Chapter 2
Fitness-Based Generative Models for Power-Law Networks

Khanh Nguyen and Duc A. Tran

Abstract Many real-world complex networks exhibit a power-law degree distribution. A dominant concept traditionally believed to underlie the emergence of this phenomenon is the mechanism of preferential attachment which originally states that in a growing network a node with higher degree is more likely to be connected by joining nodes. However, a line of research towards a naturally comprehensible explanation for the formation of power-law networks has argued that degree is not the only key factor influencing the network growth. Instead, it is conjectured that each node has a "fitness" representing its propensity to attract links. The concept of fitness is more general than degree; the former may be some factor that is not degree, or may be degree in combination with other factors. This chapter presents a discussion of existing models for generating power-law networks, that belong to this approach.

2.1 Introduction

The last decade has seen much interest in studying complex real-world networks and attempting to find theoretical models that elucidate their structure. Although empirical networks have been studied for some time, a surge in activity is often seen as having started with Watts and Strogatz's paper on "small-world networks" [23]. More recently, the major focus of research has moved from small-world networks to "scale-free" networks, which are characterized by having power-law degree

K. Nguyen • D.A. Tran (✉)
Computer Science Department, University of Massachusets, Boston, MA, USA
e-mail: knguyen@cs.umb.edu; duc@cs.umb.edu

M.T. Thai and P.M. Pardalos (eds.), *Handbook of Optimization in Complex Networks: Theory and Applications*, Springer Optimization and Its Applications 57,
DOI 10.1007/978-1-4614-0754-6_2,

Table 2.1 Real-world networks: power-law exponent (λ)

Case	λ
Mathematics coauthorship	2.2
Film actor collaborations	2.1–2.3
WWW	2.1
Internet backbone	2.15–2.2
Protein interactions	2–2.4
ER graph	NA
BA graph	3

distributions [3]; that is, if $p(k)$ is the fraction of nodes in the network having degree k (i.e., having k connections to other nodes), then (for suitably large k)

$$p(k) = ck^{-\lambda}, \tag{2.1}$$

where $c = (\lambda - 1)m^{\lambda-1}$ is a normalization factor and m is the minimum degree in the network. This distribution is observed in many real-world networks, including the WWW [1], the Internet [12], metabolic networks [15], protein networks [14], co-authorship networks [21], and sexual contact networks [19]. In these networks, there are a few nodes with high degree and many other nodes with small degree, a property not found in standard Erdós–Rényi (ER) random graphs [9].

The near ubiquity of heavy-tailed degree distributions such as the power-law (2.1) for real-world complex networks, together with the inadequacy of the ER random graphs as a theoretical model for such networks, brings into sharp relief the fundamental problem of obtaining a satisfactory theoretical explanation for how heavy-tailed degree distribution can naturally arise in complex networks.

A dominant concept traditionally believed to underlie the emergence of the power-law phenomenon is the mechanism of preferential attachment, proposed by Barabási and Albert [3]: the higher degree a node has, the more likely it is to be connected by new nodes. This model, hereafter referred to as the BA model, leads to a growing random network which simulations and analytic arguments show has a power-law degree distribution with exponent $\lambda = 3$. Despite its elegance and simplicity, a deficiency of this mechanism is due to its fixed power-law exponent. As real-world networks exhibit a wide range of exponents, typically between 2 and 3 (see Table 2.1 for examples), the BA model may only explain a small subset of complex networks.

Consequently, other mechanisms have been proposed. Some, e.g., [8], are merely formulaic without a natural interpretation. Others use different connectivity information, not merely degree, of each node to influence the formation of a network, such as the mechanism in [17]. Still, a universally accepted explanation that works for not just one network but also others remains to be found. For example, if we use the BA-based models to explain the sexual contact network studied in [19], which is known to be power-law, a new individual will prefer to have sexual contact with those individuals who already have a large number of sexual contacts, while the explanation according to [17] will infer that a new individual will have sexual contact with some existing partners of a randomly chosen individual. These explanations seem bizarre for human sexual behavior.

In searching for a more natural explanation for the formation of power-law networks in the real world, there is a line of research, [4–6,13,16,22], founded based on the conjecture that in many complex networks each node will have associated to it a “fitness” representing the propensity of the node to attract links. Using the fitness concept to explain the sexual contact network above, we could say that it is the fitness of an individual that attracts other individuals; an individual wants to have sexual contact with another individual because of the latter’s fitness, not the latter’s connectivity. The fact that a node has many contacts may just be a consequence of its high fitness. The key challenge in the design of fitness-based models is how fitness is defined; for example, what is fitness? what is it made of? what is its influence? Fitness may be just degree, or something not, or a combination of many factors, explicit or implicit. Fundamental differences in the approach to addressing these questions is discussed in the remaining sections of this chapter.

It is important to note that, as argued in [18], there are a rich variety of “emergent” topological signatures beyond mere power-law degree distributions that are also present in real-world complex networks. To date, there has been no perfect network generative model satisfying all signatures. This chapter is focused only on the power-law degree property of complex networks.

2.2 Early Network Models

One of the earliest theoretical models of a complex network was proposed and studied in detail by Erdós and Rényi [9–11] in a famous series of papers in the 1950s and 1960s. The ER random graph model consists of n nodes (or vertices) joined by links (or edges), where each possible edge between two vertices is present independently with probability p and absent with probability $1-p$. The probability $p(k)$ that a node has exactly degree k is given by the binomial distribution

$$p(k) = \binom{n-1}{k} p^k (1-p)^{n-k-1}. \tag{2.2}$$

In the limit when $n \gg kz$, where $z = (n-1)p$ is the mean degree, the degree distribution becomes the Poisson distribution

$$p(k) = \frac{z^k \mathrm{e}^{-z}}{k!}. \tag{2.3}$$

The Poisson distribution is strongly peaked about the mean z, and has a tail that decays very rapidly as $1/k!$. This rapid decay is completely different from the heavy-tailed power-law nature of the tail of the degree distribution that is observed in many real-world complex networks.

Most widely known among the theoretical models for complex networks with a degree distribution similar to real-world networks is the BA model [3], which works as follows. Starting from a small number (n_0) of nodes (which could, for example, be chosen to form a random graph), at every time-step we add a new node with $m \leq n_0$ edges linking the new node to the m distinct nodes already present in the network such that the probability Π_i that the new node will connect to node i is proportional to the degree k_i of that node; that is,

$$\Pi_i = \frac{k_i}{\sum_j k_j}. \tag{2.4}$$

Consider a node i that joins the network at time i. Denote by $k_i(t)$ the degree of this node at time t. When the network is sufficiently large, $k_i(t)$ can be represented as a continuous function of time t, which grows at the following rate:

$$\frac{\partial k_i}{\partial t} = \frac{mk_i}{k_1 + k_2 + \cdots + k_n}, \tag{2.5}$$

where $n = n_0 + t$ is the network size at time t, and k_j's are the degrees of the current nodes. For large n, we can ignore the value of n_0 and since m new links are added after each time step (resulting in $2m$ degree increase), the sum of all nodes' degrees, $k_1 + k_2 + \cdots + k_n$, can be approximated by $2mt$. Thus, (2.5) can be rewritten approximately as

$$\frac{\partial k_i}{\partial t} = \frac{mk_i}{2mt} = \frac{k_i}{2t}.$$

Hence, $k_i(t)$ must be of the form $k_i(t) = c\sqrt{t}$ for some constant c. When node i just joins the network, that is, $t = i$, its degree is m and so we have $m = c\sqrt{i}$ or $c = m/\sqrt{i}$. Consequently, the degree of node i as a function of time t is

$$k_i(t) = m\sqrt{t/i}.$$

Given k, the probability that $k_i(t)$ is less than k is

$$Pr[k_i(t) < k] = Pr\left[m\sqrt{t/i} < k\right] = Pr[i > tm^2/k^2].$$

Because node i can equally likely be any node among $(n_0 + t)$ current nodes (consisting of n_0 initial nodes and t nodes added by time t), we have

$$Pr[i > i_0] = 1 - \frac{i_0}{n_0 + t}, \tag{2.6}$$

for any given i_0. Consequently,

$$Pr[i > tm^2/k^2] = 1 - \frac{tm^2/k^2}{n_0 + t} = 1 - \frac{m^2/k^2}{n_0/t + 1} \approx 1 - m^2/k^2$$

(when t is large).

The probability density function (pdf) of node i's degree is

$$P(k) = \frac{\partial Pr[k_i(t) < k]}{\partial k} = 2m^2/k^3.$$

Thus, BA results in a power-law degree distribution with $\lambda = 3$ (independent of m, which only changes the mean degree of the network).

A different generative model for power-law networks is the copy model studied by Kumar et al. [17], in which the new node chooses an existing node at random and copies a fraction of the links of the existing node. This model assumes that the new node knows who the chosen existing node is connected to. This assumption is reasonable for web graphs for which the model is originally proposed. However, the assumption is not always reasonable in other circumstances, as is clear from the case of sexual contact networks discussed in Sect. 2.1.

2.3 Combination of Degree and Fitness

An early fitness-based model for constructing power-law networks was proposed in [5]. This model assumes that the evolution of a network is driven by two factors associated with each node, its node degree and its fitness. These factors jointly determine the rate at which new links are added to the node. While node degree represents an ability to attract links that is increasing over time, node fitness represents something attractive about the node, that is constant. For example, in the WWW network, there can be two web pages published at the same time (thus, same degree) but later one might be much more popular than the other; this might be because of something intrinsic about one page, e.g., its content, that makes it more attractive than the other page.

In the proposed model, each node i has a fitness Φ_i which is chosen according to some distribution $\rho(\Phi_i)$. The network construction algorithm generalizes the BA algorithm as follows:

- Parameters
  - n_0: the size of the initial network which can be any graph.
  - $m \leq n_0$: the number of nodes a new node connects to when it joins the network.
  - n: number of nodes in the final network.
- Procedure
  1. Initially, start with the initial network of n_0 nodes, each assigned a random fitness according to distribution ρ.

2. Each time a new node is added; stop after the nth node is added.
   - Assign a random fitness to the new node according to distribution ρ
   - Add m edges linking the new node to m distinct existing nodes such that the probability Π_i for connecting to an existing node i is taken to be proportional to its fitness Φ_i:

$$\Pi_i = \frac{k_i \Phi_i}{\sum_j k_j \Phi_j}. \tag{2.7}$$

Suppose that i is the time node i joins the network. When the network is sufficiently large, the degree of node i over time, k_i, can be represented as a continuous function of time t. The value of this function increases over time as follows:

$$\frac{\partial k_i}{\partial t} = m \left(\frac{k_i \Phi_i}{\sum_j k_j \Phi_j}. \right) \tag{2.8}$$

It can be shown that $k_i(t)$ follows a power law

$$k_i(t) = m \left(\frac{t}{i} \right)^{\Phi_i / A},$$

with A given by

$$1 = \int_0^{\Phi_{\max}} \mathrm{d}x \rho(x) \frac{1}{\frac{A}{x} - 1}.$$

(Here, $\Phi_{\max}$ defines the maximum fitness.)

The degree of a node therefore grows faster if it has a larger fitness. This allows a node with a higher fitness to enter the network late but still become more popular than nodes that have stayed in the network for a much longer period.

If every node has the same fitness, this model is identical to the orignal BA model, resulting in power-law networks of exponent $\lambda = 3$. In the case that fitness is chosen uniformly in the interval [0, 1], we have $A = 1.255$ and the fraction of nodes in the network having degree k follows a generalized power-law with an inverse logarithmic correction:

$$p(k) \propto \frac{k^{-2.255}}{\ln k}. \tag{2.9}$$

This estimation has been confirmed by numerical simulations [5].

2.4 Fitness-Based Model with Deletion

The BA model cannot explain the heavy-tail degree distribution in complex networks where there are many node deletions [20]. On the other hand, the deletion rate can be significant for certain networks, such as the WWW. In the study of [16], based on the result of year-long web crawls to track the creation and deletion of web pages on the Internet, it was found that at least 0.76 pages were removed for every new web page created. Towards an approach that accommodates deletion, a model was proposed in [16] which extends the fitness-based model of [5] discussed in the previous section by allowing a probability for deleting a node at each time step. Specifically, the network is constructed as follow:

- Parameters
  - n_0: the size of the initial network which can be any graph.
  - $m \leq n_0$: the number of nodes a new node connects to when it joins the network.
  - n: number of nodes in the final network.
- Procedure
  1. Initially, start with the initial network of n_0 nodes, each assigned a random fitness according to distribution ρ.
  2. Each time a new node is added; stop after the nth node is added.
     - Assign a random fitness to the new node according to distribution ρ.
     - Add m edges linking the new node to m distinct existing nodes such that the probability Π_i for connecting to an existing node i is taken to be proportional to its fitness Φ_i:

$$\Pi_i = \frac{k_i \Phi_i}{\sum_j k_j \Phi_j}. \tag{2.10}$$

     - With probablity c, a random node is removed, along with all its edges.

It can be shown that the evolution of the degree of node i over time follows a power law

$$k_i(t) = m\left(\frac{t}{i}\right)^{\beta(\Phi_i)},$$

where the growth exponent β is a function of fitness and deletion rate

$$\beta(\Phi_i) = \frac{\Phi_i}{A} - \frac{c}{1-c},$$

with A given by

$$1 = \int_0^{\Phi_{\max}} \mathrm{d}x \rho(x) \frac{1}{\frac{A}{x}\frac{1+c}{1-c} - 1}.$$

If fitness is identical for every node, this model is exactly the same as the BA model applied to networks with deletion, resulting in a power-law exponent $\lambda = 1 + 2/(1-c)$ for small c. As c is closer to 1 (i.e., a high deletion rate), the degree distribution diverges rapidly from power-law. This explains why the BA model is not appropriate for networks with deletion.

In the case that fitness follows a truncated exponential distribution, which has been shown to empirically characterize the fitness distribution of web pages defined as its degree growth rate, the result is that the power-law exponent is not affected by the deletion rate and stablizes around two. In other words, the network as it is growing remains power-law, regardless of the rate of node deletion.

The models discussed in this section and the previous section combine fitness with degree to influence the growth of a network. A fundamentally different model is presented in the next section, which uses fitness as the only driver for the network growth.

2.5 Preferential Attachment Using Fitness Only

In a citation network such as [21] the different nodes (i.e., papers) will have different propensities to attract links (i.e., citations). The various factors that contribute to the likelihood of a paper being cited could include the prominence of the author(s), the importance of the journal in which it is published, the apparent scientific merit of the work, the timeliness of the ideas contained in the paper, etc. Moreover, it is plausible that the overall quantity that determines the propensity of a paper to be cited depends essentially multiplicatively on such various factors. The multiplicative nature is likely in this case since if one or two of the factors happen to be very small then the overall likelihood of a paper being cited is often also small, even when other factors are not small; e.g., an unknown author and an obscure journal were enough to bury a fundamentally important scientific paper.

The lognormal fitness attachment (LNFA) model, proposed by Ghadge et al. in [13], was motivated by the observation above. In this model, the fitness Φ_i representing the property of each node i to attract links is formed multiplicatively from a number of factors $\{\phi_1, \phi_2, ..., \phi_L\}$ as follows:

$$\Phi_i = \prod_{l=1}^{L} \phi_l, \tag{2.11}$$

where each factor ϕ_l is represented as a real non-negative value. Since there may be many factors contributing to the a node's attractiveness, explicit or implicit, we assume that the number of factors ϕ_i is reasonably large and that they are statistically independent. The fitness Φ_i will therefore be lognormally distributed, irrespective of the manner in which the individual factors are distributed. Indeed, we have

$$\ln \Phi_i = \sum_{l=1}^{L} \ln \phi_l, \tag{2.12}$$

and the Central Limit Theorem implies that this sum will converge to a normal distribution. Therefore, $\ln \Phi_i$ will be normally distributed. Since a random variable X has a lognormal distribution if the random variable $Y = \ln X$ has a normal distribution, Φ_i will be lognormally distributed. The density function of the normal distribution is

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-(y-\mu)^2/(2\sigma^2)}, \tag{2.13}$$

where μ is the mean and σ is the standard derivation (i.e., σ^2 is the variance). The range of the normal distribution is $y \in (-\infty, \infty)$. It follows from the logarithmic relation $Y = \ln X$ that the density function of the lognormal distribution is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \mathrm{e}^{-(\ln x-\mu)^2/2\sigma^2}. \tag{2.14}$$

It is conventional to say that the lognormal distribution has parameters μ and σ when the associated normal distribution has mean μ and standard deviation σ. The range of the lognormal distribution is $x \in (0, \infty)$. The lognormal distribution is skewed with mean $\mathrm{e}^{\mu+\sigma^2/2}$ and variance $(\mathrm{e}^{\sigma^2} - 1)\mathrm{e}^{2\mu+\sigma^2}$.

Thus, the basic hypothesis that each of the nodes has associated to it a fitness of the form (2.11) entails that under quite general conditions this fitness will be lognormally distributed. In the LNFA model, without loss of generality, one can assume that $\mu = 0$; hence, the fitness distribution is characterized by only a single parameter σ. This lognormal distribution has mean $\mathrm{e}^{\sigma^2/2}$ and variance $(\mathrm{e}^{\sigma^2} - 1)\mathrm{e}^{\sigma^2}$; examples are shown in Fig. 2.1.

The network construction algorithm works as follows:

- Parameters
  - σ: parameter for the lognormal fitness distribution.
  - n_0: the size of the initial network which can be any graph.
  - $m \leq n_0$: the number of nodes a new node connects to when it joins the network.
  - n: number of nodes in the final network.

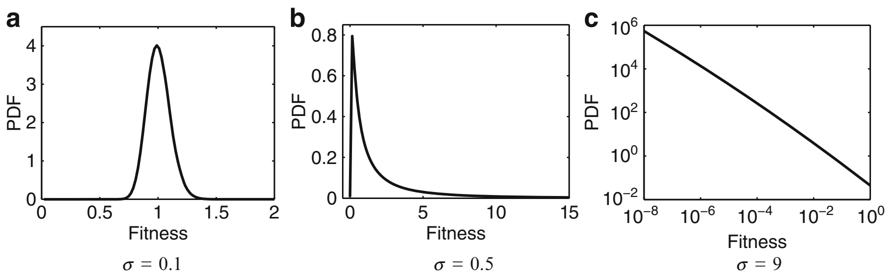

Fig. 2.1 Lognormal fitness distribution for three representative values of σ. The extreme cases are when σ is small (0.1) or large (9). In most cases, the distribution will have the shape similar to the case $\sigma = 1.5$. Figure 2.5 is plotted in the log-log scale to emphasize that all the nodes except a few have fitness zero (or extremely close) while the rest (very few) have high fitness

- Procedure
  1. Initially, start with the initial network of n_0 nodes, each assigned a random fitness according to the lognormal distribution.
  2. Each time a new node is added; stop after the nth node is added.
     - Assign a random fitness to the new node according to the lognormal distribution.
     - Add m edges linking the new node to m distinct existing nodes such that the probability Π_i for connecting to an existing node i is taken to be proportional to its fitness Φ_i:

$$\Pi_i = \frac{\Phi_i}{\sum_j \Phi_j}. \tag{2.15}$$

LNFA is almost identical to the BA model, the difference being that fitness information is used in place of degree information. Although this difference seems to be minor, it makes a fundamental shift in how the network is formed. To make this point, recall that in the BA protocol, the degree of a new node at the time it joins the network is small (m) and so it takes this node a long time before it may become a preferential choice for future new nodes to attach to. In LNFA, the new node may have a large fitness at the time it joins the network, making itself a preferential choice immediately. This is naturally reasonable because the attractiveness of a node may not result from how many nodes it is connected to; it may instead result from the "inner self" factors such as the personality of a person in a friendship network and his or her age.

As demonstrated in Fig. 2.2, LNFA can be used to generate a large spectrum of networks as seen in the real world and it is possible to do so by varying the

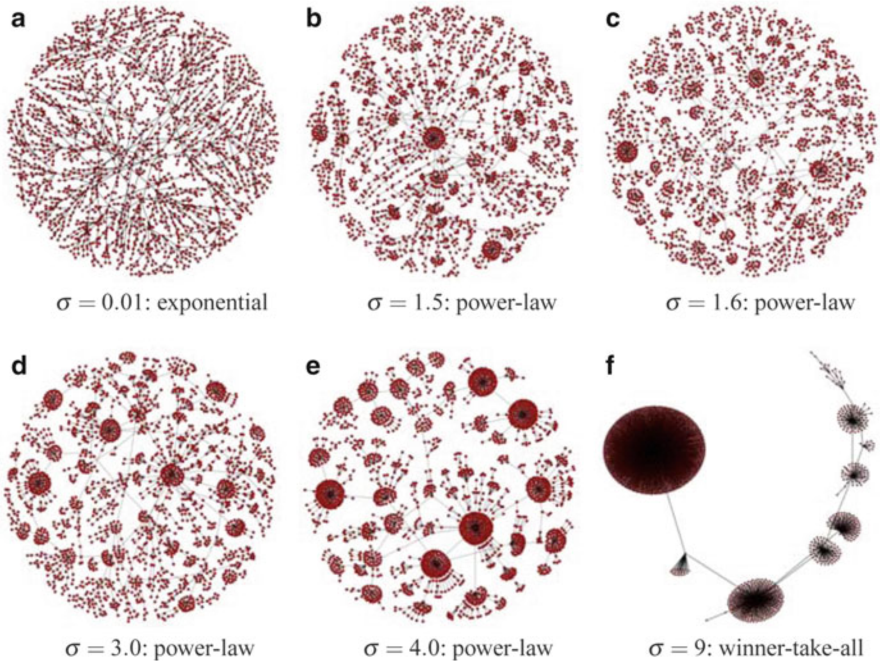

Fig. 2.2 2,000-node networks resulting from increasing σ. Transitions from exponential to power-law to winner-take-all graphs are observed

parameter σ. Consider two extreme cases of this parameter. In the first case, if σ is zero, nodes have exact same fitnesses and so, basing on Formula 2.15, each time a new node joins the network it chooses an existing node as neighbor with equal chance. This construction is simply the random graph model of [7] which yields a network with exponential degree distribution. On the other extreme, if σ is increased to reach a certain threshold, few nodes will stand out having very large fitnesses while all the other nodes will have very low fitnesses (zero or near zero; see Fig. 2.5). Consequently, an extremely high number of connections will be made to just a single node, resulting in a monopolistic network; this "winner-take-all" degree pattern is also observed in the real world [2]. Between these two extreme cases (exponential and monopolistic) we find a spectrum of power-law networks.

The transition between the regimes of degree distributions is not discontinuous. It is observed that when σ reaches 1 power-law degree distributions emerge and remain until σ is beyond 4. After that ($\sigma > 4$), the network becomes less power-law and more of a winner-take-all network. This is illustrated in Fig. 2.3 which shows the exponent λ as a function of σ. This figure plots the exponent values for 20 different networks of size ranging from 10,000 nodes to 20,000 nodes, that are constructed with the same σ value (ranging from 1 to 4.5). Here exponent values are not shown for the case $\sigma < 1$ because the corresponding resultant networks are not power-law; they are exponential random graphs. Seen from the figure, as σ increases from

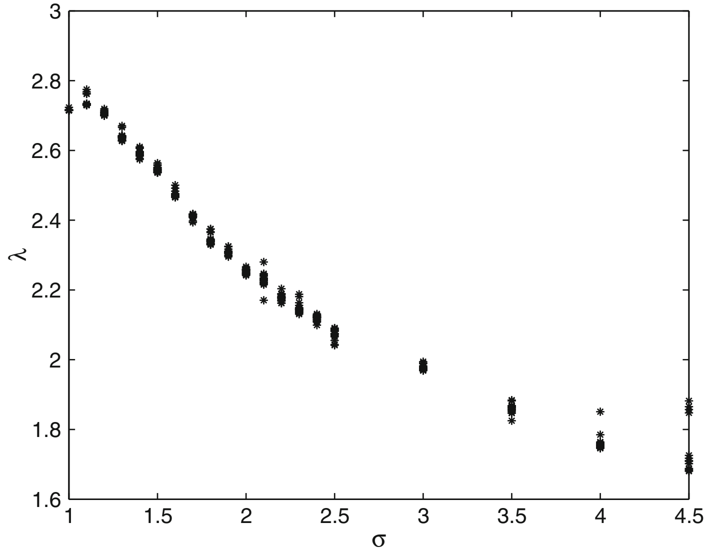

Fig. 2.3 Power-law exponent λ as a function of parameter σ. For each value of σ, 20 networks with size ranging from 10,000 to 200,000 nodes were considered. Each sample ("dot") in the plot shows the exponent for each network (for each value of σ)

1 to 3.5, there is remarkable consistency among all the 20 networks considered: they are all power-law with nearly identical exponent (it is noted that in the figure the increase in fluctuations seen at larger σ ($\sigma \geq 4$) is a consequence of the cross-over from power-law to a more monopolistic degree distribution). For example, when $\sigma = 2$, all networks have the same exponent 2.25. This implies that with LNFA the power-law degree distribution will emerge quickly as the network grows (in our observation, networks with 10,000 nodes already show their scale-freeness). It is also important to see that when the power-law pattern is observed, the exponent ranges between 1.8 and 2.8, which is close to the range [2, 3] typically observed in the real world. The exponent is monotonically decreasing as a function of σ.

2.6 Fitness-Based Model with Mutual Benefit

It is argued that in many cases of interest the power-law degree behavior is neither related to dynamical properties nor to preferential attachment. Also, the concept that the likelihood of a new node to link to an existing node depends solely on the latter's fitness might only apply to certain networks. The model proposed in [6] is suited for complex networks where it is the mutual benefit that makes two nodes link to each other. This model puts the emphasis on fitness itself without using the preferential attachment rule. Further, it is a static model building a network by growing links instead of growing nodes.

Specifically, the network construction algorithm starts with a set of n isolated nodes, where n is the size of the network to be built. Similar to the models discussed in the previous sections, each node i has a fitness Φ_i drawn from some distribution ρ. Then, for every pair of nodes, i and j, a link is drawn with a probability $f(\Phi_i, \Phi_j)$ which is some joint function of Φ_i and Φ_j. This model can be considered a generalization of the ER random graph model [9]. Rather than using an identical link probability for every pair of nodes as in the ER model, here two nodes are linked with a likelihood depending jointly on their fitnesses. A general expression for $p(k)$ can be easily derived. Indeed, the mean degree of a node of fitness Φ is

$$k(\Phi) = n \int_0^\infty f(\Phi, x)\rho(x)\,\mathrm{d}x = nF(\Phi). \tag{2.16}$$

Assuming F to be a monotonous function, and for large enough n, the following form can be obtained for $p(k)$:

$$p(k) = \rho\left[F^{-1}\left(\frac{k}{n}\right)\right]\frac{\mathrm{d}}{\mathrm{d}k}F^{-1}\left(\frac{k}{n}\right). \tag{2.17}$$

Thus, one can choose an appropriate formula for ρ and f to achieve a given distribution for $p(k)$. It is shown that a power law for $p(k)$ will emerge if fitness follows a power law and the linking probability for two nodes is proportional to the product of their fitnesses. For example, one can choose

$$f(\Phi_i, \Phi_j) = (\Phi_i \Phi_j)/\Phi_{\max}^2$$

and

$$\rho(\Phi) \propto \Phi^{-\beta},$$

(corresponding to a Zipf's behavior with Zipf coefficient $\alpha = 1/(\beta - 1)$).

In the case that fitness does not follow a power law, it is possible to find a linking function that will result in a power-law degree distribution. For example, considering an exponential fitness distribution, $\rho(\Phi) \propto \mathrm{e}^{-\Phi}$ (representing a Poisson distribution), one can choose

$$f(\Phi_i, \Phi_j) = \theta[\Phi_i + \Phi_j - z(n)]$$

where θ is the usual Heaviside step function and $z(n)$ is some threshold, meaning that two nodes are neighbors only if the sum of their fitness values is larger than the threshold $z(n)$. Using these rules, the degree distribution has a power law with exponent $\lambda = 2$. This is interesting because it shows that power-law networks can emerge even if fitness is not power-law. The same behavior also emerges if a more generic form of the linking function is used,

$$f(\Phi_i, \Phi_j) = \theta[\Phi_i^c + \Phi_j^c - z^c(n)],$$

(where c is an integer number); however, the power-law degree distribution has logarithmic corrections in some cases.

Closely related to the above model is the work of [22] which assumes the same concept that the linking probability is a joint function of the fitnesses of the end nodes. In this related work, it is concluded that for any given fitness distribution $\rho(\Phi)$ there exists a function $g(\Phi)$ such that the network generated by $\rho(\Phi)$ and $f(\Phi_i, \Phi_j) = g(\Phi_i)g(\Phi_j)$ is power-law with an arbitrary real exponent. Mutual-benefit based linking can also be combined with preferential attachment to drive the growth of a power-law network as shown in the work of [4].

2.7 Summary

This chapter has provided a review of existing models aimed at constructing power-law complex networks, that are inspired by the idea that there is some intrinsic fitness associated with a node to drive its evolution in the network. This fitness might be causal to why a node has a high degree or low, or might be an independent factor which together with the node's degree affect the node's ability to compete for links. The models discussed differ in how they interpret fitness and its influence on growing the network. It is suggested in [5, 16] that both degree and fitness jointly determine the growth rate of node degree. This may apply to complex networks such as a social network where a person's attractiveness is a combination of both his or her experience (represented by node degree) and talent (represented by fitness), or the WWW network where a web page is popular because its long time staying online (represented by node degree) and quality of its content (represented by fitness). A different approach is proposed in [6,13,22] where all the attractiveness factors associated with a node can be combined into a single factor (fitness). While the models in [6, 22] are motivated by static networks where two nodes require mutual benefit in order to make a connection, the lognormal fitness model in [13] is suitable to explain growing networks where a node wants to be a neighbor of another solely because of the latter's fitness, regardless of the former's. Although none of these models is one-size-fits-all, they do represent a vast population of complex networks. The current models, however, assume that fitness is an intrinsic factor that does not change over time. In practice, there are cases of networks where the overall attractiveness of a node might increase for one period of time and decrease for another. For the future work, it is thus an interesting research problem to explore fitness models that allow fitness to have its own evolution. The future research should also pay great attention to (in)validating theoretical models with the data collected from real-world networks.

Acknowledgements The authors would like to thank UMass Boston colleagues, Shilpa Ghadge, Bala Sundaram, and Timothy Killingback, for valuable discussions regarding the lognormal fitness model presented in this chapter.

References

1. Albert, R., Jeong, H., Barabási, A.L.: Diameter of the world-wide-web. Nature **400**, 107–110 (1999)
2. Barabási, A.: The physics of the web. Physics World **14-7**, 33–38 (2001)
3. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. Science **286**, 509–512 (1999)
4. Bedogne, C., Rodgers, G.J.: Complex growing networks with intrinsic vertex fitness. Phys. Rev. E **74**, 046,115 (2006)
5. Bianconi, G., Barabsi, A.L.: Competition and multiscaling in evolving networks. Europhysics Letters **54**, 436–442 (2001). DOI 10.1209/epl/i2001-00260-6. URL http://dx.doi.org/10.1209/epl/i2001-00260-6
6. Caldarelli, G., Capocci, A., Rios, P.D.L., Munoz, M.: Scale-free networks from varying vertex intrinsic fitness. Phys. Rev. Lett. **89-25**, 258,702 (2002)
7. Callaway, D.S., Hopcroft, J.E., Kleinberg, J.M., Newman, M.E.J., Strogatz, S.H.: Are randomly grown graphs really random? Physical Review E **64**, 041,902 (2001). URL http://www.citebase.org/abstract?id=oai:arXiv.org:cond-mat/0104546
8. Dangalchev, C.: Generation models for scale-free networks. Physica A: Statistical Mechanics and its Applications **338**, 659–671 (2004)
9. Erdós, P., Rényi, A.: On random graphs. Publicationes Mathematicae **6**, 290–297 (1959)
10. Erdós, P., Rényi, A.: On the evolution of random graphs. Publ. of Math. Inst. of the Hungarian Acad. of Sci. **5**, 17–61 (1960)
11. Erdós, P., Rényi, A.: On the strength of connectedness of a random graph. Acta Mathematia Scientia Hungary **12**, 261–267 (1961)
12. Faloutros, M., P. Faloutros, Faloutros, C.: On power-law relationship of the internet topology. ACM SIGCOM 99, Comp. Comm. Rev. **29**, 251–260 (1999)
13. Ghadge, S., Killingback, T., Sundaram, B., Tran, D.A.: A statistical construction of power-law networks. International Journal on Parallel, Emergent, and Distributed Systems **25**, 223–235 (2010)
14. Jeong, H., Mason, S., Oltvai, R., Barabási, A.: Lethality and centrality in protein networks. Nature **411**, 41 (2001)
15. Jeong, H., Tombor, B., Albert, R., Oltvai, N., Barabási, A.: The large-scale organization of metabolic networks. Nature **607**, 651 (2000)
16. Kong, J.S., Sarshar, N., Roychowdhury, V.P.: Experience versus talent shapes the structure of the web. PNAS **105**(37), 13724–13729 (2008)
17. Kumar, R., Raghavan, P., Rajagopalan, S., Sivakumar, D., Tomkins, A., Upfal, E.: Stochastic models for the web graph. In: FOCS ’00: Proceedings of the 41st Annual Symposium on Foundations of Computer Science, p. 57. IEEE Computer Society, Washington, DC, USA (2000)
18. Li, L., Alderson, D., Doyle, J., Willinger, W.: Towards a Theory of Scale-Free Graphs: Definition, Properties, and Implications. In Proceedings of Internet Mathematics (2005)
19. Liljeros, F., Edling, C.R., Amaral, L.A.N., Stanley, H.E., Aberg, Y.: The web of human sexual contacts. Nature **411**, 907 (2001)
20. Moore, C., Ghoshal, G., Newman, M.E.J.: Exact solutions for models of evolving networks with addition and deletion of nodes. Phys. Rev. E **74**(3), 036,121 (2006). DOI 10.1103/PhysRevE.74.036121
21. Neuman, M.: The structure of scientific collaboration networks. Proc. Nat. Acad. Sci. USA **98**, 404–409 (2001)
22. Vito, D., Caldarelli, G., Butta, P.: Vertex intrinsic fitness: How to produce arbitrary scale-free networks. Phys. Rev. E **70**, 056,126 (2004)
23. Watts, D., Strogatz, S.: Collective dynamics of “small-world” networks. Nature **393**, 440–442 (1998)

http://www.springer.com/978-1-4614-0753-9

Handbook of Optimization in Complex Networks
Theory and Applications
Thai, M.T.; Pardalos, P.M. (Eds.)
2012, XIV, 546 p., Hardcover
ISBN: 978-1-4614-0753-9