Weighted scale-free networks with variable power-law exponents

Takuma Tanaka

Department of Morphological Brain Science, Graduate School of Medicine, Kyoto University, Japan, Fax: +81 75 753 7370, email:ttakuma@mbs.med.kyoto-u.ac.jp

Toshio Aoyagi

Department of Applied Analysis and Complex Dynamical Systems, Graduate School of Informatics, Kyoto University, Japan

CREST, JST

Abstract

We present a weighted scale-free network model, in which the power-law exponents can be controlled by the model parameters. The network is generated through the weight-driven preferential attachment of new nodes to existing nodes and the growth of the weights of existing links. The simplicity of the model enables us to derive analytically the various statistical properties, such as the distribution of degree, strength, and weight, the degree-strength and degree-weight relationship, and the dependencies of these power-law exponents on the model parameters. Finally, we demonstrate that networks of words, coauthorship of researchers, and collaboration of actor/actresses are quantitatively well described by this model.

Key words: Weighted scale-free network ; Social network ; Self-organization

1 Introduction

Understanding the evolution and dynamics on networks has become essential to comprehend physical, biological, and sociological phenomena in our world. We ourselves live in and interact with society, which consists of people connected through relationships, that is, links, such as friendship, kinship, economic interaction, and other forms of interpersonal ties. Our society thus constitutes a network in which individuals and the connections among them can be regarded as nodes and links, respectively. Many types of real-world networks including the coauthorship of researchers [1], collaboration of actors/actresses [2], biochemical reaction in the cell [3], WWW [4], Internet [5], and words in a text [6] have been studied in physical literatures. These networks are known to belong to a class of scale-free networks. 'Scale-free' means that the distribution of the degree, i.e. the link number of the nodes, in these networks obeys a power-law and there is no characteristic degree number.

A number of papers have been published on models generating scale-free networks. The best known is the Barabási-Albert model [7], which is characterized by growth and preferential attachment. The basic idea is that the nodes with large degrees gain new links faster than the nodes with small degrees. The growth and the preferential attachment are such central ideas in the scale-free network models [8] that a large number of models of scale-free networks are based on these two mechanisms [9,10,11]. Moreover, these mechanisms have been proven to be capable of making power-law distributions since the mid-1950's [12,13]. Thus, growth and preferential attachment give a simple and powerful explanation for how these scale-free networks in our society are formed, although some other models generate the scale-free network without growth using the degree-dependent rewiring of the links [14] and the fixed fitness of the nodes [15,16].

A large body of work exists on the dynamics of scale-free networks [17], such as the spread of infectious diseases [18], opinion formation [19], strategic games [20], packet transfer [21], and synchronization of oscillators [22]. The dynamics on the network are, however, determined not only by the topological structure of the network but also by the connection strengths between the nodes. Taking the spread of the disease as an example, it is evident that the strength of the link between nodes plays a very important role in determining whether or not the disease is transmitted from one person to another. If the weight of a link is small, it is unlikely that the disease transmits itself through the link, whereas it is highly probable that the disease transmits itself if the weight is large. Hence, much effort has been focused on the examination of the statistical properties of the weighted scale-free networks [23,24] and on the modeling of the process generating them [25,26,27,28]. The weights in earlier weighted scale-free network models are determined by the degrees of the nodes [25] or their fitnesses [26,28]. In the Barrat-Barthélemy-Vespignani model [27], the weights of the links grow only when a node at the end of this link gains a new link. Although these models are interesting from the theoretical point of view, the weights in the real-world networks often grow without making new edges and are less dependent on the degrees. Taking account of these points, we will attempt to construct a simple model with the following properties in this study: (i) growth through weight-driven preferential attachment; (ii) link weights can grow larger spontaneously, that is, without making new links; (iii) growth determined by the strength of the node, not by the fitnesses of the nodes; and (iv) power-law exponents of the networks can be easily changed in controlling the model parameters. Finally, we found that our model can reproduce the degree, strength, and weight distributions of the real-world scale-free networks.

The model of the present paper is an extension of the earlier model [29].

2 Model

Before introducing our model, we define some measures to characterize weighted networks. First, the connectivity of a network can be expressed by an adjacency matrix a_{ij}, whose elements take the value 1 if the node i is connected to the node j and 0 otherwise. The degree of node i is then defined by $k_i = \sum_{j=1}^{N} a_{ij}$, where N is the total number of nodes. In addition, the weight of the link between nodes i and j is denoted by w_{ij}. Let us define the strength of node i as $s_i = \sum_{j=1}^{N} a_{ij} w_{ij}$, which is the sum of the weights of all the links connecting to node i. In this model, we assume that the links are undirected, so that the adjacency matrix a_{ij} and the weight matrix w_{ij} are symmetric.

We schematically present a set of rules for generating the network as follows (Fig. 1). The network initially starts with a single node. Rule 1: at each time step, a new node is added to the network and connections are made to m existing nodes, where the probability that the node i is chosen is proportional to $s_i + \sigma$, i.e. $(s_i+\sigma) \Big/ \sum_{j=1}^{N} (s_j + \sigma)$ (strength-driven preferential attachment), where σ is a constant parameter. The weight of this new link is then set to unity. Rule 2: at each time step, ct^{μ} pairs of the existing nodes are selected with a probability proportional to $s_i+\sigma$, i.e. $(s_i+\sigma)(s_j+\sigma) \Big/ \left(\sum_{k=1}^{N} s_k + \sigma\right)^2$. If these two nodes are not connected, they are connected by a link with a weight equal to unity (Rule 2a). If they are already connected, the weight of the corresponding link between them is incremented by one (Rule 2b). We assume $\sigma > -m$ to enable the new node to gain links in further steps because the strength of the new node is m. This rule can be regarded as a generalization of the rule in the word web growth [30]. Note that starting from a network with m nodes at time $t = m$, the total number of nodes is equal to the time t and each node can be labeled by the time u when the node is added.

As we will see in the following, the distribution of the strength, degree, and weight and the relationship among them are determined by the parameters c, m, μ, and σ. Although the power-law exponents of these distributions are determined by the parameters μ, m, and σ as we show in the following, we focus on c as an example, because c changes the structure of the network drastically. Fig. 2 shows how drastically c changes the structure of the whole network generated by the present model. The networks shown in Fig. 2 are generated by the model with parameters $c = 0.5$, $m = 1$, $\mu = 1$, and $\sigma = 0$ (left) and $c = 0.0001$, $m = 1$, $\mu = 1$, and $\sigma = 0$ (right). Although these networks share all parameters other than c and have the same number of nodes (50), they are completely different in their structure. The network with large c looks very complicated and has many links with weight larger than one (thick lines). This is because the large c tends to increase the connections between the existing nodes. On the other hand, the network with small c has a simpler, tree-like structure, because connections between the existing nodes are rare. In what follows, we will investigate various properties that our simple model exhibits.

3 Theoretical analysis

3.1 Strength distribution

To analytically obtain the statistical properties of the network generated by the above algorithm, we use a continuous approximation. Now, let us denote the averaged strength of the node at time t by $s(u,t)$, where u is the time at which this node is added to the network. Proceeding in the same way as [30], we describe the time evolution of $s(u,t)$ by

$$\frac{\partial s(u,t)}{\partial t} = (m + 2ct^{\mu})\frac{s(u,t)+\sigma}{\int_0^t dv\, s(v,t)+\sigma} \tag{1}$$

with boundary condition $s(t,t) = m$ since the node born at time t makes m connections to the existing nodes at that time. This equation means that $m+2ct^{\mu}$ ends of new edges are distributed through the preferential attachment at each time step. Total strength of the network is given by $\int_0^t dv\, s(v,t) = 2mt + \dfrac{2ct^{\mu+1}}{\mu+1}$.

Because the general model is not analytically tractable, we restrict ourselves to two cases: $\sigma = 0$ and $\mu = 1$. Setting $\sigma = 0$, we obtain the solution

$$s(u,t) = \sqrt{\frac{t(m\mu + m + ct^{\mu})^{2+\frac{1}{\mu}}}{u(m\mu + m + cu^{\mu})^{2+\frac{1}{\mu}}}}. \tag{2}$$

Note that if we set $\mu = 1$ and $m = 1$, the strength of the node born at time u has a simple form

$$s(u,t) = \sqrt{\frac{t(2+ct)^3}{u(2+cu)^3}}, \tag{3}$$

which was obtained by Dorogovtsev and Mendes [30]. To obtain the strength distribution from Eq. 2 we use the well-known equation

$$P(s) = \frac{1}{t}\left|\frac{\partial s(u,t)}{\partial u}\right|^{-1}.$$

The strength distribution $P(s)$ takes the form $P(s) \approx \frac{2(m\mu + m + ct^{\mu})^{2+1/\mu}}{(m\mu + m)^{4+2/\mu}} s^{-3}$ for $cu^{\mu} \ll 1$ because $s(u,t)$ is approximated by $\sqrt{\frac{t(m\mu + m + ct^{\mu})^{2+1/\mu}}{u(m\mu + m)^{2+1/\mu}}}$. For $cu^{\mu} \gg 1$, the approximation $s(u,t) \approx \sqrt{\frac{t(2+ct)^{2+1/\mu}}{u^{2+2\mu}c^{2+1/\mu}}}$ gives $P(s) \approx \frac{1}{1+\mu} s^{-\frac{2+\mu}{1+\mu}}$. Thus, the distribution of the strength for $\sigma = 0$ has two regions with different exponents, $-\frac{2+\mu}{1+\mu}$ and -3, separated by the crossover point $s_c = \left[\frac{2(1+\mu)(m\mu + m + ct^{\mu})^{2+1/\mu}}{(m\mu + m)^{4+2/\mu}}\right]^{(1+\mu)/(1+2\mu)}$. Fig. 3 shows the comparison of distribution of strength between theoretical and numerical results, where it is seen that the exponents obtained in the simulations agree well with theoretical ones.

When $\mu = 1$, the solution of Eq. 1 is

$$s(u,t) = -\sigma + (m+\sigma)\frac{t^{\frac{m}{2m+\sigma}}(2m+\sigma+ct)^{\frac{3m+2\sigma}{2m+\sigma}}}{u^{\frac{m}{2m+\sigma}}(2m+\sigma+cu)^{\frac{3m+2\sigma}{2m+\sigma}}}. \tag{4}$$

The strength distribution takes the form

$$P(s) \approx \frac{(2m+\sigma+ct)^{(3m+2\sigma)/m}(m+\sigma)^{(2m+\sigma)/m}}{m(2m+\sigma)^{2(m+\sigma)/m}} \times (s+\sigma)^{-(3m+\sigma)/m}$$

for $cu \ll 1$ and

$$P(s) \approx \frac{\sqrt{m+\sigma}}{2}(s+\sigma)^{-3/2}.$$

for $cu \gg 1$. Hence, the strength distribution for $\mu = 1$ also has two regions with exponents $-3/2$ and $-(3m+\sigma)/m$, and they crossover at $s_c = \frac{(m+\sigma)(2m+\sigma+ct)^2 2^{2m/(3m+2\sigma)}}{m^{2m/(3m+2\sigma)}(2m+\sigma)^{4(m+\sigma)/(3m+2\sigma)}} - \sigma$. Fig. 4 shows the comparison of strength distribution between theoretical and numerical results for $\mu = 1$.

Fig. 5 shows the dependency of the power-law exponent of the strength distribution on the parameters μ, m, and σ. If $\sigma = 0$, the strength distribution $P(s)$

obeys the exponent -3 for $cu^\mu \ll 1$ and the exponent $-\frac{2+\mu}{1+\mu}$ for $cu^\mu \gg 1$. The exponent for $cu^\mu \gg 1$ ranges from 1 to 2. If $\mu = 1$, the strength distribution obeys the exponent $-\frac{3m+\sigma}{m}$ for $cu \ll 1$ and $-3/2$ for $cu \gg 1$. Since we assumed that the inequality $\sigma > -m$ holds, the exponent of the strength distribution for $cu \ll 1$ can take any value smaller than -2. Although we assumed that parameters μ, σ, and m can take any value, we restrict our analysis to the case $\mu = 1$, $\sigma = 0$, and $m = 1$ hereafter and present only the numerical results, because the other cases are rather complicated and not analytically tractable except for the strength distribution.

3.2 Degree distribution

In order to obtain the degree distribution analytically, we limit ourselves to the model with $\mu = 1$, $\sigma = 0$, and $m = 1$. We consider a continuous version of the adjacency matrix a_{ij} as we did with s_i. Let us consider the averaged connectivity of the nodes at time t, $a(u_1, u_2, t)$, where two nodes at each end of the link are added at time u_1 and u_2. The connectivity $a(u_1, u_2, t)$ satisfies the differential equation

$$\frac{\partial a(u_1,u_2,t)}{\partial t} = 2ct\frac{s(u_1,t)s(u_2,t)}{\left(\int_0^t dv\, s(v,t)\right)^2}[1-a(u_1,u_2,t)],$$

which has a general solution

$$a(u_1,u_2,t) = 1 - \exp\left(-\frac{(2+ct)^2}{\sqrt{u_1u_2(2+cu_1)^3(2+cu_2)^3}}\right) \times F(u_1,u_2), \quad (5)$$

where $F(u_1, u_2)$ is an arbitrary function. Although the boundary condition

$$a(t,u,t) = a(u,t,t) = \frac{s(u,t)}{\int_0^t dv\, s(v,t)} = \sqrt{\frac{2+ct}{tu(2+cu)^3}}$$

cannot be satisfied, we attempt to approximately satisfy it by setting $F(u_1, u_2) = 1$ and Taylor expanding to obtain

$$\begin{aligned} a(t,u,t) = a(u,t,t) &= \sqrt{\frac{2+ct}{tu(2+cu)^3}} + \cdots \\ &= \frac{s(u,t)}{\int_0^t dv\, s(v,t)}. \end{aligned}$$

Hence, Eq. 5 is an approximate solution of connectivity if $F(u_1, u_2) = 1$.

In the same way, the average degree $k(u,t)$ is defined by

$$k(u,t) = \int_0^t dv\, a(u,v,t).$$

Splitting the integral into two regions $cu \ll 1$ and $cu \gg 1$,

$$\begin{aligned}
k(u,t) =& \int_0^t dv\, a(u,v,t) \\
\approx& \int_0^{2/c} dv \left[1 - \exp\left(-\frac{(2+ct)^2}{\sqrt{uv(2+cu)^3 2^3}}\right)\right] \\
&+ \int_{2/c}^t dv \left[1 - \exp\left(-\frac{(2+ct)^2}{\sqrt{uv(2+cu)^3 (cv)^3}}\right)\right] \\
=& t - \int_0^{2/c} dv\ \exp\left(-A/\sqrt{8v}\right) \\
&- \int_{2/c}^t dv \exp\left[-A/\left(\sqrt{c^3}v^2\right)\right],
\end{aligned} \tag{6}$$

where $A = (2+ct)^2/\sqrt{u(2+cu)^3}$. The first integral of Eq. 6 reduces to

$$\begin{aligned}
&\int_0^{2/c} dv\ \exp\left(-\frac{A}{\sqrt{8v}}\right) \\
=& \left(\frac{2}{c} - \frac{A}{2\sqrt{c}}\right) \exp\left(-\frac{\sqrt{c}A}{4}\right) + \frac{A^2}{8}\Gamma\left(0, \frac{\sqrt{c}A}{4}\right),
\end{aligned} \tag{7}$$

where $\Gamma(a,b)$ is the incomplete gamma function. To obtain the scaling behavior of this integral, we use the first-order approximation in two regimes, $\sqrt{c}A \gg 1$ and $\sqrt{c}A \ll 1$. The first term of Eq. 7 converges to zero for $\sqrt{c}A \gg 1$ and is approximated by

$$\left(\frac{2}{c} - \frac{A}{2\sqrt{c}}\right)\left(1 - \frac{\sqrt{c}A}{4}\right) \approx \frac{2}{c} - \frac{A}{\sqrt{c}}.$$

The second term of Eq. 7 $x^2\Gamma(0,x)$ converges to zero for both regimes. Eq. 7 is hence approximated by

$$\begin{cases} 0 & (\sqrt{c}A \gg 1) \\ \dfrac{2}{c} - \dfrac{A}{\sqrt{c}} & (\sqrt{c}A \ll 1). \end{cases} \tag{8}$$

The second integral of Eq. 6 reduces to

$$
\begin{aligned}
&\int_{2/c}^{t} dv \exp\left[-A/\left(\sqrt{c^3}v^2\right)\right] \\
=&\left[\exp\left(-\frac{A}{\sqrt{c^3}v^2}\right)+\sqrt{\frac{\pi A}{\sqrt{c^3}}}\mathrm{erf}\left(\sqrt{\frac{A}{\sqrt{c^3}}}\frac{1}{v}\right)\right]_{2/c}^{t} \\
=&\,t\exp\left(-\frac{A}{\sqrt{c^3}t^2}\right)+\sqrt{\frac{\pi A}{\sqrt{c^3}}}\mathrm{erf}\left(\sqrt{\frac{A}{\sqrt{c^3}}}\frac{1}{t}\right) \\
&-\frac{2}{c}\exp\left(-\frac{\sqrt{c}A}{4}\right)-\sqrt{\frac{\pi A}{\sqrt{c^3}}}\mathrm{erf}\left(\frac{\sqrt{A\sqrt{c}}}{2}\right).
\end{aligned}
\tag{9}
$$

For $\sqrt{c}A \gg 1$, Eq. 9 is approximated by

$$
\begin{aligned}
&t\left(1-\frac{A}{\sqrt{c^3}t^2}\right)+\sqrt{\frac{\pi A}{\sqrt{c^3}}}\sqrt{\frac{A}{\sqrt{c^3}}}\frac{1}{t}\frac{2}{\sqrt{\pi}}-0-\sqrt{\frac{\pi A}{\sqrt{c^3}}} \\
\approx&\, t+\frac{A}{\sqrt{c^3}t}-\sqrt{\frac{\pi A}{\sqrt{c^3}}} \\
\approx&\, t-\sqrt{\frac{\pi A}{\sqrt{c^3}}}.
\end{aligned}
$$

For $\sqrt{c}A \ll 1$, this equation is approximated by

$$
\begin{aligned}
&t\left(1-\frac{A}{\sqrt{c^3}t^2}\right)+\sqrt{\frac{\pi A}{\sqrt{c^3}}}\sqrt{\frac{A}{\sqrt{c^3}}}\frac{1}{t}\frac{2}{\sqrt{\pi}} \\
&-\frac{2}{c}\left(1-\frac{\sqrt{c}A}{4}\right)-\sqrt{\frac{\pi A}{\sqrt{c^3}}}\frac{\sqrt{A\sqrt{c}}}{2}\frac{2}{\sqrt{\pi}} \\
\approx&\, t\left(1-\frac{A}{\sqrt{c^3}t^2}\right)+\frac{2A}{\sqrt{c^3}t}-\frac{2}{c}\left(1-\frac{\sqrt{c}A}{4}\right)-\frac{A}{\sqrt{c}} \\
\approx&\, t-\frac{2}{c}-\frac{A}{2\sqrt{c}},
\end{aligned}
$$

where the last approximation uses $ct \gg 1$. Note that assuming $ct \gg 1$,

$$
\frac{A}{\sqrt{c^3}t^2} \approx \frac{c^2}{\sqrt{u(2+cu)^3}} \to 0
$$

for both regimes and $\mathrm{erf}(x) \approx \frac{2}{\sqrt{\pi}}x$ for $x \ll 1$. Eq. 9 is hence approximated by

$$\begin{cases} t - \sqrt{\dfrac{\pi A}{\sqrt{c^3}}} & (\sqrt{c}A \gg 1) \\ t - \dfrac{2}{c} - \dfrac{A}{2\sqrt{c}} & (\sqrt{c}A \ll 1). \end{cases} \tag{10}$$

Therefore, from Eqs. 6, 8, and 10 we obtain

$$k(u,t) \approx \begin{cases} \sqrt{\pi A/\sqrt{c^3}} & (\sqrt{c}A \gg 1) \\ 3A/(2\sqrt{c}) & (\sqrt{c}A \ll 1). \end{cases}$$

(i) When c is larger and $cu \gg 1$ for all u,

$$k(u,t) \approx \begin{cases} \sqrt{\dfrac{\pi}{c}}\dfrac{t}{u} & (\sqrt{c}A \gg 1) \\ \dfrac{3t^2}{2u^2} & (\sqrt{c}A \ll 1), \end{cases} \tag{11}$$

which gives two forms: $P(k) \approx \sqrt{\pi/c}/k^2$ if $\sqrt{c}A \gg 1$ and $P(k) \approx \sqrt{3/8}k^{-3/2}$ if $\sqrt{c}A \ll 1$ (crossover at $k_c = \dfrac{8\pi}{3c}$) (Fig. 6). The degree distribution of the network with $c = 0.5$, $\sigma = 0$, $m = 1$, and $\mu = 1$ is well described by Eq. 11. (ii) When c is sufficiently small and $\sqrt{c}A \ll 1$ holds for all u,

$$k(u,t) \approx 3A/(2\sqrt{c}) \approx \begin{cases} \dfrac{3\sqrt{c^3}t^2}{4\sqrt{2u}} & (cu \ll 1) \\ \dfrac{3t^2}{2u^2} & (cu \gg 1), \end{cases} \tag{12}$$

which gives $P(k) \approx 9(ct)^3/(16k^3)$ if $cu \ll 1$ and $P(k) \approx \sqrt{3/8}k^{-3/2}$ if $cu \gg 1$ (crossover at $k_c = \dfrac{3}{2^{5/3}}(ct)^2$) (Fig. 7). The degree distribution of the network with $c = 1.0 \times 10^{-4}$, $m = 1$, $\mu = 1$, and $\sigma = 0$ is in good agreement with Eq. 12.

Although we cannot derive an analytical form of the degree distribution in the general model where the conditions $\mu = 1$, $\sigma = 0$, and $m = 1$ are not satisfied, it is possible to understand how these parameters modify the degree distribution in a qualitative way. The network with small μ decreases the degree of nodes because fewer number of connections are made between existing nodes (Fig. 6). The network with $\sigma = -0.5$ has nodes with larger degree number than the maximum of the degree number in the network with $\sigma = 0$ (Fig. 7). This is because the growth rates proportional to $s_i + \sigma$ is much slower than the rate proportional to s_i if s_i is small, and thus the new connections are distributed among the nodes with large strength and degree.

From Eqs. 3, 12, and 11, we find the relationship between the degree and the strength (Figs. 8 and 9). The degree k is proportional to the strength s for $\sqrt{c}A \ll 1$, whereas $k \approx \sqrt{\pi/c}s^{1/2}$ holds for $cu \gg 1$ and $\sqrt{c}A \gg 1$. The linear relationship for $\sqrt{c}A \ll 1$ comes from the fact that the weights of almost all links between 'young' nodes equal unity. The relationship $k \propto s$ holds for the network with $\mu = 0.1$ (Fig. 8) because small μ makes the increment of the weights of the existing links less frequent and the weights of most links remain unity.

3.3 Weight distribution

As in the case of the adjacency matrix, we can define the continuous version of the weight matrix w_{ij}, $w(u_1, u_2, t)$, whose dynamics are governed by the differential equation

$$\frac{\partial w(u_1,u_2,t)}{\partial t} = 2ct\frac{s(u_1,t)s(u_2,t)}{\left(\int_0^t dv\, s(v,t)\right)^2}.$$

The solution is given by

$$w(u_1,u_2,t) = (2+ct)^2/\sqrt{u_1u_2(2+cu_1)^3(2+cu_2)^3}.$$

Note that the relationship $\int_0^t dv\, w(u,v,t) = s(u,t)$ is satisfied. Using the relationship

$$P(w,u_1,t)du_1dw = \frac{1}{t^2}du_1du_2, \tag{13}$$

we obtain

$$P(w,u_1,t) = \frac{1}{t^2}\left|\frac{\partial(u_1,u_2)}{\partial(u_1,w)}\right| = \frac{(2+cu_2)u_2}{t^2(1+2cu_2)w} \tag{14}$$

and integrating it with respect to u_1 gives the weight distribution. Assuming $cu_1 \gg 1$, $cu_2 \gg 1$, and thus $w \approx \frac{t^2}{cu_1^2u_2^2}$, the distribution $P(w,u_1,t)$ is given by

$$P(w,u_1,t) \approx \frac{cu_2^2}{2t^2cu_2w} \approx \frac{w^{-3/2}}{2\sqrt{c}u_1t}.$$

The distribution $P(w)$ hence obeys the exponent $-3/2$ for large c (Fig. 10). In the same way as the degree distribution, we can understand how the parameter

μ affects the weight distribution. In the network with $\mu = 0.1$, slow growth of the existing nodes makes the slope of the log of the weight distribution steeper. If c is small and $cu_1 \ll 1$ and $cu_2 \ll 1$ hold for the nodes which have the edges with large weight, we obtain

$$P(w, u_1, t) \approx \frac{2u_2}{t^2 w} \approx \frac{(2+ct)^4 w^{-3}}{32 u_1 t^2},$$

from which we see the power-law exponent -3 for $P(w)$ (Fig. 11). In addition, it is often observed that the average weight scales with the degrees of the nodes as $\langle w_{ij} \rangle = (k_i k_j)^\theta$ [24]. For $cu_1 \gg 1$, $cu_2 \gg 1$ and $\sqrt{c}A \gg 1$, the approximations $w(u_1, u_2, t) \approx \dfrac{t^2}{cu_1^2 u_2^2}$ and $k(u, t) \approx \sqrt{\dfrac{\pi}{c}\dfrac{t}{u}}$ give the exponent $\theta = 2$. For $cu_1 \ll 1$, $cu_1 \ll 1$, and $\sqrt{c}A \ll 1$, the approximations $w(u_1, u_2, t) \approx \dfrac{(2+ct)^2}{8\sqrt{u_1 u_2}}$ and $k(u, t) \approx \dfrac{3\sqrt{c^3}t^2}{4\sqrt{2u}}$ give the exponent $\theta = 1$. These exponents are in a good agreement with the simulation result (Figs. 12 and 13). The network with $\sigma = -0.5$ has a much larger proportion of links with large weight than the network with $\sigma = 0$ (Fig. 13). In this network, the nodes with strength 1 grow slower than in the network with $\sigma = 0$, and thus the weights are distributed to the nodes with larger strength, that is, the nodes which have links with large weight.

We summarized the power-law exponents derived analytically in Fig. 14. Although in many cases we cannot derive the exponents in an analytical form, the present model changes its power-law exponents depending on the model parameters as shown in the numerical results above.

4 Real-world networks and the model

Next, we will examine whether the simple weighted scale-free network model we have presented in this paper can reproduce the statistical properties of real-world networks. Networks of words [6], coauthorship [1] and collaboration [2] are reported to be scale-free networks. We estimate the power-law exponents of these networks and compare them with the exponents of the model networks.

In the WordNet, a node represents a word, which is connected to the preceding and succeeding words in a sentence [6]. The weight of a link is the number of the times the words corresponding to the nodes on the two ends of the link appeared side by side in the text. Dorogovtsev and Mendes [30] reported that the degree distribution of the WordNet is well-described by their network model. Since our model can be regarded as a generalization of their model, we examined whether our model could reproduce its properties including properties related to weight and strength, which were not taken into account in the Dorogovtsev and Mendes model. There are 57708 unique words and 5×10^6 word occurrences in the WordNet we reconstructed from the novels of Charles Dickens distributed by Project Gutenberg (`http://www.gutenberg.org/wiki/Main_Page`). Setting $m = 1$, $\mu = 1$, and $\sigma = 0$ as in the model of Dorogovtsev and Mendes, we obtain $c = 3.13 \times 10^{-3}$ from the summation of the strength 1.05×10^7. The model network and the WordNet show an excellent agreement except for the degree-weight relationship (Fig. 15). In the WordNet, the expected link weight $\langle w_{ij} \rangle$ tends to be slightly larger than that of the model network. It is because the connection probability of the words is not determined only by the frequency of these words. Some pairs of words such as 'in the' tends to appear side by side frequently, whereas some others such as 'an the' never appear, although the words 'an' and 'the' frequently appear in the text. This selectivity more deeply affects the relationship between two nodes than the statistical properties of single nodes. Thus, the strength, degree, and weight distributions of the WordNet are fitted by the present model even though degree-weight relationship of the WordNet cannot be explained by this model. We can safely say that our model successfully generalizes the model of Dorogovtsev and Mendes and that our model captures an important feature of the WordNet. What must be noticed is that the present model also accounts for the distribution of the weight of a link, which corresponds to the number of times the two words appear side by side, without degrading the fit to the distributions of the strength and degree.

Network of the coauthorship in the field of geology also shows a good agreement with the present model. Assuming $m = 1$, $\mu = 1$, and $\sigma = 0$, the quantity c can be estimated as $c = 1.5 \times 10^{-4}$ by the condition that, in the reconstructed coauthorship network, t is the number of researchers (100945) and $2t + ct^2$ must be equal to the summation of the strength, 1.75×10^6. The smallness of c implies that there are few research papers on which no new author contributes. The network generated by this model with the above estimated parameters exhibits scale-free properties similar to the real collaboration network of coauthorship (Fig. 16). The exponents of the various scale-free properties were simultaneously derived from a single real-measured model parameter c in the coauthorship network as well as in the WordNet. However, there exists some difference between the model network and the coauthorship network. First, degree distribution of the model network has a longer tail than the coauthorship network. This is because an author tends to write a coauthored paper with the authors s/he collaborated earlier. This tendency decreases the degree of a node and makes the degree distribution steeper in the coauthorship network. The difference in degree-strength relationship between the model and the real coauthorship results from the same tendency. Second, the weight distribution of the collaboration network has a longer tail than the model network. This is another consequence of the same tendency. Once two authors have collaborated, the number of times they collaborate increases faster than in the model network. Thus the model network cannot reproduce fully the properties of the coauthorship network because, as is the case with WordNet, there exists a correlation between the nodes.

Reconstructing the collaboration network of the actors/actresses in the same way, we found that this network also exhibits scale-free properties reproduced by the present model (Fig. 17). Since the network contains 1123031 actors/actresses and the sum of the strength equals 1.4×10^8, we set the parameters as $t = 1123031$, $m = 10$, $\mu = 1$, $\sigma = 0$, and $c = 1 \times 10^{-4}$. Although we assumed $m = 1$ for the WordNet and the coauthorship network, here we assume $m = 10$, because the model with the parameter $m = 10$ fits better to the degree and strength distribution than the model with $m = 1$. The larger m implies that a new actor/actress collaborates with many existing actors/actresses. It is not surprising because a much larger number of actors/actresses appear, on average, in a movie than in a scientific paper. Thus, the parameter m of the model reflects the structure of the real collaboration network. The difference in the degree and weight distributions and other relationships between the model and the collaboration network can be explained by the fact that an actor/actress tends to collaborate many times with the same actors/actresses, which is the same reason as the above two cases.

5 Conclusion

Extending our previous model, we constructed a model which generates weighted-scale free networks with variable power-law exponents. The model network grows through the weight-driven preferential attachment and exhibits scale-free properties. The advantage of this model is that the power-law exponents of the strength, degree, weight distribution and the relationships between them, can be controlled by changing the parameters μ and σ. This flexibility enables the model network to fit very well to real-world networks. Furthermore, our proposed model is mathematically tractable, which allows us to understand the underlying essential mechanisms. We expect that the present model becomes a theoretical tool in a wide range of studies on complex systems including oscillators, disease propagation, packet transport, and opinion formation.

Acknowledgements

This work was supported by Grants-in-Aid from the Ministry of Education, Science, Sports, and Culture of Japan: Grant numbers 18047014, 18019019, and 18300079.

References

[1] M. E. J. Newman, Proc. Natl. Acad. Sci. USA 98 (2001) 404–409.

[2] D. J. Watts, S. H. Strogatz, Nature 393 (1998) 440–442.

[3] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, A.-L. Barabási, Nature 407 (2000) 651–654.

[4] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, J. Wiener, Computer Networks 33 (2000) 309–320.

[5] M. Faloutsos, P. Faloutsos, C. Faloutsos, Computer Communications Review 29 (1999) 251–262.

[6] R. Ferrer i Cancho, R. V. Solé, Tech. Rep. 01-03-016, Santa Fe Institute (2001). URL `http://ideas.repec.org/p/wop/safiwp/01-03-016.html`

[7] A.-L. Barabási, R. Albert, Science 286 (1999) 509.

[8] R. Albert, A.-L. Barabási, Rev. Mod. Phys. 74 (2002) 47–97.

[9] S. N. Dorogovtsev, J. F. F. Mendes, A. N. Samukhin, Structure of growing networks with preferential linking, Phys. Rev. Lett. 85 (21) (2000) 4633–4636.

[10] G. Bianconi, A.-L. Barabási, Europhys. Lett. 54 (2001) 436–442.

[11] P. Holme, B. J. Kim, Growing scale-free networks with tunable clustering, Phys. Rev. E 65 (2) (2002) 026107.

[12] H. A. Simon, Biometrika 42 (1955) 425–440.

[13] D. J. de S. Price, J. Amer. Soc. Inform. Sci. 27 (1976) 292–306.

[14] K. Park, Y.-C. Lai, N. Ye, Self-organized scale-free networks, Phys. Rev. E 72 (2) (2005) 026131.

[15] G. Caldarelli, A. Capocci, P. De Los Rios, M. A. Muñoz, Scale-free networks from varying vertex intrinsic fitness, Phys. Rev. Lett. 89 (25) (2002) 258702.

[16] N. Masuda, H. Miwa, N. Konno, Analysis of scale-free networks based on a threshold graph with intrinsic vertex weights, Phys. Rev. E 70 (3) (2004) 036124. URL `http://link.aps.org/abstract/PRE/v70/e036124`

[17] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Phys. Rep. 424 (2006) 175–308.

[18] R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks, Phys. Rev. Lett. 86 (14) (2001) 3200–3203.

[19] A. T. Bernardes, S. D., J. Kertész, Epidemic spreading in scale-free networks, Eur. Phys. J. B 25 (2002) 123–127.

[20] P. Holme, A. Trusina, B. J. Kim, P. Minnhagen, Prisoners' dilemma in real-world acquaintance networks: Spikes and quasiequilibria induced by the interplay between structure and dynamics, Phys. Rev. E 68 (3) (2003) 030901.

[21] P. Echenique, J. Gomez-Gardenes, Y. Moreno, Improved routing strategies for internet traffic delivery, Phys. Rev. E 70 (5) (2004) 056105.

[22] Y. Moreno, A. F. Pacheco, Synchronization of kuramoto oscillators in scale-free networks, Europhysics Letters 68 (2004) 603.

[23] M. E. J. Newman, Phys. Rev. E 64 (2001) 016131.

[24] A. Barrat, M. Barthélemy, R. Pastor-Satorras, A. Vespignani, Proc. Nat. Acad. Sci. U.S.A. 101 (2004) 3747.

[25] S. H. Yook, H. Jeong, A.-L. Barabási, Y. Tu, Phys. Rev. Lett. 86 (2001) 5835.

[26] D. Zheng, S. Trimper, B. Zheng, P. M. Hui, Phys. Rev. E 67 (2003) 040102.

[27] A. Barrat, M. Barthélemy, A. Vespignani, Phys. Rev. E 70 (2004) 066149.

[28] S. Wang, C. Zhang, Phys. Rev. E 70 (2004) 066127.

[29] T. Tanaka, T. Aoyagi, Scale-free netwroks with self-growing weight.
URL http://arxiv.org/abs/cond-mat/0701314

[30] S. N. Dorogovtsev, J. F. F. Mendes, Proc. R. Soc. London Sect. B 268 (2001) 2603.

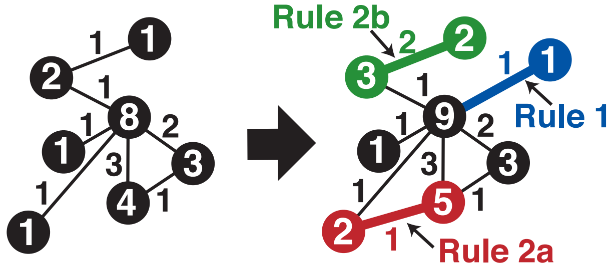

Fig. 1. Schematic explanation of the rule for generating the network. The numbers on a node and near its links indicate node strength and link weights, respectively. At each time step, a new single node (a blue circle) appears and connects to existing nodes with links of weight one (a blue link). This new link is created by preferential attachment with a probability proportional to the strength of the existing node (Rule 1 in the text). At the same time, some pairs of existing nodes are chosen on a simple strength preferential rule (see the main text for details), and a new link of weight one (a red link) is created between these chosen nodes (Rule 2a). If a link already exists between them, the weight of the link (a green link) is incremented by one (Rule 2b).

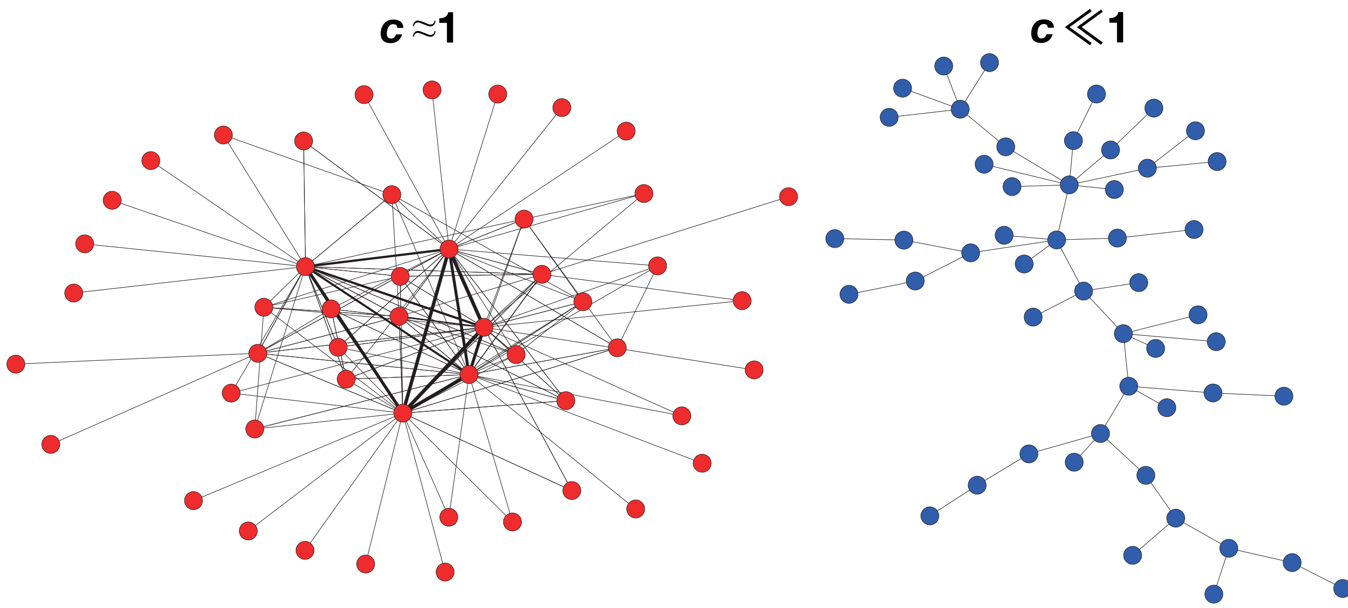

Fig. 2. Dependence of the network stricture on the parameter c. Networks with 50 nodes generated with the parameters $c = 0.5$ (left) and $c = 0.0001$ (right) are shown. The width of the links is proportional to their weights. There are many loops in the network with large c, whereas the network with small c has a tree-like structure without loops. See the main text for details.

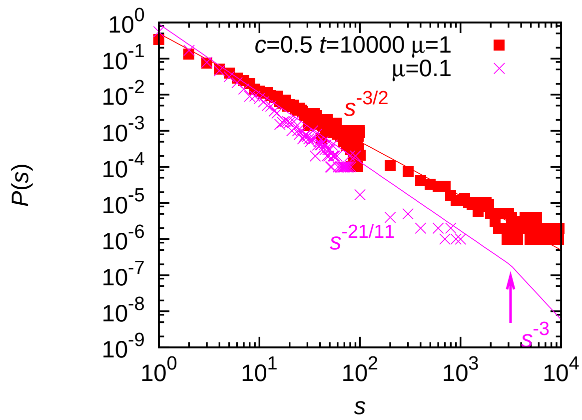

Fig. 3. Comparison of strength distribution between theoretical and numerical results for a network with large c. Strength distribution of a network with $\mu = 0.1$ shows a larger exponent (2.1/1.1) than that of a network with $\mu = 1$ (3/2). Note that the simulation results show no crossover points because they occur at strengths larger than the greatest node strength present in the network. The parameters m and σ are set to 1 and 0, respectively. The bin width is set to 1 for $s < 100$ and 100 for $s > 100$ because points are sparse in the region $s > 100$.

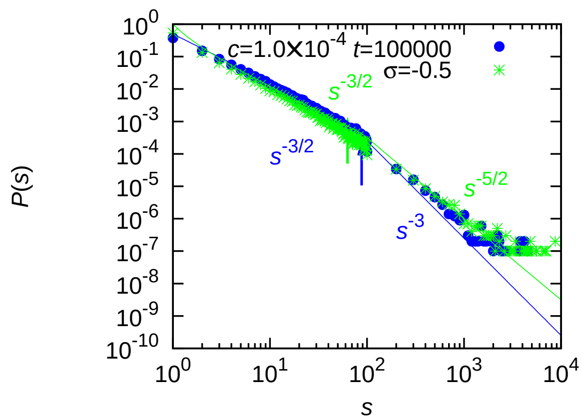

Fig. 4. Comparison of distribution strength between theoretical and numerical results for the network with small c. This figure shows the strength distribution of the network with $\sigma = 0$ (circle) and $\sigma = -0.5$ (asterisk). The network with $\sigma = -0.5$ shows a smaller exponent (5/2) than the network with $\sigma = 0$ (3) in the second regime. The parameters m and μ are set to 1. Note that for the network with $c = 1.0 \times 10^{-4}$ the power-law exponent changes at the crossover point indicated by the arrow.

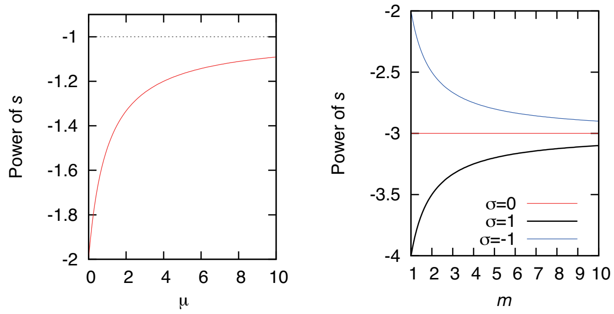

Fig. 5. The dependency of the power-law exponent of the strength distribution on the parameters μ, m, and σ. If $\sigma = 0$, the strength distribution $P(s)$ obeys the exponent $-\frac{2+\mu}{1+\mu}$ for $cu^{\mu} \gg 1$. If $\mu = 1$, the strength distribution obeys the exponent $-\frac{3m+\sigma}{m}$ for $cu \ll 1$.

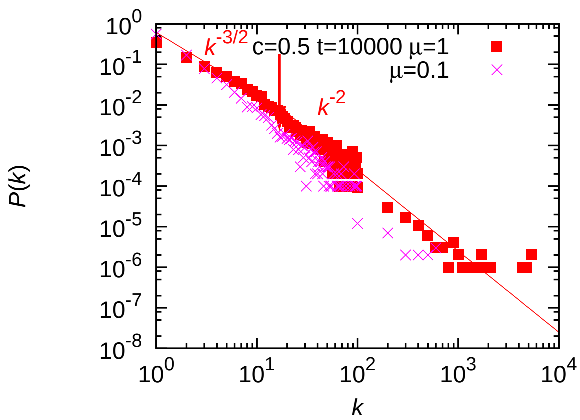

Fig. 6. Degree distribution in the network with large c. For the network with large c, $c = 0.5$ and $\mu = 1$, the theoretical results are $P(k) \approx \sqrt{3/8}k^{-3/2}$ for $(k < k_c)$ and $\sqrt{\pi/c}k^{-2}$ for $(k > k_c)$ $(k_c = 8\pi/(3c))$. The crossover point k_c is indicated by an arrow. The same parameters as in Fig. 3 are used.

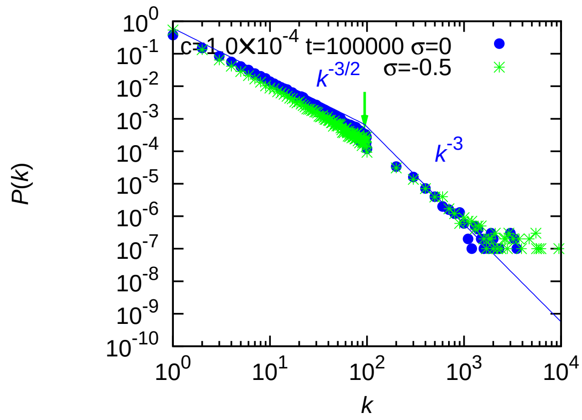

Fig. 7. Distribution of degree in the networks with small c. Degree distribution for the small c network with $c = 1.0 \times 10^{-4}$ and $\sigma = 0$, $P(k) \approx \sqrt{3/8}k^{-3/2}$ $(k < k_c)$ and $9(ct)^3/(16k^3)$ $(k > k_c)$ $(k_c = 3(ct)^2/2^{5/3})$. The crossover point k_c is indicated by an arrow. The same parameters as in Fig. 4 are used.

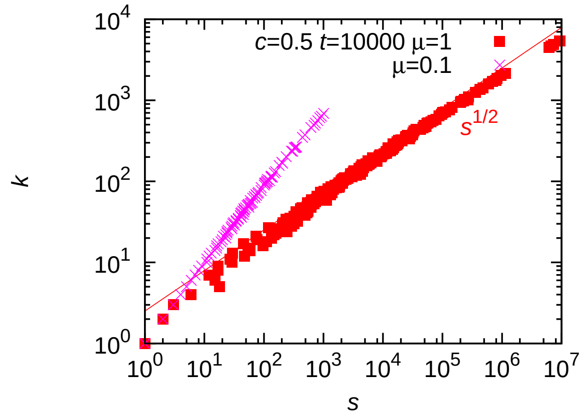

Fig. 8. Relationship between strength and degree for a network with large c. The relationships $k \propto s^{1/2}$ and $k \propto s$ hold for the networks with $\mu = 1$ and $\mu = 0.1$, respectively. The same parameters as in Fig. 3 are used.

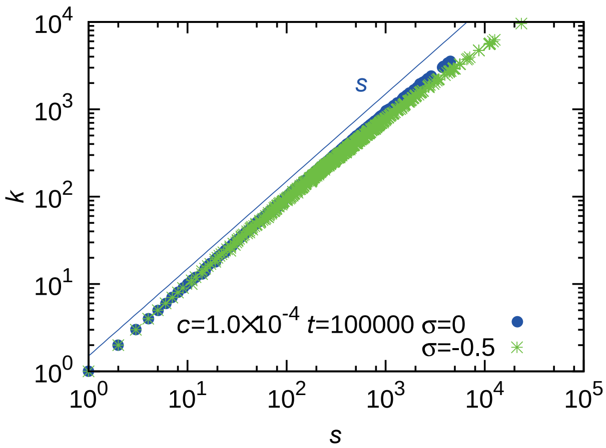

Fig. 9. Relationship between strength and degree for a network with small c. The relationship $k \sim s$ holds for the networks with $\sigma = 0$ and $\sigma = 1$. The same parameters as in Fig. 4 are used.

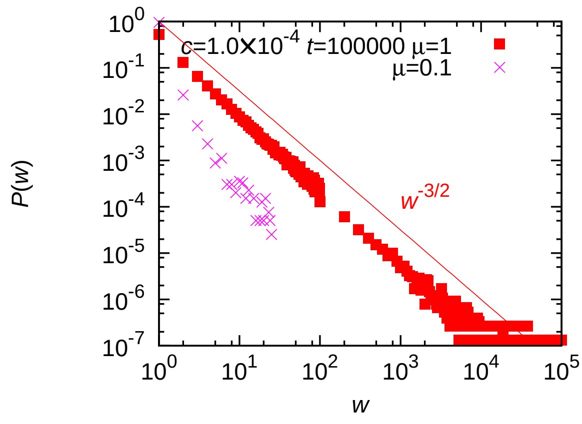

Fig. 10. Distribution of weights. No crossover behavior is observed because $cu \gg 1$ holds for almost all nodes in the network with $c = 0.5$ and $\mu = 1$. The same parameters as in Fig. 3 are used.

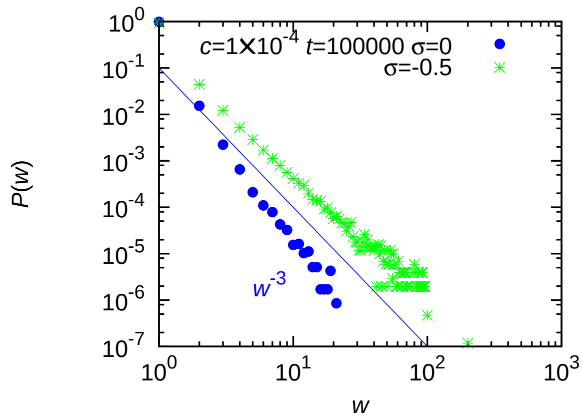

Fig. 11. Distribution of weights. No crossover behavior is observed because $cu \ll 1$ holds for almost all nodes in the network with $c = 1.0 \times 10^{-4}$. The same parameters as in Fig. 4 are used.

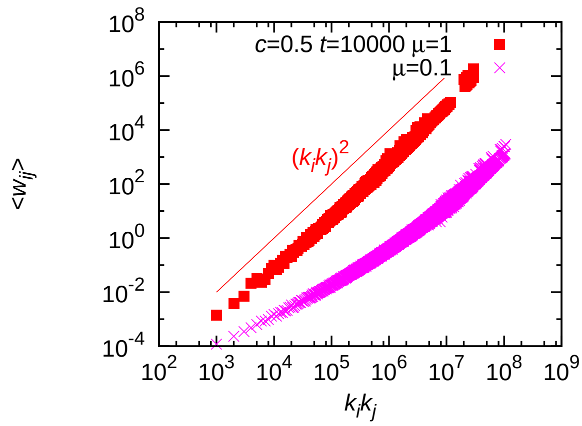

Fig. 12. The relationship between degree and weight for networks with large c. The relationship $\langle w_{ij} \rangle \sim (k_i k_j)^2$ holds for the network with $\mu = 1$. The same parameters as in Fig. 3 are used.

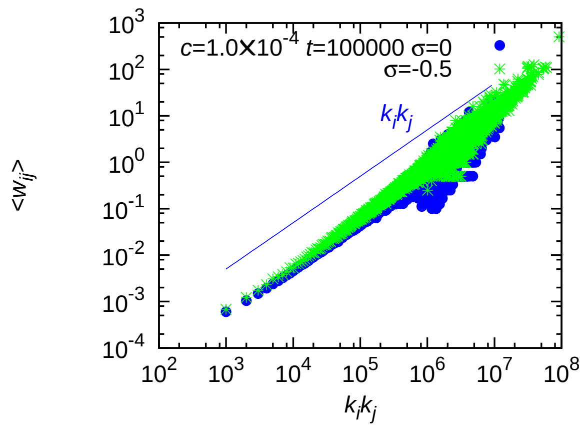

Fig. 13. The relationship between degree and weight for networks with small c. The relationship $\langle w_{ij} \rangle \sim k_i k_j$ holds for the networks with $\sigma = 0$ and $\sigma = -0.5$. The same parameters as in Fig. 4 are used.

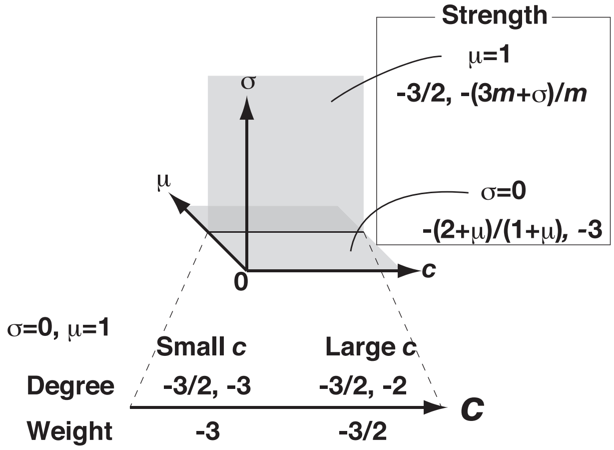

Fig. 14. Summary of the power-law exponents of the strength, degree, and weight distributions. Analysis proved that power-law exponents of the strength distribution are $-(2+\mu)/(1+\mu)$ ($cu^{\mu} \gg 1$) and -3 ($cu^{\mu} \ll 1$) if $\sigma = 0$, and $-3/2$ ($cu \gg 1$) and $-(3m+\sigma)/m$ ($cu \ll 1$) if $\mu = 1$. In the model with $\sigma = 0$ and $\mu = 1$, powers of the degree distribution are $-3/2$ ($\sqrt{c}A \ll 1$) and -2 ($\sqrt{c}A \gg 1$) if c is large, and $-3/2$ ($cu \gg 1$) and -3 ($cu \ll 1$) if c is small. In the model with same parameters, powers of the degree distribution are $-3/2$ if c is large, and -3 if c is small.

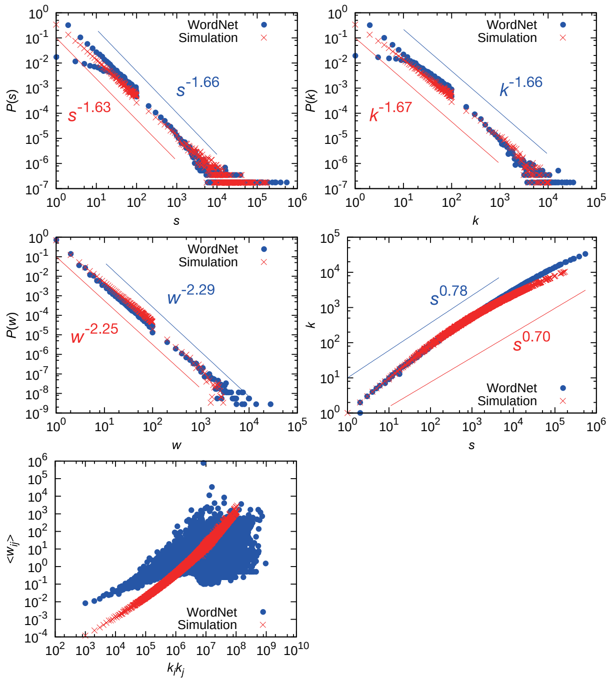

Fig. 15. Comparison of scale-free properties between the WordNet (filled circles) and the present model (cross). Strength distribution (top left), degree distribution (top right), weight distribution (middle left), strength-degree relationship (middle right), and degree-weight relationship (bottom left) are shown. Odd number of strength and degree has much smaller probability, which is manifested in the region $s < 10$ and $k < 10$, because, in most cases, a word is connected to the previous and next words in the sentence while the words at the head and the tail of the sentence are connected by the next word and the previous word only, respectively. We show the power-law exponents for the real data and the simulation results, except for the degree-weight relationship, which was difficult to fit by a power-low relationship. The WordNet is constructed from the writings of Charles Dickens in Project Gutenberg.

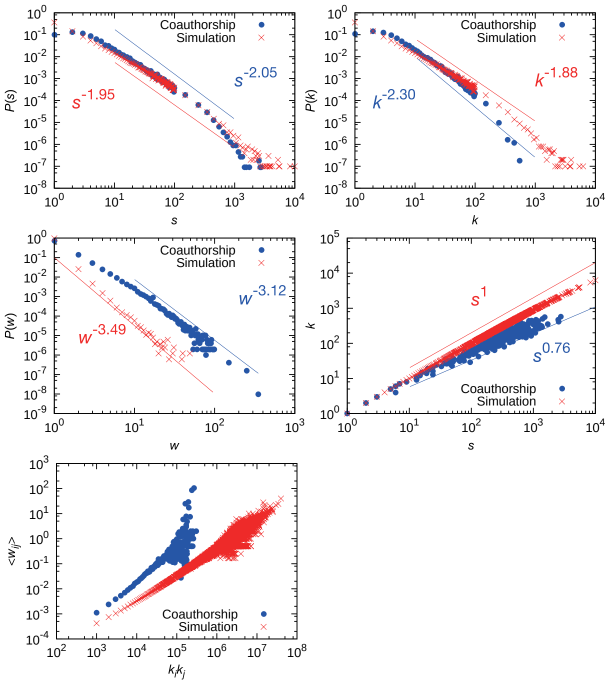

Fig. 16. Comparison of scale-free properties between the coauthorship network (filled circles) and the present model (cross). Strength distribution (top left), degree distribution (top right), weight distribution (middle left), strength-degree relationship (middle right), and degree-weight relationship (bottom left) are shown. The coauthorship network is reconstructed from the Geological Literature Search System (GEOLIS+ CD-ROM Ver.5) provided by AIST (permission number 63500-A-20070322-001).

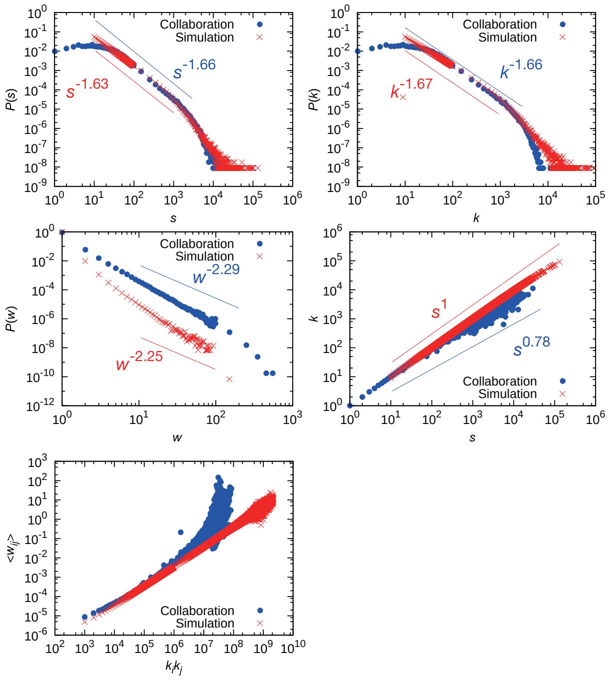

Fig. 17. Comparison of scale-free properties between the actor/actress collaboration network (filled circles) and the present model (cross). Strength distribution (top left), degree distribution (top right), weight distribution (middle left), strength-degree relationship (middle right), and degree-weight relationship (bottom left) are shown. The collaboration network is reconstructed from the data provided by The Internet Movie Database (`http://www.imdb.com/`).