

The Model of Microblogging Network with Anti-Cognitive Costed Deletion of Edges

Quan Liu *

Institute of System Engineering
Dalian University of Technology
Dalian, China
e-mail: quanshuilq2005@163.com

* Corresponding Author

Lili Rong

Institute of System Engineering
Dalian University of Technology
Dalian, China
e-mail: llrong@dlut.edu.cn

Abstract—With more and more people beginning to share information and opinions through microblogging, microblogging network has become a new research focus in the social network. In this paper, we introduced a model which gives more realistic and detailed description of the evolution process of microblogging network based on the Barabási-Albert model. The model includes the addition of new vertices with directed edges, the addition of new directed edge between old vertices selected by cognitive cost, the deletion of some old directed edges between old vertices selected by anti-cognitive cost and anti-preferential probability. We analytically calculated and simulated the degree distribution by the mean field theory. The degree distribution exhibits a multiscaling behavior analogous to the BA model. Finally, our results show that strong correlation between the in-degree and out-degree. The analytical results are supported by simulations. Our work is the basis for the study of information dissemination in the microblogging network.

Keywords—Microblogging Network; Evolution Model; Delete Edges; Anti-cognitive Cost ; Power-laws

I. INTRODUCTION

With the continuous development of Internet and Web technology, microblogging is becoming a basic tool for people's daily communication and entertainment. As a new information dissemination platform, microblogging has broad prospects in application and research, particularly in journalism, public opinion propagation, information recommendation, user behavior analysis and human dynamics, etc. As more and more researchers pay attention to microblogging, microblogging network has become a new research focus in social network.

The present research on microblogging network structure focuses on empirical analysis through various microblogging network data. Kwak et al^[1] found some characteristics of followers through analyzing the users' data of Twitter, such as non-power-law distribution, short effective diameter and low reciprocates edges, which is obviously departed from the human social networks. Java et al^[2] studied the topological and geographical properties of Twitter's social network. Huberman et al^[3] explored potential friend's relationship in Twitter, and discovered that the Twitter network is composed of mutual-attention network with high density and sparse network of real friends.

Currently, a variety of network evolution models were proposed based on the extracted features in the actual

network using complex network theory. Dorogovtsev and Mendes^[4] proposed a developing and decaying model of the undirected network through adding or removing the edge between old nodes. Deng et al^{[5]-[6]} studied the influence on scaling behavior of growing network by deleting nodes. Albert and Barabási^[7] studied the impact of adding edges and reconnecting the edges between old nodes. Bianconi and Barabási^[8] proposed a fitness model that considered the ability of node to compete for edges. Dorogovtsev and Mendes^[9] studied the growing networks that considered the age of nodes and analyzed the impact that age has on the scaling behavior. Chen et al^[10] proposed two BA extended model, one of which deleted old edges using anti-preferential probability. Literature [11] analyzed the effects that degree distribution and degree correlation of nodes have on the reciprocal sides. Literature [12] put forward a network model that add new nodes and edges, reconnect and delete edges between the old nodes, and analyzed the relation between deleting size and power law degree distribution. Some research proposed scale-free homophilic network models that combined preferential attachment and homophilic term^{[13]-[15]}. Other multiplex network models^{[16]-[18]} were presented..

The above evolution models all have their own concern, but they just cannot be directly applied to microblogging network somehow. This paper start from the perspective of complex networks, summarize the characteristics of microblogging network that were found in the existing empirical studies, put forward a evolution model of microblogging network, through theoretical analysis and simulation, we find our model is in good accordance with the actual network.

II. EXISTING MODEL

Paper [19] proposed a model of microblogging network that considered the cognitive costs of vertexes and reciprocal edges.

In Twitter, posted messages from one's "following" friends appear on his/her "timeline", which keeps updating every second. Consequently, the more the number of following users increases, the faster the timeline flows. Therefore, large amounts of following users have difficulties in following up friends' posts. In other words, the cognitive cost increases with the out-degree.

The functional form of the cognitive cost should be,

$$C(q) \propto q^{\alpha} \quad (1)$$

where q is the out-degree and α is a constant.

The algorithm of the model is as follows:

(1) In each step, a new vertex i is added to the network with probability p , then the vertex i creates a new edge to an old vertex j with probability $\Pi(k_j)$. The vertex j reciprocates the vertex i with probability ρ_j , where the constant ρ_j is determined with each vertex.

(2) On the other hand, with probability $1-p$, an old vertex i is chosen with probability Ψ_i , then the vertex i creates an edge to another old vertex j in accordance with the preferential attachment rule and the vertex j reciprocates in the same manner. Let Ψ_i depend on the cognitive cost $C(q_i)$. It is simply assumed that $\Psi_i \propto C(q_i)^{-\beta}$ where $\beta > 0$, since vertices with smaller cost should create more out-going edges. Therefore where $K = \alpha\beta > 0$.

$k_i(t)$ and $q_i(t)$ satisfies the following dynamical equation:

$$\frac{\partial k_i}{\partial t} = \Pi(k_i) + (1-p)\langle \rho \rangle \Psi(q_i) - (1-p)\Pi(k_i)\Psi(q_i) \quad (2)$$

$$\frac{\partial q_i}{\partial t} = \langle \rho \rangle \Pi(k_i) + (1-p)\Psi(q_i) - (1-p)\langle \rho \rangle \Pi(k_i)\Psi(q_i) \quad (3)$$

It gets $p(k) \propto k^{-(2+\langle \rho \rangle)}$ and $p(q) \propto q^{-(2+\langle \rho \rangle)}$ in the large-degree range.

III. AN IMPROVED MODEL OF MICROBLOGGING NETWORK

The model of microblogging network in paper [19] ignored an important characteristic of the actual microblogging network evolution, which is the cancellation of the followers' relationship. Users cannot concern over huge broadcast information when they follow an excessive numbers of other people and will choose to cancel the following of some users with their limited energy. Consider this feature of microblogging network, an improved model that considering the operation of deleting edges is proposed.

The evolution process of microblogging network can be divided into following aspects. New users join in microblogging network through registration of new accounts. Then new users choose to follow their friends' and favorite stars' microblogging to obtain their latest news and opinions. These friends are followed, also in return they may follow these new users, and it is a kind of reciprocal follow. According to the statistical result^[1], the rate of two users reciprocal follow each other reached 22.1%. Old users will follow other old users when they used microblogging. Users need to spend a lot of time to find the interesting information when followed too many people, a mass of information is pushed in the face of them. However, human cognitive ability and time spent on the microblogging is limited every day, hence some user will choose to delete some following people, those people who follow more users are more likely to delete.

The process of constructing the model for microblogging network is as follows. Microblogging network is a directed network. An individual B follows user A , we call individual B as a follower. We construct a directed graph $G(V, E)$, where V represents the set of microblogging users and E represents the set of following relations of users. In order to easily understand the spreading of information, microblogging network is defined as follows. If user B is a follower of user A , a directed edge e is created from user B to user A . If user A

and user B reciprocated follows each other, a bidirectional edge is created.

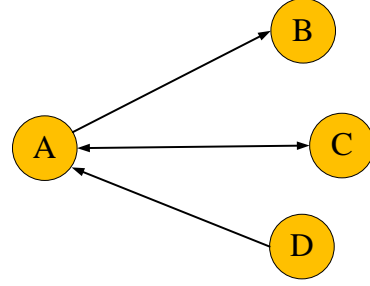


Figure 1. Follow relations of users. User A is a follower of user B , user D is a follower of user A , user A and user C reciprocated follow.

Starting with m_0 nodes and e_0 edges, and in each step we perform the following three operations:

(1) a new vertex i is added to the network with probability π_1 , m new edges are added to the network from the new vertex i to m different old nodes j . The vertex j creates a return edge to the vertex i with probability ρ . Where ρ is a constant, $\rho \in [0, 1]$. Old vertex j is selected as an end of a new edge with probability $\Pi(k_j)$.

$$\Pi(k_j) = \frac{k_j}{\sum k_j} \quad (4)$$

(2) n new edges between old vertices are added to the network with probability π_2 . A vertex i is selected as a start of a new edge with the cognitive cost $\Phi(q_i)$, a vertex j is selected as an end of a new edge with the preferential probability $\Pi(k_j)$.

$$\Phi(q_i) = \frac{q_i^{-K}}{\sum_j q_j^{-K}} \quad (5)$$

(3) h old edges are deleted in the network with probability π_3 . Where $\pi_3 = 1 - \pi_1 - \pi_2$. A vertex i is selected as a start of a deleted edge with the anti-cognitive cost $\Phi^*(q_i)$. A vertex j is selected as an end of a deleted edge with the anti-preferential probability $\Pi^*(k_j)$. Where $N(t)$ is the number of vertex. $\sum_i \Phi^*(q_i) = 1$, $\sum_i \Pi^*(k_i) = 1$. We will assume that $\pi_3 < 1/2$, $h < m$, $h < n$ so that the number of edges in the graph is indeed growing (on average).

$$\Pi^*(k_j) = \frac{1}{N(t)-1} (1 - \Pi(k_j)) \quad (6)$$

$$\Phi^*(q_i) = \frac{1}{N(t)-1} (1 - \Phi(q_i)) \quad (7)$$

IV. ANALYSIS AND SIMULATIONS OF DEGREE DISTRIBUTION

Applying mean field approximation, we get the following time evolution about the average in-degree.

$$\begin{aligned} \frac{\partial k_i}{\partial t} = & \pi_1 m \Pi(k_i) + \pi_2 n [\Phi(q_i) \rho + (1 - \Phi(q_i)) \Pi(k_i)] - \\ & \pi_3 h \left[\frac{1 - \Phi(q_i)}{\pi_1 t} \rho + \left(1 - \frac{1 - \Phi(q_i)}{\pi_1 t} \right) \frac{1 - \Pi(k_i)}{\pi_1 t} \right] \end{aligned} \quad (8)$$

We can ignore $\Phi(q_i)$ in the large-degree range^[19] and thus (8) become

$$\begin{aligned}\frac{\partial k_i}{\partial t} &= \pi_1 m \Pi(k_i) + \pi_2 n \Pi(k_i) - \frac{h\rho}{\pi_1 t} \pi_3 - \frac{h}{\pi_1 t} \pi_3 + h\pi_3 \frac{\Pi(k_i)}{\pi_1 t} \\ &= (\pi_1 m + \pi_2 n + \frac{h}{\pi_1 t} \pi_3) \Pi(k_i) - \frac{h}{\pi_1 t} (1 + \rho) \pi_3\end{aligned}\quad (9)$$

For large t , (9) become (10):

$$\frac{\partial k_i}{\partial t} \approx (\pi_1 m + \pi_2 n) \frac{k_i}{t(\pi_1 m + \pi_2 n - \pi_3 h)(1 + \rho)} - \frac{h}{\pi_1 t} (1 + \rho) \pi_3 \quad (10)$$

$$M = \frac{\pi_1 m + \pi_2 n}{(1 + \rho)(\pi_1 m + \pi_2 n - \pi_3 h)}, \quad N = \frac{h(1 + \rho)\pi_3}{\pi_1} \quad (11)$$

Thus, (10) become

$$\frac{\partial k_i}{\partial t} = \frac{M \cdot k_i}{t} - \frac{N}{t} = \frac{M \cdot k_i - N}{t}, \quad k_i(t_i) = \rho m \quad (12)$$

The solution of (12) has the form

$$k_i(t) = (\rho \cdot m - \frac{N}{M})(\frac{t}{t_0})^M + \frac{N}{M} \quad (13)$$

Then we obtain the in-degree distribution

$$p(k) \propto k^{-(2 + \rho - \frac{\pi_3 h(1 + \rho)}{\pi_1 m + \pi_2 n})} \quad (14)$$

$$\gamma = 2 + \rho - \frac{\pi_3 h(1 + \rho)}{\pi_1 m + \pi_2 n} \quad (15)$$

where the exponent γ is called the scaling exponent of in-degree, dependent of the parameter $\pi_1, \pi_2, \pi_3, m, n, h$. The scaling exponent of in-degree in this paper is different from paper [19]. These different comes from the deletion of old edges. The scaling exponent of in-degree depends on the parameter π_1, π_2, π_3 is more reasonable.

We can ignore $[[k_j]$ in the small-degree range. The process is similar to the Reference [19].

Applying continuous approximation, we get the following evolution about the average out-degree.

$$\begin{aligned}\frac{\partial q_i}{\partial t} &= \pi_1 \rho m \Pi(k_i) + \pi_2 n [\Phi(q_i) + \rho(1 - \Phi(q_i)) \Pi(k_i)] - \\ &\quad h\pi_3 \left[\frac{1 - \Phi(q_i)}{\pi_1 t} + \rho(1 - \frac{1 - \Phi(q_i)}{\pi_1 t}) \frac{1 - \Pi(k_i)}{\pi_1 t} \right]\end{aligned}\quad (16)$$

We can ignore $\Phi(q_i)$ in the large-degree range and thus (15) become

$$\begin{aligned}\frac{\partial q_i}{\partial t} &= \pi_1 \rho m \Pi(k_i) + \pi_2 \rho n \Pi(k_i) - \frac{h\pi_3}{t\pi_1} - \frac{h\rho\pi_3}{t\pi_1} + h\rho\pi_3 \frac{\Pi(k_i)}{\pi_1 t} \\ &= \rho(\pi_1 m + \pi_2 n + \frac{h\pi_3}{t\pi_1}) \Pi(k_i) - \frac{h\pi_3}{t\pi_1} (1 + \rho)\end{aligned}\quad (17)$$

We assume (16) combined with (11), we obtain,

$$\frac{\partial q_i}{\partial t} = \frac{\rho \cdot M \cdot k_i}{t} - \frac{N}{t}, \quad q_i(t_i) = m \quad (18)$$

We assume $L = \rho \cdot M (\rho \cdot m - \frac{N}{M})$

(18) combined with (13), we obtain,

$$\frac{\partial q_i}{\partial t} = L \frac{t^{M-1}}{t_0^M} + \frac{N}{t} (\rho - 1) \quad (19)$$

The solution of (19) has the form

$$q_i(t) = \frac{L}{M} (\frac{t}{t_0})^M + N(\rho - 1) \ln \frac{t}{t_0} + m - \frac{L}{M} \quad (20)$$

The out-degree has not an analytical solution.

In Fig .1, $\pi_1=0.4, \pi_2=0.5, \pi_3=0.1, m=1, n=1, h=1, t=10000, \rho=0.2$. Square means $m_0=10, e_0=10$, the edges between the initial nodes are in turn connected; Round means $m_0=10, e_0=20$, the edges between the initial nodes are connected randomly; triangle means $m_0=20, e_0=50$, the edges between the initial nodes are connected randomly. The solid line is the analytical result.

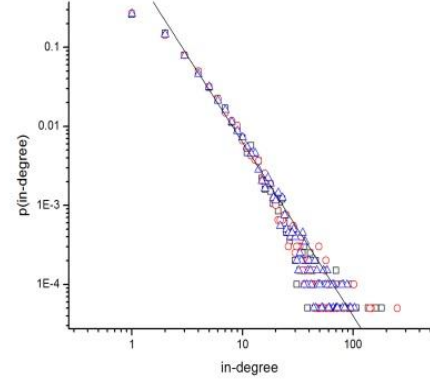


Figure 2. In-degree distributions.

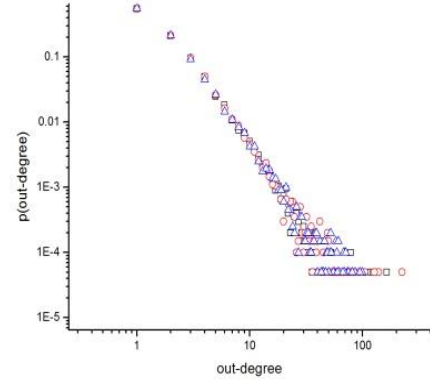


Figure 3. Out-degree distributions.

Fig .2 shows the results of numerical simulation and the theoretical analysis results of in-degree. It can be easily found that the change of the initial network does not influenced in-degree distribution from the figure. We can find that the results of numerical simulation have a good agreement with the theoretical analysis results.

In Fig .3, $\pi_1=0.4, \pi_2=0.5, \pi_3=0.1, m=1, n=1, h=1, t=10000, \rho=0.2$. Square means $m_0=10, e_0=10$, the edges between the initial nodes are in turn connected; Round means $m_0=10, e_0=20$, the edges between the initial nodes are connected randomly; triangle means $m_0=20, e_0=50$, the edges between the initial nodes are connected randomly.

Fig .3 shows the results of numerical simulation for out-degree. It can be easily find that the change of the initial network does not influenced out-degree distribution from the figure.

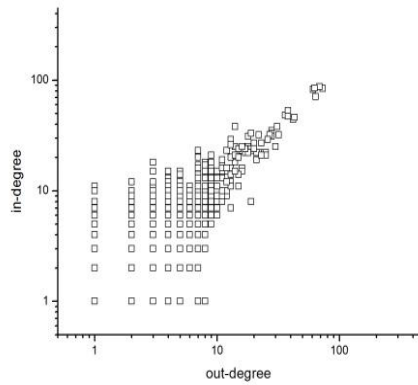


Figure 4. Log-log scatter plot between in/out-degree.

Fig .4 shows strong correlation between the in-degree and out-degree. This is in line with asymmetry between distributions of the in-degree and out-degree^[19] which is found in Twitter.

V. CONCLUSIONS

In this paper, we further explored the evolution mechanisms of microblogging networks. We have introduced the model that gives more realistic description. The model includes the addition of new nodes, the addition of new edges between old nodes, and the deletion of some edges. We have analytically calculated the degree distribution using the continuum theory.

Microblogging network is a directed network and in satisfaction with the power-law distribution. These have a lot of similarities between web network and supply chain network, so the model can be referenced and applied in these networks.

Building evolutionary models for the microblogging network is only one of the most basic works. In future we will study the propagation of information in the microblogging network.

ACKNOWLEDGMENT

This work is supported by the Natural Science Foundation of China under Project No. 71371039, 91024003 and 91024031.

REFERENCES

- [1] H. Kwak, C. Lee, H. Park and S. Moon, "What is Twitter, a social network or a news media", Proc. of the 19th Int Conf on World Wide Web, ACM, New York, pp.591-600, 2010.
- [2] A. Java, X. Song, T. Finin and B. Tseng, "Why We Twitter: Understanding Microblogging Usage and Communities, Proc of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis", ACM, New York, pp.56-65, 2007.
- [3] B. A. Huberman, D. M. Romero and F. Wu, "Social networks that matter: Twitter under the microscope", First Monday, vol.14, no.1, pp.1-5, 2009.
- [4] S. N. Dorogovtsev and J. F. F. Mendes, "Scaling behaviour of developing and decaying networks", Europhysics letters, vol.52, no.1, pp.33-39, 2000.
- [5] K. Deng and T. Yi, "Growing Networks Based on the Mechanism of Addition and Deletion", Chinese physics letters, vol.21, no.9, pp.1858-1860, 2004.
- [6] K. Deng, H. Zhao and D. Li, "Effect of node deleting on network structure", Physica A, vol.379, no.2, pp.714-726, 2007.
- [7] R. Albert and A. L. Barabási, "Topology of Evolving Networks: Local Events and Universality", Physical Review Letters, vol.85, no.24, pp.5234-5237, 2000.
- [8] G. Bianconi and A. L. Barabási, "Competition and multiscaling in evolving networks", Europhysics Letters, vol.54, no.4, pp.436-442, 2001.
- [9] S. N. Dorogovtsev and J. F. F. Mendes, "Evolution of networks with aging of sites", Physical Review E, vol.62, no.2, pp.1842-1845, 2000.
- [10] Q. Chen and D. Shi, "The modeling of scale-free networks", Physica A, vol.335, no.1-2, pp.240-248, 2004.
- [11] V. Zlatić and H. Štefančić, "Influence of reciprocal edges on degree distribution and degree correlations", Physical Review E, vol.80, pp.016117, 2009.
- [12] M. Deijfen and M. Lindholm, "Growing networks with preferential deletion and addition of edges", Physica A, vol.388, no.19, pp.4297-4303, 2009.
- [13] Ferretti L, Cortelezzi M, Yang B, et al. "Features and heterogeneities in growing network models". Physical Review E, 2012, 85(6): 066110.
- [14] Dos Santos A M, De Almeida M L, Mendes G A, et al. "Generalized scale-free homophilic network". International Journal of Modern Physics C, 2015, 26(9): 1550097.
- [15] De Almeida M L, Mendes G A, Viswanathan G M, et al. "Scale-free homophilic network". The European Physical Journal B, 2013, 86(2): 1-6.
- [16] Nicosia V, Bianconi G, Latora V, et al. "Growing multiplex networks". Physical review letters, 2013, 111(5): 058701.
- [17] Kim J Y, Goh K I. "Coevolution and correlated multiplexity in multiplex networks". Physical review letters, 2013, 111(5): 058702.
- [18] Nicosia V, Bianconi G, Latora V, et al. "Nonlinear growth and condensation in multiplex networks". Physical Review E, 2014, 90(4): 042807.
- [19] H. D. Mitsuhiro Nakamura, "Cognitive-costed agent model of the microblogging network," Proceedings of the AESCS the Sixth Inter-national Workshop on A gent-based Approaches in Economic and Social Complex Systems, Taipei, Taiwan, 2009.