Evolutional family networks generated by group-entry growth mechanism with preferential attachment and their features

Takeshi Ozeki

Department of Electrical and Electronics Engineering, Faculty of Science and Technology,

Sophia University, 7-1 Kioicho, Chiyodaku, Tokyo, 102-8554, Japan

t-ozeki@gentei.ee.sophia.ac.jp

Abstract**:-The group-entry growth mechanism with preferential attachment generates a new class of scale-free networks, from which exact asymptotic connectivity distributions and generation functions are derived. This new class of scale-free networks evolve from aristocratic networks to egalitarian networks with the asymptotic power law exponent of $\gamma = 2 + M$, depending on the size M and topology of the constituent groups. The asymptotic connectivity distribution is fitted very well with the numerical simulation, even in the region of smaller degrees. It is then demonstrated that small networks can be analysed in order to find their growth mechanism parameters using asymptotic connectivity distribution templates in the region of smaller degrees, for which it is easy to satisfy a statistical level of significance. This approach is believed to develop a new search method for scale-free networks in the real world. As an example of an evolutional family network in the real world, the Tokyo Metropolitan Railway Network is analysed.

Key-Words: *network modelling, connectivity distribution, power laws, asymptotic solution, scale-free network, shortest path, clustering coefficient, diameters, templates*

1 Introduction

Scale-free network science [1,2,3] is expected to provide potential methods by which to analyse various network characteristics of complex organizations in real-world systems, such as epidemics [4] and the dependability of social infrastructure networks [5], on the Internet. Evolutional network modelling is desirable to analyse such characteristics depending on network topologies. The Watts-Strogatz's small world [6] evolves from regular lattice networks to Erdos-Renyi's random networks by random rewiring of links by adjusting their probability [7]. The Watts-Strogatz's small world having a fixed number of nodes is discussed as a static network. On the other hand, the scale-free network of Barabasi-Albert (BA model) introduces the concept of growing networks with preferential attachment [8,9]. One of characterizations of networks is given by the connectivity distribution of $P(k)$, which is the probability that a node has k degrees (or, number of links). In the scale-free networks based on the BA model, the connectivity distribution follows the power law, in which $P(k)$ is approximated as $k^{-\gamma}$, having the exponent γ=3. Real-world complex networks are analysed in order to find various scale-free networks with various exponents, which are covered in References [1,2,3]. As an example, it is well known that social infrastructure networks, such as power grids, as egalitarian networks, follow the power law with an exponent of 4 [8]. Many trials have been reported to generate models with larger exponents for fitting these real-world networks [10-17]. Dorogovtsev *et al.* [14] modified the preferential attachment probability as $\Pi(k) \propto a m + k$ and derived the exact asymptotic solution of the connectivity distribution showing a wide range of exponents $\gamma = a + 2$, where a is the attractiveness and m is the degree of a new node. One of the static models modifying the rewiring alogorithm is successfully applied to analyse the power-law relation of the betweenness as a measure of the network load distribution, for instance [16]. The BA model modified with respect to the preferential attachment probability is practical for fitting the exponents of real-world networks, but several tasks must be accomplished in order to identify the physical causes of such preferential attachment probability. In particular, the model has difficulty in finding parametric relations with network topologies.

In the present report, the "evolutional family networks" generated by a "group entry growth mechanism" with preferential attachment is proposed. This is a modification of the BA model in the "growth mechanism", that is, the basic BA model assumed "one node" joining the network at each time step. This is the first model to generate a new kind of scale-free network with exponents depending on the topological parameters of the constituent groups. This new growth mechanism can be imagined by considering the famous Granovetter's social world [18], i.e., the social world grows with entry of a family or a group. Initially, the family members are strongly connected with each other, and then make new weak connections to the old society members, preferentially. As illustrated in Fig.1, a new family has M constituent nodes (white circles) that are initially strongly connected in a full mesh topology, for example. Each node of the new family makes one weak link connecting to an old node (black circles) that is already present in the network. At a group entry step, it is forbidden that nodes in the new family make connections to the same old nodes, to which the other nodes in the new family make connections. This requirement is essential in the group entry assumption and for the origin of equality of the network.

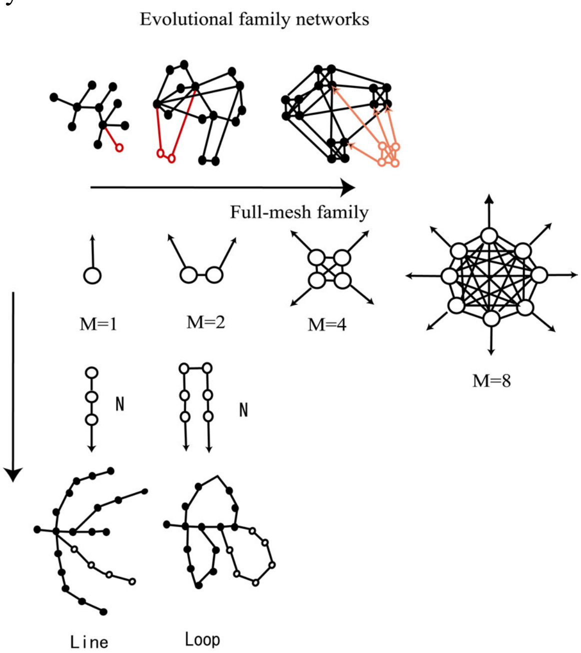


Fig. 1. The group entry units in the growth mechanism of "evolutional family networks". Horizontally, full-mesh family networks evolve with family size M. Vertically, line and loop family networks evolve with a constituent group size of N.

In the real world, this non-duplicate attachment at each time step relates to the requirement to realize efficient information links because the information obtained by each weak link of the constituent node becomes common in the new family, and duplication of information may reduce the efficiency.

The evolutional family networks with M=1 are coincident with the aristocratic network of the BA model. In contrast, the evolutional family network with larger M is an egalitarian network with larger γ, and is a new class of scale-free network that has a higher regularity and clustering coefficient, but a smaller diameter, with increasing M: These features are quite different from the Watts-Strogatz's small world. The evolutional family network can evolve with changing M through a new class of scale-free networks of various topologies from the BA model network.

2 Asymptotic connectivity distributions of full-mesh family networks

The asymptotic connectivity distributions of the full-mesh family networks shown evolving horizontally in Fig.1 are derived by the method reported by Dorogovtsev *et al.* [14]. At initial time t=1, the number of nodes in the network is M, and the number of links or edges is $M(M-1)/2$. At time $t=t$, the number of nodes in the network is $M \cdot t$ and the total number of links is $\{M^2 t - M\}/2$.

The connectivity distribution probability of node s having the degree k at time t is denoted by $p(k,s,t)$, which is given by the following master equations:

$0 \le s \le M \cdot t - 1$

$$p(k,s,t+1) = p(k,s,t)\left(1 - \frac{k}{M^2 \cdot t - M}\right)^M + p(k-1,s,t)\left(1 - \frac{k-1}{M^2 \cdot t - M}\right)^{M-1} \frac{M(k-1)}{M^2 \cdot t - M} \quad . \qquad (1)$$

$M \cdot t \le s \le M(t+1) - 1$

$$p(M,s,t+1) = 1$$

The average connectivity probability of the network is defined as the connectivity distribution as follows:

$$p(k,t) = \frac{1}{M \cdot t} \sum_{s=0}^{Mt-1} p(k,s,t) . \qquad (2)$$

Here, $p(k,t+1)$ is calculated as follows:

$$(t+1)\left(p(k,t+1) - p(k,t)\right) = -\left(1 + \frac{k}{M}\right) p(k,t) + \frac{(k-1)}{M} p(k-1,t) + 1 \cdot \delta_{k,M} + O(k^2/t^2) . \qquad (3)$$

Assuming the existence of the asymptotic connectivity distribution $\lim_{t\to\infty} p(k,t) = p(k)$ and the convergence of the left-hand side of Eq. (3) to zero, i.e., $\lim_{x\to\infty}(t+1)\left(p(k,t+1) - p(k,t)\right) = 0$, the asymptotic connectivity distribution $p(k)$ is obtained as follows:

$$p(k) = \frac{(2M)!}{M!} \frac{M \cdot (M+1)}{(M+1+k)(M+k)\cdots(k+1)k} \quad ; \; k \ge M \quad . \qquad (4)$$

For M=1, we obtain

$$p(k) = \frac{4}{k(k+1)(k+2)} \qquad k \ge 1 . \qquad (5)$$

which is coincident with the expression derived in Reference [14]. From Eq.4, the evolutional family network with M nodes approximately follows the power law with the exponent $\gamma = M + 2$ for sufficiently large k. Here, the exponent is directly related to the number of network topology parameters M.

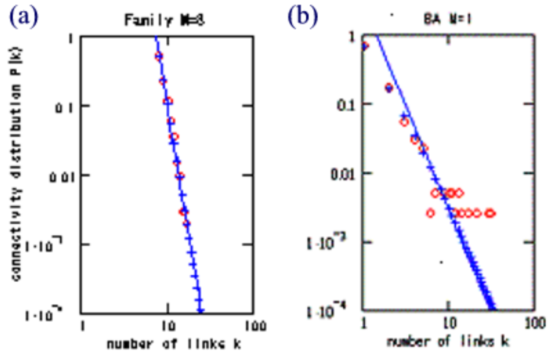


Fig.2 Connectivity distribution of full-mesh family networks with M=8 and M=1(BA). Crosses denote asymptotic distributions and white circles denote values obtained by numerical simulation.

The asymptotic connectivity distributions of the evolutional family network with M=8 and M=1 corresponding to the BA-model network are shown in Fig. 2. The exponents, measured at the degree of 8, are 7.5 and 3, respectively. One point of interest in evolutional family networks is that, even in a relatively small network, the numerically simulated connectivity distribution is fitted very well with the asymptotic connectivity distribution, as indicated in Fig. 2(a). The crosses and circles denote the asymptotic connectivity distribution and the values obtained by numerical simulation, respectively. The network size is N_0= 1000 for all cases. Figure 2(b) illustrates the fitting of the evolutional family network with M=1, which also shows excellent fitting for small degrees of less than 10. However, a larger divergence of the connectivity distribution that was numerically simulated from the asymptotic solution is clear in the distribution region below 10^{-2}. This may suggest the possibility that the asymptotic exact connectivity distribution can be applied to fit real-world networks with relatively small size and within relatively smaller degrees, for which it is easy to satisfy a statistical level of significance.

The generation functions $G_0(x,M)$ are given as follows [19]:

$$G_0(x,M)=\sum_{k=M}^{k=\infty}\frac{(2M)!M(M+1)}{M!}\frac{x^k}{k(k+1)(k+2)\cdots(k+M+1)}. \qquad (6)$$

For M=1,

$$G_0(x,1)=\frac{2(1-x)^2}{x^2}\log\frac{1}{1-x}+\frac{3x-2}{x}=4\cdot f_2(x) \;\; ; \;\; |x|<1, \qquad (7)$$

where $f_m(x)$ is a well-known function found in the mathematical formula [20], and the forms of higher M in $G_0(x,M)$ can be calculated by the following formula:

$$f_{m+1}(x)=\frac{1}{(m+1)!(m+1)}-\frac{(1-x)f_m(x)}{(m+1)x}. \qquad (8)$$

3 Asymptotic connectivity distributions with non-integer exponents

The asymptotic connectivity distribution of the family network combining lines and loops shown evolving vertically in Fig.1 is derived by the method described above, when the probabilities of the appearance of line and loop constituent networks at each entry time step are $1-\varepsilon$ and ε, respectively, as follows:

$$p(k)=p(2)\Gamma(3+2(N+\varepsilon)/(1+\varepsilon))\cdot\frac{\Gamma(k)}{\Gamma(k+1+2(N+\varepsilon)/(1+\varepsilon))} \qquad k\geq 3, \qquad (9)$$

where

$$p(1)=\frac{(1-\varepsilon)/N}{1+(1+\varepsilon)/2(N+\varepsilon)}\;,\; p(2)=\frac{0.5(1+\varepsilon)}{N+1+2\varepsilon}p(1)+\frac{1-(1-\varepsilon)/N}{1+(1+\varepsilon)/(M+\varepsilon)}. \qquad (10)$$

The asymptotic exponent is $\gamma=1+2(N+\varepsilon)/(1+\varepsilon)$, which is the non-integer exponent including the statistical weight ε.

The statistically combined growth mechanism between full-mesh family networks M and M+1 with statistical probabilities of appearance in each time step of ε and 1-ε, respectively, also generates an evolutional family network with a non-integer exponent having the following asymptotic connectivity distribution:

$$p(k)=p(k-1)\frac{k-1}{k+(M+1)(M+2-2\varepsilon)/(M+1-\varepsilon)} \;\; ; k\geq M+2 \qquad (11)$$

where $p(M)$ and $p(M+1)$ are derived by the master equation method reported in Reference [6]. The asymptotic exponent is $\gamma=1+(M+1)(M+2-2\varepsilon)/(M+1-\varepsilon)$.

As an example, Fig.3 shows the connectivity distribution of the statistically combined family networks for the case of M=4 and ε=0.5, with $\gamma=6.56$ in Eq.3, which is denoted by the solid line between the integer asymptotic exponents of $\gamma=6$ for M=4 and γ=7 for M=5. The circles denote the values obtained by numerical simulation adopting this statistically combined growth mechanism, which show good coincidence with the asymptotic connectivity distribution of Eq.3.

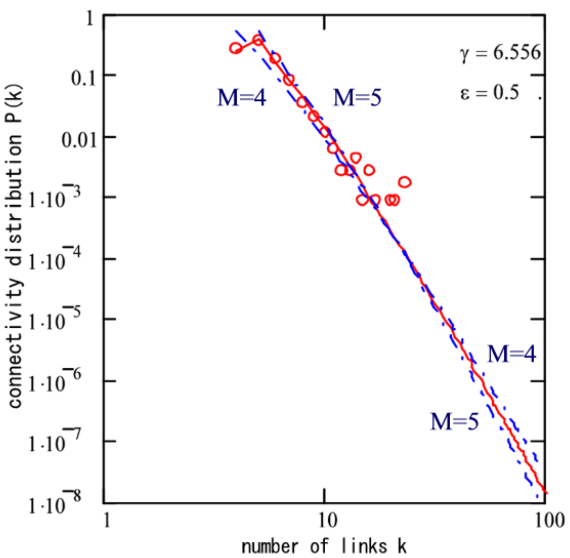


Fig.3 Statistical combination in the growth mechanism for the non-integer exponent for an evolutional family network.

For another example, in the case of changing the number of links *m* connecting the new constituent family to the old nodes, in the range of $1 \le m \le M$, at each time step, the asymptotic exact connectivity distribution is obtained as

$$P(k) = P(M)\frac{\Gamma(M+3+(M^2-M)/m)}{\Gamma(M)} \cdot \frac{\Gamma(k)}{\Gamma(k+3+(M^2-M)/m)} \quad ; \quad k \ge M+1 \qquad (12)$$

where $P(M-1)$ and $P(M)$ are obtained by a procedure similar to that of Eq.4. The exponent is $\gamma = 3 + M(M-1)/m$. Thus, these evolutional family networks can evolve from $\gamma = 2 + M$ to $\gamma = 3 + M(M-1)$ by changing *m*.

Evolutional family networks can sufficiently evolve scale-free networks with a wide range of exponents and topologies. Note that the asymptotic connectivity distribution fit well with the numerical simulation, even in relatively small networks, which suggests that we can expect such small networks to show scale-free characteristics in broad fields by the evolutional family networks.

4 Template for analysing evolutional family networks

Thus far, it has been difficult to apply scale-free network science to relatively smaller networks observed around us, because the network size should be larger than several thousands of nodes, at least, in order to satisfy the statistical level of significance in analysing power-law relations. This is true because the BA model network with a power-law exponent of 3 shows larger divergence of the connectivity distribution that is numerically simulated from the asymptotic solution in the distribution region below 10^{-2}, as shown in Fig.2.

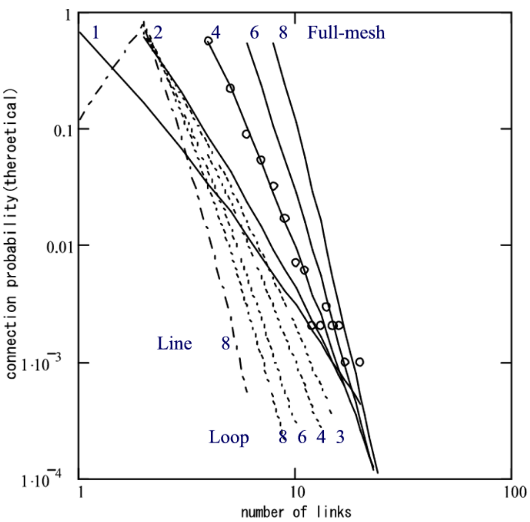


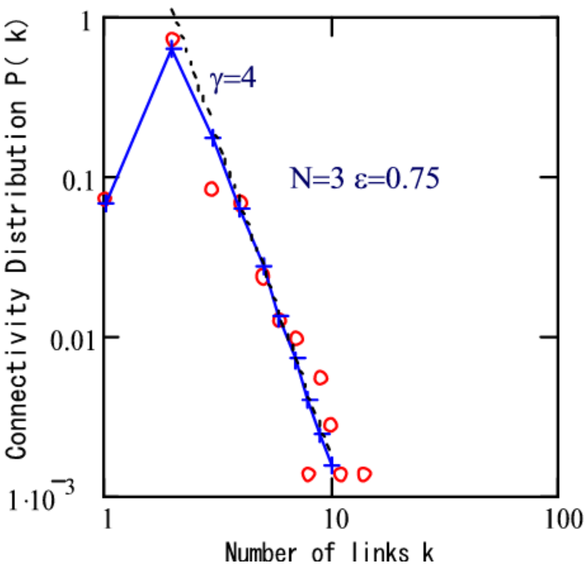
Fig.4 Growth mechanism templates for evolutional family networks.

However, it is much easier to satisfy the statistical criteria of significance when the connectivity distribution in smaller degrees in a real-world network is measurable with higher accuracy.

To demonstrate the usefulness of the growth mechanism templates, a numerically simulated connectivity distribution with a network size of 1,000 is plotted in the template of Fig.4. The white circles denote that of the full-mesh family network, which clearly shows that the circles are fitted well to the growth mechanism template of the full-mesh family network with *M*=4. The coincidence of the growth mechanism is the origin of better fitting. There appear to be various modifications of the templates corresponding to the generalization of both the growth mechanism and the probability of the preferential attachment [6,8,9,10,11].

5 Tokyo Metropolitan Railway Network

As an example of a real-world network analysis by the template, the Tokyo metropolitan railway system [21] is analysed by a statistically combined growth mechanism with line and loop family networks. Figure 5 depicts the connectivity distribution of a central part of the Tokyo Metropolitan Railway System, the number of total stations and links of which are 736 and 1762, respectively. The number of links is counted topologically, for example, we count the number of links between Tokyo and Kanda as 1, even though there are three double railways between them [22]. The fit with N=3 and ε=0.75 is excellent, which suggests that the growth mechanism is coincident with the growth mechanism of evolutional family networks. In the construction of the



railway system, a number of stations with a line or a loop topology are installed simultaneously as a group. The exponent measured at the degree of 8 is 4, which is coincident with that of the power grids of North America [8].

Fig.5 Tokyo Metropolitan Railway Networks

The number of nodes in constituent line and loop networks is N=3, which is reasonable in the central part of Tokyo with respect to its complexity [21]. A real-world network is found to be suitable for fitting networks such as the Tokyo metropolitan railway network by the modified evolutional family network model with loop topology.

6 Discussion

The asymptotic connectivity distribution function of the evolutional family network satisfies the conditions of the statistical distribution function. The network parameters, such as the average number of degrees $< k >$, the standard deviation σ, the clustering coefficient [23] and the network diameter of the evolutional family networks, are listed in Table 1. These parameters were calculated by numerical simulation with N_0=1,000. The standard deviation σ for M=1 is infinity, which corresponds to the BA model. The regularity of the network increases as M increases. The clustering coefficients increase as M increases, which is approximated by (1-2/M) for larger M. The diameter, shown in Table 1, is the number of hops to reach all of the nodes from the largest hub node, as compared to the generation function method of approximately counting the number of neighbours to estimate average path length in Reference [19].

Table 1.Full-mesh family network features

M	*1*	*2*	*4*	*8*
⟨κ⟩	*1*	*3*	*5*	*9*
σ/<κ>	∞	*0.82*	*0.37*	*0.18*
clustering	-	*0*	*0.38*	*0.78*
diameter	*9*	*7*	*6*	*5*

These characteristic parameters suggest that the evolutional family networks evolve from Barabasi-Albert scale-free network to a new class of scale-free networks, which is characterized by a larger clustering coefficient, a small diameter and a high regularity with decreasing σ/<κ> with increasing M. The evolutional family network is a new class of scale-free network that is different from the networks of the Watts-Strogatz small world, even with higher regularity for the larger constituent family size.

7 Conclusions

The group-entry growth mechanism with preferential attachment generates a new class of scale-free networks, the exact asymptotic connectivity distribution and generation function of which were derived. They evolve from aristocratic networks to egalitarian networks with an asymptotic power law exponent of $\gamma = 2 + M$, depending on the size M and the topology of the constituent groups. The asymptotic connectivity distribution fits very well with numerical simulation, even in the region of smaller degrees. Then, it is demonstrated that small networks can be analysed to find their growth mechanism parameters using asymptotic connectivity distribution templates in regions of smaller degrees, for which it is easy to satisfy a statistical level of significance. This approach is believed to develop a new search method for scale-free networks in the real world. Finally, the Tokyo Metropolitan Railway Network was analysed as an example of an evolutional family network in the real world.

Reference:

[1] A. L. Barabasi, *"Linked",* A Plume Book, (2002)

[2] Mark Buchanan, *"Nexus"*, W.W. Norton & Company Ltd., New York (2002)

[3] M. Newman A. L. Barabasi, and. J. Watts, "*The structure and Dynamics of Networks"* Princeton Univ. Press (2006)

[4] ibid., Chapter 4, pp.180-181

[5] William J. Mitchell, *"Me++",* MIT Press (2003)

[6] D. J. Watts, and S. H. Strogatz, *Collective dynamics of "small-world" Networks*, Nature 393, 440-442 (1998)

[7] P. Erdos, and A. Renyi, Publ. Math. Inst. Acad. Sci., 5,17(1960)

[8] A. L. Barabasi. and R. Albert, *Emergence of scaling in random networks,* Science 286, 509 (1999)

[9] A. L. Barabasi, R. Albert, and H. Jeong, *Mean field theory of scale free random networks*, Physica A272, 173 (1999)

[10] A. L. Barabasi, E. Ravasz, and T. Vicsek, *Deterministic scale-free networks*, Physica A299, 559-564 (2001)

[11] A. Arenas, A. Diaz-Guilera, and R. Guimera, *Communication in networks with hierarchical branching,* Phys. Rev. Letts. 86, 3196 (2001)

[12] R. Albert, A. L. Barabasi, *Topology of evolving networks* Phys. Rev. Lett., 85, 5234-5237 (2000)

[13] N. Mathias and V. Gopal, *Small Worlds: How and Why,* Phys. Rev. E-63, 021117 (2001)

[14] S. N. Dorogovtsev, J. F. F. Mendes, & A. N. Samukhin, "*Structure of growing networks with preferential linking,* Phys. Rev. Lett. 85, 4633-4636 (2000)

[15] P. L. Krapivsky, S. Redner, F. Leyvraz, *Connectivity of growing random networks,* Phys. Rev. Lett., 85, 4629 (2000)

[16] K. I. Goh, B. Kahng, and D. Kim, *Universal Behavior of Load Distribution in Scale free networks,* Phys. Rev. Lett. 87, 278701 (2001)

[17] M. J. Alava, and S. N. Dorogovtsev, *Complex networks created by aggregation,* Phys. Rev. E71, 036107 (2005)

[18] M. Granovetter, *The Strength of weak ties,* American Journal of Sociology 78, 1360-1380 (1973)

[19] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, *Scientific collaboration networks* Phys. Rev. E64, 026118 (2001)

[20] S. Moriguchi, K. Udagawa, and S. Ichimatsu, *The Mathematical Formula II*, Iwanami Pub. Co. Tokyo (1956)

[21] *Rail map of Tokyo area* (Shoubunsha Publications, Inc., Tokyo, 2004), ISBN4-398-72008-1

[22] The number of stations with k=1 includes some of stations counted as the end stations at the zone boundary for the central part of Tokyo.

[23] R. P. Satorras and A. Vespignani, "*Evolution and Structure of the Internet*", Cambridge Univ. Press (2004)