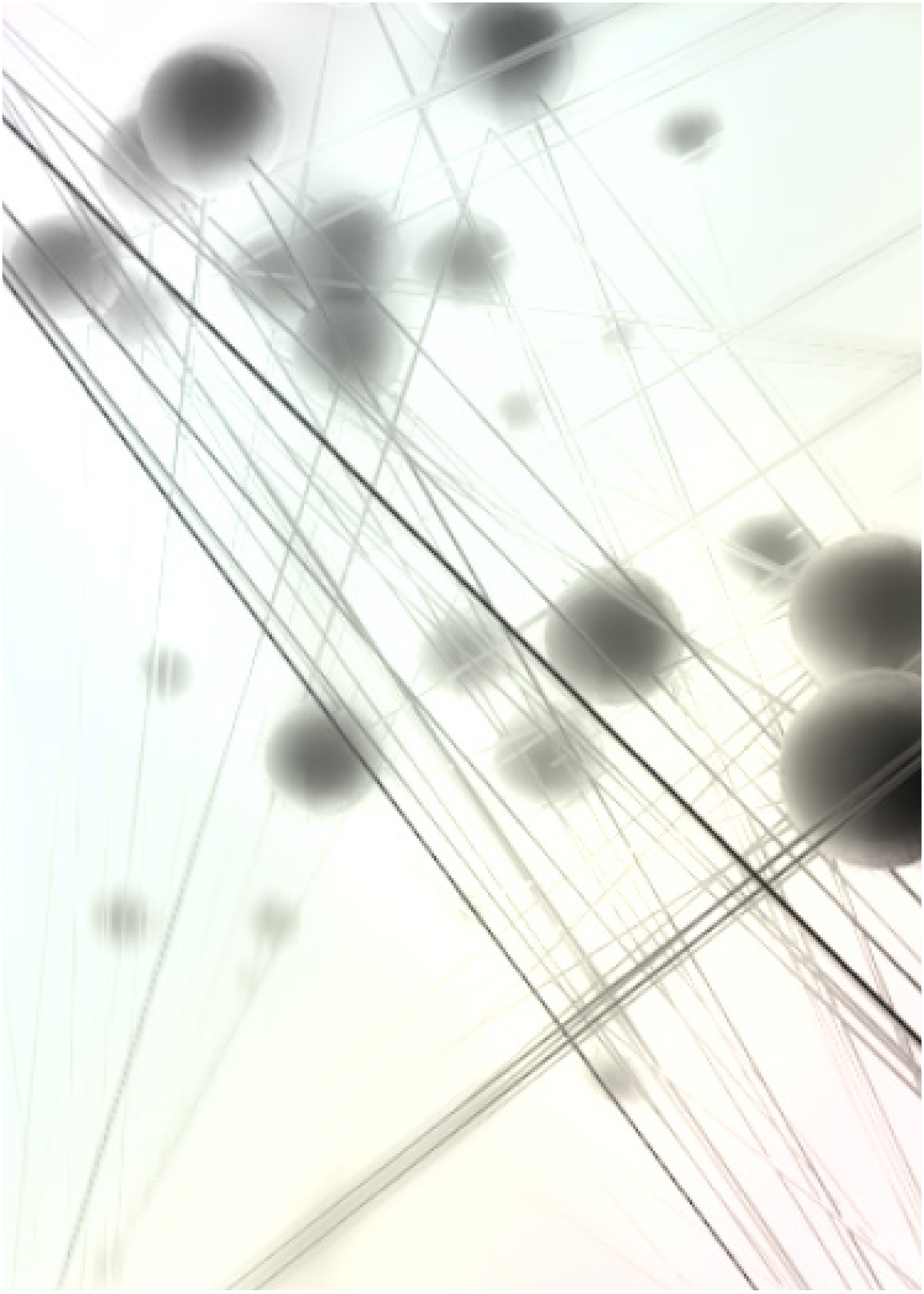
CHAPTER 5

THE BARABÁSI-ALBERT MODEL

Figure 5.0 (front cover)
Universality: G. Musella



SECTION 5.0

INTRODUCTION

Hubs represent the most striking difference between a random and a scale-free network. Their very existence raises several fundamental questions:

- Why does the random network model of Erdős and Rényi fail to reproduce the hubs and the power laws observed in many real networks?

- Why do so different systems as the WWW or the cell converge to a similar scale-free architecture?

The last question is particularly puzzling given the fundamental differences in the nature, origin, and scope of the systems that display the scale-free property:

- The nodes of the cellular network are proteins or metabolites, while the nodes of the WWW are documents, representing information without a physical manifestation.

- The links within the cell are binding interactions and chemical reactions, while the links of the WWW are URLs, or small segments of computer code.

- The history of these two systems could not be more different: the cellular network is shaped by 4 billion years of evolution, while the WWW is a few decades old.

- The purpose of the metabolic network is to chemically build the basic chemical components the cells need for life, while the purpose of the WWW is information access and delivery.

To understand why so different systems converge to a similar architecture we need to first understand the mechanism responsible for the emergence of the scale-free property. This is the main topic of this chapter. Giv-

en the major differences between the systems that display the scale-free property, the explanation must be simple and fundamental. The answers will change the way we view and model networks, forcing us to move from describing a network's topology to modeling the evolution of complex systems.

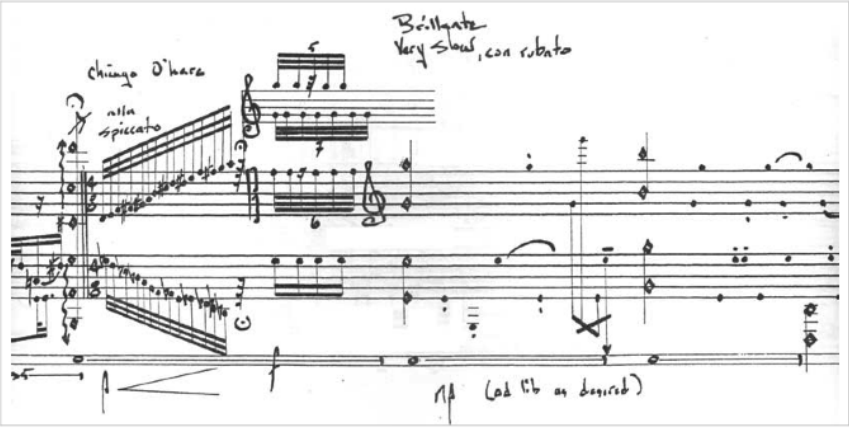


Figure 5.1
Scale-free sonata

Composed by Michael Edward Edgerton in 2003, *1 sonata for piano* was inspired by scale-free networks. The music obeys the principles of a growing network, incorporating growth and preferential attachment. The interplay between the music and networks is explained by the composer:

"6 hubs of different length and procedure were distributed over the 2nd and 3rd movements. Musically, the notion of an airport was utilized by diverting all traffic into a limited landing space, while the density of procedure and duration were varied considerably between the 6 differing occurrences."

The Image shows the beginning of what Edgerton calls Hub #5.

SECTION 5.1

GROWTH AND PREFERENTIAL ATTACHMENT

Our journey towards understanding the origin of the scale-free property by asking: why are hubs and power laws absent from the model? The answer emerged in 1999, highlighting two hidden assumptions of the Erdős-Rényi model, each of which are violated in real networks [1]. Next we discuss these two assumptions separately.

NETWORKS EXPAND THROUGH THE ADDITION OF NEW NODES

The random network model assumes that we have a *fixed* number of nodes, N. The role of the modeler is to connect these nodes, while keeping N unchanged. Yet, in most real networks the number of nodes is not fixed, but continually *grows* thanks to the addition of new nodes. Let us consider a few examples:

- In 2001 the WWW had a single node, the first webpage build by Tim Berners-Lee, the creator of the Web. Today the Web has over a trillion (10^{12}) documents, an extraordinary number that was reached through the continuous addition of new documents by millions of individuals and institutions Fig. 5.2a.

- The collaboration and citation networks continually expand through the publication of new research papers Fig. 5.2b.

- The Hollywood actor network continues to expand through the release of new movies Fig. 5.2c.

- At first the protein interaction network within our cells may appear to be static, as we inherit our genes (and hence our proteins) from our parents. Yet, it is not: the number of genes grew from a few to the over 20,000 genes present today in a human cell over four billion years.

Consequently, if we wish to model these networks, we cannot resort to a static model. Rather our approach must acknowledge that networks are the product of a steady growth process.

NODES PREFER TO LINK TO THE MORE CONNECTED NODES

The random network model assumes that we randomly choose the interaction partners of a node. In most real networks, however, new nodes prefer to link to the more connected nodes, a process called *preferential attachment*. Consider a few examples:

- We are familiar with only a tiny fraction of the trillion or more documents available on the WWW. The nodes we know are not entirely random, but we all heard about Google and Facebook, but we rarely encounter the billions of less-prominent nodes that populate the Web. As our knowledge is biased towards the more connected nodes, we are more likely to link to a high-degree node than to a node with only few links.

- With more than a million scientific papers published each year, no scientist can attempt to read them all. The more cited is a paper, the more likely that we will notice it. Therefore, our citations are biased towards the more cited publications, representing the high-degree nodes of the citation network.

- The more movies an actor has played in, the more familiar is a casting director with her skills. Hence, the higher the degree of an actor in the actor network, the higher are the chances that she will be considered for a new role.

In summary, the random network model differs from real networks in two important characteristics:

GROWTH

While the random network model assumes that the number of nodes, N, is fixed (time invariant), real networks are the result of a growth process that continuously increases N.

PREFERENTIAL ATTACHMENT

While nodes in random networks randomly choose their interaction partner, in real networks new nodes prefer to link to the more connected nodes.

There are many other differences between real and random networks, some of which will be discussed in the coming chapters. Yet, as we show next, growth and preferential attachment have a particularly important role shaping a network's degree distribution.

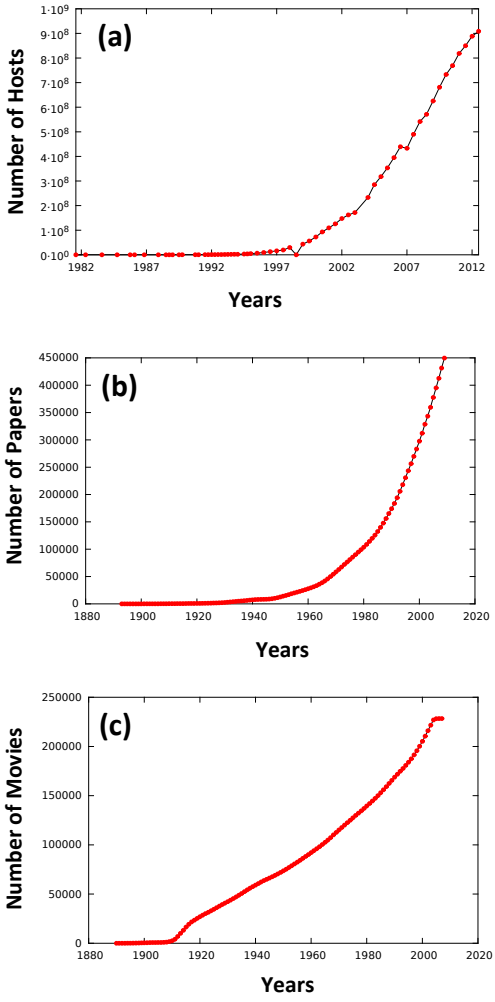


Figure 5.2
The growth of networks

(a) The evolution of the number of WWW hosts, documenting the Web's rapid growth. After *http://www.isc.org/solutions/survey/history*.

(b) The number of scientific papers published in *Physical Review* journals since the journal's funding in 1893. The observed growth drives the growth of both the science collaboration network as well as the citation network. Over the century the *Physical Review* portfolio has split several times, responding to the exponential growth of the number of research papers and to specialization. Today the corpus features *Physical Review Letters*, *Physical Review A, B, C, D, E, X* and *Reviews of Modern Physics*.

(c) Number of movies listed in IMDB.com, reflecting the growth of the Hollywood movie enterprise, and with that the growth of the actor network.

BOX 5.1

PREFERENTIAL ATTACHMENT: A BRIEF HISTORY

Preferential attachment has emerged repeatedly in mathematics and social sciences. Consequently today we can encounter it under different names in the scientific literature:

- It made its first appearance in 1923 in the celebrated urn model of the Hungarian mathematician György Pólya (1887-1985) [2], proposed to explain the nature of certain distributions. Hence, in mathematics preferential attachment is often called a *Pólya process.*

- George Udmy Yule (1871-1951) in 1925 used preferential attachment to explain the power-law distribution of the number of species per genus of flowering plants [3]. Hence, in statistics preferential attachment is often called a *Yule process*.

- Rober Gibrat (1904-1980) in 1931 proposed that the size and the growth rate of a firm are independent. Hence, larger firms grow faster [4]. Called *proportional growth*, this is a form of preferential attachment.

- George Kinsley Zipf (1902-1950) in 1941 used preferential attachment to explain the fat tailed distribution of wealth in the society [5].

- Modern analytical treatments of preferential attachment use of the master equation approach pioneered by the economist Herbert Alexander Simon (1916-2001). Simon used preferential attachment in 1955 to explain the fat-tailed nature of the distributions describing city sizes, word frequencies in a text, or the number of papers published by scientists [6].

- Building on Simon's work, Derek de Solla Price (1922-1983) used preferential attachment to explain the citation statistics of scientific publications, referring to it as *cumulative advantage* [7].

- In sociology preferential attachment is often called the *Matthew effect*, named by Robert Merton (1910-2003) [8] after a passage in the Gospel of Matthew: "For everyone who has will be given more, and he will have an abundance. Whoever does not have, even what he has will be taken from him."

- The term *preferential attachment* was introduced in the 1999 paper by Barabási and Albert [1] to explain the ubiquity of power laws in networks.

Note that the distributions characterized from Pólya to Merton describe scalar quantities, like the number of individuals with the same income or the size of cities. In contrast the Barabási-Albert model aims to describe networks. Networks have a wide array of topological characteristics that are absent from scalar distributions, but which are deeply affected by the power-law nature of the degree distribution.

THE BARABÁSI-ALBERT MODEL

The recognition that growth and preferential attachment coexist in real networks has lead to the introduction of a minimal model capable of generating networks with power-law degree distribution [1]. The model is defined as follows:

We start with m_0 nodes, the links between which are chosen arbitrarily, as long as each node has at least one link. The network develops following two steps Fig. 5.3:

(A) GROWTH

At each timestep we add a new node with m ($\leq m_0$) links that connect the new node to m nodes already in the network.

(B) PREFERENTIAL ATTACHMENT

The probability $\Pi(k)$ that one of the links of the new node connects to node i depends on the degree k_i of node i as

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}. \qquad (5.1)$$

Preferential attachment is a probabilistic rule: a new node is free to connect to any node in the network, whether it is a hub or has a single link. Eq. 5.1 implies, however, that if a new node has a choice between a degree-two and a degree-four node, it is twice as likely that it connects to the degree-four node. The model defined by steps (A) and (B) is called the *Barabási-Albert* model after the authors of the paper that introduced it in 1999 [1]. One may also encounter it in the literature as the BA model or the *scale-free* model. After t timesteps the Barabási-Albert model generates a network with $N = t + m_0$ nodes and $m_0 + mt$ links. As Fig. 5.4 shows, the network generated by the model has a power-law degree distribution, a with a degree exponent γ=3.

As Fig. 5.3 indicates, while most nodes in the network have only a few links, a few gradually turn into hubs. The hubs are the result of a *rich-gets-*

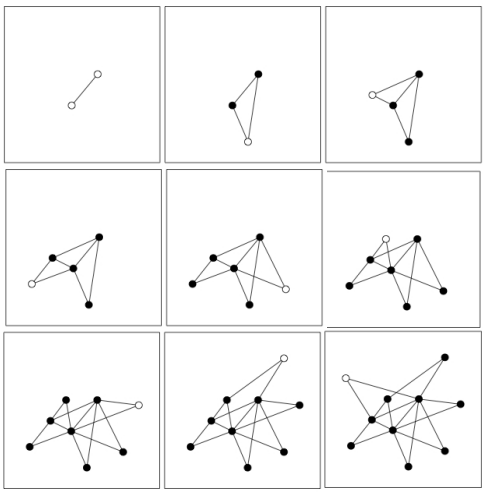


Figure 5.3
Time evolution of the Barabási-Albert model

The sequence of images shows the gradual emergence of a few highly connected nodes, or hubs, through growth and preferential attachment. White circles mark the newly added node to the network, which decides where to connect its two links (m=2) through preferential attachment Eq. 5.1. After [9].

richer phenomenon: due to preferential attachment new nodes are more likely to connect to the more connected nodes than to the smaller degree nodes. Hence, the more connected nodes will acquire links at the expense of the less connected nodes, eventually turning into hubs.

In summary, the Barabási-Albert model indicates that two simple mechanisms, growth and preferential attachment, are responsible for the emergence of networks with a power-law degree distribution. The origin of the power law and the associated hubs is a *rich-gets-richer phenomena* induced by the coexistence of these two ingredients. Yet, to understand the model's behavior and to quantify the emergence of the scale-free property, we need to describe the model's mathematical properties, which is the subject of the next section.

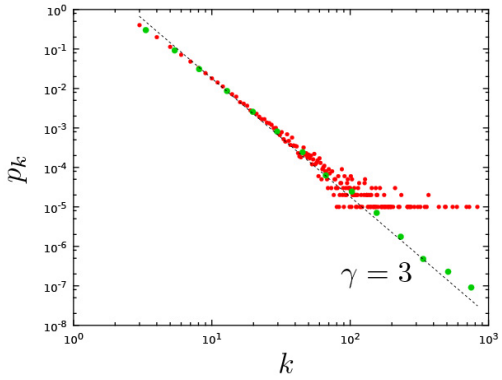


Figure 5.4
The degree distribution

The degree distribution of a network generated by the Barabási-Albert model. The plot shows p_k for a single network of size N=100,000 and m=3. It shows both the linearly-binned (red symbols) as well as the log-binned version (green symbols) of p_k. The straight line is added to guide the eye and has slope γ=3, corresponding to the resulting network's degree distribution.

To: ~~ALB~~ RÉKA <u>ALBERT</u>

from A.L. Barabási

<u>Réka</u>: Próbáld meg beprogramozni a következő kis modellt; ami esetleg megadja a várt hatványfüggvényt:

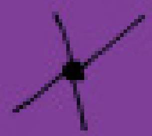
<u>MODEL 1</u>: $t=0$: (m) vertices, (no) edges.

at time (t): always add a <u>new</u> <u>vertex</u>, which has m edges (non-directed) coming out of it.
The ends of the edges will be randomly connected to the existing vertices in the system.

<u>thus</u>: $t=1$: we have $(m+1)$ vertices, and (m) edges, coming out of the new vertex.

$t=2$: $(m+2)$ vertices; $2m$ edges

$t=\tau$ $(m+\tau)$ vertices; τm edges

thus a new vertex always has (m) edges 
whose ends are connected randomly to the already present vertices in the system.

The average connectivity after time τ: $\langle k \rangle = \frac{\tau m}{\tau + m} \to \underline{m}$

Determine $P(k)$ at different times $t_1, t_2, \dots, t_n$
(where $t_1 \dots$ are large),
and see if it is exponential or power law

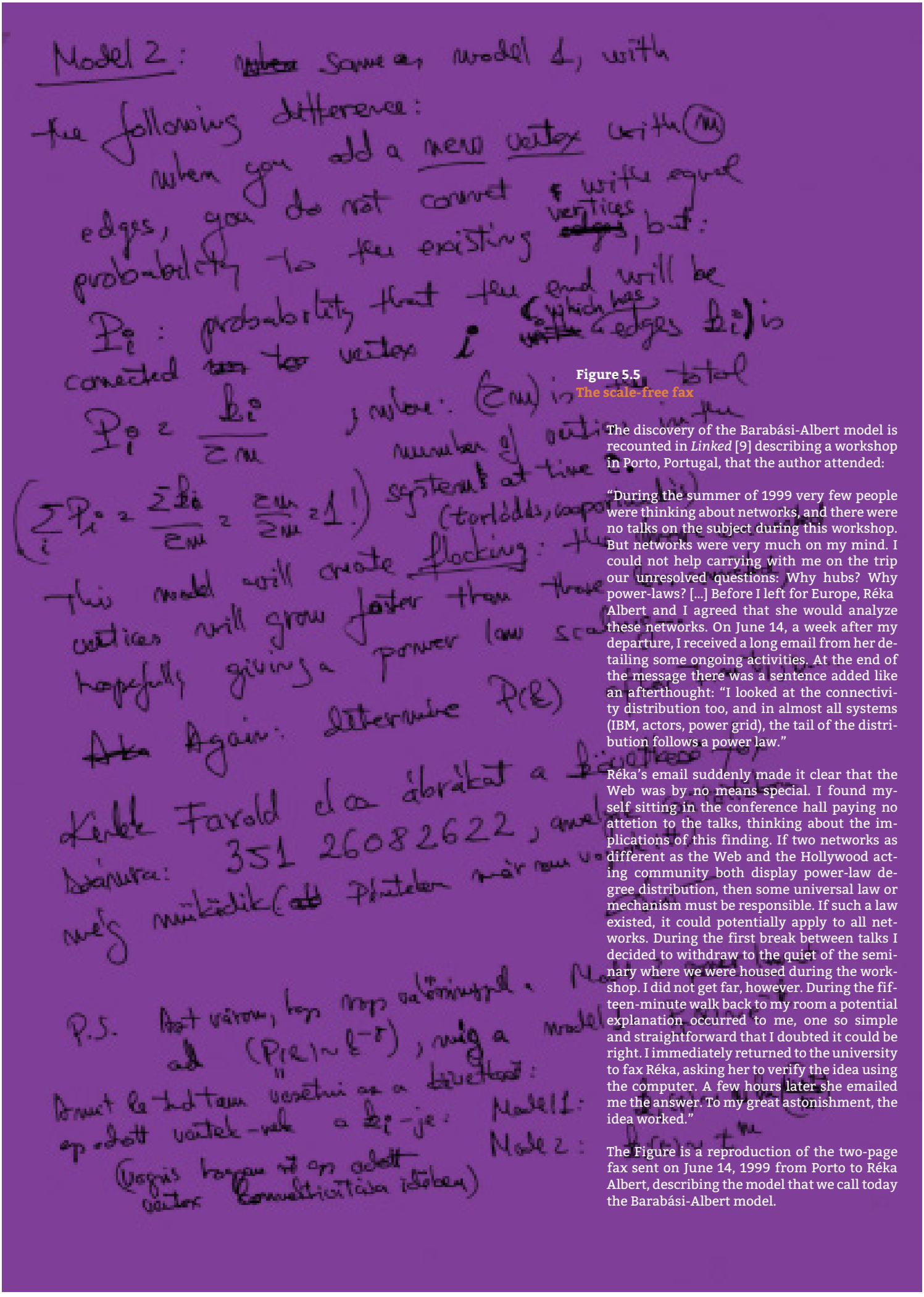


Figure 5.5
The scale-free fax

The discovery of the Barabási-Albert model is recounted in *Linked* [9] describing a workshop in Porto, Portugal, that the author attended:

"During the summer of 1999 very few people were thinking about networks, and there were no talks on the subject during this workshop. But networks were very much on my mind. I could not help carrying with me on the trip our unresolved questions: Why hubs? Why power-laws? [...] Before I left for Europe, Réka Albert and I agreed that she would analyze these networks. On June 14, a week after my departure, I received a long email from her detailing some ongoing activities. At the end of the message there was a sentence added like an afterthought: "I looked at the connectivity distribution too, and in almost all systems (IBM, actors, power grid), the tail of the distribution follows a power law."

Réka's email suddenly made it clear that the Web was by no means special. I found myself sitting in the conference hall paying no attetion to the talks, thinking about the implications of this finding. If two networks as different as the Web and the Hollywood acting community both display power-law degree distribution, then some universal law or mechanism must be responsible. If such a law existed, it could potentially apply to all networks. During the first break between talks I decided to withdraw to the quiet of the seminary where we were housed during the workshop. I did not get far, however. During the fifteen-minute walk back to my room a potential explanation occurred to me, one so simple and straightforward that I doubted it could be right. I immediately returned to the university to fax Réka, asking her to verify the idea using the computer. A few hours later she emailed me the answer. To my great astonishment, the idea worked."

The Figure is a reproduction of the two-page fax sent on June 14, 1999 from Porto to Réka Albert, describing the model that we call today the Barabási-Albert model.

BOX 5.2

THE MATHEMATICAL DEFINITION OF THE BARABÁSI-ALBERT MODEL

The definition of the Barabási-Albert model provided in this chapter leaves many mathematical details of the model unspecified:

- It does not specify the precise initial configuration of the first m_0 nodes.

- It does not specify whether the m links assigned to a new node are added one by one, independent of each other, or simultaneously. These problems were recognized by Riordan and Bollobás [10], who offered a definition that addresses these shortcomings. In contrast with the original model, Riordan and Bollobás allows for multiple edges and loops, showing later that their number will be negligible. The resulting model, called the Linearized Chord Diagram (LCD), is defined as follows:

Consider a fixed sequence of nodes v_1, v_2, ..., where the degree of the node v_i is k_i. We build a graph $(G_1^{(t)})_{t \geq 0}$ so that $G^{(t)}$ is a graph on v_i , $1 \leq i \leq t$ as follows: start with $G_1^{(0)}$ corresponding to an empty graph with no nodes, or with $G_1^{(1)}$ graph with one node and one loop. Given $G_1^{(t-1)}$ generate $G_1^{(t)}$ by adding the node v_t together with a single link between v_t and v_i, where i is chosen with probability

$$p_r(i=s)=\begin{cases} k_{G^{(t-1)}_1}\dfrac{k(v_s)}{(2t-1)}, & if\ 1 \leq s \leq t-1 \\ \dfrac{1}{(2t-1)}, & if\ s=t \end{cases} \qquad (5.2)$$

That is, we place a link from node v_t to node $v_{i,}$ where the probability that node i is chosen as the target of this new link is proportional to its degree k_i at the time, the new link already contributing to the degree of v_t. Hence, the new node v_t can also link to itself with probability $1/(2t - 1)$. For $m > 1$, we add m links from v_t one by one, counting the previous links together with the outward half of the newly added link as contributing to the degrees.

DEGREE DYNAMICS

To understand the time evolution of the Barabási-Albert model, we first focus on the time-dependent degree of a node [11]. In the model a node has a chance to increase its degree each time a new node enters the network. When a new node joins the network, it will link to m of the $N(t)$ nodes present in the system. The probability that it chooses node i is given by Eq. 5.1. Assuming that k_i is a time-dependent continuous real variable, the rate at which node i acquires links follows the equation

$$\frac{\partial k_i}{\partial t} = m\Pi(k_i) = m\frac{k_i}{\sum_{j=1}^{N-1} k_j}. \tag{5.3}$$

The coefficient m describes that each new node arrives with m links. Hence, node i has m chances to be chosen. The sum in the denominator of Eq. 5.3 goes over all nodes in the network except the newly added node, thus

$$\sum_{j=1}^{N-1} k_j = 2mt - m. \tag{5.4}$$

Therefore Eq. 5.4 becomes

$$\frac{\partial k_i}{\partial t} = \frac{k_i}{2t-1}. \tag{5.5}$$

For large t the term (-1) can be neglected in the denominator, obtaining

$$\frac{\partial k_i}{k_i} = \frac{1}{2}\frac{\partial t}{t}. \tag{5.6}$$

By integrating Eq. 5.6 and using the fact that $k_i(t_i)=m$, meaning that node i joins the network at time t_i with m links, we obtain

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta}. \tag{5.7}$$

The exponent β is the network's dynamical exponent and has the value $\beta = \frac{1}{2}$. Eq. 5.7 offers a number of predictions:

- The degree of each node increases following a power-law with the same dynamical exponent β =1/2 Fig. 5.6, implying that all nodes follow the same growth law.

- The growth in the degrees is sublinear (i.e. $\beta < 1$). In contrast in the Erdős-Rényi model $\langle k \rangle$ increases as $k_i \sim t$ if we add links one by one to the network. The sublinear nature of Eq. 5.7 is a consequence of the growing nature of the Barabási-Albert model: each new node has more nodes to link to than the previous nodes. Hence, with time each node competes for links with an increasing pool of nodes.

- The earlier node i was added, the higher is its degree $k_i(t)$. Hence, hubs are large not because they grow faster, but because they arrived earlier, a phenomenon called *first-mover advantage* in marketing and business.

- The growth rate of a node (i.e. the rate at which the node i acquires new links) is given by the derivative of Eq. 5.7

$$\frac{dk_i(t)}{dt} = \frac{m}{2}\frac{1}{\sqrt{t_i t}}, \tag{5.8}$$

  indicating that older nodes acquire more links in a unit time (as they have smaller t_i), as well as that the rate at which a node acquires links decreases with time as $t^{-1/2}$. Hence, less and less links go to a node.

Taken together, the Barabási-Albert model offers a dynamical description of a network's evolution, capturing the fact that in real networks nodes arrive one after the other, connecting to the earlier nodes. This sets up a competition for links during which the older nodes have an advantage over the younger nodes, eventually turning into hubs.

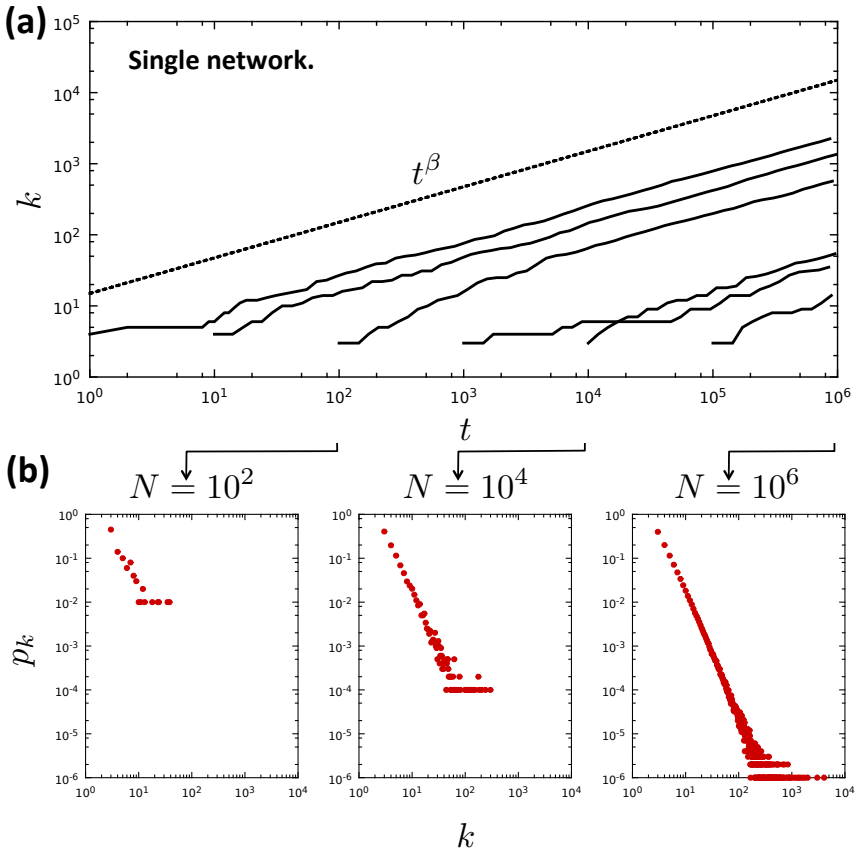


Figure 5.6
Degree dynamics

(a) Time dependence of the degrees of nodes added at time t =1, 10, 10^2, 10^3, 10^4, 10^5 (continuous lines from left to right). One can see that each node increases its degree following the same law **Eq. 5.7**. Also, at any moment the older nodes have higher degrees. The dotted line corresponds to the analytical prediction **Eq. 5.7** with $\beta = 1/2$.

(b) Degree distribution of the network after the addition of $N = 10^2$, 10^4, and 10^6 nodes, i.e. at time $t = 10^2$, 10^4, and 10^6 (illustrated by arrows in (a)). The larger the network, the more obvious is the power-law nature of the degree distribution. Note that p_k is plotted using linear binning, to better show the gradual emergence of the scale-free state.

BOX 5.3

THE MATHEMATICAL DEFINITION OF THE BARABÁSI-ALBERT MODEL

As we compare the predictions of the various network models with real data, we often have to decide how to measure time in networks. Real networks have evolved over rather different time scales: the first webpage was created in 1991, giving the WWW a history of a few decades at most. Given its trillion documents, this means that on average the WWW added more than a thousand nodes each second. In contrast the human cell is the result of 4 billion years of evolution; hence with roughly 20,000 genes, the cellular network added a node every 200,000 years. Given these enormous time-scale differences it seems impossible to use real time to compare the dynamics of these networks. Therefore, in network theory we use event time, that is, we advance time each time there is a change in the network topology. For example, in the Barabási-Albert model the addition of each new node corresponds to a new time step. Consequently in the model t=N. In more complicated models a distinct time step is assigned to each event—like the addition of a new node, the arrival of a new link, or the deletion of a node, any attempt to change the network topology. Obviously, if needed, we can establish a direct mapping between event time and the physical time.

DEGREE DISTRIBUTION

The distinguishing feature of networks generated by the Barabási-Albert model is their power-law degree distribution Fig. 5.4. In this section we calculate the functional form of p_k, helping us understand its origin. A number of analytical tools are available to calculate the model's degree distribution. The simplest is the continuum theory that we started developing in the previous section [1, 11]. It predicts that the degree distribution follows BOX 5.4,

$$p(k) \sim 2m^{1/\beta} k^{-\gamma} \tag{5.9}$$

with

$$\gamma = \frac{1}{\beta} + 1 = 3. \tag{5.10}$$

Eq. 5.9 tells us that the degree distribution follows a power law with exponent γ=3, in agreement with the numerical results shown in Fig. 5.4 and Fig. 5.7. In turn Eq. 5.10 links the degree exponent, γ, a quantity characterizing the network topology, to the dynamical exponent, β, that characterizes a node's temporal evolution. While the continuum theory predicts the correct degree exponent (γ=3), it fails to accurately predict the pre-factors of Eq. 5.9. This is why we use a proportional sign in Eq. 5.9, rather than equality. The exact degree distribution, with the correct pre-factors, can be obtained using a master [12] or rate equation [13] approach or calculated exactly using the LCD model [10] BOX 5.2. As we show in ADVANCED TOPICS 5.A, the exact form of the degree distribution of the Barabási-Albert model is

$$p_k = \frac{2m(m+1)}{k(k+1)(k+2)}. \tag{5.11}$$

Eq. 5.11 has several notable implications:

- For large k, Eq. 5.11 reduces to $p_k \sim k^{-3}$, or $\gamma = 3$, in line with Eq. 5.9 and Eq. 5.10.

- The degree exponent γ is independent of m, a prediction that agrees with the numerical results Fig. 5.7a.

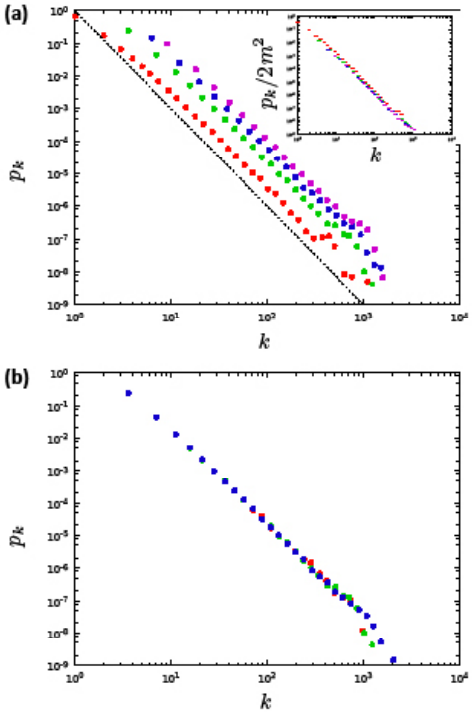


Figure 5.7
Probing the analytical predictions

(a) To show that p_k is independent of the parameters m and m_0, we generated networks with N=100,000 and m_0=m=1 (red), 3 (green), 5 (blue), and 7 (purple). The fact that the curves are parallel to each other indicates that γ is independent of m and m_0. The slope of the dashed line is -3. Inset: Eq. 5.11 predicts $p_k \sim 2m^2$, hence $p_k/2m^2$ should be independent of m. Indeed, by plotting $p_k/2m^2$ vs. k all curves in the main plot collapse into a single curve.

(b) The Barabási-Albert model predicts that p_k is independent of N. To test this we plot p_k for N = 50,000 (red), 100,000 (green), and 200,000 (blue), with m_0=m=3. The obtained p_k are practically indistinguishable, indicating that the degree distribution is time invariant.

- The power-law degree distribution observed in real networks describes systems of rather different age and size. Hence, a proper model should lead to a time-independent degree distribution. Indeed, Eq. 5.11 predicts that the degree distribution of the Barabási-Albert model is time independent, resulting in the emergence of a stationary scale-free state. Numerical simulations support this prediction, indicating that p_k observed for different t (or N) fully overlap Fig. 5.7b.

- Eq. 5.11 predicts that the coefficient of the power-law distribution is proportional to $m(m + 1)$ (or m^2 for large m), again confirmed by numerical simulations Fig. 5.7, inset).

In summary, the analytical calculations confirm that the Barabási-Albert model generates a power-law degree distribution, predicting the value of the degree exponent as γ=3. The exponent is independent of the parameters m and m_0. The calculations predict that the degree distribution is stationary (i.e. time invariant), explaining why networks with different history, size and age develop a similar degree distribution.

BOX 5.4

CONTINUUM THEORY

To calculate the degree distribution of the Barabási-Albert model we first determine the probability that the degree $k_i(t)$ of node i is smaller than a value k, i.e. $P(k_i(t) < k)$. Using Eq. 5.7, we can write

$$P(k_i(t) < k) = P\left(t_i > \frac{m^{1/\beta} t}{k^{1/\beta}}\right). \quad (5.12)$$

In the model we add the nodes at equal time intervals BOX 5.3. To capture this temporal uniformity we write the probablity that a node arrives at time ti as a random variable with a constant probability density

$$P(t_i) = \frac{1}{m_0 + t}. \quad (5.13)$$

Substituting Eq. 5.13 into Eq. 5.12 we obtain the cumulative distribution

$$P(k) = P\left(t_i \leq \frac{m^{1/\beta} t}{k^{1/\beta}}\right) = 1 - \frac{m^{1/\beta} t}{k^{1/\beta}(t + m_0)} \quad (5.14)$$

We obtain the degree distribution $p(k)$ by taking the derivative of the cumulative function, i.e.

$$p(k) = \frac{\partial P(k_i(t) < k)}{\partial k} = \frac{2m^{1/\beta} t}{m_0 + t} \frac{1}{k^{1/\beta + 1}}, \quad (5.15)$$

which for $t \gg m_0$ reduces to Eq. 5.9.

THE ABSENCE OF GROWTH OR PREFERENTIAL ATTACHMENT

The coexistence of growth and preferential attachment in the Barabási-Albert model raises an important question: are they both necessary for the emergence of the scale-free property? In other words, could we generate a scale-free network with only one of the two ingredients? To address these questions, next we discuss two limiting cases of the model, each containing only one of the two ingredients [1, 11].

MODEL A

To test the role of preferential attachment we keep the growing character of the network (ingredient A) and eliminate preferential attachment (ingredient B). Hence, Model A starts with m_0 nodes and evolves following these steps:

(1) Growth

At each time step we add a new node with $m(\leq m_0)$ links that links to m previous nodes.

(2) Random attachment

The probability that a new node links to a node with degree k_i is

$$\Pi(k_i) = \frac{1}{(m_0 + t - 1)}. \tag{5.16}$$

That is, $\pi(k_i)$ is independent of k_i, indicating that the new nodes choose randomly the nodes they link to. The continuum theory predicts that for Model A $k_i(t)$ increases logarithmically with time, i.e.

$$k(t) = m \ln\left(e \frac{m_0 + t + 1}{m_0 + t_i + 1}\right) \tag{5.17}$$

a much slower increase than the power law Eq. 5.7 derived earlier. Consequently the degree distribution becomes exponential Fig. 5.8a

$$p_k = \frac{e}{m} \exp\left(-\frac{k}{m}\right). \tag{5.18}$$

As an exponential function decays much faster than a power law, it does not support hubs. Therefore the lack of preferential attachment eliminates the network's scale-free character and the hubs.

MODEL B

To test the role of growth we next keep preferential attachment (ingredient B) and eliminate growth (ingredient A). Hence, Model B starts with N nodes and evolves following this step:

Preferential Attachment:
At each time step a node is selected randomly and connects to a node i with degree k_i already present in the network, where i is chosen with probability Eq. 5.1.

In Model B, the number of nodes remains constant during the network's evolution, but the number of links increases linearly with time. As a result the degree of each node also increases linearly with time Fig. 5.8b, inset

$$k_i(t) \simeq \frac{2}{N} t \qquad (5.19)$$

Indeed, in each time step we add a new link, without changing the number of nodes. At early times, when there are only a few links in the network (i.e. $L \ll N$), each new link connects previously unconnected nodes. In this stage the model's evolution is indistinguishable from the Barabási-Albert model with m=1. Numerical simulations show that in this regime the model develops a degree distribution with a power-law tail Fig. 5.8b. Yet, p_k is not stationary, as after roughly $T \simeq N^2$ time steps the network converges to a complete graph. Consequently, after a transient period ($t \ll N$) the node degrees start to converge to the average degree Eq. 5.19 and the degree distribution becomes peaked Fig. 5.8b. For $t \to N(N\text{-}1)/2$ the degree distribution becomes $p_k = \delta\,(N\text{-}1)$, i.e. the network turns into a complete graph in which all nodes have degree $k_{max}=N\text{-}1$. Therefore, in the absence of growth the network is not stationary, becoming a complete graph with time.

In summary, the failure of Models A and B to reproduce the empirically observed scale-free distribution indicates that growth and preferential attachment are simultaneously needed for the emergence of the scale-free property.

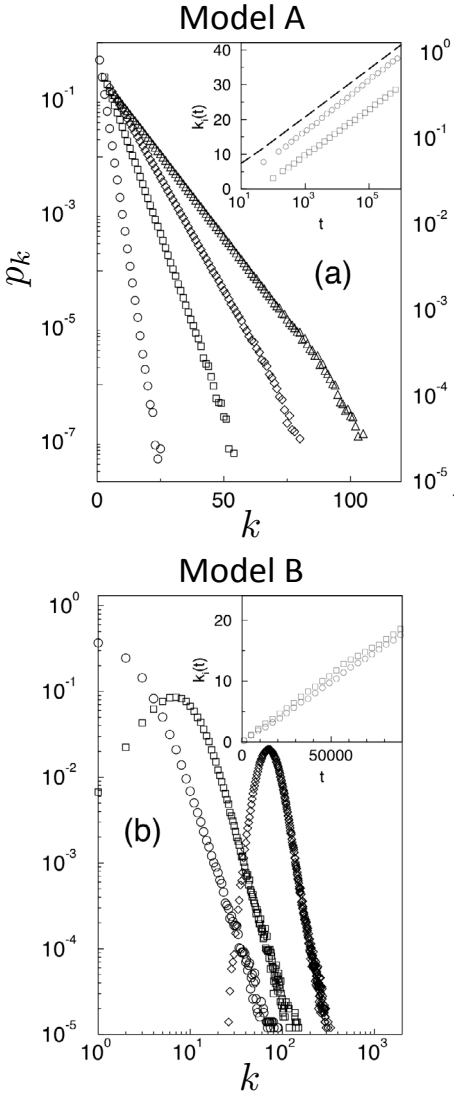


Figure 5.8
Model A and Model B

Numerical simulations of Model A and B, probing the role of growth andpreferential attachment.

(a) Degree distribution for Model A, that incorporates growth but lacks preferential attachment. The symbols correspond to m_0=m=1 (circles), 3 (squares), 5 diamonds), 7 (triangles) and N=800,000.

Inset: Time evolution of the degree of two vertices added at t^1=7 and t^2=97 for m_0=m=3. The dashed line follows $k_i(t)=m\,ln(m_0+t^{-1})$ as predicted by Eq. 5.16 for large t.

(b) Degree distribution for Model B, that lacks growth but incorporates preferential attachment, shown for N=10,000 and t=N (circles), t=5N (squares), and t=40N (diamonds).

Inset: Time dependence of the degrees of two vertices for system size N=10,000, indicating that $k_i(t)$ grows linearly, as predicted by Eq. 5.19. After [14].

MEASURING PREFERENTIAL ATTACHMENT IN REAL NETWORKS

In the previous sections we showed that growth and preferential attachment are responsible for the scale-free property. The presence of growth in real systems is obvious: all large networks arrived to their current size by continuously adding new nodes. But to convince ourselves that preferential attachment is also present in real networks, we need to detect it experimentally. In this section we show how to detect preferential attachment by measuring the $\pi(k)$ function in real networks. We start by noting that preferential attachment incorporates two hypotheses:

HYPOTHESIS 1

The likelihood to connect to a node depends on the node's degree k. This is in contrast with the random network model, for which $\pi(k)$ is independent of k.

HYPOTHESIS 2

The functional form of $\pi(k)$ is linear in k.

Both hypotheses can be tested by measuring $\pi(k)$. To be specific, we can determine $\pi(k)$ for networks for which we [14, 15] know the time at which each node joined the network, or we have at least two network maps collected at not too distant moments in time.

Consider a network for which we have two different maps, the first taken at time t and the other at time $t + \Delta t$ Fig. 5.9. During the Δt time frame some nodes did not change their degree, so for these $k(t+\Delta t) = k(t)$. For nodes that did alter their degree we measure the change $\Delta k_i = k_i\,(t+\Delta t\,) - k_i(t)$. According to Eq. 5.1, the relative change $\Delta k_i/\Delta t$ should follow

$$\frac{\Delta k_i}{\Delta t} \propto \Pi(k_i) \tag{5.20}$$

providing the functional form of preferential attachment. For Eq. 5.20 to be valid we must keep Δt small, so that the changes in Δk are relatively small. But it must not be too small so that there are still detectable differences between the two networks.

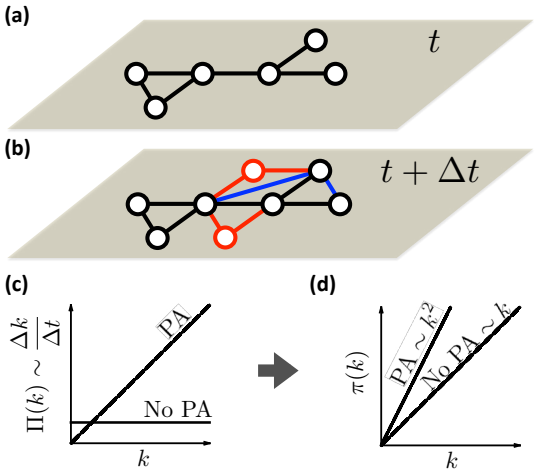


Figure 5.9
Detecting preferential attachment

If we have access to two maps of the same network, taken at time t and $t+\Delta t$, comparing them allows us to measure the $\pi(k)$ function that governs preferential attachment. Specifically we look at nodes that have gained new links thanks to the arrival of new nodes, like the two new red nodes at $t+\Delta t$. The blue lines correspond to links that connect previously disconnected nodes, called internal links. Their role is discussed in **CHAPTER 6**.

(c) In the presence of preferential attachment $\Delta k/\Delta t$ will depend linearly on a node's degree at time t.

(d) The scaling of the cumulative preferential attachment function helps us detect the presence or absence of preferential attachment.

In practice the obtained $\Delta k_i/\Delta t$ curve is typically noisy, particularly for small networks. To reduce this noise we often measure the cumulative function

$$\pi(k) \equiv \sum_{k_i=0}^{k} \Pi(k_i). \tag{5.21}$$

In the absence of preferential attachment we expect $\pi(k_i)$=constant, hence, $\pi(k) \sim k$ according to Eq. 5.21. If preferential attachment is present, i.e. $\pi(k_i)=k_i$, we expect $\pi(k) \sim k^2$. Fig. 5.10 shows the measured $\pi(k)$ for four real networks. For each system we observe a faster than linear increase in $\pi(k)$, indicating the presence of preferential attachment. Fig. 5.10 also suggests that $\pi(k)$ can be approximated with

$$\Pi(k) \sim k^{\alpha}. \tag{5.22}$$

For the Internet and citation networks we have $\alpha \simeq 1$, indicating that $\pi(k)$ depends linearly on k, as assumed in Eq. 5.2. This is in line with Hypotheses 1 and 2. For the co-authorship and the actor network the best fit provides α=0.9±0.1 indicating the potential presence of a sublinear preferential attachment.

In summary, Eq. 5.20 helps us detect the presence (or absence) of preferential attachment in real networks. The measurements show that the attachment probability depends on the node degree, in line with Hypothesis 1. Yet, we also find that while in some systems preferential attachment is linear, in others it can be sublinear, hence, Hypothesis 2 is occasionally violated. The implications of this non-linearity is discussed in the next section.

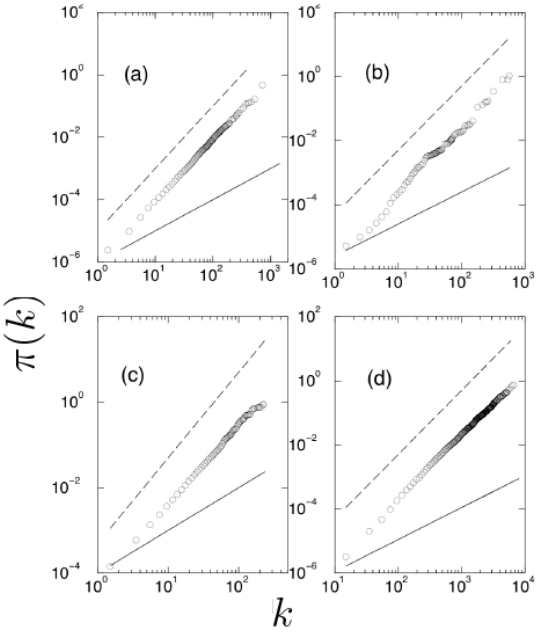


Figure 5.10
Evidence of preferential attachment

The figure shows the cumulative preferential attachment function $\pi(k)$, defined in Eq. 5.21, for several real systems:

(a) A citation network
(b) The Internet
(c) Neuroscience scientific collaboration network
(d) Actor network

In each panel we have two lines to guide the eye: the dashed line corresponds to linear preferential attachment ($\pi(k)\sim k^2$) and the continuous line indicates the absence of preferential attachment ($\pi(k)\sim k$). In line with Hypothesis 1 we detect a k-dependence in each dataset. Yet, in (c) and (d) $\pi(k)$ grows slower than k^2, indicating that for these two systems preferential attachment is sublinear, violating Hypothesis 2. Note that these measurements only consider links added through the arrival of new nodes, ignoring the addition of internal links. After [14].

NON-LINEAR PREFERENTIAL ATTACHMENT

The observation of sublinear preferential attachment in Fig. 5.9 raises an important question: what is the impact of this nonlinearity on the network topology? To answer this we replace the linear preferential attachment Eq. 5.1 with Eq. 5.21 and re-calculate the degree distribution of the Barabási-Albert model. The behavior for α=0 is clear: in the absence of preferential attachment, the model reduces to Model A discussed in SECTION 5.4. Consequently the degree distribution will follow the simple exponential function Eq. 5.17. For α = 1 we recover the Barabási-Albert model, obtaining a scale-free network with degree distribution Eq. 5.14. We next focus on the case when α ≠ 0 and α ≠ 1. The calculation, providing p_k for an arbitrary α, is presented in ADVANCED TOPICS 5.B, predicting several scaling regimes [13]:

SUBLINEAR PREFERENTIAL ATTACHMENT (0 < α < 1)
For any α > 0 new nodes favor the more connected nodes over the less connected nodes. Yet, for α < 1 the bias is not sufficient to generate a scale-free degree distribution. Instead, in this regime the degrees follow the stretched exponential distribution SECT. 4.10

$$p_k \sim k^{-\alpha} \exp\left(\frac{2\mu(\alpha)}{\langle k\rangle(1-\alpha)} k^{1-\alpha}\right) \qquad (5.23)$$

where $\mu(\alpha)$ is a function that depends only weakly on α. For α → 1 Eq. 5.22 reduces to the degree distribution of the BA model. Indeed for α=1 we have μ=2, and $\lim_{\alpha\to 1} \frac{k^{1-\alpha}}{1-\alpha} = \ln k$. Therefore $p_k \sim k^{-1}\exp(-2\ln k)=k^{-3}$. The exponential cutoff in Eq. 5.22 implies that sublinear preferential attachment limits the size and the number of the hubs.

Sublinear preferential attachment also affects the size of the largest degree, k_{max}. In CHAPTER 4 we showed that for a scale-free network the degree of the largest node scales polynomially with time Eq. 4.14. For sublinear preferential attachment we have

$$k_{\max} \sim (\ln t)^{1/(1-\alpha)} \qquad (5.24)$$

a logarithmic dependence that predicts a much slower growth of the maximum degee than the polynomial. This slower growth is the reason

why the hubs are smaller for α < 1 Fig. 5.10.

SUPERLINEAR PREFERENTIAL ATTACHMENT (α > 1)

For α > 1 the tendency to link to highly connected nodes is enhanced, accelerating the *rich-gets-richer process*. The consequence of this is most obvious for α > 2, when the model predicts a *winner-takes-all* phenomenon: almost all nodes connect to a single or a few super-hubs. Hence, we observe the emergence of a hub-and-spoke network, in which most nodes link directly to a few central nodes. The situation for 1 < α < 2 is less extreme, but similar. This winner-takes-all process impacts the time dependence of the largest hub as well, finding that Fig. 5.12.

Hence for α > 1 the largest hub links to a finite fraction of nodes in the system.

$$k_{\max} \sim t \tag{5.25}$$

In summary, nonlinear preferential attachment introduces deviations from the power law degree distribution, either limiting the size of the hubs (α < 1), or leading to super-hubs (α > 1, Fig. 5.12). Hence, $\pi(k)$ needs to depend strictly linearly on the degrees for the resulting network to have a pure power law p_k. While in many systems we do observe such a linear behavior, in others, like the scientific collaboration network and the actor network, preferential attachment is sublinear, limiting the size of the hubs. This sublinear form of $\pi(k)$ could be responsible for the systematic deviations from a pure power-law degree distribution observed in the previous chapter. Hence for these systems a stretched exponential Eq. 5.22 should offer a better fit to the degree distribution.

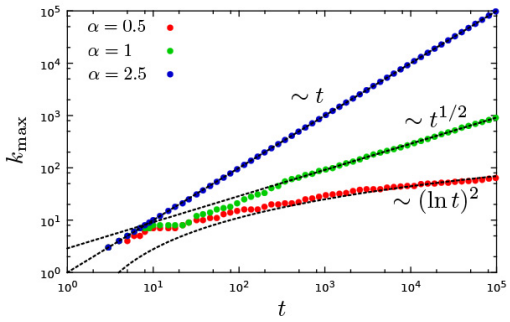


Figure 5.11
The growth of the hubs

The nature of preferential attachment affects the degree of the largest node. While in a scale-free network (α=1) the biggest hub grows as $t^{1/2}$ (green curve) Eq. 4.14, for sublinear preferential (α<1) attachment this dependence becomes logarithmic (red curve, Eq. 5.24. For superlinear preferential attachment the biggest hub grows linearly with time, always grabbing a finite fraction of all links (blue curve), Eq. 5.25. The symbols are provided by a numerical simulation; the dotted lines represent the analytical predictions.

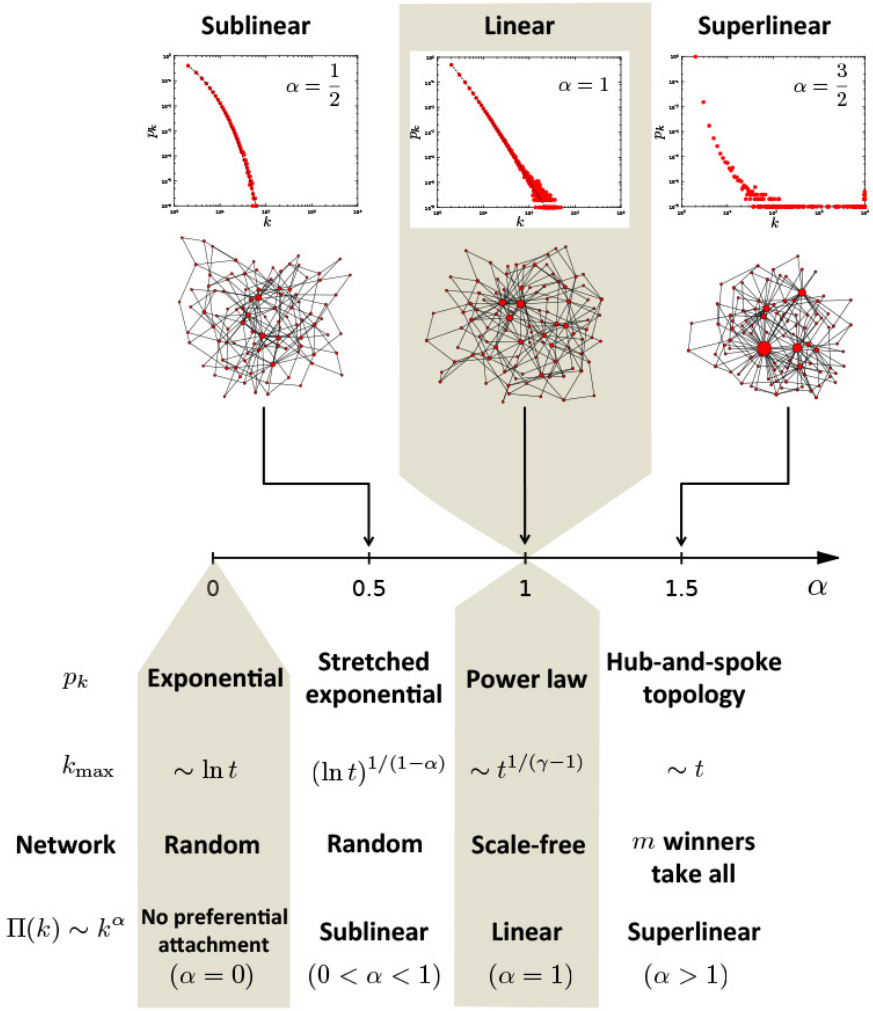


Figure 5.12

Nonlinear preferential attachment

The scaling regimes characterizing networks drivenby non-linear preferential attachment. The three toppanels show p_k for different α values (N=10⁴). The network maps under them show the obtained topologies for N=100 nodes. The theoretical results indicate the existence of four scaling regimes:

No preferential attachment (α=0)
The network has a simple exponential degree distribution following **Eq. 5.18**. Hubs are absent and the resulting network behaves like a random network.

Sublinear preferential attachment (0<α<1)
The degree distribution follows the stretched exponential **Eq. 5.23**, resulting in fewer and smaller hubs than in a scale-free network. As $\alpha \to 1$ the cutoff length increases, hence p_k follows a power law over an increasing range of degrees.

Linear preferential attachment (α=1)
This corresponds to the Barabási-Albert model, hence the degree distribution follows a power law.

Superlinear preferential attachment (α>1)
The high-degree nodes are disproportionately attractive and the network follows a winner-takes-all dynamics, developing a hub-and-spoke topology. In this configuration the earliest nodes become super hubs and all subsequent nodes link to them. Hence p_k, shown for α=1.5 (top panel), indicates the coexistence of many small nodes with a few *super hubs* in the vicinity of k=10⁴.

SECTION 5.8

THE ORIGINS OF PREFERENTIAL ATTACHMENT

Given the key the role preferential attachment plays in the evolution of real networks, we must ask, Where does preferential attachment come from? The question can be broken down to two narrower issues:

Why does $\pi(k)$ depend on k?
Why is the dependence of $\pi(k)$ linear in k?

In the past decade we witnessed the emergence of two philosophically different approaches to these questions. In the first class belong models that view preferential attachment as a result of an interplay between random events and some structural property of a network. These mechanisms do not require global knowledge of the network and rely on random actions, hence we will call them *local* or *random* mechanisms. A second class of models assume that each new node or link is preceeded by a cost-benefit analysis, balancing various needs with the available resources. These models assume familiarity with the whole network and rely on optimization principles, prompting us to call them *global* or *optimized* mechanisms. The purpose of this section is to discuss these two approaches.

LOCAL MECHANISMS

The link selection model offers perhaps the simplest example of a local or random mechanism capable of generating preferential attachment [16]. It is defined as follows:

- **Growth**
  At each time step we add a new node to the network.

- **Link selection**
  We select a link at random and connect the new node to one of the two nodes at the two ends of the selected link. This procedure is inherently local and random, as one does not need to know anything about the overall network topology to connect the new node. To show that this simple mechanism generates linear preferential attachment, we write the probability q_k that the node at the end of a randomly chosen

link has degree k as

$$q_k = Ckp_k . \tag{5.26}$$

Eq. 5.26 captures two effects:
(i) The higher the degree of a node, the higher the chance that it will be located at the end of the chosen link.

(ii) The more degree-k nodes are in the network (i.e., the higher is p_k), the more likely that a degree k node will be at the end of the link.

In Eq. 5.26 the value of C can be calculated using the normalization condition $\Sigma q_k = 1$, obtaining $C=1/\langle k\rangle$. Hence the probability that we find a degree-k node at the end of a randomly chosen link is

$$q_k = \frac{kp_k}{\langle k\rangle}. \tag{5.27}$$

a quantity called excess degree. Eq. 5.27 also represents the probability that a new node connects to a node with degree k in the link selection model, hence it plays the role of the preferential attachment $\pi(k)$. Therefore Eq. 5.26 indicates that random link selection generates preferential attachment that is linear in then degree. While link selection is perhaps the simplest mechanism for preferential attachment, it is neither the first nor the most popular in the class of models relying on local mechanisms. That distinction goes to is the *copying model*, described in Fig. 5.13.

OPTIMIZATION

A longstanding assumption of economics is that humans make rational decisions, balancing cost against benefits. Innother words, each individual aims to maximize its personal advantage. This is the starting point of rational choice theory in economics [21] and it is a hypothesis central to modern political science, sociology and philosophy. As we discuss below, such rational decisions can lead to preferential attachment as well [22, 23, 24].

Consider the Internet, whose nodes are routers or autonomous systems (AS), connected to each other via cables. Establishing a new Internet connection between two routers requires laying down a cable between them. As this can be costly, each new link is preceded by careful cost-benefit analysis. Each new router must agree with the nodes it links to that they will transmit the data packets leaving from or arriving to the new node (peering relationship). Therefore each new node will choose its link to balance access to good network performance (like proper bandwith) with the cost of laying down a new cable i.e. physical distance). This can be a conflicting desire, as the closest node does not always offer the best network performance. For simplicity let us assume that the nodes are all located on a unit square. At each time step we add a new node by randomly choosing a point within the square. When deciding where to connect the new node i to the existing nodes, we calculate the cost function [22]

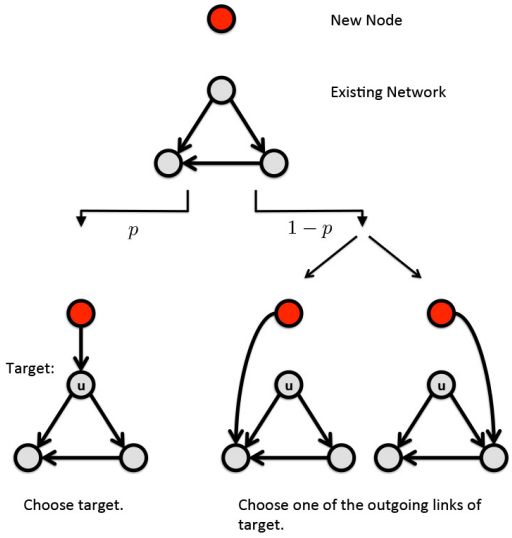


Figure 5.13
Copying model

When building a new webpage, authors tend to borrow links from webpages covering similar topics, a process captured by the copying model [17, 18]. In the model, in each time step a new node with a single link is added to the network. To choose the target node we randomly select a node u and follow a two-step procedure:

(a) Random Connection: With probability p the new node links to u.

(b) Copying: With probability 1-p we randomly choose an outgoing link of node u and link the new node to the link's target. Hence, the new node *copies* one of the links of an earlier node. For step (a) the probability of selecting a particular node is $1/N$. Step (b) is equivalent with selecting a node linked to a randomly selected link. The probability of selecting a degree-k node through the copying process of step (b) is $k/2L$ for undirected networks. That is, the likelihood that the new node will connect to a degree-k node follows preferential attachment $\pi(k) = p / N+(1- p)k/(2L)$, which is linear in k. The popularity of the copying model lies in its adaptability to real systems:

- **Social networks:** The more acquaintances an individual has, the higher is the chance that she will be introduced to new individuals by her existing acquaintances. Without friends, it is difficult to make new friends.

- **Citation Networks**: No scientist can be familiar with all papers published on a certain topic. If we assume that authors decide what to cite by randomly selecting references from the papers they have already read, then papers with more citations are more likely to be cited again.

- **Protein interaction networks**: Gene duplication, a common mechanism leading to next genes in a cell, can be mapped into the copying model, explaining the scale-free nature of protein interactions networks [19, 20].

$$C_i = \min_j [\delta d_{ij} + h_j] \tag{5.28}$$

for each node j already in the network, where d_{ij} is the Euclidean distance between i and j, and h_j is the network-based distance of node j to the first node in the network, designated as the "center " of the network Fig. 5.14. Hence h_j captures the "resources" offered by node j, in the form of its distance to the network's center. The calculations indicate the emergence of three distinct network topologies, depending on the value of the parameter δ in Eq. 5.28 and Fig. 5.15:

STAR NETWORK $\delta < (1/2)^{1/2}$

For δ = 0 the Euclidean distances are irrelevant, hence each node will simply link to the central node, turning the network into a star. This star configuration persist for any δ < $(1/2)^{1/2}$, guaranteeing that the h_{ij} term dominates over the δd_{ij} term in Eq. 5.28.

RANDOM NETWORK $\delta \geq N^{1/2}$

For very large δ the contribution provided by the distance term δd_{ij} overwhelms h_j in Eq. 5.27. In this case each new node will connect to the node closest to it. The resulting graph is a dynamic version of the Euclidean minimum spanning tree. The resulting network will have a bounded degree distribution, like a random network Fig. 5.15.

SCALE-FREE NETWORK $4 \leq \delta \leq N^{1/2}$

Numerical simulations and analytical calculations [22] indicate that for intermediate δ values the network develops a scale-free topology.

The origin of the power law is rooted in two competing mechanisms:

(i) Optimization: Each node has a basin of attraction, so that nodes landing in this basin will always link to it Fig. 5.14. The size of each basin correlate with h_j of node j at its center, which in turn correlates with the node's degree k_j Fig 5.14f.

(ii) Randomness: We choose randomly the position of the new node, ending in one of the N basins of attraction. The node with the largest degree largest basin of attraction, will gain the most new nodes and links. This leads to preferential attachment, documented in Fig. 5.15d.

In summary, the microscopic mechanisms responsible for preferential attachment can have two fundamentally different origins BOX 5.5: it can be rooted in random processes, like link selection or copying, or in optimization process, when new nodes balance conflicting criteria as they decide where to connect. These results help us understand why preferential attachment is present in so different systems as the cell or the Internet. The diversity of the mechanisms discussed in this section suggest that preferential attachment is so widespread precisely because it can come from both rational choice and random actions [25]. Most complex systems are driven by processes that have a bit of both. Hence luck or reason, preferential attachment wins either way.

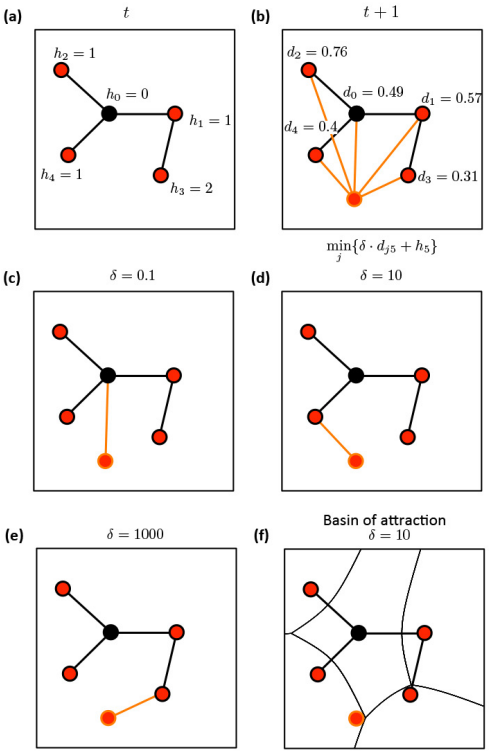


Figure 5.14
Optimization Model

(a) A small network configuration, where the h_j term in the cost function of **Eq. 5.28** is shown for each node. Here h_j represents the network-based distance of node j from node i=0. Hence h_0=0 and h_3=2.

(b) A new node, shown in orange, will choose the node j to connect to by minimizing c_j of **Eq. 5.28** . If δ =0 or δ is small the new node will connect to the central node with h_j =0.

(c)-(e) As we increase δ, the balance in **Eq. 5.28** changes, forcing the new node to connect to different nodes. The panels (c)-(e) show the choice of the new node (orange) for a different values of δ for the given network configuration.

(f) The basin of attraction for each node for δ=10. A new node arriving inside a particular basin will always link to the node at the center of the basin. The size of each basin depends on the degree of the node at its center. Indeed, the smaller is hj, the larger can be the distancento the new node while still minimizing the cost **Eq. 5.28**. Yet, the higher the degree of node j, the smaller is its expected distance to the central node h_j.

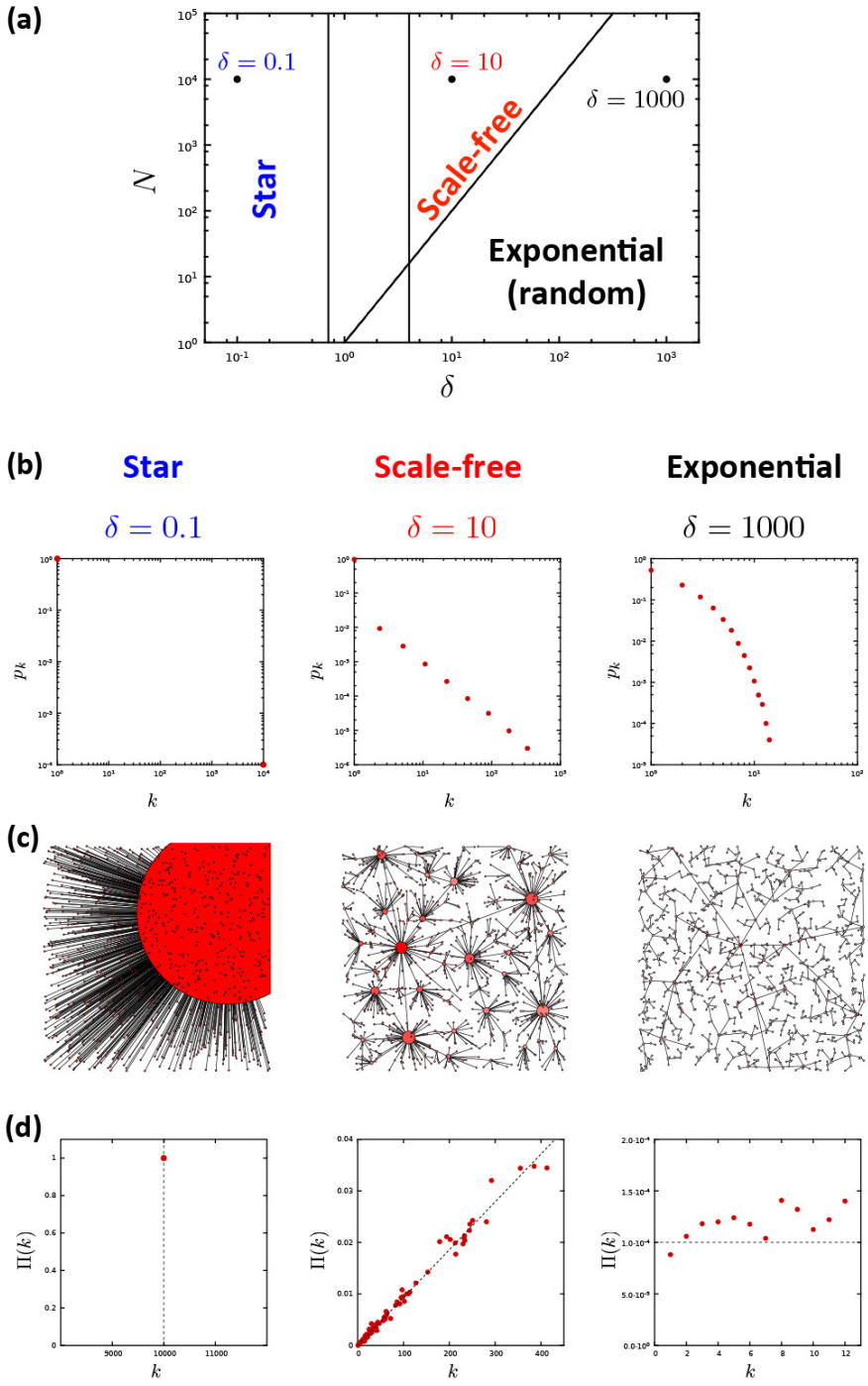


Figure 5.15
Scaling in the optimization model

(a) A schematic diagram, describing the three main classes of networks generated by the optimization model: star topology, scale-free topology and exponential networks. The structure of the network in the unmarked area is unknown. The boundary of the star configuration is given by $s=(1/2)^{1/2}$. Indeed, the maximum distance between two nodes on a square lattice with unit length, over which the model is defined, is the diagonal $2^{1/2}$. Therefore if $\delta < 1/2^{1/2}$, for any new node $\delta d_i<1$. In this case the cost of connecting to the central node is $c_i = \delta d_{ij}+0$, which is always lower than connecting to any other node at the cost of $f(i, j) = \delta d_{ij}+1$. Therefore for $\delta < (1/2)^{1/2}$ all nodes connect to node 0, resulting in a network dominated by a single hub (star network, see (c)). The oblique line making the boundary of the scale-free regime is $\delta = N^{1/2}$. Indeed, if nodes are placed randomly on the unit square, then the typical distance between neighbors decreases as $N^{-1/2}$. Hence, if $d_{ij}\sim N^{-1/2}$ then $\delta d_{ij}\geq h_{ij}$ for most node pairs. Typically the path length to the central node h_j grows slower than N (in small-world networks $h_j \sim \log N$, is scale-free networks networks $h_j\sim \ln\ln N$). Therefore C_i is dominated by the δd_{ij} term and the smallest C_i is achieved by minimizing the distance-dependent term. Note that strictly speaking the transition only occurs in the $N \to \infty$.

(b) Degree distribution for networks emerging in the three phases discussed above for $N=10^4$.

(c) Typical topologies generated by the optimization model for the selected δ values. The node size is chosen to be proportional to its degree.

(d) We used the method described in SECT. 5.7 to measure the preferential attachment function $\pi(k)$. Starting from a network with N=10,000 nodes we added a new node and measured the degree of the node that it connected to. We repeated this procedure 10,000 times, obtaining $\pi(k)$. The plots indicate the presence of a linear preferential attachment in the scale-free phase, but its absence in the star and the exponential phases.

BOX 5.5

LUCK OR REASON: AN ANCIENT DEBATE

The tension between randomness and optimization, two apparently antagonistic explanations for power laws, is by no means new: in the 1960s Herbert Simon and Benoit Mandelbrot have engaged in a fierce public dispute over this very topic. Simon proposed that randomness is responsible for the power-law nature of word frequencies. Mandelbrot, however, fiercely defended an optimization-based framework.

The debate spanned seven papers and several years and is one of the most vicious scientific disagreement on record. It is hard to know what set it off—it may have been Simon's brief note in his 1955 paper [26], dismissing Mandelbrot's explanation that power laws observed in linguistics are routed in an optimization process [27]. Mandelbrot responded with a comment [28] stating that 'Simon's model is analytically circular.' The essence of Simon's lengthy reply a year later is well summarized in its abstract: 'Dr. Mandelbrot's principal and mathematical objections to the model are shown to be unfounded' [29]. This prompted a 19 page response by Mandelbrot entitled 'Final Note [...]', stating that 'most of Simon's (1960) reply was irrelevant' [30] and, ensuring that this will not be the final note. Sure enough, Simon's subsequent reply states that 'this present "Reply" refutes the almost entirely new arguments introduced by Dr. Mandelbrot' [31].

That inspired a paper creatively entitled a "Post Scriptum to "Final Note,"" by Mandlebrot [32], stating that 'My criticism has not changed since I first had the privilege of commenting upon a draft of Simon,' Simon's final note ends but does not resolve the debate: "Dr. Mandelbrot has proposed a new set of objections to my 1955 models of Yule distributions. Like earlier objections, these are invalid." [33].

In the context of networks the argument titled in Simon's favor the power laws observed in complex networks appear to be driven by randomness and preferential attachment. Yet, as we seek to explain the origins of preferential attachment, the optimization-based ideas proposed by Mandelbrot play an important role.

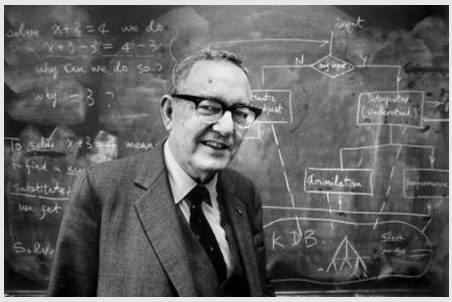


Figure 5.15a
Herbert Simon

Figure 5.15b
Benoit Mandelbrot

DIAMETER AND CLUSTERING COEFFICIENT

To complete the characterization of the Barabási-Albert model we need to discuss the behavior of two additional measures: the network diameter and the clustering coefficient. Both quantities play an important role in comparing the model predictions to the properties of real systems.

DIAMETER

The network diameter, representing the maximum distance in the Barabási-Albert network, is predicted to follow

$$D \sim \frac{\log N}{\log \log N} . \tag{5.29}$$

a result obtained independently by Cohen and Havlin [34] and Bollobás and Riordan [35], the latter also offering an exact proof. Eq. 5.29 tells us that the network diameter grows slower than log N, hence the distances in the Barabási-Albert model are smaller than the distances observed in a random graph of similar size. The difference is particularly relevant for large N. Note that while Eq. 5.29 is derived for the diameter, we expect that the average distance $\langle d \rangle$ scales in a similar fashion. The impact of the log log N the log N term captures the scaling of $\langle d \rangle$ with N, but for large $N(\geq 10^4)$ the impact of the logarithmic correction becomes noticeable.

CLUSTERING COEFFICIENT

The clustering coefficient of the Barabási-Albert model follows ADVANCED TOPICS 5.C.

$$C = \frac{m-1}{8}\frac{(\ln N)^2}{N} , \tag{5.30}$$

a result obtained by Klemm and Eguiluz [36], and proved by Bollobás [37]. The prediction Eq. 5.30 is quite different from the $1/N$ dependence obtained for the random network model Fig. 3.20. The difference comes in the $(\ln N)^2$ term, that increases the clustering coefficient for large N. Consequently the clustering coefficient of the Barabási-Albert model decays slower than expected for a random network, indicating that the obtained network is locally more clustered.

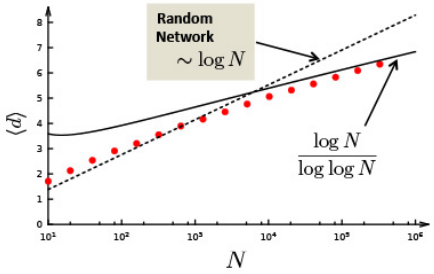


Figure 5.16
Average distance

The dependence of the average distance on the system size in the Barabási-Albert model. The continuous line corresponds to the exact result Eq. 5.29, while the dotted line corresponds to the prediction obtained in CH. 3 for a random network. Note that the analytical predictions do not provide the exact perfactors, hence the lines are not fits, but indicate only the predicted N dependent trends. The results were averaged for ten independent runs for $m = 2$.

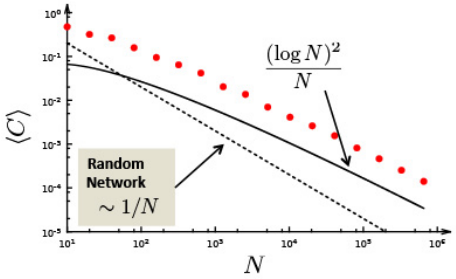


Figure 5.17
Clustering coefficient

The dependence of the average clustering coefficient on the system size N for the Barabási-Albert model. The continuous line corresponds to the analytical prediction Eq. 5.30, while the dotted line corresponds to the prediction for a random network, for which $\langle C \rangle \sim 1/N$. The results were averaged for ten independent runs for $m = 2$. The dashed and continuous curves are only drawn to indicate the N dependent trends. Hence, they do not represent a precise fit.

SUMMARY

The most important message of the Barabási-Albert model is that network structure and evolution are inseparable. Indeed, in the Erdős-Rényi, configuration or the hidden parameter models the role of the modeler is to place the links between a fixed number of nodes. Returning to our earlier analogy, the networks generated by these models relate to real networks like a photo of a painting relates to the painting itself: it may look like the real one, but the process of generating a photo is drastically different from the process of painting the original painting. The aim of the Barabási-Albert model is to capture the processes that assemble a network in the first place. Hence, it aims to paint the painting again, coming as close as feasible to the original brush strokes. consequently, the modeling philosophy of the Barabási-Albert model is simple: *to understand the topology of a complex system, we first need to describe how it came into being.*

Dynamics and network assembly take the driving role and the structural characteristics of the network, like the degree distribution, is a by-product of this modeling philosophy. Random graphs, the configuration and the hidden parameter models will continue to play an important role as we try to understand how certain network properties deviate from our expectations. Yet, if we want to explain the origin of a particular network property, we will have to use models that capture the system's genesis. In its current form the Barabási-Albert model cannot describe the wide range of network characteristics observed in real systems. This is illustrated by the model's notable limitations:

- If predicts $\gamma=3$ while the degree exponent of real networks varies between 2 and 5 Table 4.2.

- Many networks, like the WWW or citation networks, are directed, while the model generates undirected networks.

- Many processes known to occur in networks, from linking already existing nodes to the disappearance of links and nodes, are absent from the model.

- The model does not allow to distinguish between nodes based on some intrinsic characteristics, like the novelty of a research paper or the utility of a webpage. While the Barabási-Albert model is occasionally used as a model of the Internet or the cell, in reality is not designed to capture the details of any particular real network systems. It is a minimal, proof of principle model whose main purpose is to capture the basic mechanisms responsible for the emergence of the scale-free property.

Therefore, if we want to understand the evolution of systems like the Internet, the cell or the WWW, we need to incorporate the important details that contribute to the time evolution of these systems, like the directed nature of the WWW, the possibility of internal links and node and link removal. As we show in CHAPTER 6, these limitations can be systematically resolved. Finally, the results discussed in this chapter allow us to formulate the next law:

The Third Law of Networks: Growth and Preferential Attachment.

Hubs and power laws are a joint consequence of growth and preferential attachment.

Let us revisit the three criteria we used earlier to establish the validity of a network law:

(a) Quantitative formulation of the third law is provided by the Barabási-Albert model, together with its documented ability to generate scale-free networks based on growth and preferential attachment.

(b) Universality: SECTION 5.7 offers direct empirical evidence that real networks that exhibit the scale-free property are characterized by preferential attachment; SECTION 5.2 offers evidence of growth.

(c) Non-random origin: Preferential attachment is obviously absent from random networks, which is the main reason why random networks do not develop hubs and power laws.

BOX 5.6

AT A GLANCE: BARABÁSI-ALBERT MODEL

Number of nodes

$N = t$

Number of links

$N = mt$

Average Degree

$\langle k \rangle = 2m$

Degree dynamics

$k_i(t) = m\,(t/t_i)^{\beta}$

Dynamical exponent

$\beta = 1/2$

Degree distribution

$p_k \sim k^{-\gamma}$

Degree exponent

$\gamma = 3$

Average distance

$\langle d \rangle \sim \log N / \log\log N$

Clustering coefficient

$\langle C \rangle \sim (\ln N)^2 / N$

HOMEWORK

1. Calculate the degree distribution of the directed Barabási-Albert model. That is, in each time a new node arrives, that connects with a directed link to a node chosen with preferential attachment Eq. 5.1, where $\pi(k_{in})$ depends only the node's in-degree. Discuss both the in and out-degree distribution of the resulting network.

2. Use the rate equation approach described above that the directed copying model leads to a scale-free network with the incoming degree exponent $\gamma_{in} = (2 - p) / (1 - p)$, hence the degree exponent varies between $\gamma_{in} = 2$ for $p \to 0$ and $\gamma_{in} = \infty$ for $p \to 1$.

ADVANCED TOPICS 5.A
DERIVING THE DEGREE DISTRIBUTION

A number of analytical techniques are available to calculate the exact form of the degree exponent provided in Eq. 5.11. Next we derive it using the rate equation approach [12, 13]. The method is rather general, allowing us to explore the properties of a wide range of growing networks. Consequently, the calculations described here will be of direct relevance for many systems, from the WWW to protein interaction networks. Let us denote with $N(k, t)$ the number of nodes with degree k at time t. The degree distribution $p_k(t)$ relates to this quantity via $p_k(t) = N(k,t)/N$.

Since at each time-step we add a new node to the network, we have $N = t$. That is, at any moment the total number of nodes equals the number of timesteps BOX 5.3. We write preferential attachment as

$$\Pi(k) = \frac{k}{\sum_j k_j} = \frac{k}{2mt} \cdot \tag{5.31}$$

where the $2m$ term captures the fact that that in an undirected network each link contributes to the degree of two nodes. Our goal is to calculate the changes in the number of nodes with degree k after a new node is added to the network. For this we inspect the two events that alter $N(k, t)$ (and hence $p_k(t)$) following the arrival of a new node:

(i) A new node can link to a degree-k node, turning it into a degree (k+1 node, hence decreasing $N(k, t)$.

(ii) A new node can link to a degree (k-1) node, turning it into a degree k node, hence increasing $N(k, t)$.

The number of links that are expected to connect to degree k nodes after the arrival of a new node is

$$\frac{k}{2mt} \times Np_k(t) \times m = \frac{k}{2} p_k(t), \tag{5.32}$$

where the first term captures the probability that the new node will link to a degree-k node (preferential attachment); the second term provides the

total number of nodes with degree k, as the more nodes are in this category, the more likely that a new node will attach to one of them; the third term is simply the degree of the incoming node, as the higher m, the higher the chance that the new node will link to a degree-k node. We next apply Eq. 5.32 to cases (i) and (ii) above:

(i') The number of degree k nodes that acquire a new link becoming (k+1) degree nodes, is

$$\frac{k}{2}p_k(t) \tag{5.33}$$

(ii') The number of degree (k-1) nodes that acquire a new link, increasing their degree to k is

$$\frac{k-1}{2}p_{k-1}(t). \tag{5.34}$$

Combining Eq. 5.32 and Eq. 5.33 we obtain the expected number of degree-k nodes after the addition of a new node

$$(N+1)p_k(t+1) = Np_k(t) + \frac{k-1}{2}p_{k-1}(t) - \frac{k}{2}p_k(t). \tag{5.35}$$

This equation applies to all nodes with degree $k>m$. As we lack nodes with degree k=0,1, ... , m-1 in the network (each new node arrives with degree m) we need a separate equation for degree m modes. Following the arguments we used to derive Eq. 5.35, we obtain

$$(N+1)p_m(t+1) = Np_m(t) + 1 - \frac{m}{2}p_m(t)\ . \tag{5.36}$$

Eq. 5.35 and 5.36 are the starting point of the recursive process that provides p_k. Let us use the fact that we are looking for a stationary degree distribution, supported by numerical simulations Fig. 5.6. This means that in the $N = t \to \infty$ limit, $p_k(\infty) = p_k$. Using this we can write the l.h.s. of Eq. 5.35 and 5.36 as $(N+1)p_k(t+1) - Np_k(t) \to Np_k(\infty) + p_k(\infty) - Np_k(\infty) = p_k(\infty) = p_k(N+1)\ p_{mk}(t+1) - Np_m(t) \to p_m$. Therefore the rate equations Eq. 5.35 and 5.36 take the form:

$$p_k = \frac{k-1}{k+2}p_{k-1} \qquad k > m \tag{5.37}$$

$$p_m = \frac{2}{m+2} \tag{5.38}$$

Note that Fig. 5.37 can be rewritten as

$$p_{k+1} = \frac{k}{k+3}p_k \tag{5.39}$$

via a $k \to k+1$ variable change. To obtain the degree distribution, we use a recursive approach. That is, we write the degree distribution for the smallest degree, $k=m$, using Eq.. 5.38 and then use Eq. 5.39 to calculate p_k for the higher degrees:

$$\begin{aligned} p_{m+1} &= \frac{m}{m+3}p_m = \frac{2m}{(m+2)(m+3)} \\ p_{m+2} &= \frac{m+1}{m+4}p_{m+1} = \frac{2m(m+1)}{(m+2)(m+3)(m+4)} \\ p_{m+3} &= \frac{m+2}{m+5}p_{m+2} = \frac{2m(m+1)}{(m+3)(m+4)(m+5)} \end{aligned} \tag{5.40}$$

At this point we notice a simple recursive pattern: by replacing m+3 with k we obtain the probability to observe a node with degree k

$$p_k = \frac{2m(m+1)}{k(k+1)(k+2)}, \tag{5.41}$$

which represents the exact form of the degree distribution for the Barabási-Albert model. Note that:

- For large k this becomes $p_k \sim k^{-3}$, in agreement with the numerical result.

- The prefactor of Eq. 5.11 or Eq. 5.41 is different from the prefactor derived in Eq. 5.9.

This form was derived independently in [12] and [13], and the mathematical proof of its validity is provided in [10]. Note that the rate equation formalism offers an elegant continuum equation satisfied by the degree distribution of the Barabási-Albert model [16]. Starting from the equation

$$p_k = \frac{k-1}{2} p_{k-1} - \frac{k}{2} p_\infty \tag{5.42}$$

we can write

$$2p_k = (k-1)p_{k-1} - kp(k) = -p_{k-1} - k[p_k - p_{k-1}], \tag{5.43}$$

$$2p_k = -p_{k-1} - k\frac{p_k - p_{k-1}}{k-(k-1)} = -p_{k-1} - k\frac{\partial p_k}{\partial k} \tag{5.44}$$

obtaining

$$p_k = \frac{1}{2}\frac{\partial [kp_k]}{\partial k} \tag{5.45}$$

One can check that the solution of Eq. 5.45 is

$$p_k \sim k^{-3}. \tag{5.46}$$

ADVANCED TOPICS 5.B NONLINEAR PREFERENTIAL ATTACHMENT

The purpose of this section is to derive the degree distribution of an evolving networks governed by a nonlinear preferential attachment. We follow Krapivsky et al. [13]. As the results of Ref. [13] were derived for undirected networks, here we adjusted the calculation to cover undirected networks.

Strictly speaking the degree distribution only exists for $\alpha \leq 1$. For $\alpha > 1$ a few nodes attract a finite fraction of links, as explained in SECT. 5.7, and we do not have a stationary p_k. Therefore, we limit ourself to the $\alpha \leq 1$ case. We start with the Barabási-Albert model, in which at each time step a new node is added with m new links. We connect each new link to an existing node with probability

$$\Pi(k_i) = \frac{k_i^{\alpha}}{\mu(\alpha)} . \tag{5.47}$$

where k_i is the degree of node i, $0 < \alpha \leq 1$ and

$$\mu(\alpha,t) = \sum_k k^{\alpha} p_k(t) . \tag{5.48}$$

is the normalization factor. Note that $\mu(0,t) = \sum p_k(t) = 1$ and $\mu(1,t) = \sum_k k p_k(t) = \langle k \rangle = 2mt / N$ is the average degree. Since $0 < \alpha \leq 1$,

$$\mu(0,t) \leq \mu(\alpha,t) \leq \mu(1,t) . \tag{5.49}$$

Therefore in the long time limit

$$\mu(\alpha, t \to \infty) = \text{constant} . \tag{5.50}$$

whose precise value will be calculated later. For simplicity, we adopt the notation $\mu \equiv \mu(\alpha, t \to \infty)$.

Following the rate equation approach introduced in Advanced TOPICS 5.A, we write the rate equation for the network's degree distribution as

$$p_k(t+1) = \frac{m}{\mu(\alpha,t)}[(k-1)^\alpha p_{k-1}(t) - k^\alpha p_k(t)] + \delta_{k,m}. \tag{5.51}$$

The first term on the r.h.s. describes the rate at which nodes with degree (k-1) gain new links; the second term describes the loss of degree-k nodes when they gain new links, turning into (k+1) degree nodes; the last term represents the newly added nodes with degree m. Asymptotically, in the $t\to\infty$ limit we can write $p_k=p_k(t+1)=p_k(t)$. Substituting $k=m$ in Eq. 5.51 we obtain:

$$p_m = -\frac{m}{\mu} - m^\alpha p_m + 1 \tag{5.52}$$

$$p_m = -\frac{\mu/m}{\mu/m + m^\alpha}.$$

For $k > m$

$$p_k = \frac{m}{\mu}[(k-1)^\alpha p_{k-1} - k^\alpha p_k] \tag{5.53}$$

$$p_k = \frac{(k-1)^\alpha}{\mu/m + k^\alpha} p_{k-1} \tag{5.54}$$

Solving Eq. 5.53 recursively we obtain

$$p_m = \frac{\mu/m}{\mu/m + m^\alpha} \tag{5.55}$$

$$p_{m+1} = \frac{\mu/m \cdot m^\alpha}{\mu/m + (m+1)} \frac{\mu/m}{\mu/m + m^\alpha} \tag{5.56}$$

$$p_k = \frac{\mu/m}{k^\alpha} \prod_k^{j=m} \left(1 + \frac{\mu/m}{j^\alpha}\right)^{-1} \tag{5.57}$$

To determine the large k behavior of p_k we take the logarithm of (52):

$$\ln p_k = \ln(\mu/m) - \alpha \ln k - \sum_k^{j=m} \left(1 + \frac{\mu/m}{j^\alpha}\right) \tag{5.58}$$

Using the series expansion $\ln(1+x) = \sum_{n=1}^{\infty} (-1)_i^{n+1} / n \cdot x^n$ we obtain

$$\ln p_k = \ln(\mu/m) - \alpha \ln k - \sum_{i=m}^{k} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} (\mu/m)^n j^{-\alpha \cdot n} \tag{5.59}$$

We approximate the sum over j with the integral

$$\sum_k^{j=m} j_X^{-n\alpha} \approx \int_k^m x^{-n\alpha} dx = \frac{1}{1-n\alpha}(k^{1-n\alpha} - m^{1-n\alpha}) \tag{5.60}$$

which in the special case of $n\alpha$ =1 becomes

$$\sum_{j=m}^{k} j^{-1} \approx \int_m^k x^{-1} dx = \ln k - \ln m. \tag{5.61}$$

Hence we obtain

$$\ln p_k = \ln(\mu/m) - \alpha \ln k - \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \frac{(\mu/m)^n}{1-n\alpha} (k^{1-n\alpha} - m^{1-n\alpha}) \tag{5.62}$$

Consequently the degree distribution has the form

$$p_k = C_\alpha k^{-\alpha} e^{-\sum_{\infty}^{n=1} \frac{(-1)^{n+1}}{n} \frac{(\mu/m)^n}{1-n\alpha} k^{1-n\alpha}}, \tag{5.63}$$

where

$$C_\alpha = \frac{\mu}{m} e^{\sum_{\infty}^{n=1} \frac{(-1)^{n+1}}{n} \frac{(\mu/m)^n}{1-n\alpha} m^{1-n\alpha}} \tag{5.64}$$

The vanishing terms in the exponential do not influence the $k \to \infty$ asymptotic behavior, being relevant only if $1-n\alpha \geq 1$. Consequently the precise form of p_k depends on α as:

$$p_k \sim \begin{cases} k^{-\alpha} e^{\frac{-\mu/m}{1-\alpha} k^{1-\alpha}}, & 1/2 < \alpha < 1 \\ k^{-\frac{1}{2}+\frac{1}{2}(\frac{\mu}{m})^2} e^{-\frac{1}{2}\frac{\mu}{m} k^{-2}}, & \alpha = 1/2 \\ k^{-\alpha} e^{-\frac{\mu/m}{1-\alpha} k^{1-\alpha} + \frac{1}{2}\frac{(\mu/m)^2}{1-2\alpha} k^{1-2\alpha}}, & 1/3 < \alpha < 1/2 \\ \vdots \end{cases} \tag{5.65}$$

That is, for $1/2 < \alpha < 1$ the degree distribution follows a stretched exponential. As we lower α, new corrections start contributing each time α becomes smaller than $1/n$, where n is an integer. For $\alpha \to 1$ the degree distribution scales as k^{-3}, as expected for the Barabási-Albert model. Indeed for $\alpha = 1$ we have $\mu=2$, and

$$\lim_{a \to 1} \frac{k^{1-a}}{1-a} = \ln k. \tag{5.66}$$

Therefore $p_k \sim k^{-1}\exp(-2\ln k) = k^{-3}$.

Finally we need to calculate $\mu(\alpha) = \sum^{j} j^\alpha p_j$. For this we sum Eq. 5.58:

$$\sum_{k=m}^{\infty} k^\alpha p_k = \sum_{k=m}^{\infty} \frac{\mu(\alpha)}{m} \prod_{k}^{j=m} \left(1 + \frac{\mu(\alpha)/m}{j^\alpha}\right)^{-1} \tag{5.67}$$

$$1 = \frac{1}{m} \sum_{k=m}^{\infty} \prod_{i=m}^{k} \left(1 + \frac{\mu(\alpha)/m}{j^\alpha}\right)^{-1} \tag{5.68}$$

We obtain $\mu(\alpha)$ by solving Eq. 5.52 numerically.

ADVANCED TOPICS 5.C
THE CLUSTERING COEFFICIENT

The purpose of this section is to derive the average clustering coefficient, Eq. 5.30, for the Barabási-Albert model. The derivation follows the an argument proposed by Klemm and Eguiluz [36], that was supported by the exact calculation of Bollobás [37]. We aim to calculate the number of triangles expected in the model, as the number of triangles can be linked to the clustering coefficient SECT. 2.10. We denote the probability to have a link between node *i* and *j* with *P(i, j)*. Therefore, the probability that three nodes *i, j, l* form a triangle is *P(i, j) P(i, l) P(j, l)*. The expected number of triangles in which node *l* with degree k_l participates is thus given by the sum of the probabilities that node *l* participates in triangles with an arbitrary chosen node *i* and *j* in the network. This can be written as

$$Nr_l(\triangleleft) = \int_{i=1}^{N} di \int_{j=1}^{N} dj P(i,j)P(i,l)P(j,l) \tag{5.69}$$

To proceed we need to calculate *P(i,j)*, which requires us to consider how the Barabási-Albert model evolves. Let us denote the time when node *j* arrived with $t_j = j$, which we can do as in each time step we added only one new node. Hence the probability that at its arrival node *j* links to node *i* with degree k_i is given by preferential attachment:

$$P(i,j) = m\Pi(k_i(j)) = m\frac{k_i(j)}{\sum_{l=1}^{j} k_l} = m\frac{k_i(j)}{2mj}. \tag{5.70}$$

Using Eq. 5.7, we can write

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\frac{1}{2}} = m\left(\frac{j}{i}\right)^{\frac{1}{2}}, \tag{5.71}$$

where we used the fact that the arrival time of node *j* is $t_j = j$ and the arrival time of node is $t_i = i$. Hence Eq. 5.70 now becomes

$$P(i,j) = \frac{m}{2}(ij)^{-\frac{1}{2}}. \tag{5.72}$$

Using this result we can calculate the number of triangles in Eq. 5.62, writing

$$Nr_l(\triangleleft) = \int_{i=1}^{N} di \int_{j=1}^{N} dj P(i,j)P(i,l)P(j,l) \tag{5.73}$$

$$= \frac{m^3}{8} \int_{i=1}^{N} di \int_{j=1}^{N} dj (ij)^{-\frac{1}{2}} (il)^{-\frac{1}{2}} (jl)^{-\frac{1}{2}}$$

$$= \frac{m^3}{8l} \int_{i=1}^{N} \frac{di}{i} \int_{j=1}^{N} \frac{dj}{j} = \frac{m^3}{8l} (\ln N)^2 \tag{5.74}$$

The clustering coefficient can be written as $C = \frac{2Nr_l(\triangleleft)}{k_l(k_l-1)}$, hence we obtain

$$C = \frac{\frac{m^3}{4l}(\ln N)^2}{k_l(k_l-1)} \tag{5.75}$$

To simplify Eq. 5.74, we note that according to Eq. 5.7 we have

$$k_l(t) = m\left(\frac{N}{l}\right)^{\frac{1}{2}} \tag{5.76}$$

which is the degree of node l at time $t = N$. Hence, for large k_l we have

$$k_l(k_l-1) \approx k_l^2 = m^2 \frac{N}{l} \tag{5.77}$$

allowing us to write the clustering coefficient of the Barabási-Albert model as

$$C = \frac{2Nr_l(\triangleleft)}{k_l(k_l-1)} \tag{5.78}$$

Eq. 5.78, apart from a factor 2, is the result Eq. 5.30.

BIBLIOGRAPHY

[1] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. Science, 286:509-512, 1999.

[2] F. Eggenberger and G. Pólya. Über die Statistik Verketteter Vorgänge. ZAMM - Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik, 3:279-289, 1923.

[3] G. Udny Yule. A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, f.r.s. Philosophical Transactions of the Royal Society of London. Series B, 213:21-87, 1925.

[4] Gibrat R. “Les Inégalités économiques”, Paris, France, 1931.

[5] G. K. Zipf. Human behavior and the principle of least reort. Addison-Wesley Press, Oxford, England, 1949.

[6] H. A. Simon. On a class of skew distribution functions. Biometrika, 42:425-440, 1955.

[7] D. De Solla Price. A general theory of bibliometric and othercumulative advantage processes. Journal of the American Society for Information Science, 27:292-306, 1976.

[8] R. K. Merton. The Matthew effect in science. Science, 159(3810):56-63, 1968.

[9] A.-L. Barabási. Linked: The new science of networks. Perseus, New York, 2002.

[10] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. The degree sequence of a scale-free random graph process. Random Structures and Algorithms, 18:279-290, 2001.

[11] A.-L. Barabási, H. Jeong, R. Albert. Mean-field theory for scale free

random networks. Physica A, 272:173-187, 1999.

[12] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. Structure of growing networks with preferential linking. Phys. Rev. Lett., 85:4633-4636, 2000.

[13] P. L. Krapivsky, S. Redner, and F. Leyvraz. Connectivity of growing random networks. Phys. Rev. Lett., 85:4629-4632, 2000.

[14] H. Jeong, Z. Néda. A.-L. Barabási. Measuring preferential attachment in evolving networks. Europhysics Letters, 61:567-572, 2003.

[15] M. E. J. Newman. Clustering and preferential attachment in growing networks, Phys. Rev. E 64, 025102, 2001.

[16] S.N. Dorogovtsev and J.F.F. Mendes. Evolution of networks. Oxford Clarendon Press, 2002.

[17] J. M. Kleinberg, R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins. The Web as a graph: measurements, models and methods. Proceedings of the International Conference on Combinatorics and Computing, 1999.

[18] R. Kumar, P. Raghavan, S. Rajalopagan, D. Divakumar, A. S. Tomkins, and E. Upfal. The Web as a graph. Proceedings of the 19th Symposium on principles of database systems, 2000.

[19] Pastor-Satorras, R., Smith, E. & Sole, R. Evolving proteinminteraction networks through gene duplication. J. Theor. Biol. 222:m199–210, 2003.

[20] Vazquez, A., Flammini, A., Maritan, A. & Vespignani, A. Modeling of protein interaction networks. ComPlexUs 1:38–44, 2003.

[21] G.S. Becker, The economic approach to Human Behavior. Chicago, 1976.

[22] A. Fabrikant, E. Koutsoupias, and C. Papadimitriou. Heuristically optimized trade-offs: a new paradigm for power laws in the internet. In Proceedings of the 29th International Colloquium on Automata, Languages, and Programming (ICALP), pages 110-122, Malaga, Spain, July 2002.

[23] RM. D'Souza, C. Borgs, J. T. Chayes, N. Berger, and R. D. Kleinberg, Emergence of tempered preferential attachment from optimization, PNAS 104, 6112-6117, 2007.

[24] F. Papadopoulos, M. Kitsak, M. Angeles Serrano, M. Boguna, and D. Krioukov, Popularity versus similarity in growing networks, Nature, 489: 537, 2012.

[25] A.-L. Barabási Network science: luck or reason, Nature 489: 1-2, 2012.

[26] H. A. Simon. On a class of skew distribution functions. Biometrika 42:3, 425-440, 1955.

[27] B. Mandelbrot. An Informational Theory of the Statistical Structure of Languages. In Communication Theory, edited by W. Jackson, pp. 486-502. Woburn, MA: Butterworth, 1953.

[28] B. Mandelbrot. A note on a class of skew distribution function: analysis and critique of a Paper by H.A. Simon. Information and control 2: 90-99, 1959.

[29] H. A. Simon. Some Further Notes on a class of skew distribution functions. Information and Control 3: 80-88, 1960.

[30] B. Mandelbrot. Final Note on a Class of Skew Distribution Functions: Analysis and Critique of a Model due to H.A. Simon. Information and Control 4: 198-216., 1961.

[31] H. A. Simon. Reply to final note. Information and Control 4:, 217-223, 1961.

[32] B. Mandelbrot. Post scriptum to final note. Information and Control 4: 300-304 1961.

[33] H. A. Simon. Reply to Dr. Mandelbrot's Post Scriptum. Information and Control 4: 305-308, 1961.

[34] R. Cohen and S. Havlin. Scale-free networks are ultrasmall. Phys. Rev. Lett., 90:058701, 2003.

[35] B. Bollobás and O. Riordan. The diameter of a scale-free random graph. Combinatorica, 24:5-34, 2004.

[36] K. Klemm and V. M. Eguluz. Growing scale-free networks with small-world behavior. Phys. Rev. E, 65:057102, 2002.

[37] B. Bollobás, O.M. Riordan. Mathematical results on scale-free random graphs, in the Handbook of Graphs and Networks, edited by S. Bormholdt and A. G. Schuster, Wiley, 2003.