14th ICCRTS: C2 & Agility

Evaluating Dynamics of Organizational Networks via Network Entropy and Mutual Information

Topic 2: Networks and Networking

Yuan Lin[1]
Kevin C. Desouza[1,2]
Sumit Roy[2]

[1]Information School, University of Washington
[2]Electrical Engineering, College of Engineering, University of Washington

Point of Contact: Kevin C. Desouza
kdesouza@u.washington.edu

April 13, 2009

Abstract

In this work, we apply two concepts from network theory – *network entropy* and *mutual information* – to characterize the agility of network-form organizations. These concepts are developed from Shannon's work on information entropy. We define the notions of network entropy and mutual information pertinent to measuring the structural evolution of organizations. We then show the application of network entropy and mutual information in assessing the agility of networked organizational structures. We hypothesize that given the same type of topological changes (e.g., adding a link), a larger *MI* indicates a more agile organizational network. The network structure will contain more change options and there is a greater likelihood of choosing one well suited to the environment. Illustrative networks are constructed for demonstration, we investigate scenarios of adding links to a network while holding the number of nodes fixed. Areas for future research are discussed.

Acknowledgements

This research was sponsored in part by the Office of the Assistant Secretary of Defense for Networks and Information Integration (OASD-NII) through its Command & Control Research Program and the Center for Edge Power at the Naval Postgraduate School. The authors thank Mark Nissen, David Alberts, and other participants at the 6th Edge Research Workshop (Monterey, CA) and the 13th International Command and Control Research and Technology Symposium (Bellevue, WA) for helpful comments and critiques on previous versions of the paper.

1. Introduction

The past decade has seen considerable research on networked forms of organization. Networked structures are often posted as alternative to traditional hierarchical forms (Powell, 1990; Thompson, 2003).[1] Interest in these organizational forms can be attributed to their facilitation of agility, collaboration, and information sharing, which are imperative for dealing with today's uncertain environments. While research on organizational, and social, networks continues to grow at an exponential rate, studies on performance of network structures remain scant. Specifically, research is needed to examine how to measure performance of networked organizations (Desouza et al., 2008). Moreover, from a pure structural perspective, any forms of organization, including hierarchies, can be represented with the basic network analytic constructs of nodes and links (Laumann, 1991). Hence, the work conducted in this paper applies to any form of organizational structure.

Various scholarly traditions have developed their own measures for evaluating network performance. Traditional network analytical approaches use the rich mathematical tools of graph theory to arrive at measures of network typology (e.g., node degree, path length, network density). The study of social networks models relationships of humans (Wasserman & Faust, 1994) and develops measures of roles and positions (e.g., centrality, structural equivalence), as well as groups (e.g., structural cohesion, clustering coefficient). These measures focus on the configuration of nodes and links. The interested reader is referred to Brass (1995) for a good overview of these measures. Telecommunication network theory, which deals with the transport of information between technical devices, focuses measures on the quality and quantity of information (e.g., error rate, throughput, channel capacity) and the transport cost (e.g., delay, packet loss) across networks. These measures focus more on the functionality of nodes and links. The interested reader can refer to Schwartz (1987) and van Mieghem (2006) for more details.

In an uncertain environment, a favorable organizational structure should be agile: it should be adaptive to innovations, flexible, responsive to changes, yet also robust and resilient to damages. Currently to the best of our knowledge, few measures can adequately capture the structural agility of organization (Alberts and Hayes, 2003; Dove 1994; Kumar and Motwani, 1995; Metes et al. 1998). We treat agility as a structural property of the organization since our focus is on organizational structure (as opposed to an operational property[2]). Structural properties can be more easily measured since there is no need to predict the specific change the system will encounter (Giachetti et al., 2003), which is necessary for assessing organizations with unanticipated environments.

This paper investigates the utility of two network performance measures – *network entropy* (*NE*) and *mutual information* (*MI*). They are developed from Shannon's information entropy, which is a fundamental basis of the theory of electronic network communication, but their application to organizational networks and social networks has yet to be explored. We are interested in organizations which have networked structures (i.e. have distributed entities that are connected for the purposes of sharing information and collaborating on actions) and reside in dynamic contexts. We define these new measures based on network degree distribution and graph isomorphism (Section 2), providing strategies for computing these measures. We then apply them to evaluate the agility of organizations (Section 3). The comprehensiveness of these two measures is useful due to the complexity of contemporary organizational structure and the frequency of reorganization in capricious environments.

---

[1] We use the term networked organizational structures and network structures interchangeably. The work presented here can be applied to network structures of various types. These networks can be purely technical (i.e. involving artificial devices), human (i.e. social networks), or a socio-technical (i.e. machine-human).

[2] Agility can be ascertained through study of *operational* properties. These properties cover the specific operational activities of an organization and its specific operating context.

2. Network Entropy

The classic definition of *thermodynamic entropy* attributed to Boltzmann measures the "disorder" in a system (Haddad et al., 2005). Given a set of macroscopic states that are observed, the thermodynamic entropy measures the amount of uncertainty about the underlying system configuration that is spread over different possible micro-states corresponding to the same observed macro-state. This uncertainty is captured by a probability distribution over the possible microstates. Shannon (1949) formalized this in reverse order by beginning with a (discrete) set of symbols $\{x_i\}$ (the microstates) such that the occurrence is described by a probability mass function $p(x_i) = P(X = x_i)$ (the probability of occurrence of the *i*-th symbol). The *information entropy* over the ensemble of symbols is given by

$$H(X) = -E[\log p(x)] = -\sum_i p(x_i) \log p(x_i) \tag{2.1}$$

where $\sum_i p(x_i) = 1$.

Within a network context, a key determinant of structure is the variability of the degree distribution of the nodes. The degree of a node in a network is the number of one-hop neighbors or edges of the node. The degree distribution p_i of a network is then defined to be the fraction of nodes in the network with degree *i*. In other words, if there are N nodes in total in a network and N_i of them have degree *i*, then

$$p_i = N_i / N \tag{2.2}$$

is the probability that a node is degree *i*.

Network entropy (*NE*) is developed from information entropy based on this probability distribution (Losee, 1990). To illustrate, consider the subnet G_1 in Figure 1 (a) with nodes v_1, v_2, v_3, v_4, v_5, and v_6. The *network entropy* of G_1 is

$$NE(G_1) = H_{G_1}(P) = -\sum_i p_i \log p_i = -p_1 \log p_1 - p_2 \log_2 p_2 - p_3 \log_2 p_3 \approx 1.46$$

The computation results are summarized in Table 1. If a link is added between v_1 and v_6, as shown in Figure 1 (b), we get a new subnet G_2 with

$$NE(G_2) = H_{G_2}(P) = -p_1 \log_2 p_1 - p_2 \log_2 p_2 - p_3 \log_2 p_3 \approx 1.58$$

If the original link between v_3 and v_6 in G_1 is broken, as in Figure 1 (c), for the new subnet G_3

$$NE(G_3) = H_{G_3}(P) = -p_0 \log_2 p_0 - p_1 \log_2 p_1 - p_2 \log_2 p_2 - p_3 \log_2 p_3 \approx 1.79$$

We can also join two subsets, say, G_1 and G_2 together to obtain the subset G_4 (see Figure 1 (d)) with total number of nodes N =12 and

$$NE(G_4) = H_{G_4}(P) = -p_1 \log_2 p_1 - p_2 \log_2 p_2 - p_3 \log_2 p_3 - p_4 \log_2 p_4 \approx 1.83$$

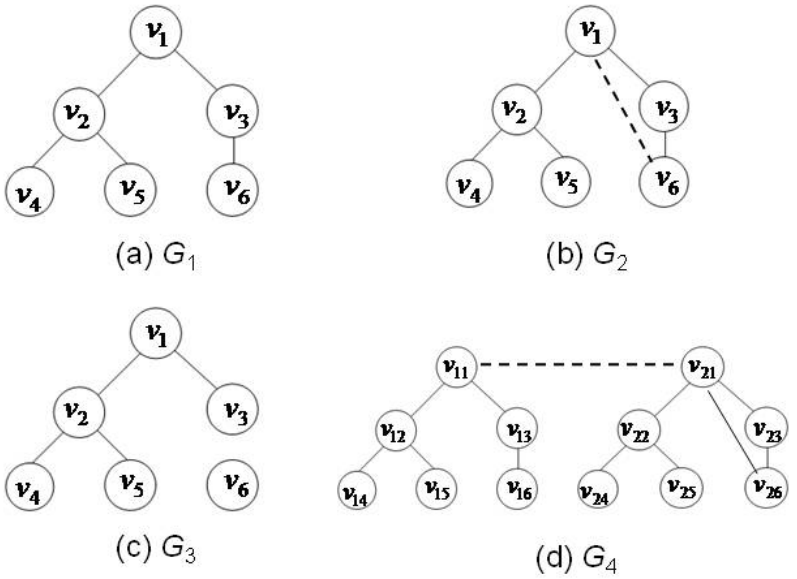

Figure 1: Subnets

Table 1: The calculation of network entropy

Subnet	P_0	$\log_2 P_0$	P_1	$\log_2 P_1$	P_2	$\log_2 P_2$	P_3	$\log_2 P_3$	P_4	$\log_2 P_4$	P_5	$\log_2 P_5$	*NE*
G_1	0/6	-	3/6	-1	2/6	-1.58	1/6	-2.58	0/6	-	0/6	-	1.46
G_2	0/6	-	2/6	-1.58	2/6	-1.58	2/6	-1.58	0/6	-	0/6	-	1.58
G_3	1/6	-2.58	3/6	-1	1/6	-2.58	1/6	-2.58	0/6	-	0/6	-	1.79
G_4	0/12	-	5/12	-1.26	3/12	-2	3/12	-2	1/12	-3.58	0/12	-	1.83

Like information entropy, *NE* attains its maximum value (log *N*) when $p_i = 1/N$ for any node degree *i*, and its minimum value 0 when some $p_j = 0$ and some $p_i = 1$. The former case corresponds to the most uncertain situation since there are many alternatives. The degree of a randomly picked node could be 0, 1,…, or $N-1$. The latter case corresponds to the least uncertain situation since there is just one possibility wherein every node has the same degree. *NE* depends continuously on the probabilities p_i so that two networks with similar degree distributions have close values of *NE*. In Figure 1, G_2 is obtained by connecting v_1 and v_6 of G_1. The degree of v_1 changes from 2 to 3, giving rise to a new degree alternative (degree = 3). The increased uncertainty results in an addition to *NE*: $NE(G_2) > NE(G_1)$. G_1 is obtained by connecting v_3 and v_6 of G_3. The degree of v_6 changes from 0 to 1. As a result, there is no longer an alternative of 0 degree in the network, which results in the decline of *NE*: $NE(G_1) < NE(G_3)$. G_4 is obtained by connecting G_1 and G_2. Not only does a new alternative appear, 4 degree, but the total number of degree alternatives is increased from 6 to 12. Uncertainty is captured by a growing *NE*: $NE(G_4) > NE(G_1)$, $NE(G_4) > NE(G_2)$.

3. Improving structural agility: maximizing mutual information

In a dynamic setting, *NE* alone cannot measure the uncertainty in a network structure, because network topological changes contribute to variances in uncertainty. To illustrate, consider adding a link to our original network. Depending on which nodes we connect, we will arrive at different degree distributions (2.2). For example, by connecting the nodes v_1 and v_6 in G_1 (Figure 1 (a)), we get G_2 (Figure 1 (b))

whose degree distribution is: $P_0 = P_4 = P_5 = 0$, $P_1 = P_2 = P_3 = 1/3$. However, if connecting another pair of nodes, say v_5 and v_6, we will get a degree distribution: $P_0 = P_4 = P_5 = 0$, $P_1 = P_3 = 1/6$, $P_2 = 2/3$. A new measure is needed to capture this uncertainty. In information theory, *mutual information* (*MI*) is defined as the change in the amount of uncertainty of the desired variable (*X*) by observing a related variable (*Y*). We use *MI* as a measure of the uncertainty caused by network topological change. The inclusion of the correlation between the two graphs helps to overcome the common yet unrealistic assumption that the growth phases of networks are independent of one another.

Assume any node in graph *G* whose degree changes from *X* to *Y*. Network-based *MI* is developed from the more general definition of mutual information $I(X;Y)$ given below, by replacing general entropy (*H*) with network entropy (*NE*):

$$I(X;Y) = H(X) - H(X \mid Y) = H(Y) - H(Y \mid X) = H(X) + H(Y) - H(X,Y) \tag{2.3}$$

The Venn diagram in Figure 2 shows the logic relations between entropy and mutual information.

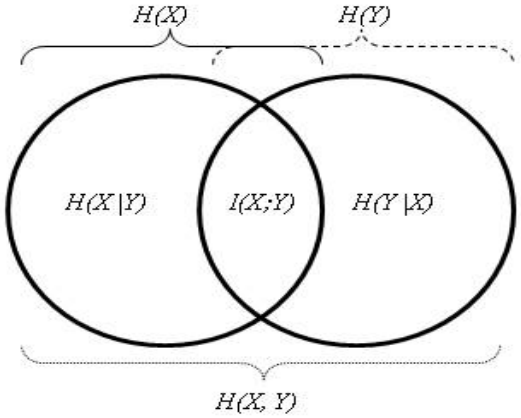

Figure 2: The logic relations between entropy and mutual information

$H(X,Y)$ is the joint entropy of *X* and *Y*, formalizing their correlation:

$$H(X,Y) = -\sum_i \sum_j p(x_i, y_j) \log p(x_i, y_j) \tag{2.4}$$

So we can also calculate $MI(X;Y)$ as below:

$$MI(X;Y) = \sum_i \sum_j p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x_i)\, p(y_j)} \tag{2.5}$$

where $p(x_i, y_j)$ is the value of the joint probability of variables *X* and *Y*, $p(X,Y)$,when $X = x_i$ and $Y = y_j$. $p(X,Y)$ is the product of $p(X)$, which is the probability of *X*, and $p(Y \mid X)$, which is the conditional probability of *Y* given *X*. Specifically,

$$p(x_i, y_j) = p(y_j \mid x_i)\, p(x_i) \tag{2.6}$$

where $p(x_i)$ is exactly the p_i in (2.2).

We propose to apply *MI* in improving the agility of organizational networks. Our hypothesis is: given the same type of topological changes (e.g., adding a link), a larger *MI* indicates a more agile organizational network, which has more change options. Thus, we are more likely to choose the one that best fits the environment. From the perspective of information theory, it means there are more possibilities (i.e., uncertainty) related to a certain type of topological changes, and more information is needed to determine which one is the case. Alberts and Hayes (2003) insightfully describes this relationship in their statement, "the road to agility is paved with information".

Let's investigate a scenario of randomly adding links to an *N*-node graph which has no links at the beginning. We will add one link a time until all nodes are fully connected. In other words, if *m* links exist at the former growth phase, there are *m* + 1 links at the following phase ($0 \le m \le N(N-1)/2-1$). To compute *MI*, three data structures are maintained for each possible state during network evolution. The first is an *N*-dimension vector indicating current degree distribution: $\rho = <\rho[0], \rho[1],..., \rho[N-1]> = <p_0, p_1,..., p_{N-1}>$, where p_k is the probability of the occurrence of a *k*–degree node. Different states have different degree distribution vectors. The second is the adjacency matrix *A*, an $N \times N$ matrix in which a non-diagonal entry $A[u][v]$ is the number of edges from node *u* to node *v* (the nodes are sequentially labeled as 0, 1…$N-1$). Each diagonal entry $A[u][u]$ has a value of 0. $A[u][v]=1$ if and only if nodes *u* and *v* are linked. We avoid linking two already linked nodes by checking the corresponding entry in the adjacency matrix. The third data structure is another *N*-dimension vector δ which records the degree of each node: $\delta = <\delta[0], \delta[1],..., \delta[N-1]> = <d(0), d(1),..., d(N-1)>$. After picking and linking a certain pair of nodes, the three data structures are updated accordingly.

We keep linking two nodes until any of the following requirements cannot be met: (1) the degree of either node is less than the maximally possible value (by checking the vector δ); (2) the nodes have not yet been linked (by checking the adjacency matrix *A*); (3) the produced graphs are non-isomorphic. Two graphs *X* and *Y* are isomorphic to each other if a relabeling of the vertices of *X* (i.e., a permutation of the labels) yields *Y* and vice-versa. Non-isomorphic graphs thus have different degree distribution (by checking the vector ρ). We do not differentiate individual nodes because *NE* and *MI* are organizational-level measures; they are computed based on the entire network. In other words, our unit of analysis is organizations. The links connecting the nodes and the overall network structure generated are of more interest. Consider three-node graphs ($N = 3$). There can be $\binom{\frac{N(N-1)}{2}}{0}+\binom{\frac{N(N-1)}{2}}{1}+\ldots+\binom{\frac{N(N-1)}{2}}{\frac{N(N-1)}{2}} = 2^{\frac{N(N-1)}{2}} = 8$ different labeled graphs (Figure 3). $\frac{N(N-1)}{2} = 3$ is the maximal number of links that a graph with $N = 3$ nodes can have and in this case the graph is fully-connected. $\binom{\frac{N(N-1)}{2}}{m}$ is the number of different graphs each of which has *N* nodes and *m* links. However, only four out of the eight graphs are non-isomorphic to one another, as shown in Figure 4.

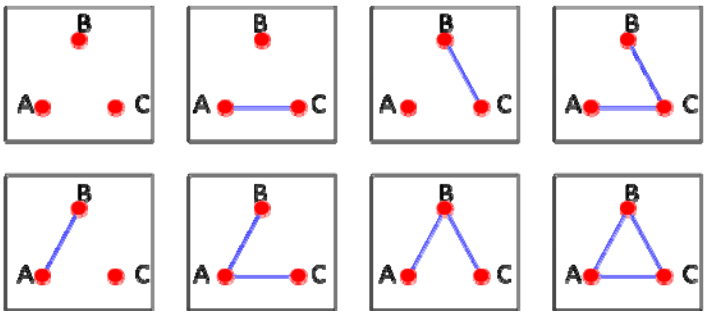

Figure 3: All eight graphs on three nodes

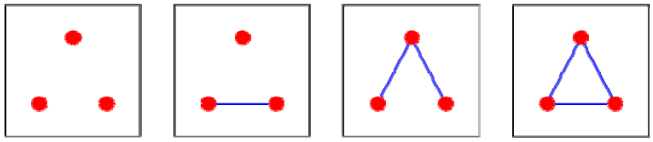

Figure 4: All four non-isomorphic graphs on three nodes

Now let's show the principle of maximizing *MI* using the evolution of 4-node networks as an example. In these networks, the number of links $m \in [0,5]$, and each node has a degree $d \in [0,3]$. The degree distribution satisfies $\sum_{k=0}^{3} p_k = 1$. $p(x_i)$ [$p(y_j)$] corresponds to p_i (p_j), the occurrences of an i–degree (j–degree) node. The whole process (adding one more link at a time) is displayed in Table 2 and summarized in Figure 5. In Table 2, the labels in the first column index different graphs. The second column illustrates different non-isomorphic graphs and the third column is the number of links in each type of them. The fourth column shows the cardinality of corresponding isomorphic graph families[3]. The fifth to the eighth column shows the degree distribution of certain graphs, and the final column shows their respective *NE*. The values of *MI* are shown in Figure 5, a flowchart of network evolution that moves from a fully unconnected graph (without any link) to a fully connected graph (with all possible links). Each circled number represents a possible state (a network structure with certain degree distribution) in the evolution. The number next to each state is the corresponding *NE*. Subsequent states are linked with arrows pointing from the predecessors to the successor. The number next to each arrow is the corresponding *MI*. We discuss the calculation details below.

Table 2: Link addition procedure on a 4-node graph

#	*ill.*	*m*	*C*	*p*(0)	*p*(1)	*p*(2)	*p*(3)	*NE*[4]
00		0	1	4/4	0	0	0	0
10		1	6	2/4	2/4	0	0	1

[3] It is the number of isomorphic graphs. For example, the cardinality of 3-node, 2-link isomorphic graphs is 3, as shown in Figure 2.

[4] It is calculated using log base 2.

Table 2: Link addition procedure on a 4-node graph (continue)

20		2	3	0	4/4	0	0	0
21		2	12	1/4	2/4	1/4	0	1.5
30		3	12	0	2/4	2/4	0	1
31		3	4	0	3/4	0	1/4	0.81
32		3	4	1/4	0	3/4	0	0.81
40		4	1	0	0	4/4	0	0
41		4	14	0	1/4	2/4	1/4	1.5
50		5	6	0	0	2/4	2/4	1
60		6	1	0	0	0	4/4	0

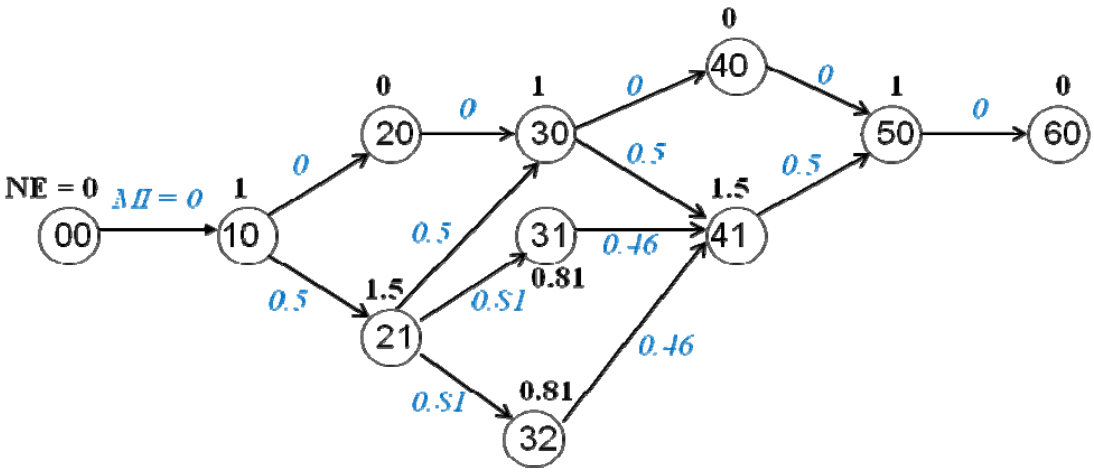

Figure 5: Link addition procedure on a 4-node graph (branch-growing pattern)

Suppose we are adding a link to the graph "10" in Table 2, which results in either "20" or "21." We first need to calculate the value of $MI(10;20)$ and $MI(10;21)$. We then label the nodes clockwise from 0 to 3. Since "10" has two 0-degree nodes and two 1-degree nodes, $p(x_0) = p(x_1) = 1/2$. The graph does not have any nodes with higher degree. $\rho(10) =< 2/4, 2/4, 0, 0 >$, $\delta(10) =< 1, 1, 0, 0 >$, $A(10) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

According to the above requirements and $\rho(10)$, we can link two 0-degree nodes, which produce the graph "20", or link a 0-degree node and a 1-degree node, which produces the graph "21". Each node in

graph "20" has a degree of 1, $\rho(20) =< 0,4/4,0,0 >$. $p(y_1) = 1$, $p(y_0) = p(y_2) = p(y_3) = 0$. Any node in "10", regardless of its original degree, becomes a 1-degree node in "20," i.e., $p(y_1 \mid x_0) = p(y_1 \mid x_1) = 1$. $p(x_0, y_1) = p(y_1 \mid x_0)\ p(x_0) = 1/2$, $p(x_1, y_1) = p(y_1 \mid x_1)\ p(x_1) = 1/2$. $MI(10;20) = p(x_0, y_1)\log\frac{p(x_0, y_1)}{p(x_0)p(y_1)} + p(x_1, y_1)\log\frac{p(x_1, y_1)}{p(x_1)p(y_1)} = 0$. There are one 0-degree node, two 1-degree node, and one 2-degree nodes in "21", $\rho(21) =< 1/4, 2/4, 1/4, 0 >$. $p(y_0) = p(y_2) = 1/4$, $p(y_1) = 2/4$, $p(y_3) = 0$. A 0-degree node in "10" is about to have a degree of 0 or 1 in "21," i.e., $p(y_0 \mid x_0) = p(y_1 \mid x_0) = 1/2$. A 1-degree node in "10" is about to have a degree of 1 or 2 in "21," i.e., $p(y_1 \mid x_1) = p(y_2 \mid x_1) = 1/2$. Therefore, $p(x_0, y_0) = p(y_0 \mid x_0)\ p(x_0) = 1/4$, $p(x_0, y_1) = p(y_1 \mid x_0)\ p(x_0) = 1/4$, $p(x_1, y_2) = p(y_2 \mid x_1)\ p(x_1) = 1/4$, $p(x_1, y_1) = p(y_1 \mid x_1)\ p(x_1) = 1/4$.

$$MI(10;21) = p(x_0, y_0)\log\frac{p(x_0, y_0)}{p(x_0)p(y_0)} + p(x_0, y_1)\log\frac{p(x_0, y_1)}{p(x_0)p(y_1)} + p(x_1, y_1)\log\frac{p(x_1, y_1)}{p(x_1)p(y_1)} + p(x_1, y_2)\log\frac{p(x_1, y_2)}{p(x_1)p(y_2)} = 0.5.$$

According to our hypothesis, $MI(10;21) > MI(10;20)$ means a change from the graph "10" to "21" is greater than that to "20". This is intuitively verified by seeing the topological transformation (Table 2). In comparison with "10", "20" can be interpreted as increasing the number of unconnected pairs, whereas "21" introduces a hierarchy. Based on the principle of *MI* maximization, we can find the best path of network evolution with the largest sum of *MI*. In other words, it is the "longest" path in terms of *MI* in the network evolution graph (e.g., Figure 4 for 4-node networks). In this example it is "00→10→21→31(32) →41→50→60" and the sum of *MI* is 0 + 0.5 + 0.81 + 0.81 + 0.5 + 0 = 2.62. This is a dynamic programming (Bellman, 1966) problem in which optimal solutions to sub-problems are used to find the optimal solutions to the overall problem. For a network with certain number of nodes, we can first construct its evolution graph and track the opposite number of every *MI* and then use the classic Bellman–Ford algorithm (Bellman, 1958; Ford & Fulkerson, 1956) to find the shortest path in terms of the opposite number of *MI*, which is the longest path in terms of *MI*. It remains to be explained how to construct the evolution graph for any *N*-node network.

We propose to solve this problem using the tentative algorithm shown below, based on the three data structures mentioned above: degree distribution vector ρ, node degree vector δ, the adjacency matrix A. By repeatedly running this algorithm at every growth stage, we can obtain the whole evolving graph of the network and find the path that best supports agile network evolution.

BEGIN

1. Set up ρ_0, δ_0, A_0 for the precedent graph *X*; set a counter $c = 0$
2. Set $r = 0$
3. If $r < N - 1$, go on; otherwise, go to the end
4. If $\rho[r] = 0$, do $r + 1$ and go to Step 3; otherwise, go on
5. Set $t = r$
6. If $t < N - 1$, go on; otherwise, do $r + 1$ and go to Step 3.

7. If $\rho[t] = 0$, do $t+1$ and go to Step 6; otherwise, go on.
8. If $\rho[r] \geq 2/N$ (when $r = t$) or $\rho[r] \geq 1/N \wedge \rho[t] \geq 1/N$ (when $r \neq t$), go on; otherwise, do $t+1$ and go to Step 6.
9. Sequentially traverse the dimensions of vector δ from $\delta[0]$ until finding a pair of indices *u* and *v* which satisfy: (i) $\delta[u] = r$ and $\delta[v] = t$; (ii) $A[u][v] = 0$ (i.e., there is no link between them).
10. If such a pair of indices does not exist, do $t+1$ and go to Step 6; otherwise, go on
11. Construct new ρ_c, δ_c, A_c for one of the possible following graph Y_c: first copy the ρ, δ, A of X (get $\rho_c = \rho_0$, $\delta_c = \delta_0$, $A_c = A_0$); then do $\rho_c[r] - 1/N$, $\rho_c[t] - 1/N$, $\rho_c[r+1] + 1/N$, $\rho_c[t+1] + 1/N$, $A_c[u][v] + 1$, $\delta_c[u] + 1$, and $\delta_c[v] + 1$.
12. Compute $NE(Y_c)$ and $MI(X;Y_c)$ based on ρ_0 and ρ_c and the preceding formulas.
13. Do $A_0[u][v] + 1$, $A_0[v][u] + 1$, $r+1$, $c+1$ and go back to Step 3.

END

Admittedly, the correctness of this algorithm needs to be examined against the results of some well-established mathematical solution, e.g., the exact number of all different graphs (in terms of degree distribution) computed using *Pólya's enumeration theorem*. Moreover, the complexity of this algorithm can be refined. An organizational network in real life usually has huge number of nodes which keep changing all the time. It is impossible to calculate the probability of its appearance like above since *N* is infinite and the vectors and matrixes will have too many dimensions to work on.

Regardless, this algorithm is a start, which we tried on some small-scale simulated and real datasets. Figure 6 shows the entire link addition procedure and the longest path starting with a 6-node, 9-link random network, whose topology is shown as the first one in Figure 7. Figure 8 shows the procedure and the longest path starting with a 6-node, 9 link scale-free network, whose topology is shown as the first one in Figure 9. Figure 7 and 9 show the respective topological evolution of these two networks[5]. The end point of either procedure is a fully connected 6-node network, the last one in Figure 7 or Figure 9.

[5] They are all generated using *Pajek*, a program for Windows used for analysis and visualization of large networks. It is freely available, for noncommercial use, at its download page. (http://vlado.fmf.uni-lj.si/pub/networks/pajek/)

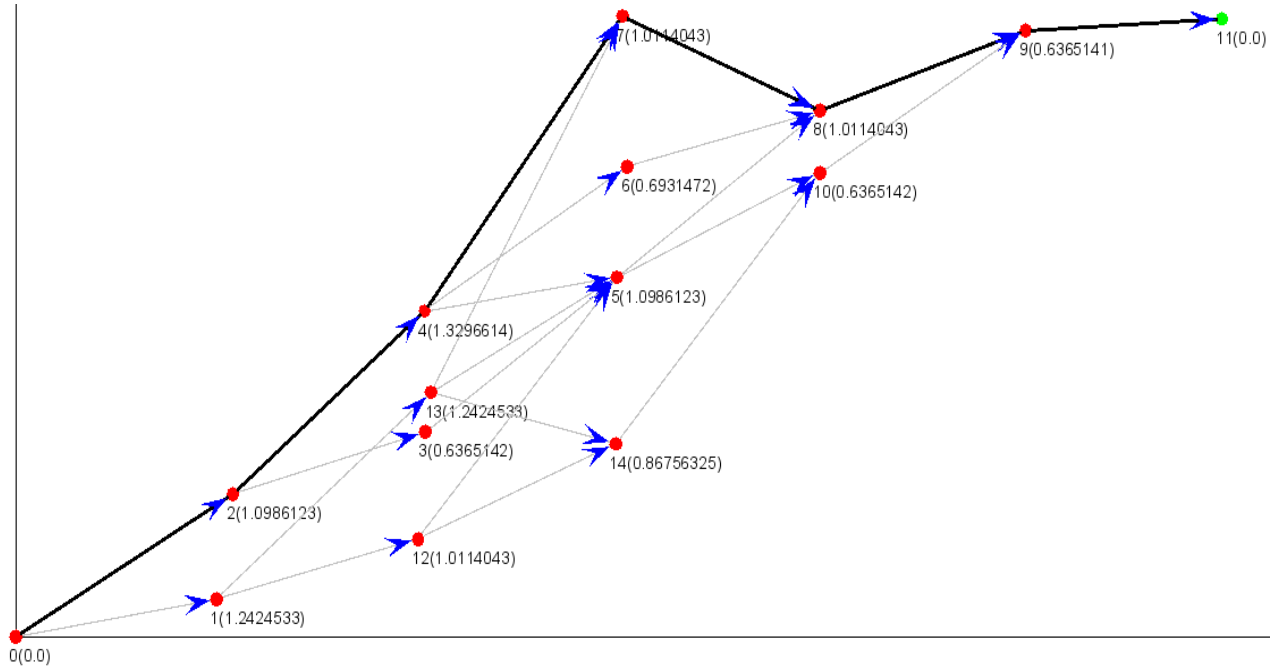

Figure 6: Adding 6 links to a 6-node, 9-link random network (flowchart)[6]

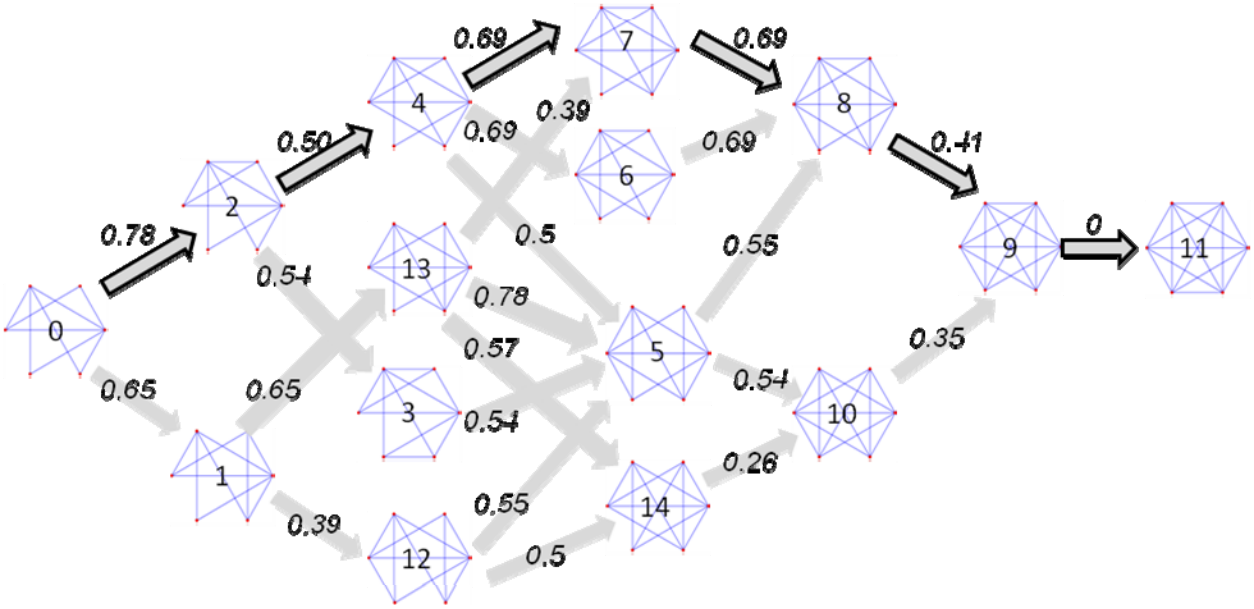

Figure 7: Adding 6 links to a 6-node, 9-link random network (topology evolution)[7]

[6] The numbers in the parentheses are the network entropy of corresponding graphs (calculated using natural logarithm).

[7] The numbers on the arrows are the mutual information (calculated using natural logarithm).

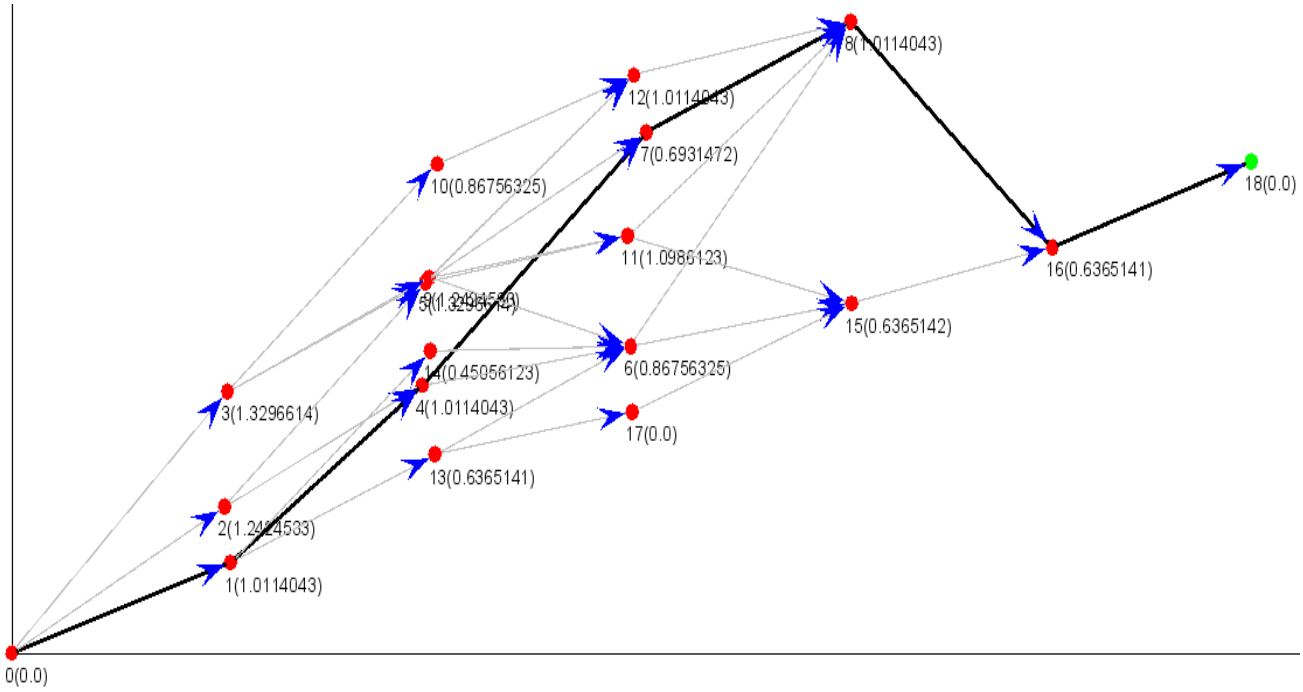

Figure 8: Adding 6 links to a 6-node, 9-link scale-free network (flowchart)

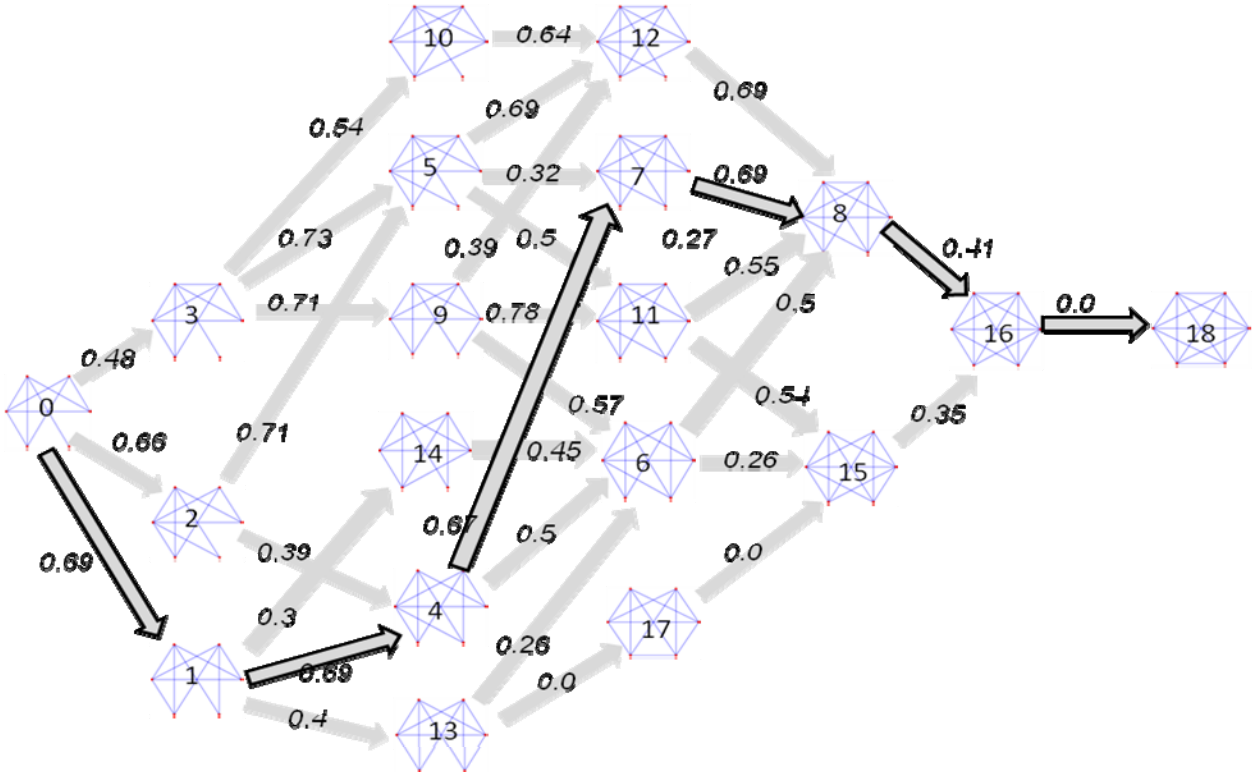

Figure 9: Adding 6 links to a 6-node, 9-link scale-free network (topology evolution)

Figure 10 shows the procedure and the longest path starting with an 11-node, 32-link real network. Four links are added to the original topology. This network is adapted from the data used in a previous study

by Hlebec (1993). The analyzed network consisted of communication interactions among twelve members and advisors of the Student Government at the University in Ljubljana. Data were collected with face to face interviews which lasted from 20 to 30 minutes and were conducted in May 1992.

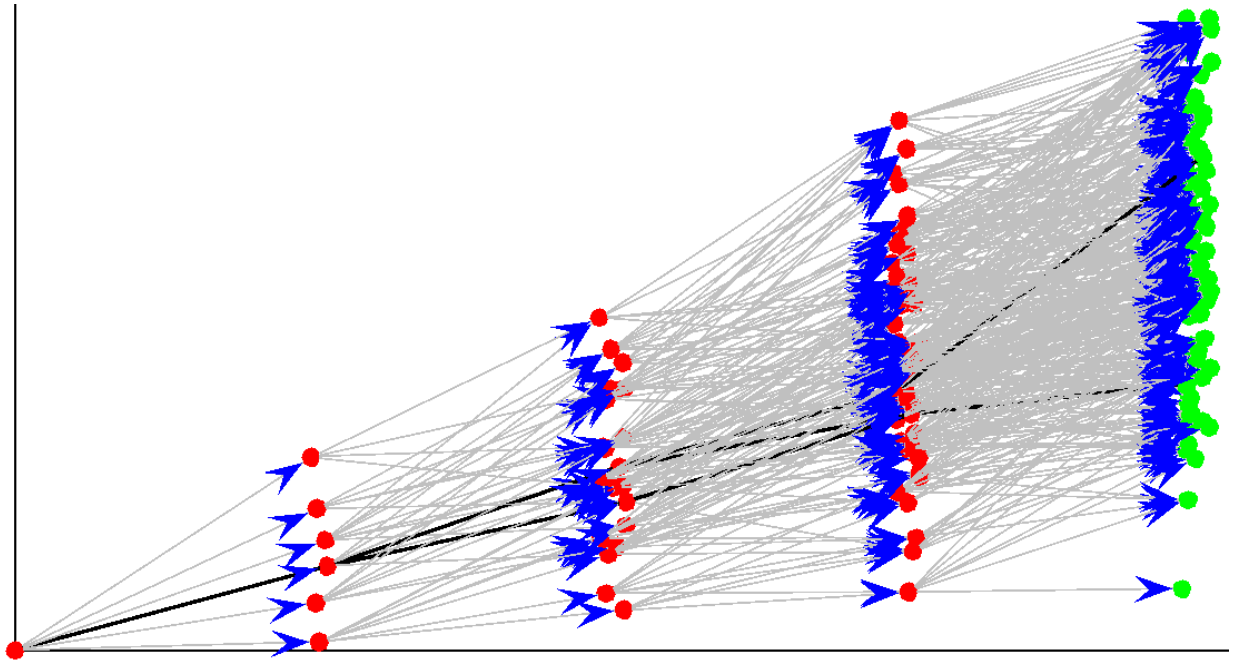

Figure 10: Adding 4 links to an 11-node, 32-link real-data network (flowchart)

Figure 11 shows the procedure and the longest path starting with a 10-node, 33-link real network. Six links are added to the original topology. This network is adapted from the data used in a previous study by Knoke and Kuklinski (1982). They selected a subset of 10 organizations and two relationships, of which we modeled the information exchange relationship.

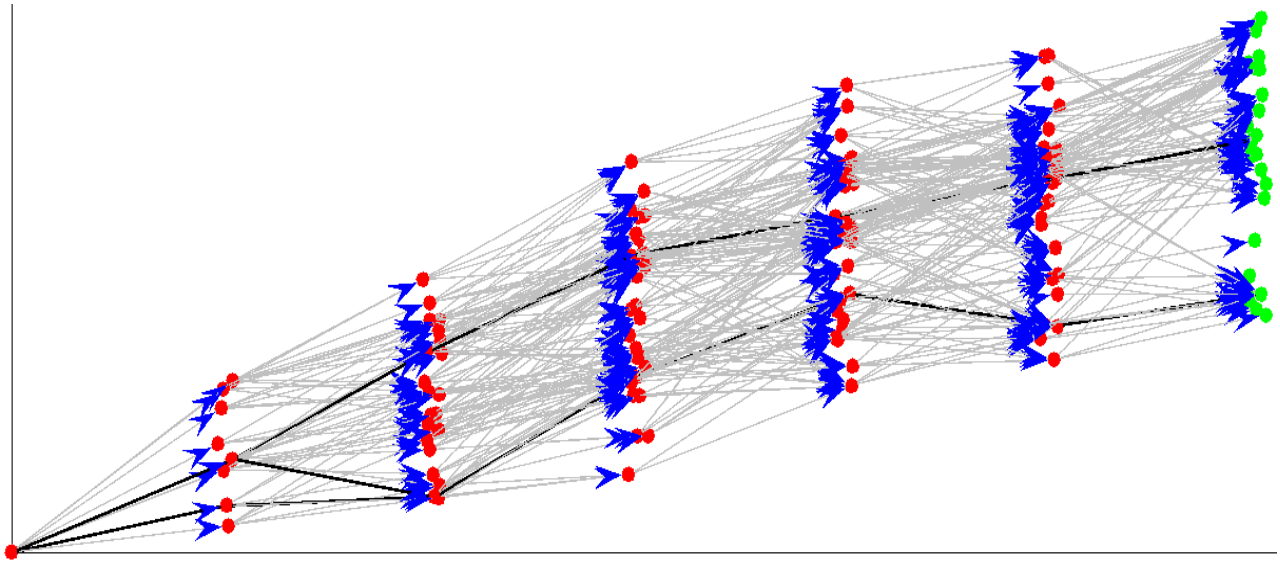

Figure 11: Adding 6 links to a 10-node, 33-link real-data network (flowchart)

This algorithm is also adaptable to other scenarios of topological changes. The first is the multiple-link case, i.e., more than one link ($l \geq 1$) is added to the network at the same time; we can then simply repeat our approach in Section 2.2 for l times. Second, the deletion of links is achievable by reversing the process, i.e., replacing addition with subtraction. The three rules need to be modified in this case. For example, the degree of each node should not be negative. Third, there are times the links are not added (or deleted) randomly. For instance, in a "rich-get-richer" scale-free network, the nodes with higher degree (i.e., hubs) are more likely to have nodes link to them. This can be implemented by introducing a weight function. A graph yielded by linking an a-degree node and a b-degree node in the original graph appears with a probability of $w(a,b)$, where

$$w(a,b) = \frac{a+b+0.1}{\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(i+j+0.1)}$$

(i, j) is the degree of any pair of nodes in the original graph which, if connected, produce one of the non-isomorphic graphs. (a,b) is an instance. A value of "0.1" (or other small values) is added in the case that both a and b are equal to 0. Take the evolution of the four-node network in Table 2 again. There are two non-isomorphic graphs "40" and "41" that occur from adding a link to "30". Graph 40 is formed by linking a 2-degree node and a 1-degree node, whereas "41" is obtained by linking two 1-degree nodes. The weight on "40" is given by $\frac{1+2+0.1}{(1+2+0.1)+(1+1+0.1)} = \frac{3.1}{5.2} \approx \frac{3}{5}$, which is larger than that on "41" as $\frac{1+1+0.1}{5.2} = \frac{2.1}{5.2} \approx \frac{2}{5}$

5. Conclusion and future work

This paper proposes two new measures –*NE* and *MI* – to evaluate the agility of organizational networks in a contemporary changeable environment. We argue that a more agile organizational network should exhibit a larger *MI* given the same kind of topological changes. We investigate scenarios of adding links to a network while holding the number of nodes fixed. We used small-scale simulated and real networks for demonstration. This exploratory study shows the potential of applying simple structural elements, such as degree distribution, to the investigation of a dynamic organizational network.

This work contributes to our understanding of organizational agility (see Figure 12). Specifically, assume Point *S* represents the current state of an organization, which is trying to reorganize in face of certain environmental shock. It has two choices: to centralize or decentralize. Then it has to decide the extent of centralization (or decentralization) by considering its goal and the cost.

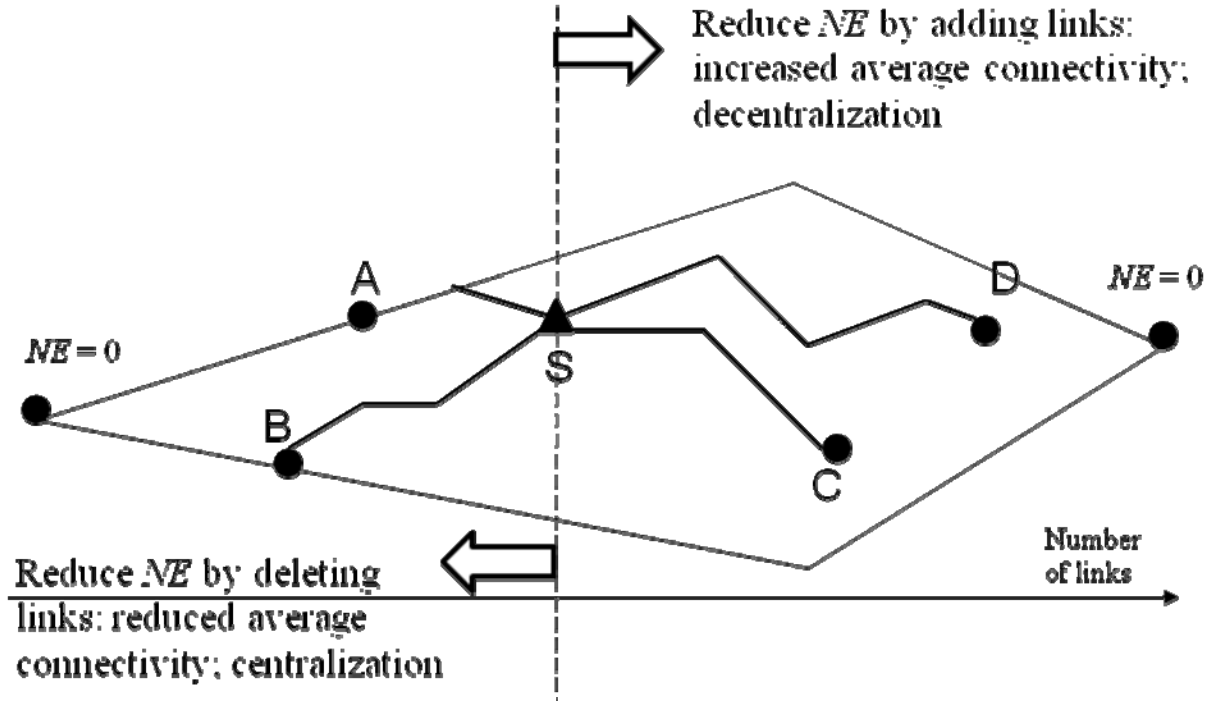

Figure 12: Making reorganization decisions using *NE* and *MI*

The value of *NE* indicates the equality of most nodes in a network. Low *NE* means most nodes have the same status in terms of degree. However, it could be that: (1) most nodes connect to a few hubs and are thus separated from each other, in which case there are relatively fewer links in the graph (the left arrow in Figure 12); (2) most nodes connect to each other, in which case there are relatively more links in the graph (the right arrow).

No matter which side the organization chooses, it is possible that there are more than one states which are equal in *NE*, e.g., States *A* and *B* on the left side of S, or States *C* and *D* on the right side of *S*. To choose from them, the organization may want to examine how many changes it needs to make and the accompanying cost. At this point, *MI* becomes an important index, which we argue should be a representation of the change cost. Then we find out the shortest paths in terms of *MI* between, say, (1) *S* and *A* (2) *S* and *B*. If the aggregate *MI* of path *SA* is bigger than that of path *SB*, the organization probably wants to change to State *B* instead of *A*.

This work can be extended and refined through further research. A social network in the real world typically has huge number of human nodes which keep changing all the time. Our current computation strategy is infeasible when network dimensionality grows significantly. We note that in the literature, there already exist models for the degree distribution of large scale networks, e.g., that of scale-free networks[8] whose degree distribution follows well-known power laws such as the *Zipf* distribution. Since *NE* and *MI* are defined based on degree distribution, a possible solution is approximation by use of these models. However, future work will also call for demonstrating validity and utility of the measures through

[8] Networks that typically consist of a few nodes (hubs) with large degrees and many others with smaller degrees

computer simulations. Networked organizational structures of different types and size will need to be studied, as do various scenarios of network evolution (e.g., adding or deleting links, combining or separating networks).

References

Albert, R. and Barabási, A. L. (2002). Statistical mechanics of networks. *Reviews of Modern Physics 74*, 1, 47 – 97.

Alberts, D. S., & Hayes, R. E. (2003). *Power to the Edge: Command, Control in the Information Age*. Washington, D.C: U.S. Department of Defense, Command and Control Research Program.

Arteta, B. M., & Giachetti, R. E. (2004). A measure of agility as the complexity of the enterprise system. *Robotics and Computer-Integrated Manufacturing, 20*(6), 495-503.

Bellman, R. (1958). On a routing problem. *Quarterly of Applied Mathematics*, 16:87–90.

Bellman, R. (1966). Dynamic Programming. *Science, 153*(3731), 34-37.

Bergstrom, C. T., & Lachmann, M. (2004). *Shannon information and biological fitness.* Paper presented at the Information Theory Workshop, 2004.

Brass, D. J. (1995). A social network perspective on human resource management. In K. M. Rowland & G. R. Ferris (Eds). Research in Personnel and Human Resource Management: Vol, 13: 39-79. Greenwich, CT: JAI Press.

Cover, T. M., & Thomas, J. A. (1991). *Elements of information theory*. New York: Wiley.

Desouza, K. C., Roy, S., & Lin, Y. (2008). Performance Measures for Edge Organizations. Paper presented at the 13th International Command and Control Research and Technology Symposium (ICCRTS).

Dove, R. (1994). The meaning of life and the meaning of agile, *Prod Mag* 106 (11), 14–15.

Ford, L. R. & Fulkerson, D. R. (1956). Maximal flow through a network. *Canadian Journal of Mathematics*, *8*(3):99–404.

Haddad, Wassim M.; Chellaboina, VijaySekhar; Nersesov, Sergey G. (2005). *Thermodynamics - A Dynamical Systems Approach*. Princeton University Press.

Hlebec, Valentina. (1993). Recall Versus Recognition: Comparison of the Two Alternative Procedures for Collecting Social Network Data. In A. Ferligoj & A. Kramberger (Eds). Developments in Statistics and Methodology. Metodološki zvezki 9, FDV, Ljubljana, pp. 121-129.

Hubbard, D. W. (2007). How to measure anything: finding the value of "intangibles" in business. Hoboken, N.J.: John Wiley & Sons. 46.

Knoke D. and Kuklinski J. (1982). Network analysis, Beverly Hills, CA: Sage.

Kumar, A. & Motwani, J. (1995). A methodology for assessing time-based competitive advantage of manufacturing firms. *Int J Oper Prod Manage* 15 (2), 36–53.

Laumann E. O. (1991). Comment on "The Future of Bureaucracy and Hierarchy in Organizational Theory: A Report from the Field." In P. Bourdieu & J. S. Coleman (Eds). Social Theory for a Changing Society: 1991, Boulder, CO: Westview. pp. 90-93.

Losee, R. M. (1990). *The science of information: Measurement and applications*. San Diego: Academic Press.

Metes, G., Gundry, J. & Bradish, P. (1998). *Agile networking: competing through the Internet and Intranets*. Englewood Cliffs, NJ: Prentice-Hall.

Nissen, M. E. (2007). Enterprise command, control, and design: Bridging C2 practice and CT research. *The International C2 Journal, 1*(1), 61–112.

Powell, W. W. (1990). Neither market nor hierarchy: Network forms of organization. *Research in Organizational Behavior, 12*: 295–336.

Schwartz, M. (1987). *Telecommunication networks: Protocols, modeling, and analysis.* Massachusetts: Addison-Wesley.

Shannon, C. E. & Weaver, W. (1949). *A mathematical theory of communication*. Urbana: University of Illinois Press.

Thompson, G. (2003). *Between hierarchies and markets: the logic and limits of network forms of organization*. Oxford; New York: Oxford University Press.

Van Mieghem, P. (2006). *Performance analysis of communications networks and systems*. New York: Cambridge University Press.

Wasserman, S. & Faust, K. (1994). *Social networks analysis: Methods and applications.* Cambridge: Cambridge University Press.