Scale-Free Directed Networks from Threshold Model with Two Intrinsic Node Variables

Shigeo SHIODA
Urban Environment and Systems, Faculty of Engineering, Chiba University
1-33 Yayoi, Inage, Chiba 263-8522, Japan
Tel/Fax: +81-43-290-3237. E-mail: shioda@faculty.chiba-u.jp

1 Introduction

Complex networks such as the world-wide web (WWW) or social networks are often self-organized by the actions of a large number of individual components. Although the local interactions between individual components in real complex networks would be diverse and complex, real complex networks have common simple topological properties like small-world property, scale-free property, or clustering. Various network models have been proposed in order to understand how these properties arise.

Real complex networks can be categorized into two types: undirected and directed. There are several directed complex networks, such as the WWW, e-mail networks, or citation networks, which exhibit interesting small-world or scale-free property. The aim of this work is to provide a network model for directed networks to understand the topological properties typical to complex networks. For this purpose, we use a non-growing network model proposed by Caldarelli *et al.* [1] and Masuda *et al.* [2]. In their model, which we call the *threshold model* in this paper, each node has an intrinsic variable, which is referred to as *weight*. The connectivity between a pair of nodes is determined by whether the sum of weights of the pair exceeds a given threshold.

In this work, we make a very simple extension to the threshold model; each node has two intrinsic variables (outgoing weight and incoming weight). A directed link is established from node i to node j if the sum of outgoing weight of node i and the incoming weight of node j exceeds a predetermined threshold.

Surprisingly, this simple extension gives much richer characteristics to the model. For example, in the proposed model, the distributions of out-degree and in-degree respectively have power-law tails, and their scaling exponent are controllable in the range $(1, \infty)$. Note that the scaling exponent of the original threshold model is fixed at 2. The average clustering coefficient of nodes with out-degree (or in-degree) n in the proposed model also has a power-law tail as a function of n, and its scaling exponent is variable depending on the correlation between two node weights. This also contrasts with the original threshold model, whose clustering coefficient has constant scaling exponent 2. The proposed model is highly analytically tractable, and thus we can analytically find the above mentioned characteristics without using numerical methods.

2 Proposed Model

2.1 Link establishing mechanism

Each node has two intrinsic variables. One variable, which is referred to as *outgoing weight* in this paper, represents the potential of the node to establish directed links to other nodes, and the other, which is referred to as *incoming weight*, represents the potential of the node to have directed links from other nodes. We let w_i^{out} and w_i^{in} respectively denote of the incoming and outgoing weights of node i. The outgoing (incoming) weight of a node is randomly and independently assigned based on the (cumulative) distribution function $F(x)$ ($G(x)$). We assume that the directed link from node i to node j is established if and only if $w_i^{out} + w_j^{in} \geq \theta$.

2.2 Degree distribution

In this work, the number of outgoing (incoming) links of a node is referred to as *out-degree* (*in-degree*) of the node. If the outgoing and incoming weights follow exponential distributions such as

$$F(x) = 1 - e^{-\lambda_1 x}, \quad G(x) = 1 - e^{-\lambda_2 x}, \tag{1}$$

then the complementary cumulative distribution function (CCDF) of the out-degree, $P_c^{out}(n)$, has the following analytical expression;

$$P_c^{out}(n) = e^{-\lambda_1\theta}\left(\frac{N}{n}\right)^{\lambda_1/\lambda_2}.$$

This expression shows that the out-degree has a power-law distribution and its scaling exponent, γ_d^{out}, is equal to $(\lambda_1+\lambda_2)/\lambda_2$. Since

$$1 < \gamma_d^{out} = \frac{\lambda_1+\lambda_2}{\lambda_2} < \infty,$$

the scaling exponent of the out-degree distribution in our model could take arbitrary values larger than 1. We also find that the CCDF of the in-degree $P_c^{in}(n)$ has the following analytical expression:

$$P_c^{in}(n) = e^{-\lambda_2\theta}\left(\frac{N}{n}\right)^{\lambda_2/\lambda_1}.$$

That is, the in-degree also has a power-law distribution its scaling exponent γ_d^{in} is equal to $(\lambda_1+\lambda_2)/\lambda_1$. There is a following relationship between γ_d^{out} and γ_d^{in}:

$$\gamma_d^{in} = \frac{\gamma_d^{out}}{\gamma_d^{out}-1}, \quad \gamma_d^{out} = \frac{\gamma_d^{in}}{\gamma_d^{in}-1}.$$

In particular

$$\gamma_d^{in} > 2 \Leftrightarrow \gamma_d^{out} < 2, \quad \gamma_d^{in} < 2 \Leftrightarrow \gamma_d^{out} > 2.$$

2.3 Clustering coefficient

First we assume that incoming and outgoing weights of a node are independent from each other. If the outgoing and incoming weights follow exponential distributions, then the average clustering coefficient of nodes with out-degree n, $C_{out}(n)$, has the following analytical expression;

$$C_{out}(n) = \begin{cases} \frac{\lambda_1}{\lambda_1-\lambda_2}e^{-\lambda_2\theta}\frac{N}{n} + \frac{\lambda_2}{\lambda_2-\lambda_1}e^{-\lambda_1\theta}\left(\frac{N}{n}\right)^{\lambda_1/\lambda_2} & \lambda_1 \neq \lambda_2, \\ e^{-\lambda_1\theta}\frac{N}{n}\left(1+\log\frac{n}{N}+\lambda_1\theta\right) & \lambda_1 = \lambda_2. \end{cases} \tag{2}$$

This result indicates that the clustering coefficient with respect to the out-degree, $C_{out}(n)$, follows a power-law distribution and its scaling exponent γ_c^{out} is

$$\gamma_c^{out} = \begin{cases} 1 & \lambda_1 \geq \lambda_2, \\ \lambda_1/\lambda_2 & \lambda_1 < \lambda_2, \end{cases}$$

confirming that $\gamma_c^{out} \leq 1$. We also find that

$$C_{in}(n) = \begin{cases} \frac{\lambda_2}{\lambda_2-\lambda_1}e^{-\lambda_1\theta}\frac{N}{n} + \frac{\lambda_1}{\lambda_1-\lambda_2}e^{-\lambda_2\theta}\left(\frac{N}{n}\right)^{\lambda_2/\lambda_1} & \lambda_1 \neq \lambda_2, \\ e^{-\lambda_1\theta}\frac{N}{n}\left(1+\log\frac{n}{N}+\lambda_1\theta\right) & \lambda_1 = \lambda_2, \end{cases} \tag{3}$$

and thus the scaling exponent of $C_{in}(n)$, γ_c^{out}, is given by

$$\gamma_c^{in} = \begin{cases} \lambda_2/\lambda_1 & \lambda_1 \geq \lambda_2, \\ 1 & \lambda_1 < \lambda_2, \end{cases}$$

also confirming that $\gamma_c^{in} \leq 1$.

If incoming and outgoing weights of a node are correlated, the clustering coefficients, $C_{out}(n)$ and $C_{in}(n)$, have different analytical expressions with (2) and (3). Here, let us consider the case where the correlation coefficient between incoming and outgoing weights is one. In this case

$$\begin{aligned} &C_{out}(n) = C_{in}(n) \\ &= \begin{cases} \frac{\lambda_2}{\lambda_2-\lambda_1}e^{-\lambda_1\theta}\left(\frac{N}{n}\right)^{\frac{\lambda_1+\lambda_2}{\lambda_2}} + \frac{\lambda_1}{\lambda_1-\lambda_2}e^{-\lambda_2\theta}\left(\frac{N}{n}\right)^{\frac{\lambda_1+\lambda_2}{\lambda_1}} & \lambda_1 \neq \lambda_2, \\ \frac{N^2}{n^2}e^{-\lambda_1\theta}\left(1+2\log\frac{n}{N}+\lambda_1\theta\right) & \lambda_1 = \lambda_2, \end{cases} \end{aligned}$$

and thus

$$\gamma_c^{out} = \gamma_c^{in} = \min\{\gamma_d^{in}, \gamma_d^{out}\},$$

indicating that $1 < \gamma_c^{out} = \gamma_c^{in} \leq 2$.

3 Conclusion

In this work, we propose a scale-free network model for directed networks. For the lack of space, in this article, we only show results of the case where the incoming and outgoing weights are exponentially distributed, but we can derive analytical expressions for the degree distribution and for the clustering coefficient when both weights follow more general distributions.

References

[1] G. Caldarelli, A. Capocci, P. D. L. Rios, and M. Munoz. Phys. Rev. Lett., **89**, 258702 (2002).

[2] N. Masuda, H. Miwa, and N. Konno. Phys. Rev. E, **70**, 036124 (2004).