A Weighted Network Model Based on the Preferential Selection of Edges¹

Shenghui Chen and Qinghua Chen²

School of Mathematics and Computer Science Fujian Normal University Fuzhou 350007, China

Abstract

A model for the growth of weighted networks is proposed. The model is based on the edge preferential selection. By the master equation approach, the distributions of strength, weight and degree are provided and results show that each distribution has a power-law tail. Particularly, the network has a high clustering coefficient.

Keywords: Weighted networks; Strength distribution; Weight distribution; Degree distribution; Clustering coefficient

1 Introduction

In the past few years, much attention has been focused on the study of complex networks. For example, World Wide Web (WWW) [2,18], Internet [29], scientific collaboration networks (SCN) [26,10], world wide airport networks (WAN) [17,7], etc. Many of these systems share a common scale-free feature that the degree distribution P(k) decays as a power law, i.e., $P(k) \propto k^{-\gamma}$, and the exponent γ is scattered between 2 and 3. The idea of incorporating preferential attachment in a growing network, first introduced by Barabási and Albert [4,5], has lead to a considerable number of models for scale-free

¹This research is supported by the Project of Science and Technology of Fujian Province (Grant No. 2006F5035), and the Natural Science Foundation of Fujian Province, China (Grant No. Z0511020).

²Corresponding author: School of Mathematics and Computer Science, Fujian Normal University, Fuzhou 350007, China.

E-mail address: shenghuichen123@163.com (S. Chen), qhdchen@yahoo.com.cn (Q. Chen).

networks [1,11-15,19,20,22,24,27,31,32]. Researchers have mainly focused on the topological property of the networks, that is, unweighted networks. However, many technological, biological and social systems are best described by weighted networks, whose properties and dynamics depend not only on their structures but also on the connection weights between their nodes. For instance, the number of coauthored papers of two scientists is very important in the understanding of the web of scientists with collaborations [26], and the diversity of the predator-prey interaction and of metabolic reactions is considered a critical ingredient of ecosystems [28,21] and metabolic networks[3], and the number of available seats in flights between two airports is an important quantity in the characterization of WAN.

A network is usually expressed by a graph, whose nodes are the elements of the system and edges represent the interactions between them. For a topological network, the graph can be expressed via its adjacency matrix W, whose element $w_{ij} = 1$ if node i and j are connected, and $w_{ij} = 0$ otherwise. Similarly, a weighted network can be described by a weighted adjacency matrix W, whose element w_{ij} represents the weight on the edge connecting node i and j. For the sake of simplicity, we only consider undirected networks in this paper, where the weights are symmetric, i.e., $w_{ij} = w_{ji}$. As a generalization of the degree, the strength s_i of node i, defined as $s_i = \sum_{j \in \nu(i)} w_{ij}$, where the sum runs over the set $\nu(i)$ of neighbors of node i, is an important quantity in weighted networks. The strength of a node integrates the information about its connectivity and the weights of its links. For instance, the strength in WAN provides the actual traffic going through a node and indicates the size and importance of an airport in a certain extent. For the SCN, the strength is a measure of scientific productivity. Recent studies [7,23,16] have shown that the distributions of node strength and edge weight are heavy tailed in many weighted networks. Many models have been proposed to investigate the mechanism responsible for the properties found in many natural weighted networks [8,9,33,35,36,37].

In most previous network models, the probability Π_i that node i is chosen to be connected to the new node often relates directly to the quantities of node i, like node degree k_i [4,5], strength s_i [8], fitness η_i [37,6], etc. Differentially, Dorogovtsev et al. proposed an evolving network model where the new node will be linked to both ends of an existing link selected randomly [15]. By the master equation approach [14], the degree distribution of the network was provided,

$$P(k) = \frac{12}{k(k+1)(k+2)}. (1)$$

Liu et al. indicated that, generally, a realistic network grows in time according to an attachment rule that is neither completely preferential nor completely random. In terms of the quantity Π_i , it should contain both a deterministic component reflecting preferential attachment, and a random component as well [25]. In particular, they assumed

$$\Pi_i = \frac{(1-q)k_i + q}{\sum_{j} [(1-q)k_j + q]},\tag{2}$$

where $0 \le q \le 1$ and the summation is over the whole network at a give time. Controlled by a simple parameter q, the model can produce a scale-free network with the degree distribution

$$P(k) \propto k^{-\left(3 + \frac{q}{m(1-q)}\right)},\tag{3}$$

where the parameter m represents the number of links a new node possess.

In present paper, we introduce a weighted network model based on edge preferential selection. At each time step, a new node is added with two edges that connect the new node to both ends of a preferentially selected link. Meanwhile, the weight of the link selected preferentially will be strengthened. Using the master equation approach, we analyze the properties of the network produced by present model. Results show that the model can produce a network with the power-law distributions of strength, weight and degree, and the clustering coefficient of the network shows a high value at the same time.

2 The model

Inspired by the work in Ref. [15,25], we propose a model to study the self-organization of weighted evolving networks. The algorithm goes as follows:

- (i) Initial condition: Starting with three nodes connected each other. The weight of each link is assigned 1.
- (ii) Growth: Add a new node with two edges that connect the new node to both ends of an existing link. The preferential probability that edge e_{ij} will be selected is given by:

$$P(w_{ij}) = \frac{(1-q)w_{ij} + q}{\sum_{e_{kl} \in E} [(1-q)w_{kl} + q]},$$
(4)

where $0 \le q \le 1$ and E represents the set of all edges. The weight of each new edge is assigned 1.

- (iii) Weight evolution: Add 1 to the weight w_{ij} of edge e_{ij} selected in the previous step.
- (iv) The whole process is repeated from step (ii), until the desired size of the network is reached.

After t time steps, the model leads to a network with t+3 nodes, 2t+3 edges, and the total weight of the network is $\sum_{e_{kl} \in E} w_{kl} = 3t+3$.

3 Strength, weight and degree distributions

We will investigate the strength, weight and degree distributions by the analysis. At each time step, if an edge e_{ij} of node i is selected preferentially, then node i will be connected to the new node and the degree k_i of node i will increases by 1, meanwhile, the weight w_{ij} of edge e_{ij} increases by 1, therefore, the strength s_i of node i increases by 2. Consequently, the strength s_i of node i displays the linear property:

$$s_i = 2k_i - 2. (5)$$

At time t, the probability that the weight w_{ij} of edge e_{ij} increases by 1 is given by:

$$p_1(w_{ij}) = \frac{(1-q)w_{ij} + q}{\sum\limits_{e_{kl} \in E} [(1-q)w_{kl} + q]} = \frac{(1-q)w_{ij} + q}{(3-q)t + 3}.$$
 (6)

Thus, the probability that the strength s_i of node i increases by 2 is

$$p_2(s_i) = \sum_{j \in \nu(i)} p_1(w_{ij}) = \frac{(1-q)s_i + qk_i}{(3-q)t + 3} = \frac{(1-\frac{q}{2})s_i + q}{(3-q)t + 3}.$$
 (7)

We denote p(s, i, t) the probability that at time t node i has a strength s. Let

$$p(s,t) = \sum_{i=1}^{t} p(s,i,t)/t.$$
 (8)

Thus, the strength distribution can be defined as

$$P(s) = \lim_{t \to \infty} p(s, t). \tag{9}$$

Similarly, we have the definitions of p(k, i, t), p(k, t) and p(k) for the node degree. We denote i_1 , i_2 the two nodes linked by node i when it enter the system. Let

$$P(w,t) = \frac{\sum_{i=1}^{t} [p(w, e_{ii_1}, t) + p(w, e_{ii_2}, t)]}{2t},$$
(10)

where $p(w, e_{ij}, t)$ represents the probability that at time t edge e_{ij} has a weight w. Thus, the weight distribution can be defined as

$$P(w) = \lim_{t \to \infty} p(w, t). \tag{11}$$

The main properties of present model are the following.

Theorem 3.1 If $\lim_{t\to\infty} p(s,t)$ exists and $\lim_{t\to\infty} t[p(s,t+1)-p(s,t)]=0$, then

$$P(s) = \frac{(3-q)\Gamma(\frac{7-2q}{2-q})\Gamma(\frac{s}{2} + \frac{q}{2-q})}{(5-q)\Gamma(\frac{2}{2-q})\Gamma(\frac{s}{2} + \frac{5-q}{2-q})},$$
(12)

thus,

$$P(s) \sim A_1 s^{-\frac{5-2q}{2-q}}, \text{ for large } s$$
 (13)

where s are even numbers and A_1 is a constant.

Proof. p(s, i, t) satisfies the following master equation by utilizing the total probability formula:

$$p(s, i, t+1) = p_2(s-2)p(s-2, i, t) + [1 - p_2(s)]p(s, i, t).$$
(14)

From Eq. (14), we have

$$(t+1)p(s,t+1) - p(s,t+1,t+1)$$

$$= p_2(s-2)tp(s-2,t) + [1-p_2(s)]tp(s,t).$$
(15)

Note that $p(s,t+1,t+1) = \delta_{s,2}$ and $\lim_{t\to\infty} t[p(s,t+1) - p(s,t)] = 0$. Inserting Eq. (7) into Eq. (15) and letting $t\to\infty$, we derive the following recursive equation:

$$P(s) - \delta_{s,2} = \frac{\left(1 - \frac{q}{2}\right)(s-2) + q}{3-q}P(s-2) - \frac{\left(1 - \frac{q}{2}\right)s + q}{3-q}P(s),\tag{16}$$

with solution

$$P(s) = \begin{cases} \frac{(2-q)(s-2)+2q}{(2-q)s+6} P(s-2), & s \ge 4\\ & . \\ \frac{3-q}{5-q}, & s = 2 \end{cases}$$
 (17)

Using the Stirling formula, we obtain the following from Eq. (17):

$$P(s) = \frac{(3-q)\Gamma(\frac{7-2q}{2-q})\Gamma(\frac{s}{2} + \frac{q}{2-q})}{(5-q)\Gamma(\frac{2}{2-q})\Gamma(\frac{s}{2} + \frac{5-q}{2-q})}$$

$$\sim \frac{(3-q)\Gamma(\frac{7-2q}{2-q})\sqrt{2\pi}(\frac{s}{2} + \frac{q}{2-q})^{\frac{s}{2} + \frac{q}{2-q} - \frac{1}{2}}e^{-(\frac{s}{2} + \frac{q}{2-q})}}{(5-q)\Gamma(\frac{2}{2-q})\sqrt{2\pi}(\frac{s}{2} + \frac{5-q}{2-q})^{\frac{s}{2} + \frac{5-q}{2-q} - \frac{1}{2}}e^{-(\frac{s}{2} + \frac{5-q}{2-q})}}$$

$$\sim A_1 s^{-\frac{5-2q}{2-q}}, \tag{18}$$

where $A_1 = \frac{(3-q)\Gamma(\frac{7-2q}{2-q})}{(5-q)\Gamma(\frac{2}{2-q})}2^{\frac{5-2q}{2-q}}$, i.e., the strength distribution of the network follows a power law for large s, with the exponent γ_s : $\frac{5}{2} \leq \gamma_s = \frac{5-2q}{2-q} \leq 3$. The proof is completed.

Theorem 3.2 If $\lim_{t\to\infty} p(w,t)$ exists and $\lim_{t\to\infty} t[p(w,t+1)-p(w,t)]=0$, then

$$P(w) = \frac{(3-q)\Gamma(w + \frac{q}{1-q})\Gamma(\frac{5-2q}{1-q})}{(4-q)\Gamma(w + \frac{4-q}{1-q})\Gamma(\frac{1}{1-q})},$$
(19)

thus,

$$P(w) \sim A_2 w^{-\frac{4-2q}{1-q}}, \text{ for large } w$$
 (20)

where A_2 is a constant and $0 \le q < 1$.

Proof. $p(w, e_{ij}, t)$ satisfies the following master equation by utilizing the total probability formula:

$$p(w, e_{ij}, t+1) = p_1(w-1)p(w-1, e_{ij}, t) + [1 - p_1(w)]p(w, e_{ij}, t).$$
 (21)

From Eq. (22), we have

$$2(t+1)p(w,t+1) - 2\delta_{w,1} = 2p_1(w-1)tp(w-1,t) + 2[1-p_1(w)]tp(w,t).$$
 (22)

Note that $\lim_{t\to\infty} t[p(w,t+1)-p(w,t)]=0$. Inserting Eq. (6) into Eq. (22) and letting $t\to\infty$, we obtain the following recursive equation:

$$P(w) - \delta_{w,1} = \frac{(1-q)(w-1) + q}{3-q} P(w-1) - \frac{(1-q)w + q}{3-q} P(w),$$
 (23)

with solution

$$P(w) = \begin{cases} \frac{(1-q)(w-1)+q}{(1-q)w+3} P(w-1), & w \ge 2\\ \frac{3-q}{4-q}, & w = 1 \end{cases}$$
 (24)

Similarly, when $0 \le q < 1$, from Eq. (24) we have

$$P(w) = \frac{(3-q)\Gamma(\frac{5-2q}{1-q})\Gamma(w+\frac{q}{1-q})}{(4-q)\Gamma(\frac{1}{1-q})\Gamma(w+\frac{4-q}{1-q})}$$

$$\sim A_2 w^{-\frac{4-2q}{1-q}}, \tag{25}$$

where $A_2 = \frac{(3-q)\Gamma(\frac{5-2q}{1-q})}{(4-q)\Gamma(\frac{1}{1-q})}$. Consequently, when $0 \le q < 1$, the weight distribution of the network follows a power law for large w, with the exponent γ_w : $4 \le \gamma_w = \frac{4-2q}{1-q} < \infty$. The proof is completed.

Particularly, when q = 1, from Eq. (24) we have $P(w) = 2 \cdot 3^{-w}$, i.e., the weight distribution P(w) decays exponentially. Indeed, when q = 1, the selection of edge is completely random, which may result into the homogeneous property of the edge weight of the network.

Theorem 3.3 If $\lim_{t\to\infty} p(k,t)$ exists and $\lim_{t\to\infty} t[p(k,t+1)-p(k,t)]=0$, then

$$P(k) = \frac{(3-q)\Gamma(\frac{7-2q}{2-q})\Gamma(k+\frac{2q-2}{2-q})}{(5-q)\Gamma(\frac{2}{2-q})\Gamma(k+\frac{3}{2-q})},$$
(26)

thus,

$$P(k) \sim A_3 k^{-\frac{5-2q}{2-q}}, \text{ for large } k,$$
 (27)

where A_3 is a constant.

Proof. Obviously, the probability $p_3(k_i)$ that the degree k_i of node i increases by 1 is equal to the probability $p_2(s_i)$ that the strength s_i of node i increases by 2, i.e.,

$$p_3(k_i) = p_2(s_i) = \frac{\left(1 - \frac{q}{2}\right)(2k_i - 2) + q}{(3 - q)t + 3} = \frac{(2 - q)k_i + 2q - 2}{(3 - q)t + 3}.$$
 (28)

Similarly, p(k, i, t) satisfies the following equation:

$$p(k, i, t+1) = p_3(k-1)p(k-1, i, t) + [1 - p_3(k)]p(k, i, t).$$
 (29)

From Eq. (29), we have

$$(t+1)p(k,t+1) - p(k,t+1,t+1) = p_3(k-1)tp(k-1,t) + [1-p_3(k)]tp(k,t).$$
 (30)

Note that $\lim_{t\to\infty} t[p(k,t+1)-p(k,t)]=0$ and $p(k,t+1,t+1)=\delta_{k,2}$. Inserting Eq. (28) into Eq. (30) and letting $t\to\infty$, we derive the following recursive equation:

$$p(k) - \delta_{k,2} = \frac{(2-q)(k-1) + 2q - 2}{3-q}p(k-1) - \frac{(2-q)k + 2q - 2}{3-q}p(k), \quad (31)$$

with solution

$$P(k) = \begin{cases} \frac{(2-q)(k-1)+2q-2}{(2-q)k+q+1} P(k-1), & k \ge 3\\ \frac{3-q}{5-q}, & k = 2 \end{cases}$$
(32)

Similarly, from Eq. (32) we have

$$P(k) = \frac{(3-q)\Gamma(\frac{7-2q}{2-q})\Gamma(k+\frac{2q-2}{2-q})}{(5-q)\Gamma(\frac{2}{2-q})\Gamma(k+\frac{3}{2-q})}$$

$$\sim A_3 k^{-\frac{5-2q}{2-q}}, \tag{33}$$

where $A_3 = \frac{(3-q)\Gamma(\frac{7-2q}{2-q})}{(5-q)\Gamma(\frac{2}{2-q})}$, it means that the degree distribution also follows a power law for large k, with the exponent γ_k : $\frac{5}{2} \leq \gamma_k = \frac{5-2q}{2-q} \leq 3$. The proof is completed.

4 Clustering coefficient

Along with the scale-free property of the networks, another significant quantity C, namely clustering coefficient, is widely used to analyze the structure of the systems. Actually, in many real networks, especially in social networks, the clustering coefficient C shows a high value. The clustering coefficient of node i [34] is defined as

$$C_i = \frac{2E_i}{k_i(k_i - 1)},\tag{34}$$

where E_i denotes the number of the existing links between all neighbors of node i, and k_i is the degree of node i. Then, the clustering coefficient of the whole network is the average:

$$C = \frac{1}{N} \sum_{i=1}^{N} C_i, \tag{35}$$

where N is the size of the network. Obviously, the clustering coefficient C_i measures the local cohesiveness in the neighborhood of node i and the average clustering coefficient C expresses the statistical level of cohesiveness measuring the global density of interconnected vertex triplets in the network.

The WS model [34] shows a high clustering but without the power-law degree distribution, while the BA model [4,5] with the scale-free nature does not possess the high clustering. In present model, at each time step the new

node is linked to both ends of an existing ling, then a new triad is formed, which may result in a high value of the clustering coefficient. We can calculate the clustering coefficient using a slight variation of the rate equation approach [30]: for each vertex $\partial E(k)/\partial k = 1$ with the initial condition E(2) = 1, resulting in E(k) = k - 1. Thus, the clustering coefficient of a vertex of degree k equals

$$C(k) = \frac{2E(k)}{k(k-1)} = \frac{2}{k}. (36)$$

The average clustering coefficient is obtained by using the degree distribution:

$$C = \sum_{k=2}^{\infty} P(k)C(k), \tag{37}$$

yielding a numerical value of C = 0.793, 0.769 and 0.739 for q = 0, 0.5 and 1, respectively. Consequently, the network produced by present model shows a high value of the clustering coefficient.

5 Conclusions

In present paper, we have proposed a weighted network model based on edge preferential selection. The model results in scale-free behavior for the strength, weight and degree distributions, and the exponents are controlled by a parameter q. Particularly, at each time step a new triad is formed by the newly added node, so the network shows a high value of the clustering coefficient C.

References

- [1] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, Rev. Mod. Phys., 74 (2002), 47-97.
- [2] R. Albert, H. Jenog and A.-L. Barabási, Diameter of the world-wide web, Nature, 401 (1999), 130-131.
- [3] E. Almaas, B. Kovács, T. Viscek, Z.N. Oltval and A.-L. Barabási, Global organization of metabolic fluxes in the bacterium Escherichia coli, Nature, 427 (2004), 839-843.
- [4] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, Science, 286 (1999), 509-512.

- [5] A.-L. Barabási, R. Albert and H. Jeong, Mean-field theory for scale-free random networks, Physica A, 272 (1999), 173-187.
- [6] G. Bianconi and A.-L. Barabási, Competition and multiscaling in evolving networks, Europhys. Lett., 54 (2001), 436-442.
- [7] A. Barrat, M. Barthélemy, R. Pastor-satorras and A. Vespignani, The architecture of complex weighted networks, Proc. Natl. Acad. Sci. U.S.A., 101 (2004), 3747-3752.
- [8] A. Barrat, M. Barthélemy and A. Vespignani, Weighted Evolving Networks: Coupling Topology and Weight Dynamics, Phys. Rev. Lett., 92 (2004), 228701.
- [9] A. Barrat, M. Barthálemy and A. Vespignani, Modeling the evolution of weighted networks, Phys. Rev. E, 70 (2004), 066149.
- [10] A.-L. Barabási, H. Jenog, Z. Néda, E. Ravasz, A. Schubert and T. Vicsek, Evolution of the social network of scientific collaborations, Physica A, 311 (2002), 590-614.
- [11] Q. Chen and D. Shi, The modeling of scale-free networks, Physica A, 335 (2004), 240-248.
- [12] Q. Chen and D. Shi, Markov chains theory for scale-free networks, Physica A, 360 (2006), 121-133.
- [13] S.N. Dorogovtsev and J.F.F Mendes, Evolution of networks, Adv. Phys., 51 (2002), 1079-1187.
- [14] S.N. Dorogovtsev, J.F.F Mendes and A.N. Samukhin, Structure of Growing Networks with Preferential Linking, Phys. Rev. Lett., 85 (2000), 4633-4636.
- [15] S.N. Dorogovtsev, J.F.F Mendes and A.N. Samukhin, Size-dependent degree distribution of a scale-free growing network, Phys. Rev. E, 63 (2001), 062101
- [16] D. Garlaschelli, S. Battiston, M. Castri, V. D. P. Servedio and G. Caldarelli, The scale-free topology of market investments, (2003) cond-mat/0310503.

- [17] R. Guimera, S. Mossa, A. Turtschi and L.A.N. Amearal, The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles, Proc. Natl. Acad. Sci. U.S.A., 102 (2005), 7794-7799.
- [18] B.A. Huberman and L.A. Adamic, Growth dynamics of the World-Wide Web, Nature, 401 (1999), 131-131.
- [19] K. Klemm and V.M. Eguíluz, Highly clustered scale-free networks, Phys. Rev. E, 65 (2002), 036123.
- [20] K. Klemm and V.M. Eguiluz, Growing scale-free networks with small-world behavior, Phys. Rev. E, 65 (2002), 057102.
- [21] A.E. Krause, K.A. Frank, D.M. Mason, R.E. Ulanowicz and W.W. Taylor, Compartments revealed in food-web structure, Nature, 426 (2003), 282-285.
- [22] P.L. Krapivsky, S. Redner and F. Leyvraz, Connectivity of Growing Random Networks, Phys. Rev. Lett., 85 (2000), 4629-4632.
- [23] C. Li and G. Chen, Network connection strengths: Another power-law?, (2003) cond-mat/0311333.
- [24] X. Li and G. Chen, A local-world evolving network model, Physica A, 328 (2003), 274-286.
- [25] Z. Liu, Y.-C. Lai, N. Ye and P. Dasgupta, Connectivity distribution and attack tolerance of general networks with both preferential and random attachments, Physics Letters A, 303 (2002), 337-344.
- [26] M.E.J. Newman, The structure of scientific collaboration networks, Proc. Natl. Acad. Sci. U.S.A., 98 (2001), 404-409.
- [27] M.E.J. Newman, The structure and function of complex networks, SIAM Review, 45 (2003), 167-256.
- [28] S.L. Pimm, Food Webs (University of Chicago Press, Chicago, 2002), 2nd ed.
- [29] R. Pastor-satorras and A. Vespignani, Evolution and structure of the Internet: Astatistical Physics Approach, Cambridge University Press, Cambridge, England, 2004.

- [30] G. Szabó, M. Alava and J.kertész, Structural transitions in scale-free networks, Phys. Rev. E, 67 (2003), 056102.
- [31] D. Shi, Q. Chen and L. Liu, Markov chain-based numerical method for degree distributions of growing networks, Phys. Rev. E, 71 (2005), 036140.
- [32] J. Saramäki and K. Kaski, Scale-free networks generated by ramdom walkers, Physica A, 341 (2004), 80-86.
- [33] W.-X. Wang, B. Hu, T. Zhou, B.-H. Wang and Y.-B. Xie, Mutual selection model for weighted networks, Phys. Rev. E, 72 (2005), 046140.
- [34] D.J. Watts and S.H. Strogatz, Collective dynamics of small-world networks, Nature, 393 (1998), 440.
- [35] W.-X. Wang, B.-H. Wang, B. Hu, G. Yan and Q. Ou, General Dynamics of Topology and Traffic on Weighted Technological Networks, Phys. Rev. Lett., 94 (2005), 188702.
- [36] S.H. Yook, H. Jeong, A.-L. Barabási and Y. Tu, Weighted evolving networks, Phys. Rev. Lett., 86 (2001), 5835-5838.
- [37] D. Zheng, S. Trimper, B. Zheng and P. M. Hui, Weighted scale-free networks with stochastic weight assignments, Phys. Rev. E, 67 (2003), 040102.

Received: October 26, 2006