A Brief Introduction to Network Economics

Network Economics

Wei-Torng Juang

Institute of Economics, Academia Sinica

March 2020

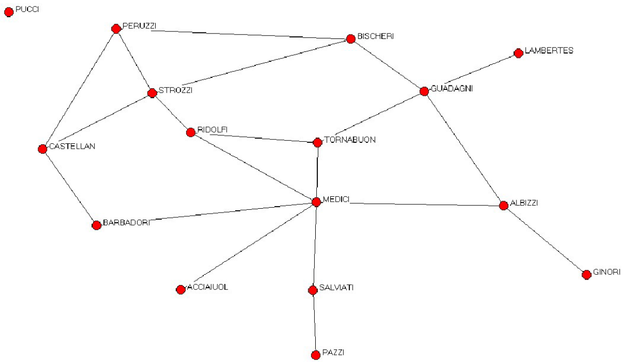
1.1. Why Model Networks?

- Transmission of information about job opportunities
- Trades of many goods and services
- Basis for the provision of mutual insurance in developing countries
- How diseases spread, which products we buy, which languages we speak, how we vote, whether we become criminals, how much education we obtain, and our likelihood of succeeding professionally
- Thus it is critical to understand
    1. How social network structures affect behavior

1.1. Why Model Networks?

- Transmission of information about job opportunities
- Trades of many goods and services
- Basis for the provision of mutual insurance in developing countries
- How diseases spread, which products we buy, which languages we speak, how we vote, whether we become criminals, how much education we obtain, and our likelihood of succeeding professionally
- Thus it is critical to understand
  1. How social network structures affect behavior
  2. Which network structures are likely to emerge in a society

- The Medici have been called the "godfathers of the Renaissance" in Florence during the 1400s.



- The Strozzi had both greater wealth and more seats in the local legislature, and yet the Medici rose to eclipse them
- The key to understanding the family's rise can be seen in the network structure:
    - How many families a given family is linked to through marriage
        - The Medici come out on top
        - But only edging out the next highest families, the Strozzi and the Guadagni, by a ratio of 3 to 2
- Need to look a bit closer at the network structure, say, the "Betweenness"

The "Betweenness"

- Let $P(i,j)$ denote the number of shortest paths connecting family i to family j
- Let $P_k(i,j)$ denote the number of these path that includes family k
- For example,
  - The shortest path between the Barbadori and the Guadagni has three links in it:

The "Betweenness"

- Let $P(i, j)$ denote the number of shortest paths connecting family i to family j
- Let $P_k(i, j)$ denote the number of these path that includes family k
- For example,
  - The shortest path between the Barbadori and the Guadagni has three links in it:
    - Barbadori-Medici-Albizzi-Guadagni

The "Betweenness"

- Let $P(i,j)$ denote the number of shortest paths connecting family i to family j
- Let $P_k(i,j)$ denote the number of these path that includes family k
- For example,
  - The shortest path between the Barbadori and the Guadagni has three links in it:
    - Barbadori-Medici-Albizzi-Guadagni
    - Barbadori-Medici-Tornabuon-Guadagni

The "Betweenness"

- Let $P(i,j)$ denote the number of shortest paths connecting family i to family j
- Let $P_k(i,j)$ denote the number of these path that includes family k
- For example,
  - The shortest path between the Barbadori and the Guadagni has three links in it:
    - Barbadori-Medici-Albizzi-Guadagni
    - Barbadori-Medici-Tornabuon-Guadagni
  - Let $i =$ Barbadori and $j =$ Guadagni, then $P(i,j) = 2$

1. 2. A Set of Examples

1.2.1. Florentine Marriages

The "Betweenness"

- Let $P(i,j)$ denote the number of shortest paths connecting family i to family j
- Let $P_k(i,j)$ denote the number of these path that includes family k
- For example,
  - The shortest path between the Barbadori and the Guadagni has three links in it:
    - Barbadori-Medici-Albizzi-Guadagni
    - Barbadori-Medici-Tornabuon-Guadagni
  - Let $i =$ Barbadori and $j =$ Guadagni, then $P(i,j) = 2$
    - Set $k =$ Medici, then $P_k(i,j) = 2$

The "Betweenness"

- Let $P(i,j)$ denote the number of shortest paths connecting family i to family j
- Let $P_k(i,j)$ denote the number of these path that includes family k
- For example,
  - The shortest path between the Barbadori and the Guadagni has three links in it:
    - Barbadori-Medici-Albizzi-Guadagni
    - Barbadori-Medici-Tornabuon-Guadagni
  - Let i = Barbadori and j = Guadagni, then $P(i,j) = 2$
    - Set k = Medici, then $P_k(i,j) = 2$
    - Set k = Strozzi, then $P_k(i,j) = 0$

The "Betweenness"

- Let $P(i,j)$ denote the number of shortest paths connecting family i to family j
- Let $P_k(i,j)$ denote the number of these path that includes family k
- For example,
  - The shortest path between the Barbadori and the Guadagni has three links in it:
    - Barbadori-Medici-Albizzi-Guadagni
    - Barbadori-Medici-Tornabuon-Guadagni
  - Let $i =$ Barbadori and $j =$ Guadagni, then $P(i,j) = 2$
    - Set $k =$ Medici, then $P_k(i,j) = 2$
    - Set $k =$ Strozzi, then $P_k(i,j) = 0$
    - Set $k =$ Albizzi, then $P_k(i,j) = 1$

The "Betweenness"

- Let $P(i,j)$ denote the number of shortest paths connecting family i to family j
- Let $P_k(i,j)$ denote the number of these path that includes family k
- For example,
  - The shortest path between the Barbadori and the Guadagni has three links in it:
    - Barbadori-Medici-Albizzi-Guadagni
    - Barbadori-Medici-Tornabuon-Guadagni
  - Let $i =$ Barbadori and $j =$ Guadagni, then $P(i,j) = 2$
    - Set $k =$ Medici, then $P_k(i,j) = 2$
    - Set $k =$ Strozzi, then $P_k(i,j) = 0$
    - Set $k =$ Albizzi, then $P_k(i,j) = 1$
- The Medici is the key family in connecting the Barbadori and the Guadagni

- To gain intuition about how central a family is, we can ask, for each pair of families, on what fraction of the total number of shortest paths between the two the given family lies
- This number is 1 for the fraction of the shorstest paths the Medici lie between the Barbadori and the Guadagni, and 1/2 for the corresponding fraction that the Albizzi lie on.
- Averaging across all pairs of other families gives a betweenness or power measure for a given family
- In particular, we can calculate

$$\sum_{i,j:i\neq j,k\notin\{i,j\}} \frac{P_k(i,j)/P(i,j)}{(n-1)(n-2)/2} \tag{1.1}$$

for each family k, where $P_k(i,j)/P(i,j) = 0$ if there is no paths connecting i and j

- This measure of betweenness for the Medici is 0.522
- The second-highest family is the Guadagni with a betweenness of 0.255
- A similar calculation for the Strozzi yields 0.103
- To the extent that marriage relationships were keys to communicating information, brokering business deals, and reaching political decisions, the Medici were much better positioned than other families
- Network structure can provide important insights beyond those found in other political and economic characteristics
- The network structure is important for more than a simple count of how many social ties each member has: different measures of betweenness or centrality will capture different aspects of network structure

1. 2. A Set of Examples

1.2.2. Roamnces among High School Students

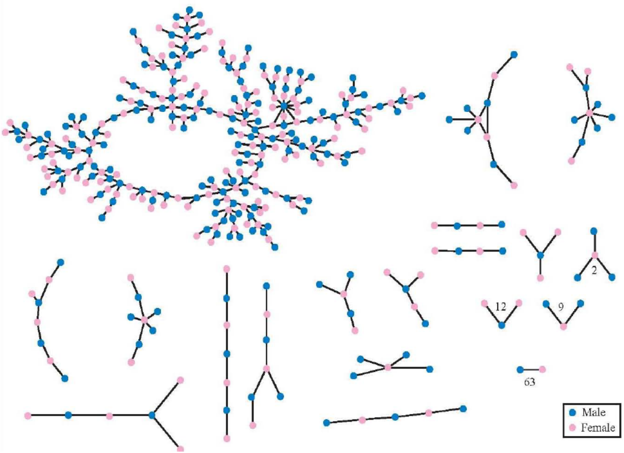


Figure 1.2: A Figure from Bearman, Moody and Stovel [47] based the Add Health Data

- A bipartite network: the nodes can be divided into two groups, male and female, so that links only lie between groups (with a few exceptions)
- The distribution of the degrees of the nodes (number of links each node has) turns out to closely match a network in which links are formed uniformly at random
- A giant component consisting of more than 100 students, with the next largest one only 10 students in it
- This component structure has important implications for the diffuse of disease, information, and behaviors
- The network is quite treelike: there are few loops or cycles in it (the giant component)
- The treelike structure contrasts with the denser friendship network next: many cycles and a shorter distance between nodes

1.2.2. Friendships among High School Students

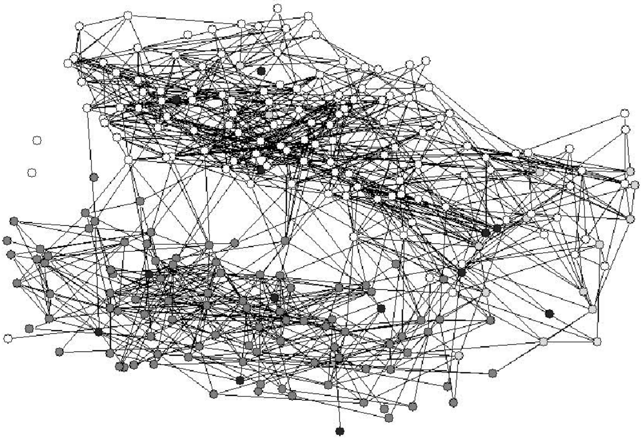


Figure 1.3: "Add Health" Friendships among High School Students Coded by Race: Hispanic=Black, White=White, Black=Grey, Asian and Other = Light Grey.

- Homophily: a bias toward similar individuals (the race)
    - 52 percent of the students are white and yet 85 percent of white students' friendships are with other whites
    - 38 percent of the students are black and yet 85 percent of black students' friendships are with other blacks
    - Hispanics are more integrated, comprising 5 percent of the population but having only 2 percent of their friendships with other Hispanics
- The students are somewhat segregated by race, which affects the spread of information, learning and the speed with which things propagate through the network

Suppose a completely random process is responsible for the formation of links in a network.

- Consider a set of nodes $N = \{1, ..., n\}$.
- Let a link between any two nodes, i and j, be formed with probability $p, 0 < p < 1$.
- Any given network that has m links on n nodes has a probability of

$$p^m (1-p)^{\frac{n(n-1)}{2}-m} \tag{1.2}$$

  of forming under this process.
- The degree of a node is the number of links that the node has.
- The probability that any given node i has exactly d links is

$$\binom{n-1}{d} p^d (1-p)^{n-1-d} \tag{1.3}$$

- For large n and small p, this binomial expression is approximated by a Poisson distribution, so that the fraction of nodes that have d links is approximately

$$\frac{e^{-(n-1)p}((n-1)p)^d}{d!}. \tag{1.4}$$

- The class of random graphs is often referred to as the class of Poisson random networks.
- As an example, let $n = 50$ and $p = 0.02$.
- We should expect about 37.5% of the nodes to be isolated, which is roughly 18 or 19 nodes.
- There happen to be 19 isolated nodes in the network in Figure 1.4.

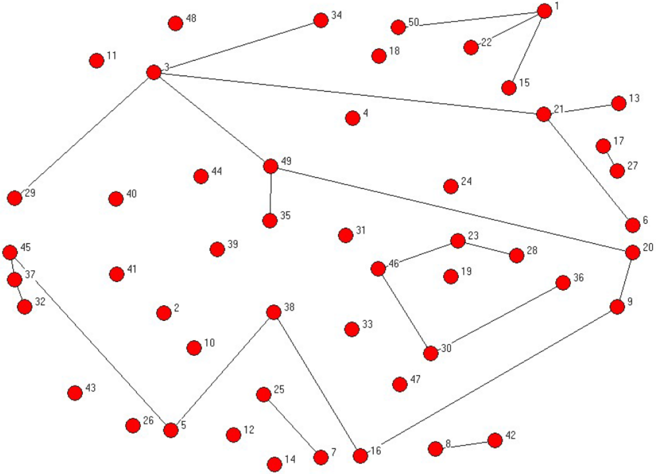


Figure 1.4: A Randomly Generated Network with Probability .02 on each Link

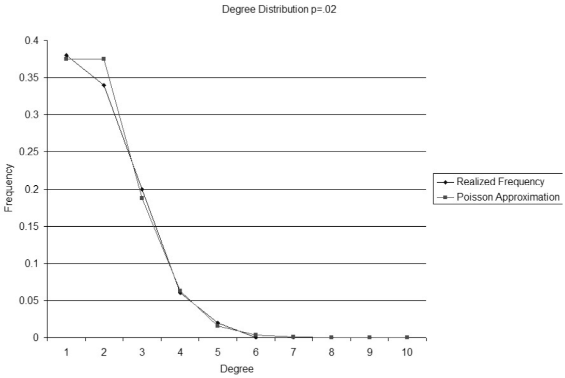


Figure 1.5: Frequency Distribution of a Randomly Generated Network and the Poisson

- When $p = \log(50)/50 = 0.078$, this is roughly the threshold at which isolated nodes should disappear.
- Based on (1.4), with $n = 50$ and $p = 0.08$, we should expect about 2% of the nodes to be isolated (with degree 0), or rough 1 node out of 50.
- This is exactly what occurs in the realized network in Figure 1.6 (again, by chance).

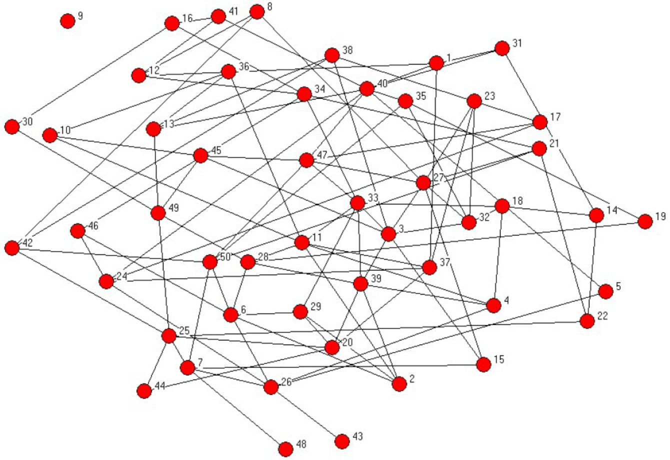


Figure 1.6: A Randomly Generated Network with Probability .08 of each Link

1. 2. A Set of Examples

1.2.3. Random Graphs and Networks

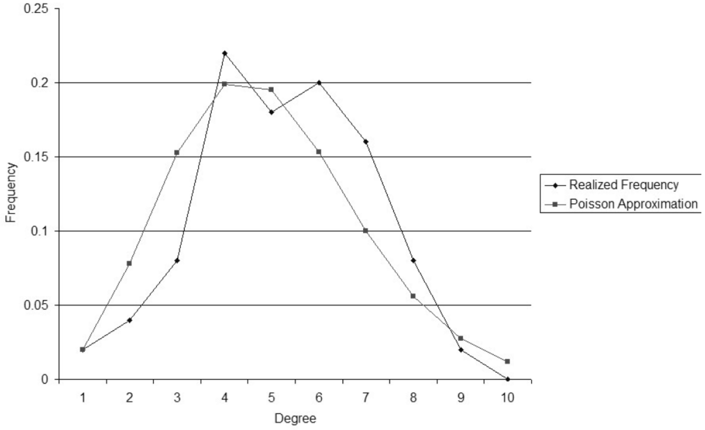


Figure 1.7: Frequency Distribution of a Randomly Generated Network and the Poisson Approximation for a Probability of .08 on each Link

- Given n, p, what is the fraction of nodes that are completely isolated (degree $d = 0$)?
- By (1.4), this number is approximated by $e^{-(n-1)p}$ for large networks, provided$(n-1)p$ is not too large.
- Examine the threshold at which this fraction is just such that we expect to have one isolated node on average, or

$$e^{-(n-1)p} = \frac{1}{n}.$$

- This is equivalent to

$$p(n-1) = \log(n),$$

where the average degree is $p(n-1)$.

- If the average degree is substantially greater than $\log(n)$, then the probability of having any isolated nodes tends to 0.
- While if the average degree is substantially less then $\log(n)$, then the probability of having at least some isolated nodes tends to 1.

- Contrast with random network-formation model, in the Florentine marriage network marriages were carefully arranged.
- The next example addresses:
  - (1) Which networks would maximize the welfare of a society and
  - (2) Which network will arise if the players have discretion in choosing their links.

- Jackson and Wolinsky (1996):
  - Links represent social relationships, for instance friedships, between players.
  - These relationships offer benefits in terms of favors, information, and the like, and also involves some costs.
  - Players also benefits from indirect relationships.
  - The benefit deteriorates with the distance of the relationship by a factor δ between 0 and 1.
  - Players only pay costs for maintaining their direct relationships.

Given a network g, the net utility $u_i(g)$ that player i receives from g is the sum of benefits from his direct and indirect connections to other players less the cost of maintaining his direct links:

$$u_i(g) = \sum_{j \neq i : ij \in g} \delta^{\ell_{ij}(g)} - d_i(g)c,$$

where $\ell_{ij}(g)$ is the number of links in the shortest path between i and j, $d_i(g)$ is the number of links that i has (i's degree), and c is the cost for maintaining a link.

Figure 1.8: The utilities to the players in a three-link four-player network in the symmetric connections model.

Definitions

A network is efficient if it maximizes the total utility to all players in the society. That is, g is efficient if it maximizes $\sum_i u_i(g)$.

- If $c < \delta - \delta^2$: adding a link between any two agents always increases total welfare.

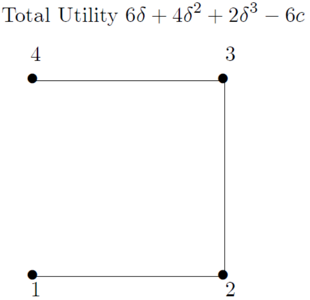


Figure 1.9: The Gain in Total Utility from Changing a “Line” into a “Star”.

- A simple characterization of the set of efficient networks:
  - (1) costs are so low that it makes sense to add all links;
  - (2) costs are so high that no links make sense; or
  - (3) cost are in a middle range, and the unique efficient architecture is a star network.

Definition

Pairwise stability: (1) no agent can raise his payoff by deleting a link that he is directly involved in; and (2) no two agent can both benefit (at least one strictly) by adding a link between themselves.

- If $c < \delta - \delta^2$, then the complete network is both pairwise stable and efficient.
- If $\delta > c > \delta - \delta^2$, then a star network will be both pairwise stable and efficient.
- If $c > \delta$, then the efficient (star) network is not pairwise stable.