A COMPARISON OF CENTRALITY MEASURES FOR EVALUATING THE PROJECT COORDINATION PERFORMANCE

A Thesis Submitted to
the Graduate School of Engineering and Sciences of
İzmir Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in Architecture

by
Pelin İpek TİMURCAN

September 2014
İZMİR

We approve the thesis of **Pelin İpek TİMURCAN**

Examining Committee Members:

Assist. Prof. Dr. Sevgi Zeynep DOĞAN
Department of Architecture, İzmir Institute of Technology

Prof. Dr. Serdar KALE
Department of Architecture, İzmir Institute of Technology

Assoc. Prof. Dr. Suat GÜNHAN
Department of Architecture, The University of Texas at San Antonio

5 September 2014

Assist. Prof. Dr. Sevgi Zeynep DOĞAN
Supervisor, Department of Architecture, İzmir Institute of Technology

Assoc. Prof. Dr. Şeniz ÇIKIŞ
Head of the Department of Architecture

Prof. Dr. R. Tuğrul SENGER
Dean of the Graduate School of Engineering and Sciences

ACKNOWLEDGEMENTS

I wish to thank, first and foremost, my thesis supervisor Assist. Prof. Dr. Sevgi Zeynep DOĞAN who supported me throughout my thesis process with her valuable knowledge and expert guidance. She always encouraged me to go further in my studies and helped me to look things from a different perspective.

I am equally grateful to Prof. Dr. Serdar KALE, who has read numerous drafts of the thesis, for his special assistance, suggestions and comments to my study.

I would also like to thank my thesis committee member Assoc. Prof. Dr. Suat GÜNHAN for his irreplaceable guidance and support.

I especially wish to express my deep appreciation and sincere gratitude to my fiancee Yiğit TEMURCAN for his continuous support, patience and optimism throughout the process.

Finally, it is a pleasure to thank to my parents Neşe and Engin TİMURCAN and my brother Mithat Can TİMURCAN for their encouragement, emotional support, trust and great patience throughout my education and in every moment of my life.

ABSTRACT

A COMPARISON OF CENTRALITY MEASURES FOR EVALUATING THE PROJECT COORDINATION PERFORMANCE

Evaluation of coordination performance in a project network requires reliable measures and monitoring methods for effective management. Recent literature includes studies addressing the relationship between coordinative activity and the configuration of communication networks. In these works, the role of network centrality is investigated through the basic standard centrality measures of degree, betweenness and closeness. Current social network analysis research emphasizes new formulations of centrality measures for robust structural analysis of project networks.

This study presents a novel approach for measuring network centrality using the concept of information centrality. It is based on the idea that all paths carry information. The significance of information centrality values for the actors in two major projects of Enron Corporation is investigated.

There are three major findings from these analyses. First finding suggests that highly centralized actors show more coordination. Second finding suggests that information centrality is the most potent predictor for coordination in undirected networks. Third finding is that out-centrality measures are better predictors for coordination performance in directed networks. The implications of these findings have potentials for assessing the coordination performance in construction management research, and it is promising for the structural analysis of project communication networks.

ÖZET

PROJE KOORDİNASYON BAŞARIMININ DEĞERLENDİRİLMESİ İÇİN MERKEZİLİK ÖLÇÜMLERİNİN BİR KARŞILAŞTIRMASI

Projenin koordinasyon performansının etkin yönetilebilmesi için güvenilir ölçüm ve izleme yöntemlerine gerek vardır. Son yıllarda yapılan araştırmalar, koordinasyon faaliyeti ile iletişim ağının yapılanması arasındaki ilişkiyi ele almaktadır. Bu çalışmalarda ağ merkeziliğinin rolü, temel merkezilik ölçümleri olan derece, arada olma ve yakınlık merkeziliği değerleri ile incelenmiştir. Sosyal ağ analizi ile ilgili yeni araştırmalar, proje ağlarına ait güvenilir yapısal analizlerin yeni tanımlanan merkezilik ölçümleri ile yapılabileceğini kanıtlamıştır.

Bu çalışmada, ağ merkeziliğinin güvenilir ölçümlerine ulaşabilmek amacıyla, bilgi merkeziliği kavramı kullanılarak yeni bir yaklaşım sunulmuştur. Bu yaklaşım, bütün yörüngelerin bilgi taşıdığı fikrine dayanır. Çalışmada bilgi merkeziliği değerlerinin önemi araştırılmıştır. Bilgi merkeziliği değerleri, Enron şirketinin iki büyük projesinin aktörleri için hesaplanmıştır.

Bu analizlerden üç önemli bulgu elde edilmiştir. İlk bulgu merkezi olarak konumlanan aktörlerin daha fazla koordinasyon içinde olduğunu göstermiştir. İkinci bulgu, koordinasyon için bilgi merkeziliğinin yönsüz ağlarda en güçlü belirleyici olduğunu göstermiştir. Üçüncü bulgu, yönlü ağlarda koordinasyon performansı için dış-merkezilik ölçümlerinin iç-merkezilik ölçümlerinden daha iyi belirleyiciler olduğunu göstermiştir. Bu bulguların inşaat yönetimi araştırmalarında koordinasyon performansını değerlendirmek için potansiyelleri olduğu görülmektedir ve bu sonuçlar proje iletişim ağının yapısal analizi için umut vericidir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

SNA	: Social Network Analysis
E-mail	: electronic mail

CHAPTER 1

INTRODUCTION

In this 'Introduction' chapter, definition of the problem area where this thesis work proposes to make a contribution and the framework of the study are clarified. Arguments of previous research are reviewed. Objectives are listed as primary and secondary. The procedure of the study is presented under 'Research Methodology.' Finally the limitations of the study are briefly explained under the subtitle 'Limitations.'

1.1. Definition of the Problem

Communication networks and the information exchange play a critical role in the coordinative activity of project participants (Hossain et al. 2006; Hossain 2009a,b; Hossain and Wu 2009). The social network characteristic of centrality has the potential of identifying project participants in coordinator roles. Traditionally, contractual links within the project participant network define the coordinator role but recent findings argue that a central position in the communication network of a project organization may also be an indication of a party's coordinator role (Dogan et al. 2012; Dogan et al. 2013; Hossain et al. 2006; Hossain 2009a,b; Hossain and Wu 2009). The correlation between network centrality and coordination is widely discussed in social studies (Bavelas 1950; Freeman 1979; Leavitt 1951). However, the construction research domain lacks such discussion apart from the work of Hossain et al. (2006), Hossain (2009a) and Dogan et al. (2012). Recently, Dogan et al. (2013) proposed an uncomplicated, quantitative method to predict a firm's coordination score via a centrality index based on social network analysis. In this recent study, three standard centrality measures of the firms in a wayfinding signage project at a major airport construction were calculated and a centrality index was defined for each firm using the average of these centrality measures. Findings show that the coordination scores of the firms are highly correlated with the centrality measures.

Centrality measures the distribution of relationships and describes the influence of a node based on how well connected the node is in a network. Three basic measures, degree, betweenness and closeness are proposed by sociologists to determine the centrality of a node in a social network (Freeman 1979). Different measures of centrality capture different aspects of a node according to its role within the network. As Freeman argued in his seminal paper (1979) the degree of centrality indexes a node's activity, whereas betweenness centrality measures a node's control, and closeness centrality measures its communication efficiency. Previous empirical studies provide nodes that exhibit high centrality and thus achieve high levels of performance (Ahuja et al. 2003; Pryke 2004; Schilling and Phelps 2007; Tallberg 2004; Wasserman and Faust 1994). However, over the years, researchers have proposed different measures, such as information centrality (Stephenson and Zelen 1989) focusing on different features of central nodes.

This thesis aims to introduce a novel approach for evaluating coordination using network centrality based on the concept of information centrality that has been overlooked for the analysis of communication networks in construction research. The measure of information centrality weighs all paths between a pair of nodes assuming all paths carry information and quantifies the relevance of each of the nodes in the network. It has a rationale and provides readily interpretable data. The calculations are relatively straightforward and can be done for large networks. The use of quantifiable values of information centrality can also illustrate communication structures, patterns and relationships in an e-mail network (Stephenson and Zelen 1989).

The thesis is organized as follows: In this chapter, first the initial idea and the framework of the study are clarified. In the second chapter, coordination theory, coordination studies in the construction management research and social network analysis method are reviewed. The procedure of the study is explained in the "Research Methodology" chapter. In the fourth chapter findings and discussion are presented and finally the contents of the study were briefly explained and concluding remarks are made in the "Conclusion" chapter.

1.2. Aim of the Study

The aim of this research is to determine the effects of network centrality on coordination. During the research two main objectives are composed; primary and secondary objectives. The primary objectives of the research are;

- to investigate the relationship between coordination performance and information centrality building upon the established coordination and social network theories,
- to find out which measure of centrality is the most useful predictor for coordination ability.

The secondary objectives of the study are;

- to investigate the relationship between different centrality measures and coordination,
- to classify recent research streams on SNA which are promising for construction research.

1.3. Research Methodology

This study was built on existing coordination theory defined by Malone and Crowston (1994). This theory is based on studying the interdependence between activities. The coordination key phrases data were derived from a precedent study of Hossain (2009a) for a broader examination of the importance of information centrality for capturing and assessing the coordinative activity within a project network. The proposed approach was further investigated in two case studies involving electronic information exchange between the employees of Enron Corporation. The computational procedures are provided with theoretical bases and methods.

The study was carried out in two consecutive phases regarding correlations: the five measures centrality measures were associated with the actors' coordination scores were thoroughly investigated and the centrality measures were analyzed in order to illustrate the reliability dimensions which may indicate the potential coordination performance.

The study used the concept of information centrality in conjunction with other standard centrality measures to explore the correlation between network centrality and coordination of two major projects for Enron Corporation.

Findings from this study motivate that the thorough analyses using information centrality measure in communication networks may result in more robust and illustrative findings for the coordination performance and information reliability. From a theoretical standpoint, the method for measuring information centrality proposed by

Stephenson and Zelen (1989) makes use of all paths between nodes rather than geodesic paths. The calculations based on this framework can be easily adopted and organized for larger project networks.

It can be claimed that the approach presented in this study associated with the proposed computational procedures has the application potentials for rapid analysis of complex project organizations, robustness of information exchange or specifically changes in project networks in the case of new potential network nodes added or subtracted (Stephenson and Zelen 1989). The adopted approach has advantages over complex and expensive computational methods and graphic techniques for analyzing project networks and organizations, in terms of simplicity and practicality.

The methodology of the study involves two parts: (1) Conducting social network analysis for the email dataset obtained from the Enron e-mail dataset; calculating centrality measures; and investigating correlations among the different types of centrality measures; and (2) Evaluating correlation between coordination scores and network centrality indices of project participants.

1.4. Limitations and Assumptions

The primary limitation of the dataset is that non-email communication data are not considered in this study. Since some of the centrality measures such as eigenvector centrality and information centrality cannot be defined for directed networks, the analysis restricted to undirected networks for these measures.

It is worth noting that this research was restricted to two Enron projects as case studies. The social network analysis of the Enron e-mail data can be used for further research on the correlation between centrality measures and coordination performance.

1.5. Outline

In the "Introduction" chapter, definition of the problem, objectives, methodology and limitations of the research are explained briefly. The "Literature Review" chapter involves the review of the coordination and social network analysis theories. The summary tables of the related researches are presented for each of them. The "Research Methodology" chapter includes information about the Enron email data and the case studies, data extraction process and coordination data and network centrality analysis procedures. The "Research Findings and Discussion" chapter presents the results of network centrality and coordination analysis of the case studies and statistical analysis and discussions of the research results. The last chapter is the "Conclusion" chapter. Chapter 5 involves a brief summary of the study and concluding remarks.

CHAPTER 2

LITERATURE REVIEW

In this chapter, a literature review on social network analysis and coordination theory is conducted. This review of the literature will discuss the concepts and mechanisms used in the studies of SNA and coordination. The literature review of this study includes two parts: (1) First part includes the Social Network theory, social network analysis centrality metrics and SNA research in construction management (2) Second part involves coordination theory (3) Third part includes coordination in construction management research

2.1. Social Network Analysis

The concept of social network analysis was first introduced in the 1930s (Moreno 1934). The original studies focused on the social and political relationships between individuals. Since its emergence from sociology in 1930s (Moreno 1934), social networks have been gaining wide attention from theoretical exploration to practical application.

Barnes (1954) started using the term "social network" to denote patterns of ties, concepts usually used by social scientists bounded groups (e.g., tribes, families) and social categories (e.g., gender, ethnicity).

Graphs and sociograms were created with nodes representing individuals and the links between the nodes representing relationships between individuals. In this context sociograms were put forward as a fundamental tool for investing the fabric of interpersonal relationships within groups of individuals.

Park et al. (2011) defined SNA as "a methodology used to identify the conditions of social structures by analyzing the interactions and interrelationships of a set of actors".

2.1.1. Fundamental Concepts in Network Analysis

A *network* is a set of *nodes* connected by a set of *ties*. The *nodes* can be anything persons/individuals, teams, organizations, concepts, patents, etc. Networks which are only made of one type of nodes are *homogeneous*, if the network consists of different types of nodes it is called *heterogeneous.* Whereas ties connect pairs of nodes and can be *directed* or *undirected* according to the type of the interaction between the two nodes. Also ties can be *dichotomous* (present or absent) or *weighted* (the scale can be measured).All ties are weighted or have values, even if the tie exist and is assigned a value of 1 or it doesn't and it is assigned a value of 0. When network analysts collect data on ties from a set of nodes, they call it *relational data*. Relational data can be visualized in matrix form or in graphic form (Coulon unpublished paper 2005). Table 2.1 below, summarizes network terminology.

Table 2.1. Important terms and definitions.

Network Analysis Terms	**Definitions**
Actor/Node	The basic element of a network
Tie/Edge /Line	A set of two nodes
Directed Tie	The interaction is directed to a source node to a destination node
Weighted/Valued Tie	The interaction between the two nodes is associated with a measure of strength
Network	A set of nodes connected by a set of ties
Network Size	The total number of nodes of a network
Network Density	The amount of interactions between the network members
Network Centrality	The distribution of relationships through the network
Geodesic Distance	The distance between the two nodes with the greatest separation in a network
Adjacency matrix	The representation of which nodes of a graph are adjacent to which other nodes
Network Centralization	A measure of the most central node in a network in comparison to every other node
Dyad	Subgraphs of size 2 consisting of a pair of actors and all ties between them
Triad	Subgraphs of size 3 consisting of a triple of actors and all ties among them
Ego	A node which receives particular focus
Alter	The set of nodes that is connected with ego excluding the ego itself

Nodes that are not adjacent may nevertheless be reachable from one to the other. A walk from node *i* to node *j* is an unrestricted sequence of adjacent nodes that begins with *i* and ends with *j*. A trail is a walk in which no link is repeated. A path is a trail in which no node is visited more than once (Borgatti and Everett 2006).

The length of a walk is defined as the number of links it contains, and the shortest path between two nodes is known as a geodesic. The length of a geodesic path between two nodes is known as the geodesic distance between them. The geodesic distances between all pairs of nodes can be represented as a matrix *D* in which *dij* gives the length of the shortest path from node *i* to node *j*. (Borgatti and Everett 2006)

2.1.2. Network Centrality

The centrality has been studied particularly in social sciences. It is a concept that describes nodes' prominence and/or importance in a network (Lee 2006). It addresses which individuals are best connected to other or have most influence. Actors who are the most important or the most prominent are usually located in strategic locations within the network. In order to quantify this relative importance, different centrality measures have been proposed in social science. All such measures attempt to describe and measure properties of "actor location" in a social network (Wasserman and Faust 1994). Among them, degree, closeness, betweenness and the eigenvector centrality are the most commonly used centrality measures.

2.1.2.1. Degree Centrality

Degree centrality is the most basic of all measures that is the extent to which a person is connected to its immediate environment and neighbors. As defined by Freeman (1979), degree centrality is a count of the number of ties incident upon a given node. Also it is the sum of each row in the adjacency matrix representing the network. The defining equation of degree centrality is,

$$C_i^D = \sum_{j=1}^{n} a_{ij} \tag{1}$$

Where n is the number of nodes in the network, and $a_{ij} = 1$ if the nodes i and j are connected by a line, $a_{ij} = 0$ otherwise (Freeman 1979).

Users who communicate with the greater number of people obtain the greater out-degree centrality value. Actors with high out-degree are recognized by other network members as a crucial cog that occupies a central location in a network. On the other hand users who have low out-degree centrality are not very open to the external world and do not communicate with many members. In-degree centrality, called also degree prestige, is based on the in-degree number so it takes into account the number of members that are adjacent to a particular member of the community (Musiat et al. 2009).

According to Loosemore (1998) a person's in-degree and out-degree centralities represent the degree to which it is a receiver or sender of information from or to its neighbors, respectively. At socio-centric level the in-degree index reflects the extent to which one or a few people are the focus of information supply in the whole network. The out-degree index reflects the extent to which information supply is controlled by one or a few people (Loosemore 1998).

In-degree centrality indicates a person's popularity in a network whereas out-degree centrality indicates a person's control or leadership upon a network. However the degree centrality provides information about the local view of relationships between nodes, and does not reflect the overall network structure (Loosemore 1998).

2.1.2.2. Closeness Centrality

Closeness centrality is based on the notion of distance. The measure focus on how close an actor is to all the other actors in the network. Closeness centrality measures independence or efficiency. A person who is closed to many others can have difficulties in acting independently without others knowing. Freeman (1979) argued that a node's interdependence is determined by its closeness to other nodes in the network.

As defined by Freeman (1979), a node's closeness centrality is the sum of graph-theoretic distances from all other nodes, where the distance from a node to another is defined as the length (in links) of the shortest path from one to the other. In the case of information flows, Borgatti (2005) stated that nodes with low closeness scores as being well-positioned to obtain novel information early, when it has the most value. The closeness centrality of point i is (Freeman, 1979):

$$C_i^C = (L_i)^{-1} = \frac{N-1}{\sum_{j \in G} d_{ij}} \qquad (2)$$

Minimum distance of Geodesic is d_{ij} is the minimum number of edges traversed to get from i to j. The closeness cenrality of point i is where L_i the average distance from actor i to all other actors. In a directed graph, the geodesic distance between two actors may differ with the nodal order. “in” and “out” closeness scores are computed separately for a nonsymmetrical matrix (Wasserman and Faust 1994). Closeness was calculated by inverting the distance matrix and taking the row average for out-closeness and the column average for in-closeness (Freeman 1979).

2.1.2.3. Betweenness Centrality

Betweenness centrality measures the extent to which a particular node lies between the various other nodes in the network. In other words betweenness centrality is able to identify *boundary spanners*, people that act as bridges between two or more communities that otherwise would not be able to communicate to each other. Freeman argued that people with high degree of betweenness have potential to control others by controlling the information flows between them. Betweenness centrality is based on the assumption that an individual may gain power if he presides over a communication bottleneck. These people are powerful because of their critical positions in maintaining the open information flow in the network (Loosemore 1998).

Betweenness centrality is defined as the share of times that a node i needs a node k (whose centrality is being measured) in order to reach a node j via the shortest path. Specifically, if g_{ij} is the number of geodesic paths from i to j, and g_{ikj} is the number of these geodesics that pass through node k, then the betweenness centrality of node k is given by,

$$C_i^B = \sum_i \sum_j \frac{g_{ikj}}{g_{ij}},\ i \neq j \neq k \qquad (3)$$

Betweenness centrality is the frequency at which a point occurs on the geodesic that connects pairs of points. Thus, any point that falls on the shortest path between other points can potentially control the transmission of information or effect exchange by being an intermediary. “It is this potential for control that defines the centrality of these points” (Freeman 1979).

2.1.2.4. Eigenvector Centrality

Based on the idea that an actor is more central if it is in relation with actors that are themselves central, Bonacich (1972) argued that the centrality of some node does not only depend on the number of its adjacent nodes, but also on their value of centrality. Bonacich (1972) suggested that the eigenvector of the largest eigenvalues an adjacency matrix could make a reliable network centrality measure. Unlike degree, which weights every contact equally, the eigenvector weights ties with others according to their centralities. 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. Eigenvector centrality is defined as the principal eigenvector of the adjacency matrix defining the network (Bonacich 2007). Eq. (4) describes eigenvector centrality v in two equivalent ways, as a matrix equation and as a sum. The centrality of a vertex is proportional to the sum of the centralities of the vertices to which it is connected. λ is the largest eigenvalue of A and n is the number of vertices:

$$\lambda \mathrm{v} = \mathrm{A}\mathrm{v} \qquad \lambda \mathrm{v}_\mathrm{i} = \sum_{\mathrm{j}=1}^{\mathrm{n}} \mathrm{a}_{\mathrm{ij}} \mathrm{v}_\mathrm{i}, \quad \mathrm{i} = 1, \dots, \mathrm{n} \tag{4}$$

Where A is the adjacency matrix of the graph, lambda (λ) is a constant (the eigenvalue) and v is the eigenvector. The idea is that even if a node influences just one other node, who subsequently influences many other nodes (who themselves influence still more others), then the first node in that chain is highly influential. (Bonacich 2007)

Another way of interpreting the walk-based measures is that a person's centrality should be a function of the centrality of the people he or she is associated with. In other words, rather than measure the extent to which a given actor "knows everybody", the extent to which the actor "knows everybody who is anybody" should be measured. (Borgatti and Everett 2006)

2.1.2.5. Information Centrality

Freeman's (1979) betweenness centrality encompasses the betweenness counts focusing only on geodesics of the given network paths. This leads to the neglect of the paths with distances greater than the minimum path length attained by the geodesics. In communication relations, actors might choose message paths that are longer than the geodesics under the influence of strong reliability reasons or intrinsic system characteristics. Information might be deliberately channeled through various communication paths created via many intermediaries (Stephenson and Zelen 1989). In these form of networked communications, information robustness in a network with its all-possible paths is vital where information might take an indirect route.

The index of centrality concept developed by Stephenson and Zelen (1989) comprises all the paths between network actors by assigning a weight coefficient to the each path depending on its length. In this framework, a weighted function of the paths is calculated using the inverses of the lengths of the paths as weights. Geodesics are given as weights of unity, while longer paths than the geodesic length receive smaller weights based on the contained information (Wasserman and Faust 1994). The information of a path is defined quite simply as the inverse of its length.

The concept of information is used extensively both in the communication and the statistical estimation theory. Information is statistically defined as the inverse of the variance of an estimator. If an estimator has a small variance, it has large information which is considered positive. The opposite case is also a valid argument where poor estimators with large variances have little information. This approach is applied to the network centrality framework by extending betweenness on geodesics to all possible paths and weighting according to the information contained by these particular paths (Stephenson and Zelen 1989).

Stephenson and Zelen (1989) argued that the information centrality of an actor is a function of all the information for the paths flowing out from that specific actor. The chosen function is the harmonic average. The procedural arguments given below presents Stephenson and Zelen's (1989) approach in mathematical details:

- If (i, j) refers to a pair of nodes in the graph, the first phase is the identification of all possible paths connecting i and j. Suppose that there are k_{ij} paths connecting i and j; accordingly the existing paths for (i, j) are $P_{ij}(1), P_{ij}(2), \dots, P_{ij}(k_{ij})$.
- Then topological distance (d_{ij}) is defined for all the denoted paths, which is the sum of the number of existing links on the specified path. Distance calculations, i.e. $d_{ij}(1), d_{ij}(2), ., d_{ij}(k)$,are done for the each path $P_{ij}(1), P_{ij}(2), ., P_{ij}(k_{ij})$.

- The information measure between two nodes i and j (I_{ij}) is defined as the reciprocal of the topological distance (d_{ij}) between them: $I_{ij} = 1/d_{ij}$. The information of node (i) is the sum of all information content with regard to all other nodes $(I_{i1}, I_{i2} \dots \dots . I_{in})$ in the network: $\sum_{j=1}^{n} \frac{1}{I_{ij}}$; n indicating the total number of nodes in the network.
- The centrality information measure (I_i) of node (i) is defined by the harmonic average of the information flowing from (i) to the all other nodes in the network:

$$\mathrm{I_i} = \frac{\mathrm{n}}{\sum_{\mathrm{j=1}}^{\mathrm{n}} {}^{1}/_{\mathrm{I_{ij}}}} \tag{5}$$

Stephenson and Zelen (1989) stated that the information centrality calculations could be practically conducted by simply inverting an incidence matrix rather than the principal information centrality formula (Eq. 1) as previously introduced. A crucial component of this later formula is the sum of the strengths or values for the lines incident with a node (Wasserman and Faust 1994). For a network with n nodes, n×n matrix $A = (a_{ij})$ is defined:

$$\mathrm{a_{ii}} = 1 + (\text{sum of the weights of all the links of node i}), \tag{7}$$

and off-diagonal elements

$$a_{ij} = \begin{cases} 1 & \text{if nodes } i \text{ and } j \text{ are not linked} \\ 1 - w_{ij} & \text{if nodes } i \text{ and } j \text{ are linked.} \end{cases} \tag{8}$$

a_{ii} Represents the diagonal elements of matrix A; a_{ij} represents the off-diagonal elements of matrix A; w_{ij} is the weight of the link between the nodes i and j. Then, information centrality is calculated by inverting the matrix A. The matrix $C \equiv A^{-1}$ is defined: $C = (c_{ij})$.

For the information measure, two intermediate quantities are required. These are $T = \sum_{j=1}^{n} c_{jj}$ and $R = \sum_{j=1}^{n} c_{ij}$, where T is the sum of the diagonal entries of the matrix, while R is any one of the row sums (all the row sums are equal). With these two quantities, and the elements of C, finally the information centrality I_i for node i is calculated as follows:

$$I_i = \frac{1}{c_{ii} + (T - 2R)/n} \tag{9}$$

2.1.3. Network Centrality Classification

The graph-theoretic typology presented by Borgatti and Everet (2006) is a partial answer to the commonly-asked question of how to choose among centrality measures. The typology essentially divides measures into groups that are more competitive with each other than with other measures. According to Borgatti and Everet (2006) the measures within the same box in Table 2.2. are similar enough on key attributes that they can be thought of as competitive, i.e., as potentially substitutable alternatives for each other. Among measures within each box, it is reasonable to ask which is better. In contrast, measures in different boxes differ in fundamental ways, and are perhaps best viewed as complementary.

Table 2.2. Cross-classification of centrality measures.
(Source: Borgatti and Everett 2006)

	Radial	Medial
Volume	Freeman degree, Bonacich eigenvector,	Freeman betweenness,
Length	Freeman closeness, Stephenson-Zelen information	

Borgatti and Everet (2006) explains that Freeman degree, Freeman betweenness and Bonacich eigenvector centrality measures count the number or volume of walks (of some kind) joining each node to all others. They call these measures as *volume measures*. Freeman closeness and Stephenson-Zelen information centrality measures assess the *lengths* of the walks that a node is involved in. They call these measures as *length measures*. The distinction between volume measures and length measures forms a classificatory dimension, which they call *Walk Property*. It refers to what property of paths (their number or their length) is being measured (Borgatti and Everett 2006).

All of the measures considered so far except Freeman Betweenness centrality assess walks that emanate from or terminate with a given node. Borgatti and Everett

(2006) call these centrality measures as *radial* measures. Another class of centrality measures exists which are based on the number of walks that pass through a given node. They are called *medial* measures. The distinction between radial and medial measures forms another classificatory dimension, which is *Walk Position.*

2.1.4. Network Centralization

Network centralization is a key measure that reflects the distribution of relationships through the network. Centralization provides a measure on the extent to which a whole network has a centralized structure.

The general procedure involved in any measure of graph centralization is to look at the differences between the centrality scores of the most central point and those of all other points. Centralization, then, is the ratio of the actual sum of differences to the maximum possible sum of differences. The three different ways of operationalizing this general measure which Freeman discusses follow from the use of one or other of the three concepts of point centrality (Scott 2000).

In a highly centralized network, a small percentage of the nodes will have a high percentage of relationships with other nodes in the network. In contrast, a network with low centrality will have relatively equal distribution of relationships through the network.

According to Ruan et al. (2012) network centralization is used to measure the overall network as a whole, showing how unequal the distribution of centrality is. In a highly centralized network most of the actual ties are connected to a central node or connections made mainly between certain units.

2.1.5. Density

The density of a binary network (where the values of 0 or 1 are ascribed) is the number of actual ties divided by the maximum number of possible ties in the network. For a valued network density is defined as the sum of the ties divided by the number of possible ties. The density value varies between 0 (no connection in the network) and 1 (every node are interrelated) and if it is close to 1 the network is said to be **dense**, otherwise it is **sparse.** Equation is like following;

$$Density(\Delta) = \frac{l}{[n*(n-1)]} \quad (6)$$

Δ Represents the density of the network; l represents the number of existent lines; n is the number of existent nodes (Park et al. 2011).

Hanneman and Riddle (2005) suggested that density of a network give insights into such phenomena as the speed at which information diffuses among the nodes, and the extent to which actors have high levels of social capital and/or social constraint.

Centralization and density are important complementary measures. Whereas density describes the general level of connectedness in a network; centralization describes the extent to which this connectedness is organized around particular focal nodes (Scott 2000).

2.2. Current Status of Research on SNA in Construction

Since its emergence from sociology in 1930s (Moreno 1934), SNA have been gaining wide attention from theoretical exploration to practical application. However, the research topics explored by the SNA method in the construction domain are diverse and demand an analysis of the SNA related issues. An analytical review of the reported literature within the SNA domain may lead the way for future researchers to gain an understanding of the topic and to conduct associated research more thoroughly and efficiently. Retrieval from academic journals can be regarded as the most effective approach for particularly new researchers, to gain an in-depth insight into the research trends about a specific topic. Tsai and Wen (2005) stated that a systematic analysis of papers published in academic journals would help researchers explore the current status and future trend of a chosen topic. However, in the field of construction, no such critical analysis of SNA research has been undertaken to date. Therefore, this paper attempts to comprehensively review the SNA related literature in the three leading journals on construction management and to investigate the research trend of SN related studies. The first paper on SNA was published in construction journals in 1998. Therefore SNA literature review is conducted from 1998 to 2014, inclusive.

SNA is a body of theory and methodology for the analysis of systems as networks of relationships. It is defined as a broad strategy for investigating social structure rather than formal theory (Otte and Rousseau 2002).

The application of SNA is relatively new in the construction industry. Nevertheless, the volume of SNA research in construction management has increased radically in recent years, as it has in many disciplines. In this paper, we first review and analyze the emerging SNA paradigm in construction research. We begin with a conventional review of recent SNA research published in recognized construction management related journals, namely *Journal of Construction Engineering and Management, ASCE (JCEM); Construction Management and Economics (CME); and Engineering, Construction and Architectural Management (ECAM).* (Search code was "social network analysis" in the title, abstract or keywords.)

Next, we analyze this research, developing a set of dimensions along with network studies conducted, including theoretical contribution, units of analysis and adopted software. 21 papers in relation to SNA studies published in above mentioned journals between 1998 and 2014 are presented in Table 2.1.

Table 2.3. SNA in Construction Management Research.

AUTHORS and *YEAR*	**ARTICLE in *JOURNAL***	**THEORETICAL CONTRIBUTION**	**UNITS of ANALYSIS**	**ANALYSIS via *SOFTWARE***
Loosemore, M. 1998	Social network analysis: using a quantitative tool within an interpretative context to explore the management of construction crisis in ECAM	Provides an universal model for construction crisis by association of quantitative and qualitative methods	The interactions between actors in a leisure centre project during a construction crisis.	Degree centrality, closeness centrality, betweenness centrality via UCINET
Loosemore, M. 1998	The influence of communication structure upon management efficiency *in CME*	Provides an investigation of the relationship between the structure of communication networks which specifically emerge in response to construction crisis and the efficiency of crisis management efforts	Four construction projects	Degree centrality, closeness centrality, betweenness centrality via UCINET 1
Thorpe, T. Mead, S. 2001	Project-specific web sites: friend of foe? in JCEM	Provides a theoretical background on project-specific Web sites (PSWS), and describes SNA that was used to understand how PSWS affects information push and pull	Three construction projects using project-based internet systems	Centrality, via Krackplot tool via UCINET
Pryke, S. 2004	Analyzing construction project coalitions: exploring the application of social network analysis in CME	Provides an exploration of the application of SNA as a new quantitative approach is construction industry	Application to the analysis of UK construction procurement	Point centrality, degree centrality, closeness centrality
Pryke, S. 2005	Towards a social network theory of project governance in CME	Provides an exploration of a very specific framework for the examination of the governance of construction coalitions using SNA	Four UK construction projects	Network density, actor point centrality, via UCINET 6
Hossain, L. 2009	Communications and coordination in construction projects in CME	Provides an exploration of the association between network centrality and coordination for a construction project	Networks, Relationship between actors Companies	Network centrality via UCINET 6
Chinowsky, P., Diekmann, J., O'Brien, J. 2010	Project Organizations as Social Networks in JCEM	Provides a social network model of construction introduced a dual-focus approach for enhancing professional trust and strong communication	Networks consist of actors from 4 engineering compa-nies that engaged in both management of design and construction	Network density, centrality, betweenness and power via UCINET
El-Sheikh, A., Pryke, S. 2010	Network gaps and project success in CME	Provides a combined application of gap analysis and SNA can help practitioners to exceed client expectations	Two under-construction projects in the Academy Program of the DfES	Network density, centrality betweenness and power via UCINET 6
Park, H., Seung H.H., Rojas, E.M., JeongWook, S., Jung, W. 2011	Social Network Analysis of Collaborative Ventures for Overseas Construction Projects in JCEM	Provides collaboration strategies considering relevant network patterns and their different levels of performance in the different networks	389 cases of overseas construction projects executed by Korean firms that involved collaboration from 1990 to 2006.	Density, direct/indirect ties, in-degree/out-degree betweenness and closeness centrality triad, via Pajek via NetDraw

(Cont. on next page)

Table 2.3. (Cont.)

Larsen, G.D. 2011	Understanding the early stages of the innovation diffusion process: awareness, influence and communication networks in CME	Provides the use of SNA to map the communication networks in order to understand how actors become aware of an innovation	389 cases of overseas construction projects executed by Korean firms that involved collaboration from 1990 to 2006.	Density, direct/indirect ties, in-degree/out-degree betweenness and closeness centrality triad via Pajek via NetDraw
Ruan, X., Ochieng, E.G, Price, A.D.F., Egbu, C.O. 2012	Knowledge integration process in construction project a social network analysis approach to compare competitive and collaborative working in CME	Provides a network perspective to analyze the knowledge integration for understanding the affects of collaborative working on overall project performance	A comparison of knowledge supporting networks from two different procurement systems: overall partnering and project partnering	Network centralization, network density, normalized degree centrality, via UCINET
Alsamadani, R. Hallowell, M. Javernick-Will, A. N. 2012	Measuring and modeling safety communication in small work crews in the US using social network analysis in CME	Provides an exploration of SNA as a tool to measure and model safety communication patterns and determine effective and ineffective safety networks	Nine construction firms in the Denver Metropolitan region of the US	Density, centrality, betweenness, via UCINET via NetDraw
Wambeke, B.W. Liu,M. Hsiang, S. M. 2012	Using Pajek and centrality analysis to identify a social network of construction trades in JCEM	Provides a SNA application as a method of identifying organizational SN of trades and analyzing the network to identify the key trades of a construction project	A data center construction project involving a general contractor and 43 trades	Degree centrality, Eigenvector centrality, via Pajek
Comu, S., Iorio, J., Taylor, J. E., Dossick, C.S., 2013	Quantifying the Impact of Facilitation on Transactive Memory System Formation in Global Virtual Project Networks in JCEM	Provides a SNA application for measuring the impact of facilitators on performance of global virtual project networks engaged in the task work	Two facilitated and two non-facilitated global virtual project networks	Dyadic task relation, Subgroup formation, via UCINET via Statnet
Solis, F. Sinfield, J. V. Abraham, D. M. 2013	Hybrid approach to the study of inter-organization high performance teams in JCEM	Provides a hybrid approach that proposes the use of SNA as a complementary methodology to understand the interorganizational teamwork of construction projects	A cardiovascular center project in the state of California	Density, centrality, structural equivalence via Pajek
Alsamadani, R. Hallowell, M. Javernick-Will, A. Cabello, J. 2013	Relationships among Language Proficiency, Communication Patterns, and Safety Performance in Small Work Crews in the United States in JCEM	Provides an investigation of safety-related knowledge exchange of the crew level using SNA	14 construction crews, each with less than 40 workers in the Denver Metro Region of the United States	In-degree/out-degree centrality, in and out betweenness centrality via UCINET
Zhang, L. He, J. Zhou, S. 2013	Sharing Tacit Knowledge for Integrated Project Team Flexibility: Case Study of Integrated Project Delivery in JCEM	Provides a theoretical guide for the integrated project team to improve its ability to survive by sharing tacit knowledge in the dynamic environment of construction projects	20 team members from the integrated project team	Network density, cohesiveness, degree centralization, betweenness centralization via UCINET 6

(Cont. on next page)

Table 2.3. (Cont.)

Aljassmi, H. Han, S. Davis, S. 2014	Project Pathogens Network: New Approach to Analyzing Construction-Defects-Generation Mechanisms in JCEM	Provides a methodology that is deliberated to providing both mathematical and visualization analyses needed to address the complex mechanisms of defect generation.	7 interviewees including four clients, two engineers from the contracting company and one engineer from the designer company of 4 residential projects in Dubai	Closeness, reachability via UCINET
Arriagada D., R. Alarcón C., L. 2014	Knowledge Management and Maturation Model in Construction Companies *in JCEM*	Provides a model that allows an adequate organizational characterization and uses the model to recommend paths of knowledge maturity and management inside construction companies	9 key actors in a construction project	Out-degree/in-degree centrality Network centralization via UCINET 6
West, J. 2014	Collaborative Patterns and Power Imbalance in Strategic Alliance Networks *in JCEM*	Provides an investigation of collaborative patterns in the resources and mining sector and their impact on firm performance and resource quality using social network analysis.	9 mining (specialist) companies and 32 generalist companies	Density, direct ties, indirect ties, degree centrality, betweenness centrality, closeness centrality, point connectivity, Bonacich power, and eigenvector centrality via UCINET 6
Sanaei, M., Javernick-Will, A.N., Chinowsky, P. 2013	The influence of generation on knowledge sharing connections and methods in construction and engineering organizations headquartered in the US	Provides an analysis of existing knowledge sharing connections based upon generational attributes to determine whether generation influences the knowledge sharing structure within the organization	734 employees within three communities of practice in two construction and engineering organizations in the US	Randomized controlled trial (RCT) Chi-Square Test via UCINET 6 via NetMiner

2.3. Coordination Theory

Coordination philosophy was introduced in 1916 by Henry Fayol and his research was published as a book. This book was translated in English by Constance Storss in 1949 (Wood and Wood 2002). Fayol (1949) argues that coordination plays a significant role in management activity. For Fayol (1949), to coordinate is to "harmonize all the activities of a concern so as to facilitate its working and its success" (Lamond 1998). His research emphasizes that coordination provides the best cooperation between team members; hence improving the communication, integration and team working. Many scholars and researchers have inspired from the initial findings of Fayol (1949) and have focused on investigating the concept of coordination. The concepts differed from each other depending on how coordination is perceived by research scholars in their empirical studies.

Van de Ven (1976) claims that coordination is the mode of linking together different parts of an organization to perform a set of collective tasks.

Thomas Malone (1988) defined coordination as "the additional information processing performed when multiple, connected actors pursue goals that a single actor pursuing the same goals would not perform." This definition of coordination implies the following components: (1) a set of (two or more) *actors,* (2) who perform *tasks,* (3) in order to achieve *goals.* The components of coordination and the coordination processes associated with them are summarized by Malone and Crowston (1994) are presented in Table 2.4.

Table 2.4. Components of Coordination.
(Source: Malone and Crowston 1990)

Components of coordination	Associated coordination processes
Goals	Identifying goals
Activities	Mapping goals to activities (e.g., goal decomposition)
Actors	Selecting Actors
	Assigning activities to actors
Interdependencies	Managing interdependencies

Malone (1988) also defines coordination as a body of principles about how the activities of separate actors can be coordinated. Crowston and Malone (1990) redefined the coordination as "body of principles about how activities can be coordinated, that is, about how actors can work together harmoniously". They explained that on the aspects of the element of coordination that was implied by the word “harmoniously” were interdependencies.

Crowston and Malone’s definitions (1990) of coordination give a prominent role to interdependence. According to these definitions if there is no interdependence, there is nothing to coordinate. Depending on this statement Malone and Crowston's Theory of Coordination is extended by focusing on the kinds of interdependencies between activities and possible management of different interdependency types. Table 2.5

presents a preliminary list of types of interdependencies and coordination processes that can be used to manage them proposed by Malone and Crowston (1990).

Table 2.5. Examples of kinds of interdependencies.
(Source: Malone and Crowston 1990)

Kinds of Interdependence	**Common object**	**Examples of coordination process for managing interdependencies**
Prerequisite constraints	Output of one activity which is required by the next activity	Ordering activities, moving information from one activity to the next
Shared resources	Resource required by multiple activities	Allocating resources
Simultaneity	Time at which more than one must occur	Synchronizing activities

2.3.1. Dependency Types

The process of coordination was broken down into four key coordination processes as defined by Malone and Crawston (1994). They further defined the dependency types of the Theory of Coordination by characterizing and identifying the dependencies and coordination process that can be used to manage them. (Table 2.6.) Since coordination is managing dependencies as defined in the study of Malone et al (1994), if there is no interdependence, there is nothing to coordinate.

Managing shared resources: It is defined as the management and control of the limited resources (e.g., money, storage, space, or actor's time) to be intimately connected with personal and organizational power. Malone et al. (1994) state that a resource allocation process is needed to manage interdependencies among multiple activities sharing some limited resource. Hossain (2009a) interpreted this resource allocation process as introducing or suggesting a person to perform a task.

Managing Producer/Consumer Relationship: This is a common type of dependency between activities is a "producer/consumer" relationship which is usage of a product of one activity by another activity. Hossain (2009a) interpreted this dependency in his study as the creation or dissemination of information. The three kinds of dependencies of producer/consumer relationships defined by Malone and Crowston (1994) are as follows:

Prerequisite constrains**:** It is a very common dependency between "producer" activity and a "consumer" activity is that the producer activity must be completed before the consumer activity can begin.

***Transfer*:** It is a process that when one activity produces something that is used by another activity, the thing produced must be transferred from the "producer" activity to "consumer" activity.

***Usability*:** It is a dependency that must often be managed in a producer/ consumer relationship is that whatever is produced should be usable by the activity that receives it.

Managing simultaneity constrains: This is a type of dependency among activities is that they need to occur at the same time (or cannot occur at the same time). Hossain (2009a) interpreted these dependencies in his study as synchronizing tasks between actors, taking possible times for an event, allocating a time for a particular event and passing information about the time of an event.

Managing task/subtask dependencies: It is a common type of dependency among activities is that group of activities are all "subtasks" for achieving some overall goal.

Table 2.6. Examples of Common Dependencies between Activities and Alternative Coordination Processes for Managing Them (Source: Malone and Crowston 1994).

Dependency	**Examples of coordination processes for managing dependency**
Shared resources	"First come/first serve", priority order, budgets, managerial decision, market-like bidding
Task assignments	(same as for "Shared resources")
Producer / consumer relationships	
Prerequisite constraints	Notification, sequencing, tracking
Transfer	Inventory management (e.g., "Just In Time", "Economic Order Quantity")
Usability	Standardization, ask users, participatory design
Design for manufacturability	Concurrent engineering
Simultaneity constraints	Scheduling, synchronization
Task / subtask	Goal selection, task decomposition

According to Malone and Crowston's Coordination Theory (1990) actors in organizations face coordination problems that arise from dependencies that constrain how tasks can be performed (Crowston 1997). To overcome these coordination problems they suggested that actors must perform additional activities which are called coordination mechanisms (Malone et al. 1990).

Crowston (1997) used the framework developed by Malone and Crowston (1994). He applies coordination theory to show how task processes can be decomposed, documented and altered to create new forms of organizing work.

Malone and Crowston's research (1994) on coordination theory will serve as the basis of our theory study and later the empirical analysis. The coordination theory literature review summary is presented in Table 2.7.

Table 2.7. Coordination Theory Literature.

AUTHORS and *YEAR*	**ARTICLE and *SOURCE***	**THEORETICAL CONTRIBUTION**
Fayol, H. (1949)	General and Industrial Management, *Book-Pitman.*	Provides an identification of six groups of activities or essential functions to which all industrial undertakings give rise - technical, commercial, financial, security, accounting, and managerial activities.
Van de Ven, A.H. (1976)	On the nature, formation, and maintenance of relations among organizations, *Academy of Management Review*	Provides a theory for explaining how and why relationships among two or more human service organizations voluntarily emerge, and how they function over time.
Malone, T. W. (1988)	What is coordination theory? *National Science Foundation Coordination Theory Workshop*	Provides a body of scientific theory, which is called "coordination theory", about how the activities of separate actors can be coordinated.
Malone, T. W. , Crowston, K. (1990)	What is coordination theory and how can it help design cooperative work systems? *Proceedings of the 1990 ACM conference on Computer-supported cooperative work*	Provides the interdisciplinary study of coordination perspective that focuses on how people work together now and how they might do so differently with new information technologies.
Malone, T. W. , Crowston, K. (1994)	The interdisciplinary study of coordination, *ACM Computing Surveys*	Provides a key insight of the framework of coordination as the process of managing dependencies among activities and identifies the different kinds of dependencies and coordination processes that can be used to manage them.
Herbsleb, J. D. Mockus, A. (2003)	Formulation and preliminary test of an empirical theory of coordination in software engineering, *ACM SIGSOFT Software Engineering Notes*	Provides key themes emergent from the qualitative analysis that highlight integral cooperative and social aspects to software development: proximity, artifact management patterns and uncooperative behaviors.
Crowston, K. Rubleske, J. Howison, J. (2006)	Coordination theory: A ten-year retrospective *Book-Human-Computer Interaction in Management Information Systems*	Provides the interdisciplinary study of coordination perspective that focuses on how people work together now and how they might do so differently with new information technologies.
Begel, A. (2008)	Effecting change: Coordination in large-scale software development, *Proceedings of the 2008 international workshop on Cooperative and human aspects of software engineerin*g	Provides an identification of the coordination problems that go along with differences of location, time zone and culture and possible solutions for those problems.
Deng, X., Chen, T., Pan, D. (2008)	Organizational Coordination Theory and Its Application in Virtual Enterprise *Book-Research and Practical Issues of Enterprise Information Systems II*	Provides a summary of the frame of the organizational theory by refining the meaning of coordination and studies the virtual enterprise's coordination mechanism in strategy and task layers
Tellioglu, H. (2010).	Coordination of work: towards a typology, *Proceedings of the 11th International Conference on Computer Systems and Technologies*	Provides identification of certain types of coordination and their characteristics based on the study of different coordination theories and approaches.
Dan, S. (2013)	Theorizing coordination: towards a novel theoretical framework EGPA Permanent Study Group VI	Provides a theoretical overview of coordination and argues the most prominent theories in the field of political science and public administration

2.4. Coordination in Construction

Hai et al. (2012) stated that in order to understand the coordination in construction, coordination philosophy interpreted into the construction management processes must be explored in detail. Hai et al. (2012) explored the characteristics and nature of construction in response to the coordination environment and coordination principles adopted in the construction industry.

Xue et al. (2007) explained the process of coordination starting from demands of the client early in the conceptual phase, then design and construction phases to maintenance, replacement and eventually decommissioning of building. According to Hossain (2009a) the process of construction, depending on the complexity of the finished structure, requires a high level of coordination among all the professionals and trade persons from design office to the construction site, until the project is finished. The complex process of construction project consists of numerous activities. Therefore, the ideal coordination environments rarely exist. The coordination problems of the construction industry have attracted many researchers to focus and facilitate coordination in construction (Pocock et al. 1996; Chitkara 1998; Saram and Ahmed 2001; Saram 2002; Iyer and Jha 2005; Hossain et al. 2006; Kubicki et al. 2006; Jha and Iyer 2007; Jha and Misra 2007; Hossain 2009a,b; Hossain and Wu 2009; March 2009).

Kubicki et al. (2006) state that coordination is a vital activity in the building construction process. They also emphasize that coordination and cooperation are a prerequisite to build up effective and efficient processes of construction. Kubicki et al. (2006) also suggest that the success of a construction project is based on the relations between project participants. Pocock et al. (1996) and Higgin and Jessop (1965) have the same opinion that sufficient level of interaction between designers and constructors bring success to the overall project performance. Therefore, coordination is essential to improve the separate working environments and build up teamwork by integrating design and construction phases (Higgin and Jessop 1965). Iyer and Jha (2005) support that coordination between project participants is the most significant factor with the greatest impact on project success and cost performance.

Another issue studied by researchers is focusing on the changing nature of temporary construction project. Construction projects have limited time frames, hence it leads frequent changes of project participants and resources. March (2009) claims that the changing of construction participants has brought the impact of less opportunity for them to develop long-term relationships. Chitkara (1998) states that due to the nature of temporary project organizations, incorporating the coordination principles in the various departments is essential. He also points out that coordination between project participants may facilitate information exchange, communication in the construction project and interaction of the participants.

On the other hand, Badiru (2012) suggests organizational structure of construction is to get developed based on the project. In each organization, team members consist of varying roles for different functions and other interlinked roles. Hence, a successful construction organization requires a high level of coordination among project team members throughout the construction process (Badiru 2012). The study of Hossain (2009a) supports this idea and points out that high level of coordination is a prerequisite for achieving higher degree of operational efficiency. Hossain et al. (2006) and Hossain (2009a) argue that existing coordination theory defined by Malone and Crowston (1990, 1994) allows the application of social network analysis. He also points out that this analysis method makes available to measure coordination quantitatively. Hossain and Wu (2009) have focused in determining the effects of network centrality and coordination. They investigated the differences in coordinative activity between individuals with high and low level of network centralization (Hossain and Wu 2009).

Soh and Wang (2000) claims that high level of coordination of all project teams from office to construction site is crucial for project success. Saram (2002) focuses on how day-to-day coordination on a construction project is achieved and he points out that this approach provides a better understanding of the level of coordination in a construction project. His study (2002) has identified the total of 64 construction activities and these have been ranked in accordance with its relative importance in order to explore day-by-day coordination. Jha and Mısra (2007) further developed the findings of Saram and Ahmed (2001). Jha and Mısra (2007) argue that coordination contributes to the outcome of the project. This study identified coordination activities and ranked them based on four project performance criteria (schedule, cost, quality and dispute) in order to get a better understanding of relative importance of the activities. The coordination studies in the construction management research are presented in Table 2.8.

Table 2.8. Coordination in Construction Management Research.

AUTHORS and *YEAR*	**ARTICLE and *SOURCE***	**THEORETICAL CONTRIBUTION**
Higgin, G. Jessop, N. (1965)	Communications in the Building Industry: The Report of a Pilot Study, *Routledge*	Provides a picture of the industry as reflected by a Sociological Operational Research mirror and an exploration of what lies behind the industry's communication difficulties
Pocock, J. B., Hyun, C. T., Liu, L. Y., Kim, M. K. (1996)	Relationship between project interaction and performance indicators, *JCEM*	Provides a method for measuring a project's degree of interaction (DOI), and verifies the relationship between DOI and performance indicators such as cost growth, schedule growth, and number of modifications
Chitkara, K. K. (1998)	Construction Project Management, *Tata McGraw-Hill*	Provides a thorough understanding of construction project management topics with the help of various concepts, practical insight, real-life examples and skills to execute large and small projects
Soh, C. K. Wang, Z. (2000)	Parametric coordinator for engineering design, *Journal of Computing in Civil Engineering*	Provides a novel approach that can facilitate the coordination of design information through managing design changes with the help of a parametric coordinator
Saram, D. D. d., Ahmed, S. M. (2001)	Construction coordination activities: What is important and what consumes time, *Journal of Management in Engineering*	Provides identification of which activities are performed to achieve coordination and which among those are the most important and more time consuming for a construction coordinator
Saram, D. D. d., (2002)	Measuring the quality of contractors' co-ordination activities during the construction process, *Doctoral degree of Civil and Structural Engineering (Management) Thesis, Hong Kong*	Provides an exploration of Critical Incident Technique as a practical method for measuring the quality of construction co-ordination processes
Iyer, K. C. Jha, K. N. (2005)	Factors affecting cost performance: evidence from Indian construction projects, *International Journal of Project Management*	Provides identification of critical success factors and the factors affecting cost performance of Indian construction Projects
Kubicki, S. Bignon, J. C. Halin, G. Humbert, P. (2006)	Assistance to building construction coordination–towards a multi-view cooperative platform, *Electronic Journal of Information Technology in Construction*	Provides a new assistance tool for coordination (building construction multi-view interface) taking into account the analysis of coordination modes and IT potentialities
Hossain, L., Wu, A., Chung,K. K. S. (2006)	Actor centrality correlates to project based coordination, Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work	Provides exploration of how project team members interact when working towards a common goal by using network centrality concepts and coordination theory
Jha, K. N. Misra, S. (2007)	Ranking and classification of construction coordination activities in Indian projects, *CME*	Provides identification of the most important coordination activities corresponding to schedule, cost, quality and no dispute performance criteria of civil engineering projects.
Xue, X., Wang, Y., Shen, Q., Yu, X. (2007)	Coordination mechanisms for construction supply chain management in the Internet environment, *International Journal of Project Management*	Provides a general framework or Construction Supply Chain (CSC) coordination and identification of suitable coordination mechanisms that can promote the effective coordination in CSC
Hossain, L. (2009a)	Communications and coordination in construction projects, *CME*	Provides an exploration of the association between network centrality and coordination for a construction project

(Cont. on next page)

Table 2.8. (Cont.)

Hossain, L. (2009b)	Effect of organizational position and network centrality on project coordination, *International Journal of Project Management*	Provides an exploration of the effect of the employee's organizational position and network centrality on project based coordination
Hossain, L. , Wu, A. (2009)	Communications network centrality correlates to organizational coordination *International Journal of Project Management*	Provides an exploration of correlation between actor centrality and project-based coordination
March, C. (2009)	Business Organization for Construction, *Taylor & Francis*	Provides the tools required to skillfully and successfully operate a business in today's construction industry
Shen, F., Chang, A. (2011)	Exploring Coordination Goals of Construction Projects, *Journal of Management in Engineering*	Provides an exploration of coordination goals by proposing a way of evaluating goal achievements
Hai, T. K., Yusof, A. M., Ismail, S., Wei, L. F. (2012)	A Conceptual Study of Key Barriers in Construction Project Coordination, *Journal of Organizational Management Studie*s	Provides an outline of the basic ideas to improve existing poor performance of construction through emphasizing concerns towards the key barriers of coordination
Badiru, A. B. (2012)	Triple C model of project management: Communication, cooperation, and coordination, *CRC Press*	Provides practical steps and techniques for tracking, managing, and controlling project costs as well as implementing the project management body of knowledge
Dogan, S. Z. Gunhan, S. Erbasaranoglu, B. (2012)	Coordination Process and Network Centrality in ISGI Airport's Wayfinding Project, *Construction Research Congress 2012, ASCE*	Provides exploration of the effects of network centrality on coordination performance.
Dogan, S. Z., Arditi, D. Gunhan, S. Erbasaranoglu, B. (2013)	Assessing Coordination Performance Based on Centrality in an Email Communication Network, *Journal of Management in Engineering*,	Provides a procedure for monitoring the coordinative performance of project participants.

CHAPTER 3

RESEARCH METHODOLOGY

This chapter contains three subsections, namely the Enron dataset, the case studies and the procedure. Enron e-mail dataset is used as the coordination and network analysis dataset. The case studies are Dabhol Power Corporation and Azurix Corporation. The procedure is the coordination and communication analysis and network centrality measurements involving sentence extraction, assigning the coordination weight and measuring centrality through the communication matrix.

3.1. Enron Email Dataset

The Enron Corporation was an energy trading, natural gas and electric utilities company based in Houston, Texas, which employed around 21 000 people by mid- 2001. Enron was involved in the distribution of electricity and gas throughout the United States and the development and operation of power plants, pipelines, and other infrastructure worldwide. After the expansion of business front into energy trading and securities trading, Enron was named "America's Most Innovative Company" by Fortune magazine for five consecutive years from 1996 to 2000. On December 2, 2001 the Enron Company filed a bankruptcy after a wave of accounting scandals and the US Justice Department investigated whether Enron defrauded investors by concealing information about its finances.

E-mail communication logs from the Enron Corporation between 1997 and 2002 were made public during their legal investigation by Federal Energy Regulatory Commission (FERC). These logs became known as Enron dataset or Enron Corpus. The raw Enron corpus contains 619,446 messages belonging to 158 users. The release of the data provided researchers a unique glimpse of e-mail communications inside a major corporation. The Enron corpus has been used for many purposes. Since then, the researchers have created many different versions of the dataset by modifying the corpus according to their needs or using a particular subset of the corpus for different purposes. According to Kessler (2010), the Enron email dataset has been investigated previously to examine email classification and threading behavior among email users (Klimt and

Yang 2004), social networking analysis (Corrada-Emmanuel et al. 2005), document classification (Bekkerman et al. 2005), and linking trends (Shetty and Adibi 2004).

3.1.1. Original Corpus (As distributed by William Cohen)

The original Enron Corpus was distributed by William Cohen in March 2004. This dataset was collected and prepared by the CALO Project and SRI International. This version of the corpus is almost identical to the one made public by the FERC except the attachments of the e-mails. Although the attachments were excluded, it is still a huge corpus, containing 517,431 distinct e-mail messages. The dataset has many duplicate and corrupt messages. It contains e-mail messages exchanged between 151 users mostly senior management of Enron. Every user has a folder named after him/her in the original corpus. Within this folder the individual foldering strategy of the user has been maintained.

3.1.2. Klimt and Yang Corpus

Klimt and Yang (2004) from Carnegie Mellon University were amongst the first people to work on the Enron corpus. They wrote a paper providing a brief introduction and analysis of the dataset. Their goal was to analyze the suitability of this corpus for exploring how to classify messages as organized by a human (Klimt and Yang 2004). For this reason they went through the entire corpus and eliminated the duplicate messages by removing certain folders from each user. These messages were mostly in the computer generated folders and these folders did not appear to be used directly by the users. In their cleaned Enron corpus (2004), there were a total of 200,399 messages belonging to 158 users with an average of 757 messages per user. The cleaned version was one-third the size of the original corpus which means approximately 62 percent of the original corpus is made up of duplicate emails (Naimisha 2008).

3.1.3. Bekkerman Corpus

This version of the Enron Corpus was created by Bekkerman et al. (2004) from the University of Massachusetts. For their research they used a subset of the original

William Cohen corpus (2004). The aim of their study (2004) was to discuss the challenges that arise from differences between email foldering and traditional document classification, comparing the classification techniques on email foldering.

Bekkerman et al. (2004) explored the classification of emails, such as the organization of messages in user-defined folders and thread detection (Diesner and Carley 2005). They used the email directories of seven former Enron employees that have large amount of messages. They removed the non-topical folders and flatten all the folder hierarchies. They also removed folders that contained less than three e-mail messages, since they were very small and would not help either in training or testing. Bekkerman corpus (2004) now contains a total of 273 Folders and 20,581 e-mail messages. The smallest folders contain 3 e-mail messages, whereas the largest folder contains 1398 e-mails (Naimisha 2008).

3.1.4. Corrada-Emmanuel Corpus

The Corrada-Emmanuel corpus was derived from the original Cohen corpus by Corrada-Emmanuel et al. (2005) from the University of Massachusetts, Amherst. He explored the dataset by using the MD5 digest of the body of the emails (Diesner and Carley 2005). He found out that the corpus contains 250,484 unique messages from 147 people. Corrada Emmanuel et al. (2005) did not model email messages that were not received by at least one of the 147 and the total number of email messages traded among these users was 23,488. Corrada-Emmanuel et al. (2005) created various mappings between e-mails within the Enron corpus. These include mappings of e-mails to relative paths, authors and recipients (Naimisha 2008).

3.1.5. Shetty and Adibi Corpus

This version of the corpus was created by Shetty and Adibi (2004) from the University of Southern California. Shetty and Adibi (2004) provide information on quantitative features of the corpus, such as the distribution of the number of emails per user and over time (months, years) (Diesner and Carley 2005). Shetty and Adibi (2004) aimed to investigate the types of inter-personal relationships between Enron employees; who corresponded with whom, the level of communication between top management and other employees. They created a social network that represents 151 Enron employees. In this network each exchange of at least 5 emails between any pair of agents across the entire time range (1998 to 2002) was considered as a link (Diesner and Carley 2005).

Shetty and Adibi (2004) used the corpus distributed by William Cohen (2004) and created a MySQL database for the entire corpus to catalyze the statistical analysis of the data. This corpus contains 252,759 e-mail messages exchanged between 151 users (Naimisha 2008). These e-mail messages are present in around 3000 user-defined folders. Figure 3.1.shows the schema of the database. The Enron database contains four tables namely, Employee List, Message, Recipient Info and Reference Info. The first table contains information of each of the 151 employee. The second table contains the information of the email message the sender, subject, text and other information. The third table contains the recipient's information. It contains the email address of the recipient and the type (To, CC, BCC) in which the message was sent to the recipient. The fourth table contains information of all those messages that have been referenced after being sent once, either as a forward or reply (Shetty and Adibi 2004).

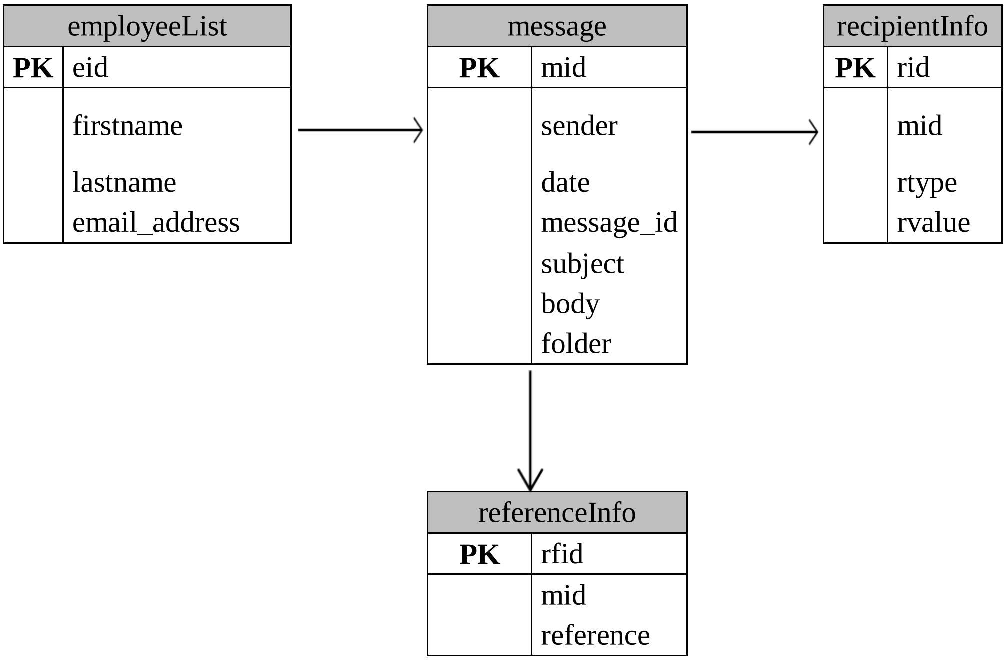
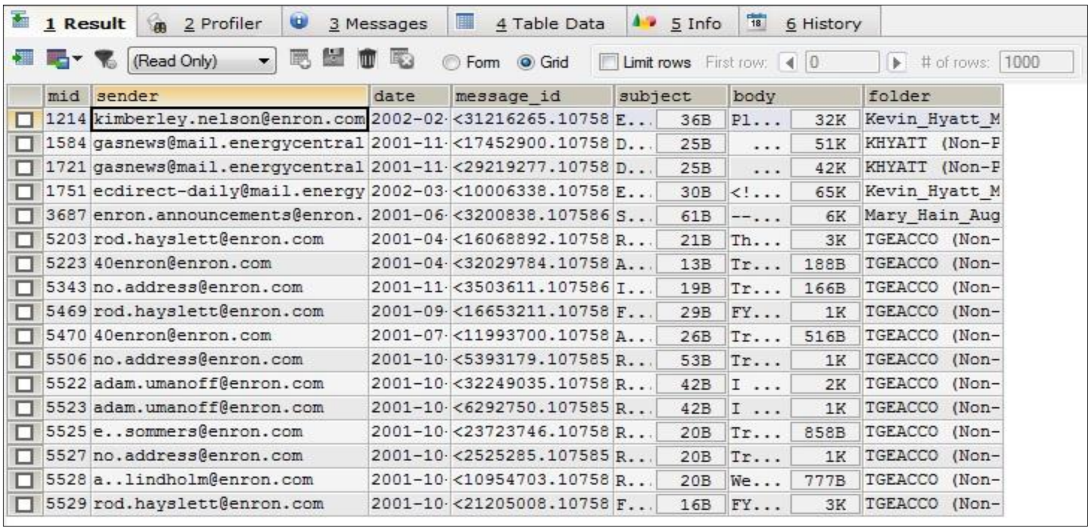


Figure 3.1. Enron Database Schema.
(Source: Shetty and Adibi 2004)

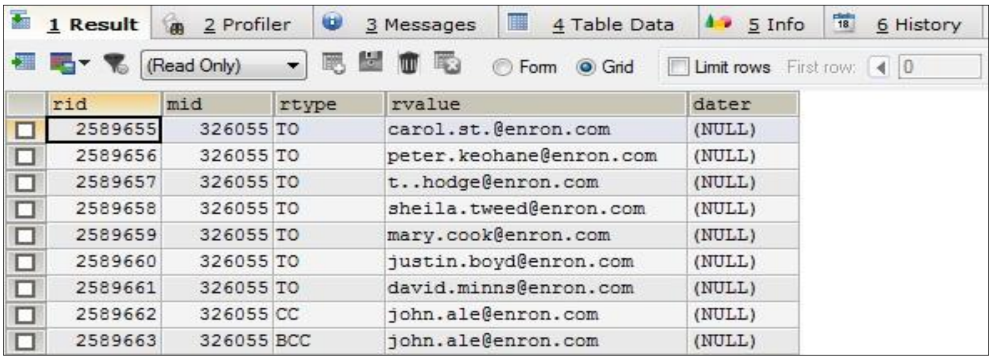
3.1.6. The Dataset of the Study

In this study, the version of the corpus created by Shetty and Adibi (2004) was used for extracting the evidence of coordination and performing network centrality measurements. As the corpus was a structured database, it allows extensible queries to be run on the dataset. This provides flexibility such as partitioning the dataset into e-mail based on project scope (Hossain 2009a). The data is extracted from the MySQL database by writing a query. Figure 3.2. shows the results of the query for the Dabhol Dataset. Figure 3.3. shows the recipient info of an e-mail from the Dabhol Dataset. The version of the corpus created by Shetty and Adibi (2004) is available at http://www.isi.edu/~adibi/Enron/Enron.htm.



mid	sender	date	message_id	subject		body		folder
1214	kimberley.nelson@enron.com	2002-02	<31216265.10758	E...	36B	Pl...	32K	Kevin_Hyatt_M
1584	gasnews@mail.energycentral	2001-11	<17452900.10758	D...	25B	...	51K	KHYATT (Non-P
1721	gasnews@mail.energycentral	2001-11	<29219277.10758	D...	25B	...	42K	KHYATT (Non-P
1751	ecdirect-daily@mail.energy	2002-03	<10006338.10758	E...	30B	<!...	65K	Kevin_Hyatt_M
3687	enron.announcements@enron.	2001-06	<3200838.107586	S...	61B	--...	6K	Mary_Hain_Aug
5203	rod.hayslett@enron.com	2001-04	<16068892.10758	R...	21B	Th...	3K	TGEACCO (Non-
5223	40enron@enron.com	2001-04	<32029784.10758	A...	13B	Tr...	188B	TGEACCO (Non-
5343	no.address@enron.com	2001-11	<3503611.107586	I...	19B	Tr...	166B	TGEACCO (Non-
5469	rod.hayslett@enron.com	2001-09	<16653211.10758	F...	29B	FY...	1K	TGEACCO (Non-
5470	40enron@enron.com	2001-07	<11993700.10758	A...	26B	Tr...	516B	TGEACCO (Non-
5506	no.address@enron.com	2001-10	<5393179.107585	R...	53B	Tr...	1K	TGEACCO (Non-
5522	adam.umanoff@enron.com	2001-10	<32249035.10758	R...	42B	I ...	2K	TGEACCO (Non-
5523	adam.umanoff@enron.com	2001-10	<6292750.107585	R...	42B	I ...	1K	TGEACCO (Non-
5525	e..sommers@enron.com	2001-10	<23723746.10758	R...	20B	Tr...	858B	TGEACCO (Non-
5527	no.address@enron.com	2001-10	<2525285.107585	R...	20B	Tr...	1K	TGEACCO (Non-
5528	a..lindholm@enron.com	2001-10	<10954703.10758	R...	20B	We...	777B	TGEACCO (Non-
5529	rod.hayslett@enron.com	2001-10	<21205008.10758	F...	16B	FY...	3K	TGEACCO (Non-

Figure 3.2. MySQL Screenshot of the Dabhol Dataset.



rid	mid	rtype	rvalue	dater
2589655	326055	TO	carol.st.@enron.com	(NULL)
2589656	326055	TO	peter.keohane@enron.com	(NULL)
2589657	326055	TO	t..hodge@enron.com	(NULL)
2589658	326055	TO	sheila.tweed@enron.com	(NULL)
2589659	326055	TO	mary.cook@enron.com	(NULL)
2589660	326055	TO	justin.boyd@enron.com	(NULL)
2589661	326055	TO	david.minns@enron.com	(NULL)
2589662	326055	CC	john.ale@enron.com	(NULL)
2589663	326055	BCC	john.ale@enron.com	(NULL)

Figure 3.3. MySQL Screenshot of the Recipient Information of an E-mail from Dabhol Dataset.

In this study Dabhol Power Corporation and Azurix Corporation were selected as case studies. The key project keywords and coordination key phrases defined in Hossain's previous work (2009a) were adopted for e-mail extraction of Dabhol and Azurix datasets. In addition to the project name, common names associated with the project are also used. In some cases, such as Maharashtra, the name of an Indian region is used because Enron's sole dealing with the region is through the Dabhol project. The list of project names and the alternate associations is provided in Table 3.1.

According to Hossain (2009a) the motivation for studying coordination on a project-based scope is to better capture the coordinative processes as the employees work towards a common goal. Their definition of project scope goes beyond the pattern of messaging and takes into account the reason for messaging. Based on the definitions of Hossain (2009a) the e-mails meaningful and oriented toward the project goal were included in this study.

Emails containing at least one of the project keywords and one of the coordination key phrases were included in the project scope. Using these project keywords, the employees found to have either sent or received an email matching one of these was extracted. Each project scope was treated separately and so the operation was repeated two times. The list of coordination key phrases is provided in Table 3.2.

The results contained the sender and recipient email addresses, date, subject and body of the email and the email's folder information. The results are transferred in to Ms Excel files. The adjacency matrix of the employees was created for each project using the transferred data. As a part of the data cleansing the employees with an in- and out-degree fewer than two were removed.

Seven hundred and twenty one people were extracted as a part of Azurix project scope. Nine hundred and four people were extracted as a part of the Dabhol project scope. The coordination score of each actor was calculated by the summation of the weights of key coordination phrases extracted from their sent e-mails. 180 people were found to have demonstrated coordination in the Dabhol scope and 157 people were found to have demonstrated coordination in the Azurix scope. The e-mail addresses that sending company announcements were removed from the dataset in order to reduce the effect of mass-mailers. As a part of the data cleansing the employees with an in and out-degree fewer than two were removed. In the Dabhol dataset this process eliminated 31 nodes. From the original 180 nodes, the data cleansing process totally eliminated 79 nodes, leaving 101 employees with the genuine data. In the Azurix dataset the second phase of the cleansing process eliminated 54 nodes. From the original 157 nodes, the data cleansing process totally eliminated 70 nodes, leaving 87 employees with the genuine data. Also the employees who have more than one e-mail addresses were counted only once and their total coordination score was calculated by the summation of each addresses' coordination score. This process removed 48 nodes from the Dabhol dataset and 16 nodes from the Azurix dataset.

Table 3.1. Project names and alternate association.
(Source: Hossain 2009a)

Formal project/company name	**Alternate associations**
Dabhol Power Company	Dabhol Power Company Dabhol, DPC, Maharashtra, MSEB (Maharashtra State Electricity Board)
Azurix Water Company	Azurix Water Company Azurix, Wessex Water, BOT Contract, WaterDesk.com, Water2Water.com, American Water Works

3.2. Dabhol Power Corporation and Azurix Corporation

Dabhol Power Corporation (DPC) was created in 1997 by Enron for the purchase and sale of electricity in Maharashtra, India. Enron International's energy plan included a new Power Plant and pipeline from Dabhol to Hazira (Hossain 2009a). The estimated cost of the project was $2.8 billion. The project faced two major problems. First one was the violent protests owing to the environmental impact of the new plant. Second one was the newly elected local state government threat to cancel the deal because of its high price tag and the alleged corruption by the previous government that negotiated the project (Hossain and Wu 2009). The project involved two phases; the construction of the power plant (740 megawatts), and the eventual expansion of its output capacity (1444 megawatts) (Hossain 2009a). The phase one was completed and the DPC reported profits of $42 million during the first year of its operations by May 1999. However, phase two of the project stagnated and in December 2001, Enron filed for Chapter 11 bankruptcy before the project was completed (Hossain 2009a).

In 1998, Enron moved into the water sector, creating the Azurix Corporation, which it part-floated on the New York Stock Exchange (NYSE) in June 1999. Azurix struck a major deal to operate the water and sewage for two regions of Argentina's Buenos Aires Province. Azurix paid $439 million for the 30 year concessions, which served just fewer than 2 million people. With operations in Argentina, England and Mexico, Azurix was a globe-spanning company. Although Azurix was profitable, it wasn't living up to the majestic expectations of Enron. The company reported net income of $37.7 million in 1999, on revenues of $618 million. Overall, Azurix failed to break into the water utility market, and in April 2001, Enron announced its intention to break up Azurix and sell its assets (Hossain 2009a).

3.3. Coordination Score Calculation Procedure

The study of Hossain (2009a) builds on existing coordination theory and presents a new approach for exploring organizational processes. According to Hossain (2009a) by identifying and mapping these four key coordination processes defined by Malone and Crowston (1994) the study is able to identify the specific instances of coordination. Using these discrete instances, it becomes possible to measure and compare levels of coordination. He interpreted and operationalized the four processes were then for the study of the e-mail corpus. These four coordination processes are: managing shared resources, managing producer–consumer relationships, managing simultaneity constraints and managing task/subtask dependencies. Using text mining techniques, Hossain (2009a) operationalized these four processes into key phrases to be extracted from the e-mail dataset. The final list of coordination key phrases is shown in Table 3.2.

Coordination is measured with the application of text mining techniques. A query was written to extract the coordination data bounded by project scope from the e-mail dataset. By running the query in MySQL software, the frequency of the coordination key phrases and the information of senders using each coordination phrase were extracted. The results of the extraction were organized by using Pivot table option in Ms. Excel software. The procedure of calculation of coordination scores is summarized in Figure 3.4.

1) Add tables including project keywords and coordination key phrases to the MySQL data.

2) Run SQL script code to search for the project keywords and coordination key phrases within the project scope.

3) Run SQL query to extract the coordination data from the MySQL database.

4) Import the SQL results to Microsoft Excel and calculate each key phrase's weight

   4a) Use sender data of coordinative emails to identify the used key phrases and their frequencies

   4b) Use calculated weights and frequencies to measure the total coordination score for each employee

Figure 3.4. Procedure of Calculation of Coordination Scores.

Table 3.2. Coordination Key Phrases.
(Source: Hossain 2009a)

Resource allocation	**Producer/consumer relationships**	**Simultaneity constraints**	**Tasks/Subtasks**
Help coordinate	Are as follows	As we move closer	I have considered
Please allow	Attached is a	Please allow time	I recommend
Please communicate	Attached please find	On track	I suggest
Please coordinate	The bottom line is	Sufficient time	I wanted to
Please do	The purpose is	Take the time	I would like to
Please get	For your information	Agenda	I would suggest
Please make arrangements	FYI	Follow up	We can discuss
Please make sure		On time	We can then
Please update		Make a schedule	We have seriously
Do you want to			We need to
I request			We should
I would appreciate			To ensure that
I would like to			It will need
I would like your			I am changing
I would ask			I believe
Look into			Let me know if
Make sure that			Please let me know
Please see			We have had
Please speak			Would probably be
Please work			I believe you are
Put this together			Which brings me to
You will be			We have begun
You work with			
Ensure that			
We can go			

Each coordination phrase was assigned a weight based on their level of significance. The weight was determined by the number of people that uses the keyword and the frequency with which they use it. The weight of the words is equal to the base two log of the sum of the usage frequency of the words. A word used more commonly was assigned a greater weight. The reason for using the base two log of the frequency was to capture effect of words with higher frequency without creating substantial outliers. This creates a normal distribution of the coordination weights and reduces the outliers. The weights of the words varied from 0 to 7.68 with an average of 3.30 for Dabhol dataset.

Microsoft Excel software was used to calculate each member's coordination score. The weights of the coordination phrases were calculated using the formula explained above. The keywords were processed individually. The dataset was queried to identify the employees that have used the key phrase in an email (Table 3.3.).

For each employee, the list of the used keywords was identified. The weighted coordination score for each keyword was calculated by multiplying the frequency of the keyword with its calculated weight. The final coordination score of each person was collated by aggregating the calculated scores based on the keyword matches (Table 3.4.). The coordination score for each employee was measured within the project scope. In this process, only the emails that are within the project scope was extracted and counted. The coordination scores ranged from 3.32 to 3618.01 with an average of 88,58 in the Dabhol dataset. The coordination scores ranged from 3.32 to 1265.34 with an average of 56.60 in the Azurix dataset.

Table 3.3. Weight Calculation of the Keyword "Agenda".

Employee	Coordination Key Phrase	Frequency
amr.ibrahim@enron.com	Agenda	2
ann.schmidt@enron.com	Agenda	31
bgrizzle@capricornholdings.com	Agenda	1
brian.redmond@enron.com	Agenda	1
capitol-news@list.asme.org	Agenda	1
carol.howes@enron.com	Agenda	1
courtney.votaw@enron.com	Agenda	9
dan.masters@enron.com	Agenda	1
david.port@enron.com	Agenda	1
henry.means@enron.com	Agenda	8
inja.chun@enron.com	Agenda	1
j..kean@enron.com	Agenda	1
j.kaminski@enron.com	Agenda	1
jeffrey.shankman@enron.com	Agenda	1
john.hardy@enron.com	Agenda	6
jonathan.whitehead@enron.com	Agenda	1
joseph.deffner@enron.com	Agenda	1
kimberley.nelson@enron.com	Agenda	6
laura.glenn@enron.com	Agenda	1
linda.robertson@enron.com	Agenda	1
m..schmidt@enron.com	Agenda	21
maureen.mcvicker@enron.com	Agenda	1
mike.mcconnell@enron.com	Agenda	1
miyung.buster@enron.com	Agenda	1
nikita.varma@enron.com	Agenda	8
paul.y barbo@enron.com	Agenda	2
pmadpr@worldnet.att.net	Agenda	8
rbw@mrwassoc.com	Agenda	1
richard.shapiro@enron.com	Agenda	1
rick.buy@enron.com	Agenda	2
rob.walls@enron.com	Agenda	1
rod.hayslett@enron.com	Agenda	1
sandeep.kohli@enron.com	Agenda	4
sarah.palmer@enron.com	Agenda	4
sharonda.stephens@enron.com	Agenda	4
steven.kean@enron.com	Agenda	1
susan.mara@enron.com	Agenda	1
tom.hoatson@enron.com	Agenda	1
vince.kaminski@enron.com	Agenda	1
wayne.perry@enron.com	Agenda	1
yaqoobalzadjali@omanlng.co.om	Agenda	1
	Sum of the coordination key phrase frequency =	142
	Weight of the coordination key phrase =	7,15

Table 3.4. An Example of Coordination Score Calculation.

Employee	Coordination Key Phrase	Frequency	Weight of the phrase	Weighted Score
james.derrick@enron.com	Agenda	1	5.58	5.58
james.derrick@enron.com	FYI	4	6.86	27.43
james.derrick@enron.com	I would ask	1	1.00	1.00
james.derrick@enron.com	Please see	1	5.36	5.36
james.derrick@enron.com	You will be	1	4.32	4.32
			Coordination Score:	43.70

3.4. Measuring coordination through centrality

This study measures centrality on two axes: (1) Measurement; and (2) directional analysis. Five measures of centrality are used in this study: (1) degree; (2) closeness; (3) betweenness; (4) eigenvector and (5) information. Based on Freeman's (1979) definitions, degree centrality denotes the number of nodes connected to one node in particular; betweenness centrality is the extent to which a firm lies between other pairs of firms (it is the proportion of all the shortest paths (i.e., geodesic distances) between pairs of other firms that pass through the firm); closeness centrality is based on the sum of the geodesic distances from each node to all other nodes. Eigenvector centrality is defined as the principal eigenvector of the adjacency matrix defining the network (Bonacich 1972). Then, following Stephenson and Zelen (1989)'s definition information centrality is the harmonic average of all the information included in all the paths flowing from that firm. Centrality was calculated using UCINET 6 for Windows (Borgatti et al. 2002).

The central and more visible actors in the network are more likely to be potential allies for other powerful actors, and thereby appearing even more powerful. In addition, individuals situated in most centralized position of a network were likely to emerge as the leader and participate more in task solution. To measure the effect of centrality on coordination processes by collecting data from the actors which enact these coordination processes. Then the data can be compared with its relative centralities to determine if a correlation exists. (Hossain 2009a).

Hypothesis 1: Centrally positioned actors show more coordinative activity.

Hossain's previous work (2009a) suggests that betweenness centrality is the most powerful independent predictor of the effects of centrality on coordination. Freeman (1980) concludes that "betweenness and closeness based measures of point centrality are determined by the same structural elements of a communication network". Since both are functions of local pair dependency, all measures have in common the same structural element: the geodetic pathway. If one assumes that communication only occurs long the shortest possible path, then communication channels are by default the geodesics. This fundamental assumption of Freeman's betweenness and closeness centralies neglects measuring communication occurring along reachable, non-geodetic pathways (Stephenson and Zelen 1989). Bonacich's eigenvector centrality (1972) associated with the largest characteristic eigenvalue of the adjacency matrix does not make use of geodesic paths. However, this approach neglects multiple shared paths between points in a network. Stephenson and Zelen's (1989) information centrality uses all paths, but gives them relative weighting as a function of the "information" they contain. This study aims to investigate if information centrality is a better predictor for coordination among other centrality measures.

Hypothesis 2: Information Centrality is the best measurement (structural characteristics) for predicting coordination in undirected graphs.

A social network can be directed or undirected. A network is directional if the ties are oriented from one actor to another (Wasserman and Faust 1994). In order to investigate the relation between directed centrality measures and coordination, five directed centrality measures were also calculated; out-degree, in-degree, out-closeness, in-closeness and betweenness. In social network applications out-centrality measures the expansiveness and in-centrality measures the receptivity or popularity. In-centrality regarded as an indicator for the prominence of an actor, whereas out-centrality measures the influence of an actor (Hossain 2009a).

Hypothesis 3: Out-centrality measures are the most potent predicate for coordination in directed graphs.

3.5. Statistical Methods

Pearson product moment test, Spearman rank test and the Kendall rank test were used to measure correlation between centrality measures. The r-estimate is the correlation found using the Spearman rank test and Pearson product moment test. Kendall's Tau (τ) is the correlation found using the Kendall rank test. The p-value indicated that the probability of receiving correlations by chance alone. Statistical significance level used a p-value of 0.05. The Pearson product moment test is parametric, meaning the normal distribution (bell-shape) is assumed. Owing to the non-normal distribution of the data, using the non-parametric Spearman and Kendall rank tests would give more accurate results compared with Pearson product moment test. However, the central limit theorem states that when a dataset is large (eg. $n>80$) the mean will follow the normal distribution even if the respective variable is not normally distributed in the population. As a result both parametric and non-parametric tests were used for the hypothesis testing (Hossain et al. 2006).

The correlations of centrality measures are examined in two groups: directed centrality measures and undirected centrality measures. Undirected measures are degree, closeness, betweenness, eigenvector and information centrality. Directed measures are in-degree, out-degree, in-closeness, out-closeness and betweenness. The correlations between the centrality measures and coordination score are also measured.

The Mann-Whitney U test was used to determine if there is a substantial difference in coordination scores of people with high and low centrality. The Mann-Whitney U-test is non-parametric and does not assume a normal distribution. The null hypothesis is that there is no statistical difference in coordination scores between the low and high groups of centrality. The Mann Whitney-U test was performed separately for each centrality measure. In order to test the statistical difference in coordination between high and low centrality groups, the data is ordered by each centrality measure and splitted into two groups. They were divided by both mean and median centrality measurement. If the high and low groups of centrality were shown to be statistically different for all centrality measures, the null hypothesis will be rejected and the hypothesis of centrality positioned individuals show more coordination will be accepted.

The Hypothesis 2 is that information centrality is the most potent predicate for coordination in undirected graphs. The Mann Whitney U test investigated Hypothesis 2 by again using the statistical difference between high and low groups of centrality and statistical significances were compared against the each other.

The one-way analysis of variance (ANOVA) is also used to determine whether there are any significant differences between the means of three or more independent groups. The one-way ANOVA compares the means between the groups and determines whether any of those means are significantly different from each other. The null hypothesis is that the means of the groups are equal. If the hypothesis is true, then the "between group variance" will be equal to the "within group variance." If, however, the one-way ANOVA returns a significant result, we accept the alternative hypothesis (H_A), which is that there are at least 2 group means that are significantly different from each other. A one-way ANOVA test was conducted to test the differences in coordinative activity between the high and low groups of information centrality.

CHAPTER 4

RESEARCH FINDINGS AND DISCUSSION

This chapter presents the quantitative findings from two case studies discussed in the previous chapter. Statistical analysis of centrality measures and coordination scores are presented.

4.1. Network Properties Results

The density of the Azurix's information exchange relation matrix is 0.0502. This means %5 of all the possible ties are present. The density of Dabhol's information exchange relation matrix is 0.0954 which means %9.5 of all the possible ties are present.

Table 4.1. Network Properties Results.

Dataset	Network Size	Network Density	Network Centralization (symmetrized)	Network Centralization (in-degree)	Network Centralization (out-degree)
Azurix	87	0.0502	3.18%	1.921%	3.161%
Dabhol	101	0.0954	1.57%	1.076%	1.591%

Network centralization measures were calculated for both directed and undirected (symmetrized) networks. The results imply that networks are not highly centralized. This explains that the connections are not made between certain units. Ruan et al. (2012) suggested that in collaborative working projects the knowledge supportive activities happened relatively equally between all units and members supported one another using their knowledge and expertise.

4.2. Distribution of Coordination Phrases

Table 4.2. Coordination Key Phrase Distribution for Dabhol Project.

Distribution of Total Phrases for Each Dependency Type - Dabhol Project				
Resource Allocation Dependency	Producer/Consumer Relationships Dependency	Simultaneity Constraints Dependency	Tasks/Subtasks Relationship Dependency	Total Phrase in sent emails
404	293	385	629	1711

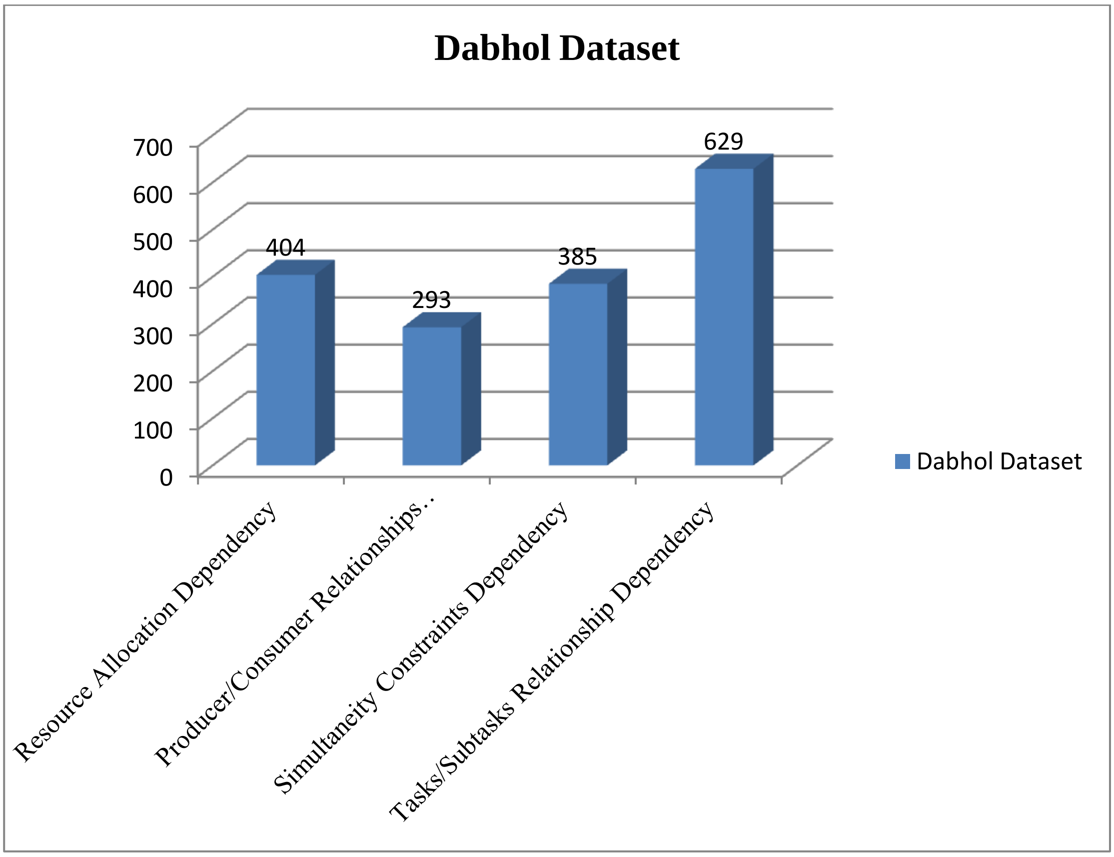


Figure 4.1. Coordination Key Phrase Distribution for Dabhol Project.

The actors used 1711 coordination phrases in e-mail communication of Dabhol project. Resource Allocation Dependency related phrases were used 404 times, Producer/Consumer Relationships Dependency related phrases were used 293 times, Simultaneity Constraints Dependency related phrases were used 385 times and Tasks/Subtasks Relationship Dependency related phrases were used 629 times.

Table 4.3. Coordination Key Phrase Distribution for Azurix Project.

Distribution of Total Phrases for Each Dependency Type - Azurix Project				
Resource Allocation Dependency	Producer/Consumer Relationships Dependency	Simultaneity Constraints Dependency	Tasks/Subtasks Relationship Dependency	Total Phrase in sent emails
345	200	208	501	1254

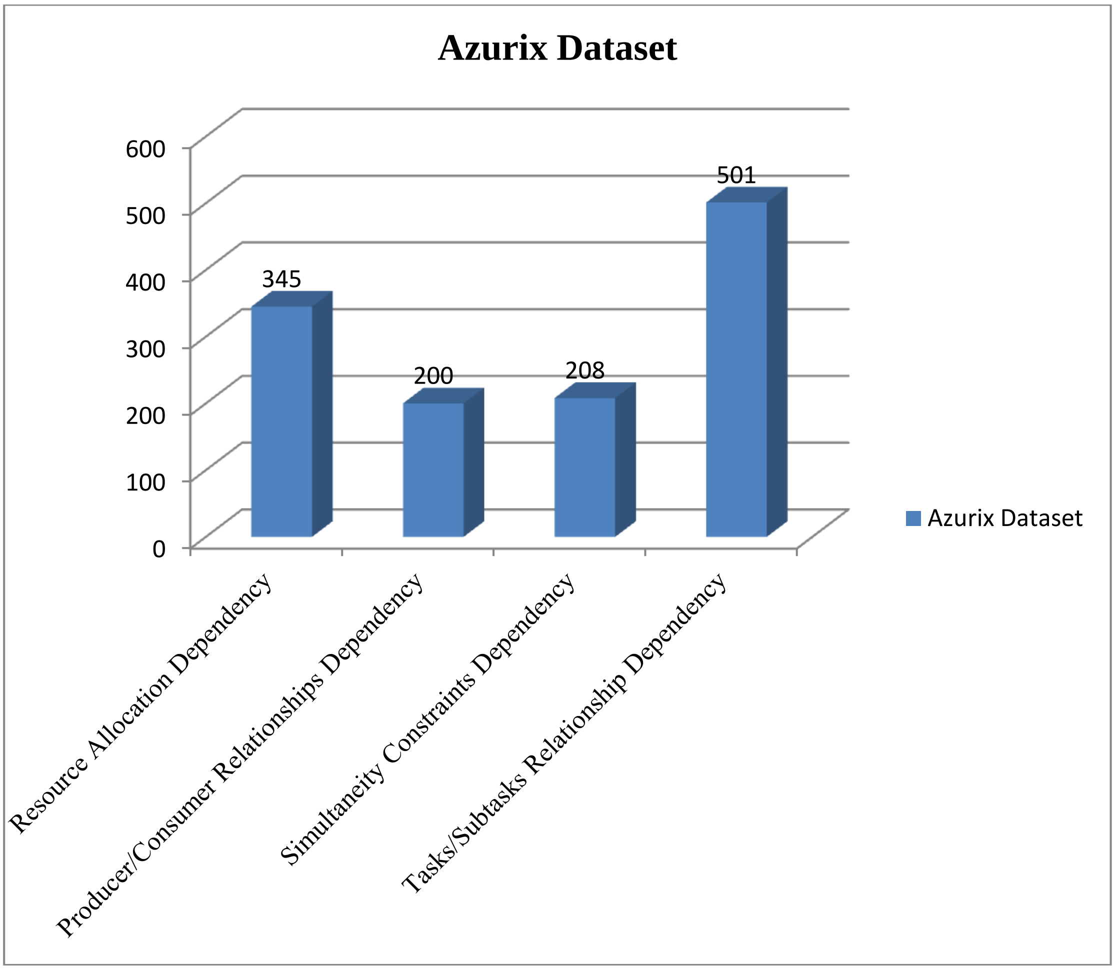
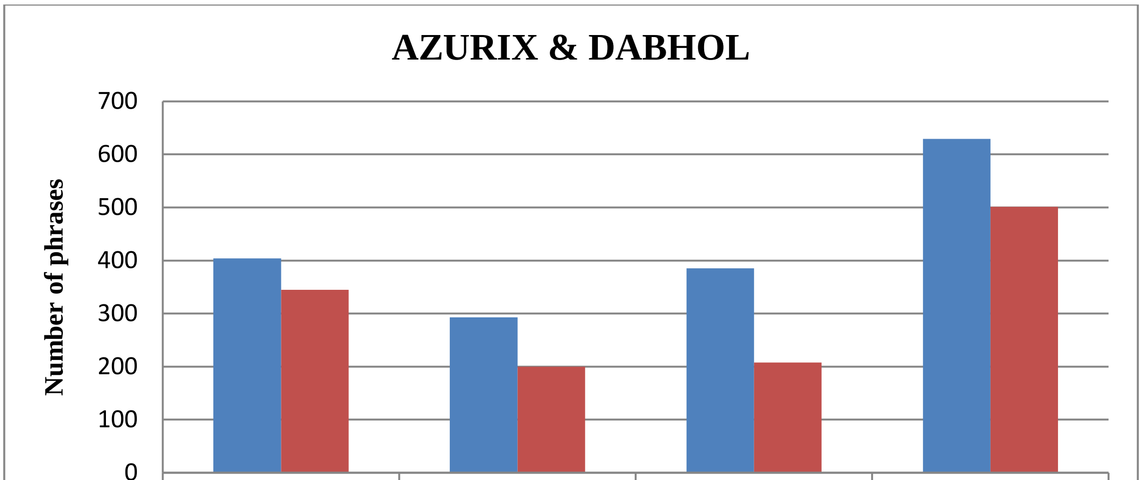


Figure 4.2. Coordination Key Phrase Distribution for Dabhol Project.

The actors used 1254 coordination phrases in e-mail communication of Azurix project. Resource Allocation Dependency related phrases were used 345 times, Producer/Consumer Relationships Dependency related phrases were used 200 times, Simultaneity Constraints Dependency related phrases were used 208 times and Tasks/Subtasks Relationship Dependency related phrases were used 501 times.



	Resource Allocation Dependency	Producer/Consumer Relationships Dependency	Simultaneity Constraints Dependency	Tasks/Subtasks Relationship Dependency
Dabhol	404	293	385	629
Azurix	345	200	208	501

Figure 4.3. Total Coordination Key Phrase Distribution.

The actors used 2965 coordination phrases in e-mail communication of Azurix and Dabhol project totally. Resource Allocation Dependency related phrases were used 749 times, Producer/Consumer Relationships Dependency related phrases were used 493 times, Simultaneity Constraints Dependency related phrases were used 593 times and Tasks/Subtasks Relationship Dependency related phrases were used 1130 times. Tasks/Subtasks Relationship Dependency related phrases were the most used coordination phrases for both Azurix and Dabhol projects. Coordination phrases were used more often in Dabhol project compared to Azurix project.

4.3. Network Centrality Correlations

Eigenvector centrality and information centrality formulations assume that the relational structure is non-directional, the analysis restricted to undirected networks for these measures. UCINET gives the option of whether to treat data as symmetric or asymmetric while computing degree centrality. Therefore, degree, betweennes, closeness and eigenvector centrality were calculated using a symmetrized and binarized data matrix for the undirected measures. For the directed and weighted network of Dabhol and Azurix datasets, UCINET eliminates the smaller weight indicating either the sent or the received number of emails on the adjacency matrix; and reconstructs the weighted matrix with the greater weight values assigned to the incoming or outgoing link of the node.

4.3.1. Undirected Network Correlations

Correlation coefficients (Pearson's r) among five undirected centrality measures and coordination scores of Azurix dataset are presented in Table 4.4.

Table 4.4. Correlation coefficients (Pearson's r) among the undirected centrality measures of Azurix network.

	Degree	Closeness	Betweenness	Eigenvector	Information	LnCoordination
Degree	1					
Closeness	0.393**	1				
Betweenness	0.827**	0.259*	1			
Eigenvector	0.826**	0.368**	0.505**	1		
Information	0.757**	0.495**	0.527**	0.750**	1	
LnCoordination	0.358**	-0.022	0.285**	0.379**	0.362**	1

* . $P< 0.05$, **. $P < 0.01$

There is some variability in the correlation among measures between two case studies. For the Azurix dataset, the highest product moment correlation was between degree and betweenness centrality measures ($r = 0.827$, $p < 0.01$). The next highest correlation was between eigenvector and betweenness ($r = 0.826$, $p < 0.01$) followed by information and degree centrality ($r = 0.757$, $p < 0.01$). The lowest correlation observed was between closeness and degree centrality measures ($r = 0.259$, $p < 0.05$).

Correlation coefficients (Pearson's r) among five undirected centrality measures and coordination scores of Dabhol dataset are presented in Table 4.5.

For the Dabhol dataset, eigenvector and degree are strongly correlated ($r = 0.924$, $p < 0.01$). The next highest correlation was between degree and betweenness ($r = 0.766$, $p < 0.01$) followed by information and degree centrality measures ($r = 0.757$, $p < 0.01$). The correlation between closeness and betweenness measures is also the lowest ($r = 0.333$, $p < 0.01$) for the Dabhol dataset.

Correlations between network centrality measures and coordination scores are examined for the two case studies in order to find out which measure of centrality is the most useful predictor for coordination ability. Network centrality measures are weakly correlated but statistically significant with the coordination scores for the two datasets. However, for the Azurix dataset eigenvector centrality ($r = 0.379$, $p < 0.01$) and for the Dabhol dataset information centrality ($r = 0.420$, $p < 0.01$) have the highest correlation with coordination scores among other centrality measures.

Table 4.5. Correlation coefficients (Pearson's r) among the undirected centrality measures of Dabhol network.

	Degree	Closeness	Betweenness	Eigenvector	Information	LnCoordination
Degree	1					
Closeness	0.514**	1				
Betweenness	0.766**	0.333**	1			
Eigenvector	0.924**	0.504**	0.559**	1		
Information	0.746**	0.585**	0.467**	0.712**	1	
LnCoordination	0.368**	0.265**	0.409**	0.304**	0.420**	1

* . P< 0.05, **. P < 0.01

Correlation coefficients (Spearman's r) among five undirected centrality measures and coordination scores of Azurix dataset are presented in Table 4.6.

Table 4.6. Rank order correlation coefficients (Spearman's r) among the undirected centrality measures of Azurix network.

	Degree	Closeness	Betweenness	Eigenvector	Information	LnCoordination
Degree	1					
Closeness	0.850**	1				
Betweenness	0.849**	0.719**	1			
Eigenvector	0.790**	0.920**	0.568**	1		
Information	0.851**	0.826**	0.659**	0.827**	1	
LnCoordination	0.274*	0.301**	0.300**	0.270*	0.383**	1

* . P< 0.05, **. P < 0.01

For the Azurix network, the highest rank order correlation was between eigenvector and closeness centrality measures ($r = 0.920$, $p < 0.01$). The next highest correlation was between information and degree centrality ($r = 0.851$, $p < 0.01$). Closeness and degree ($r = 0.850$, $p < 0.01$) and betweennes and degree ($r = 0.849$, $p < 0.01$) are also strongly correlated. The lowest correlation observed was between betweenness and eigenvector centrality measures ($r = 0.568$, $p < 0.01$).

Correlation coefficients (Spearman's r) among five undirected centrality measures and coordination scores of Dabhol dataset are presented in Table 4.7.

For the Dabhol dataset, eigenvector and closeness are highly correlated ($r = 0.924$, $p < 0.01$) similar to the Azurix dataset. The next highest correlation was between degree and closeness ($r = 0.882$, $p < 0.01$) followed by eigenvector and degree centrality measures ($r = 0.880$, $p < 0.01$). The lowest correlation observed was between eigenvector and betweenness centrality measures ($r = 0.571$, $p < 0.01$).

Table 4.7. Rank order correlation coefficients (Spearman's r) among the undirected centrality measures of Dabhol network.

	Degree	Closeness	Betweenness	Eigenvector	Information	LnCoordination
Degree	1					
Closeness	0.882**	1				
Betweenness	0.779**	0.704**	1			
Eigenvector	0.880**	0.922**	0.571**	1		
Information	0.839**	0.801**	0.673**	0.793**	1	
LnCoordination	0.291**	0.326**	0.382**	0.231*	0.409**	1

* . P< 0.05, **. P < 0.01

Network centrality measures are weakly correlated with the coordination scores for the two datasets. It is also worth noting that rank order correlation coefficients are statistically more significant compared to the Pearson correlation coefficients. Information centrality has the highest correlation with coordination scores for both the Azurix dataset ($r = 0.383$, $p < 0.01$) and Dabhol dataset ($r = 0.409$, $p < 0.01$).

Table 4.8. Kendall's tau (τ) coefficient among the undirected centrality measures of Azurix network.

	Degree	Closeness	Betweenness	Eigenvector	Information	LnCoordination
Degree	1					
Closeness	0.709**	1				
Betweenness	0.713**	0.554**	1			
Eigenvector	0.638**	0.777**	0.416**	1		
Information	0.721**	0.661**	0.484**	0.666**	1	
LnCoordination	0.195*	0.200**	0.210**	0.185*	0.246**	1

* . P< 0.05, **. P < 0.01

For the Azurix network, the highest Kendall rank order correlation was between eigenvector and closeness centrality measures ($r = 0.777$, $p < 0.01$). The next highest correlation was between information and degree centrality ($r = 0.721$, $p < 0.01$). Betweennes and degree ($r = 0.713$, $p < 0.01$) and closeness and degree ($r = 0.709$, $p < 0.01$) are also strongly correlated. The lowest correlation observed was between betweenness and eigenvector centrality measures ($r = 0.416$, $p < 0.01$).

For the Dabhol dataset, eigenvector and closeness are highly correlated ($r = 0.770$, $p < 0.01$) similar to the Azurix dataset. The next strongest correlation was between degree and eigenvector ($r = 0.746$, $p < 0.01$) followed by closeness and degree centrality measures ($r = 0.736$, $p < 0.01$). The lowest correlation was between eigenvector and betweenness centrality measures ($r = 0.422$, $p < 0.01$).

Network centrality measures are weakly correlated with the coordination scores for the two datasets. Information centrality has the highest correlation with coordination scores for both the Azurix dataset ($r = 0.246$, $p < 0.01$) and Dabhol dataset ($r = 0.279$, $p < 0.01$).

Table 4.9. Kendall's tau (τ) coefficient among the undirected centrality measures of Dabhol network.

	Degree	Closeness	Betweenness	Eigenvector	Information	LnCoordination
Degree						
Closeness	0.736**					
Betweenness	0.617**	0.543**				
Eigenvector	0.746**	0.770**	0.422**			
Information	0.706**	0.639**	0.515**	0.630**		
LnCoordination	0.206**	0.221**	0.272**	0.157*	0.279**	

* . P< 0.05, **. P < 0.01

4.3.2. Directed Network Correlations

A directional analysis is further performed to investigate the differences between in- and out- centrality and the effects on coordination ability. In-centrality seems to indicate the prominence of an actor, whereas out-centrality measures the influence of an actor. (Wasserman and Faust 1994) Out-degree, in-degree, out-closeness, in-closeness and betweenness (non-symmetric) centrality measures are calculated for the directed data. The option of whether to treat data as asymmetric was chosen while calculating in-degree and out-degree measures. The data matrix was dichotomized ($X_{ij} > 0$ was recoded to 1) while calculating in-closeness and out-closeness measures. UCINET cannot handle valued data while calculating Freeman betweenness centrality, so it binarizes the data automatically. However it does handle directed (non-symmetric) data, so it does not symmetrize.

For the Azurix dataset, the highest product moment correlation was between in-degree and betweenness centrality measures ($r = 0.678$, $p < 0.01$). The next highest correlation was between out-closeness and out-degree ($r = 0.451$, $p < 0.01$) followed by betweenness and out-degree centrality ($r = 0.415$, $p < 0.01$). The correlations between in-closeness and out-closeness ($r = -0.202$, $p > 0.05$) and in-closeness and out-degree measures ($r = -0.015$, $p > 0.05$) are negative and non-significant for the Azurix data.

Table 4.10. Correlation coefficients (Pearson's r) among the directed centrality measures of Azurix network.

	Outdegree	Indegree	Outcloseness	Incloseness	Betweenness	LnCoordination
Outdegree	1					
Indegree	0.139	1				
Outcloseness	0.451**	0.200	1			
Incloseness	-0.015	0.249*	-0.202	1		
Betweenness	0.415**	0.678**	0.306**	0.159	1	
LnCoordination	0.540**	0.244*	0.128	0.057	0.296**	1

* . P< 0.05, **. P < 0.01

For the Dabhol dataset, betweenness and in-degree has the highest correlation ($r = 0.529$, $p < 0.01$). The next highest correlation was between in-closeness and betweenness ($r = 0.397$, $p < 0.01$) followed by in-closeness and in-degree centrality measures ($r = 0.364$, $p < 0.01$). The correlation between in-closeness and out-closeness measures is negative and non-significant ($r = -0.094$, $p > 0.05$) for the Dabhol data.

Table 4.11. Correlation coefficients (Pearson's r) among the directed centrality measures of Dabhol network.

	Outdegree	Indegree	Outcloseness	Incloseness	Betweenness	LnCoordination
Outdegree	1					
Indegree	0.053	1				
Outcloseness	0.264**	0.096	1			
Incloseness	0.038	0.364**	-0.094	1		
Betweenness	0.247*	0.529**	0.171	0.397**	1	
LnCoordination	0.611**	0.207*	0.245*	0.070	0.355**	1

* . P< 0.05, **. P < 0.01

Correlations between network centrality measures and coordination scores are examined for the two case studies in order to find out which directed measure of centrality is the most useful predictor for coordination ability. Network centrality measures are weakly correlated with the coordination scores for the two datasets mostly according to the Pearson product moment correlation. However, out-degree centrality has a strong correlation with coordination scores for both the Azurix data ($r = 0.540$, $p < 0.01$) and Dabhol data ($r = 0.611$, $p < 0.01$).

Table 4.12. Rank order correlation coefficients (Spearman's r) among the directed centrality measures of Azurix network.

	Outdegree	Indegree	Outcloseness	Incloseness	Betweenness	LnCoordination
Outdegree	1					
Indegree	0.024	1				
Outcloseness	0.760**	-0.042	1			
Incloseness	-0.141	0.616**	-0.240*	1		
Betweenness	0.574**	0.490**	0.444**	0.300**	1	
LnCoordination	0.291*	0.174	0.095	0.020	0.273*	1

* . P< 0.05, **. P < 0.01

For the Azurix data, the highest Spearman rank order correlation was between out-closeness and out-degree centrality measures ($r = 0.760$, $p < 0.01$). The next highest correlation was between in-closeness and in-degree ($r = 0.616$, $p < 0.01$) followed by betweenness and out-degree centrality ($r = 0.574$, $p < 0.01$). The correlations between in-closeness and out-closeness ($r = -0.240$, $p > 0.05$), in-closeness and out-degree measures ($r = -0.141$, $p > 0.05$) and out-closeness and in-degree ($r = -0.042$, $p > 0.05$) are negative and non-significant for the Azurix data.

For the Dabhol dataset, in-degree and in-closeness has the highest correlation ($r = 0.749$, $p < 0.01$). The next highest correlation was between in-degree and betweenness ($r = 0.698$, $p < 0.01$) followed by in-closeness and betweenness centrality measures ($r = 0.584$, $p < 0.01$). The correlation between in-closeness and out-closeness measures is negative and significant ($r = -0.345$, $p < 0.01$) while the correlation between out-closeness and in-degree ($r = -0.116$, $p > 0.05$) is negative and non-significant for the Dabhol data.

Network centrality measures are weakly correlated with the coordination scores for the two datasets mostly according to the Spearman correlation coefficients. However, out-degree centrality has the highest correlation with coordination scores for both the Azurix data ($r = 0.291$, $p < 0.05$) and Dabhol data ($r = 0.544$, $p < 0.01$).

Table 4.13. Rank order correlation coefficients (Spearman's r) among the directed centrality measures of Dabhol network.

	Outdegree	Indegree	Outcloseness	Incloseness	Betweenness	LnCoordination
Outdegree	1					
Indegree	0.285**	1				
Outcloseness	0.535**	-0.116	1			
Incloseness	0.121	0.749**	-0.345**	1		
Betweenness	0.533**	0.698**	0.149	0.584**	1	
LnCoordination	0.544**	0.227*	0.200*	0.060	0.378**	1

* . P< 0.05, **. P < 0.01

Table 4.14. Kendall's tau (τ) coefficient among the directed centrality measures of Azurix network.

	Outdegree	Indegree	Outcloseness	Incloseness	Betweenness	LnCoordination
Outdegree	1					
Indegree	0.015	1				
Outcloseness	0.625**	-0.026	1			
Incloseness	-0.104	0.485**	-0.209**	1		
Betweenness	0.480**	0.396**	0.366**	0.233**	1	
LnCoordination	0.215*	0.127	0.066	0.016	0.204*	1

* . P< 0.05, **. P < 0.01

For the Azurix data, the highest Kendall rank order correlation was between out-closeness and out-degree centrality measures ($r = 0.625$, $p < 0.01$). The next highest correlation was between in-closeness and in-degree ($r = 0.485$, $p < 0.01$) followed by betweenness and out-degree centrality ($r = 0.480$, $p < 0.01$). The correlation between in-closeness and out-closeness ($r = -0.209$, $p < 0.01$) is negative and significant, while the correlations between in-closeness and out-degree ($r = -0.104$, $p > 0.05$) and out-closeness and in-degree ($r = -0.026$, $p > 0.05$) are negative and non-significant for the Azurix data.

For the Dabhol dataset, in-degree and in-closeness has the highest correlation ($r = 0.601$, $p < 0.01$). The next highest correlation was between in-degree and betweenness ($r = 0.539$, $p < 0.01$) followed by in-closeness and betweenness centrality measures ($r = 0.469$, $p < 0.01$). The correlation between in-closeness and out-closeness measures is negative and significant ($r = -0.232$, $p < 0.01$) while the correlation between out-closeness and in-degree ($r = -0.068$, $p > 0.05$) is negative and non-significant for the Dabhol data.

Network centrality measures are weakly correlated with the coordination scores for the two datasets mostly according to Kendall rank correlation coefficients. However, out-degree centrality has the highest correlation with coordination scores for both the Azurix data ($r = 0.215$, $p < 0.05$) and Dabhol data ($r = 0.408$, $p < 0.01$).

Table 4.15. Kendall's tau (τ) coefficient among the directed centrality measures of Dabhol network.

	Outdegree	Indegree	Outcloseness	Incloseness	Betweenness	LnCoordination
Outdegree	1					
Indegree	0.197^{**}	1				
Outcloseness	0.409^{**}	-0.068	1			
Incloseness	0.087	0.601^{**}	-0.232^{**}	1		
Betweenness	0.406^{**}	0.539^{**}	0.117	0.469^{**}	1	
LnCoordination	0.408^{**}	0.164^{*}	0.137^{*}	0.037	0.272^{**}	1

* . P< 0.05, **. P < 0.01

For the Dabhol Project, the highest product moment correlation was between eigenvector and degree centrality measures ($r = 0.924$, $p < 0.01$). The next highest correlation was between degree and betweenness ($r = 0.766$, $p < 0.01$) followed by directed betweenness and undirected betweenness centrality ($r = 0.777$, $p < 0.01$). The correlation between in-closeness and out-closeness centrality ($r = -0.094$, $p > 0.05$) is negative and non-significant for the Dabhol data. The next lowest correlation was between in-degree and out-degree ($r = 0.053$, $p > 0.05$) and it is also not statistically significant. The results are presented in (Table 4.16).

Out-degree centrality was highly correlated with the coordination scores for Dabhol project according to the Pearson correlation test ($r = 0.611$, $p < 0.01$) and the lowest correlation was between in-closeness and coordination ($r = 0.070$, $p > 0.05$).

The highest Spearman rank order correlation was between eigenvector and closeness measures ($r = 0.927$, $p < 0.01$). The next highest correlation was between degree and closeness ($r = 0.889$, $p < 0.01$) followed by degree and eigenvector centrality ($r = 0.880$, $p < 0.01$). The correlation between in-closeness and out-closeness centrality

($r = -0.345$, $p < 0.01$) is negative and significant, while the correlation between in-degree and out-closeness ($r = -0.116$, $p > 0.05$) is negative and non-significant for the Dabhol data. The next lowest correlation was between and in-closeness and out-degree it is also not statistically significant. The Spearman rank correlation results are presented in (Table 4.17).

Network centrality measures were weakly correlated with the coordination scores for Dabhol project according to Spearman rank correlation test. Out-degree centrality had the highest correlation with coordination scores ($r = 0.544$, $p < 0.01$) and the lowest correlation was between in-closeness and coordination ($r = 0.060$, $p > 0.05$).

For the Dabhol data, the highest Kendall rank order correlation was between closeness and eigenvector centrality measures ($r = 0.778$, $p < 0.01$). Degree centrality had the same correlation with closeness and eigenvector centrality ($r = 0.746$, $p < 0.01$) .The correlation between in-closeness and out-closeness ($r = -0.232$, $p < 0.01$) is negative and significant, while the correlation between out-closeness and in-degree ($r = -0.068$, $p > 0.05$) is negative and non-significant for the Dabhol data. The Kendall's rank correlation results are presented in (Table 4.18).

Network centrality measures were weakly correlated with the coordination scores for Dabhol project according to Kendall's rank correlation test. Out-degree centrality had the highest correlation with coordination scores ($r = 0.408$, $p < 0.01$) and the lowest correlation was between in-closeness and coordination ($r = 0.037$, $p > 0.05$).

Table 4.16. Pearson correlation coefficients among the directed and undirected centrality measures of Dabhol data.

		1	2	3	4	5	6	7	8	9	10	11
1	Degree	1										
2	Closeness	0.490^{**}	1									
3	Betweenness	0.766^{**}	0.300^{**}	1								
4	Eigenvector	0.924^{**}	0.460^{**}	0.559^{**}	1							
5	Outdegree	0.381^{**}	0.198^{*}	0.322^{**}	0.351^{**}	1						
6	Indegree	0.578^{**}	0.241^{*}	0.560^{**}	0.481^{**}	0.053	1					
7	Outcloseness	0.428^{**}	0.539^{**}	0.204^{*}	0.383^{**}	0.264^{**}	0.096	1				
8	Incloseness	0.339^{**}	0.362^{**}	0.232^{*}	0.350^{**}	0.038	0.364^{**}	-0.094	1			
9	Betweenness-2	0.708^{**}	0.288^{**}	0.777^{**}	0.601^{**}	0.247^{*}	0.529^{**}	0.171	0.397^{**}	1		
10	Information	0.746^{**}	0.597^{**}	0.467^{**}	0.712^{**}	0.388^{**}	0.483^{**}	0.458^{**}	0.473^{**}	0.448^{**}	1	
11	LnCoordination	0.368^{**}	0.265^{**}	0.409^{**}	0.304^{**}	0.611^{**}	0.207^{*}	0.245^{*}	0.070	0.355^{**}	0.420^{**}	1

* . $P< 0.05$, **. $P < 0.01$

Table 4.17. Spearman rank correlation coefficients among the directed and undirected centrality measures of Dabhol data.

		1	2	3	4	5	6	7	8	9	10	11
1	Degree	1										
2	Closeness	0.889**	1									
3	Betweenness	0.779**	0.718**	1								
4	Eigenvector	0.880**	0.927**	0.571**	1							
5	Outdegree	0.693**	0.685**	0.587**	0.588**	1						
6	Indegree	0.710**	0.641**	0.660**	0.603**	0.285**	1					
7	Outcloseness	0.361**	0.375**	0.265**	0.365**	0.535**	-0.116	1				
8	Incloseness	0.478**	0.494**	0.483**	0.445**	0.121	0.749**	-0.345**	1			
9	Betweenness-2	0.682**	0.668**	0.754**	0.519**	0.533**	0.698**	0.149	0.584**	1		
10	Information	0.839**	0.812**	0.673**	0.793**	0.686**	0.756**	0.190	0.458**	0.597**	1	
11	LnCoordination	0.291**	0.321**	0.382**	0.231*	0.544**	0.227*	0.200*	0.060	0.378**	0.409**	1

* . P< 0.05, **. P < 0.01

Table 4.18. Kendall's tau (τ) rank correlation coefficients among the directed and undirected centrality measures of Dabhol data.

		1	2	3	4	5	6	7	8	9	10	11
1	Degree	1										
2	Closeness	0.746**	1									
3	Betweenness	0.617**	0.555**	1								
4	Eigenvector	0.746**	0.778**	0.422**	1							
5	Outdegree	0.552**	0.525**	0.439**	0.447**	1						
6	Indegree	0.559**	0.482**	0.489**	0.453**	0.197**	1					
7	Outcloseness	0.272**	0.278**	0.184**	0.274**	0.409**	-0.068	1				
8	Incloseness	0.350**	0.366**	0.351**	0.319**	0.087	0.601**	-0.232**	1			
9	Betweenness-2	0.529**	0.506**	0.605**	0.381**	0.406**	0.539**	0.117	0.469**	1		
10	Information	0.706**	0.648**	0.515**	0.630**	0.544**	0.598**	0.144*	0.336**	0.446**	1	
11	LnCoordination	0.206**	0.217**	0.272**	0.157*	0.408**	0.164*	0.137*	0.037	0.272**	0.279**	1

* . P< 0.05, **. P < 0.01

For the Azurix Project, the highest product moment correlation was between degree and betweenness centrality measures ($r = 0.827$, $p < 0.01$). The next highest correlation was between degree and eigenvector ($r = 0.826$, $p < 0.01$) followed by directed betweenness and undirected betweenness centrality ($r = 0.773$, $p < 0.01$). In-closeness had negative and non-significant correlations with out-closeness ($r = -0.202$, $p > 0.05$), eigenvector ($r = -0.040$, $p > 0.05$) and out-degree ($r = -0.015$, $p > 0.05$). The Pearson correlation results are presented in (Table 4.19).

Out-degree centrality was highly correlated with the coordination scores for Azurix project according to the Pearson correlation test ($r = 0.540$, $p < 0.01$) and the lowest correlation was between closeness and coordination ($r = -0.022$, $p > 0.05$).

The highest Spearman rank order correlation was between eigenvector and closeness measures ($r = 0.920$, $p < 0.01$). The next highest correlation was between information and degree centrality ($r = 0.851$, $p < 0.01$) followed by degree and closeness centrality ($r = 0.850$, $p < 0.01$). The correlation between in-closeness and out-closeness centrality ($r = -0.240$, $p <0.05$) is negative and significant, while the correlations between out-degree and in-closeness ($r = -0.141$, $p > 0.05$) and in-degree and out-closeness ($r = -0.042$, $p > 0.05$) are negative and non-significant for the Azurix data. The next lowest correlation was between and in-degree and out-degree and it is also not statistically significant. ($r = 0.024$, $p > 0.05$) The Spearman rank correlation results are presented in (Table 4.20).

Network centrality measures were weakly correlated with the coordination scores for Azurix project according to Spearman rank correlation test. Information centrality had the highest correlation with coordination scores ($r = 0.383$, $p < 0.01$) and the lowest correlation was between in-closeness and coordination ($r = 0.020$, $p > 0.05$).

For the Azurix data, the highest Kendall rank order correlation was between eigenvector and closeness centrality measures ($r=0.777$, $p<0.01$). Degree centrality had high correlations with information centrality ($r=0.721$, $p<0.01$) and betweenness centrality ($r=0.713$, $p<0.01$). The correlation between in-closeness and out-closeness ($r=-0.209$, $p<0.01$) is negative and significant, while the correlations between in-closeness and out-degree ($r=-0.104$, $p>0.05$) and out-closeness and in-degree ($r=-0.026$, $p>0.05$) are negative and non-significant for the Azurix data. The Kendall's rank correlation results are presented in (Table 4.21).

Network centrality measures were weakly correlated with the coordination scores for Azurix project according to Kendall's rank correlation test. Information centrality had the highest correlation with coordination scores ($r = 0.246$, $p < 0.01$) and the lowest correlation was between in-closeness and coordination ($r = 0.016$, $p > 0.05$).

Table 4.19. Pearson correlation coefficients among the directed and undirected centrality measures of Azurix data.

		1	2	3	4	5	6	7	8	9	10	11
1	Degree	1										
2	Closeness	0.393**	1									
3	Betweenness	0.827**	0.259*	1								
4	Eigenvector	0.826**	0.368**	0.505**	1							
5	Outdegree	0.566**	0.185	0.351**	0.561**	1						
6	Indegree	0.605**	0.213*	0.491**	0.618**	0.139	1					
7	Outcloseness	0.537**	0.289**	0.338**	0.671**	0.451**	0.200	1				
8	Incloseness	0.069	0.342**	0.150	-0.040	-0.015	0.249*	-0.202	1			
9	Betweenness-2	0.745**	0.198	0.773**	0.635**	0.415**	0.678**	0.306**	0.159	1		
10	Information	0.757**	0.495**	0.527**	0.750**	0.482**	0.484**	0.504**	0.185	0.457**	1	
11	LnCoordination	0.358**	-0.022	0.285**	0.379**	0.540**	0.244*	0.128	0.057	0.296**	0.362**	1

* . P< 0.05, **. P < 0.01

Table 4.20. Spearman rank correlation coefficients among the directed and undirected centrality measures of Azurix data.

		1	2	3	4	5	6	7	8	9	10	11
1	Degree	1										
2	Closeness	0.850^{**}	1									
3	Betweenness	0.849^{**}	0.719^{**}	1								
4	Eigenvector	0.790^{**}	0.920^{**}	0.568^{**}	1							
5	Outdegree	0.595^{**}	0.510^{**}	0.449^{**}	0.468^{**}	1						
6	Indegree	0.546^{**}	0.470^{**}	0.469^{**}	0.463^{**}	0.024	1					
7	Outcloseness	0.493^{**}	0.531^{**}	0.316^{**}	0.563^{**}	0.760^{**}	-0.042	1				
8	Incloseness	0.159	0.165	0.236^{*}	0.032	-0.141	0.616^{**}	-0.240^{*}	1			
9	Betweenness-2	0.624^{**}	0.580^{**}	0.636^{**}	0.504^{**}	0.574^{**}	0.490^{**}	0.444^{**}	0.300^{**}	1		
10	Information	0.851^{**}	0.826^{**}	0.659^{**}	0.827^{**}	0.563^{**}	0.532^{**}	0.534^{**}	0.130	0.574^{**}	1	
11	LnCoordination	0.274^{*}	0.301^{**}	0.300^{**}	0.270^{*}	0.291^{**}	0.174	0.095	0.020	0.273^{*}	0.383^{**}	1

* . $P< 0.05$, **. $P < 0.01$

Table 4.21. Kendall's tau (τ) rank correlation coefficients among the directed and undirected centrality measures of Azurix data.

		1	2	3	4	5	6	7	8	9	10	11
1	Degree	1										
2	Closeness	0.709**	1									
3	Betweenness	0.713**	0.554**	1								
4	Eigenvector	0.638**	0.777**	0.416**	1							
5	Outdegree	0.485**	0.383**	0.333**	0.354**	1						
6	Indegree	0.455**	0.360**	0.370**	0.344**	0.015	1					
7	Outcloseness	0.382**	0.391**	0.232**	0.432**	0.625**	-0.026	1				
8	Incloseness	0.117	0.098	0.169*	-0.015	-0.104	0.485**	-0.209**	1			
9	Betweenness-2	0.531**	0.460**	0.552**	0.384**	0.480**	0.396**	0.366**	0.233**	1		
10	Information	0.721**	0.661**	0.484**	0.666**	0.438**	0.430**	0.392**	0.076	0.447**	1	
11	LnCoordination	0.195*	0.200**	0.210**	0.185*	0.215**	0.127	0.066	0.016	0.204*	0.246**	1

* . P< 0.05, **. P < 0.01

4.4. Coordination Processes and Network Centrality Correlations

Correlations between network centrality measures and coordination scores are examined for the two case studies in order to find out which measure of centrality is the most useful predictor for each coordination process. Network centrality measures are weakly correlated but statistically significant with the four coordination scores for the two datasets.

4.4.1. Undirected Network Correlations

Correlation coefficients (Pearson's r) among five undirected centrality measures and four coordination scores of Azurix dataset are presented in Table 4.22.

Table 4.22. Correlation coefficients (Pearson's r) between coordination processes and the undirected centrality measures of Azurix network.

	Degree	Closeness	Betweenness	Eigenvector	Information
Resource Allocation	0.274*	0.031	0.094	0.357**	0.282**
Producer Consumer	0.351**	0.088	0.247*	0.337**	0.336**
Task Subtask	0.341**	0.039	0.186	0.408**	0.310**
Simultaneity Constraints	0.222*	0.051	0.027	0.321**	0.236*

* . P< 0.05, **. P < 0.01

There is some variability in the correlation among measures between two case studies. For the Azurix dataset, the highest product moment correlation was between task subtask coordination scores and eigenvector centrality ($r = 0.408$, $p < 0.01$). The next highest correlation was between resource allocation coordination scores and eigenvector centrality ($r = 0.357$, $p < 0.01$) followed by producer consumer coordination scores and degree centrality ($r = 0.351$, $p < 0.01$). The correlation between simultaneity constraints coordination scores and eigenvector centrality ($r = 0.321$, $p < 0.01$) was also the highest correlation among other centrality measures. The lowest correlation observed was between simultaneity constraints coordination scores and betweenness centrality ($r = 0.027$, $p < 0.05$).

Correlation coefficients (Pearson's r) among five undirected centrality measures and four coordination scores of Dabhol dataset are presented in Table 4.23

For the Dabhol dataset, the highest correlation was between producer consumer coordination scores and information centrality ($r = 0.254$, $p < 0.05$). The correlation between task subtask coordination scores and betweenness centrality ($r = 0.242$, $p < 0.01$) was the highest among other centrality measures. The highest correlation observed for simultaneity constraints coordination scores was between information centrality ($r = 0.197$, $p < 0.01$). The correlation between resource allocation coordination score and information centrality was the highest among other centrality measures ($r = 0.194$, $p < 0.01$). The correlation between resource allocation coordination score and closeness was the lowest ($r = 0.076$, $p > 0.05$) for the Dabhol dataset.

Table 4.23. Correlation coefficients (Pearson's r) between coordination processes and the undirected centrality measures of Dabhol network.

	Degree	Closeness	Betweenness	Eigenvector	Information
Resource Allocation	0.101	0.076	0.149	0.078	0.194
Producer Consumer	0.189	0.113	0.253*	0.156	0.254*
Task Subtask	0.186	0.107	0.242*	0.149	0.234*
Simultaneity Constraints	0.119	0.077	0.147	0.096	0.197*

* . P< 0.05, **. P < 0.01

Table 4.24. Spearman rank correlation coefficients between coordination processes and the undirected centrality measures of Azurix network.

	Degree	Closeness	Betweenness	Eigenvector	Information
Resource Allocation	0.241*	0.252*	0.254*	0.212*	0.292**
Producer Consumer	0.264*	0.217*	0.322**	0.118	0.291**
Task Subtask	0.172	0.190	0.164	0.198	0.287**
Simultaneity Constraints	0.103	0.183	0.092	0.224*	0.286**

* . P< 0.05, **. P < 0.01

Correlation coefficients (Spearman's r) among five undirected centrality measures and four coordination scores of Azurix dataset are presented in Table 4.24.

For the Azurix network, the highest rank order correlation was between producer consumer coordination scores and betweenness centrality ($r = 0.322$, $p < 0.01$). The next highest correlation was between resource allocation coordination scores and information centrality ($r = 0.292$, $p < 0.01$) followed by task subtask coordination scores and information centrality ($r = 0.287$, $p < 0.01$) and simultaneity constraints coordination scores and information centrality ($r = 0.286$, $p < 0.01$). The lowest correlation observed was between simultaneity constraints coordination scores and betweenness centrality ($r = 0.092$, $p > 0.05$).

Correlation coefficients (Spearman's r) among five undirected centrality measures and four coordination scores of Dabhol dataset are presented in Table 4.25.

For the Dabhol dataset, the highest correlation was between producer consumer coordination scores and information centrality ($r = 0.409$, $p < 0.01$). The next highest correlation was between simultaneity constraints coordination scores and information centrality ($r = 0.397$, $p < 0.01$) followed by task subtask coordination scores and betweenness centrality ($r = 0.327$, $p < 0.01$) and resource allocation coordination scores and information centrality ($r = 0.325$, $p < 0.01$). The lowest correlation observed was between resource allocation coordination scores and eigenvector centrality ($r = 0.092$, $p > 0.05$).

Table 4.25. Spearman rank correlation coefficients between coordination processes and the undirected centrality measures of Dabhol network.

	Degree	Closeness	Betweenness	Eigenvector	Information
Resource Allocation	0.147	0.193	0.212*	0.120	0.325**
Producer Consumer	0.289**	0.321**	0.336**	0.248*	0.409**
Task Subtask	0.211*	0.240*	0.327**	0.143	0.307**
Simultaneity Constraints	0.282**	0.276**	0.261**	0.245*	0.397**

* . P< 0.05, **. P < 0.01

Table 4.26. Kendall's tau (τ) coefficient between coordination processes and the undirected centrality measures of Azurix network.

	Degree	Closeness	Betweenness	Eigenvector	Information
Resource Allocation	0.196*	0.192*	0.183*	0.160*	0.225**
Producer Consumer	0.209*	0.167*	0.251**	0.087	0.222**
Task Subtask	0.123	0.135	0.124	0.139	0.196**
Simultaneity Constraints	0.085	0.142	0.072	0.174*	0.225**

* . P< 0.05, **. P < 0.01

Correlation coefficients Kendall's tau (τ) among five undirected centrality measures and four coordination scores of Azurix dataset are presented in Table 4.26.

For the Azurix network, the highest Kendall rank order correlation was between producer consumer coordination scores and betweenness centrality ($r = 0.251$, $p < 0.01$). The next highest correlations were between resource allocation coordination scores and information centrality ($r = 0.225$, $p < 0.01$) and simultaneity constraints coordination scores and information centrality ($r = 0.225$, $p < 0.01$) followed by task subtask coordination scores and information centrality ($r = 0.196$, $p < 0.01$). The lowest correlation observed was between simultaneity constraints coordination scores and betweenness centrality ($r = 0.072$, $p > 0.05$).

Correlation coefficients Kendall's tau (τ) among five undirected centrality measures and four coordination scores of Dabhol dataset are presented in Table 4.27.

For the Dabhol dataset, the highest correlation was between simultaneity constraints coordination scores and information centrality ($r = 0.307$, $p < 0.01$). The next highest correlation was between producer consumer coordination scores and information centrality ($r = 0.302$, $p < 0.01$) followed by task subtask coordination scores and betweenness centrality ($r = 0.242$, $p < 0.01$) and resource allocation coordination scores and information centrality ($r = 0.232$, $p < 0.01$). The lowest correlation observed was between resource allocation coordination scores and eigenvector centrality ($r = 0.086$, $p > 0.05$).

Table 4.27. Kendall's tau (τ) coefficient between coordination processes and the undirected centrality measures of Dabhol network.

	Degree	Closeness	Betweenness	Eigenvector	Information
Resource Allocation	0.105	0.130	0.154*	0.086	0.232**
Producer Consumer	0.217**	0.234**	0.250**	0.184*	0.302**
Task Subtask	0.146*	0.170*	0.242**	0.100	0.216**
Simultaneity Constraints	0.221**	0.207**	0.196**	0.186*	0.307**

* . P< 0.05, **. P < 0.01

4.4.2. Directed Network Correlations

A directional analysis is further performed to investigate the correlation between directed centrality measures and four coordination processes. Correlation coefficients (Pearson's r) among five directed centrality measures and four coordination scores of Azurix dataset are presented in Table 4.28.

Table 4.28. Correlation coefficients (Pearson's r) between coordination processes and the directed centrality measures of Azurix network.

	Outdegree	Indegree	Outcloseness	Incloseness	Betweenness
Resource Allocation	0.818**	0.072	0.183	-0.037	0.186
Producer Consumer	0.808**	0.143	0.205	0.052	0.281**
Task Subtask	0.703**	0.323**	0.206	0.057	0.373**
Simultaneity Constraints	0.858**	0.029	0.197	-0.020	0.176

* . P< 0.05, **. P < 0.01

For the Azurix dataset, the highest product moment correlation was between simultaneity constraints coordination scores and out-degree centrality ($r = 0.858$, $p < 0.01$). The next highest correlation was between resource allocation coordination scores and out-degree centrality ($r = 0.818$, $p < 0.01$) followed by producer consumer coordination scores and out-degree centrality ($r = 0.808$, $p < 0.01$). The correlation between task subtask coordination scores and out-degree centrality ($r = 0.703$, $p < 0.01$) was also the highest correlation among other directed centrality measures. The lowest correlation observed was between resource allocation coordination scores and in-closeness centrality ($r = -0.037$, $p > 0.05$).

Correlation coefficients (Pearson's r) among five directed centrality measures and four coordination scores of Dabhol dataset are presented in Table 4.29

For the Dabhol dataset, the highest correlation was between task subtask coordination scores and out-degree centrality ($r = 0.906$, $p < 0.01$) followed by simultaneity constraints coordination scores and out-degree centrality ($r = 0.897$, $p < 0.01$). The highest correlations observed for producer consumer coordination scores and resource allocation were identical and between out-degree centrality ($r = 0.874$, $p < 0.01$). The correlation between simultaneity constraints coordination scores and in-degree centrality was the lowest ($r = -0.021$, $p > 0.05$) for the Dabhol dataset.

Table 4.29. Correlation coefficients (Pearson's r) between coordination processes and the directed centrality measures of Dabhol network.

	Outdegree	Indegree	Outcloseness	Incloseness	Betweenness
Resource Allocation	0.874**	-0.018	0.053	0.062	0.143
Producer Consumer	0.874**	0.131	0.098	0.110	0.233*
Task Subtask	0.906**	0.022	0.086	0.069	0.223*
Simultaneity Constraints	0.897**	-0.021	0.076	0.041	0.126

* . P< 0.05, **. P < 0.01

Table 4.30. Spearman rank correlation coefficients between coordination processes and the directed centrality measures of Azurix network.

	Outdegree	Indegree	Outcloseness	Incloseness	Betweenness
Resource Allocation	0.194	0.160	-0.008	-0.021	0.163
Producer Consumer	0.209	0.212*	0.051	0.143	0.264*
Task Subtask	0.163	0.117	0.102	0.064	0.125
Simultaneity Constraints	0.244*	0.097	0.108	-0.076	0.138

* . P< 0.05, **. P < 0.01

Correlation coefficients (Spearman's r) among five directed centrality measures and four coordination scores of Azurix dataset are presented in Table 4.30.

For the Azurix dataset, the highest rank correlation was between producer consumer coordination scores and directed betweenness centrality ($r = 0.264$, $p < 0.05$). The next highest correlation was between simultaneity constraints coordination scores and out-degree centrality ($r = 0.244$, $p < 0.05$) followed by resource allocation coordination scores and out-degree centrality ($r = 0.194$, $p > 0.05$) and task subtask coordination scores out-degree centrality ($r = 0.163$, $p > 0.05$). The lowest correlation observed was between simultaneity constraints coordination scores and in-closeness centrality ($r = -0.076$, $p > 0.05$).

Correlation coefficients (Spearman's r) among five directed centrality measures and four coordination scores of Dabhol dataset are presented in Table 4.31

For the Dabhol dataset, the highest correlation was between task subtask coordination scores and out-degree centrality ($r = 0.481$, $p < 0.01$). The highest correlations observed for producer consumer coordination scores and simultaneity constraints coordination scores were identical and between out-degree centrality (r =

0.411, $p < 0.01$). The correlation between resource allocation coordination scores and out-degree centrality ($r = 0.389$, $p < 0.01$) was also the highest correlation among other centrality measures. The correlation between resource allocation coordination scores and in-closeness centrality was the lowest ($r = 0.013$, $p > 0.05$) for the Dabhol dataset.

Table 4.31. Spearman rank correlation coefficients between coordination processes and the directed centrality measures of Dabhol network.

	Outdegree	Indegree	Outcloseness	Incloseness	Betweenness
Resource Allocation	0.389**	0.125	0.095	0.013	0.166
Producer Consumer	0.411**	0.291**	0.128	0.149	0.290**
Task Subtask	0.481**	0.151	0.153	0.055	0.295**
Simultaneity Constraints	0.411**	0.165	0.064	0.053	0.191

* . P< 0.05, **. P < 0.01

Table 4.32. Kendall's tau (τ) coefficient between coordination processes and the directed centrality measures of Azurix network.

	Outdegree	Indegree	Outcloseness	Incloseness	Betweenness2
Resource Allocation	0.157	0.124	-0.003	-0.018	0.132
Producer Consumer	0.170*	0.175*	0.035	0.106	0.213*
Task Subtask	0.118	0.086	0.074	0.039	0.094
Simultaneity Constraints	0.203*	0.082	0.082	-0.061	0.116

* . P< 0.05, **. P < 0.01

Correlation coefficients Kendall's tau (τ) among five directed centrality measures and four coordination scores of Azurix dataset are presented in Table 4.32.

For the Azurix network, the highest Kendall rank order correlation was between producer consumer coordination scores and directed betweenness centrality ($r = 0.213$, $p < 0.05$). The next highest correlation was between simultaneity constraints coordination scores and out-degree centrality ($r = 0.203$, $p < 0.05$) followed by resource allocation coordination scores and out-degree centrality ($r = 0.157$, $p > 0.05$) and task subtask coordination scores out-degree centrality ($r = 0.118$, $p > 0.05$). The lowest correlation observed was between simultaneity constraints coordination scores and in-closeness centrality ($r = -0.061$, $p > 0.05$).

Correlation coefficients Kendall's tau (τ) among five directed centrality measures and four coordination scores of Dabhol dataset are presented in Table 4.33.

For the Dabhol dataset, the highest correlation was between task subtask coordination scores and out-degree centrality ($r = 0.356$, $p < 0.01$). The next highest correlation was between simultaneity constraints coordination scores and out-degree centrality ($r = 0.328$, $p < 0.01$) followed by producer consumer coordination scores and out-degree centrality ($r = 0.305$, $p < 0.01$) and resource allocation coordination scores and out-degree centrality ($r = 0.300$, $p < 0.01$). The correlation between resource allocation coordination scores and in-closeness centrality was the lowest ($r = 0.008$, $p > 0.05$) for the Dabhol dataset.

Table 4.33. Kendall's tau (τ) coefficient between coordination processes and the directed centrality measures of Dabhol network.

	Outdegree	Indegree	Outcloseness	Incloseness	Betweenness2
Resource Allocation	0.300**	0.080	0.071	0.008	0.125
Producer Consumer	0.305**	0.214**	0.097	0.110	0.223**
Task Subtask	0.356**	0.114	0.109	0.034	0.218**
Simultaneity Constraints	0.328**	0.130	0.049	0.040	0.149

* . P< 0.05, **. P < 0.01

4.5. Mann-Whitney U Test Results

The coordination scores from the group high in centrality were compared to those low in centrality. Mann-Whitney U-test is used to determine whether the two groups are statistically different from each other. The coordination scores and centrality measurements were sorted by centrality measurements and both mean and median were found.

4.5.1. Dabhol Project Results

4.5.1.1. Degree Centrality

Degree centrality measurement for the Dabhol data shows a ($p = 0.018$) for median and a ($p = 0.021$) for mean which are both statistically significant. There is a significant difference between two groups and the null hypothesis is rejected.

Table 4.34. The results of Mann-Whitney U test for degree centrality (median).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (median)	Independent-Samples Mann-Whitney U Test	0.018	Reject the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

Table 4.35. The results of Mann-Whitney U test for degree centrality (mean).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (mean)	Independent-Samples Mann-Whitney U Test	0.021	Reject the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

4.5.1.2. Closeness Centrality

Closeness centrality measurement for the Dabhol data shows a strong statistically significance ($p = 0.01$) for median. There is a significant difference between two groups and the null hypothesis is rejected. However it is not statistically significant for mean ($p = 0.097$). This means there is no statistical difference between the two groups and the null hypothesis is retained.

Table 4.36. The results of Mann-Whitney U test for closeness centrality (median).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (median)	Independent-Samples Mann-Whitney U Test	0.001	Reject the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

Table 4.37. The results of Mann-Whitney U test for closeness centrality (mean).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (mean)	Independent-Samples Mann-Whitney U Test	0.097	Retain the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

4.5.1.3. Betweenness Centrality

Betweenness centrality measurement for the Dabhol data shows a strong statistically significance for both median (p = 0.007) and for mean (p = 0.000). There is a significant difference between two groups and the null hypothesis is rejected.

Table 4.38. The results of Mann-Whitney U test for betweenness centrality (median).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (median)	Independent-Samples Mann-Whitney U Test	0.007	Reject the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

Table 4.39. The results of Mann-Whitney U test for betweenness centrality (mean).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (mean)	Independent-Samples Mann-Whitney U Test	0.000	Reject the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

4.5.1.4. Eigenvector Centrality

Eigenvector centrality measurement for the Dabhol data is not statistically significant for both median (p = 0.381) and mean (p = 0.051). This means there is no statistical difference between the two groups and the null hypothesis is retained.

Table 4.40. The results of Mann-Whitney U test for eigenvector centrality (median).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (median)	Independent-Samples Mann-Whitney U Test	0.381	Retain the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

Table 4.41. The results of Mann-Whitney U test for eigenvector centrality (mean).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (mean)	Independent-Samples Mann-Whitney U Test	0.051	Retain the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

4.5.1.5. Information Centrality

Information centrality measurement for the Dabhol data shows a strong statistically significance for both median ($p = 0.000$) and mean ($p = 0.002$). There is a significant difference between two groups and the null hypothesis is rejected.

Table 4.42. The results of Mann-Whitney U test for information centrality (median).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (median)	Independent-Samples Mann-Whitney U Test	0.000	Reject the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

Table 4.43. The results of Mann-Whitney U test for information centrality (mean).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (mean)	Independent-Samples Mann-Whitney U Test	0.002	Reject the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

Among the Dabhol centrality measures information centrality shows the strongest statistical significance among others ($p = 0.000$) for median and betweenness shows the strongest statistical significance ($p = 0.000$) for mean.

Table 4.44. The results of Mann-Whitney U test for undirected centrality measures (median).

Network Centrality Measure: Dabhol Project (Median)	
Degree	p = 0.018
Closeness	P = 0.001
Betweenness	P = 0.007
Eigenvector	P = 0.381
Information	P = 0.000

Table 4.45. The results of Mann-Whitney U test for undirected centrality measures (mean).

Network Centrality Measure: Dabhol Project (Mean)	
Degree	p = 0.021
Closeness	P = 0.097
Betweenness	P = 0.000
Eigenvector	P = 0.051
Information	P = 0.002

4.5.2. Azurix Project Results

4.5.2.1. Degree Centrality

Degree centrality measurement for the Azurix data shows a (p = 0.002) for median and a (p = 0.018) for mean which are both statistically significant. There is a significant difference between two groups and the null hypothesis is rejected.

Table 4.46. The results of Mann-Whitney U test for degree centrality (median).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (median)	Independent-Samples Mann-Whitney U Test	0.002	Reject the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

Table 4.47. The results of Mann-Whitney U test for degree centrality (mean).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (mean)	Independent-Samples Mann-Whitney U Test	0.018	Reject the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

4.5.2.2. Closeness Centrality

Closeness centrality measurement for the Azurix data shows a strong statistically significance (p = 0.014) for median. There is a significant difference between two groups and the null hypothesis is rejected. However it is not statistically significant for mean (p = 0.512). This means there is no statistical difference between the two groups and the null hypothesis is retained.

Table 4.48. The results of Mann-Whitney U test for closeness centrality (median).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (median)	Independent-Samples Mann-Whitney U Test	0.014	Reject the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

Table 4.49. The results of Mann-Whitney U test for closeness centrality (mean).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (mean)	Independent-Samples Mann-Whitney U Test	0.512	Retain the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

4.5.2.3. Betweenness Centrality

Betweenness centrality measurement for the Azurix data shows a strong statistically significance (p = 0.031) for median. There is a significant difference between two groups and the null hypothesis is rejected. However it is not statistically significant for mean (p = 0.066). This means there is no statistical difference between the two groups and the null hypothesis is retained.

Table 4.50. The results of Mann-Whitney U test for betweenness centrality (median).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (median)	Independent-Samples Mann-Whitney U Test	0.031	Reject the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

Table 4.51. The results of Mann-Whitney U test for betweenness centrality (mean).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (mean)	Independent-Samples Mann-Whitney U Test	0.066	Retain the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

4.5.2.4. Eigenvector Centrality

Eigenvector centrality measurement for the Azurix data shows a strong statistically significance (p = 0.007) for mean. There is a significant difference between two groups and the null hypothesis is rejected. However it is not statistically significant for median (p = 0.056). This means there is no statistical difference between the two groups and the null hypothesis is retained.

Table 4.52. The results of Mann-Whitney U test for eigenvector centrality (median).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (median)	Independent-Samples Mann-Whitney U Test	0.056	Retain the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

Table 4.53. The results of Mann-Whitney U test for eigenvector centrality (mean).

Null Hypothesis	Test	Sig.	Decision
The distribution of (LnCoordination) is the same across categories of low centrality and high centrality (mean)	Independent-Samples Mann-Whitney U Test	0.007	Reject the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

4.5.2.5. Information Centrality

Information centrality measurement for the Azurix data shows a strong statistically significance for both median (p = 0.001) and mean (p = 0.018). There is a significant difference between two groups and the null hypothesis is rejected.

Table 4.54. The results of Mann-Whitney U test for information centrality (median).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (median)	Independent-Samples Mann-Whitney U Test	0.001	Reject the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

Table 4.55. The results of Mann-Whitney U test for information centrality (mean).

Null Hypothesis	Test	Sig.	Decision
The distribution of coordination (LnCoordination) is the same across categories of low centrality and high centrality (mean)	Independent-Samples Mann-Whitney U Test	0.001	Reject the null hypothesis

Asymptotic significances are displayed. The significance level is 0.05.

Coordination scores of high and low groups of centrality were used and the statistical significance for each centrality measurement was compared against others. For Azurix centrality measures information centrality shows the strongest statistical significance among others (p = 0.001).

Table 4.56. The results of Mann-Whitney U test for undirected centrality measures (median).

Network Centrality Measure: Azurix Project (Median)	
Degree	p = 0.002
Closeness	p = 0.014
Betweenness	p = 0.031
Eigenvector	p = 0.056
Information	p = 0.001

Table 4.57. The results of Mann-Whitney U test for undirected centrality measures (mean).

Network Centrality Measure: Azurix Project (Mean)	
Degree	p = 0.018
Closeness	p = 0.512
Betweenness	p = 0.066
Eigenvector	p = 0.007
Information	p = 0.001

4.6. ANOVA Test Results

The one-way analysis of variance (ANOVA) is used to determine whether there are any significant differences between the means of information centrality and coordination scores.

4.6.1. Dabhol Project Results

A one-way ANOVA was conducted to test the differences in coordinative activity between the high and low groups of information centrality. Coordination scores differed significantly across the low and high groups of information centrality for median division as determined by one-way ANOVA, (F (1, 99) = 24.465, p = 0.000), indicating that the centrality positioned individuals show more coordination activity. Coordination scores also differed significantly across the low and high groups of information centrality for mean division as determined by one-way ANOVA, (F (1, 99) = 13.908, p = 0.000).

Table 4.58. The results of ANOVA test for the Dabhol dataset (median).

Dependent Variable: LnCoordination					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	27.297[a]	1	27.297	24.465	0.000
Intercept	1186.158	1	1186.158	1063.126	0.000
Median	27.297	1	27.297	24.465	0.000
Error	110.457	99	1.116		
Total	1320.467	101			
Corrected Total	137.753	100			

a. R Squared = 0.198 (Adjusted R Squared = 0.190)

Table 4.59. The results of ANOVA test for the Dabhol dataset (mean).

Dependent Variable: LnCoordination					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	16.968[a]	1	16.968	13.908	0.000
Intercept	1127.236	1	1127.236	923.925	0.000
Mean	16.968	1	16.968	13.908	0.000
Error	120.785	99	1.220		
Total	1320.467	101			
Corrected Total	137.753	100			

4.6.2. Azurix Project Results

Table 4.60. The results of ANOVA test for the Azurix dataset (median).

Dependent Variable: LnCoordination					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	15.905[a]	1	15.905	13.078	0.001
Intercept	818.560	1	818.560	67.093	0.000
Median	15.905	1	15.905	13.078	0.001
Error	103.370	85	1.216		
Total	935.321	87			
Corrected Total	119.275	86			

a. R Squared = 0.133 (Adjusted R Squared = 0.123)

A one-way ANOVA was conducted to test the differences in coordinative activity between the high and low groups of information centrality. Coordination scores differed significantly across the low and high groups of information centrality for median division as determined by one-way ANOVA, (F (1, 85) = 13.078, p = 0.001), indicating that the centrality positioned individuals show more coordination activity. Coordination scores also differed significantly across the low and high groups of information centrality for mean division as determined by one-way ANOVA, (F (1, 85) = 12.202, p = 0.001).

Table 4.61. The results of ANOVA test for the Azurix dataset (mean).

Dependent Variable: LnCoordination					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	14.973[a]	1	14.973	12.202	0.001
Intercept	813.400	1	813.400	662.874	0.000
Mean	14.973	1	14.973	12.202	0.001
Error	104.302	85	1.227		
Total	935.321	87			
Corrected Total	119.275	86			

a. R Squared = 0.126 (Adjusted R Squared = 0.115)

4.7. Discussions

4.7.1. Network Properties

The *network centralization* concept defined by Freeman (1979) gives an impression about the distribution of the centrality values. If some nodes have very high centrality scores and the vast majority of the nodes have very small centrality socres, the value of network centralization is much higher than in cases the centrality scores are almost equally distributed. The network centralization is defined as the sum of differences between the most central node and all other nodes, divided by a possible maximum of sum of differences. For calculating the network centralization un-normalized scores of degree centrality are used. Network centralization values are calculated for the undirected networks of Azurix and Dabhol projects. Due to the different sizes of the Axurix and Dabhol networks normalized degree centralization values are used in order to make comparison. The characteristics which affect network centralization are *network size*, *allow/ignore self-loops*, *directed/undirected networks* and *binary/weighted networks*. In the case studies, network centralization measures are calculated for both directed and undirected networks by using weighed networks instead of binary networks and ignoring the self-loops. Different combinations of these network characteristics result in different formulas and different values. Azurix project network has a 3.18% network centralization while Dabhol project network has a 1.57% network centralization. These results indicate that even though the network size of the Azurix project (87) is smaller than the network size of the Dabhol project (101) , the number of highly centralized nodes in the Azurix network are more than the ones in the Dabhol network.

The total degree centrality is created by summation of in- and out- degree centralies. In case of directed networks of Azurix and Dabhol projects the network centralization is separated into in-degree centralization and out-degree centralization. Azurix project has a 3.161% out-degree centralization and a 1.921% in-degree centralization. The network centralization is calculated depending on the sum of differences between the most central node and all other nodes. The results indicate that the sum of the differences between the highest out-degree centrality score and other out-degree centrality scores is higher than the sum of the differences between highest in-degree centrality scores and other in-degree centrality scores of Azurix network. Dabhol project has a 1.591% out-degree centralization and a 1.076% in-degree centralization which indicate the sum of the differences between the most central node and all other nodes are similar.

4.7.2. Network Centrality Correlations

Strong correlations are found among the 10 centrality measures presented in this research. According to Valente et al. (2008) too high correlation would indicate redundancy and too low correlation would indicate that the variables measured different things. The level of correlation between degree, betweenness, closeness, eigenvector and information centrality indicates that these measures are conceptually related. The correlations for the undirected measures were quite different from those for the directed versions.

In this study, Pearson product moment correlation, Spearman rank correlation and Kendall rank correlation tests were used. According to Rothenberg et al (1995) the rank order correlation tests demonstrate wider divergence than product moment correlation test.

According to Borgatti and Everet (2006) the degree and eigenvector measures are both radial and volume measures and they are similar enough on key attributes that they can be thought of as competitive that as potentially substitutable alternatives for each other. Closeness centrality and information centrality are radial and length measures and they are also similar on key attributes according to this classification. Among these measures it is reasonable to ask which is better. Borgatti and Everet (2006) explain that measures in different classes differ in fundamental ways, and are perhaps best viewed as complementary. Betweenness centrality is a medial and a volume measure (Borgatti and Everett 2006).

The findings of this study are in line with the study of Valente et al. (2008). The correlations between measures were also quite varied. In their study the highest correlation was between eigenvector centrality and degree (average r = 0.92). The next highest correlation was between symmetrized betweenness and degree (average r=0.85).

The lowest correlation between measures is between out-closeness and in-closeness (average r = 0.01). According to Valente et al. (2008) this is surprising, indicating that the direction of the calculation matters more than the property being measured by the algorithm. In this case, results of the Pearson product moment test show that the lowest correlation was between measures out-closeness and in-closeness centralities for Dabhol (r = -0.094) and Azurix datasets (r = -0.202). The Spearman and Kendall rank correlation tests also support that the lowest correlation is between in-closeness and out-closeness centrality but differentiated in the value of correlation coefficients.

Results of the Pearson product moment test show degree and eigenvector centralities are highly correlated for Dabhol (r = 0.924) and Azurix datasets (r = 0.826). These two centrality measures are also strongly correlated according to the Spearman and Kendall rank correlation test results but less well correlated compared to Pearson test results.

According to Bonacich (2007) eigenvector centrality capitalizes on how differences in degree can propagate through a network. It will not show differences if all nodes have the same degree. If one believes that differences in degree centrality, status, or power, than eigenvector centrality is called for. For example, sociometric popularity may be ultimately be based on being chosen by popular others. If each individual is limited to a fixed number of relationships one should not expect to find differences in eigenvector popularity (Bonacich 2007). In this case, depending on the correlation results between degree and eigenvector centrality it can be interpreted that individuals in the two project networks have similar degrees. Otherwise, degree and eigenvector centrality measures would be different from each other and they would be less correlated. Since degree and eigenvector centralities are substitutable alternatives for each other according to Borgatti and Everet (2006) the high correlation between these two centrality measures confirms their classification.

Rothenberg et al. (1995) found high correlation between degree and betweenness centralities (r = 0.961). They also found that information centrality was less well correlated with degree (r = 0.860) and betweenness centrality (r = 0.780) according to Pearson product moment test results. However the Spearman rank order correlation between degree and betweenness is the lowest observed (r = 0.423). Correlations of information centrality with degree (r = 0.829) and betweenness centrality (r = 0.600) were also lower than Pearson test results (Rothenberg et al. 1995)

The Pearson product moment test findings of this study show that degree and betweenness centralities are highly correlated for Dabhol (r = 0.766) and Azurix (r = 0.827) datasets.

They are also less well correlated with information centrality measure; Dabhol dataset degree (r = 0.746) and betweenness (r = 0.467) for Dabhol dataset, degree (r = 0.757) and betweenness (0.527) for Azurix dataset. The Spearman rank correlation test findings of this study show that degree and betweenness centralities are also highly correlated for Dabhol (r = 0.779) and Azurix (r = 0.849) datasets. Two correlations were observed to increase in Spearman test results. They are also strongly correlated with information centrality measure; degree (r = 0.839) and betweenness (r = 0.673) for Dabhol dataset, degree (r = 0.821) and betweenness (r = 0.659) for Azurix dataset. The findings of this study support the work of Rothenberg et al. (1995) and verify that these three centrality measures are strongly correlated among themselves. However in their study, Pearson Product moment test demonstrated higher correlations compared to Spearman rank order correlation tests. In this study depending on the non-parametric nature of the datasets, Spearman rank correlation test demonstrated higher correlations.

Eigenvector and closeness centrality measures are strongly correlated according to the Spearman and Kendall rank correlation test results for both Dabhol and Azurix networks. Spearman test results demonstrate high correlation between closeness and eigenvector centrality for Dabhol (r = 0.927) and Azurix (r = 0.920) datasets. Kendall test results demonstrate high correlation between closeness and eigenvector centrality for Dabhol (r = 0.778) and Azurix (r = 0.777) datasets. Eigenvector and closeness centrality measures are different in terms of their walk property (Borgatti and Everett 2006). Since they are not alternatives of each other as Borgatti and Everet (2006) explain, these measures are viewed as complementary.

Information and closeness centrality measures are moderately correlated according to Pearson product moment test results for Dabhol (r = 0.597) and Azurix (r = 0.495) datasets. However Spearman rank correlation results show high correlation between information and closeness centrality for Dabhol (r = 0.812) and Azurix (r = 0.826) datasets. In view of the fact that they are both length measures according to Borgatti and Everet (2006) they have similar on key attributes. Among these measures it is reasonable to investigate their relation with coordination. For the Dabhol dataset closeness centrality and coordination (r = 0.265) is lowly correlated according to Spearman rank correlation results nevertheless information centrality is moderately correlated with coordination (r = 0.420). For the Azurix dataset both of the closeness (r = 0.257) and information centrality (r = 0.336) is lowly correlated with coordination. Depending on these findings information centrality is better compared to closeness centrality in terms of relation with coordination.

According to all three correlation test results, information centrality has the highest correlation with coordination among other centrality measures with one exception. Only Pearson product moment correlation between eigenvector centrality and coordination (r = 0.379) is higher than the correlation between information centrality and coordination (r = 0.362) for Azurix project. In this case it is reasonable to question which correlation test might be more appropriate for this type of data.

According to Chok (2010) the Pearson product moment correlation is a natural parameter of association for a bivariate normal distribution. He points out that a statistical test based on the Pearson’s correlation coefficient is likely to be the most powerful for this type of data than similar tests on the other correlation coefficients. However, Pearson product moment correlation is influenced by outliers, unequal variances, non-normality and non-linearity. Thus, if the data contains outliers in one of both of the variables, Spearman’s rank order correlation coefficient is considered more appropriate. As the Spearman’s rank correlation coefficient is calculated by applying the Pearson correlation formula to the ranks of the data rather than to the actual data values themselves, many of the distortions of the Pearson correlation are reduced considerably.

Chok (2010) also indicates that the Pearson's correlation coefficient takes into account both the number and degree of concordances and discordances, whereas Kendall's tau correlation coefficient reflects only the numbers of concordances and discordances regardless of their degree. Thus, Kendall's tau is even less sensitive to outliers and is often preferred due to its simplicity and ease of interpretation. On the other hand Spearman's correlation is in between of the Pearson's and Kendall's, reflecting the degree of concordances and discordances on the rank scale. Due to all these deficiencies of the Pearson's correlation coefficient and Kendall's tau, the appropriateness of Spearman's statistical test for any type of interval data makes Spearman's correlation coefficient overall more preferable (Chok 2010).

This study aims to determine which centrality measures are relevant for analyzing coordination in a communication network. Previous researchers are also interested in finding out which centrality measure is the best predictor for coordination ability. However the SNA researches in the construction domain lacks information centrality researches apart from the work of Timurcan and Dogan (2014). They proposed a quantitative method to predict a firm's coordination score via a centrality index based on social network analysis. In this recent study, degree, closeness, betweenness and information centrality measures of the firms in a wayfinding signage project at a major airport construction were calculated and a centrality index was defined for each firm using the average of these centrality measures. The findings were compared to the previous work of Dogan et al. (2013). The correlation results show that the coordination scores of the firms are highly correlated with the centrality measures. Their work (2014) also proves that the utility of information centrality in the construction management domain is incontrovertible.

Hossain and Wu (2009) also used Spearman Rank test to measure correlation between coordination and directed measures of centrality. They argue that the Dabhol project shows clear results highlighting that out-centrality measurements correlated with coordination much stronger than in-centrality types. The correlation test results of this study successfully validates that out-degree centrality is the most potent predictor for coordination in directed networks. Out-degree centrality has the highest correlation with coordination according to each of the three correlation test results. The highest correlation ($r = 0.611$) is Pearson product moment correlation between out-degree centrality and coordination. In testing Hypothesis 3, out-degree centrality measure was found to be the best predicate for coordination.

4.7.3. Association Between Network Centrality and Coordination Processes

Correlations between network centrality measures and coordination scores are examined for the two case studies in order to find out which measure of centrality is the most useful predictor for each coordination process.

According to Malone and Crowston (1994) if coordination is defined as managing dependencies, then further progress should be possible by characterizing different kinds of dependencies and identifying the coordination processes that can be used to manage them. Managing shared resources, managing producer/consumer relationships, managing simultaneity constraints, managing task/subtask dependencies are defined as coordination processes by Malone and Crowston (1994). In this study association between network centrality measures and each coordination process is investigated in order to find the appropriate centrality measure to make inferences about each coordination process. Pearson product moment test, Spearman rank correlation and Kendall rank correlation test were used to measure the correlation between directed and undirected network centrality measures and coordination processes.

Coordination scores for each coordination process are calculated individually for each employee. Depending on the non-parametric nature of the data, non-parametric correlation tests demonstrated more accurate results. According to Spearman rank correlation test results information centrality has the highest correlation with resource allocation ($r = 0.325$), producer/consumer ($r = 0.409$), and simultaneity constraints ($r = 0.327$) among other centrality measures for the Dabhol project. The highest correlation for task/subtask dependencies is the one with the betweenness centrality ($r = 0.327$).

Information centrality has the highest correlation with resource allocation ($r = 0.292$), task/subtask ($r = 0.287$), and simultaneity constraints ($r = 0.286$) among other centrality measures for the Azurix project. The highest correlation differentiated for producer/consumer relationship and betweenness centrality ($r = 0.322$). Kendall test results also verified the Spearman rank correlation test results.

According to Spearman rank correlation test results out-degree has the highest correlation with each coordination processes: resource allocation (r = 0.389), producer/consumer (r = 0.411), task/subtask dependencies (r = 0.481) and simultaneity constraints (r = 0.411) among other centrality measures for the Dabhol project.

Out-degree centrality has the highest correlation with resource allocation (r = 0.194), task/subtask (r = 0.163), and simultaneity constraints (r = 0.244) among other centrality measures for the Azurix project. The highest correlation differentiated for producer/consumer relationship and directed betweenness centrality (r = 0.264). Kendall test results also verified the Spearman rank correlation test results.

These findings also support the aim of the study to investigate the importance of information centrality for predicting coordination. Information centrality is the best predictor for each coordination process among undirected centrality measures. Out-degree centrality is also the best predictor among directed centrality measures.

4.7.4. Mann-Whitney U and ANOVA Tests

Previous research carried out by Hossain and Wu (2006) and Hossain (2009) have found that all tests of undirected centrality used in their study were shown to be statistically different between high and low groups (Hossain et al. 2006). In our study the Mann-Whitney U test results of Dabhol dataset indicated that degree, closeness, betweenness and information centrality were statistically significant. The null hypothesis for Hypothesis 1 is that there is no statistical difference between the two groups is rejected for these centrality measures.

The eigenvector centrality was the weakest in terms of statistical strength for median division in this study. Degree, betweenness and closeness centralities were statistically significant for median division for both Dabhol and Azurix datasets and this supports the findings of Hossain (2009a). However, closeness centrality was the weakest in terms of statistical strength for mean division in this study for both Dabhol and Azurix datasets. Since Hossain (2009a) used only median division making any comparison is not possible. The difference between the results of mean and median division of closeness centrality could be depending on the distribution function of the data. The betweenness and degree centrality results are in the line with the findings of Hossain and Wu (2009).

Hossain and Wu (2009) suggest that betweenness is best independent predictor for coordination. Our findings indicate that the difference between coordination scores of high and low groups of information centrality is statistically more significant than high and low groups of betweenness centrality. It makes sense that information centrality uses all paths, not only geodesics but betweenness centrality assumes that communication only occurs along the shortest possible path, the geodesics. This assumption of betweenness centrality neglects measuring communication occurring along reachable, non-geodetic pathways. According to Stephenson and Zelen (1989) it is quite possible that information will take a more circuitous route either by random communication or may be intentionally channeled through many intermediaries in order to “hide” or “shield” information in a way not captured by geodesic paths.

ANOVA test is only performed for information centrality measure to test Hypothesis 2 as a complementary test for Mann Whitney U tests. AVOVA test results indicate that the centrality positioned individuals show more coordination activity. Coordination scores also differed significantly across the low and high groups of information centrality for the two project networks.

4.7.5. Limitations

Pentland (1995) argues that the lack of a universal organizational grammar meant that processes such as coordination are best studied using a context specific taxonomy. Depending on this argument, Hossain and Wu (2009) built a context specific taxonomy from the Enron corpus. In this study, the taxonomy built by Hossain and Wu (2009) was used. One of the limitations of this study is the integrity errors about the email addresses of the dataset. The e-mails of the employees are extracted from the "To" and "From" fields of the dataset to measure centrality and coordination scores. In some instances a single employee could have several e-mail addresses. For instance, steven.kean@enron.com and j..kean@enron.com are belong to the same person. Hossain et al. (2006) treated these e-mail addresses as discrete nodes. In this study they are treated as one single node.

Although the coordination key phrases and project keywords were adopted from a previous study of Hossain (2009a), different data extraction procedure and data cleansing process were used in this study. Therefore, sample sizes are different between the two studies. Hossain (2009a) found 173 people demonstrated coordination in the Dabhol project scope and after data cleansing 80 employees were included in the data. In this study 101 employees were included in the Dabhol dataset. In Hossain's study (2009a) 159 people found to demonstrated coordination in the Azurix project scope and after the data cleansing 97 employees were included in the Azurix dataset. In this study 87 employees were included in the Azurix dataset.

Another limitation of this study is that non-email communications which are not reflected in the data. Face-to-face and telephone communications and mailing should be taken into consideration for thorough results.

CHAPTER 5

CONCLUSION

This study analyzes the electronic communication and coordination in Enron Corporation's two major projects. Project participants' email communication data are used for analysis. Analysis revealed the coordination and communication performance depending on the coordination theory and the social network method. The study used the concept of information centrality in conjunction with other four standard centrality measures to examine the relation between the network centrality and coordination performance.

This study stresses the importance of information centrality for capturing and assessing the coordinative activity within a project network. The proposed approach was further investigated in an in-depth case study involving electronic information exchange between participants of two major projects of Enron Corporation.

Malone and Crowston (1990) defined the coordination as "body of principles about how activities can be coordinated, that is, about how actors can work together harmoniously". In today's information age, the e-mail communications constituted a good source of data for measuring the network relationships and the coordination patterns during the project process.

This study adopted Hossain's (2009a) three-phased methodology for coordination measure: (1) sentence extraction, (2) key phrase cataloguing – (3) weighted score assignment. The coordination key phrases data was derived from a precedent study (Hossain 2009a) for a broader examination of the relationship between information centrality and coordination. The computational procedures are provided with theoretical bases and methods.

The study was carried out in two consecutive phases: the centrality measures associated with the participants' coordination scores were thoroughly investigated and the statistical difference of coordination scores between high and low groups of centrality measures were analyzed and compared against each other in order to illustrate the reliability dimensions which may indicate the potential coordination performance.

The empirical evidence from the case study shows that the coordination scores are moderately correlated with centrality measures. This study also validates the procedure suggested by Hossain (2009a) for measuring the coordinative activity through network centrality.

All measures of centrality used in this study except eigenvector centrality were shown to be statistically significant between the high and low groups for median division. Also degree, betweenness and information centrality were found to be statistically different between high and low groups for mean division. Closeness and eigenvector centrality were not statistically different for mean division.

Findings from this study motivate that the thorough analyses using information centrality in communication networks may result more robust and illustrative findings for the coordination performance and information reliability. From a theoretical standpoint, the method for measuring information centrality proposed by Stephenson and Zelen (1989) makes use of all paths between nodes rather than geodesic paths. The calculations based on this framework can be easily adopted and organized for larger project networks.

In conclusion, Hypothesis 1 centrality positioned individuals in a network show more coordination is true for degree, betweenness and information centrality. Hypothesis 2 holds true that information centrality is the best predictor for coordination ability in the undirected graphs. The experiments on directed networks show that out-centrality measures are better predictor for coordination than in-centrality measures.

It can be claimed that the approach presented in this thesis associated with the proposed computational procedures has the application potentials for rapid analysis of complex project organizations, robustness of information exchange or specifically changes in project networks in the case of new potential network nodes added or subtracted (Stephenson and Zelen 1989). The proposed approach has advantages over complex and expensive computational methods and graphic techniques for analyzing project networks and organizations, in terms of simplicity and practicality.

This study investigated the coordinative activities and network centrality of only two major projects of Enron Corporation. More design and construction projects could be analyzed for evaluating the coordination performance of a project. Besides, communication and coordination performance of the Dabhol and Azurix projects were analyzed depending on the e-mail data. For further analysis face-to-face communications, telephone communications and mailing should be taken into consideration for thorough results. Researchers can create or can come up with new hypotheses and analyze the case of specific actors.

Advances in SNA research emphasize that there is more potential for SNA studies in the construction industry (Timurcan and Dogan 2013). Below summarizes Timurcan and Dogan's presented work in Creative Construction Conference in 2013. Three categories in SNA advances, which are promising for construction research, are classified to include (1) SNA automation, (2) team formation, and (3) performance evaluation.

With the recent developments in telecommunication and information technologies, virtual teams have widely emerged in the construction industry as well. Participants from different disciplines, organizations and even countries utilized technological innovations to collaborate irrespective of their physical locations. These novel design and construction teams use virtual collaboration tools such as e-mail, teleconference, videoconference, and virtual workspaces. Adoption of BIM technology has also become an inter-linkage of construction project actors. It might be possible to discover interaction patterns of the virtual construction teams via SNA augmented BIM. The project manager can monitor each member's performance in the network and could easily assign specific tasks and roles. Automating SNA in a BIM environment could provide a solid decision support system and even self-regulation for the realization of the whole network.

Application of SNA on team formation is an emerging area of interest across many disciplines. Current SNA related construction research allows us to see which project participants are already collaborating; how often the participants are working together; and how many others in the project team a given participant can reach through intermediary participants (Wambeke al. 2012; Hossain 2009a). Building projects are sets of links that organize professionals, teams, and firms in a connected way around a common purpose. During the implementation of a project, several problems related to the project network may arise, such as very isolated participants or peripheral members of a network may occur; or groups that fully concentrate on relationships may arise; or firms who are the only link between two distinct groups may appear; or the accumulation of team members in isolated points may take place (Monclar et al. 2011). Those inconveniences result in a series of communication and coordination problems, which ultimately cause knowledge losses to the project (Hossain 2009a). Thus, it is vital for project success to conduct a series of collaboration and coordination analysis via

SNA, by which we can monitor current coordination success; and then assist in the formation of new teams. The construction industry is based on temporary networks where project participants are regrouped on almost every project with little regard to past network connections. The use of SNA metrics can provide valuable information about the informal structure of the organization to understand and verify how collaboration occurs in a multi-disciplinary, multi-team project, which is vital when creating new teams (Cheatham and Cleereman 2006; Monclar et al. 2011).

Traditional performance measurement techniques are insufficient when the examined team represents a complex and dynamic environment like virtual collaboration or BIM setting (Arling 2007). SNA has been used for measuring the performance of project teams in many other disciplines (Damian et al. 2007; Fitsilis et al. 2009). Nowadays performance measurement has become more vital and complicated with the increase of virtuality in construction projects.

SNA can capture the virtual work that is done to complete tasks by identifying the contacts of each member and frequencies (Arling 2007). SNA provides quantitative data on existing coordination patterns between the project participants. The ratio between the physical communication and electronic communication can facilitate determining the efficient percentage of virtuality during the building design process. The type of the electronic communication between the project participants can be identified and also analyzed to find out the effective types of communication in order to accomplish tasks. The ability to compare the types and the amounts of communication will contribute to the performance measurement of virtuality in construction projects and examine the effects of who to connect whom via what type of communication on the project outcomes (Arling 2007).

The motivation for understanding coordination is to minimize the costs associated with coordination. A coordinating node must invest time and effort to maintain the networks and communication infrastructure. Studies have shown that vertically hierarchical network structures increase efficiency and reduce coordination costs. This is because a hierarchical structure has a chain of command, and thus there is an established medium on which to pass strategic information (Hossain 2009b).

According to Hossain (2009b) a free-forming dynamic network structure in which an actor has many adjacent nodes would result in the actor having to invest more time and effort to build the network in order to maintain its social capital. In reality however, people tend to communicate through these dynamic network structures, thus limiting the effectiveness of the formal hierarchy. A study of network centrality aims to harness the benefits of these dynamic networks and to improve coordinative activities.

This study sets out to determine the extensibility of these ideas about coordination at the construction management area. In the future, the researcher would like to expand this work by creating a tool to enable more systematic collaborative team formation. This tool will allow a user to specify what concepts the team needs to be familiar with and to what degree. This type of information can be gathered from the social network and fed into this tool to suggest an appropriate team based on the user's parameters. The participants who would work well in a group setting can be preferred or a successful specialist from a previous project network can be included in the new project team.

In addition, users should be able to specify the importance of the features such as; team members are trusted experts, have worked together in the past, etc. SNA offers insight in assigning the roles and tasks to team members depending on the type and level of coordination in the team. The inferences both at the organization level and individual level might make it available to examine the optimal level of virtual communication to form a highly coordinated construction team.

BIBLIOGRAPHY

Ahuja, M. K., Galletta, D. F. and Carley, K. M. (2003). "Individual centrality and performance in virtual R&D groups: An empirical study." *Management Science,* 49(1), 21-38.

Aljassmi, H., Han, S. and Davis, S. (2014). "Project pathogens network: new approach to analyzing construction-defects-generation mechanisms." *Journal of Construction Engineering and Management*, 140(1), 04013028.

Alsamadani, R., Hallowell, M. and Javernick-Will, A. N. (2012). "Measuring and modelling safety communication in small work crews in the US using social network analysis." *Construction Management and Economics,* 31(6), 568-579.

Alsamadani, R., Hallowell, M., Javernick-Will, A. N. and Cabello, J. (2013). "Relationships among language proficiency, communication patterns, and safety performance in small work crews in the United States." *Journal of Construction Engineering and Management*, 139(9), 1125-1134.

Arling, P. (2007). "Redefining and Measuring Virtual Work in Teams: An Application of Social Network Analysis." *Proc., 40th Annual Hawaii International Conference on System Sciences*, Waikoloa, HI, 42-52.

Arriagada D., R. E. and Alarcón C., L. F. (2014). "Knowledge management and maturation model in construction companies." *Journal of Construction Engineering and Management*, 140(4), B4013006.

Badiru, A. B. (2012). *Triple C model of project management: Communication, cooperation, and coordination*, CRC Press.

Barnes, J. A. (1954). *Class and committees in a Norwegian island parish*. Plenum.

Bavelas, A. (1950). "Communication patterns in task oriented groups. " *Journal of the Acoustical Society of America*, 57, 271-282.

Begel, A. (2008). "Effecting change: Coordination in large-scale software development." *Proc., International workshop on Cooperative and Human Aspects of Software Engineering,* ACM.

Bekkerman, R., McCallum, A., Huang, G. (2004). "Automatic categorization of email into folders: Benchmark experiments on Enron and SRI corpora." *Computer Science Department Faculty Publication Series,* Paper 218.

Bonacich, P. (1972). "Factoring and weighting approaches to status scores and clique identification." *Journal of Mathematical Sociology*, 2(1), 113-120.

Bonacich, P. (1987). "Power and centrality: A family of measures." *American Journal of Sociology*, 1170-1182.

Bonacich, P. (2007). "Some unique properties of eigenvector centrality." *Social networks*, 29(4), 555-564.

Bonacich, P. and Lloyd, P. (2001). "Eigenvector-like measures of centrality for asymmetric relations." *Social networks*, 23(3), 191-201.

Borgatti, S. P., Everett, M. G. and Freeman, L. C. (2002). "UCINET 6 for Windows: Software for social network analysis." <http://www.analytictech.com> (Dec. 5, 2013)

Borgatti, S. P. (2005). "Centrality and network flow." *Social networks*, 27(1), 55-71.

Borgatti, S. P. and Everett, M. G. (2006). "A graph-theoretic perspective on centrality." *Social networks,* 28(4), 466-484.

Cheatham, M. and Cleereman, K. (2006). "Application of Social Network Analysis to Collaborative Team Formation." *Proc., International Symposium on Collaborative Technologies and Systems (CTS),* 306-311.

Chitkara, K. K. (1998). *Construction project management*, Tata McGraw-Hill.

Chinowsky, P. S., Diekmann, J. and Galotti V. (2010). "Project organizations as social networks." *Journal of Construction Engineering and Management*, 136(4), 452-458

Choc, N. S. (2010). "Pearson's versus Spearman's and Kendall's Correlation Coefficients for Continuous Data." M. S. Thesis, School of Public Health, University of Pittsburgh.

Comu, S., Iorio, J., Taylor, J. and Dossick, C. (2013). "Quantifying the impact of facilitation on transactive memory system formation in global virtual project networks." *Journal of Construction Engineering and Management,* 139(3), 294-303.

Corrada-Emmanuel, A., McCallum, A., Smyth, P., Steyvers, M. and Chemudugunta, C. (2005). "Social network analysis and topic discovery for the Enron email dataset." *Proc., SIAM International Conference in Data Mining*, Newport Beach, CA.

Crowston, K. (1997). "A coordination theory approach to organizational process design." *Organization Science,* 8(2), 157-175.

Crowston, K., Rubleske, J. and Howison, J. (2006). "Coordination theory: A ten-year retrospective." *Human-computer Interaction and Management Information Systems: Foundations*, M. E. Sharpe, Inc., 120-138.

Dan, S. (2013). "Theorizing coordination: towards a novel theoretical framework." *Proc., European Group for Public Administration*, Edinburgh.

Damian, D., Marczak, S. and Kwan, I. (2007). "Collaboration Patterns and the Impact of Distance on Awareness in Requirements-Centred Social Networks." *Proc., 15th IEEE International Requirements Engineering Conference*, 59-68.

Deng, X., T. Chen and Pan, D. (2008). "Organizational coordination theory and its application in virtual enterprise." *Research and Practical Issues of Enterprise Information Systems II Volume 1*, 254: 311-316.

Diesner, J. and Carley, K. M. (2005). "Exploration of communication networks from the Enron email corpus." *Proc., SIAM International Conference in Data Mining*, Newport Beach, CA.

Dogan, S. Z., Gunhan, S. and Erbasaranoglu B. (2012). "Coordination process and network centrality in ISGI Airport's wayfinding project." *Proc., Construction Research Congress 2012: Construction Challenges in a Flat World*, 708-717.

Dogan, S. Z., Arditi, D., Gunhan S. and Erbasaranoglu, B. (2013). "Assessing coordination performance based on centrality in an E-mail communication network." *Journal of Management in Engineering*, Underreview.

El-Sheikh, A. and Pryke, S. (2010). "Network gaps and project success." *Construction Management and Economics*, 28(12), 1205-1217.

Fayol, H. (1949). *General and industrial management*, Pitman.

Fitsilis, P., Gerogiannis, V., Anthopoulos, L. and Kameas, A. (2009). "Using social network analysis for software project management." *Proc., International Conference on the Current Trends in Information Technology (CTIT)*, Dubai, 1-6.

Freeman, L. C. (1979). "Centrality in social networks conceptual clarification." *Social networks*, 1(3), 215-239.

Hai, T. K., Yusof, A. M., Ismail, S. and Wei, L. F. (2012). "A Conceptual Study of Key Barriers in Construction Project Coordination." *Journal of Organizational Management Studies*, 2012, doi: 10.5171/2012.795679.

Hanneman, R. A. and Riddle, M. (2005). *Introduction to social network methods*, University of California, Riverside, CA.

Herbsleb, J. D. and Mockus, A. (2003). "An empirical study of speed and communication in globally distributed software development." *Software Engineering*, 29(6), 481-494.

Higgin, G. and Jessop, N. (1965). *Communications in the building industry: The report of a pilot study*, Routledge.

Hossain, L., Wu, A. and Chung, K. K. S. (2006). "Actor centrality correlates to project based coordination." *Proc., 20th anniversary conference on Computer supported cooperative work*, ACM, Banff, Alberta, Canada, 363-372.

Hossain, L. (2009a). "Communications and coordination in construction projects." *Construction Management and Economics*, 27(1), 25-39.

Hossain, L. (2009b). "Effect of organisational position and network centrality on project coordination." *International Journal of Project Management*, 27(7), 680-689.

Hossain, L. and Wu, A. (2009). "Communications network centrality correlates to organisational coordination." *International Journal of Project Management,* 27(8), 795-811.

Iyer, K. C. and Jha, K. N. (2005). "Factors affecting cost performance: evidence from Indian construction projects." *International Journal of Project Management,* 23(4), 283-295.

Jha, K. and Iyer, K. (2007). "Commitment, coordination, competence and the iron triangle." *International Journal of Project Management*, 25(5), 527-540.

Jha, K. N. and Misra, S. (2007). "Ranking and classification of construction coordination activities in Indian projects." *Construction Management and Economics,* 25(4), 409-421.

Kessler, G. (2010). "Virtual business: An Enron email corpus study." *Journal of Pragmatics,* 42(1), 262-270.

Klimt, B. and Yang, Y. (2004). *Introducing the Enron corpus,* CEAS.

Kubicki, S., Bignon, J. C., Gilles, H. and Pascal, H. (2006). "Assistance to building construction coordination–towards a multi-view cooperative platform." *Electronic Journal of Information Techonology in Construction*, 11, 565-586.

Lamond, D. (1998). "Back to the future: Lessons from the past for a new management era." *Management Theory and Practice*, Macmillan Education Australia, South Yarra, 3-14.

Larsen, G. D. (2011). "Understanding the early stages of the innovation diffusion process: awareness, influence and communication networks." *Construction Management and Economics,* 29(10), 987-1002.

Leavitt, H. J. (1951). "Some effects of certain communication patterns on group performance." *Journal of Abnormal and Social Psychology,* 46, 38–50.

Lee, C. Y. (2006). "Correlations among centrality measures in complex networks." *arXiv preprint physics,* 13.

Loosemore, M. (1998a). "Social network analysis: using a quantitative tool within an interpretative context to explore the management of construction crises." *Engineering, Construction and Architectural Management*, *5*(4), 315-326.

Loosemore, M. (1998b). "The influence of communication structure upon management efficiency." *Construction Management and Economics,* 16(6), 661-671.

Malone, T. W. (1988). "What is coordination theory?" (Working Paper #2051-88), Massachusetts Institute of Technology, Sloan School of Management, Cambridge, MA.

Malone, T. W. and Crowston, K. (1990). "What is coordination theory and how can it help design cooperative work systems?" *Proc., ACM conference on Computer-supported cooperative work,* ACM, Los Angeles, California, USA, 357-370.

Malone, T. W. and Crowston, K. (1994). "The interdisciplinary study of coordination." *ACM Computing Surveys,* 26(1), 87-119.

March, C. (2009). *Business organisation for construction*, Taylor & Francis.

McLean, B. and Elkind P. (2004). *The smartest guys in the room: The amazing rise and scandalous fall of Enron*, Penguin Books.

Moreno, J. L. (1934). *Who shall survive?: a new approach to the problem of human interrelations*, Nervous and Mental Disease Publishing Co, Washington.

Monclar, R. S., Oliveira, J., Faria, F. F. d., Ventura, L., Souza, J. M. d. and Campos, M. L. M. (2011). "Using social networks analysis for collaboration and team formation identification." *Proc., 15th International Conference on Computer Supported Cooperative Work in Design (CSCWD),* 562-569.

Musiat, K., Kazienko, P. and Bródka, P. (2009). "User position measures in social networks." *Proc., Workshop on Social Network Mining and Analysis*, ACM, Paris, France, 1-9.

Naimisha, K. (2008). "Applications of social network analysis to community dynamics. " M. S. Thesis, Supercomputer Education and Research Centre Indian Institute of Science.

Otte, E. and Rousseau, R. (2002). "Social network analysis: a powerful strategy, also for the information sciences." *Journal of information Science*, 28(6), 441-453.

Park, H., Han, S. H., Rojas, E., Son, J. and Jung, W. (2011). "Social network analysis of collaborative ventures for overseas construction projects." *Journal of Construction Engineering and Management*, 137(5), 344-355.

Pentland, B. T. (1995). "Grammatical models of organizational processes." *Organization Science*, 6(5), 541-556.

Pocock, J. B., Hyun, C. T. Liu, L. Y. and Kim, M. K. (1996). "Relationship between project interaction and performance indicators." *Journal of Construction Engineering and Management* 122(2), 165-176.

Pryke, S. D. (2004). "Analysing construction project coalitions: exploring the application of social network analysis." *Construction Management and Economics,* 22(8), 787-797.

Pryke, S. D. (2005). "Towards a social network theory of project governance." *Construction Management and Economics,* 23(9), 927-939.

Rothenberg, R. B., Potterat, J. J., Woodhouse, D. E., Darrow, W. W., Muth, S. Q. and Klovdahl, A. S. (1995). "Choosing a centrality measure: Epidemiologic correlates in the Colorado Springs study of social networks." *Social Networks,* 17(3–4), 273-297.

Ruan, X., Ochieng, E. G., Price, A. D. F. and Egbu, C. O. (2012). "Knowledge integration process in construction projects: a social network analysis approach to compare competitive and collaborative working." *Construction Management and Economics*, 30(1), 5-19.

Sanaei, M., Javernick-Will, A. N. and Chinowsky, P. (2013). "The influence of generation on knowledge sharing connections and methods in construction and engineering organizations headquartered in the US." *Construction Management and Economics*, 31(9), 991-1004.

Saram, D. D. d. and Ahmed, S. M. (2001). "Construction coordination activities: What is important and what consumes time." *Journal of Management in Engineering* 17(4), 202-213.

Saram, D. D. d. (2002). "Measuring the quality of contractors' co-ordination activities during the construction process." Doctoral dissertation, The Hong Kong Polytechnic University.

Schilling, M. A. and Phelps, C. C. (2007). "Interfirm collaboration networks: The impact of large-scale network structure on firm innovation." *Management Science,* 53(7), 1113-1126.

Scott, J. (2000). *Social network analysis: A handbook*, SAGE Publications.

Shen, F. and Chang, A. (2011). "Exploring coordination goals of construction projects." *Journal of Management in Engineering*, 27(2), 90-96.

Shetty, J. and Adibi, J. (2004). *The Enron email dataset database schema and brief statistical report,* Information Sciences Institute Technical Report, University of Southern California.

Soh, C. K. and Wang, Z. (2000). "Parametric coordinator for engineering design." *Journal of Computing in Civil Engineering,* 14(4), 233-240.

Solis, F., Sinfield, J. and Abraham, D. (2013). "Hybrid approach to the study of inter-organization high performance teams." *Journal of Construction Engineering and Management,* 139(4), 379-392.

Stephenson, K. and Zelen, M. (1989). "Rethinking centrality: Methods and examples." *Social Networks,* 11(1), 1-37.

Tallberg, J. (2004). "The power of the presidency: Brokerage, efficiency and distribution in EU negotiations." *Journal of Common Market Studies,* 42(5), 999-1022.

Tellioglu, H. (2010). "Coordination of work: towards a typology." *Proc., 11th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing on International Conference on Computer Systems and Technologies*, ACM, Sofia, Bulgaria, 311-316.

Thorpe, T. and Mead, S. (2001). "Project-specific web sites: Friend or foe?" *Journal of Construction Engineering and Management*, 127(5), 406-413.

Timurcan, P. I. and Dogan, S. Z. (2013). "A Review of Research Trends in Social Network Analysis." *Proc., Creative Construction Conference 2013,* Budapest, Hungary, 776-786.

Timurcan, P. I. and Dogan, S. Z. (2014). "Evaluating the Project Coordination Performance Using Information Centrality." *Proc., Construction Research Congress 2014,* ASCE, Atlanta, Georgia, 1119-1128.

Tsai, C. C. and Wen, M. L. (2005). "Research and trends in science education from 1998 to 2002: a content analysis of publication in selected journals." *International Journal of Science Education*, 27(1), 3-14.

Van de Ven, A. H. (1976). "On the nature, formation, and maintenance of relations among organizations." *Academy of Management Review,* 1(4), 24-36.

Valente, T. W., Coronges, K., Lakon, C. and Costenbader, E. (2008). "How correlated are network centrality measures?" *Connections (Toronto, Ont.),* 28(1), 16-26.

Wambeke, B., Liu, M. and Hsiang, S. (2012). "Using pajek and centrality analysis to identify a social network of construction trades." *Journal of Construction Engineering and Management,* 138(10), 1192-1201.

Wasserman, S. and Faust, K. (1994). "Centrality and prestige." *Social network analysis : methods and applications,* Cambridge University Press, Cambridge, New York, 169-219.

West, J. (2014). "Collaborative patterns and power imbalance in strategic alliance networks." *Journal of Construction Engineering and Management,* 140(6), 04014010.

Wood, J. C. and Wood, M. C. (2002). *Henri Fayol: Critical evaluations in business and management*, Routledge.

Xue, X., Wang, Y., Shen, Q. and Yu, X. (2007). "Coordination mechanisms for construction supply chain management in the internet environment." *International Journal of Project Management*, 25(2), 150-157.

Zhang, L., He, J. and Zhou, S. (2013). "Sharing tacit knowledge for integrated project team flexibility: Case study of integrated project delivery." *Journal of Construction Engineering and Management* 139(7), 795-804.

APPENDIX A

NETWORK CENTRALITY AND COORDINATION RESULTS

Table A.1. Network Centrality Results of Azurix Dataset.

Employee	Degree	Closeness	Betweenness	Eigenvector	Information	Out-degree	In-degree	Out-closeness	In-closeness	Betweenness-2
A-ACT1	1.163	8.285	0.000	0.212	0.796	2.000	1.000	1.249	2.402	0.000
A-ACT2	6.977	9.338	4.400	4.498	1.264	3.000	3.000	1.347	1.704	153.250
A-ACT3	1.163	8.366	0.000	0.218	0.473	1.000	0.000	1.163	1.149	0.000
A-ACT4	8.140	9.588	0.611	31.825	1.525	14.000	0.000	3.439	1.149	0.000
A-ACT5	12.791	9.641	2.763	37.872	1.663	68.000	4.000	3.300	1.576	205.667
A-ACT6	5.814	9.081	0.863	3.366	1.165	3.000	3.000	1.176	1.641	27.000
A-ACT7	1.163	9.034	0.000	5.753	0.916	1.000	2.000	3.229	1.578	0.000
A-ACT8	2.326	8.423	0.009	0.555	0.750	2.000	0.000	1.176	1.149	0.000
A-ACT9	2.326	8.875	0.000	1.453	1.075	0.000	2.000	1.149	1.724	0.000
A-ACT10	5.814	9.430	1.175	12.758	1.283	2.000	5.000	1.163	1.739	7.667
A-ACT11	3.488	9.024	0.614	2.561	1.145	2.000	4.000	1.249	2.422	13.333
A-ACT12	3.488	8.921	0.536	2.320	0.894	3.000	3.000	1.330	1.620	29.833
A-ACT13	10.465	9.430	1.690	31.289	1.588	21.000	2.000	3.410	1.163	2.333
A-ACT14	1.163	9.062	0.000	5.815	0.948	1.000	0.000	3.383	1.149	0.000
A-ACT15	4.651	9.409	0.164	12.775	1.190	1.000	3.000	1.176	1.601	8.100
A-ACT16	2.326	9.168	0.000	5.981	0.904	0.000	2.000	1.149	1.599	0.000
A-ACT17	16.279	9.461	14.929	7.687	1.454	14.000	0.000	3.754	1.149	0.000
A-ACT18	2.326	9.297	1.733	2.614	0.896	0.000	2.000	1.149	1.614	0.000
A-ACT19	3.488	9.062	0.000	2.163	1.147	2.000	4.000	1.348	1.639	0.000
A-ACT20	1.163	9.034	0.000	5.753	0.633	1.000	1.000	3.229	1.578	0.000
A-ACT21	2.326	9.043	0.056	2.636	0.886	0.000	2.000	1.149	1.615	0.000
A-ACT22	2.326	8.129	0.000	0.281	0.657	2.000	0.000	1.176	1.149	0.000

(Cont. on next page)

Table A.1. (Cont.)

Employee	Degree	Closeness	Betweenness	Eigenvector	Information	Out-degree	In-degree	Out-closeness	In-closeness	Betweenness-2
A-ACT23	1.163	8.285	0.000	0.212	0.796	1.000	2.000	1.249	2.402	0.000
A-ACT24	4.651	8.758	1.075	1.256	0.914	0.000	4.000	1.149	1.205	0.000
A-ACT25	17.442	9.641	6.639	29.442	1.551	15.000	2.000	3.324	1.570	78.383
A-ACT26	11.628	9.609	2.555	35.386	1.598	21.000	0.000	3.836	1.149	0.000
A-ACT27	12.791	9.762	6.560	20.730	1.495	11.000	2.000	3.290	1.579	181.333
A-ACT 28	8.140	9.257	0.136	28.453	1.542	3.000	16.000	3.277	1.575	23.000
A-ACT29	6.977	9.588	4.871	12.094	1.227	0.000	7.000	1.149	1.683	0.000
A-ACT30	6.977	9.503	4.830	5.718	1.374	9.000	4.000	1.349	1.647	73.083
A-ACT31	18.605	9.862	12.132	48.784	1.654	10.000	38.000	3.346	1.576	460.617
A-ACT32	2.326	8.903	0.000	1.436	1.081	1.000	5.000	1.249	2.429	0.000
A-ACT33	5.814	9.471	0.320	18.634	1.459	5.000	10.000	1.163	1.629	2.333
A-ACT34	2.326	9.227	2.257	1.760	1.050	2.000	1.000	1.264	1.723	43.750
A-ACT35	4.651	9.419	0.893	16.138	1.486	8.000	4.000	3.280	1.576	43.000
A-ACT36	3.488	8.977	0.534	8.521	1.457	0.000	12.000	1.149	1.648	0.000
A-ACT37	3.488	9.043	2.176	1.833	0.877	0.000	3.000	1.149	1.763	0.000
A-ACT38	3.488	9.399	0.112	10.304	1.143	1.000	4.000	1.163	1.738	4.400
A-ACT39	2.326	9.168	0.000	5.981	0.904	0.000	2.000	1.149	1.599	0.000
A-ACT40	19.767	9.954	17.518	44.237	1.622	15.000	9.000	3.329	1.581	614.717
A-ACT41	2.326	8.921	0.378	1.500	0.999	3.000	1.000	1.264	1.638	5.000
A-ACT42	3.488	9.338	0.044	11.733	1.091	1.000	2.000	3.231	1.575	1.750
A-ACT43	1.163	1.190	0.000	0.000	0.826	0.000	1.000	1.149	1.176	0.000
A-ACT44	2.326	8.949	0.018	6.575	1.069	1.000	2.000	1.163	1.163	0.333
A-ACT45	6.977	9.806	10.903	18.787	1.368	4.000	2.000	3.259	1.576	111.583
A-ACT46	3.488	8.704	0.381	1.469	0.922	1.000	2.000	1.263	1.176	8.000

(Cont. on next page)

Table A.1. (Cont.)

Employee	Degree	Closeness	Betweenness	Eigenvector	Information	Out-degree	In-degree	Out-closeness	In-closeness	Betweenness-2
A-ACT47	6.977	9.358	0.817	18.966	1.396	5.000	3.000	3.284	1.572	21.900
A-ACT48	8.140	9.503	3.544	11.704	1.386	1.000	7.000	1.362	1.681	96.500
A-ACT49	5.814	9.091	2.108	3.269	1.156	0.000	6.000	1.149	1.883	0.000
A-ACT50	3.488	9.062	0.000	2.163	1.258	5.000	1.000	1.348	1.638	0.000
A-ACT51	2.326	8.652	0.085	0.702	0.789	1.000	2.000	1.176	1.754	34.000
A-ACT52	3.488	8.793	0.043	1.467	0.946	2.000	1.000	1.204	1.163	0.000
A-ACT53	3.488	9.062	0.000	9.098	1.340	1.000	7.000	3.229	1.579	0.000
A-ACT54	2.326	9.062	0.000	8.231	0.917	2.000	0.000	3.371	1.149	0.000
A-ACT55	2.326	9.043	0.056	2.636	0.886	0.000	2.000	1.149	1.615	0.000
A-ACT56	4.651	9.409	0.089	13.475	1.383	3.000	9.000	1.163	1.628	1.000
A-ACT57	2.326	1.190	0.055	0.000	0.455	3.000	0.000	1.190	1.149	0.000
A-ACT58	2.326	8.839	0.000	1.637	0.718	2.000	0.000	1.176	1.149	0.000
A-ACT59	0.000	0.000	0.000	0.000	0.095	0.000	0.000	1.149	1.149	0.000
A-ACT60	2.326	1.190	0.055	0.000	0.827	1.000	1.000	1.163	1.163	1.000
A-ACT61	3.488	9.062	0.000	9.098	1.165	4.000	1.000	3.234	1.578	22.000
A-ACT62	5.814	9.492	1.615	13.026	1.279	2.000	3.000	1.280	1.601	18.333
A-ACT63	4.651	9.348	1.236	9.540	1.224	1.000	4.000	1.378	1.621	2.050
A-ACT64	12.791	9.588	2.494	39.607	1.601	1.000	24.000	3.229	1.583	22.583
A-ACT65	13.953	9.630	17.430	10.267	1.480	8.000	21.000	1.249	2.458	252.500
A-ACT66	4.651	9.287	0.192	15.070	1.196	0.000	4.000	1.149	1.628	0.000
A-ACT67	3.488	8.793	0.415	1.384	0.916	0.000	3.000	1.149	1.688	0.000
A-ACT68	9.302	9.409	6.675	7.496	1.336	8.000	0.000	3.600	1.149	0.000
A-ACT69	0.000	0.000	0.000	0.000	1.451	0.000	0.000	1.149	1.149	0.000
A-ACT70	1.163	1.190	0.000	0.000	0.371	0.000	2.000	1.149	1.163	0.000

(Cont. on next page)

Table A.1. (Cont.)

Employee	Degree	Closeness	Betweenness	Eigenvector	Information	Out-degree	In-degree	Out-closeness	In-closeness	Betweenness-2
A-ACT71	3.488	9.198	0.005	14.548	1.260	2.000	3.000	3.393	1.163	0.500
A-ACT72	8.140	9.399	3.692	19.770	1.453	0.000	10.000	1.149	1.658	0.000
A-ACT73	3.488	8.749	1.018	1.099	0.826	1.000	2.000	1.163	1.176	0.000
A-ACT74	6.977	9.534	5.649	9.438	1.260	5.000	4.000	1.330	1.628	108.333
A-ACT75	2.326	8.839	0.000	1.637	0.834	2.000	1.000	1.163	1.163	0.000
A-ACT76	3.488	9.034	0.410	2.292	0.994	1.000	3.000	1.248	2.429	14.000
A-ACT77	3.488	9.101	0.000	13.928	1.082	3.000	1.000	3.277	1.568	0.000
A-ACT78	19.767	9.829	11.632	48.257	1.654	14.000	43.000	3.300	1.585	582.900
A-ACT79	1.163	1.163	0.000	0.000	0.647	2.000	2.000	1.163	1.163	0.000
A-ACT80	3.488	8.802	0.277	1.390	0.941	2.000	2.000	1.176	1.744	6.500
A-ACT81	1.163	8.678	0.000	0.894	0.581	0.000	1.000	1.149	1.163	0.000
A-ACT82	4.651	9.297	0.113	11.857	1.269	4.000	2.000	1.397	1.595	3.050
A-ACT83	2.326	8.696	0.233	2.532	0.818	2.000	0.000	1.294	1.149	0.000
A-ACT84	1.163	8.285	0.000	0.212	0.796	1.000	2.000	1.249	2.402	0.000
A-ACT85	6.977	8.949	6.259	1.776	1.300	17.000	12.000	1.250	2.433	169.833
A-ACT86	11.628	9.630	10.835	14.429	1.439	8.000	2.000	1.506	1.595	201.550
A-ACT87	1.163	1.163	0.000	0.000	0.956	2.000	2.000	1.163	1.163	0.000

Table A.2. Network Centrality Results of Dabhol Dataset.

Employee	Degree	Closeness	Betweenness	Eigenvector	Information	Out-degree	In-degree	Out-closeness	In-closeness	Betweenness-2
D-ACT1	9.000	14.793	0.092	16.285	2.587	7.000	3.000	2.308	4.072	0.000
D-ACT2	7.000	14.837	0.171	10.864	2.716	6.000	5.000	2.294	4.431	72.945
D-ACT3	13.000	15.083	0.570	19.781	2.840	12.000	1.000	2.379	1.020	0.000
D-ACT4	12.000	15.748	5.724	15.532	3.500	192.000	5.000	2.324	4.283	434.619
D-ACT5	2.000	12.315	1.919	0.061	1.374	1.000	3.000	1.000	4.513	82.000
D-ACT6	1.000	11.976	0.000	0.032	0.433	1.000	0.000	2.402	0.990	0.000
D-ACT7	16.000	16.129	2.502	28.342	3.047	18.000	3.000	2.331	4.462	24.927
D-ACT8	1.000	12.330	0.000	0.192	1.548	3.000	0.000	1.000	0.990	0.000
D-ACT9	16.000	16.181	8.860	20.164	3.329	28.000	21.000	2.301	4.744	612.032
D-ACT10	8.000	14.903	0.468	14.412	2.843	6.000	8.000	2.312	4.566	534.873
D-ACT11	11.000	15.924	6.025	14.054	2.953	6.000	12.000	2.478	1.110	214.459
D-ACT12	5.000	13.947	0.218	2.035	1.764	5.000	2.000	2.705	1.031	0.000
D-ACT13	1.000	13.405	0.000	0.514	0.773	0.000	1.000	0.990	4.726	0.000
D-ACT14	9.000	15.244	6.180	4.789	2.838	13.000	8.000	2.300	4.577	455.461
D-ACT15	9.000	15.175	1.124	14.911	2.557	9.000	3.000	2.316	4.462	2.533
D-ACT16	8.000	14.970	1.638	12.962	3.176	26.000	7.000	2.361	1.010	3.200
D-ACT17	6.000	14.104	0.183	1.085	2.157	3.000	6.000	2.262	4.527	80.751
D-ACT18	8.000	14.815	4.010	2.633	2.550	7.000	7.000	2.281	4.634	556.079
D-ACT19	0.000	0.000	0.000	0.000	0.794	0.000	0.000	0.990	0.990	0.000
D-ACT20	2.000	15.267	0.152	5.635	1.918	4.000	0.000	2.358	0.990	0.000
D-ACT21	10.000	15.456	10.034	5.450	2.573	9.000	2.000	2.414	1.020	168.667
D-ACT22	3.000	14.368	0.000	6.327	1.627	3.000	1.000	2.298	4.562	0.000
D-ACT23	9.000	15.175	0.127	17.884	3.035	18.000	0.000	2.371	0.990	0.000

(Cont. on next page)

Table A.2. (Cont.)

Employee	Degree	Closeness	Betweenness	Eigenvector	Information	Out-degree	In-degree	Out-closeness	In-closeness	Betweenness-2
D-ACT24	3.000	14.684	1.204	3.948	1.683	4.000	0.000	2.799	0.990	0.000
D-ACT25	1.000	12.579	0.000	0.118	1.241	2.000	0.000	2.295	0.990	0.000
D-ACT26	5.000	13.947	0.218	2.035	1.826	5.000	4.000	2.705	1.031	34.000
D-ACT27	18.000	15.823	3.767	28.157	3.488	65.000	5.000	3.227	1.000	0.700
D-ACT28	4.000	14.451	0.030	6.328	2.260	2.000	4.000	2.295	4.413	443.209
D-ACT29	23.000	16.313	14.441	26.366	3.353	7.000	32.000	2.310	4.764	650.663
D-ACT30	8.000	14.903	0.146	14.288	3.083	1.000	19.000	2.293	4.316	6.667
D-ACT31	7.000	15.361	2.992	6.724	3.149	11.000	32.000	2.295	4.684	188.657
D-ACT32	12.000	15.576	0.781	21.558	2.856	11.000	5.000	2.325	4.521	83.900
D-ACT33	11.000	15.748	0.914	20.124	3.071	17.000	5.000	2.325	4.454	78.483
D-ACT34	15.000	15.480	5.489	20.737	3.458	9.000	133.000	2.308	4.203	328.698
D-ACT35	8.000	15.267	2.005	7.740	2.350	5.000	5.000	2.467	1.110	89.117
D-ACT36	7.000	15.601	3.985	4.988	2.891	5.000	13.000	2.289	4.658	136.249
D-ACT37	11.000	15.291	0.257	23.142	3.057	4.000	14.000	2.307	4.498	3.667
D-ACT38	11.000	15.361	0.456	21.943	3.023	1.000	16.000	2.292	4.655	21.529
D-ACT39	23.000	16.529	9.369	38.333	3.472	39.000	24.000	2.329	4.737	670.886
D-ACT40	2.000	13.755	0.007	2.177	2.894	0.000	16.000	0.990	4.329	0.000
D-ACT41	13.000	16.077	2.329	24.165	3.356	30.000	10.000	2.320	4.699	314.686
D-ACT42	4.000	14.663	1.400	2.308	2.347	7.000	2.000	2.286	4.376	89.313
D-ACT43	14.000	15.898	3.370	24.360	3.097	4.000	18.000	2.313	4.735	748.763
D-ACT44	3.000	14.327	1.785	1.797	1.982	4.000	1.000	2.414	1.000	14.167
D-ACT45	1.000	1.000	0.000	0.000	1.079	0.000	1.000	0.990	1.000	0.000
D-ACT46	10.000	15.152	2.494	11.641	2.973	8.000	8.000	2.671	1.020	9.000
D-ACT47	2.000	14.286	0.000	4.023	1.224	2.000	0.000	2.345	0.990	0.000

(Cont. on next page)

Table A.2. (Cont.)

Employee	Degree	Closeness	Betweenness	Eigenvector	Information	Out-degree	In-degree	Out-closeness	In-closeness	Betweenness-2
D-ACT48	8.000	14.925	0.179	14.224	2.980	2.000	15.000	2.290	4.577	18.144
D-ACT49	13.000	16.000	1.083	25.344	3.109	17.000	6.000	2.328	4.619	415.865
D-ACT50	4.000	13.966	2.029	2.514	3.128	0.000	34.000	0.990	5.112	0.000
D-ACT51	14.000	15.949	2.450	26.569	3.265	3.000	27.000	2.298	4.735	181.057
D-ACT52	5.000	15.129	0.398	4.764	2.811	7.000	13.000	2.285	4.673	112.265
D-ACT53	6.000	14.859	0.116	11.395	2.660	7.000	7.000	2.305	4.575	208.275
D-ACT54	2.000	13.514	1.919	0.419	0.747	1.000	2.000	2.380	1.020	61.000
D-ACT55	6.000	14.706	0.140	11.063	2.863	6.000	9.000	2.301	4.568	5.836
D-ACT56	1.000	11.025	0.000	0.005	0.586	0.000	1.000	0.990	4.550	0.000
D-ACT57	5.000	14.368	0.057	5.273	2.064	0.000	5.000	0.990	4.787	0.000
D-ACT58	2.000	13.550	0.000	2.555	1.274	2.000	0.000	2.331	0.990	0.000
D-ACT59	3.000	14.306	0.000	4.356	1.757	3.000	2.000	2.292	4.564	0.000
D-ACT60	2.000	12.563	0.000	0.264	0.710	2.000	1.000	1.010	1.000	0.000
D-ACT61	6.000	14.993	0.341	3.822	2.666	6.000	11.000	2.283	4.664	111.448
D-ACT62	15.000	15.244	0.879	22.658	2.930	14.000	1.000	2.438	1.010	5.986
D-ACT63	1.000	12.579	0.000	0.118	3.385	24.000	0.000	2.295	0.990	0.000
D-ACT64	12.000	14.286	6.598	1.551	3.211	7.000	42.000	2.273	4.552	241.782
D-ACT65	3.000	14.493	0.759	4.222	2.375	1.000	6.000	1.000	4.787	75.000
D-ACT66	1.000	1.000	0.000	0.000	1.079	1.000	0.000	1.000	0.990	0.000
D-ACT67	10.000	15.873	1.922	17.801	2.892	4.000	12.000	2.303	4.713	193.028
D-ACT68	4.000	13.123	0.061	0.768	1.719	3.000	5.000	2.659	1.031	6.000
D-ACT69	15.000	15.480	1.337	24.727	3.330	2.000	35.000	2.292	4.640	31.822
D-ACT70	4.000	15.038	1.076	3.678	2.321	5.000	6.000	2.311	4.521	234.721
D-ACT71	4.000	14.245	3.798	3.194	1.296	0.000	4.000	0.990	5.322	0.000

(Cont. on next page)

Table A.2. (Cont.)

Employee	Degree	Closeness	Betweenness	Eigenvector	Information	Out-degree	In-degree	Out-closeness	In-closeness	Betweenness-2
D-ACT72	21.000	16.667	12.404	30.751	3.399	23.000	38.000	2.317	4.801	901.214
D-ACT73	5.000	13.947	0.218	2.035	1.764	5.000	3.000	2.705	1.031	34.000
D-ACT74	5.000	14.225	0.322	3.880	2.050	3.000	4.000	2.450	1.110	6.967
D-ACT75	5.000	14.881	0.125	10.707	2.099	1.000	4.000	2.290	4.602	1.800
D-ACT76	9.000	15.456	0.729	17.013	2.901	9.000	9.000	2.315	4.486	137.243
D-ACT77	5.000	15.038	0.000	10.266	2.576	8.000	2.000	2.303	4.545	0.000
D-ACT78	7.000	14.793	1.041	5.979	2.435	7.000	1.000	2.670	1.020	0.000
D-ACT79	1.000	13.755	0.000	2.267	2.330	0.000	1.000	0.990	1.010	0.000
D-ACT80	20.000	15.823	5.412	29.679	3.532	71.000	14.000	3.233	1.000	4.917
D-ACT81	6.000	14.577	0.497	8.042	2.498	0.000	8.000	0.990	4.344	0.000
D-ACT82	5.000	14.903	0.247	8.893	2.225	4.000	3.000	2.311	4.564	92.646
D-ACT83	0.000	0.000	0.000	0.000	0.794	0.000	0.000	0.990	0.990	0.000
D-ACT84	6.000	13.908	4.085	0.791	2.457	13.000	2.000	2.278	4.474	163.500
D-ACT85	9.000	15.244	1.412	11.378	2.832	3.000	13.000	2.306	4.688	91.608
D-ACT86	5.000	14.728	0.687	3.640	2.716	7.000	12.000	2.264	4.617	4.000
D-ACT87	34.000	16.502	17.336	47.392	3.536	19.000	80.000	2.322	4.735	1453.094
D-ACT88	3.000	14.556	0.027	3.141	1.595	3.000	0.000	2.342	0.990	0.000
D-ACT89	8.000	14.556	0.480	12.714	2.876	2.000	13.000	2.283	4.078	2.591
D-ACT90	3.000	14.306	0.000	4.356	1.487	3.000	0.000	2.346	0.990	0.000
D-ACT91	2.000	13.263	0.000	1.346	1.222	0.000	2.000	0.990	1.031	0.000
D-ACT92	2.000	12.563	0.000	0.264	0.710	2.000	1.000	1.010	1.000	0.000
D-ACT93	2.000	14.245	0.000	2.884	1.270	2.000	0.000	2.339	0.990	0.000
D-ACT94	2.000	13.793	0.000	1.721	1.846	4.000	0.000	2.463	0.990	0.000
D-ACT95	7.000	14.286	1.901	1.058	2.338	6.000	2.000	2.331	1.010	30.652

(Cont. on next page)

Table A.2. (Cont.)

Employee	Degree	Closeness	Betweenness	Eigenvector	Information	Out-degree	In-degree	Out-closeness	In-closeness	Betweenness-2
D-ACT96	2.000	13.038	0.000	0.262	1.632	0.000	4.000	0.990	4.677	0.000
D-ACT97	3.000	13.774	0.000	0.627	1.561	3.000	0.000	2.441	0.990	0.000
D-ACT98	4.000	14.881	0.079	4.564	3.332	8.000	5.000	2.300	4.331	2.309
D-ACT99	19.000	16.207	3.278	31.123	3.306	15.000	24.000	2.316	4.789	829.663
D-ACT100	7.000	14.306	0.495	1.204	2.524	3.000	10.000	2.262	4.440	18.670
D-ACT101	1.000	12.579	0.000	0.118	1.560	3.000	0.000	2.295	0.990	0.000

Table A.3. Coordination Scores of Azurix Dataset.

Employee	Resource allocation	Producer/consumer relationships	Tasks/Subtasks	Simultaneity constraints	Coordination Score	Ln-Coordination
A-ACT1	21.089	0.000	12.259	20.380	53.727	3.984
A-ACT2	0.000	6.977	0.000	0.000	6.977	1.943
A-ACT3	9.908	0.000	12.259	0.000	22.167	3.099
A-ACT4	16.378	0.000	39.807	0.000	56.185	4.029
A-ACT5	443.118	218.607	230.262	373.360	1265.347	7.143
A-ACT6	0.000	20.932	18.686	0.000	39.618	3.679
A-ACT7	0.000	0.000	0.000	6.492	6.492	1.871
A-ACT8	0.000	6.977	0.000	0.000	6.977	1.943
A-ACT9	0.000	3.700	0.000	0.000	3.700	1.308
A-ACT10	0.000	0.000	8.714	0.000	8.714	2.165
A-ACT11	0.000	0.000	6.022	0.000	6.022	1.795
A-ACT12	0.000	0.000	6.129	0.000	6.129	1.813
A-ACT13	203.829	117.531	155.573	72.667	549.600	6.309
A-ACT14	0.000	0.000	7.140	0.000	7.140	1.966
A-ACT15	10.755	0.000	0.000	0.000	10.755	2.375
A-ACT16	0.000	6.977	0.000	0.000	6.977	1.943
A-ACT17	27.297	34.886	17.297	0.000	79.481	4.376
A-ACT18	42.238	28.333	66.009	20.965	157.545	5.060
A-ACT19	0.000	3.700	0.000	0.000	3.700	1.308
A-ACT20	0.000	0.000	12.458	0.000	12.458	2.522
A-ACT21	16.925	3.700	10.715	0.000	31.340	3.445
A-ACT22	0.000	6.977	0.000	0.000	6.977	1.943
A-ACT23	0.000	6.977	12.152	0.000	19.129	2.951

(Cont. on next page)

Table A.3. (Cont.)

Employee	Resource allocation	Producer/consumer relationships	Tasks/Subtasks	Simultaneity constraints	Coordination Score	Ln-Coordination
A-ACT24	5.044	11.065	0.000	0.000	16.109	2.779
A-ACT25	14.863	3.700	0.000	17.265	35.828	3.579
A-ACT26	0.000	0.000	27.548	16.680	44.228	3.789
A-ACT27	4.585	6.977	0.000	0.000	11.562	2.448
A-ACT28	4.248	0.000	7.140	0.000	11.387	2.433
A-ACT29	117.170	27.909	48.218	5.755	199.052	5.294
A-ACT30	14.493	6.977	16.728	0.000	38.199	3.643
A-ACT31	17.066	0.000	0.000	3.700	20.766	3.033
A-ACT32	0.000	0.000	12.152	0.000	12.152	2.497
A-ACT33	44.290	11.101	32.171	18.911	106.473	4.668
A-ACT34	0.000	13.955	3.459	0.000	17.414	2.857
A-ACT35	0.000	21.355	48.607	0.000	69.962	4.248
A-ACT36	9.366	0.000	12.152	0.000	21.518	3.069
A-ACT37	4.954	0.000	0.000	0.000	4.954	1.600
A-ACT38	6.170	0.000	6.129	0.000	12.299	2.510
A-ACT39	0.000	0.000	6.129	6.492	12.621	2.535
A-ACT40	0.000	20.932	12.358	0.000	33.290	3.505
A-ACT41	0.000	3.700	0.000	10.340	14.040	2.642
A-ACT42	0.000	0.000	6.229	0.000	6.229	1.829
A-ACT43	0.000	0.000	6.229	0.000	6.229	1.829
A-ACT44	16.378	0.000	0.000	32.774	49.153	3.895
A-ACT45	5.044	10.344	0.000	0.000	15.389	2.734
A-ACT46	0.000	0.000	8.551	0.000	8.551	2.146

(Cont. on next page)

Table A.3. (Cont.)

Employee	Resource allocation	Producer/consumer relationships	Tasks/Subtasks	Simultaneity constraints	Coordination Score	Ln-Coordination
A-ACT47	0.000	6.977	12.152	0.000	19.129	2.951
A-ACT48	9.539	3.700	12.045	0.000	25.284	3.230
A-ACT49	5.459	0.000	0.000	0.000	5.459	1.697
A-ACT50	4.954	0.000	20.508	0.000	25.462	3.237
A-ACT51	0.000	6.977	6.229	0.000	13.206	2.581
A-ACT52	0.000	3.700	0.000	0.000	3.700	1.308
A-ACT53	0.000	13.955	0.000	0.000	13.955	2.636
A-ACT54	4.585	3.700	0.000	0.000	8.285	2.114
A-ACT55	11.124	13.955	39.699	0.000	64.777	4.171
A-ACT56	13.755	6.977	0.000	3.700	24.433	3.196
A-ACT57	22.429	0.000	0.000	0.000	22.429	3.110
A-ACT58	0.000	0.000	12.152	0.000	12.152	2.497
A-ACT59	18.340	0.000	66.327	0.000	84.666	4.439
A-ACT60	0.000	0.000	12.458	0.000	12.458	2.522
A-ACT61	11.276	6.644	19.498	5.170	42.588	3.752
A-ACT62	7.459	0.000	12.259	0.000	19.718	2.982
A-ACT63	20.700	30.304	24.410	0.000	75.415	4.323
A-ACT64	0.000	0.000	13.368	5.755	19.123	2.951
A-ACT65	16.552	80.072	24.403	0.000	121.027	4.796
A-ACT66	20.700	20.932	24.609	0.000	66.241	4.193
A-ACT67	0.000	0.000	7.140	0.000	7.140	1.966
A-ACT68	0.000	0.000	3.322	0.000	3.322	1.201
A-ACT69	52.151	27.909	40.348	16.680	137.088	4.921

(Cont. on next page)

Table A.3. (Cont.)

Employee	Resource allocation	Producer/consumer relationships	Tasks/Subtasks	Simultaneity constraints	Coordination Score	Ln-Coordination
A-ACT70	22.429	0.000	0.000	0.000	22.429	3.110
A-ACT71	10.755	0.000	16.493	6.492	33.740	3.519
A-ACT72	0.000	0.000	12.458	0.000	12.458	2.522
A-ACT73	0.000	13.955	0.000	0.000	13.955	2.636
A-ACT74	21.762	0.000	63.299	6.492	91.553	4.517
A-ACT75	0.000	0.000	0.000	12.247	12.247	2.505
A-ACT76	2.000	6.977	6.229	0.000	15.206	2.722
A-ACT77	4.248	0.000	7.140	0.000	11.387	2.433
A-ACT78	63.485	45.564	201.538	26.872	337.459	5.821
A-ACT79	9.908	0.000	0.000	0.000	9.908	2.293
A-ACT80	5.459	27.909	22.146	0.000	55.514	4.017
A-ACT81	0.000	0.000	12.458	0.000	12.458	2.522
A-ACT82	44.857	0.000	49.034	0.000	93.892	4.542
A-ACT83	0.000	0.000	6.229	0.000	6.229	1.829
A-ACT84	0.000	0.000	12.152	0.000	12.152	2.497
A-ACT85	27.428	28.756	140.257	14.625	211.067	5.352
A-ACT86	32.757	0.000	0.000	0.000	32.757	3.489
A-ACT87	23.055	24.722	33.854	23.020	104.651	4.651

Table A.4. Coordination Scores of Dabhol Dataset.

Employee	Resource allocation	Producer/consumer relationships	Tasks/Subtasks	Simultaneity constraints	Coordination Score	Ln-Coordination
D-ACT1	0.000	7.679	0.000	0.000	7.679	2.039
D-ACT2	13.615	0.000	60.659	14.299	88.573	4.484
D-ACT3	14.101	0.000	40.326	0.000	54.427	3.997
D-ACT4	951.058	482.823	1092.708	1091.423	3618.011	8.194
D-ACT5	20.056	0.000	20.089	0.000	40.145	3.692
D-ACT6	0.000	23.038	20.017	0.000	43.056	3.762
D-ACT7	17.523	1.585	13.877	6.248	39.233	3.670
D-ACT8	18.523	0.000	67.724	11.673	97.920	4.584
D-ACT9	4.585	17.395	64.324	0.000	86.305	4.458
D-ACT10	0.000	0.000	24.264	0.000	24.264	3.189
D-ACT11	11.392	7.679	6.248	7.150	32.469	3.480
D-ACT12	2.322	7.679	12.132	0.000	22.133	3.097
D-ACT13	0.000	15.359	0.000	0.000	15.359	2.732
D-ACT14	19.586	24.124	24.112	6.248	74.070	4.305
D-ACT15	0.000	0.000	0.000	12.496	12.496	2.525
D-ACT16	37.332	26.739	154.253	84.728	303.052	5.714
D-ACT17	0.000	0.000	6.087	7.150	13.237	2.583
D-ACT18	0.000	11.514	31.723	0.000	43.237	3.767
D-ACT19	6.807	0.000	6.248	7.066	20.121	3.002
D-ACT20	0.000	0.000	17.466	0.000	17.466	2.860
D-ACT21	6.170	15.359	39.809	7.109	68.447	4.226
D-ACT22	0.000	0.000	12.786	0.000	12.786	2.548

(Cont. on next page)

Table A.4. (Cont.)

Employee	Resource allocation	Producer/consumer relationships	Tasks/Subtasks	Simultaneity constraints	Coordination Score	Ln-Coordination
D-ACT23	0.000	7.814	0.000	0.000	7.814	2.056
D-ACT25	0.000	15.359	0.000	0.000	15.359	2.732
D-ACT26	2.322	7.679	12.132	0.000	22.133	3.097
D-ACT27	31.630	38.397	76.441	97.140	243.608	5.496
D-ACT28	5.858	0.000	21.974	0.000	27.832	3.326
D-ACT29	5.858	0.000	7.700	0.000	13.558	2.607
D-ACT30	0.000	15.359	0.000	0.000	15.359	2.732
D-ACT31	24.702	61.436	43.679	18.741	148.557	5.001
D-ACT32	0.000	0.000	0.000	12.496	12.496	2.525
D-ACT33	64.209	52.321	77.126	24.992	218.648	5.387
D-ACT34	0.000	69.115	0.000	0.000	69.115	4.236
D-ACT35	0.000	0.000	3.322	0.000	3.322	1.201
D-ACT36	31.823	54.057	64.635	8.150	158.664	5.067
D-ACT37	0.000	15.359	0.000	0.000	15.359	2.732
D-ACT38	8.340	17.530	0.000	0.000	25.870	3.253
D-ACT39	0.000	42.782	110.740	0.000	153.523	5.034
D-ACT40	7.665	0.000	6.672	0.000	14.338	2.663
D-ACT41	15.147	15.359	24.053	49.965	104.524	4.649
D-ACT42	0.000	0.000	6.087	7.150	13.237	2.583
D-ACT43	0.000	23.038	0.000	0.000	23.038	3.137
D-ACT44	16.574	0.000	0.000	0.000	16.574	2.808
D-ACT45	0.000	0.000	6.672	0.000	6.672	1.898
D-ACT46	18.802	11.101	6.044	64.227	100.175	4.607

(Cont. on next page)

Table A.4. (Cont.)

Employee	Resource allocation	Producer/consumer relationships	Tasks/Subtasks	Simultaneity constraints	Coordination Score	Ln-Coordination
D-ACT47	6.807	7.679	0.000	11.673	26.160	3.264
D-ACT48	0.000	0.000	0.000	7.150	7.150	1.967
D-ACT49	0.000	11.728	0.000	12.496	24.224	3.187
D-ACT50	4.644	0.000	0.000	0.000	4.644	1.536
D-ACT51	0.000	0.000	16.591	0.000	16.591	2.809
D-ACT55	9.615	9.716	42.557	0.000	61.888	4.125
D-ACT53	31.539	23.038	6.248	0.000	60.825	4.108
D-ACT54	0.000	7.679	33.896	0.000	41.575	3.728
D-ACT55	0.000	0.000	12.089	0.000	12.089	2.492
D-ACT56	8.340	0.000	0.000	0.000	8.340	2.121
D-ACT57	0.000	0.000	0.000	7.150	7.150	1.967
D-ACT58	0.000	7.679	0.000	0.000	7.679	2.039
D-ACT59	0.000	15.359	6.700	4.524	26.583	3.280
D-ACT60	9.285	7.679	13.373	0.000	30.338	3.412
D-ACT61	21.388	19.818	49.311	8.150	98.667	4.592
D-ACT62	4.700	7.679	13.442	7.150	32.972	3.496
D-ACT63	98.764	57.457	65.248	83.008	304.478	5.719
D-ACT64	18.416	11.586	45.067	26.795	101.866	4.624
D-ACT65	0.000	7.679	0.000	0.000	7.679	2.039
D-ACT66	0.000	0.000	13.345	0.000	13.345	2.591
D-ACT67	12.510	0.000	0.000	0.000	12.510	2.527
D-ACT68	0.000	0.000	26.159	0.000	26.159	3.264
D-ACT69	0.000	7.679	0.000	7.150	14.829	2.697

(Cont. on next page)

Table A.4. (Cont.)

Employee	Resource allocation	Producer/consumer relationships	Tasks/Subtasks	Simultaneity constraints	Coordination Score	Ln-Coordination
D-ACT70	0.000	15.359	6.087	0.000	21.446	3.066
D-ACT71	11.716	27.896	20.224	14.299	74.136	4.306
D-ACT72	30.094	82.532	129.777	7.150	249.553	5.520
D-ACT73	2.322	7.679	12.132	0.000	22.133	3.097
D-ACT74	11.392	46.077	12.892	7.150	77.511	4.350
D-ACT75	0.000	0.000	0.000	9.047	9.047	2.202
D-ACT76	31.770	35.576	50.710	28.599	146.655	4.988
D-ACT77	14.198	7.401	29.842	0.000	51.440	3.940
D-ACT78	0.000	0.000	4.459	0.000	4.459	1.495
D-ACT79	40.605	15.359	41.625	34.847	132.436	4.886
D-ACT80	30.800	24.278	79.747	55.227	190.052	5.247
D-ACT81	0.000	0.000	13.345	0.000	13.345	2.591
D-ACT82	0.000	0.000	6.700	12.496	19.196	2.955
D-ACT83	0.000	23.038	0.000	0.000	23.038	3.137
D-ACT84	4.585	7.679	6.741	6.248	25.254	3.229
D-ACT85	11.716	4.858	13.442	0.000	30.016	3.402
D-ACT86	10.028	16.238	12.717	7.066	46.049	3.830
D-ACT87	38.059	38.161	86.561	31.319	194.100	5.268
D-ACT88	4.585	3.700	13.373	0.000	21.658	3.075
D-ACT89	0.000	7.679	0.000	7.150	14.829	2.697
D-ACT90	6.807	0.000	6.248	4.524	17.579	2.867
D-ACT91	0.000	0.000	13.345	0.000	13.345	2.591
D-ACT92	9.285	7.679	0.000	0.000	16.965	2.831

(Cont. on next page)

Table A.4. (Cont.)

Employee	Resource allocation	Producer/consumer relationships	Tasks/Subtasks	Simultaneity constraints	Coordination Score	Ln-Coordination
D-ACT93	6.807	0.000	0.000	0.000	6.807	1.918
D-ACT94	16.574	0.000	8.285	7.150	32.009	3.466
D-ACT95	0.000	3.907	7.459	6.248	17.614	2.869
D-ACT96	0.000	7.679	0.000	0.000	7.679	2.039
D-ACT97	4.700	0.000	0.000	0.000	4.700	1.548
D-ACT98	79.675	46.077	58.455	23.347	207.554	5.335
D-ACT99	20.536	5.170	36.317	12.496	74.518	4.311
D-ACT100	0.000	0.000	6.087	7.150	13.237	2.583
D-ACT101	11.716	0.000	6.700	7.150	25.566	3.241

Table A.5. Coordination Key Phrases of Dabhol Dataset.

Dependency Type	Coordination Key Phrases	Frequency	Weight
Producer/consumer relationships	Are as follows	13	3.700
	Attached is a	29	4.858
	Attached please find	15	3.907
	For your information	6	2.585
	FYI	205	7.679
	The bottom line is	22	4.459
	The purpose is	3	1.585
Resource allocation	Do you want to	4	2.000
	Ensure that	112	6.807
	Help coordinate	2	1.000
	I request	3	1.585
	I would appreciate	18	4.170
	I would like to	58	5.858
	I would like your	2	1.000
	Look into	91	6.508
	Make sure that	29	4.858
	Please allow	2	1.000
	Please communicate	1	0.000
	Please do	24	4.585
	Please get	5	2.322
	Please make sure	4	2.000
	Please see	26	4.700
	Please update	5	2.322
	Please work	1	0.000
	Put this together	1	0.000
	We can go	7	2.807
	You will be	9	3.170
Simultaneity constraints	Agenda	142	7.150
	As we move closer	1	0.000
	Follow up	23	4.524
	On time	76	6.248
	On track	134	7.066
	Please allow time	2	1.000
	Sufficient time	1	0.000
	Take the time	6	2.585
Tasks/Subtasks	I am changing	1	0.000
	I believe	102	6.672
	I believe you are	1	0.000
	I have considered	2	1.000
	I recommend	7	2.807
	I suggest	9	3.170
	I wanted to	22	4.459
	I would suggest	3	1.585
	It will need	11	3.459
	Let me know if	66	6.044
	Please let me know	68	6.087
	To ensure that	76	6.248
	We can discuss	10	3.322
	We can then	2	1.000
	We have had	20	4.322
	We have seriously	1	0.000
	We need to	107	6.741
	We should	104	6.700
	Which brings me to	1	0.000
	Would probably be	16	4.000

Table A.6. Coordination Key Phrases of Azurix Dataset.

Dependency Type	Coordination Key Phrases	Frequency	Weight
Producer/consumer relationships	Are as follows	17	4.087
	Attached is a	13	3.700
	Attached please find	10	3.322
	For your information	13	3.700
	FYI	126	6.977
	The bottom line is	21	4.392
Resource allocation	Do you want to	3	1.585
	Ensure that	72	6.170
	I request	1	0.000
	I would appreciate	24	4.585
	I would ask	2	1.000
	I would like to	31	4.954
	I would like your	2	1.000
	Look into	77	6.267
	Make sure that	19	4.248
	Please coordinate	2	1.000
	Please do	33	5.044
	Please get	5	2.322
	Please make sure	5	2.322
	Please see	44	5.459
	We can go	2	1.000
	You will be	20	4.322
Simultaneity constraints	Agenda	54	5.755
	As we move closer	1	0.000
	Follow up	13	3.700
	On time	36	5.170
	On track	90	6.492
	Take the time	14	3.807
Tasks/Subtasks	I believe	75	6.229
	I recommend	6	2.585
	I suggest	11	3.459
	I wanted to	10	3.322
	I would suggest	4	2.000
	It will need	1	0.000
	Let me know if	65	6.022
	Please let me know	70	6.129
	To ensure that	41	5.358
	We can discuss	2	1.000
	We have had	5	2.322
	We need to	141	7.140
	We should	70	6.129
	Would probably be	3	1.585

APPENDIX B

FREQUENCY DISTRIBUTIONS OF NETWORK CENTRALITY MEASURES

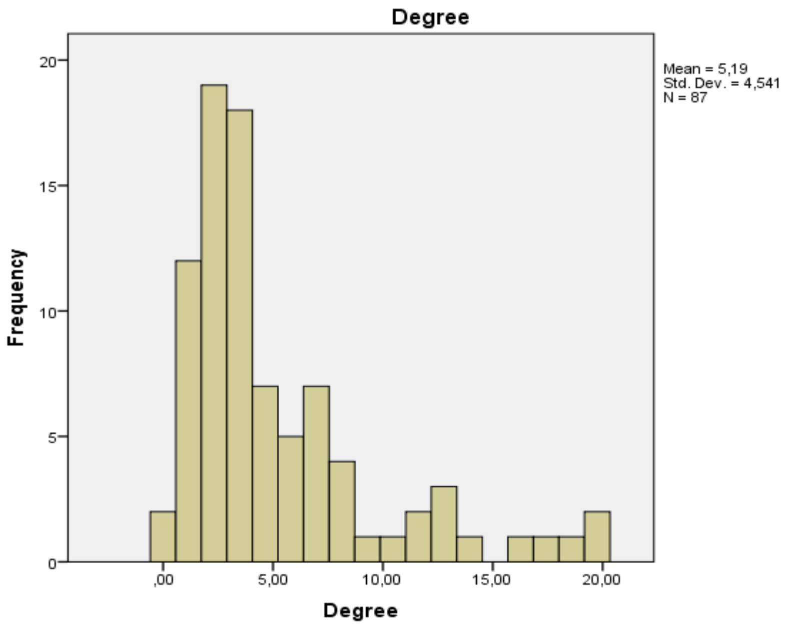


Figure A.1 Degree Centrality Distribution of Azurix Dataset.

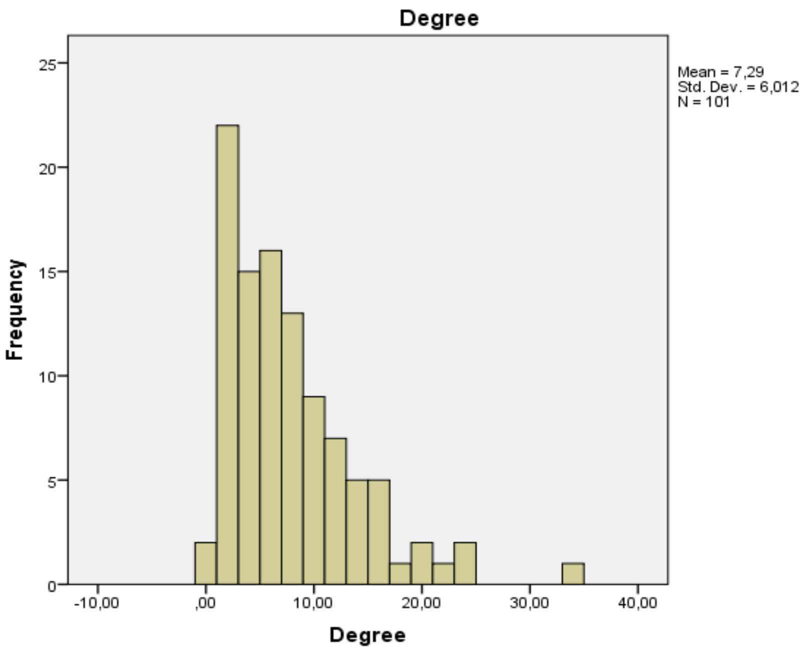


Figure A.2. Degree Centrality Distribution of Dabhol Dataset.

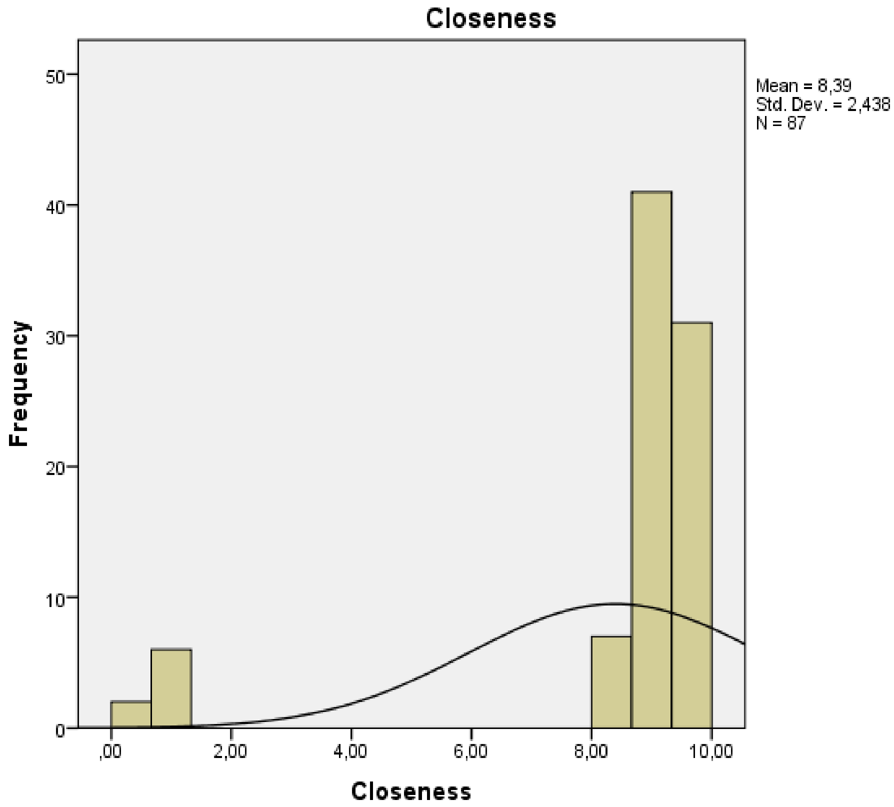


Figure A.3. Closeness Centrality Distribution of Azurix Dataset.

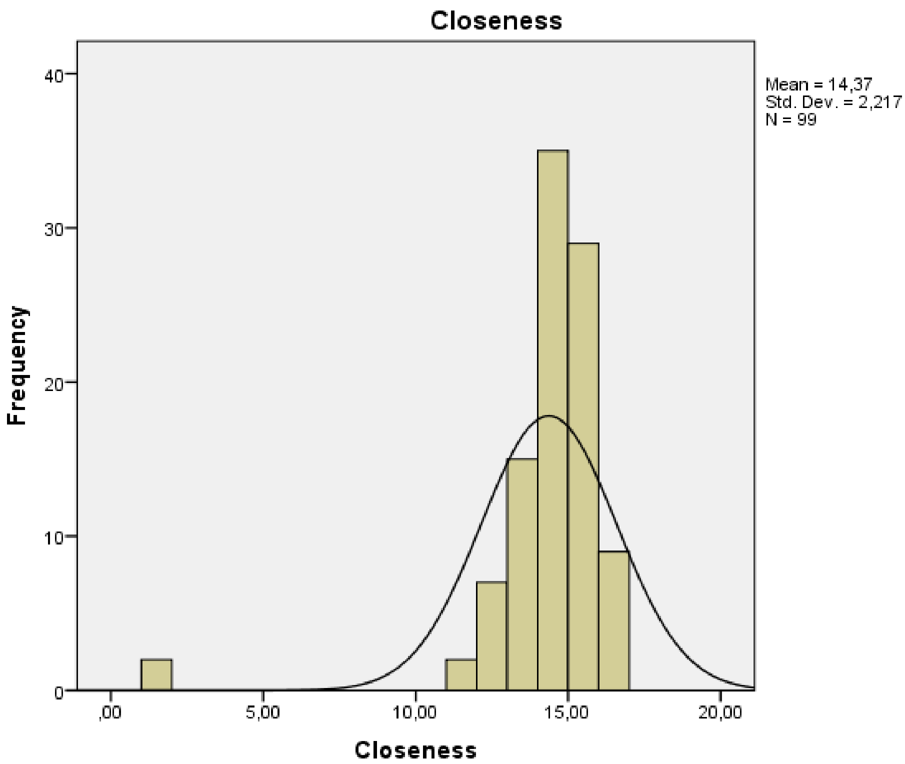


Figure A.4. Closeness Centrality Distribution of Dabhol Dataset.

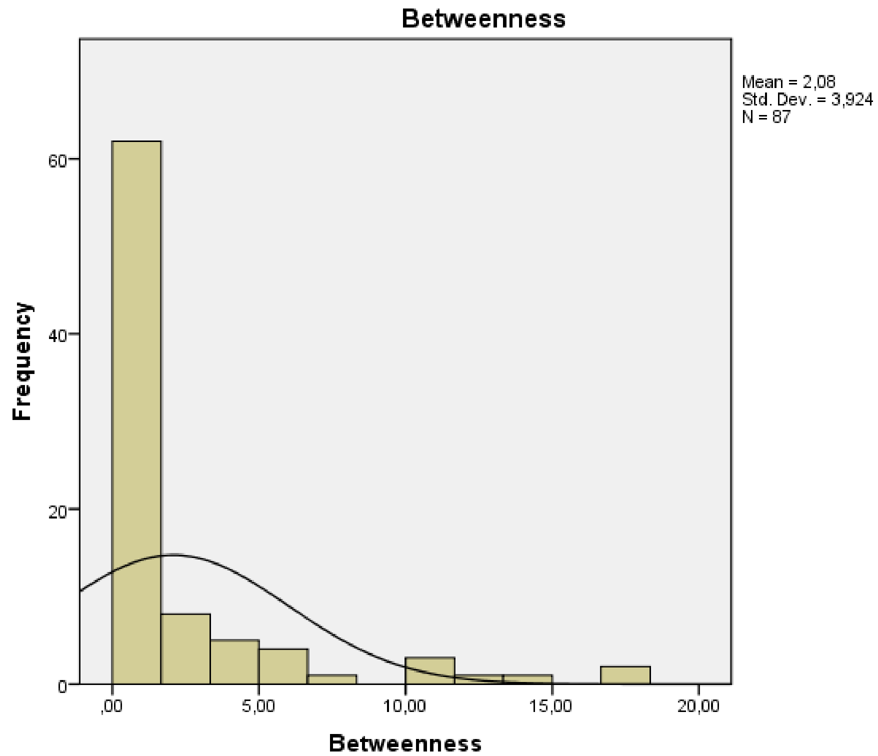


Figure A.5. Betweenness Centrality Distribution of Azurix Dataset.

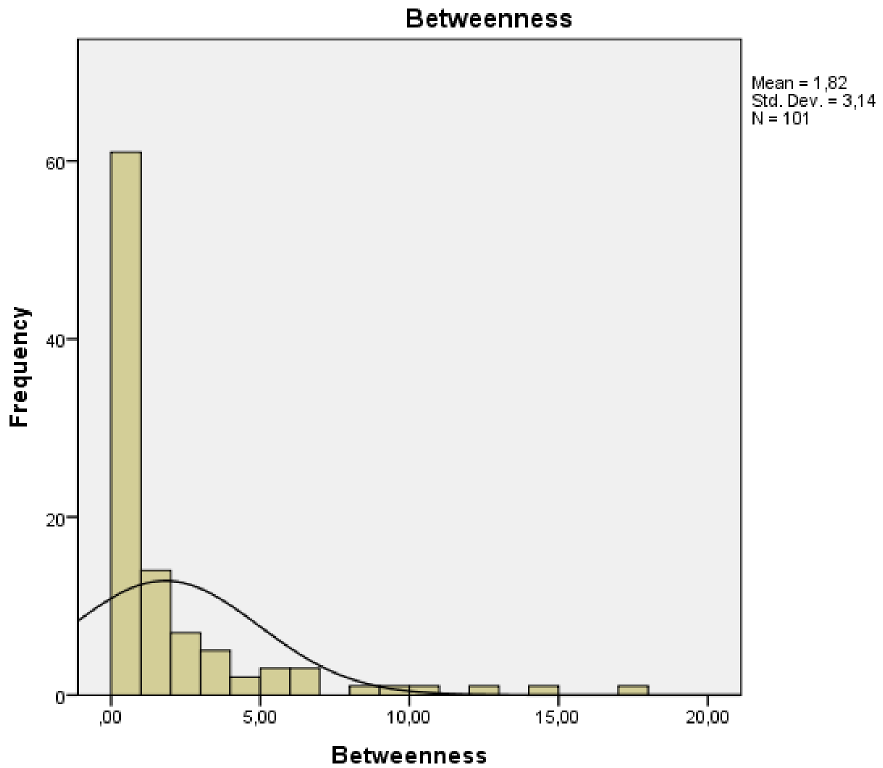


Figure A.6. Betweenness Centrality Distribution of Dabhol Dataset.

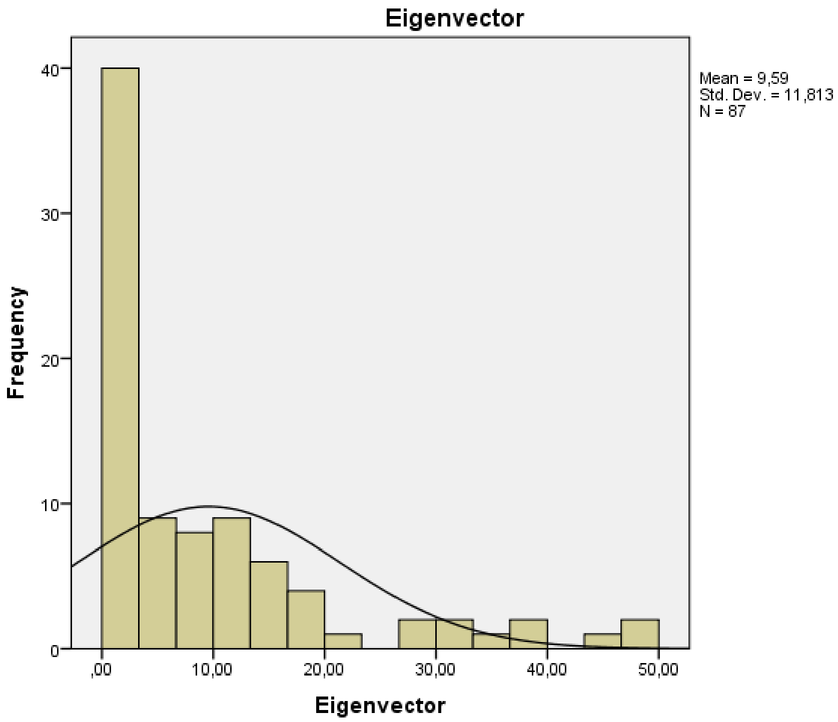


Figure A.7. Eigenvector Centrality Distribution of Azurix Dataset.

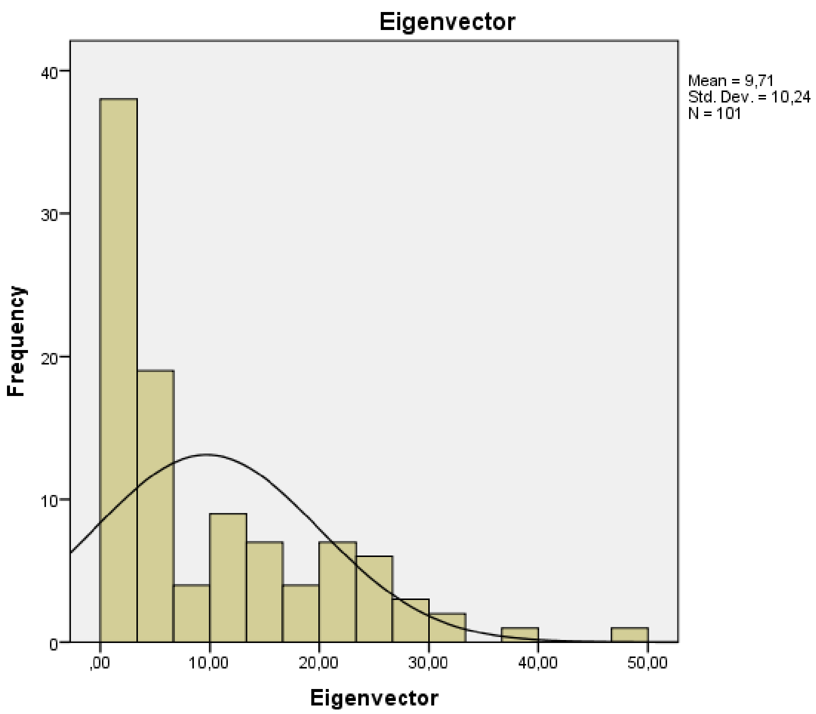


Figure A.8. Eigenvector Centrality Distribution of Dabhol Dataset.

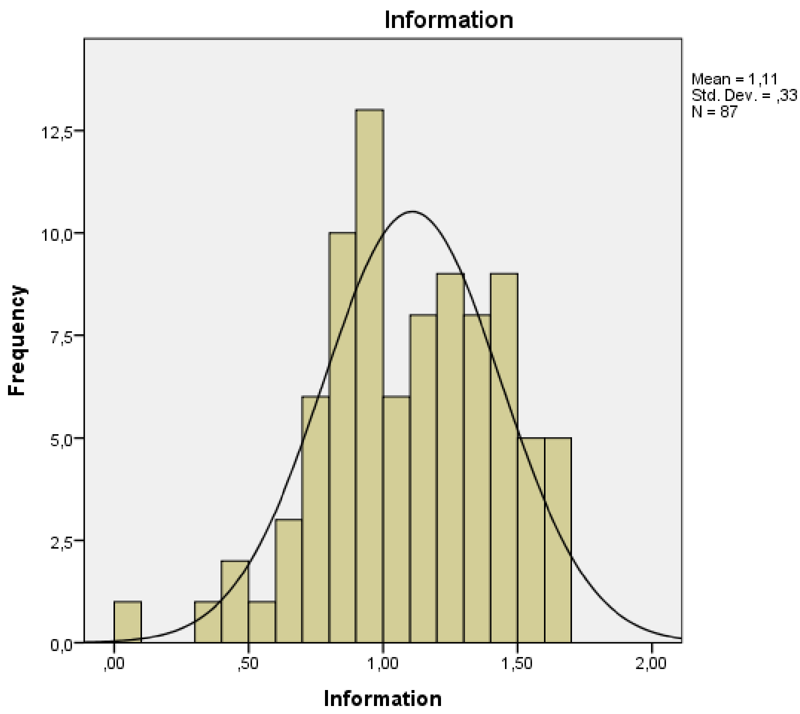


Figure A.9.Information Centrality Distribution of Azurix Dataset.

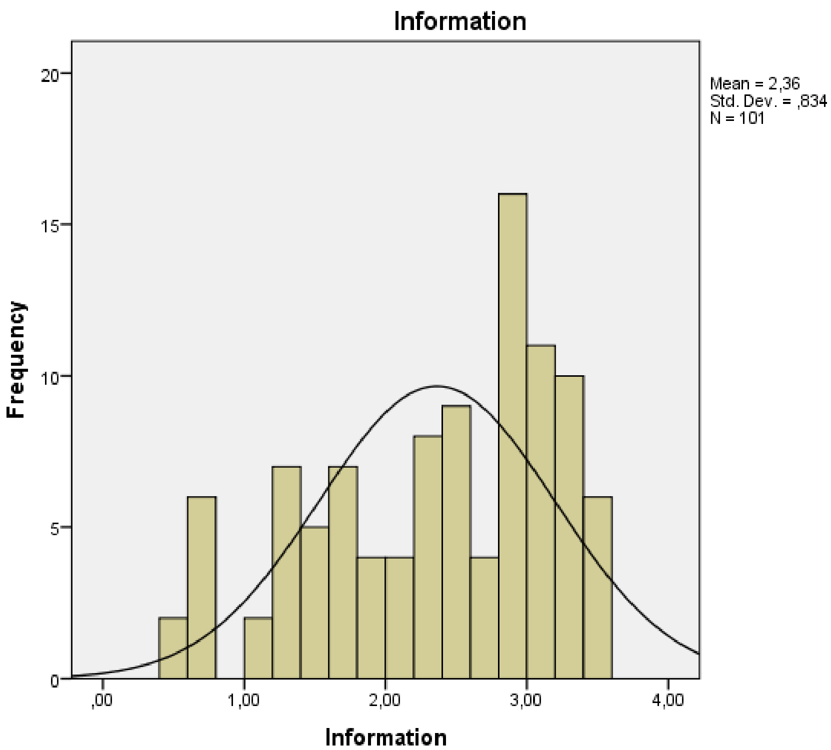


Figure A.10.Information Centrality Distribution of Dabhol Dataset.

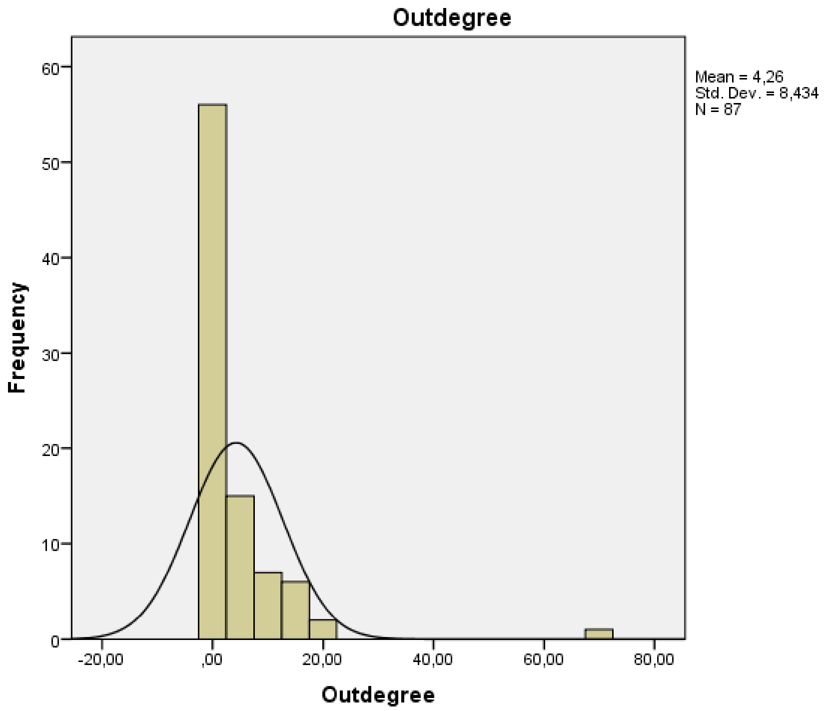


Figure A.11. Out-degree Centrality Distribution of Azurix Dataset.

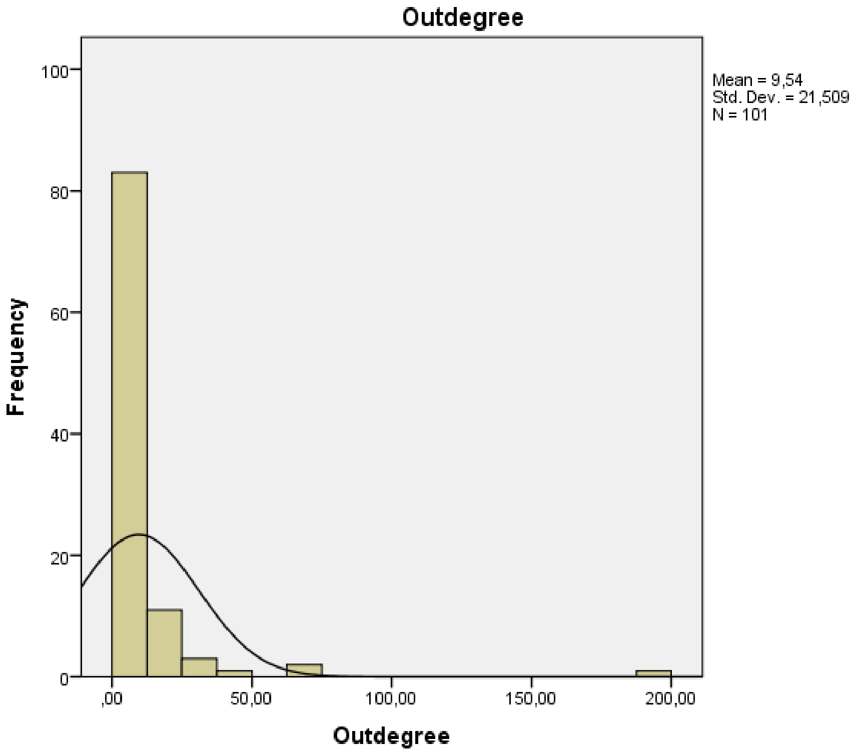


Figure A.12. Out-degree Centrality Distribution of Dabhol Dataset.

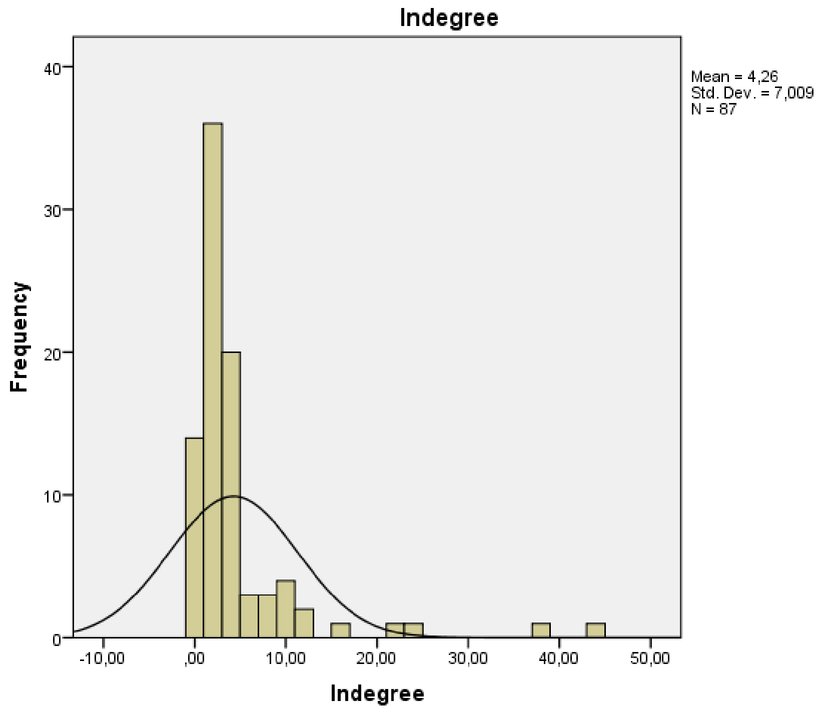


Figure A.13. In-degree Centrality Distribution of Azurix Dataset.

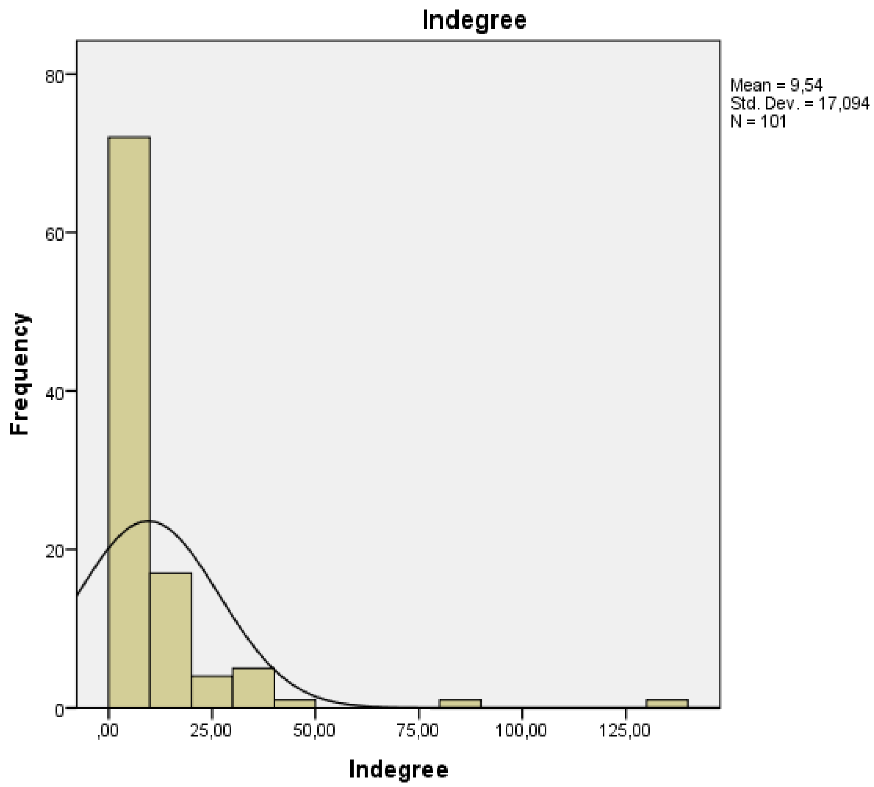


Figure A.14. In-degree Centrality Distribution of Dabhol Dataset.

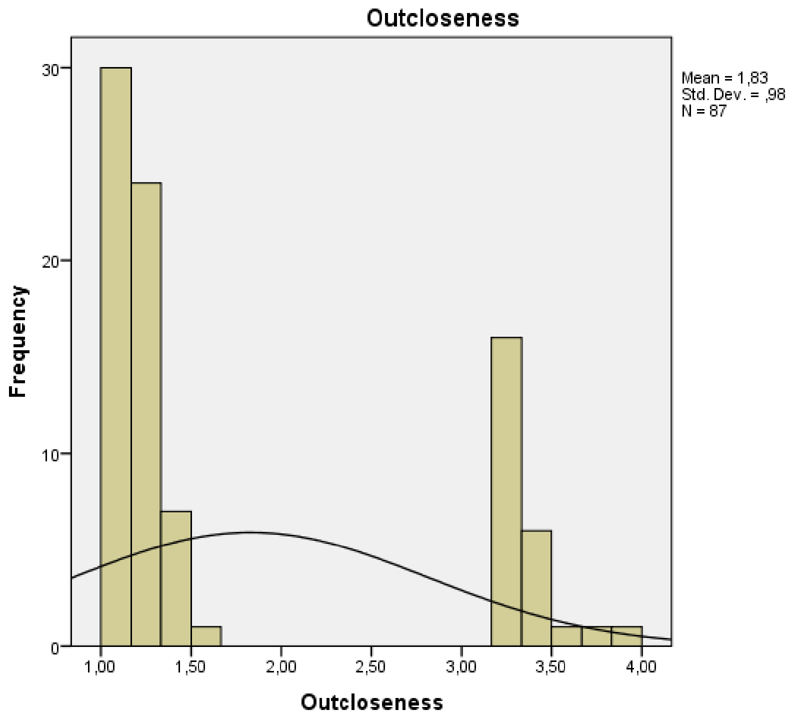


Figure A.15. Out-closeness Centrality Distribution of Azurix Dataset.

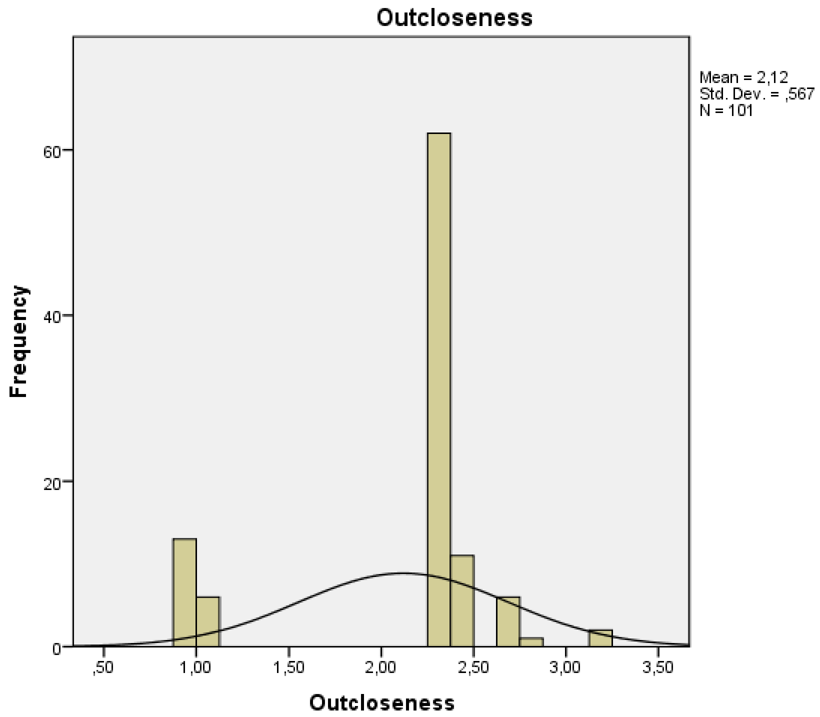


Figure A.16. Out-closeness Centrality Distribution of Dabhol Dataset.

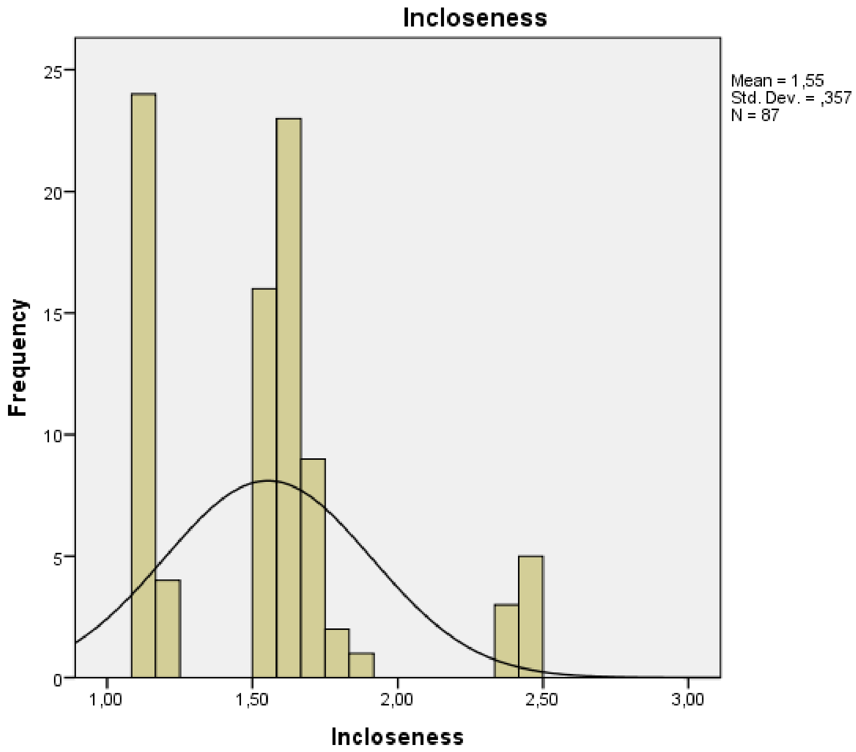


Figure A.17. In-Closeness Centrality Distribution of Azurix Dataset.

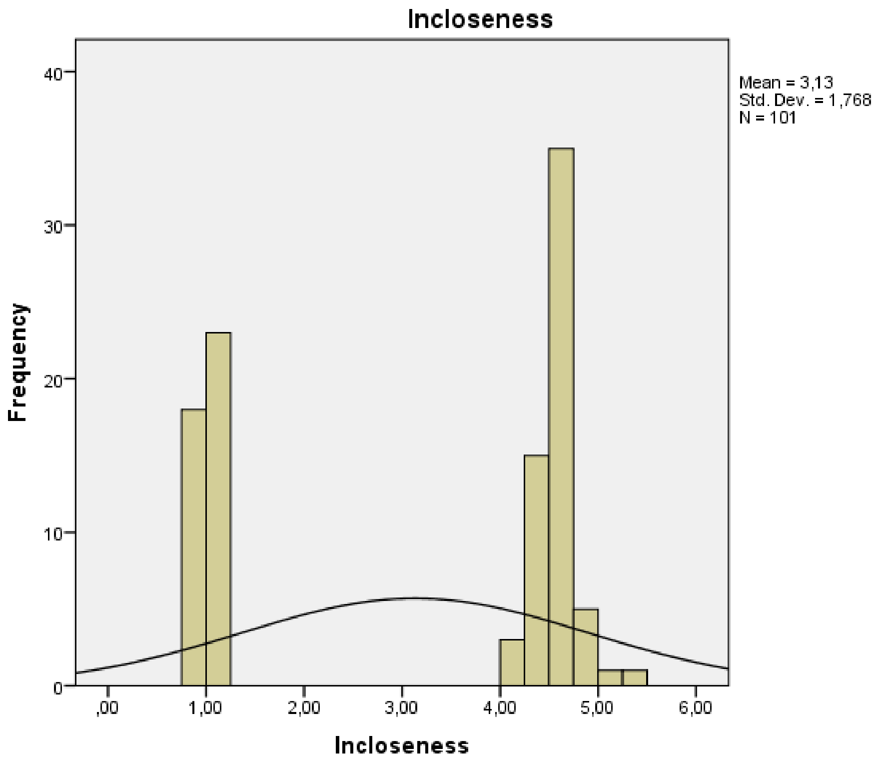


Figure A.18. In-Closeness Centrality Distribution of Dabhol Dataset.

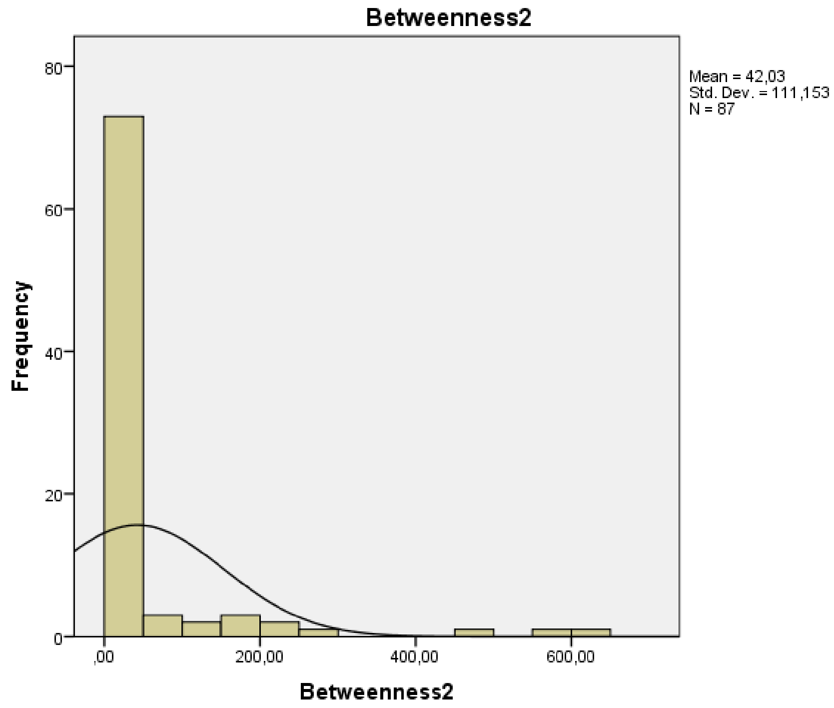


Figure A.19. Betweenness Centrality (directed) Distribution of Azurix Dataset.

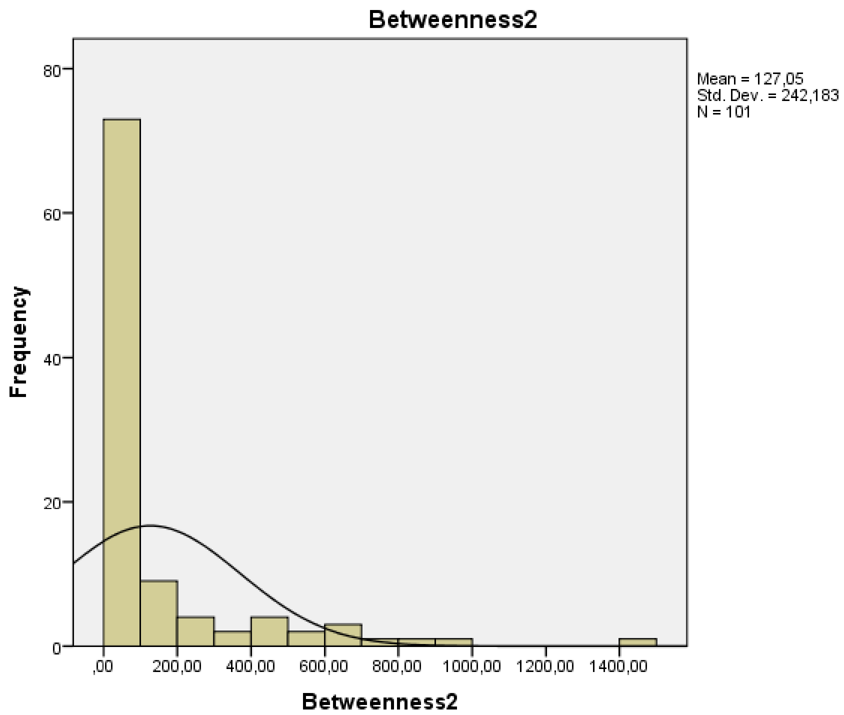


Figure A.20. Betweenness Centrality (directed) Distribution of Dabhol Dataset.

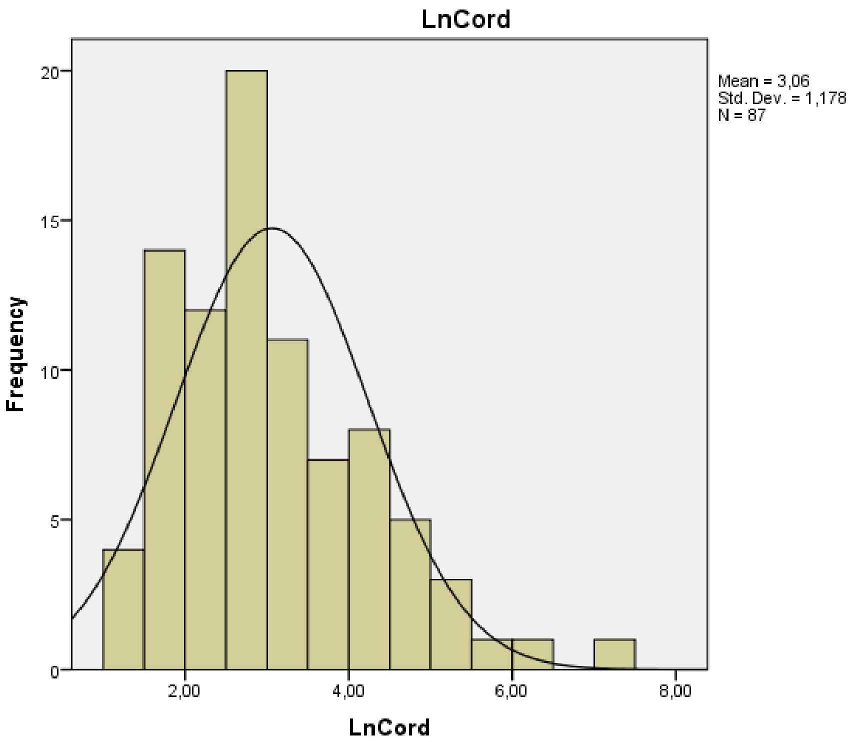


Figure A.21. Coordination Score Distribution of Azurix Dataset.

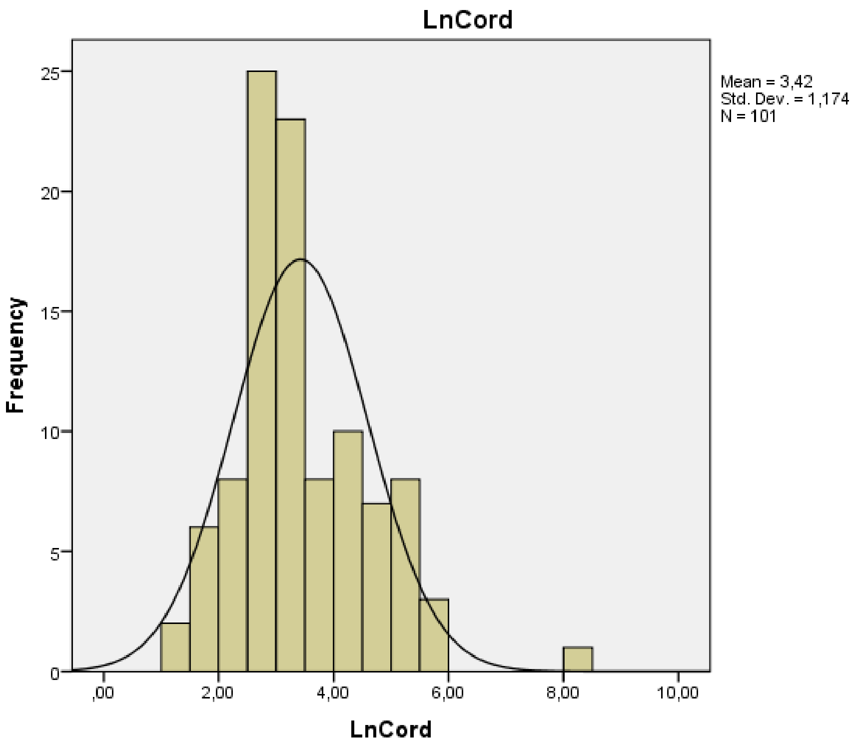


Figure A.22. Coordination Score Distribution of Dabhol Dataset.