A whirlwind tour of random graphs *

Fan Chung †

April 1, 2008

Contents

*Part of this survey is adapted from [16]

†University of California, San Diego

1 Introduction

Nowadays we are surrounded by assorted large information networks. For example, the phone network has all users as vertices which are interconnected by phone calls from one user to another. The Web can be viewed as a network with webpages as vertices which are then linked to other webpages. There are various biological networks arising from numerous databases, such as the gene network which represents the regulatory effect among genes. Of interest are many social networks expressing various types of social interactions. Some noted examples include the Collaboration graph (denoting coauthorship among mathematicians) and the Hollywood graph (consisting of actors/actresses and their joint appearances in feature films), among others.

How are these networks formed? What are basic structures of such large networks? How do they evolve? What are the underlying principles that dictate their behaviors?

To answer these questions, graph theory comes into play. Random graphs have a similar flavor as these large information networks in a natural way. For example, the phone network is formed by making random phone calls while a random graph results from adding a random edge one at a time. Although the classical random graphs can not directly be used to model real networks and seem to exhibit different 'shapes', the methods and approaches in random graph theory provides useful tools for the modeling and analysis of these information networks.

In this article, we will start with some basic graph theory in Section 2. We then introduce the main themes of random graphs in Section 3. Then we consider the classical random graph theory in Section 4 before we proceed to describe some general random graph models with given degree distributions, in particular, the power law graphs in Section 5. In Section 6, we will cover two types of "on-line" graph models, including the model of preferential attachment and the duplication model.

Although random graphs can be used to analyze various aspects of realistic networks, we wish to point out that there is no silver bullet to answer all the difficult problems about these large complex networks. In the last section we will put things in perspective by clarifying what random graphs can and can not do.

2 Some basic graph theory

All the information networks that we have mentioned can be formulated in terms of graphs. A graph G consists of a vertex set, denoted by $V = V(G)$ (which

contains all the objects that we wish to deal with) and an edge set $E = E(G)$ which consists of of specified pairwise relations between vertices. For example, a friendship graph has the vertex set consisting of people of interest and the edge set denoting the pairs of people who are friends. In Table 1 we list a number of graphs associated with various networks.

Graph	Vertices	Edges
Flight schedule graph	cities	flights
Phone graph	telephone numbers	phone calls
Collaboration graph	authors in Math Review	coauthorship
Web graph	Webpages	links
Biological graph	genes	regulatory effects

Table 1: Graph models for several networks.

As an introduction to graph theory, we describe the so-called *party problem*:

> *Among six people in a party, show that there are at least three people who know each other or there are three people who do not know each other.*

This can be said in graph-theoretical terms:

> *Any graph on* 6 *vertices must contain a triangle or contain three independent vertices with no edge among them.*

Indeed, 6 is the smallest number for this to occur since there is a graph on 5 vertices that contain neither a triangle nor three independent vertices. Such a graph is a cycle on 5 vertices, denoted by C_5, as seen in Figure 1.

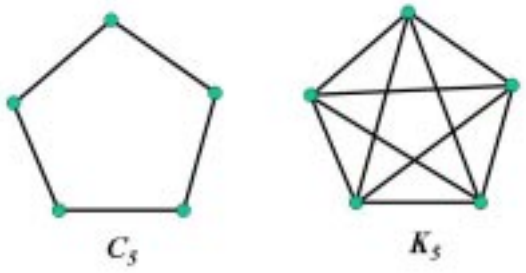

Figure 1: A five cycle C_5 and a complete graph K_5.

Let K_n denote a complete graph on n vertices which has all $\binom{n}{2}$ edges. For example, a triangle is K_3 which turns out is also C_3. The above party problem is

a toy case of the so-called *Ramsey theory* which deals with unavoidable patterns in large graphs. In 1930, Ramsey [40] showed the following:

> *For any two positive integers k and l, there is an associated number $R(k,l)$ such that any graph on $n \geq R(k,l)$ vertices must contain either K_k as a subgraph or contain l independent vertices.*

For example, $R(3,3) = 6$ as stated in the party problem. It is not too difficult to show that $R(4,4) = 17$. However, the value of $R(5,5)$ is not yet determined (in spite of the huge computational power we have today). All that is known is $43 \leq R(5,5) \leq 49$ (see [25] and [36]). Relatively few exact Ramsey numbers $R(k,l)$ are determined. For an extensive survey on this topic, the reader is referred to the dynamic survey in the Electronic Journal of Combinatorics at http://www.combinatorics.org/.

In 1947, Erdős wrote an important paper [21] that helped start two areas including combinatorial probabilistic methods and Ramsey theory. He established the following lower bound for the Ramsey number $R(k,k)$ by proving

$$R(k,k) \geq 2^{k/2}, \tag{1}$$

the argument is quite simple and elegant:

Suppose we wish to find a graph on n vertices that does not contain K_k or an independent subset of k vertices. How large can n be? For a fixed integer n, there are all together $2^{\binom{n}{2}}$ possible graphs on n vertices. We say a graph is *bad* if it contains K_k or an independent subset of k vertices. How many bad graphs can there be? There are $\binom{n}{k}$ ways to choose k out of n vertices. So, there are at most $2\binom{n}{k}2^{\binom{n}{2}-\binom{k}{2}}$ bad graphs. Therefore there is a graph on n vertices that is not bad if

$$2^{\binom{n}{2}} \geq 2\binom{n}{k}2^{\binom{n}{2}-\binom{k}{2}}.$$

So, for $n \geq 2^{k/2}$, there must be a graph on n vertices that is not bad, which implies (1).

We note that for the upper bound there is an inductive proof to show that $R(k,k) \leq \binom{2k-2}{k-2}$ which is about 4^k. In the previous five decades, there have been some improvements only by a factor of a lower order for both the upper and lower bounds [42, 19]. It remains unsettled (with Erdős award unclaimed) to determine if $\lim_{k\to\infty}(R(k,k))^{1/k}$ exists or what value it should be.

A basic notion in graph theory is "adjacency". A vertex u is said to be *adjacent* to another vertex v if $\{u,v\}$ is an edge. Or, we say u is a *neighbor* of v. Equivalently, v is a neighbor of u. The *degree* of a vertex u is the number of edges containing u. If we restrict ourselves to *simple* graphs (i.e., at most one edge between any pairs of vertices), then the degree of u is just the number of

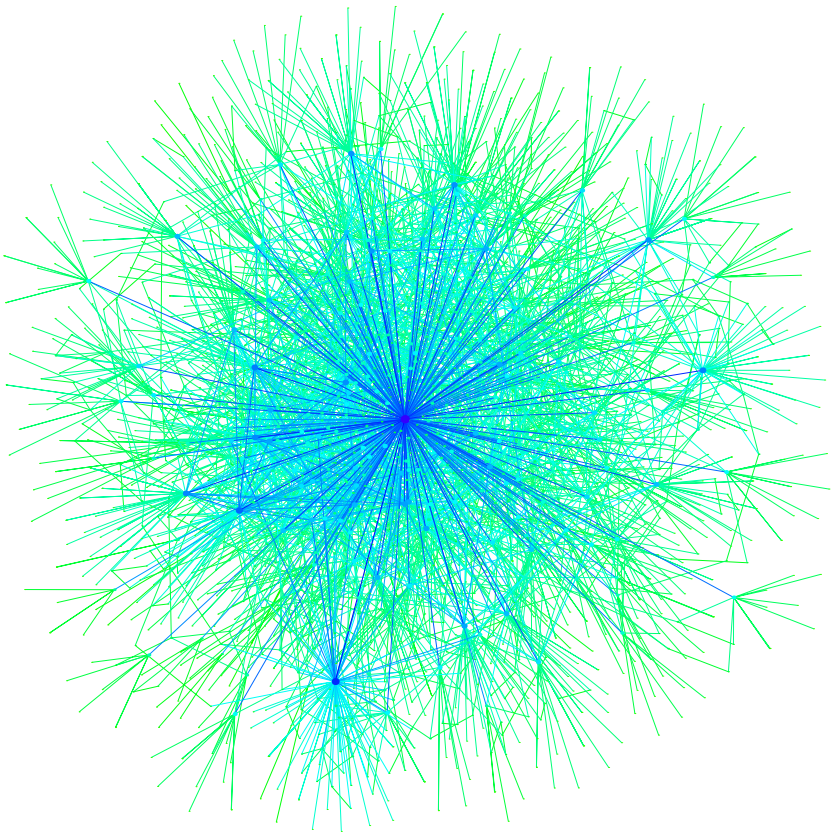

Figure 2: The collaboration graph.

neighbors that u has. Suppose that in a graph G the vertex v_i has degree d_i for $1 \leq i \leq n$. Then $(d_1, d_2, \ldots, d_n)$ forms a degree sequence for G. Sometimes, we organize the degree sequence so that $d_1 \geq d_2 \geq \ldots \geq d_n$. Here comes a natural question on graph realization: For what values d_i is the sequence $(d_1, d_2, \ldots, d_n)$ a degree sequence of some graph?

To answer this question, first we observe that the sum of all d_i's must be even since that is exactly twice the number of edges. This is the folklore "Handshake Theorem".

In a 1961 paper, Erdős and Gallai [24] answered the above question. They gave a necessary and sufficient condition by showing that a sequence $(d_1, d_2, \ldots, d_n)$, where $d_i \geq d_{i+1}$, is a degree sequence of some graph if and only if the sum of d_i's is even and for each integer $r \leq n-1$,

$$\sum_{i=1}^{r} d_i \leq r(r-1) + \sum_{i=r+1}^{n} \min\{r, d_i\}.$$

Another way to keep track of the degrees of a graph is to consider the degree distribution as follows: Let n_k denote the number of vertices having degree k. Instead of writing down the degree sequence (which consists of n numbers and n can be a very large number), we just use n_k. Therefore, the number of values that we need to keep does not exceed the maximum degree. If all degrees are the same value, we say the graph is regular. In this case, only one of the n_k's is nonzero.

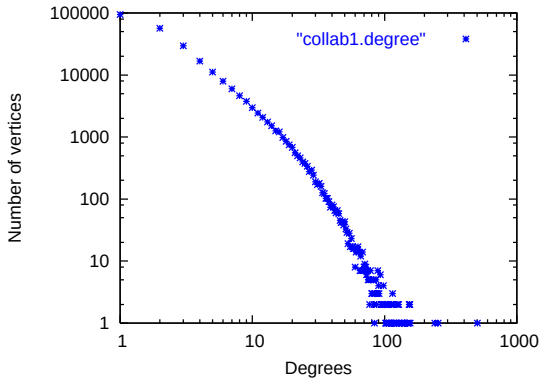

Figure 3: The number of vertices for each possible degree for the collaboration graph.

Many real-world graphs have degree distribution satisfying the so-called "*power law*". Namely, the number n_k of vertices of degree k is proportional to $k^{-\beta}$ for some fixed positive value β. For example, the Collabration graph, as illustrated in Figure 2, can be approximated by a power law with exponent $\beta = 2.46$. The degree distribution of the Collaboration graph is included in Figure 3 in log-log scale.

In a graph G, a *path* is a sequence of vertices $v_0, v_1, \ldots, v_k$ such that v_{i-1} is adjacent to v_i for $i = 1, \ldots, k$. The length of a path is the number of edges in the path. For example, the above mentioned path has length k joining v_0 and v_k. If $v_0 = v_k$, the path is said to be a *cycle*. A graph which contains no cycle is called a *tree*. A graph is *connected* if any two vertices can be joined by a path. For a graph G, a maximum subset of vertices each pair of which can be joined by paths is called a *connected component*. Thus, a graph is connected if there is only one connected component. In a connected graph, the *distance* between two vertices u and v is the length of a shortest path joining u and v. The maximum distance among all pairs of vertices is called the *diameter* of a graph.

In 1967, the psychologist Stanley Milgram [37] conducted a series of experiments which indicated that any two strangers are connected by a chain of intermediate acquaintances of length at most six. Since then, the so-called "*small world phenomenon*" has long been a subject of anecdotal observation

and folklore. Recent studies have suggested that the phenomenon is pervasive in numerous networks arising in nature and technology, and in particular, in the structural evolution of the World Wide Web [5, 33, 44].

In addition to "six degrees of separation", various numbers have emerged with many networks. In 1999, Barabási et al. [5] estimated that any two webpages are at most 19 clicks away from one another (in certain models of the Internet). Broder et al. [11] set up crawlers on a webgraph of 200 million nodes and 1.5 billion links and reported that the average path length is about 16. A mathematician who has written joint papers is likely to have Erdős number at most eight [29] (i.e., with a chain of coauthors with length at most 8 connecting to Erdős). The majority of actors or actresses have the so-called "Kevin Bacon number" two or three.

Before we make sense of these numbers, some clarification is in order: There are in fact two different interpretations of 'short distance' in a network. One notion is the diameter of the graph. Another notion is the *average distance* (which might be closer to what was meant by these experiments). We will discuss the small world phenomenon further in a later section.

3 Random graphs in a nutshell

What does a "random graph" mean? Before proceeding to describe random graphs, some clarification for "random" is in order. According to the Cambridge Dictionary, "random" means "happening, done or chosen by chance rather than according to a plan". Quite contrary to this explanation, our random graphs have precise meanings and can be clearly defined. Using the terminology in probability, a random graph is a random variable defined in a probability space with a probability distribution. In layman's terms, we first put all graphs on n vertices in a lottery box and then the graph we pick out of the box is a random graph. (In this case, all graphs are chosen with equal probability.)

What do we want from our random graphs? Well, we would like to say that a random graph (in some given model) has certain properties (e.g., having small diameter). Such a statement means that with probability close to 1 (as the number n of vertices approaches infinity), the random graph we pick out of the lottery box satisfies the property that we specified. In other words, a random graph has a specified property means that almost all graphs of interest have the desired property. Note that this is quite a strong implication! Any statement about a random graph is really about almost all graphs! The beauty of random graphs lies in being able to use relatively few parameters in the model to capture the behavior of almost all graphs of interest (which can be quite numerous and complex).

In the early days of the subject, Erdős and Rényi introduced two random

graph models. The first one is a random graph $\mathcal{F}(n,m)$ defined on all graphs with n vertices and m edges each of which is chosen with equal probability. The second is the celebrated Erdős-Rényi random graph $\mathcal{G}(n,p)$ defined on all graphs on n vertices and each edge is chosen independently with probability p. Consequently, a graph on n vertices and x edges is chosen with probability $p^x(1-p)^{\binom{n}{2}-x}$ in $\mathcal{G}(n,p)$.

The advantage of the Erdős-Rényi model is the independence of choices for the edges (i.e., each pair of vertices has its own dice for determining being chosen as an edge). Since the probability of two independent events is the product of probabilities of two events, we can compute with ease. For example, the probability of a random graph in $\mathcal{G}(n,p)$ containing a fixed triangle is $1/8$ for $p=1/2$. It is possible to compute such a probability for a random graph in $\mathcal{F}(n,m)$, e.g., $\binom{\binom{n}{2}-3}{m-3}/\binom{\binom{n}{2}}{m}$, which is a more complicated expression. For many problems, such as the diameter problem, it can be quite nontrivial for $\mathcal{F}(n,m)$ because the dependency among edges is getting in the way.

To model real graphs, there are some obvious difficulties. For example, the random graph $\mathcal{G}(n,p)$ has all degrees very close to pn if the graph is not so sparse, (i.e., $p \geq \log n/n$). The distribution of the degrees follows the same bell curve for every vertex. As we know, many real-world graphs satisfy the power law which is very different from the degree distribution of $\mathcal{G}(n,p)$. In order to model real-world networks, it is imperative to consider random graphs with general degree distribution and, in particular, the power law distribution.

There are basically two types of random graph models for general degree distributions. The *configuration model* is a take-off from random regular graphs [6]. The way to define random regular graphs $\mathcal{G}_k$ of degree k on n vertices is to consider all possible matchings in a complete graph K_{kn}. Note that a matching is a maximum set of vertex-disjoint edges. Each matching is chosen with equal probability. We then get a random k-regular graph by partitioning the vertices into subsets of size k. Each k- subset then is associated with a vertex in a random regular graph $\mathcal{G}_k$. Although such a random regular graph might contain loops (i.e., an edge having both endpoints the same vertex), the probability of such an event is of a lower order and can be controlled. It is then obvious to define random graphs with general degrees. Instead of partitioning the vertex set of the large graph into equal parts, we choose a random matching of a complete graph on $\sum_i d_i$ vertices which are partitioned into subsets of sizes $d_1, d_2, \ldots, d_n$. Then we form the random graph by associating each edge in the matching with an edge between associated vertices.

In the configuration model, there are nontrivial dependencies among the edges. As a generalization of the Erdős-Rényi model, there is a *random graph model for given expected degrees*. Let $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ denote the specified degrees. The $\mathcal{G}(\mathbf{w})$ model yields random graphs with expected degrees $\mathbf{w}$. The edge between v_i and v_j is independently chosen with probability $w_i w_j/W$

where $W = \sum_i w_i$. In other words, each pair of vertices has its own dice with probability assigned so that the expected degree at vertex v_i is exactly w_i. The Erdős-Rényi model is just the case with all w_i's equal to pn. Since the $\mathcal{G}(\mathbf{w})$ model inherits the robustness and independence of the Erdős-Rényi model, many strong properties can be derived. We will discuss some of these further in Section 5, especially when $\mathbf{w}$ satisfies power laws.

All the random graph models mentioned above are *off-line* models. Since real-world graphs are dynamically changing — both in adding and deleting vertices and edges — there are several *on-line random graph models* in which the probability spaces are changing at the tick of the clock. In fact, in the study of complex real-world graphs, the on-line model came to attention first.

There are a large number of research papers, surveys and books on random graphs, mostly about the Erdős-Rényi model $\mathcal{G}(n,p)$. After the year 2000, the study of real-world graphs has led to interesting directions and new methods for analyzing random graphs with general degree distributions. Many on-line models have been proposed and published. Here we will only be able to cover the main ones — the preferential attachment schemes and the duplication model.

4 Classical random graphs

In early 60's, Erdős and Rényi wrote a series of influential papers on random graphs. Their modeling and analysis are thorough and elegant. Their approaches and methods are powerful and have had enormous impact up to this day. In this section, we will give a brief overview. First we will describe the classical results on the evolution of random graphs $\mathcal{G}(n,p)$ of the Erdős-Rényi model. Then we will discuss the diameter of $\mathcal{G}(n,p)$ as the edge density ranges from 0 to 1.

4.1 The evolution of the Erdős-Rényi graph

What does a random graph in $\mathcal{G}(n,p)$ look like? Erdős and Rényi [23] gave a full answer for the edge density p ranging from 0 to 1.

At the start, there is no edge and the edge density is 0. We have isolated vertices.

As p increases, the expected number $p\binom{n}{2}$ of edges gets larger. When there are about $\sqrt{n}$ edges, how many connected components are there and what sizes and structures are they? For $0 < p \ll 1/n$, Erdős and Rényi [23] showed that the random graph G is a disjoint union of trees. Furthermore, they gave a beautiful formula. For $p = cn^{-k/(k-1)}$, the probability that j is the number of connected components in G formed by trees on k vertices is $\lambda^j e^{-\lambda}/j!$ where $\lambda =$

$(2c)^{k-1}k^{k-2}/k!$. For example, when we have about $\sqrt{n}$ edges, the probability that the random graph G contains j trees on 3 vertices is close to $2^j/j!$ (if n is large enough).

As we have more edges, cycles start to appear. When the graph has a linear number of edges, i.e., $p = c/n$, with $c < 1$, almost all vertices are in connected components of trees and there are only a small number of cycles. Namely, the expected number of cycles is $\frac{1}{2}\log\frac{1}{1-c} - \frac{c}{2} - \frac{c^2}{4}$.

When a random graph has edges ranging from slightly below $n/2$ to slightly over $n/2$ edges, i.e., $p = (1 + o(1))/n$, there is an unusual phenomenon, called "*double jumps*". What are double jumps and why is it so unusual? In the study of "threshold function" or, "phase transition" that happens in natural or evolving systems, it is of interest to identify the critical point, below which the behavior is dramatically different from what is above. Erdős and Rényi [23] found that as p is smaller than $1/n$, the largest component in G has size $O(\log n)$ and all components are either trees or unicyclic (i.e., each component contains at most one cycle). If p is $(1 + \mu)/n$ and $\mu > 0$, then the giant component emerges. However, when $p = 1/n$, the largest component is of size $O(n^{2/3})$. There has been detailed analysis examining this tricky transition in details (see [7] and [31]).

When $p = c/n$ for $c > 1$, the random graph G has one giant component and all others are quite small, of size $O(\log n)$. Also, Erdős and Rényi [23] determined the number of vertices in the giant connected component to be $f(c)n$ where

$$f(c) = 1 - \frac{1}{c}\sum_{k=1}^{\infty}\frac{k^{k-1}}{k!}(ce^{-c})^k. \tag{2}$$

Finally when $p = c\log n/n$ and $c > 1$, the random graph G is almost always connected. When c goes to infinity, G is not only connected but is almost regular. Namely, all vertices have degrees close to pn.

4.2 The diameter of the Erdős-Rényi graph

We consider the diameter of a random graph G in $\mathcal{G}(n,p)$ for all ranges of p including the range for which $\mathcal{G}(n,p)$ is not connected. For a disconnected graph G, the diameter of G is defined to be the diameter of its largest connected component.

Roughly speaking, the diameter of a random graph in $\mathcal{G}(n,p)$ is of order $\frac{\log n}{\log(np)}$ if the expected degree np is at least 1. Note that this is best possible in the following sense. For any graph with degrees at most d, the number of vertices that can be reached within distance k is at most $1 + d + d(d-1) + d(d-1)^2 +$

$\ldots + d(d-1)^{k-1}$. This sum should be at least n if k is the diameter. Therefore we known that the diameter, denoted by $diam(G)$ is at least $(\log n)/\log(d-1)$.

To be precise, it can be shown the diameter of a random graph G in $\mathcal{G}(n,p)$ is $(1+o(1))\frac{\log n}{\log np}$ if the expected degree np goes to infinity as n approaches infinity. When $np \geq c > 1$, the diameter $diam(G)$ is within a constant factor of $\frac{\log n}{\log np}$ where the constant depends only on c and is independent of n. When $np = c < 1$, the random graph is surely disconnected and $diam(G)$ is equal to the diameter of a tree component.

In fact, the diameter of a graph G in $\mathcal{G}(n,p)$ is quite predictable as follows. The values for the diameter of $\mathcal{G}(n,p)$ is almost surely concentrated on at most two values around $\frac{\log n}{\log np}$ if $\frac{np}{\log n} = c > 8$. When $\frac{np}{\log n} = c > 2$, the diameter of $\mathcal{G}(n,p)$ is almost surely concentrated on at most three values. For the range $2 \geq \frac{np}{\log n} = c > 1$, the diameter of $\mathcal{G}(n,p)$ is almost surely concentrated on at most four values.

Range	$diam(\mathcal{G}(n,p))$	Reference
$\frac{np}{\log n} \to \infty$	Concentrated on at most 2 values	[9]
$\frac{np}{\log n} = c > 8$	Concentrated on at most 2 values	[12]
$8 \geq \frac{np}{\log n} = c > 2$	Concentrated on at most 3 values	[12]
$2 \geq \frac{np}{\log n} = c > 1$	Concentrated on at most 4 values	[10]
$1 \geq \frac{np}{\log n} = c > c_0$	Concentrated on at most $2\lfloor\frac{1}{c_0}\rfloor + 4$ values	[12]
$\log n > np \to \infty$	$diam(\mathcal{G}(n,p)) = (1+o(1))\frac{\log n}{\log(np)}$	[12]
$np \geq c > 1$	The ratio $\frac{diam(\mathcal{G}_{n,p})}{\frac{\log n}{\log(np)}}$ is finite (between 1 and $f(c)$)	[12]
$np < 1$	$diam(\mathcal{G}(n,p))$ equals the diameter of a tree component if $(1-np)n^{1/3} \to \infty$	[35]

Table 2: The diameter of random graphs $\mathcal{G}(n,p)$.

It is of particular interest to consider random graphs $\mathcal{G}(n,p)$ for the range of $np > 1$ and $np \leq c\log n$ for some constant c since this range includes the emergence of the unique giant component. Because of a phase transition in connectivity at $p = \log n/n$, the problem of determining the diameter of $\mathcal{G}(n,p)$ and its concentration seems to be difficult for certain ranges of p. If $\frac{np}{\log n} = c > c_0$ for any (small) constant c and c_0, then the diameter of $\mathcal{G}(n,p)$ is almost surely concentrated on finitely many values, namely, no more than $2\lfloor\frac{1}{c_0}\rfloor + 4$ values.

These facts are summarized in Table 2 with references listed. As we can see from the table, numerous questions remain.

5 Random power law graphs

5.1 Parameters for modeling power law graphs

A large realistic network usually has a huge number of parameters with complicated descriptions. By "modeling a realistic network", we mean cutting down the number of parameters to relatively few and still capture a good part of the character of the network.

To choose the parameters for modeling a real network, the exponent β of the power law is relatively easy to select. We can plot the log-degree versus log-frequency table and choose a good approximation of the slope.

In a graph G, suppose that there are y vertices of degree x. Then G is considered to be a power law graph if x and y satisfy (or can be approximated by) the following equation:

$$\log y = \alpha - \beta \log x. \tag{3}$$

In other words, we have

$$\mid \{v|deg(v) = x\} \mid \quad \approx y = \frac{e^{\alpha}}{x^{\beta}}.$$

Basically, α is the logarithm of the volume of the graph and β can be regarded as the log-log growth rate of the graph.

To take a closer look of the degree distribution of a typical realistic graph, several impediments obviously exist.

(a) When we fit the power law model, there are discrepancies especially when the degree is very small or very large. There is almost always a heavy tail distribution at the upper range and there seems to be scattering at the lower range. For example, for the collaboration graph, should we or shouldn't we include the data point for isolated vertices (an author with no coauthors)? Should we stay with the largest component or include all small components (including the isolated vertices)?

(b) The power law states that the number of vertices of degree k is proportional to $k^{-\beta}$. We can approximate the number of vertices of degree k by the function $f(k) = ck^{-\beta}$ for some constant c. However, $f(k)$ is usually not an integer. By taking either the ceiling or floor of $f(k)$, some errors are inevitable. In fact, such errors are acute when k or $f(k)$ is small.

(c) The power law model is usually a better fit in the middle range (than at either end). Still, in many examples, there is a visible slight "hump" in the curve instead of the straight line representing the power law in the log-log table.

Among the above three points, (c) is mainly due to first-order approximations. The straight line with slope β is a linear approximation of the actual plotted data. Thus the power law model is an important and necessary step for more complicated real cases. Here, we will first discuss (b) and then (a).

Item (b) concerns rounding errors which can be checked by the following basic calculations about the power law graphs according to (3).

(1) The maximum degree of the graph is at most $e^{\frac{\alpha}{\beta}}$. Note that $0 \leq \log y = \alpha - \beta \log x$.

(2) The number of vertices n can be computed as follows (under the assumption that the maximum degree is $e^{\frac{\alpha}{\beta}}$). By summing $y(x)$ for x from 1 to $e^{\frac{\alpha}{\beta}}$, we have

$$n = \sum_{x=1}^{e^{\frac{\alpha}{\beta}}} \frac{e^{\alpha}}{x^{\beta}} \approx \begin{cases} \zeta(\beta)e^{\alpha} & \text{if } \beta > 1, \\ \alpha e^{\alpha} & \text{if } \beta = 1, \\ \frac{e^{\frac{\alpha}{\beta}}}{1-\beta} & \text{if } 0 < \beta < 1, \end{cases}$$

where $\zeta(t) = \sum_{n=1}^{\infty} \frac{1}{n^t}$ is the Riemann Zeta function.

(3) The number of edges E can be computed as follows:

$$E = \frac{1}{2} \sum_{x=1}^{e^{\frac{\alpha}{\beta}}} x \frac{e^{\alpha}}{x^{\beta}} \approx \begin{cases} \frac{1}{2}\zeta(\beta-1)e^{\alpha} & \text{if } \beta > 2, \\ \frac{1}{4}\alpha e^{\alpha} & \text{if } \beta = 2, \\ \frac{1}{2}\frac{e^{\frac{2\alpha}{\beta}}}{2-\beta} & \text{if } 0 < \beta < 2. \end{cases}$$

(4) The differences of the real numbers in (1)-(3) and their integer parts can be estimated as follows: For the number n of vertices, the error term is at most $e^{\frac{\alpha}{\beta}}$. For $\beta \geq 1$, it is $o(n)$, which is a lower order term. For $0 < \beta < 1$, the error term for n is relatively large. In this case, we have

$$n \leq \frac{e^{\frac{\alpha}{\beta}}}{1-\beta} - e^{\frac{\alpha}{\beta}} = \frac{\beta e^{\frac{\alpha}{\beta}}}{1-\beta}.$$

As can be seen, n can have the same magnitude as $\frac{e^{\frac{\alpha}{\beta}}}{1-\beta}$. Therefore the rounding error can be of the same order of magnitude. For the number E of edges, similar situations occur. For $\beta \geq 2$, the rounding error term of E is $o(E)$, a lower order term. For $0 < \beta < 2$, the error of E has the same magnitude as in the formula of item (3). Thus, one is advised to exercise caution when dealing with the case $0 < \beta < 2$.

To deal with the concerns mentioned above in (a), we need additional parameters.

- The average degree w is a useful parameter.
- The second order average degree $\tilde{w} = \sum_i w_i^2 / \sum_i w_i$.

- The maximum degree $m = d_{\max}$ and also the minimum degree $d_{\min}$ denote the range that the power law degree distribution fits (within acceptable approximation). In other words, the maximum degree $m = d_{\max}$ and the minimum degree $d_{\min}$ are meant to be the largest and the least degrees in a power law subgraph of G. Often, $d_{\min}$ is taken to be 1 unless otherwise specified.

With these parameters, we are ready to define a random power law graph. For random graphs with given expected degree sequences satisfying a power law distribution with exponent β, we may assume that the expected degrees are $w_i = ci^{-\frac{1}{\beta-1}}$ for i satisfying $i_0 \leq i < n + i_0$. Here c depends on the average degree and i_0 depends on the maximum degree m, namely, $c = \frac{\beta-2}{\beta-1} w n^{\frac{1}{\beta-1}}, i_0 = n(\frac{w(\beta-2)}{m(\beta-1)})^{\beta-1}$.

The power law graphs with exponent $\beta > 3$ are quite different from those with exponent $\beta < 3$ as evidenced by the value of $\tilde{w}$ (assuming $m \gg w$).

$$\tilde{w} = \begin{cases} (1+o(1))w\frac{(\beta-2)^2}{(\beta-1)(\beta-3)} & \text{if } \beta > 3, \\ (1+o(1))\frac{1}{2}w\ln\frac{2m}{w} & \text{if } \beta = 3, \\ (1+o(1))d^{\beta-2}\frac{(\beta-2)^{\beta-1}m^{3-\beta}}{(\beta-1)^{\beta-2}(3-\beta)} & \text{if } 2 < \beta < 3. \end{cases}$$

The above values of $\tilde{w}$ are quite useful in the study of average distance and diameter of random graphs.

5.2 The evolution of random power law graphs

A natural question concerning the configuration model is how the random graphs evolve for power law distributions. Can we mimic the classical analysis as in the Erdős-Rényi random graph model?

Here we consider a configuration model with degree distribution as in the (α, β)-graph. As it turns out, the evolution only depends on β and not on α as follows.

1. When $\beta > \beta_0 = 3.47875\ldots$, the random graph almost surely has no giant component where the value $\beta_0 = 3.47875\ldots$ is a solution to
$$\zeta(\beta-2) - 2\zeta(\beta-1) = 0.$$
When $\beta < \beta_0 = 3.47875\ldots$, there is almost surely a unique giant component.

2. When $2 < \beta < \beta_0 = 3.47875\ldots$, the second largest component is almost surely of size $\Theta(\log n)$. For any $2 \leq x < \Theta(\log n)$, there is almost surely a component of size x.

3. When $\beta = 2$, almost surely the second largest component is of size $\Theta(\frac{\log n}{\log\log n})$. For any $2 \leq x < \Theta(\frac{\log n}{\log\log n})$, there is almost surely a component of size x.

4. When $1 < \beta < 2$, the second largest component is almost surely of size $\Theta(1)$. The graph is almost surely not connected.

5. When $0 < \beta < 1$, the graph is almost surely connected.

6. When $\beta = \beta_0 = 3.47875\ldots$, the situation is complicated. It is similar to the double jump of the random graph $\mathcal{G}(n,p)$ with $p = \frac{1}{n}$. For $\beta = 1$, there is a nontrivial probability for either case that the graph is connected or disconnected.

A useful tool in configuration model is a result of Molloy and Reed [38, 39]:

For a random graph with $(\gamma_i + o(1))n$ vertices of degree i, where γ_i are nonnegative values which sum to 1 and n is the number of vertices, the giant component emerges when $Q = \sum_{i\geq 1} i(i-2)\gamma_i > 0$, provided that the maximum degree is less than $n^{1/4-\epsilon}$ and some "smoothness" conditions are satisfied. Also, there is almost surely no giant component when $Q = \sum_{i\geq 1} i(i-2)\gamma_i < 0$ and the maximum degree is less than $n^{1/8-\epsilon}$.

Let us consider Q for our (α,β)-graphs with $\beta > 3$.

$$\begin{aligned} Q &= \frac{1}{n}\sum_{x=1}^{e^{\frac{\alpha}{\beta}}} x(x-2)\lfloor\frac{e^\alpha}{x^\beta}\rfloor \\ &\approx \frac{1}{\zeta(\beta)}\left(\sum_{x=1}^{e^{\frac{\alpha}{\beta}}}\frac{1}{x^{\beta-2}} - 2\sum_{x=1}^{e^{\frac{\alpha}{\beta}}}\frac{1}{x^{\beta-1}}\right) \\ &\approx \frac{\zeta(\beta-2)-2\zeta(\beta-1)}{\zeta(\beta)} \end{aligned}$$

Hence, we consider the value $\beta_0 = 3.47875\ldots$, which we recall is a solution to $\zeta(\beta-2)-2\zeta(\beta-1) = 0$. If $\beta > \beta_0$, we have

$$\sum_{x=1}^{e^{\frac{\alpha}{\beta}}} x(x-2)\lfloor\frac{e^\alpha}{x^\beta}\rfloor < 0.$$

We remark that for $\beta > 8$, Molloy and Reed's result immediately implies that almost surely there is no giant component. When $\beta \leq 8$, additional analysis is needed to deal with the degree constraints [2].

It can be shown that the second largest component almost surely has size $\Theta(\log n)$. Furthermore, the second largest component has size at least $\Theta(\log n)$.

5.3 $\mathcal{G}(\mathbf{w})$ model for power law graphs

In the Erdős-Rényi model $\mathcal{G}(n,p)$, the threshold function for the phase transition of the giant component is at $p = 1/n$. Namely, when the average degree pn is less than 1, all connected components are small (of size $O(\log n)$) and there is no giant component. When the average degree is more than 1, the giant component emerges in full swing. (There is a "double jump" which takes place when the average degree is close to 1 as discussed in Section 4.)

For the random graph model $\mathcal{G}(\mathbf{w})$, with given expected degrees $\mathbf{w}$, it is natural to ask the same question:

> What parameter in which range will trigger the (sudden) emergence of the giant component?

In addition to w, the expected average degree, we have scores of parameters, e.g., $\tilde{w}$ and higher order average degrees. Which parameter w, $\tilde{w}$ or others is critical for the rise of the giant component?

These questions were answered in [13]:

Suppose that G is a random graph in $\mathcal{G}(\mathbf{w})$ with expected degree sequence $\mathbf{w}$. If the expected average degree w is strictly greater than 1, then the following holds:
(1) Almost surely G has a unique giant component. Furthermore, the volume of the giant component is at least $(1 - \frac{2}{\sqrt{we}} + o(1))\mathrm{Vol}(G)$ if $w \geq \frac{4}{e} = 1.4715\ldots$, and is at least $(1 - \frac{1+\log w}{w} + o(1))\mathrm{Vol}(G)$ if $w < 2$.
(2) The second largest component almost surely has size at most $(1+o(1))\mu(w)\log n$, where

$$\mu(w) = \begin{cases} \frac{1}{1+\log w - \log 4} & \text{if } w > 4/e; \\ \frac{1}{w-1-\log w} & \text{if } 1 < w < 2. \end{cases}$$

Moreover, with probability at least $1 - n^{-k}$, the second largest component has size at most $(k+1+o(1))\mu(w)\log n$, for any $k \geq 1$.

There is a sharp asymptotic estimate for the volume of the giant component for a random graph in $\mathcal{G}(\mathbf{w})$. In [15], it was proved that if the expected average degree is strictly greater than 1, then almost surely the giant component in a graph G in $\mathcal{G}(\mathbf{w})$ has volume $\lambda_0\mathrm{Vol}(G) + O(\sqrt{n}\log^{3.5} n)$, where λ_0 is the unique nonzero root of the following equation:

$$\sum_{i=1}^{n} w_i e^{-w_i\lambda} = (1-\lambda)\sum_{i=1}^{n} w_i. \tag{4}$$

Because of the robustness of the $\mathcal{G}(\mathbf{w})$ model, many properties can be derived for appropriate degree distributions, including power law graphs.

Average distance and the diameter:

A random graph G in $\mathcal{G}(\mathbf{w})$ has average distance almost surely $(1+o(1))\frac{\log n}{\log \tilde{w}}$, if $\mathbf{w}$ satisfies certain conditions (called admissible conditions in [14]). The diameter is almost surely $\Theta(\frac{\log n}{\log \tilde{w}})$. In addition to studying the average distance and diameter, the structure of a random power law graph is very interesting, especially for the range $2 < \beta < 3$ where the power law exponents β for numerous real networks reside. In this range, the power law graph can be roughly described as an "octopus" with a dense subgraph having small diameter $O(\log \log n)$, as the core while the overall diameter is $O(\log n)$ and the average distance is $O(\log \log n)$. When $\beta > 3$ and the average degree w is strictly greater than 1, almost surely the average distance is $(1+o(1))\frac{\log n}{\log \tilde{w}}$ and the diameter is $\Theta(\log n)$. A phase transition occurs at $\beta = 3$ and then the graph has diameter almost surely $\Theta(\log n)$ and average distance $\Theta(\log n/\log \log n)$.

Eigenvalues:

Eigenvalues of graphs are useful for controlling many graph properties and consequently have numerous algorithmic applications including clustering algorithms, low rank approximations, information retrieval and computer vision. In the study of the spectra of power law graphs, there are basically two competing approaches. One is to prove analogues of Wigner's semi-circle law (such as for $\mathcal{G}(n,p)$) while the other predicts that the eigenvalues follow a power law distribution [27]. Although the semi-circle law and the power law have nothing in common, both approaches are essentially correct if one considers the appropriate matrices. there are in fact several ways to associate a matrix to a graph. The usual adjacency matrix A associated with a (simple) graph has eigenvalues quite sensitive to the maximum degree (which is a *local* property). The combinatorial Laplacian $D - A$ with D denoting the diagonal degree matrix is a major tool for enumerating spanning trees and has numerous applications. Another matrix associated with a graph is the (normalized) Laplacian $L = I - D^{-1/2}AD^{-1/2}$ which controls the expansion/isoperimetrical properties (which are *global*) and essentially determines the mixing rate of a random walk on the graph. The traditional random matrices and random graphs are regular or almost regular so the spectra of all the above three matrices are basically the same (with possibly a scaling factor or a linear shift). However, for graphs with uneven degrees, the above three matrices can have very different distributions.

Here we state bounds for eigenvalues for random graphs in $\mathcal{G}(\mathbf{w})$ with a general degree distribution from which the results on random power law graphs then follow [18].

1. The largest eigenvalue of the adjacency matrix of a random graph with a given expected degree sequence is determined by m, the maximum degree, and $\tilde{w}$, the weighted average of the squares of the expected degrees. In this case the largest eigenvalue of the adjacency matrix is almost surely $(1+o(1))\max\{\tilde{w}, \sqrt{m}\}$ provided some minor conditions are satisfied. In addition, if the k^{th} largest expected degree m_k is significantly larger than

$\tilde{w}^2$, then the k^{th} largest eigenvalue of the adjacency matrix is almost surely $(1+o(1))\sqrt{m_k}$.

2. For a random power law graph with exponent $\beta > 2.5$, the largest eigenvalue of a random power law graph is almost surely $(1+o(1))\sqrt{m}$ where m is the maximum degree. Moreover, the k largest eigenvalues of a random power law graph with exponent β have power law distribution with exponent $2\beta - 1$ if the maximum degree is sufficiently large and k is bounded above by a function depending on β, m and w, the average degree. When $2 < \beta < 2.5$, the largest eigenvalue is heavily concentrated at $cm^{3-\beta}$ for some constant c depending on β and the average degree.

3. The eigenvalues of the Laplacian satisfy the semi-circle law under the condition that the minimum expected degree is relatively large ($\gg$ the square root of the expected average degree). This condition contains the basic case when all degrees are equal (the Erdős-Rényi model). If we weaken the condition on the minimum expected degree, we can still have the following strong bound for the eigenvalues of the Laplacian which implies strong expansion rates for rapid mixing,

$$\max_{i \neq 0} |1 - \lambda_i| \leq (1 + o(1)) \frac{4}{\sqrt{w}} + \frac{g(n) \log^2 n}{w_{\min}}$$

where w is the expected average degree, $w_{\min}$ is the minimum expected degree and $g(n)$ is any slow growing function of n.

6 On-line random graphs

6.1 Preferential attachment schemes

The preferential attachment scheme is often attributed to Herbert Simon. In his paper [41] of 1955, he gave a model for word distribution using the preferential attachment scheme and derived *Zipf's law* (i.e., the probability of a word having occurred exactly i times is proportional to $1/i$).

The basic setup for the preferential attachment scheme is a simple *local* growth rule which leads to a *global* consequence — a power law distribution. Since this local growth rule gives preferences to vertices with large degrees, the scheme is often described by "*the rich get richer*". Of interest is to determine the exponent of the power law from the parameters of the local growth rule.

There are two parameters for the preferential attachment model:

- A probability p, where $0 \leq p \leq 1$.

- An initial graph G_0, that we have at time 0.

Usually, G_0 is taken to be the graph formed by one vertex having one loop. (We consider the degree of this vertex to be 1, and in general a loop adds 1 to the degree of a vertex.) Note, in this model multiple edges and loops are allowed.

We also have two operations we can do on a graph:

- *Vertex-step* — Add a new vertex v, and add an edge $\{u, v\}$ from v by randomly and independently choosing u in proportion to the degree of u in the current graph.
- *Edge-step* — Add a new edge $\{r, s\}$ by independently choosing vertices r and s with probability proportional to their degrees.

Note that for the edge-step, r and s could be the same vertex. Thus loops could be created. However, as the graph gets large, the probability of adding a loop can be well bounded and is quite small.

The random graph model $G(p, G_0)$ is defined as follows:

Begin with the initial graph G_0.
For $t > 0$, at time t, the graph G_t is formed by modifying G_{t-1} as follows :
with probability p, take a vertex-step,
otherwise, take an edge-step.

When G_0 is the graph consisting of a single loop, we will simplify the notation and write $G(p) = G(p, G_0)$.

There were quite a number of papers analyzing the preferential attachment model $G(p)$, usually having similar conclusions of power law degree distribution. However, many of these analyses are heuristics without specifying the ranges for the power law to hold. Heuristics often run into the danger of incorrect deductions and incomplete conclusions. It is quite essential to use rigorous proofs which help specify the appropriate conditions and ranges for the power law. The following statement was proved in [16].

For the preferential attachment model $G(p)$, almost surely the number of vertices with degree k at time t is

$$M_k t + O(2\sqrt{k^3 t \ln(t)}).$$

where $M_1 = \frac{2p}{4-p}$ and $M_k = \frac{2p}{4-p} \frac{\Gamma(k)\Gamma(1+\frac{2}{2-p})}{\Gamma(k+1+\frac{2}{2-p})} = O(k^{-(2+\frac{p}{2-p})})$, for $k \geq 2$. In other words, almost surely the graphs generated by $G(p)$ have the power law degree distribution with the exponent $\beta = 2 + \frac{p}{2-p}$.

6.2 Duplication models

Networks of interactions are present in all biological systems. The interactions among species in ecosystems, between cells in an organism and among molecules in a cell all lead to complex biological networks. Using current technological advances, extensive data of such interactions has been acquired. To find the underlying structure in these databases, it is of great importance to understand the basic principles of various genetic and metabolic networks.

It has been observed that many biological networks have power law graphs with exponents β less than 2. The ranges for the exponents of the power law for biological networks are quite different from the ranges for nonbiological networks. Various examples, such as the WWW-graphs, call graphs, and various social networks, among others, are power law graphs with the exponent β between 2 and 3. Table 3 lists the exponents of a variety of biological and nonbiological networks with associated references. As we saw in Section 6.1, the preferential attachment model generates graphs with power law degree distribution with exponents β between 2 and 3. Therefore there is a need to consider alternative models for biological networks.

Biological networks	exponent β	references
Yeast protein-protein net	1.6,1.7	[20, 43]
E. Coli metabolic net	1.7, 2.2	[3, 28]
Yeast gene expression net	1.4–1.7	[20]
Gene functional interaction	1.6	[30]
Nonbiological networks		
Internet graph	2.2 (indegree), 2.6 (outdegree)	[4, 27, 34]
Phone call graph	2.1–2.3	[1, 2]
Collaboration graph	2.4	[29]
Hollywood graph	2.3	[4]

Table 3: Power law exponents for biological and nonbiological networks.

The duplication of the information in the genome — genes and their controlling elements — is a driving force in evolution and a determinative factor of biological networks. The process of duplication is quite different from the preferential attachment process that is regarded by many as the basic growth rule for most nonbiological networks.

Here we consider a duplication model. If we only allow pure duplication, the resulting graph depends heavily on the initial graph and does not satisfy the power law. So we consider a duplication model that allows randomness within the duplication step as defined below. We will see that this duplication model generates power law graphs with exponents in the range including the interval between 1 and 2 and therefore is more suitable for modeling complex biological

networks.

There are two basic parameters for the duplication model:

- A selection probability p, where $0 \leq p \leq 1$.
- An initial graph G_0, that we have at time 0.

Usually, G_0 is taken to be the graph formed by one vertex. However, G_0 can be taken to be any finite simple connected graph. Unlike the preferential attachment model, in this model the generated random graph is always a simple graph.

There is one basic operation:

Duplication step: A *sample* vertex u is selected randomly and uniformly from the current graph. A new vertex v and edge $\{u, v\}$ is added to the graph. For each neighbor w of u, with probability p, $\{v, w\}$ is added as a new edge.

The edge $\{u, v\}$ in the duplication step is called a *basic* edge. The vertex u is said to be the *parent* of v and v is called a *child* of u. We note that a vertex can have several children or no child at all and that each vertex not in the initial graph G_0 has a parent. All basic edges from children to parents form a forest where the vertices in G_0 are roots of component trees. All terms like "leafs" and "descendants", if not defined, refer to this forest.

The duplication step can be further decomposed into two parts — vertex-duplication and edge-duplication as follows:

Vertex-duplication: At time t, randomly select a sample vertex u and add a new vertex v and an edge $\{u, v\}$.

Edge-duplication: At time t, for the vertex v created, its parent u and each neighbor w of u, with probability p, add an edge $\{v, w\}$ to w.

For any vertex v, a *descendant* of v can only be connected to descendants of v's neighbors (including v itself). An edge $\{x, y\}$ is said to be a descendant of an edge $\{u, v\}$, if "x is a descendant of u and y is a descendant of v" or "x is a descendant of v and y is a descendant of u".

We remark that having the basic edges $\{u, v\}$ makes the graph G always connected. This helps avoid degenerate cases such as having mostly isolated vertices.

For the above duplication model, it can be shown [17] that its degree distribution obeys a power law with the exponent β of the power law satisfying the following equation:

$$1 + p = p\beta + p^{\beta-1}. \tag{5}$$

We remark that the solutions for (5) that are illustrated in figure 4 consist

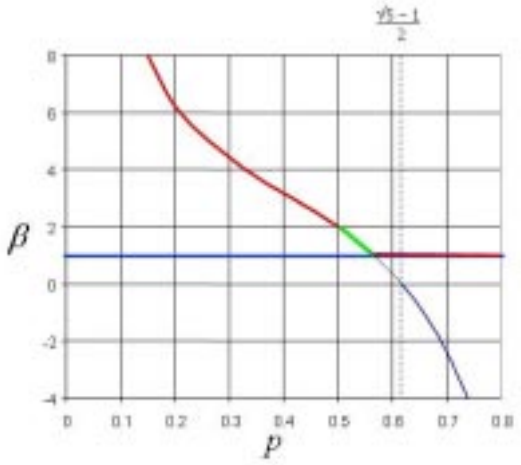

Figure 4: The value of β as a function of p.

of two parts. One is the line $\beta = 1$. The other is a curve which is a monotonically decreasing function of p. The two curves intersect at $(x, 1)$ where $x = 0.56714329\ldots$ the solution of $x = -\log x$. One very interesting range for β is when p is near $1/2$. To get a power-law with exponent 1.5, for example, one should choose $p = 0.535898\ldots$. Also we see that the second curve intersects zero at $p = (\sqrt{5} - 1)/2$, an intriguing number (the "golden mean"). At $p = 1/2$, one solution for β is 2. Although there are two solutions for each p, the stable solutions are on the curve when $p < 0.56714329\ldots$ and $\beta = 1$ for $p > 0.56714329\ldots$.

7 Remarks

The small world phenomenon, that occurs ubiquitously in numerous existing networks, refers to two similar but different properties:

Small distance — Between any pair of nodes, there is a short path.

The clustering effect — Two nodes are more likely to be adjacent if they share a common neighbor.

There have been various approaches to model networks that exhibit the small world phenomenon. In particular, the aspect of small distances can be well explained by using random graphs with general degree distributions which include the power law distribution. However, the other feature concerning the clustering effect seems much harder to model.

To model the clustering effect, a typical approach is to add random edges to a grid graph or the like ([26, 32, 44]). Such grid-based models are quite restrictive and far from satisfactory for modeling biological networks or collaboration graphs, for example. On the other hand, random power law graphs are good for modeling small distance, but fail miserably for modeling the clustering effect. In a way, the aspect of small distances is about neighborhood expansion while the aspect of the clustering effect is about neighborhood density. The related graph-theoretical parameters seem to be of an entirely different scale. For example, while the clustering effect is quite sensitive to average degree, the small distance effect is not.

The heart of the problem can be quite simply stated: For a given network, what is its true geometry? How can we capture the geometry of the network (without invoking too many parameters)?

References

[1] J. Abello, A. Buchsbaum, and J. Westbrook, A functional approach to external graph algorithms, *Proc. 6th European Symposium on Algorithms*, Springer, Berlin, 1998, 332–343.

[2] W. Aiello, F. Chung and L. Lu, A random graph model for massive graphs, *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing*, ACM Press, New York, 2000, 171–180.

[3] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, *Review of Modern Physics* **74** (2002), 47–97.

[4] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, *Science* **286** (1999), 509–512.

[5] A.-L. Barabási, R. Albert, and H. Jeong, Scale-free characteristics of random networks: the topology of the world-wide web, *Physica A* **281** (2000), 69–77.

[6] E. A. Bender and E. R. Canfield, The asymptotic number of labeled graphs with given degree sequences. *J. Combinatorial Theory* Ser. A **24** (1978), 296–307.

[7] B. Bollobás, The evolution of random graphs, *Trans. Amer. Math. Soc.* **286** (1984), 257–274.

[8] B. Bollobás, The diameter of random graphs, *Trans. Amer. Math. Soc.* **267** (1981), 41–52.

[9] B. Bollobás, *Random Graphs*, Academic Press, New York, 1985, xvi+447pp.

[10] B. Bollobás, The evolution of sparse graphs, *Graph Theory and Combinatorics* (Cambridge 1983), Academic Press, London-New York, 1984, 35–57.

[11] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins and J. Wiener, Graph Structure in the Web, *Proceedings of the WWW9 Conference*, May, 2000, Amsterdam. Paper version appeared in *Computer Networks* **33** (2000), 309–320.

[12] F. Chung and L. Lu, The diameter of sparse random graphs, *Advances in Applied Math.* **26** (2001), 257–279.

[13] F. Chung and L. Lu, Connected components in random graphs with given expected degree sequences, *Annals of Combinatorics* **6** (2002), 125–145.

[14] F. Chung and L. Lu, The average distances in random graphs with given expected degrees, *Proceeding of National Academy of Science* **99** (2002), 15879–15882.

[15] F. Chung and L. Lu, The volume of the giant component of a random graph with given expected degrees, *SIAM J. Discrete Math.*, **20** (2006), 395–411.

[16] F. Chung and L. Lu, *Complex Graphs and Networks*, *CBMS* Lecture Series, No. 107, AMS Publications, 2006, vii + 264pp.

[17] F. Chung, L. Lu, G. Dewey and D. J. Galas, Duplication models for biological networks, *J. Computational Biology* **10** no. 5 (2003), 677–687.

[18] F. Chung, L. Lu and V. Vu, The spectra of random graphs with given expected degrees, *Proceedings of National Academy of Sciences* **100** no. 11 (2003), 6313–6318.

[19] D. Conlon, A new upper bound for diagonal Ramsey numbers, *Annals of Mathematics*, to appear.

[20] S. N. Dorogovtsev and J. F. F. Mendes, Scaling properties of scale-free evolving networks: Continuous approach, *Phys. Rev. E* **63** 056125 (2001), 19 pp.

[21] P. Erdős, Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.* **53** (1947), 292–294.

[22] P. Erdős and A. Rényi, On random graphs, I. *Publ. Math. Debrecen* **6** (1959), 290–297.

[23] P. Erdős and A. Rényi, On the evolution of random graphs, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** (1960), 17–61.

[24] P. Erdős and T. Gallai, Gráfok előírt fokú pontokkal (Graphs with points of prescribed degrees, in Hungarian), *Mat. Lapok* **11** (1961), 264–274.

[25] G. Exoo, A lower bound for $R(5, 5)$, *Journal of Graph Theory*, **13** (1989), 97–98.

[26] A. Fabrikant, E. Koutsoupias and C. H. Papadimitriou, Heuristically optimized trade-offs: a new paradigm for power laws in the internet, *Automata, languages and programming*, *Lecture Notes in Computer Science* **2380**, Springer, Berlin, 2002, 110–122.

[27] M. Faloutsos, P. Faloutsos, and C. Faloutsos, On power-law relationships of the Internet topology, *Proceedings of the ACM SIGCOM Conference*, ACM Press, New York, 1999, 251–262.

[28] R. Friedman and A. Hughes, Gene duplications and the structure of eukaryotic genomes, *Genome Res.* **11** (2001), 373–381.

[29] J. Grossman, P. Ion, and R. De Castro, The Erdős Number Project. `http://www.oakland.edu/enp`

[30] Z. Gu, A. Cavalcanti, F.-C. Chen, P. Bouman and W.-H. Li, Extent of gene duplication in the genomes of drosophila, nematode, and yeast, *Mol. Biol. Evol.* **19** (2002), 256–262.

[31] S. Janson, D. E. Knuth, T. Łuczak and B. Pittel, The birth of the giant component, *Random Structures & Algorithms* **4** (1993), 231–358.

[32] J. Kleinberg, The small-world phenomenon: An algorithmic perspective, *Proc. 32nd ACM Symposium on Theory of Computing*, (2000), 163-170.

[33] J. Kleinberg, R. Kumar, P. Raphavan, S. Rajagopalan and A. Tomkins, The web as a graph: Measurements, models and methods, *Proceedings of the International Conference on Combinatorics and Computing*, *Lecture Notes in Computer Science* **1627**, Springer, Berlin, 1999, 1–17.

[34] R. Kumar, P. Raghavan, S. Rajagopalan and A. Tomkins, Trawling the web for emerging cyber communities, *Proceedings of the 8th World Wide Web Conference*, Toronto, 1999.

[35] T. Łuczak, Random trees and random graphs, *Random Structures & Algorithms* **13** (1998), 485–500.

[36] B. D. McKay and S. P. Radziszowski, Subgraph counting identities and Ramsey numbers, *Journal of Combinatorial Theory* (B) **61** (1994), 125–132.

[37] S. Milgram, The small world problem, *Psychology Today* **2** (1967), 60–67.

[38] M. Molloy and B. Reed, A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms* **6** (1995), 161–179.

[39] M. Molloy and B. Reed, The size of the giant component of a random graph with a given degree sequence, *Combin. Probab. Comput.* **7** (1998), 295–305.

[40] F. P. Ramsey, On a problem of formal logic, *Proc. London Math. Soc.* **30** (1930), 264–286.

[41] H. A. Simon, On a class of skew distribution functions, *Biometrika* **42** (1955), 425–440.

[42] J. Spencer, Ramsey's theorem — a new lower bound, *J. Comb. Theory* (A) **18** (1975), 108–115.

[43] L. Stubbs, Genome comparison techniques, *Genomic Technologies: Present and Future*, (eds. D. Galas and S. McCormack), Caister Academic Press, 2002.

[44] D. J. Watts and S. H. Strogatz, Collective dynamics of 'small world' networks, *Nature* **393** (1998), 440–442.