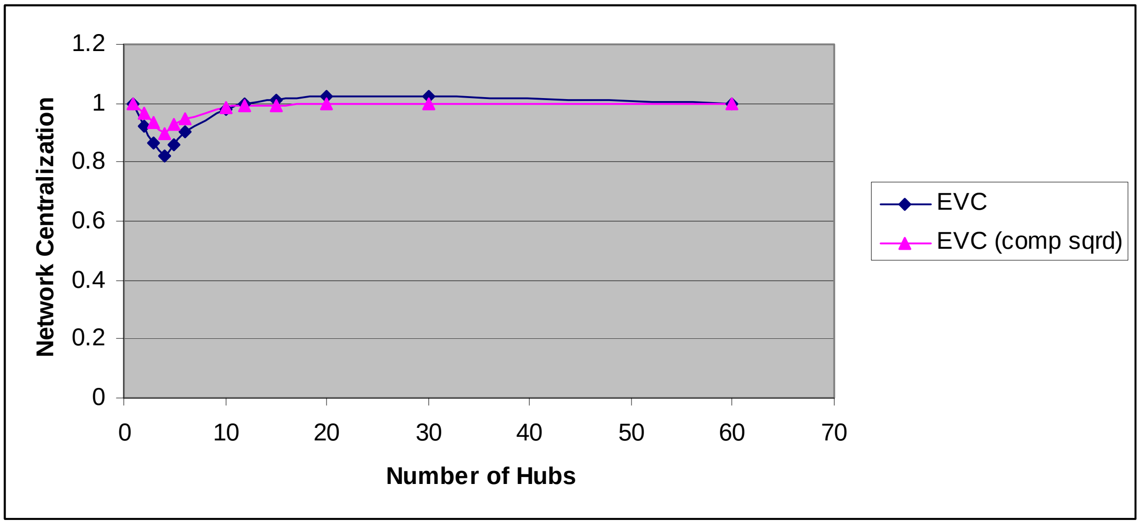

Figure 5: Variation of network centralization (eigenvector) with number of hubs (N=60)

For different types of network configurations with varying level of decentralization, the relative level of the four types of network centralization values changes differently as shown in Figure 6. This shows that different types of communication and interaction mechanism may have different level of effectiveness at various levels of decentralization. We observe that the sensitivity of the eigenvector and betweenness centrality measures is low compared to the degree and closeness centrality. Among the four centrality measures, closeness centrality has the maximum sensitivity to the level of decentralization. Therefore, it may work as a better measure for quantifying the level of decentralization in a network.

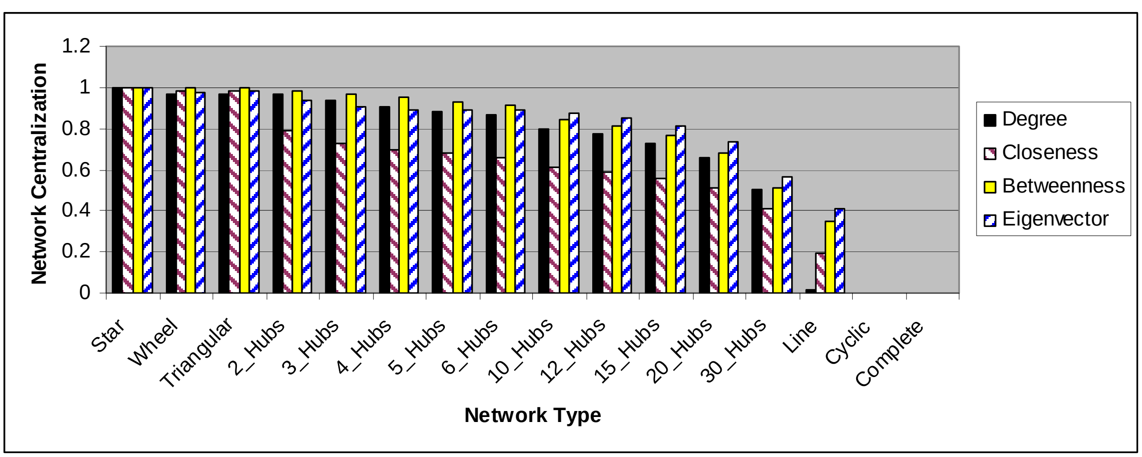

Figure 6: Network Centralization for various network types (N= 60)

Since the variance of node centrality in the network can also be used as a measure of network centralization (Wasserman & Faust 1994), we have compared its behavior for all the four types of centrality measures as shown in Figure 7. There is a monotonous relationship between the variance of node centrality and the network centralization. We notice that the variance of all the types of node centralities increases with the level of decentralization. This measure of centralization has some counter-intuitive behaviour. For example, though cyclic and complete graphs have the least variance (zero) of node centrality, these two graphs have zero network centralization. Therefore, this cannot be used as a reliable measure of network centralization of any generic network configuration.

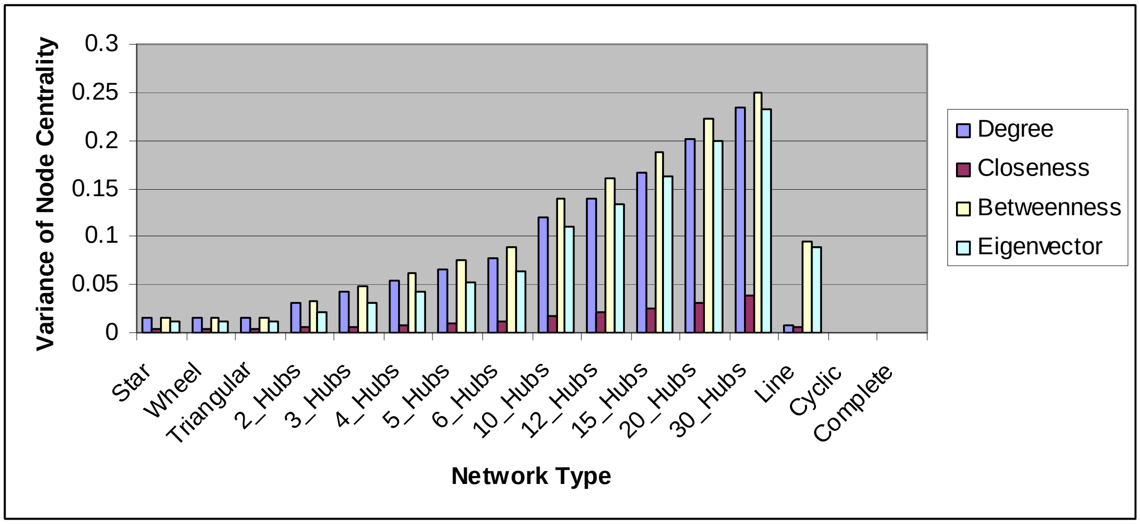

Figure 7: Variance of node centrality for various network types (N= 60)

There is no unanimity on which type of network centrality measure should be used to measure the level of centralization. These centrality measures may be treated as expected centrality value of a particular node if the interaction between nodes is governed by a particular kind of flow process. In reality several mechanisms may be simultaneously present; therefore the node centrality should be computed as the weighted average of these measures. The weights should account for the relative intensity of particular type of flow process prevalent in the network. We have computed the relative centrality of the central node and hub nodes for the network configuration shown in figure 1 (L) with size N = 5041. Figure 8 and 9 show the variation of the centrality values for different levels of decentralization. Since the possible number of hub nodes (n) is large, it has been presented in logarithmic scale. We notice that when the number of hubs is relatively small compared to the number of leaf nodes at each of the hub nodes, the hub nodes retain supremacy over the central node on the basis of the degree and eigenvector centralities. The central node retains its supremacy over the hub nodes on the betweenness and closeness centrality for all levels of decentralization.

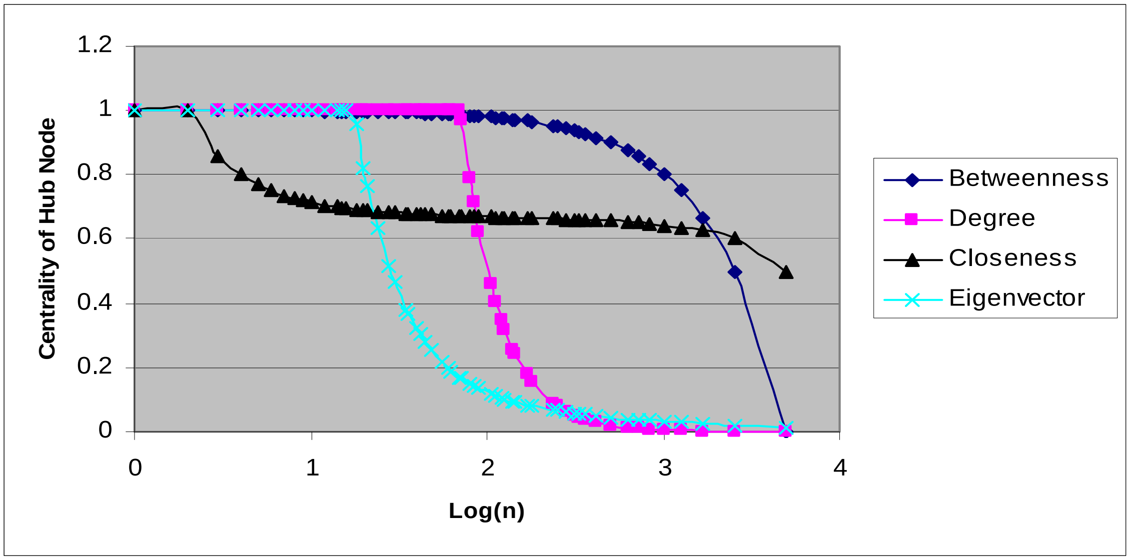

Figure 8: Normalized centrality of Hub nodes vs Log(n)

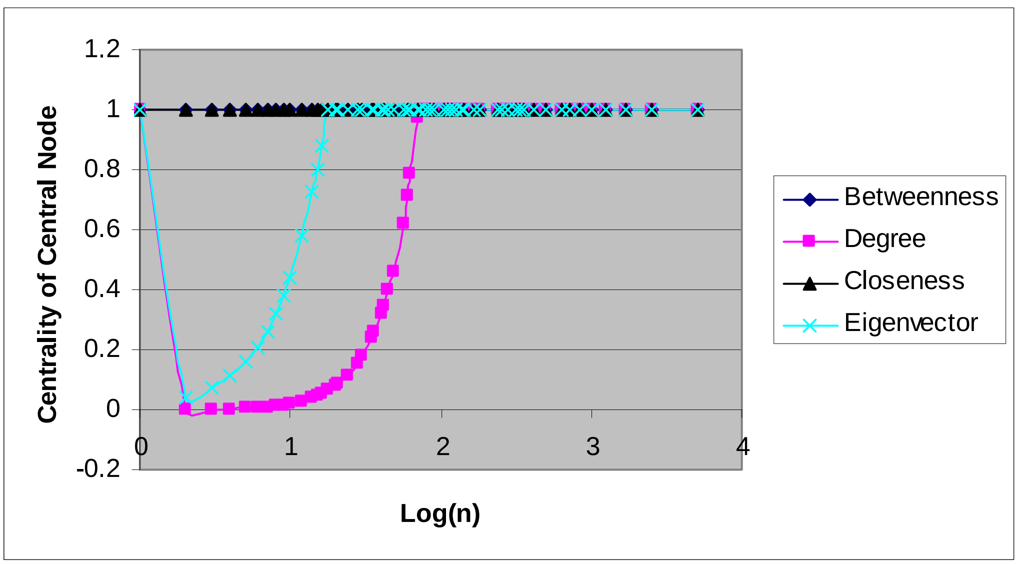

Figure 9: Normalized centrality of central node vs Log(n)

Correlation between various centrality measures

(Perra & Fortunato 2008) has compared the centrality measures based on the spectral properties of graph matrices and shown that for graphs with special properties, the measures are often correlated with each other but for real graphs with less regular structure, the centrality measures are far less correlated. We investigated how the centrality values are correlated for the two types of configurations. The correlation between the five centrality measures discussed in this paper as the number of hub nodes increases in the network has been shown in table 1 and table 2. We find that the centrality measures are strongly correlated for the configuration without any central node while there is a varying level of correlation for the configuration with a central node. The two types of eigenvector centralities show strong correlation though not equal to unity. Degree centrality has relatively high correlation with both types of eigenvector centralities compared to the betweenness and closeness centralities. An interesting finding is that the betweenness centrality is negatively correlated with all other centralities for such configuration.

Table 1: Correlation for network with a central node and hubs

	Degree	Betweenness	Closeness	EVC	EVC (Comp Sqrd)
Degree	1				
Betweeness	-0.37899	1			
Closeness	0.09161	-0.19487	1		
EVC	0.750984	-0.52004	0.240882	1	
EVC (Comp Sqrd)	0.721989	-0.41398	0.134161	0.96637	1

Table 2: Correlation for network without a central node and hubs

	Degree	Betweenness	Closeness	EVC	EVC (Comp Sqrd)
Degree	1				
Betweeness	0.999901	1			
Closeness	0.911722	0.90862	1		
EVC	0.996507	0.996831	0.906238	1	
EVC (Comp Sqrd)	0.999882	0.999995	0.908671	0.997067	1

5. Concluding Remarks

The sensitivity of various centrality measures on the level of decentralization of a network has been discussed. The issue with the derivation of network centralization based on eigenvector centrality of nodes of the network has been discussed and a modification in the eigenvector centrality measure for the nodes and normalization constant have been suggested in order to derive the network centralization in a unified manner compared to the other three measures. We have derived the theoretical formulae for the various network centralization measures of some standard network topologies with varying level of decentralization. Further, the impact of the level of decentralization on the network centralization measures based on all the four centrality measures has been discussed. The variance of node centrality corresponding to each of the four centrality measures namely degree, betweenness, closeness and eigenvector has also been analyzed. This paper contributes to the social network analysis literature by providing a theoretical derivation of various network centralization measures for some standard network configurations of any size that may serve as benchmarks for comparing the centralization of empirical networks. Further, there is no established way in the literature to estimate the number of hubs in a real world network based on their centrality values. We observe that the four centrality measures have different characteristics response to the level of centralization. There is a possibility to fill this void by mapping the observed centrality measures of empirical network on these theoretical expected values to estimate the effective number of hubs in the network. A comparative analysis of the four centralization measure may point to the level of decentralization in the network. The result also shows at a particular level of decentralization which of the known flow mechanisms could be more effective in a decentralized network. The proposed approach and the findings in this research may be used for approximating centrality values for large networks.

References

Baker, W.E., 1984 The Social Structure of a National Securities Market. The American Journal of Sociology 89, 775-811

Bardhan, P., 2002. Decentralization of Governance and Development. The Journal of Economic Perspectives 16(4), 185-205

Bonacich, P., 1987. Power and Centrality: A Family of Measures. The American Journal of Sociology 92

Bonacich, P., Lloyd, P., 2004. Calculating status with negative relations. Social Networks 26, 331–338

Borgatti, S.P., 2005. Centrality and network flow. Social Networks 27, 55–71

Borgatti, S.P., 2006. Identifying sets of key players in a social network. Comput Math Organiz Theor 12, 21–34

Borgatti, S.P., Everett, M.G., 2006. A Graph-theoretic perspective on centrality. Social Networks 28, 466-484

Borgatti, S.P., Jones, C., Everett, M.G., 1998. Network Measures of Social Capital. CONNECTIONS 21(2), 27-36

Burt, R.S., 1999. The Social Capital of Opinion Leaders. Annals of the American Academy of Political and Social Science 566, 37-54

Burt, R.S., 2001. Structural Holes versus Network Closure as Social Capital. Book Chapter 2 in Social Capital: Theory and Research edited by Nan Lin, Karen S. Cook, and Ronald S. Burt, Transaction Pulishers, New Jersey, 31-56

Burt, R.S., 2004. Structural Holes and Good Ideas. American Journal of Sociology 110(2), 349-399

Christakis, N., Fowler, J., 2008. The collective dynamics of smoking in a large social network. New England Journal of Medicine 358, 2249-2258

Costenbader, E., Valente, T.W., 2003. The stability of centrality measures when networks are sampled. Social Networks 25, 283–307

Doerfel, M.L., 1998. What Constitutes Semantic Network Analysis? A Comparison of Research and Methodologies. CONNECTIONS 21(2), 16-26

Edelman, A., Rao, N.R., 2005. Random Matrix Theory. Acta Numerica DOI: 10.1017/S0962492904000236, 1-65

Ercsey-Ravasz, M., Toroczkai, Z., 2010. Centrality scaling in large networks. Physical Review Letters 105, 038701, arXiv:1003.0692v2

Freeman, L.C., 1979. Centrality in social networks: Conceptual clarification. Social Networks 1, 215-239

Friedkin, N.E., 1991. Theoretical Foundations for Centrality Measures. The American Journal of Sociology 96, 1478-1504

Friedkin, N.E., 1993. Structural Bases of Interpersonal Influence in Groups: A Longitudinal Case Study. American Sociological Review 58(6), 861-872

Gargiulo, M., Benassi, M., 2000. Trapped in Your Own Net? Network Cohesion, Structural Holes, and the Adaptation of Social Capital. Organization Science 11, 183-196

Guckenheimer, J., Ottino, J.M., 2008. Foundations for Complex Systems Research in the Physical Sciences and Engineering. In: Report from an NSF Workshop. Cornell University, Northwestern University

Gupta, A.K., Govindarajan, V., 1991. Knowledge Flows and the Structure of Control within Multinational Corporations. The Academy of Management Review 16(4), 768-792

Hanneman, R.A., Shelton, C.R., 2011. Applying modality and equivalence concepts to pattern finding in social process-produced data. Social Network Analysis and Mining 1(1), 59-72

Heimo, T., Tibely, G., Saramaki, J., Kaski, K., Kertesz, J., 2008. Spectral methods and cluster structure in correlation-based networks. Physica A 387, 5930-5945

Helbing, D., 2008. Managing Complexity: Insights, Concepts, Applications. Springer.

Inkpen, A.C., Tsang, E.W.K., 2005. Social Capital, Networks, and Knowledge Transfer. Academy o? Management Review 30(1), 146-165

Kim, D.-H., Jeong, H., 2005. Systematic analysis of group identification in stock markets. Physical Review E 72, 046133

Kolaczyk, E.D., 2009. Statistical Analysis of Network Data: Methods and Models. Springer.

Lin, M., Li, N., 2010. Scale-free network provides an optimal pattern for knowledge transfer. Physica A: Statistical Mechanics and its Applications 389, 473-480

Lohmann, G., Margulies, D., Horstmann, A., Pleger, B., al, J.L.e., 2010. Eigenvector Centrality Mapping for Analyzing Connectivity Patterns in fMRI Data of the Human Brain. PLoS ONE 5(4): e10232, 1-8

Madhavan, R., Koka, B.R., Prescott, J.E., 1998. Networks in transition: how industry events (re)shape interfirm relationships. Strategic Management Journal 19, 439–459
Mishra, N., Schreiber, R., Stanton, I., Tarjan, R.E., 2009. Finding Strongly-Knit Clusters in Social Networks. Internet Mathematics November 2009
Newman, M., 2008. The physics of networks. Physics Today, 33-38
Newman, M.E.J., 2003. The Structure and Function of Complex Networks. SIAM Review 45, 167--256
Newman, M.E.J., 2006. Finding community structure in networks using the eigenvectors of matrices. Physical Review E 74, 036104
Perra, N., Fortunato, S., 2008. Spectral centrality measures in complex networks. Physical Review E 78, 036107; arXiv:0805.3322v2 [physics.soc-ph]
Renneboog, L., Zhao, Y., 2011. Us knows us in the UK: On director networks and CEO compensation. Journal of Corporate Finance doi:10.1016/j.jcorpfin.2011.04.011
Restrepo, J., Ott, E., Hunt, B., 2007. Approximating the largest eigenvalue of network adjacency matrices. Physical Review E Pt 76(5), 056119
Rivkin, J.W., Siggelkow, N., 2007. Patterned Interactions in Complex Systems: Implications for Exploration. Management Science 53(7), 1068-1085
Rosen, D., Barnett, G.A., Kim, J.H., 2010. Social networks and online environments: when science and practice co-evolve. Social Network Analysis and Mining 1, 27-42
Roy, R.B., Sarkar, U.K., 2011a. Identifying influential stock indices from global stock markets: A social network analysis approach. In: The 2nd International Conference on Ambient Systems, Networks and Technologies (ANT), Ontario, Canada
Roy, R.B., Sarkar, U.K., 2011b. A social network approach to examine the role of influential stocks in shaping interdependence structure in global stock markets. In:International Conference on Advances in Social Network Analysis and Mining (ASONAM), Kaohsiung, Taiwan
Uzzi, B., 1997. Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness. Administrative Science Quarterly 42(1), 35-67
Wasserman, S., Faust, K., 1994. Social Network Analysis: Methods and Applications. Cambridge University Press, 461-502