The Evolution of Social Network: Theory and Application

2006. 10.21

고길곤
Social Network Analysis

- SNA provides a strong methodological foundation for the systematic analysis of “relation” among “units” of social system

  ↠ John Almack (1922): “The influence of intelligence on the selection of associates”
  ↠ Sociology, anthropology, communication, marketing, political science, organizational studies, education, psychology, physics, mathematics, and computer science
The Mathematics of NETWORKS

Mathematics Awareness Month April 2004
Relations

- Resource exchange (value chain)
- Information transmissions
- Power relations (Organizational Chart)
- Emotional relations (friendship)
- Sexual Relationship

- “Unit” of network
  - Individuals
  - Groups, team, division, department
  - Organizations (Interorganizational network)
  - Society
  - Country (World Trade Networks)
Issues in Dynamic Networks (1)

● Mechanisms of the Network Evolution
  → Which rules govern the formation of network structure?
  → How social groups emerge from individual interactions? How superstructure is formed through group interactions?
  • Preferential attachment mechanism
  • Rational Actors Model: utility maximization approach
  • Random Markov Process
Diffusions:
how ideas, information or goods travel through a network. In almost all cases, the diffusion models rest on a network that is assumed to be constant. What are the implications for diffusion when the network structure change?
Issues in Dynamic Networks (3)

- System Complexity
  - Can we manage the network structure?
    - If we control the “leverage nodes” or “leverage edges”, can we get an “intended” network structure?
Random Graph Theory

1. Emergence of a Large Component
   \[ p = \frac{\lambda}{N}, \text{ where } \lambda \text{ is a constant.} \]
   \[ \text{If } \lambda < 1, \text{ network will have a component with small size} \]
   \[ \text{If } \lambda > 1 \text{ and satisfies } p \geq \frac{\ln(N)}{N}, \text{ we will get a giant component connecting an entire network} \]

Eg. If there are 1,000 nodes, \( p = 0.007 \) will be enough to get an connected random graph
Mechanism of Networking

- Implications
  - Theoretically, it is rare to observe disconnected nodes or components
  - So, when you do network analysis, check why some nodes are isolated
  - Measurement error? Hidden relationship? Or immature Network?
2. Scale Dependent

\[ P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k} \]

- Average degree (\(Np\)) is dependent on the size of network

![The degree distribution of ER graph (N=1000, p=0.03)](image)
2. Short Distance and Small Worlds

\[ \bar{\ell} \sim \ln N / \ln[pN] \]

- As the critical value of forming a giant component is \( p=\ln(N)/N \), if we replace the \( p \) in the above equation, we will get the maximum distance of network of a connected network.

Eg. \( N=1,000,000 \), the average distance of a network will be smaller than 5.27
Scale-free network of BA

- BA’s Preferential Attachment Rules
  1. Dependency of Relations
     -> The choice of friends or partners is not random but depends on pre-existing relations
     -> Higher Clustering Coefficients
- Ko, Lee and Park’s closeness attachment rule
  -> Sub-optimization problems
Ko, Lee, and Park’s Closeness Rules (1)

- To maintain direct ties, actors have to pay costs. BA’s preferential attachment rule underestimates such costs.
- Assumptions: Actors strategically exploit the pre-existing structure by making ties with influential actors who bridges themselves to others with short distance (Closeness rule).
Ko, Lee, and Park’s Closeness Rule (2)

- Finding 1
  - Size Dependency

![Graph showing the relationship between network size and average normalized closeness centrality. The graph includes a line for the mean, Q1(25%), and Q3(75%) values.](graph.png)
Finding 2: Weak Correlation between degree and closeness centrality
Finding 3: Sub-optimization in CPN and different structures by attachment rules

<table>
<thead>
<tr>
<th>Maximum degree</th>
<th>Average distance</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td></td>
<td>–</td>
<td>–</td>
<td>Random Network</td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td>–</td>
<td>CPN</td>
<td>–</td>
</tr>
<tr>
<td>High</td>
<td>DPN</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
Economic Efficiency Model

- Cost of Networking
  - Cost of maintaining relations
  - Cost of information processing
  - Cost of coordination

- Benefit of Networking
  - Accessibility to information, resources, etc.
  - Sharing the risk, resource, and responsibilities

- Jackson and Wolinsky (A strategic model of social and economic networks, Journal of Economic Theory, 71, 44–74 ), Doreian (Actor network utilities and network evolution, Social Networks 28, 137-164)
JW Model

- **Stability:** Does the evolution of Network lead a stable status?
  - The formation of link is a mutual decision but severance is unilateral decision

\[
\begin{align*}
  (i) & \quad \text{for all } ij \in g, \; Y_i(g, v) \geq Y_i(g - ij, v) \quad \text{and} \quad Y_j(g, v) \geq Y_j(g - ij, v) \\
  \text{and} & \\
  (ii) & \quad \text{for all } ij \notin g, \; \text{if } Y_i(g, v) < Y_i(g + ij, v) \text{ then } Y_j(g, v) > Y_j(g + ij, v). 
\end{align*}
\]

\( Y_i(g, v) \) is the payoff to player i from graph g with the value function v.

- **Efficiency:** Is Network efficient?

\[
u_i(g) = w_{ii} + \sum_{j \neq i} \delta^{tij} w_{ij} - \sum_{j: ij \in g} c_{ij},
\]
Network should not be too dense or too sparse (N=16, l=18)
Implications

- Based on the utility function, we can identify the stable network structure
  - Stable network does not necessarily mean efficient network
- Given empirical networks, we can identify which edges can be removable or addable to enhancing efficiency
- The evolutionary process depends on the sequence of networking
- We should note that “relations” are not the results of structural benefits or costs but also a function of attributes.
Simulation Approach

1. Define behavior rules of making ties
   - Structural property (e.g., Degree, closeness, information) and attributes

2. Define the entrance and exit rule of nodes

3. Simulate the initial networks following the rules defined 1 and 2

   Eg. Integrated Model of ABS (Agent-Based Simulation) and Network Theory
   - ABS shows the interaction of agents forms an efficient component rapidly without central management

- Sight range: 5
- Prob. of exchange when agents meet: 0.1
- Discount rate: 100
Implications

- Results of ABS suggest that an efficient information sharing network can be formed through flexible interaction.
- Network structures emerging from ABS are not necessarily a scale-free structure (in our simulation, log-normal distribution).
- Despite the homogenous initial conditions, flexible interactions create heterogeneous network positions and roles.
Strictly speaking, P* model (Wasserman & Pattison 1996) is not for constructing dynamic network but for detecting the processes of evolution that explain the structure of observed networks.

Configuration: A configuration is a subset of nodes and some of the edges that may be contained in a triad.

P* model enables us to build up statistical models which identify the statistically significant configurations in the observed network.
<table>
<thead>
<tr>
<th>Triad:</th>
<th>Name:</th>
<th>Positions: (ITP = Matrix of Individual Triad Positions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>003</td>
<td>$i = j = k =$ isolates. ITP$[1,1]$</td>
</tr>
<tr>
<td>j k</td>
<td>012</td>
<td>$j=$Sender, ITP$[2,1]$; $i =$ Reciver, ITP$[3,1]$; $k=$isolate ITP$[4,1]$</td>
</tr>
<tr>
<td>j k</td>
<td>102</td>
<td>$i = j =$ Dyad, ITP$[5,1]$; $k =$ isolate ITP$[6,1]$</td>
</tr>
<tr>
<td>j k</td>
<td>021D</td>
<td>$i=$Sender, ITP$[7,1]$; $j = k =$ Reciever ITP$[8,1]$</td>
</tr>
<tr>
<td>j k</td>
<td>021U</td>
<td>$j = k =$ Sender, ITP$[9,1]$; $i=$reciver ITP$[10,1]$</td>
</tr>
<tr>
<td>j k</td>
<td>021C</td>
<td>$j =$ Sender, ITP$[11,1]$; $i =$ Bridge ITP$[12,1]$; $k =$ End, ITP$[13,1]$</td>
</tr>
<tr>
<td>j k</td>
<td>111D</td>
<td>$i =$ Sender, ITP$[14,1]$; $k =$ Bridge ITP$[15,1]$; $j =$ End, ITP$[16,1]$</td>
</tr>
<tr>
<td>j k</td>
<td>111U</td>
<td>$j =$ Sender, ITP$[17,1]$; $k =$ Bridge, ITP$[18,1]$; $i =$ End, ITP$[19,1]$</td>
</tr>
<tr>
<td>j k</td>
<td>030T</td>
<td>$j =$ Sender, ITP$[20,1]$; $k =$ Bridge, ITP$[21,1]$; $i =$ End, ITP$[22,1]$</td>
</tr>
<tr>
<td>j k</td>
<td>030C</td>
<td>$i = j = k =$ Member, ITP$[23,1]$</td>
</tr>
<tr>
<td>j k</td>
<td>201</td>
<td>$i = k =$ Sender, ITP$[24,1]$; $j =$ Bridge, ITP$[25,1]$</td>
</tr>
<tr>
<td>j k</td>
<td>120D</td>
<td>$i =$ Sender, ITP$[26,1]$; $j = k =$ Reciver, ITP$[27,1]$</td>
</tr>
<tr>
<td>j k</td>
<td>120U</td>
<td>$i =$ Reciver, ITP$[28,1]$; $j = k =$ Sender, ITP$[29,1]$</td>
</tr>
<tr>
<td>j k</td>
<td>120C</td>
<td>$j =$ Sender, ITP$[30,1]$; $i =$ Bridge, ITP$[31,1]$; $k=$End, ITP$[32,1]$</td>
</tr>
<tr>
<td>j k</td>
<td>210</td>
<td>$j =$ Sender, ITP$[33,1]$; $i =$ Bridge, ITP$[34,1]$; $k=$End, ITP$[35,1]$</td>
</tr>
<tr>
<td>j k</td>
<td>300</td>
<td>$i = j = k =$ Member, ITP$[36,1]$</td>
</tr>
</tbody>
</table>
Let’s assume some configurations
→ Making ties with actors in the same group
→ Making ties with actors in the different groups
→ Mutuality in the same group
→ Mutuality in the different group
→ Transitivity
Using logistic regression, we can estimate the degree of contribution of each configuration in increasing the odds ratio of having edges.

Wasserman & Pattison (1996) model is incomplete because of the failure in controlling the multicollinearity among configurations.

However, there is much room for developing more precise statistical models which allow us to infer statistically significant configurations forming observed network structure.

\[
\log\left( \frac{P(X = 1 \mid X_{ij}^c)}{P(X = 0 \mid X_{ij}^c)} \right) = \theta' [z(X_{ij}^+) - z(X_{ij}^-)] + \epsilon
\]
Implications

- Once we estimate the coefficients of each configurations governing the evolution of network under the assumption of Markov random process, we are able to detect errors in observed networks (e.g., which edges are not less likely to be observed, while observed networks have)
Empirical Approaches

- Collect Network Data over Years
  - Authorship networks
  - Drug use
  - Policy networks
- Graphical visualization
- Statistical Analysis
Data on drug users in Colorado Springs, over 5 years
Data on drug users in Colorado Springs, over 5 years
Data on drug users in Colorado Springs, over 5 years
Data on drug users in Colorado Springs, over 5 years
Data on drug users in Colorado Springs, over 5 years
Implications

- Internally cohesive subgroups
- The connection between groups still rely on couple of edges
- The structural advantages of cutpoints and bridges last for a long time
Connectivity refers to how actors in one part of the network are connected to actors in another part of the network.

Structure: Diffusion depends on the network structure
→ Number of paths, the length of distance
→ The degree of connectivity

Timing: Diffusion depends on sequence and timing

Attribute:
→ Internally cohesive and externally closed relationship
Probability of Influence

by distance and number of paths, assume a constant $p_{ij}$ of 0.6

Path distance

<table>
<thead>
<tr>
<th>Path distance</th>
<th>10 paths</th>
<th>5 paths</th>
<th>2 paths</th>
<th>1 path</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>1.0</td>
<td>0.8</td>
<td>0.6</td>
<td>0.4</td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Probability of Influence
Integrated Network Analysis Program

- Data Collections
- Basic Statistical Analyses
- Network Data Construction
- Social Network Analysis and Statistical Inferences
- Sensitivity Analyses and Simulations
Conclusions

- Structure of social system may be stable
  - Evolution of social network theory has to explain how such a network is formed
  - The meaning of “evolution” is not always the same to “progress”
  - The debates on “adaptive” or “selective” mechanism are still on going
- Structure of social system may be unstable
  - The entrances of new actors or creations of new relations
  - If it is true, which groups or actors play critical roles in the dynamic changes of network
Other issues

- How to define the costs and benefits
  - Myopic vs. perfect rationality (do actors rely on the shortest path?)
  - Degree of uncertainty: stochastic or deterministic process
- The adequate size of networks
- The integration of structural and attribute information
- The accumulation of empirical studies on network evolution and organizational performance
- Universal law?