Proceedings of DETC '06
ASME 2006 International Design Engineering Technical Conferences and
18th International Conference on Design Theory and Methodology (DTM)
Philadelphia, Pennsylvania USA, September 10-13, 2006

DETC2006-99443

ON THE USE OF NETWORK ANALYSIS IN PRODUCT DEVELOPMENT TEAMS

Joe A. Bradley[+] & Ali A. Yassine
Product Development Research Laboratory
Department of Industrial and Enterprise Systems Engineering
University of Illinois at Urbana-Champaign
Urbana, IL 61801

ABSTRACT

In this paper, we consider the product development process as a network of interacting elements (e.g. development participants, physical subsystems/components, or development activities) exchanging information in order to achieve the common goal of developing a successful product or service. Drawing from network analysis (NA) techniques, we consider three network measures: single-node centrality, group-centrality, and the key-player problem. Using these measures, we determine a small subset of nodes within the network that is most important to the flow of information. That is, these nodes significantly control (i.e. receive, distribute, or process) more information than any other node in the network. Identification of this subset of nodes is essential to devise improved management strategies for information flow within the product development process.

We find that when using NA techniques to analyze product development processes it is important to consider all three analysis measures because the different measures produce a different subset of top scoring nodes. We discuss some of the underlying reasons for these differences and conclude that the convenient measure(s) to use should be based on the particular development environment and the underlying managerial objectives. We demonstrate these measures and results by studying the development process of a large commercial aircraft engine.

(*Keywords: Product Development, Network Analysis, Graph Theory, Fragmentation, Information Flow, Product Architecture).*

1. INTRODUCTION

A survey of U.S. companies indicates that 35% of all annual revenue is the result of new products (Quey, 2004). These are products commercialized by the organization within the last three years. Therefore, it is important for firms to understand and streamline their product development (PD) process so that new high quality products can be introduced on timely and economic basis. Many researchers agree that the PD process requires continual communication amongst the entities involved (Sosa et al., 2003). An understanding of information flow opens new avenues for improving the project management, organizational design, and product architecture of the development process (Yassine and Braha, 2003).

One method for analyzing the flow of information is the Design Structure Matrix (DSM) (Steward, 1981; Browning, 2001; Sharman and Yassine, 2004). The DSM tool is useful for visualizing the relationship amongst a set of entities and analyzing several different situations depending on the type and nature of the information being exchanged. However, the DSM does not provide the capability or sophistication for uncovering and understanding the statistical properties of a complex development network.

We can enhance our understanding of PD networks by borrowing concepts from network theory. Network theory focuses on the relationships amongst entities and on the patterns and implications of these relationships. In this paper, we use the network analysis (NA) techniques embodied within this theoretical frame to study PD team networks.

[+] Corresponding Author: jabradly@uiuc.edu.

The motivation of this research is to consider the structural and statistical properties of the PD team network and to understand how these properties can impact information flow and consequently impact product development decisions (e.g. product architecture, team formation, and project management) over the project duration. We employ a series of measures (individual-centrality, group-centrality, and key-player identification) developed within network theory to identify information hubs. We define these hubs as a set of nodes ranking high on a NA measure. We identify the high scoring nodes in the network and explore how controlling these nodes can affect the overall performance of the network. We subsequently used the results of the analysis to suggests how managerial decisions can impact design changes, speed of development, and candidates for outsourcing. An analysis of this PD team network provides insight regarding the control of information in the network.

In the next section, we present a brief survey of previous studies in the PD area that are most relevant to our work. In Section 3, we introduce the network analysis framework and the network analysis concepts of individual-centrality, group-centrality, and the key-player problem used as well as provide an explanation of how these measures are applied to the PD network. Section 4 introduces the PD case study that we analyze and a discussion of results highlighting differences and similarities observed based on the different measures. In Section 5, we provide managerial insights and implications. Section 6 recaps the main conclusions of this paper.

2. LITERATURE REVIEW

Researchers have begun to explore the use of NA techniques to better understand PD processes. Battallas and Yassine (2006) consider a scale-free structure of PD networks and the implications on management. A distinguishing feature of the scale free network is that there exists a distribution with a high percentage of the nodes with a few links and a much lower number of nodes with a large amount of links toward them. This gives rise to the presence of highly connected nodes or hubs within the network structure. They consider each member of the development team is represented by a node and an edge between nodes represents the exchange of information. Based on node centrality and brokerage measures, they introduce the concept of an Information Leaders Team (ILT). The ILT nodes were based on their overall ranking on the centrality and brokerage scores. This group is believed to be key information handlers in the network.

Sosa et al. (2005) employ NA techniques to study component modularity in complex product design. They introduce three measures: (1) Degree modularity, (2) Distance modularity, and (3) Bridge modularity. These concepts are highly parallel to concepts discussed in general NA terminology. This framework provides insight into how component interactions can impact design decisions. Sosa et al. (2004) employs statistical network analysis to study the discontinuities in product architecture and the integrated organizational structure.

Braha and Bar-Yam (2004) also utilize NA techniques to analyze the properties of product development teams. They assert two significant features of the PD network. First a PD network exhibits the small-world property. Secondly, PD task networks are dominated by a few highly central entities. These findings reiterate the concept of an Information Leaders Team introduced by Battallas and Yassine (2006).

Whitney (2003, 2004) considers the connectivity of mechanical assemblies. He shows that the properties of a mechanical assembly's network are not necessarily scale-free. He remarks that there are a number of fundamental constraints that impacts the assembly/component network. As a result of these constraints, a naturally occurring network (i.e. biological networks) will differ from a mechanical assembly/component network. If using a network analysis approach, it is necessary to consider additional factors due to engineering constraints.

Additional related literature is in the area of collaborative design. Buehler et al. (2005) studied the optimistic biasness in group collaboration task. They find that groups have an enhanced optimistic outlook of the success of a project versus that of the individual. Seitamaa-Hakkarainen et al. (2000) looks at collaborative design in a networked learning environment. Furthermore, many cognitive researchers have intensively studied design processes over the past few decades. Whereas many of these researchers have considered the impact of social behaviors on the design process, we consider the social influence only in terms of whether or not two teams interact (exchange information). We do not consider the many additional human factors such as gender, race, age, or others in this study.

The literature discussed in the proceeding paragraphs provides a brief assessment of the current research in using NA techniques to explore PD networks. The focus of the aforementioned literature is on understanding how the structure of the network impact specific network properties for a given set of nodes, we extend this research by comparing the results of different NA measures to identify highly-connected nodes within the product development network.

3. NETWORK ANALSYSIS TECHNIQUES

The network analysis that we are using in this paper borrows many of its mathematical concepts from graph theory and social network analysis (Harary, 1969; Wasserman et al., 1994). A benefit of network analysis is that it considers not only the influence of the individual elements but also the relationship amongst them. In the case of product development, we consider two networks, a team network (individual design teams and groups) and a physical network (components and subsystems). We can construct a graph defined by nodes (actors or development participants) and edges (relationships between actors) to represent a PD network, see Figure 1. If the graph of Figure 1 represents a team network (i.e. organizational architecture), then each node

A through E are the development participants and an edge between them represents the exchange of information. On the other hand, if the graph represents a physical network (i.e. product architecture), then nodes A through E represent product components or subsystems and the edges represents a relationship (dependence) between them (e.g., subsystem A interfaces with subsystem B).[1] In this paper, we primarily focus on the team network, but consider a joint analysis of both networks in section 4.4.

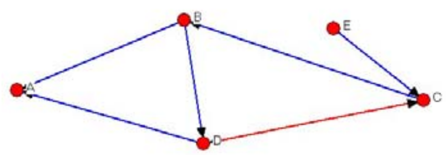

Figure 1: A five node graph. The nodes represent participants in the networks. The connecting edges represent the exchange of information or dependences between a pair of participants.

When a few nodes have the majority of control (i.e. receiving, distributing, or processing information) over the entire network, this subset of nodes becomes critical to the network's behavior and performance because of their strategic location and relevant relations to other nodes. Furthermore, peripheral nodes in the network rely on these nodes to reduce the length of communication paths within the network.

There are several indexes used to identify the most significant nodes and quantify there level of importance based on numerical measures. In this investigation, we consider the following measures: Centrality (individual and group) and the Key-Player Problem (KPP-1 and KPP-2). These concepts and formulations will be described in the following subsections.

3.1 Individual Node Centrality

We consider three different individual centrality approximations: (1) Degree-Centrality, (2) Closeness-Centrality, and (3) Betweenness-Centrality.

Degree-Centrality measures the number of out-links connecting a node to its neighbors and the number of in-links that a node is receiving from adjacent nodes. In some cases there is a distinction between out-degree and in-degree centrality. In the case of a PD team network, there is no significant distinction between in and out degree centrality. For example given node *i* and node *j* connected in a network, we would suspect that information flows indiscriminately between the two nodes (Morelli et al., 1995). Either *i* or *j* can initiate the present state of communication flow. Because this measure has the tendency to scale in proportion to the size of the network (i.e., increase in the number of nodes), it becomes complicated to use for comparing different networks. However, the degree centrality can be standardized by dividing by the maximum number of actors that a node can be connected to (Wasserman et al., 1999). If we assert that the number of nodes in a network is *n* then a node can be connected to at most all other nodes in the network but not itself. Therefore there exists *n-1* connections. The standardize centrality formulation is as follows:

$$C'_D(n_i) = d(n_i) = \frac{\sum_{\forall j \neq i} x_{ij}}{n-1} \qquad (1)$$

where, $C'_D(n_i)$ = degree centrality of node *i*,

$d(n_i)$ = degree of node *i*, and

x_{ij} = 1 if *i* is incident to *j* and 0 otherwise.

A limitation to the degree-centrality measure is that it only considers nodes that are directly joined to the node of interest. Although this is a significant measure, it does not consider indirect ties by which a node can reach others using paths available in the network. Therefore, we introduce an additional measure, closeness-centrality. A node ranking high on closeness-centrality can reach other nodes through short distances. In network literature, a shortest path between two nodes is defined as a geodesic[2] (Wasserman et al., 1999). We adopt this definition for our study. To determine the geodesic, we must first determine all the distances form a single node, call it *i* and all the other nodes in the network call them *j*. The shortest path to each of the nodes will be the geodesic. The geodesic is determined for all the nodes in the network. We define the geodesic between node *i* and *j* as follows:

$$d = (n_i n_j) \qquad (2)$$

Using the geodesic, we can formulate the standardized closeness-centrality measure for *n* nodes. The formulation is as follows:

$$C_c(n_i) = \frac{n-1}{\sum_{j=1}^{n} d(n_i n_j)} \qquad (3)$$

where, $C_c(n_i)$ = closeness centrality of node *i*, and

$d(n_i n_j)$ = geodesic between *i* and *j*.

The closeness-centrality measure requires that there exist a path from every node to every other node, meaning that the network is connected. This could be a drawback when studying some network structures. A high closeness-central node can play a significant role in the speed of information transmission within the network. This is very important for effective communication throughout the design teams.

A third way of measuring centrality is by considering nodes that lie along the path between other nodes. This is defined as betweenness-centrality. The importance of these nodes is for control over the integrity and reliability of the information in

[1] An additional PD network can be constructed from considering the development activities. This will lead to a project network similar to traditional project networks such as CPM and PERT networks.

[2] We define distance as the number of links between a pair of nodes. Each link is given a weight of one.

the network. These nodes regulate and greatly influence the transmission of information within the network (Wasserman et al., 1999). Between-central nodes are storages of power and knowledge in a network.

In order to calculate an index that measures betweenness-centrality three assumptions are required. First, it is assumed that if an actor (node) wants to reach another actor, it will prefer the shortest path available. This assumption makes sense in the real-world because any actor within a network will attempt to minimize the number of additional actors it must interact with prior to reaching the desired destination. A second assumption is that, if two or more geodesics are available, the actor chooses between them with equal probability. The third assumption is that if actor *i* communicates with actor *j* then a symmetric relationship holds, were *j* must communicate with *i*. A formulation of the betweenness-centrality index will require information about the geodesic of nodes containing the node of interest. The mathematical formulation for betweenness-centrality is as follows:

$$C_B(n_i) = \frac{\sum_{j<k, i\neq j, j\neq k} \frac{g_{jk}(n_i)}{g_{jk}}}{\frac{((n-2)(n-1))}{2}} \qquad (4)$$

where, $C_B(n_i)$ = betweenness centrality of node *i*,

$g_{jk}(n_i)$ = number of geodesics linking *j* and *k* that contains *i* in between, and

g_{jk} = total number of geodesics linking *j* and *k*.

The aforementioned measures consider the individual properties of nodes in the network. However, there may exist a group of nodes that demonstrate a high centrality within the network. In the following section, we will explore this behavior.

3.2 Group Centrality

The measures discussed in section 3.1 consider the properties of a single/individual node. We now discuss a measure that considers a group of nodes that are most central to the network. Everett and Borgatti (1999) discuss a concept of group centrality that is derived from the individual measures of centrality. That is, the centrality of a group is computed directly from the relationships among its individual nodes. Everett and Borgatti (1999) state that using this approach eliminates the problem with overlapping groups - instances where one individual can belong to many groups.

One approach to determining group centrality is to average or sum the individual centrality scores. This approach has limitations. For example if two nodes *i* and *j* have the same individual centrality score and they are independently added to a node *k* forming two groups (*i*,*k*) and (*j*,*k*), the total centrality would be the same for both groups. Consider the condition where the set of nodes that are highly connected to *k* are contained within and equal to the set of nodes highly connected to *i*; however, the nodes connected to node *j* are highly different. Since node *j* brings in a new set of node connections to group (*j*,*k*) the centrality score for group (*j*,*k*) should be greater than that for group (*i*,*k*). Consequently, the additive approach does not take into account like-neighbor connections.

Everett and Borgatti (1999) define group degree centrality as the characterization of the maximum number of non-group nodes that are connected to a set of group members. Multiple ties to the same node are counted only once. The formulation for group centrality is as follows:

$$CG = \frac{\sum_{\forall j\neq i; i, j\in P} x_{ij} + \sum_{\forall i\neq j; j, i\in P} x_{ji} - \sum_{j\neq i}\sum_{m=n} y x_{im} x_{jn}}{2(n-1)\sum_{m,n; j\notin P} x_{mn}} \qquad (5)$$

where, CG =group degree-centrality,

P = a set containing the group of selected nodes.

$\sum_{\forall j\neq i; i, j\in P} x_{ij} + \sum_{\forall i\neq j; j, i\in P} x_{ji}$ = the summation of nodes adjacent to the group identified in set.

$\sum_{j\neq i}\sum_{m=n} y x_{im} x_{jn}$ = the term for non-group nodes connected to more than one group node. y = 1 if there is node overlapping and 0 otherwise.

$\sum_{m,n; j\notin P} x_{mn}$ = the term for all non-group nodes connected to group member.

3.3 Key Players Metric

The final network measure that we consider in this paper is the Key Player Problem (KPP) (Borgatti, 2001). Borgatti (2001) proposes two types of KPP conditions:

- KPP 1 - Given a network, find a set of k nodes (called a kp-set of order k) which if removed, would maximally disrupt communication among the remaining nodes (fragmentation).
- KPP 2 - Given a network, find a kp-set that is maximally connected to all other nodes (reach).

The formulations for KPP-1 and KPP-2 used in this paper are described as follows (Borgatti, 2001):

KPP-1:

$$F = 1 - \frac{2\sum_{i>j} \frac{1}{d_{ij}}}{n(n-1)} \qquad (6)$$

where, F = percent fragmentation,

n = number of nodes in the network, and

d_{ij} = distance between pairs of nodes *i* and *j*.

KPP-2:

$$R = \frac{\sum_j \frac{1}{d_{sj}}}{n} \quad (7)$$

where, R = percent reach,
n = number of nodes in the network, and
d_{sj} = distance from a key player set S to node j.

The KPP-1 method identifies a group of nodes that are most important in holding an otherwise fragmented network together. The KPP-2 method identifies a group of nodes that have maximum overall connectivity, directly or indirectly, in the network.

4.0 CASE STUDY-JET ENGINE PD TEAMS

The PD case study examined in this paper is based on previous work by Sosa et al. (2000, 2003). They presented a study for a large commercial aircraft engine developed at Pratt & Whitney Aircraft Company. This is a complex product composed of 6 modular chunks (groups A to F): A. Fan, B. Low Pressure Compressor (LPC), C. High Pressure Compressor (HPC), D. Combustion Chamber (CC), E. High Pressure Turbine (HPT), F. Low Pressure Turbine (LPT), and 2 integral subsystems (groups G & H): G. Mechanical Components and H. External and Controls, see Figure 2. The chunks and subsystems are divided into a total of 54 major components. Each of these 54 components was assigned to a design team. The overall PD network consists of 54 cross-functional teams working simultaneously on the design of each component. The analysis in this paper is based on the interactions of the 54 teams working on component designs. These team interactions and component interfaces were identified using a Design Structure Matrix (DSM) construction. To develop the product design interface DSM, Sosa et al. (2000) interviewed design experts who have a deep understanding of the architecture of the product. They identified how the product was decomposed into systems and how these systems are subsequently decomposed into components. Based on the design dependencies the final DSM was constructed. To develop the design team interaction DSM, Sosa et al. (2000) identified the teams responsible for developing each of the components. They surveyed the key members of each team to understand the interaction (communication) amongst the various teams. Based on the team interaction survey, the final DSM was constructed. We have used these DSM models to determine the structure of the PD team network and to also identify component interfaces (dependencies).

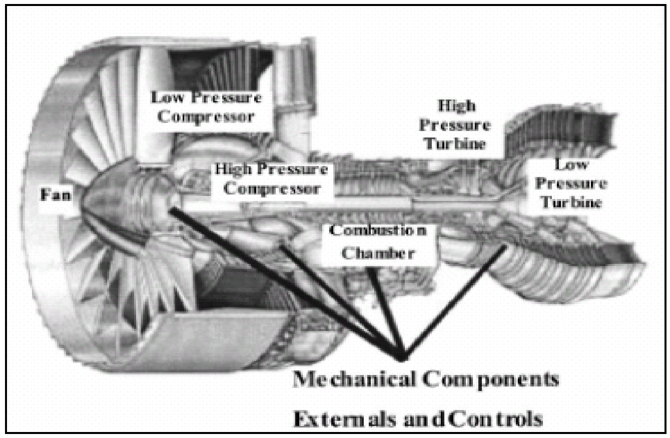

Figure 2: Commercial Aircraft Engine-identifying the key subsystems and modular chunks (Sosa et al., 2000 & 2003)

4.1 Analysis of Case Study

Computations were performed using Ucinet Software (Borgatti et al., 2002). This software implements the mathematical formulations for the network measures described in section 3. Table 1 shows the highest ranking 15 out of 54 nodes based only on the three different definitions of individual node centrality. Each measure is expressed as a percentage of the maximum possible value for that measure based on the PD network. The higher the score for a measure signals a high connectedness and consequently a critical role in informational control. We define informational control as receiving, distributing, and processing the information within the network. Successful product development requires frequent information exchanges; therefore, high-scoring nodes, based on these measures, will be very important to proper management of the PD process.

node	Degree	Node	closeness	node	Betweenness
G1	58.49	G1	70.67	G1	22.48
H3	45.28	H3	64.63	D2	9.04
H10	45.28	H10	64.63	H10	7.39
D2	43.40	D2	63.10	D1	7.05
E5	35.85	G2	58.24	H3	6.36
G2	35.85	B7	57.61	E5	6.07
D1	33.96	H1	57.61	A1	5.72
B7	32.08	D1	56.99	G2	3.24
H8	32.08	H8	56.99	B7	3.20
H1	30.19	E5	55.79	H8	3.05
A1	28.30	E4	54.08	C6	2.76
H4	28.30	H4	54.08	A3	2.75
H7	28.30	H7	54.08	C4	2.63
B6	22.64	F1	53.54	E4	2.04
G7	22.64	A1	53.00	F1	2.00

Table 1: Individual Node Centrality Measures-Identification of the 10 highest scoring nodes in rank-order of measurement

In this study, we arbitrarily identify the top 10 scoring nodes of the jet engine PD team network as shown in Table 1. Based only on the individual node centrality, we have 90% overlap (highlighted nodes) amongst the top 10 in each measure, see Table 2. We define overlap as the presence of that node in the top 10 scoring set for each individual centrality measure (e.g., degree[3], closeness, and betweenness). Our results suggest a high correlation amongst the measures, which implies that if a node scored high in degree centrality it will likely score high in both closeness and betweenness centrality. If we consider these centrality measures only, it appears that we have successfully identified the top 10 scoring nodes in the PD network. However, the high correlation observed in this network cannot be generalized to any networks; so additional measures for identifying the top scoring nodes must be considered.[4]

High Centrality Overlapping Nodes
G1
H3
H10
D2
E5
G2
D1
B7
H8

Table 2: Identification of overlapping nodes found in the top 10 for each individual node centrality measure. H1 & A1 only appears in two of the three centrality categories for the top 10 scoring nodes. They are not included in this overlapping list.

Another approach for identifying the top scoring nodes is to consider the group-centrality measure. In contrast to the individual centrality measures, this method considers a group as a starting set. For this paper, we only consider group degree-centrality. Table 3 shows the node set based on the group degree-centrality measure. This set reaches 100% of the other nodes in the network. Furthermore, we observe that it is only necessary to have the top seven to connect 100% of the nodes in the network. We can speculate that there is redundancy in choosing a top 10 based on this particular measure. The addition of nodes F2, D4, or H9 do not improve the result as shown in Table 3. We can speculate that it would be a waste of resources if we included these three additional teams in a group that is responsible for managing a PD project. Communication can reach all participants in the network with a group containing only the first seven nodes, Table 3. This measure could be usefully in identifying resource over- or under- utilization. However, the top seven or even ten nodes selected by other measures do not necessary reach 100% of the other nodes in the network. We will continue to use the top 10 nodes set selection as a baseline group for comparison in this study.

Group Degree Central-ity	% Reach
G1	59.3
A1	70.4
B1	74.1
C3	79.6
E5	94.4
H6	98.1
H3	100
F2	100
D4	100
H9	100

Table 3: Nodes based on group degree centrality measures. Shows the cumulative reach as additional nodes are added

When we compare individual degree and group degree centrality, we observe 30% overlap (highlighted nodes) amongst the top 10 nodes given by the rank ordering shown in Table 4. Recalling the formulation for group centrality, the measure counts a connection to the group only once regardless if multiple group members are connected to the same non-group member node. This observation might suggests that some of the top scoring nodes based on the individual centrality measures are all connected to many of the same nodes. Table 5 shows a comparison of the reach between the group selected by group degree-centrality and the group selected by individual degree-centrality. A top 10 group selected using individual degree centrality reaches only 92.6% of the other nodes in the network. We can speculate that information will not reach 8.4% of the teams. This is approximately five teams based on the 54 considered in this study.

Individual Degree Centrality	Group Degree Centrality
G1	G1
H3	A1
H10	B1
D2	C3
E5	E5
G2	H6
D1	H3
B7	F2
H8	D4
H1	H9

Table 4: Overlapping matrix comparing individual and group degree centrality

[3] This measure will be used for comparison to group degree-centrality throughout the paper.

[4] We have experimented with other networks and we did not find strong correlations between the various measures.

	% of Nodes Reached
Group Degree Centrality	100%
Individual Degree Centrality	92.60%

Table 5: Reach comparison for top 10 nodes based on group and individual degree centrality

Figure 3 and Figure 4 show a graphical representation of the top 10 nodes for individual degree and group degree centrality, respectively. By observation, the group centrality graph appears to be more extrovert, meaning that there is less interplay amongst the group's members. We speculate that the top 10 nodes selected are more connected to other teams in the network than to each other. The individual degree-centrality graph appears to be highly introvert; meaning there is a high level of interplay between the group's members. These groups are more connected to each other than other teams in the network. These speculations are based on the number of links amongst the top 10 scoring nodes selected by each measure. As can be easily observed, the number of links amongst the top 10 scoring nodes selected using the individual-centrality is greater than the number of links amongst the nodes selected based on group-centrality, see Figures 3 and 4. Whether a group is characterized as introverted or extroverted can have a significant impact on the project. Hollingshead (1998) found that the time to solution was much faster in groups that frequently work together or communicate frequently. This is a possible trade-off encountered that is dependent upon the measure used for selecting a top 10 group. Later in the discussion, we will suggest some managerial implications for selecting a group based on these measures.

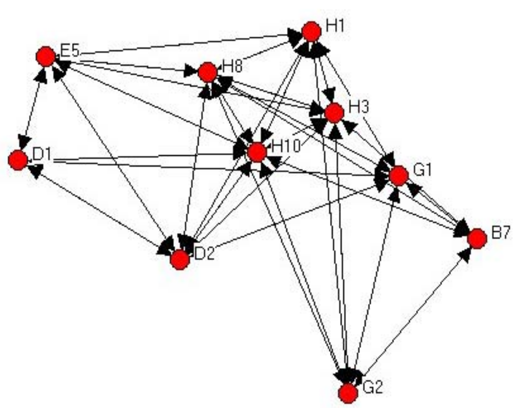

Figure 3: Individual Centrality Graphical Representation

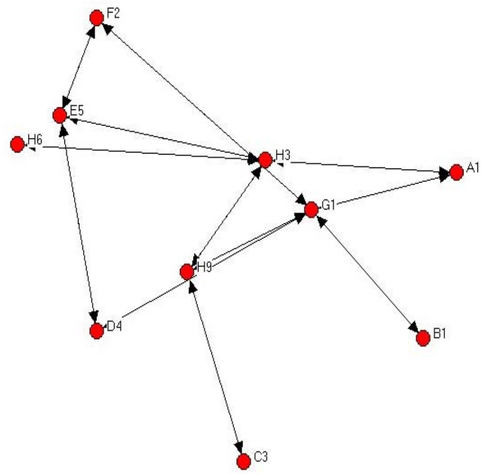

Figure 4: Group Centrality Representation

The last measure considered in this study is the Key Player Problem (KPP) (Borgatti, 2001). We observe both similarities and differences between the KPP-1 & KPP-2 node selections. Tables 6 and 7 show the node sets for KPP-1 and KPP-2, respectively. If the PD teams identified using KPP-1 are removed from the network, 66.9% of the remaining network is disconnected. We speculate that this would result in 66.9% communication breakdown. On the other hand, a set of nodes identified using the KPP-2 measure identifies a group of well-connected, non-redundant nodes. As a result, these nodes reach 100% of all the other nodes in the network. Furthermore, these nodes are "gatekeepers" of information for distinct segments of the network. If any one of these nodes fails, a part of the network is completely isolated. There is no redundancy; no other node (participants) will begin communicating with this now isolated group. However, the KPP-2 method could improve the identification of communication failures within the network because the node(s) that failed is from a small identifiable group.

KPP-1	% Fragmentation
G1	46.20
D2	47.40
G2	49.40
H10	51.00
D1	54.10
H3	55.50
E5	59.80
F1	63.70
B7	65.30
H8	66.90

Table 6: KPP-1 node selection. If removed, 66.9% of network becomes disconnected. Shows cumulative percent fragmentation with addition of nodes to selected set

KPP-2
D1
A1
F1
E4
B7
H3
A1
C4
D2
G2

Table 7: KPP-2 node selection. Has 100% connectivity, but no redundancy.

Figures 5 and 6 show a graphical representation of KPP-1 and KPP-2, respectively. The KPP-1 graph looks similar to the individual-centrality graphical representation. We speculate that the same argument for introvert characteristics that holds for individual centrality also holds for the KPP-1. We are not able to propose a similar argument for the KPP-2 graph. The KPP-2 graph has a more distinctive behavior in comparison to the other three graphs. These observations are based on visual inspection of the graphs.

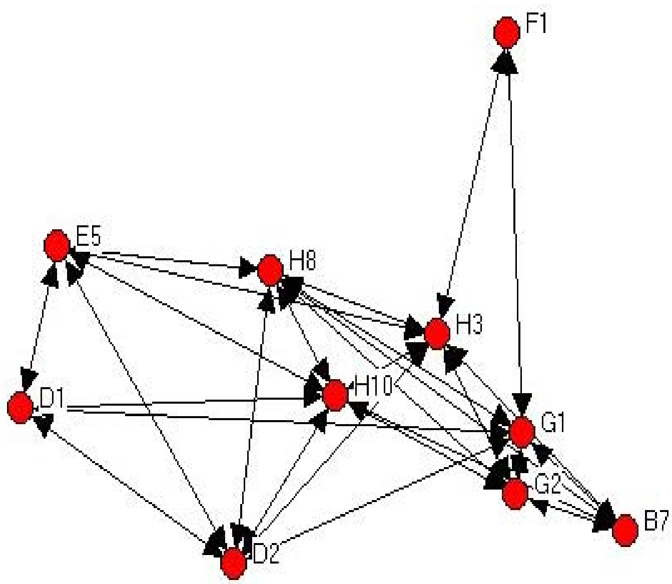

Figure 5: The KPP-1 graphical representation

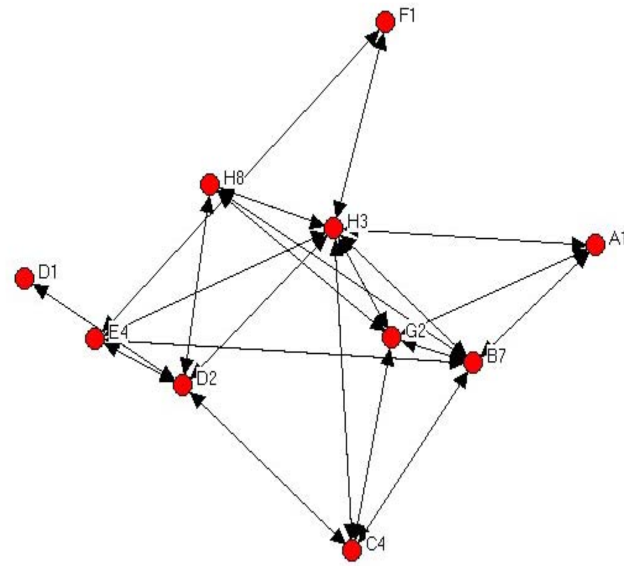

Figure 6: KPP-2 graphical representation

To summarize the different NA measures, we construct a correlation matrix to quantify the differences and similarities observed. Tables 8 and 9 show the top 10 scoring node sets for each measure and a correlation matrix, respectively. The correlation is based on the presence of a given node in the top 10 scoring set for each measure. We observe that there is a .80 correlation between individual degree-centrality and KPP-1. There is a -.60 correlation between group degree centrality and KPP-2. These results reiterate some of the differences and similarities encountered using the different measures.

Individual Degree Centrality	Group Degree Centrality	KPP -1	KPP -2
G1	G1	G1	D1
H3	A1	D2	A1
H10	B1	G2	F1
D2	C3	H10	E4
E5	E5	D1	B7
G2	H6	H3	H3
D1	H3	E5	C4
B7	F2	F1	D2
H8	D4	B7	G2
H1	H9	H8	H8

Table 8: Node sets for each NA measures

Pearson Correlations				
	Individual Degree Centrality	Group Degree Centrality	KPP-1	KPP-2
Individual Degree Centrality	1	-0.4	0.8	0.2
Group Degree Centrality		1	-0.4	-0.6
KPP-1			1	0.4
KPP-2				1

Table 9: NA measures correlation matrix

4.2 Discussion of Results from the Network Measures Perspective

In Section 4.1, we provide a general overview of results and comment on the observed differences and similarities. In this section, we reflect on the results in light of the different NA measures considered in this study.

One immediate observation is the direct representation or lack thereof of each design group within the selected top 10 node set. Recalling that there are 8 groups (6 component design groups and 2 integration groups), each of these groups is divided into component design teams. We have labeled these groups (A1, A2….A7 through H10). We observe that the nodes selected based on group centrality and the KPP-2 method has representation from all 8 groups, meaning there is a node from each group A to H contained within the selected set. However, the selected nodes for each set are very different which is confirmed by a negative correlation, Table 9. This suggests that these methods would seek input from all groups but could identify different information managers. This is a result of the mathematical construct of the measures. That is, the KPP-2 node set has no redundant connections amongst group members selected. However, the set selected by group degree centrality can have multiple group members that are connected to the same non-member node(s); it only counts this connection once in the calculation.

The nodes selected based on KPP-1 are missing representation from groups A and C. Furthermore, the nodes selected based on individual degree-centrality are missing representation from groups A, C, and F. If we choose the top 10 nodes based on KPP-1 or individual node centrality, some groups will be only indirectly connected to this group of top scoring nodes if connected at all. In some cases this could be problematic from a project management perspective, since it could lead to occurrences were information is not properly communicated to all teams involved in the PD project, creating a negative impact on the project.

We observe both similarities and difference amongst the sets, depending on the NA measure used for analysis. As stated earlier the individual-centrality and the KPP-1 measures neglected some teams. From a managerial perspective, it seems important that all designs teams are involved in design decisions. If some groups are not involved in the decisions, there could be conflict and miscommunication when information passes through the product development network. However, a benefit of choosing a set using one of these methods is that the selected set is very cohesive (several links amongst selected set of nodes) and information variation will be reduced for the teams represented by this selected set. Because of frequent communication amongst the nodes contain within the selected sets, design problems can be recognized at early stages of the project. This is the result of fewer steps to reach others. There is lots of interplay amongst this node set. One caveat, because groups selected using these measures communicate so frequently, the volume of information flow could be counter-productive, meaning that decision-making is much longer than would be expected. An additional implication could be lack of creativity within this set. We speculate that the project will have a high level of centralized control.

Group degree-centrality or KPP-2 will select a set of nodes that is representative of all the design groups. Information would most likely diffuse through the *entire* network more efficiently -correct and consistent. This could impact time to development and subsequently time to market. This is a key benefit of using either of these two methods in identifying the top 10 information nodes in the network. However, there is less cohesiveness amongst this node set. This could results in less frequent communication amongst the selected set (fewer links amongst set). Each group would operate nearly autonomously. Some design problems may not be recognized until the later integration stage of the PD project.

These results further support our assumption that the choice of the NA measures used should be based on the objectives for a particular situation. These measures should be considered as a part of a family of network analysis tools for improving product development processes.

4.3 Discussion of Results from a Design Team Perspective

In Section 4.2, we provide a general overview of statistical results and comment on the observed differences and similarities based on the NA perspective. In this section, we reflect on the numerical results while considering the role of the various teams involved in product development. Table 10 below shows the node-labels and the actual engineered component for all nodes appearing in a top 10 selection set based on at least one of the measures used in this paper.

Node Label	Engineered Component
A1	Fan Containment Case
B1	LPC Airfoils
B7	LPC Intermediate Case
C3	HPC Variable Vanes
C4	HPC Fixed Stators/Cases
D1	CC Burner
D2	CC Diffuser
D4	CC Diffuser Tubes
E4	HPT Rotor
E5	HPT Case & Blade Outer Air Seal
F1	LPT Shaft
F2	LPT Case
G1	MC Main shaft
G2	MC Gearbox
H1	EC Tubes
H3	EC Electrical Controls
H6	EC Airsystem
H8	EC Ignition
H9	EC Sensor
H10	EC Mechanical Control

Table 10: Engineered components of the jet engine appearing in the top 10 for at least one measure

One immediate observation is that node G1 (Main-shaft) is ranked number one[5] for overall individual-centrality, group-centrality, and the KPP-1 method. This suggests that the main shaft design team is very important to the PD process. From the engineering design perspective does this really make sense? In this network, G1 represents the main-shaft, this seems like a feasible result. The main-shaft could act as a bus by which the additional components are attached (Sharman and Yassine, 2004).

Considering H10, another high ranking node in the network, demonstrates another feasible possibility. This node represents the electrical controls. We can speculate that all electronic and electromechanical components will need to interface with this component or subsystem; therefore, the electrical controls design team will exchange information with several of the other teams.

Some additional results arise when considering nodes that did not appear in the highest scoring top 10 node set, but are important from an engineering perspective. An immediate observation is node D5, which represents the Combustion Chamber Fuel Nozzle. If we only considered individual centrality, this node is ranked 37th based on individual degree-centrality, 46th based on closeness-centrality, and 45th based on betweenness-centrality out of a total of 54. This suggests that the fuel nozzle design team does not have strong informational control within the network. As a result, it appears that the design of the fuel nozzle does not significantly impact the overall design of the jet engine and can be designed independently. So would these results for G1, H10, or D5 be supported based on the product architecture?

4.4 Discussion of Results from a Product Architecture Perspective

So far we have made all the calculations based on the product development team network. In this section, we consider the product architecture (physical network) and perform the same analysis instead. We find that the fuel nozzle component interfaces or has dependences with very few other components. Therefore, it makes sense that the fuel nozzle design team is ranked low based on team interaction network analysis. We observe that it is ranked low based on both product design team and product architecture network measures. Based on these results, we suggest that the fuel nozzle design changes do not impact a large number of other components. In contrast, we find that G1 (main-shaft) and H10 (electrical controls) are ranked high based on both product architecture and product design team analyses. We speculate that any change to the main-shaft or electrical controls will have far-reaching impact on the network from both a product design and an information flow perspective.

Other more interesting results arise when considering nodes that rank high based on team interaction analysis but rank low when analyzing the product architecture network, or vice versa. One node that exhibits high product team connectivity, but low product interfacing or dependences is B7 (Low Pressure Compressor Intermediate Case). This design team is ranked in the top 10 scoring nodes in three of four measures, which suggests that this team interacts (exchange information) with a number of other teams in the network. However, this component interfaces with only four other components. One node that exhibits low product team connectivity, but high product interfacing is H7 (EC Harness). This component interfaces with fifteen other components in the network, but is not selected in the top 10 scoring nodes based on any of the measures. We would have to select at least the top fifteen nodes for it to show up in two of the measures considered in this study.

When comparing both networks (design team and product architecture), we find that there is a 66.9% chance that if two components interface or have dependences in the product architecture network, their corresponding teams in the design team network exchange information[6]. If we assume this value to imply efficiency, we can suggest that this PD network has 66.9% efficiency in information flow and project

[5] This ranked was calculated by assuming that we desire to select only one team for all NA measures.

[6] This is based on a simple matching algorithm between the product team interaction and product architecture matrices. This is performed automatically using UCInet.

management. Recalling the previous definition of efficiency was the dissemination of correct and consistent information, would this efficiency be characteristic of the real-world product development network, meaning are deadlines met 66.9% of the time and delays occur 33.1% of the time for any task undertaking within the PD process? Is 33.1% of total development time used correcting errors or fixing design problems? We can speculate that the network efficiency will impact critical product development issues such as time to market and development time. Sosa et al. (2004) show evidence of misalignment between the product architecture and product organizational structure, which our findings support. The observations suggest that there is an informal network of communication that could possibly be impacting the PD process, either negatively or positively. Also, there may be communication channels missing that should be present in the network. There are components that interface or have dependencies, but their design teams are not exchanging information. This could suggest a re-evaluation of the current assignment of components to the different design teams and/or perhaps a better product architectural configuration.

5.0 MANAGERIAL INSIGHTS

A motivating factor for using network analysis to study product development processes is to gain an understanding of relationships amongst subsystems/components and their corresponding design teams. An additional benefit is that these methods can also identify non-obvious relationships that might exist. Understanding these relationships is important for the proper management of the PD process, such as managing overlapping product development activities.[7] Krishnan et al. (1997) suggests that by removing the coupling between development activities the overall process can be accelerated. This acceleration is the results of more frequent exchanges of preliminary information in contrast to a one-time transfer of information at boundaries. Network analysis can be useful in identifying which nodes in the network are most highly connected. The overlapping sequences can be controlled by processing information through these highly-connected nodes.

A second implication is in the outsourcing decision for subsystems and/or components. If a specific set of nodes consistently score high on most of these network measures, this could lead to a decision to make in-house as opposed to buy (i.e. outsource). In contrast, a group that is peripheral (not highly connected) in overall connectivity could be more effectively managed with an outsourcing partnership. For example, in the PD network studied in this paper the mainshaft (G1) is integral to the final product (i.e. based on the NA methods used, this component interfaces with a large number of other components and its design team interacts with a large number of other teams). This would lead to a decision to make this component because of the criticality of this component to the overall jet engine development. The NA methods could augment more traditional methods for making outsourcing decisions-based on cost efficiency and core competencies.

A third implication is in the decision of product architecture. Although this case analysis looks more in depth at team interaction, these interactions can be mapped to the actual product architecture. We observe considerable evidence of relationship agreement between the PD team network and the product architecture network (66.9% overlapping). Typically, product architecture protocols strongly favor modularity. The idea is to make modules as independent as possible. This reduces the number of interfaces and is believed to make product assembly much easier. NA methods can highlight informal relationships between subsystems, which could lead to new discussions regarding product architectural configurations for future products. An important question is deciding where to place module boundaries. For example, if we consider the NA betweenness-centrality measure (set of nodes connecting non-adjacent nodes), we can identify nodes that regulate and greatly influence the transmission of information within the network. Nodes ranking high on this measure can be considered to be at the interface of two non-adjacent nodes. This could suggest the location for redrawing product architecture boundaries.

6.0 SUMMARY

In this study, we investigate the use of network analysis techniques for studying information flows within a product development team network. As a case study, we consider the PD process for a large commercial aircraft engine. We use two key measures, centrality (individual and group) and the key-player method (KPP-1 and KPP-2) to select a set of ten top scoring nodes that are important to the flow of information within the network. The choice of which measure(s) to use for selection is dependent upon the managerial objectives for the PD process. These objectives could be (1) universal team involvement, (2) high level of managerial control, (3) high level of autonomy, or (4) outsourcing candidates to name a few. Furthermore, we consider the benefits of NA methods to gain more insight into the relationship between organizational structure (i.e. team network) and the product architecture (i.e. physical network). There could be future implications of these network analyses on the upstream organizational design and managerial decision making for early stages in product development processes.

REFERENCES

Batallas, D.A., Yassine, A., 2006. Information Leaders in Product Development Organizational Networks: Social Network Analysis of the Design Structure Matrix. Forthcoming, IEEE Trans. Eng. Mgmt.

Borgatti, S.P., 2001. The Key Player Problem, Carroll School of Management, Chesnut Hill, MA.

[7] Sequential design activities that can be executed in parallel.

Borgatti, S.P., M.G. Everett, and L.C. Freeman, 2002 UCInet for Windows: Software for Social Analysis, Harvard Analytic Technologies.

Braha D., Bar-Yam, Y., 2004., Information flow structure in large-scale product development organizational networks. Journal of Information Technology (19) 244-253.

Browning, T.R., 2001. Applying the Design Structure Matrix to System Decomposition and Integration Problems: A Review and New Directions, IEEE Transactions on Engineering Management, Vol. 48, No.3.

Buehler, R., Messervey, D., Griffin, D., 2005. Collaborative Planning and Prediction: Does Group Discussion Affect Optimistic Biases in Time Estimation? Organizational Behavior and Human Decision Processes. Vol 97, pp. 47-63.

Everett, M.G., Borgatti, S.P., 1999. The Centrality of Groups and Classes., Journal or Mathematical Sociology. 23(3): 181-201.

Harary, F. 1969. Graph Theory, Reading, MA. Addison-Wesley.

Holtta, K., Otto, K., 2003 Incorporating design complexity measures in architectural assessment. In Proc of ASME Design Engineering Technical Conferences, Chicago, IL. September 2-6.

Krishnan, V., Eppinger, S.D., Whitney, D.E., 1997. A Model-Based Framework to Overlap Product Development Activities. Management Science. Vol. 43, No.4, pp 437-450.

Morelli, M.D., Eppinger, S.D., Gulati, R.K., 1995. Predicting Technical Communication in Product Development Organizations. IEEE Transactions on Engineering Management, Vol. 42, No. 3, August.

Moreno, J.L., 1934. Who Shall Survive?: Foundations of Sociometry, Group Psychotherapy, and Sociodrama. Washington, D.C., Nervous and Mental Disease Publishing Co Reprinted in 1953 (Second Edition) and in 1978 (Third Edition) by Beacon House, Inc., Beacon, NY.

Quey, T., 2004. New Product Development Process: BADM 425, University of Illinois-Urbana/Champaign.

Seitamaa-Hakkarainen, P., Lahti, H., Muukkonen, H., Hakkarainen, K., Collaborative Designing in a Networks Learning Environment. Collaborative Design Proceedings of CoDesigning 2000. Chp 38, pp. 411-420.

Sharman, D.M., Yassine, A.A., 2004. “Characterizing Complex Product Architecture. Systems Engineering. Vol. 7, Issues 1, pp 35-60.

Sosa, M., Eppinger, S.D., Rowles, C.M., Designing Modular and Integrative Systems. Proceedings of DETC ’00. ASME 2000 International Design Engineering Technical Conference and Computers and Information in Engineering Conference, Baltimore, Maryland, September 10-13, 2000.

Sosa, M.E., Eppinger, S.D., Rowles, C.M., 2003. Identifying Modular and Integrative Systems and Their Impact on Design Team Interactions, Transactions of the ASME, (125), 240-252.

Sosa, M.E., Agrawal, A., Eppinger, S.D., Rowles, C. M., 2005. A Network Approach to Define Modularity of Product Components. Proceedings of IDETC/CIE, ASME.

Sosa, M.E., Eppinger, S.D., Rowles, C.M., 2004. The Misalignment of Product Architecture and Organizational Structure in Complex Product Development. Management Science (50), No. 12, December, 1674 - 1689.

Steward, D. 1981. The Design Structure Matrix: A Method for Managing the Design of Complex Systems. EM-28(3) 71-74.

Wasserman, S., Faust, K. 1994. Social Network Analysis. Cambridge University Press, New York.

Whitney, D.E., 2003. Links Between Strucutre and Behavior in Complex Networks: What the Literature Says. ESD-WP-LIT-2003-02. Massachusetts Institute of Technology.

Whitney, D.E., 2004 Connectivity Limits of Mechanical Assemblies Modeled as Networks. ESD-WP-2004-07. Massachusetts Institute of Technology.

Yassine, A., Braha, D., 2003. Complex Concurrent Engineering and the Design Structure Matrix Method. Concurrent Engineering Research and Applications, September 2003, Vol. 11, No. 3, pp. 165-176.