Evaluating dynamics of Organizational Networks via Network Entropy and Mutual Information

Yuan Lin α
Kevin C. Desouza α, β
Sumit Roy β

α The Information School
β Electrical Engineering, College of Engineering
University of Washington

Command and Control Research and Technology Symposium 2009
Organizational performance & network structure

• Organizational performance depends on timely access to information and the ability to use this information to make appropriate decisions.

• The structure of organizational network (formal & informal) impacts communication patterns and thus information diffusion.
Organizational adaptation & Network dynamics

• Uncertain environment asks for continuous organizational adaptation

• Organizational adaptation depends on the structural agility of organizational networks

• Structural agility means conducting intended network evolution efficiently.

• What is “intended” and what is “efficient”?
Research Question

• What measures can we use to evaluate network evolution in terms of effectiveness and efficiency?

• Prospective measures are expected to
  – Capture primary structural features as they are pertinent to organizational performance
  – Provide a lens on network evolution, viewing it as a process of related stages
  – Be easily implemented
Information Entropy & Mutual Information

- Shannon (1949)

- **Entropy** $H(X)$: the amount of uncertainty about a random variable ($X$), captured by a probability distribution over possible microstates.

- **Mutual information** $I(X;Y)$: change in the amount of uncertainty about the desired variable ($X$) by observing a related variable ($Y$).
Entropy & Mutual Information for networks

• Uncertainty in network structure: the degree distribution
• Node degree: the number of one-hop neighbors of the node.
• Network degree (probability) distribution
  – If the network has $N$ nodes and $N_i$ of them have degree $i$, then the probability that a node with degree $i$ is $p_i = N_i / N$. 

$p_i$
Network Entropy (NE)

• Definition:
Assume a network $X$. $NE(X) = -E[\log p(X)] = -\sum p(x_i) \log p(x_i)$, where $p(x_i) = N(x_i) / N(X)$. There are $N(X)$ nodes in $X$, among which $N(x_i)$ nodes has the degree of $i$.

• Example: a 6-node network $X$

<table>
<thead>
<tr>
<th>$i$</th>
<th>$0$</th>
<th>$1$</th>
<th>$2$</th>
<th>$3$</th>
<th>$4$</th>
<th>$5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(x_i)$</td>
<td>$0/6$</td>
<td>$3/6$</td>
<td>$2/6$</td>
<td>$1/6$</td>
<td>$0/6$</td>
<td>$0/6$</td>
</tr>
<tr>
<td>$\log_2 p(x_i)$</td>
<td>-</td>
<td>-1</td>
<td>-1.58</td>
<td>-2.58</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

$NE(X) = -\sum_{i=0}^{5} p(x_i) \log p(x_i) = 1.46$
Mutual Information ($MI$)

- **Definition**

Assume a network whose degree distribution changes from $X$ to $Y$. $MI(X;Y) = \sum_i \sum_j p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x_i)p(y_j)}$, where $p(x_i, y_j) = p(y_j \mid x_i)p(x_i)$ is the joint probability of $X$ and $Y$, when $X = x_i$ and $Y = y_j$. As previously defined, $p(x_i) = N(x_i)/N(X)$, $p(y_j) = N(y_j)/N(Y)$.
Mutual Information (cont.)

- Example: a 4-node network changes from Stage $X$ to Stage $Y$

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>0</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$j$</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>$P(x_i)$</td>
<td>2/4</td>
<td>2/4</td>
<td>2/4</td>
<td>2/4</td>
</tr>
<tr>
<td>$P(y_j)$</td>
<td>1/4</td>
<td>2/4</td>
<td>2/4</td>
<td>1/4</td>
</tr>
<tr>
<td>$p(y_j</td>
<td>x_i)$</td>
<td>1/2</td>
<td>1/2</td>
<td>1/2</td>
</tr>
<tr>
<td>$p(x_i, y_j)$</td>
<td>1/4</td>
<td>1/4</td>
<td>1/4</td>
<td>1/4</td>
</tr>
</tbody>
</table>

\[
MI(X;Y) = p(x_0, y_0) \log \frac{p(x_0, y_0)}{p(x_0)p(y_0)} + p(x_0, y_1) \log \frac{p(x_0, y_1)}{p(x_0)p(y_1)} + p(x_1, y_1) \log \frac{p(x_1, y_1)}{p(x_1)p(y_1)} + p(x_1, y_2) \log \frac{p(x_1, y_2)}{p(x_1)p(y_2)} = 0.5
\]
Measuring Network Evolution

**Example:** Adding links to a 4-node empty network until it becomes fully connected, one link at a time.

<table>
<thead>
<tr>
<th>#</th>
<th>topology</th>
<th>m</th>
<th>p(0)</th>
<th>p(1)</th>
<th>p(2)</th>
<th>p(3)</th>
<th>NE</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td></td>
<td>0</td>
<td>4/4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>1</td>
<td>2/4</td>
<td>2/4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>2</td>
<td>0</td>
<td>4/4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td>2</td>
<td>1/4</td>
<td>2/4</td>
<td>1/4</td>
<td>0</td>
<td>1.5</td>
</tr>
<tr>
<td>30</td>
<td></td>
<td>3</td>
<td>0</td>
<td>2/4</td>
<td>2/4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td>3</td>
<td>0</td>
<td>3/4</td>
<td>0</td>
<td>1/4</td>
<td>0.81</td>
</tr>
<tr>
<td>32</td>
<td></td>
<td>3</td>
<td>1/4</td>
<td>0</td>
<td>3/4</td>
<td>0</td>
<td>0.81</td>
</tr>
<tr>
<td>40</td>
<td></td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>4/4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>41</td>
<td></td>
<td>4</td>
<td>0</td>
<td>1/4</td>
<td>2/4</td>
<td>1/4</td>
<td>1.5</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>2/4</td>
<td>2/4</td>
<td>1</td>
</tr>
<tr>
<td>60</td>
<td></td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4/4</td>
<td>0</td>
</tr>
</tbody>
</table>

Graph of Network Evolution
Measuring Network Evolution

- **NE** measures the start and end network states (effectiveness).
- Small **NE** implies most nodes are similar in degree. Yet there are two probabilities:
  - **a. Centralized structure**: Most nodes connect to a few hubs and are thus separated from each other. There are relatively fewer links in the network.
  - **b. Decentralized structure**: Most nodes connect to each other. There are relatively more links in the network.
Measuring Network Evolution

- $MI$ measures the changing process (efficiency)
- Large $MI$ implies more changes in network degree distribution, which can be interpreted as
  a. Agility (bigger step to intended structure)
  b. High change cost

$MI = 0.81$

$NE = 1$

$NE = 1.5$

$NE = 0.81$

Level of centralization: $(1) < (2) < (3)$
The Best Path & the Agility of Organizational Network

• Given the same type of network evolution (e.g., link addition), a path with large sum of MI indicates an agile organizational network, which moves between centralization and decentralization in the biggest magnitude

• The best path: the longest path in terms of MI in the graph of network evolution

• Find the best path
  – Construct the graph of network evolution
  – Associate each link in the evolution graph with the opposite number of MI
  – Find the shortest path using Bellman–Ford algorithm
Example 1
Adding 6 links to a 6-node, 9-link random network
Example 2
Adding 6 links to a 6-node, 9-link scale-free network
**Example 3**
Adding 6 links to a 10-node, 33-link real-data network (data adapted from Knoke & Kuklinski, 1982)

**Example 4**
Adding 6 links to an 11-node, 32-link real-data network (data adapted from Hlebec, 1993)
Future Work: Combination of $NE$ & $MI$

Reduce $NE$ by adding links: increased average connectivity; decentralization

Reduce $NE$ by deleting links: reduced average connectivity; centralization

Number of links
Conclusions

- Two measures—NE & MI—for evaluating the dynamics of organizational networks
  - Built on network degree distribution
  - See network evolution as a process of related stages
- The evolution path with large sum of MI indicates an agile organizational network
- Together they show the relative advantage of different organizational adaptation strategies, regarding the intended topological state and the evolution path an organization should take
Thank You

For more information:

Yuan Lin  linly@u.washington.edu
Kevin C. Desouza  kdesouza@u.washington.edu
Sumit Roy  sroy@u.washington.edu