Structure Formation of Social Network

DU Nan[1], FENG Hui[2], HUANG Zigang[3],
Sally MAKI[4], WANG Ru(Ruby)[5], and ZHAO Hongxia (Melissa)[6]

[1]Beijing University of Posts and Telecommunications, China
[2]Fudan University, China
[3]Lanzhou University, China
[4]University of California, Berkeley, USA
[5]Central China Normal University, China
[6]Institute of Automation, Chinese Academy of Sciences, China

Abstract: We consider how individual's behaviors affect structure formation of social network. A dynamic model is designed, in which nodes move and build interpersonal relationship with other nodes according to its personality factor. Agent-Based modeling shows evolutional structure with small-world and power-law properties.
Keywords: Social network; complex networks; Agent-Based modeling

1 Introduction

Social network is the backbone of social relationship, which is an attracting interdisciplinary research field within social science, statistic and mathematics [1]. The structure of social networks has distinct properties from random networks and regular networks. The well-known six-degree experiment done by Milgram in 1969 [2] revealed the astonishing short average distance compare with the large scale of social network, which depicts the small-world network [3], as well as large clustering coefficient[1]. Besides small-world property, many natural complex networks present scale-free distribution [4], e.g. social network, biological network, WWW, in which the degree distribution of network follows power-law. It is also different from the single-scale distribution in random networks and regular networks.

Although several computational models could generate small-world or scale-free networks [3]-[5], these models do not consider detailed backgrounds of different networks owning similar properties. For instance, statistical methods have given some common characteristic of social network, while what are the reasons behind the behind those characteristic. [6]-[8] use game theory to evaluate the influence on network structure by node pair's game. In [9], links are built imitating social communication rules within node pair selected randomly.

However, random paring methods do not consider the continuous space-time evolution procedure in real social scene. The differences of nodes are not considered, either. We designed an evolutionary model to evaluate the influence of individual's behaviors on formation of social network structure, in which personality factor is key point on agents' moving and intercommunication.

1 Large clustering coefficient means one nodes' neighbors' are more like to connect with each other.

2 Evolution model of social network

2.1 The weighted graph model of social network

Weighted graph $G = (N, L, W)$ is proper to describe interpersonal relationship with a view of different social intimacy among peoples, in which N is the set of nodes, L is the set of links and W is the set of weights on links. We use $N \times N$ adjacent matrix to represent the link relationship between nodes, in which $a_{ij}=1$ if there's a link present between node i and j, $a_{ij}=0$ if node i and j are not adjacent. There is no self-circle in our model, in other words, all diagonal elements of A are zero. Similarly, we use $N \times N$ weights matrix W for weighted graph, in which w_{ij} is the weight of the link connecting from node i to node j. $w_{ij} = 0$ if node i and j are not connected, and a positive integer is assigned to w_{ij} if node i and j are adjacent.

Each node has three properties: personality, step-size and energy. The personality is a real number in [0, 1], which represents how extroverted (small personality factor) or introverted (large personality factor) one acts in social communication. Extroverted nodes prefer to making new friends, while introverted nodes would like to stay with old friends. Nodes move inside an area with periodic boundary and would meet other nodes[2]. New link a_{ij} will be built if node i and j meet each other and no link before them. The initial weight is set to 1. w_{ij} will be added by 1[3] at each time when node i and j meet each other. We give energy factor E_i to node i, which will be consumed by $cost_{ij} = cost_{ji} = p_i \cdot p_j$ when new link between node i and j is built. Each node's neighbors will not be added if residual energy cannot afford new links, then only weights update all the time. The cost of new link depends on mutual personality factor. Extroverted people have more friends and less cost of building new link, which follows social common sense.

2.2 Dynamic evolution rules

Discrete model is suitable for computer simulation. Nodes are distributed randomly in some area and move around according to the state of last epoch. The moving method depends on angle and step size.

The moving angle of node i points to targeted position (x_i^{new}, y_i^{new}) consisting of deterministic part and random part[4] as Equ.(1). The deterministic part is the weighted average position of nearest neighbors, which stands for the influence from friends. The random part, x_{random} and y_{random}, are the random position inside the area, which expresses the desire for exploring new social relationship.

$$x_i^{new} = \left[p_i \cdot \bar{x} + (1 - p_i) \cdot x_{random}\right]$$
$$y_i^{new} = \left[p_i \cdot \bar{y} + (1 - p_i) \cdot y_{random}\right], \qquad (1)$$

where $\bar{x} = \sum_{j \in N_i} w_{ij} x_j \Big/ \sum_{j \in N_i} w_{ij}$, $\bar{y} = \sum_{j \in N_i} w_{ij} y_j \Big/ \sum_{j \in N_i} w_{ij}$, and N_i is the set of node i's nearest neighbors. The more extroverted a node is, the more randomly it moves and it would like to make new friends.

The step-size s_i of node i also depends on its personality factor p_i, extroverted nodes behave

[2] Two or more nodes meet each other if they locate the same grid in simulation zone.
[3] Similar count method as in [10]。
[4] Similar design as linear Vicsek model [11].

actively in social communication and take more aggressive move, as

$$step_i = (1 - p_i) \cdot step_{\max} , \quad (2)$$

in which $step_{max}$ is the preset maximum step size.

3 Simulation of social network formation

We didn't give rigorous dynamics of the iterative model in Section 2.2, in that it is too difficult to deduce in mathematical way. Instead, we use Agent-Based simulation method invented in 1940s, which did not become a conventional method in research on social network, distributed processing, macro-economics etc. until 1990s subject to computation capability. Each agent in Agent-Based simulation follows certain rules and move according to local information. By lots of iteration, some statistical result may emerge. In this work, we use NetLogo[5], a widely used Agent-Based tool, to simulate our model, which can provide friendly visualization. Figure 1 shows the initial state of once simulation.

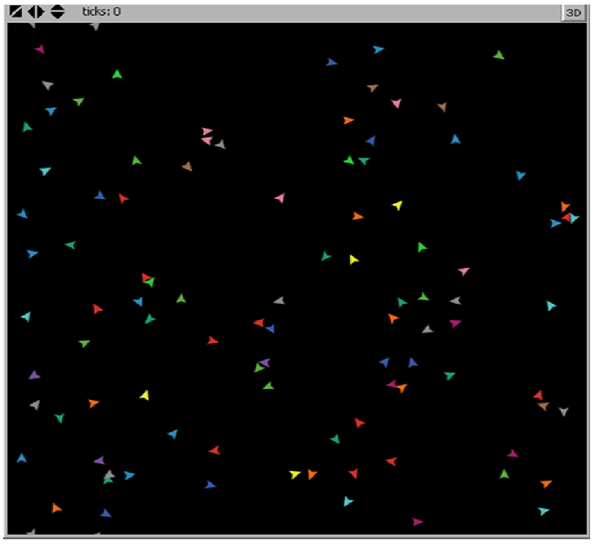


Figure 1 Initial state of once simulation of 100 nodes for E_i=1 and $Step_{max}$=5 in 30*30 periodic boundary[6] rectangle.

The personality factor p_i has great effects on network evolution in our model. To illustrate an intuitive way, we present two simulation snapshots as Figure 2, where all p_i is fixed to ones and zeros. In all ones case, all the nodes prefer to stay with their nearest neighbors and don't like to make new friends, so we can observe many small groups in the snapshot. The all zeros case gives the opposite phenomenon, where each node has random links with others.

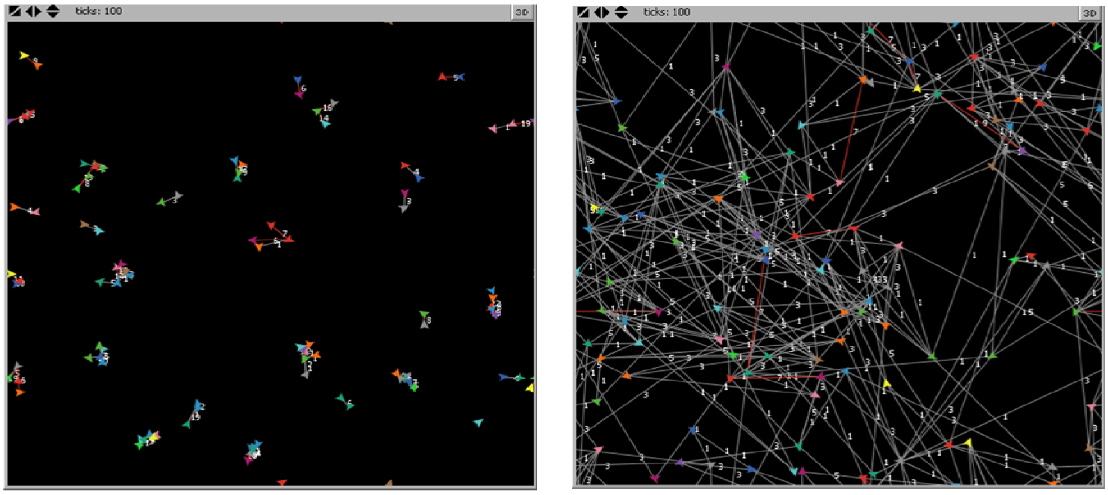
Since extreme cases never happen in real world, we use limited Gaussian distribution[7] instead in later simulation. Figure 2 gives once simulation snapshot with Gaussian personality factor, which presents topology status between two extreme cases.

---

[5] http://ccl.northwestern.edu/netlogo/

[6] When an object passes through one boundary of the area, it reappears on the opposite boundary with the same velocity. Periodic boundary conditions are particularly useful for simulating a part of a bulk system with no surfaces present.

[7] This distribution is derived from a small survey over about 40 students in CSSS2008. Not strictly, of course.



(a) (b)

Figure 2 Evolution snapshots with fixed personality factor, the other parameters are same as in Figure 1. (a) is the snapshot in T=100 with all p_i=1; (b) is the snapshot in T=100 with all p_i=0.

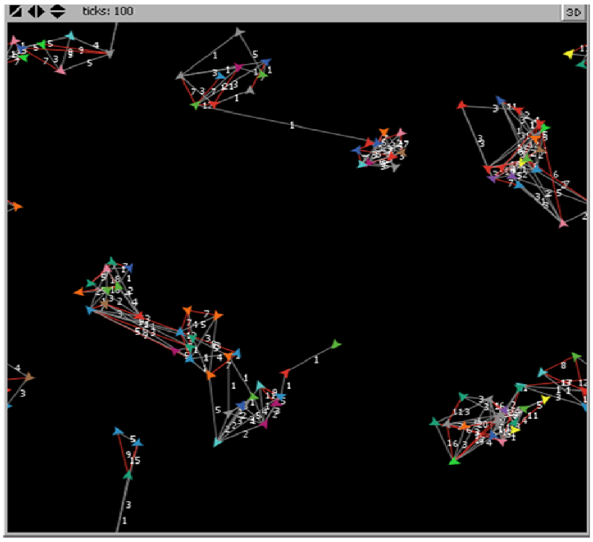


Figure 3 Simulation snapshot with limited Gaussian personality between 0.4 and 1, the other parameters are same as in Figure 1.

4 Results and analysis

4.1 Evaluation metrics for weighted graph

We use special metrics [13][14] for weighted graph to evaluate the network structure, as well as some metrics [15] for non-weighted graph if necessary.

1) weighted clustering coefficient

The weighted clustering coefficient of node i is defined as

$$c_i = \frac{1}{s_i(k_i - 1)} \sum_{j,m} \frac{(w_{ij} + w_{im})}{2} a_{ij} a_{jm} a_{mi} . \tag{3}$$

The weighted clustering coefficient of whole graph is the average of all the nodes.

$$C = \frac{1}{N}\sum_i c_i \qquad (4)$$

2) weighted distance

First, we define the length l_{ij} between node i and j as the reciprocal of the weight, $l_{ij}=1/w_{ij}$. It is then possible to define the weighted distance d_{ij} as the smallest sum of the edge lengths throughout all the possible paths in the graph from node i to j. The distance in weighted graph can be calculated using well-known Dijkstra algorithm.

3) Degree distribution

The degree k_i of a node i is the number of edges incident with the node, and is defined in terms of the adjacency matrix A as

$$k_i = \sum_{j \in N_i} a_{ij} , \qquad (5)$$

where N_i is the set of the nearest neighbors of node i. The degree distribution $P(k)$ is defined as the fraction of nodes in the graph having degree k.

4) Strength Distribution

In weighted graph, the strength of node i is the sum of weights between node i and its nearest neighbors, defined as

$$s_i = \sum_{j \in N_i} w_{ij} . \qquad (6)$$

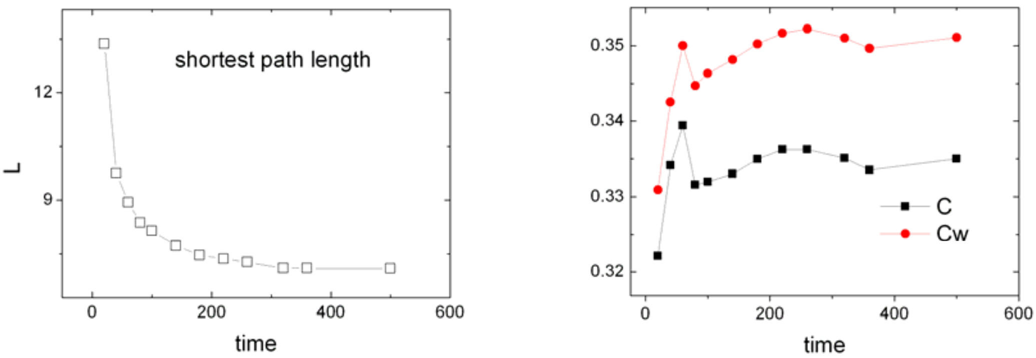
The degree distribution $P(s)$ is defined as the fraction of nodes in the graph having strength s.

5) Weight Distribution

The degree distribution $P(w)$ is defined as the fraction of nodes in the graph having strength w.

4.2 Analysis

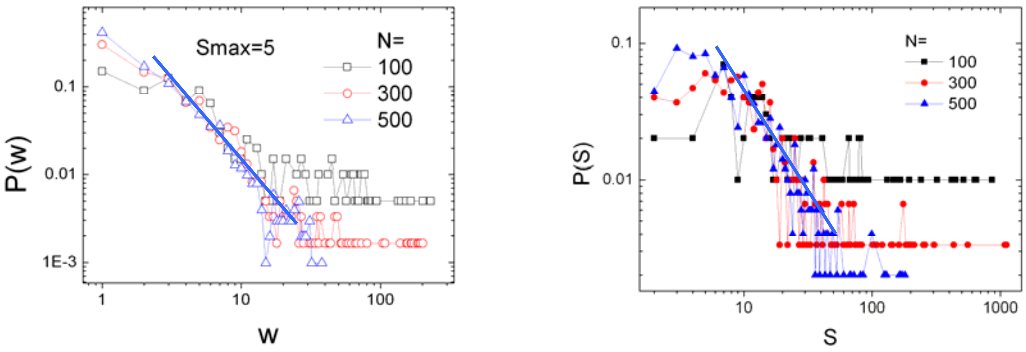
Figure 4 shows the average distance and clustering coefficient as functions of simulation epochs. When topology become steady, we observe typical small-world properties with short distance and large coefficient.



(a) (b)

Figure 4 Small-world properties of 500 nodes simulation results for E_i=0.5, the other parameters are same as in Figure 1. (a) Average distance decreases with simulation epochs. The steady shortest path length is rather small. (b) The clustering coefficient does not change a lot with simulation epochs and remains a rather large number.

Figure 5 shows the power-law distribution of weights and strength, which becomes more clearly as increasing number of nodes.



(a) (b)

Figure 5 Scale-free distribution of (a) weights and (b) strength when topology becomes steady, where the blue line is the fitting power law. All the parameters are same as in Figure 4.

5 Conclusion

Our work intends to simulate the influence of interpersonal relationship on social network structure. We observe large clustering coefficient, short average distance, and power-law distributed weights and strength, which presents certain regular hierarchy and community structure coinciding with the real world.

However, this model is a toy rather than real research. Some assumption is too arbitrary to convincing, such as the energy limitation, maximum step size and so on. And we did not implement large-scale simulation due to the capability of simulation tool, so the result is not sufficient in statistical way. Also we did not have enough time to validate the community structure and modularity [16] observed on simulation, nor the join-and-quit mechanism.

6 Acknowledgements

This work was done in the Complex System Summer School (CSSS) held by Santa Fe Institute and Institute of Theoretical Physics, Chinese Academy of Science in 2008. We thank Aaron Clauset for his project advice. We also thank the organizers and sponsors of this summer school.

References:

[1] J. Scott, Social network analysis: A handbook, Sage, 2000.

[2] Travers, Jeffrey and Stanley Milgram. "An Experimental Study of the Small World Problem." Sociometry, Vol. 32, No. 4, 1969, pp. 425-443.

[3] D.J. Watts and S.H. Strogatz, “Collective dynamics of `small-world' networks,” Nature, vol. 393, Jun. 1998, p. 440.

[4] A. Barabási and R. Albert, “Emergence of Scaling in Random Networks,” Science, vol. 286, 1999, pp. 509-512.

[5] M.E.J. Newman and D.J. Watts, “Renormalization group analysis of the small-world network

model," Physics Letters A, vol. 263, 1999, pp. 341-346.

[6] B. Skyrms and R. Pemantle, "A dynamic model of social network formation," Proceedings of the National Academy of Sciences, vol. 97, 2000, pp. 9340-9346..

[7] M.O. Jackson, "A survey of models of network formation: Stability and efficiency," Group Formation in Economics: Networks, Clubs and Coalitions, 2003.

[8] H. Ohtsuki, C. Hauert, E. Lieberman, and M.A. Nowak, "A simple rule for the evolution of cooperation on graphs and social networks," Nature, vol. 441, 2006, pp. 502-505.

[9] E.M. Jin, M. Girvan, and M.E.J. Newman, "Structure of growing social networks," Physical Review E, vol. 64, Oct. 2001, p. 46132.

[10] M. Li, D. Wang, Y. Fan, Z. Di, and J. Wu, "Modelling weighted networks using connection count," New Journal of Physics, vol. 8, 2006, pp. 355–364.

[11] A. Jadbabaie, J. Lin, and A.S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," IEEE Transactions on Automatic Control, 2003.

[12] E. Bonabeau, "Agent-based modeling: Methods and techniques for simulating human systems," Proceedings of the National Academy of Sciences, 2002.

[13] M. Barthélemy *et al.*, "Characterization and modeling of weighted networks," Physica A, vol. 346, Feb. 2005, pp. 34-43.

[14] A. Barrat *et al.*, "The architecture of complex weighted networks," Proceedings of the National Academy of Sciences, vol. 101, 2004, pp. 3747-3752.

[15] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. Hwang, "Complex networks: Structure and dynamics," Physics Reports, vol. 424, Feb. 2006, pp. 175-308.

[16] M.E.J. Newman, "Modularity and community structure in networks," Proceedings of the National Academy of Sciences, vol. 103, 2006, pp. 8577-8582.