

The Continuous Time Markov Processes and Degree Distribution of Evolving Networks*

Dong Han and Qing Zhou

Department of Mathematics, Shanghai Jiaotong University,
Shanghai 200030, P.R. China.

(Received March 28, 2006)

Dedicated to Prof. Fuji Zhang on the occasion of his 70th birthday

Abstract

In this paper a continuous-time Markov process is presented to describe the evolution of networks, and the degree distributions of several processes are discussed.

INTRODUCTION: Complex networks describe a wide range of systems in nature and society, it has brought many attentions recently (see [1, 4, 6, 10, 17]). The recent study of networks shifts the main focus from classical graph theory to very large networks which may dynamically evolving in time. This movement is triggered by the paper of Watts and Strogatz [21] about the ‘small-world phenomenon’ and the paper of Barabási and Albert [3] about ‘scale-free’ nature of networks.

One of most popular topology structure, a scale-free probability distribution for network connectivity, has been reported in many systems as diverse as protein reaction networks ([12, 13]), metabolic reaction networks ([14, 19, 20]) and ecological food webs ([18]). However, the common evolutionary mechanism (‘scale-free’ nature): preferential attachment based on the number of existing connections, is confirmed mainly by heuristic methods or computer simulation in the literatures. In fact, as pointed out in ([5]) that the results on ‘scale-free’ networks are largely heuristic and experimental studies with “rather little rigorous mathematical work; what there is sometimes confirms and sometimes contradicts the heuristic results”.

*Partially supported by National Natural Science Foundation of China (10371074, 10531070)

Thus, increasing the level of rigor in this field becomes a challenge and it will improve our understanding on complex networks.

The purpose of this paper is not only to attempt to prove some theorems, but also to draw the attention of theoretical bio-chemists to this interesting topic. Thus some basic results which we obtained recently are presented in the paper.

Most of evolving networks can be regarded as a graph valued discrete time Markov chain [16]. In the paper, we will embed the evolving process of networks into a continuous time Markov process. The advantage of this consideration is that the corresponding master equation becomes a differential equation, which is much easy to deal with comparing with the difference master equation in the case of discrete time processes. Using the framework of continuous time Markov process, we discussed the degree distribution of several evolving networks.

This paper is organized as follows. In the first section, we introduce a continuous time Markov process which describes evolving networks, and prove a result which plays the same role as Azuma inequality in discrete time processes. In second section, we discuss some continuous time preferential growth model of evolving networks.

§1 GRAPH VALUED MARKOV PROCESSES

Denote the set of natural numbers by \mathbb{N} and

$$E_0 = \left\{ x = ((x_{ij}), n_x) \in \{0, 1\}^{\mathbb{N}^2} \times \mathbb{N} \mid x_{ij} = x_{ji}, x_{ii} = 0 \text{ and } x_{ij} = 0 \text{ if } \max(i, j) > n_x \right\}.$$

An element $x = ((x_{ij}), n_x) \in E_0$ can be regarded as an adjacency matrix of a simple graph G with all n_x -vertices labelled, which may contains isolated vertices. In this sense, the space E_0 can be considered as the set of finite vertex labelled simple graphs.

For $x, y \in E_0$, if $n_x = n_y$, these two vertex labelled graphs have the same order, and furthermore $\|x - y\| = \sum_{i,j} |x_{ij} - y_{ij}|$ is twice of the number of different edges of two vertex labelled graphs $x, y \in E_0$, and which measures the difference between two vertex labelled graphs of the same order. We also use $|x| = \|x\|/2$ for the total number of edges of the graph $x \in E_0$, and $d_i(x) = \sum_j x_{ij}$ is the degree of the node i .

The continuous-time Markov chain $X(t) \in E_0$ we will considered in this paper is a jump process defined by the following one-step jump probability matrix:

$$p(x, y) \geq 0, x, y \in E_0, \quad (1)$$

where $p(x, x) = 0, \sum_y p(x, y) = 1$, and the sojourn time in any state is exponentially distributed with rate 1.

Note that, if for all $i, j \in \mathbb{N}$, $x_{ij} \leq y_{ij}$ and $n_x \leq n_y$, we can think that the graph y is obtained from x by adding some new nodes and/or edges, or conversely, x is obtained from y

by deleting some existing nodes and/or edges. So the Markov chain defined above describes evolving networks, it will be called a graph valued process.

Let τ_k be the k -th transition time of the process, that is,

$$\tau_0 = 0, \tau_k = \inf\{t > \tau_{k-1} | X(t) \neq X(\tau_{k-1})\}, k \geq 1,$$

and $T_k = \tau_k - \tau_{k-1}$ denote sojourn times, $N(t)$ be the number of jumps in the time interval $[0, t]$, i.e., $N(t) = \max\{n | \tau_n \leq t\}$. Then $\{T_k, k \geq 1\}$ is a sequence of i.i.d. random variable with the exponentially distribution:

$$P\{T_k > t | X(0) = x\} = P_x\{T_k > t\} = e^{-t}$$

for all $k \geq 1$ and $t \geq 0$, and therefore, $N(t)$ is subject to Poisson distribution with rate 1, that is,

$$P\{N(t) = n | X(0) = x\} = P_x\{N(t) = n\} = e^{-t} \frac{t^n}{n!}, n \geq 0,$$

for $x \in E_0$.

Note that $\tau_{N(t)} \leq t < \tau_{N(t)+1}$ and then

$$\frac{\sum_{k=1}^{N(t)} T_k}{N(t)} \leq \frac{t}{N(t)} < \frac{\sum_{k=1}^{N(t)+1} T_k}{N(t)}.$$

It follows from the strong law of large number that

$$\lim_{t \rightarrow \infty} \frac{N(t)}{t} = \frac{1}{\mathbb{E}_x(T_k)} = 1, \text{ a.s. } -P_x. \quad (2)$$

The continuous-time transition probabilities

$$P(t; x, y) = P\{X(t) = y | X(0) = x\} = P_x\{X(t) = y\}$$

is generated by a linear bounded operator $q(x, y) = p(x, y) - \delta_{x, y}$, which is called the infinitesimal transition rates, since

$$P(t; x, y) = q(x, y)t + o(t).$$

The transition probabilities $P(t; x, y)$ are differentiable and they satisfy the Kolmogorov differential equations:

$$P'(t; x, y) = \sum_{z \in E_0} q(x, z)P'(t; z, y) = \sum_{z \in E_0} P'(t; x, z)q(z, y).$$

Let $f : E_0 \rightarrow \mathbb{R}$ be a Lipschitz function: i.e. there is a constant M , such that for all $x, y \in E_0$, we have $|f(x) - f(y)| \leq M\|x - y\|$. Then $f(X(t))$ is a stochastic process, and

$$\begin{aligned} \mathbb{E}'_x[f(X(t))] &= \sum_y P'(t; x, y)f(y) \\ &= \sum_{y, z} P(t; x, z)q(z, y)f(y) = \mathbb{E}_x[Qf(X(t))] \end{aligned} \quad (3)$$

where

$$Qf(x) = \sum_{y \in E_0} q(x, y)f(y) = \sum_{y \in E_0} p(x, y)[f(y) - f(x)].$$

Let $\sigma_t = \sigma\{X(s) | 0 \leq s \leq t\}$ be the smallest σ -algebra which includes all events occurred before t , f a Lipschitz function on E_0 . Then fix $t > 0$, for all $0 \leq s \leq t$, the stochastic process $Y_f(s) = \mathbb{E}_x[f(X(t)) | \sigma_s]$ is a martingale. The analogous to the Azuma inequality [2, 11] for the discrete time process is the following theorem.

Theorem 1 *Let $\{X(t), t \geq 0\}$ be a graph valued process. For any fixed $t > 0$, let $Y_f(s) = \mathbb{E}_x[f(X(t)) | \sigma_s]$ ($0 \leq s \leq t$), where $x \in E_0$ and f is Lipschitz. If there is a constant $a > 0$, such that $\mathbb{E}_x[(Y_f(s_1) - Y_f(s_2))^2] \leq a$ for any $|s_1 - s_2| \leq 1$, then*

$$\mathbb{E}_x \left[\left(\frac{f(X(t))}{t} - \frac{\mathbb{E}_x[f(X(t))]}{t} \right)^2 \right] \leq \frac{a(t+1)}{t^2} \quad (4)$$

Proof. Since $Y_f(s), 0 \leq s \leq t$ is a martingale, therefore, for all $0 \leq s_1 \leq s_2 \leq s_3 \leq s_4 \leq t$, we have

$$\begin{aligned} \mathbb{E}_x[(Y_f(s_4) - Y_f(s_3))(Y_f(s_2) - Y_f(s_1))] &= \\ &= \mathbb{E}_x[(Y_f(s_2) - Y_f(s_1))\mathbb{E}_x[Y_f(s_4) - Y_f(s_3) | \sigma_{s_3}]] = 0. \end{aligned}$$

So we get required inequality,

$$\begin{aligned} \mathbb{E}_x[(f(X(t)) - \mathbb{E}_x[f(X(t))])^2] &= \mathbb{E}_x[(Y_f(t) - Y_f(0))^2] = \\ &= \mathbb{E}_x \left[\left(Y_f(t) - Y_f([t]) + \sum_{k=0}^{[t]-1} (Y_f(k+1) - Y_f(k)) \right)^2 \right] = \\ &= \mathbb{E}_x[(Y_f(t) - Y_f([t]))^2] + \sum_{k=0}^{[t]-1} \mathbb{E}_x[(Y_f(k+1) - Y_f(k))^2] \leq a(t+1). \end{aligned}$$

□

As a corollary, we have

Corollary 2 *If $\{X(t), t \geq 0\}, x \in E_0, f : E_0 \rightarrow \mathbb{R}$ as in Theorem 1. $Y_f(s) = \mathbb{E}_x[f(X(t)) | \sigma_s]$ satisfies $\mathbb{E}_x[(Y_f(s_1) - Y_f(s_2))^2] \leq a$ for any $|s_1 - s_2| \leq 1$, and in addition, if the limit*

$$\lim_{t \rightarrow \infty} \frac{\mathbb{E}_x[f(X(t))]}{t} = A,$$

then

$$\lim_{t \rightarrow \infty} \frac{f(X(t))}{t} = A \text{ a.s. } -P_x$$

§2 CONTINUOUS TIME BARABÁSI-ALBERT MODEL

In this section we will apply the graph valued process to discuss the degree distribution in some continuous time preferential growth network models. The model of growth networks is as follows: starting with an initial state $x \in E_0$, at every one-step jump, we add a new node and one edge that links the new node to one of the nodes already present in the graph.

Let $e(i, j) \in \{0, 1\}^{\mathbb{N}^2}$, $i \neq j$ defined by $e_{kl}(i, j) = \delta_{ki}\delta_{lj} + \delta_{kj}\delta_{li}$. Then for $x \in E_0$, the one-step jump probability matrix

$$p(x, y) = \begin{cases} \frac{W(d_i(x))}{\sum_{j=1}^{n_x} W(d_j(x))} & n_y = n_x + 1, y_{kl} = x_{kl} + e_{kl}(i, n_x + 1) \text{ and } 1 \leq i \leq n_x \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where $W(d)$ is a positive increasing function, which describes the preferential attachment. This is our graph valued process description of the preferential growth networks.

Let $S(x) = \sum_{i=1}^{n_x} W(d_i(x))$, and as before $D_k(x) = \#\{i | d_i = k\}$ denotes the number of nodes with degree k .

Theorem 3 *Suppose that there is a constant c such that $\sup\{W(k+1) - W(k) : k \geq 1\} \leq c$ and*

$$\lim_{t \rightarrow \infty} \frac{\mathbb{E}_x[S(X(t))]}{t} = S > 0. \quad (6)$$

Then, the probability that a node has the degree k in the graph valued process $X(t)$ is

$$P_k = \lim_{t \rightarrow \infty} \frac{D_k(X(t))}{N(t)} = \frac{S}{W(k) + S} \prod_{i=1}^{k-1} \frac{W(i)}{W(i) + S}, \quad a.s.-P_x. \quad (7)$$

Proof. For any fixed $t > 0$, let $Y_S(s) = \mathbb{E}_x[S(X(t))|\sigma_s]$ for $0 \leq s \leq t$. Since for all $|s_1 - s_2| \leq 1$, we have

$$\begin{aligned} \mathbb{E}_x[(Y_S(s_1) - Y_S(s_2))^2] &\leq \mathbb{E}_x[(\sup\{|W(k+1) - W(k)|\})N(|s_1 - s_2|)^2] \leq \\ &\leq c^2 \mathbb{E}_x[N^2(1)] = 2c^2, \end{aligned}$$

and $N(t)/t \rightarrow 1$ a.s.- P_x as $t \rightarrow \infty$, it follows from Corollary 2 that

$$\lim_{t \rightarrow \infty} \frac{S(X(t))}{t} = S, \quad a.s.-P_x. \quad (8)$$

On the other hand, for all $t \geq 0$, we have

$$\mathbb{E}'_x[D_k(X(t))] = W(k-1)\mathbb{E}_x\left[\frac{D_{k-1}(X(t))}{S(X(t))}\right] - W(k)\mathbb{E}_x\left[\frac{D_k(X(t))}{S(X(t))}\right] + \delta_{k,1}. \quad (9)$$

Let $D_k(t) = \mathbb{E}_x[D_k(X(t))]$. By (8) the equation (9) can be rewritten as

$$D'_k(t) = \frac{W(k-1)D_{k-1}(t)}{St} - \frac{W(k)D_k(t)}{St} + \delta_{k,1} + \epsilon_k(t), \quad (10)$$

where

$$\begin{aligned} \epsilon_k(t) = W(k-1)\mathbb{E}_x \left[D_{k-1}(X(t)) \frac{S - S(X(t))/t}{S(X(t))S} \right] - \\ - W(k)\mathbb{E}_x \left[D_k(X(t)) \frac{S - S(X(t))/t}{S(X(t))S} \right] = o(1). \end{aligned} \quad (11)$$

Solving the differential equation (10) we get

$$D_k(t) = t \frac{S}{W(k) + S} \prod_{i=1}^{k-1} \frac{W(i)}{W(i) + S} + o(1)t.$$

So we have

$$\lim_{t \rightarrow \infty} \frac{\mathbb{E}_x[D_k(X(t))]}{t} = \frac{S}{W(k) + S} \prod_{i=1}^{k-1} \frac{W(i)}{W(i) + S}.$$

Let $Y_{D_k}(s) = \mathbb{E}_x[D_k(X(t))|\sigma_s]$. It is easy to check that, for any $|s_1 - s_2| \leq 1$,

$$\mathbb{E}_x[(Y_{D_k}(s_1) - Y_{D_k}(s_2))^2] \leq 4\mathbb{E}_x[N^2(1)] = 8.$$

Thus, by Corollary 2, we can conclude that the probability that a node has the degree k is given by (7). \square

Remark 4 *In fact, the equation (10) is just the rate equation introduced by Krapivsky, Redner, and Leyvraz in [15], which describes the rate of the average number of nodes with k edges at time.*

As an application of Theorem 3, we consider the following three examples of evolving networks.

Example 5 (*Uniform attachment*)

Let $W(k) \equiv 1$, we have

$$\lim_{t \rightarrow \infty} \frac{\mathbb{E}_x[S(X(t))]}{t} = \lim_{t \rightarrow \infty} \frac{\mathbb{E}_x[N(t)]}{t} = 1, \quad (12)$$

i.e., $S = 1$, and therefore, by (7) the probability that a node has the degree k is

$$P_k = \frac{1}{2^k}. \quad (13)$$

Example 6 (*Barabási-Albert model with initial attractiveness*)

Let $W(k) = k + \beta$ with $\beta > -1$. Then

$$S = \lim_{t \rightarrow \infty} \frac{\mathbb{E}_x[S(X(t))]}{t} = \lim_{t \rightarrow \infty} \frac{\mathbb{E}_x[2N(t) + \beta(N(t) + 1)]}{t} = 2 + \beta. \quad (14)$$

By (7), it follows that

$$P_k = \frac{2 + \beta}{k + 2(\beta + 1)} \prod_{i=1}^{k-1} \frac{i + \beta}{i + 2(\beta + 1)} \sim (2 + \beta)e^{-\beta} \frac{\Gamma(2\beta + 3)}{\Gamma(\beta + 1)} \frac{1}{k^{3+\beta}}$$

for large k . This degree distribution agrees with the result in [7]

Example 7 (*Sub-linear preferential attachment*)

Let $W(k) = k^\alpha$, $0 < \alpha < 1$, and suppose that the limit

$$S = \lim_{t \rightarrow \infty} \frac{\mathbb{E}_x[S(X(t))]}{t} \quad (15)$$

exists. Although it is likely that the limit exists, it has not been verified yet. Since $1 \leq W(k) = k^\alpha \leq k$, therefore, $N(t) \leq S(X(t)) \leq 2N(t)$. It follows that $1 \leq S \leq 2$, if the limit (15) exists. Then by (7) we have

$$P_k \sim \frac{S}{k^\alpha} \exp \left\{ -\frac{Sk^{1-\alpha}}{1-\alpha} \right\} \quad (16)$$

for large k .

Notice that, in [8], the same asymptotic formula (16) has been obtained but under a more strong assumption: both (15) and all limits $\lim_{t \rightarrow \infty} \mathbb{E}_x[D_k(X(t))]/t$ exist.

References

- [1] R. Albert & A.L. Barabasi, Statistical Mechanics of Complex Networks, *Rev. Mod. Phys.*, **74** (2002), 47.
- [2] K. Azuma. Weighted sums of certain dependent random variables, *Tohoku Math. J.*, **19** (1967), 357.
- [3] A.L. Barabási & R. Albert, Emergence of scaling in random networks, *Science*, **286** (1999), 509.
- [4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez & D.U. Hwang, Complex networks: Structure and dynamics, *Phys. Reports*, **424** (2006), 175.
- [5] B. Bollobás & O. Riordan, Robustness and vulnerability of scale-free random graphs *Internet Math.* **1**, (2003), 1.
- [6] B. Bollobás & O. Riordan, Mathematical results on scale-free random graphs, *Handbook of Graphs and Networks*, Wiley-VCH, Berlin, 2002.
- [7] B. Bollobás, O. Riordan, J. Spencer & G. Tusnady, The degree sequence of a scale-free random graph process, *Random Structures and Algorithms*, **18** (2001), 279.
- [8] F. Chung, S. Handjani & D. Jungreis, Generalizations of Polya's urn problem, *Ann. Combin.*, **7** (2003), 141.
- [9] R. Diestel, *Graph Theory*, Springer-Verlag, Electronic Edition 2000.

- [10] S.N. Dorogovtsev & J.F.F. Mendes, Evolution of networks, *Adv. Phys.*, **51** (2002), 1079.
- [11] W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Assoc.*, **58** (1963), 13.
- [12] L.H. Hartwell, J.J. Hopfield, S. Leibler & A.W. Murray, From molecular to modular cell biology, *Nature*, **402** (1999), 47.
- [13] H. Jeong, S.P. Mason & A.L. Barabási, Lethality and centrality in protein networks, *Nature*, **411** (2001), 41.
- [14] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai & A.L. Barabási, The large-scale organization of metabolic networks, *Nature*, **407** (2000), 651.
- [15] P.L. Krapivsky, S. Redner & F. Leyvraz, Connectivity of growing random networks, *Phys. Rev. Lett.*, **85** (2000), 4629.
- [16] Z.M. Ma, Random networks and graph valued Markov processes, *Invited talk on the China-Japan Joint Conference on Discrete Geometry, Combinatorics and Graph Theory*, Nov 18-24, 2005.
- [17] M.E.J. Newman, The structure and function of complex networks, *SIAM Rev.* **45** (2003), 167.
- [18] R.V. Solé, R. Ferrer-Cancho, J.M. Montoya & S. Valverde, Selection, tinkering, and emergence in complex networks, *Complexity*, **8** (2003), 20.
- [19] A. Wagner, How the global structure of protein interaction networks evolves, *Proc. R. Soc. Lond.*, **B**, 270 (2003), 457.
- [20] A. Wagner & D.A. Fell, The small world inside large metabolic networks, *Proc. R. Soc. Lond.*, **B**, 268 (2001), 1803.
- [21] D.J. Watts & S.H. Strogatz, Collective dynamics of 'small-world' networks, *Nature*, **393** (1998), 440.