

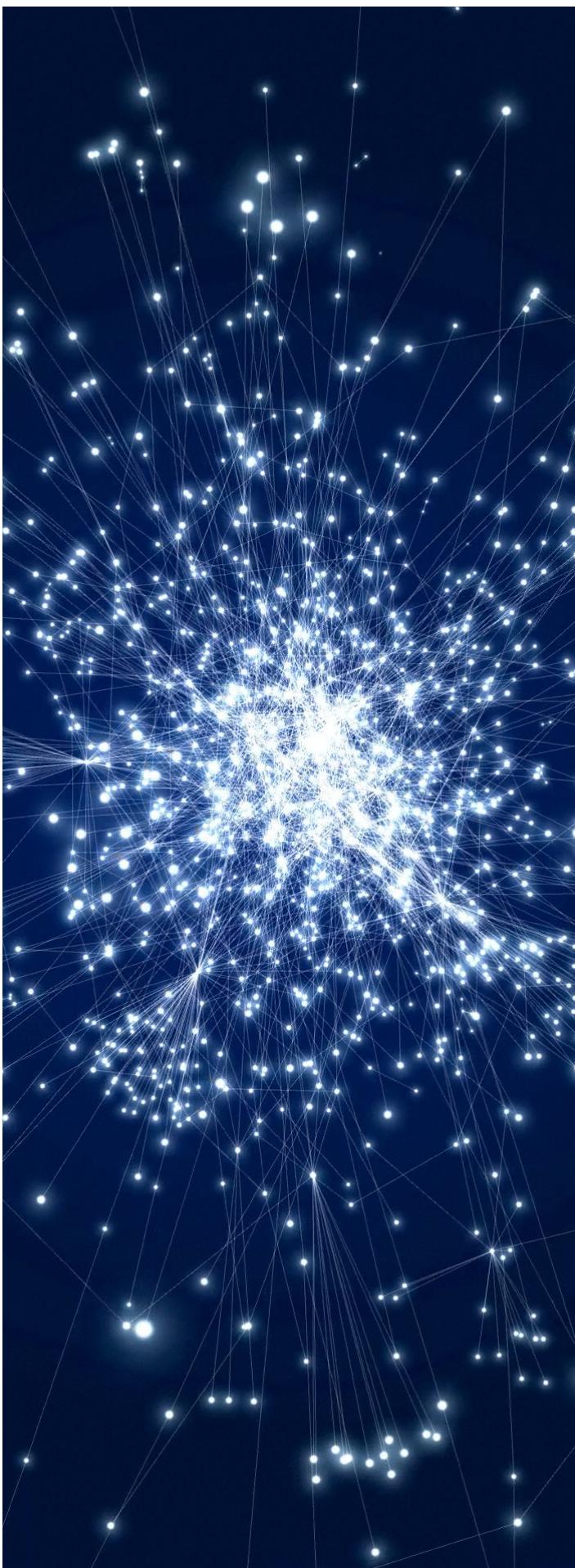
Network Science

Albert-László Barabási

Data Visualization by Mauro Martino
Data Analysis by Márton Pósfai

PDF VERSION:
NOVEMBER 2012

RANDOM NETWORKS CHAPTER 3



INTRODUCTION

THE RANDOM NETWORK MODEL

THE NUMBER OF LINKS IS VARIABLE

DEGREE DISTRIBUTION

REAL NETWORKS DO NOT HAVE A
POISSON DEGREE DISTRIBUTION

THE EVOLUTION OF A RANDOM
NETWORK

REAL NETWORKS ARE SUPERCRITICAL

THE SMALL WORLD PROPERTY

CLUSTERING COEFFICIENT

REAL NETWORKS ARE NOT RANDOM

ADVANCED TOPICS 3.A:
DERIVING THE POISSON DEGREE DIS-
TRIBUTION

ADVANCED TOPICS 3.B:
THE MAXIMUM AND THE MINIMUM
DEGREE

ADVANCED TOPICS 3.C:
GIANT COMPONENT

ADVANCED TOPICS 3.D:
COMPONENT SIZES

ADVANCED TOPICS 3.E:
SUPERCRITICAL REGIME

ADVANCED TOPICS 3.F:
PHASE TRANSITIONS

ADVANCED TOPICS 3.G:
CORRECTION TO SMALL WORLDS

BIBLIOGRAPHY

SECTION 1

INTRODUCTION

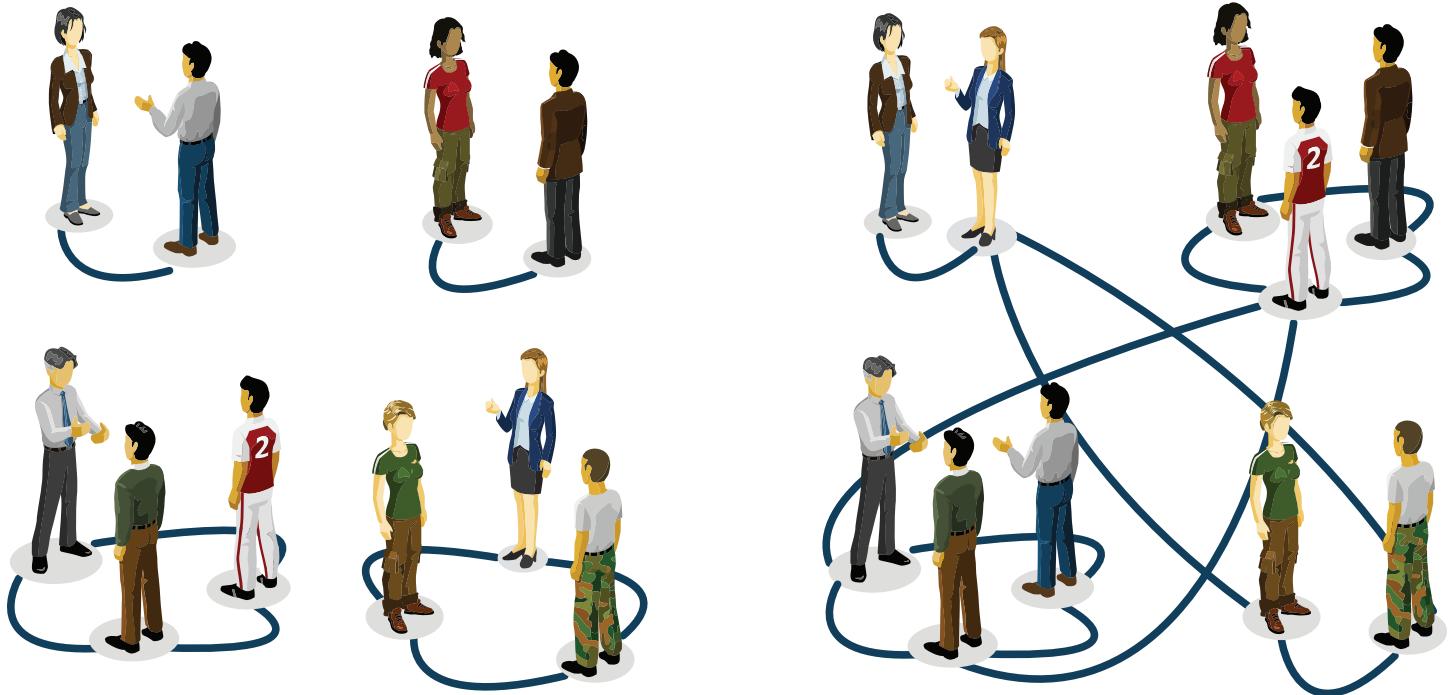


Image 3.1

From a cocktail party to random networks.

The emergence of an acquaintance network through random encounters at a cocktail party.

Imagine organizing a party for a hundred guests who initially do not know each other [1]. Offer them wine and cheese and you will soon have dozens of chatting groups of two to three. Now mention to Mary, one of your guests, that the red wine in the unlabeled dark green bottles is a rare vintage, much better than the one with the fancy red label. If she shares this information only with her acquaintances, you know that your expensive wine is safe, because she only had time to meet a few others in the room. However, the guests will continue to mingle, creating subtle paths between individuals that may still be strangers to each other. For example, while John has not yet met Mary, they have both met Mike, so now there is an invisible path from John to Mary through Mike. As time goes on, the guests will be increasingly interwoven by such intangible links. With that the secret of the unlabeled bottle will be passed from Mary to Mike and from Mike to John, slowly escaping into a rapidly expanding group.

To be sure, when all guests had gotten to know each other, everyone would be pouring the superior wine. But if each encounter took only ten minutes, meeting all ninety-nine others would take about sixteen hours. Thus, you could reasonably hope that a few drops of the better wine would be left for you to enjoy once the party is over.

Yet, you will be wrong. The purpose of this chapter is to show you why. We will see that the party maps into a classic model in network science called the random network model. And random network theory tells us that we do not have to wait until all individuals get to know each other for our expensive wine to be in danger. Rather, soon after each person meets at least one other guest, an invisible network will form that will allow the information to reach most guests. Hence in no time everyone will be drinking the better wine.

THE RANDOM NETWORK MODEL

An important goal of network science is to build models that accurately reproduce the properties of real networks observed in real systems. Most networks we encounter in nature do not have the comforting regularity of a crystal lattice or the predictable radial architecture of a spider web. Rather, at first inspection most real networks look as if they were spun randomly. Random network theory embraces this apparent randomness by constructing networks that are *truly random*.

From a modeling perspective a network is a relatively simple object, consisting of only nodes and links. The real challenge, however, is to place the links between the nodes in a way to reproduce the complexity and apparent randomness of real systems. In this context the philosophy behind a random network is simple: it assumes that this goal is best achieved by placing the links randomly between the nodes. With that we arrive to the definition of a random network:

A random network consists of N labeled nodes where each node pair is connected with the same probability p .

Two definitions of random networks.

There are two equivalent ways of defining a random network:

- **$G(N,L)$ model:** N labeled nodes are connected with L randomly placed links. Erdős and Rényi (Erdős & Rényi, 1959) used this definition in their string of articles on random networks.
- **$G(N,p)$ model:** Each pair of N labeled nodes is connected with probability p , a model introduced by Gilbert (Gilbert, 1959).

Hence the $G(N,p)$ model fixes the probability p that two nodes are connected and the $G(N,L)$ model fixes the total number of links L . While in the $G(N,L)$ model the average degree of a node is simply $\langle k \rangle = 2L/N$, other network characteristics are easier to calculate in the $G(N,p)$ model. Throughout this book we will explore the $G(N,p)$ model, not only for the ease that it allows us to calculate key network characteristics, but also because its construction is closer to the way real systems develop. Indeed, in real networks the number of links is rarely fixed, but we can instead determine the probability that two nodes link to each other.

To construct a random network, denoted with $G(N, p)$ ([Box 3.1](#)):

1. Start with N isolated nodes.
2. Select a node pair, and generate a random number between 0 and 1. If the random number exceeds p , connect the selected node pair with a link, otherwise leave them disconnected.
3. Repeat step (2) for each of the $N(N-1)/2$ node pairs.

The network obtained through this procedure is called a random graph or a random network. Two mathematicians, Pál Erdős and Alfréd Rényi, have played an important role in understanding the properties of random networks. In their honor a random network is often called the Erdős-Rényi network ([Box 3.2](#)).

A brief history of random networks.

Anatol Rapoport (1911–2007), a Russian immigrant to the United States, was the first to explore the properties of a random network. Trained as a pianist, Rapoport's interests turned to mathematics after realizing that a successful career as a concert pianist would require a wealthy patron. He became interested in mathematical biology at a time when mathematicians and biologists hardly spoke to each other. In a paper written with Ray Solomonoff in 1951 [28], Rapoport demonstrated that if we increase the average degree of a network, we will observe an abrupt transition from a collection of disconnected nodes to a state in which the graph contains a giant component. Despite its pioneering ideas, Rapoport's paper remains relatively unknown.

The study of random networks reached prominence thanks to the fundamental work of Pál Erdős and Alfréd Rényi. In a sequence of eight papers published between 1959 and 1968 [8–15], they merged probability theory and combinatorics with graph theory, establishing random graph theory, a new branch of mathematics [5].

The random network model was independently introduced by Gilbert [18] the same year Erdős and Rényi published their first paper on the subject. Yet, the impact of Erdős and Rényi's work is so overwhelming that they are rightly considered the fathers of random networks.

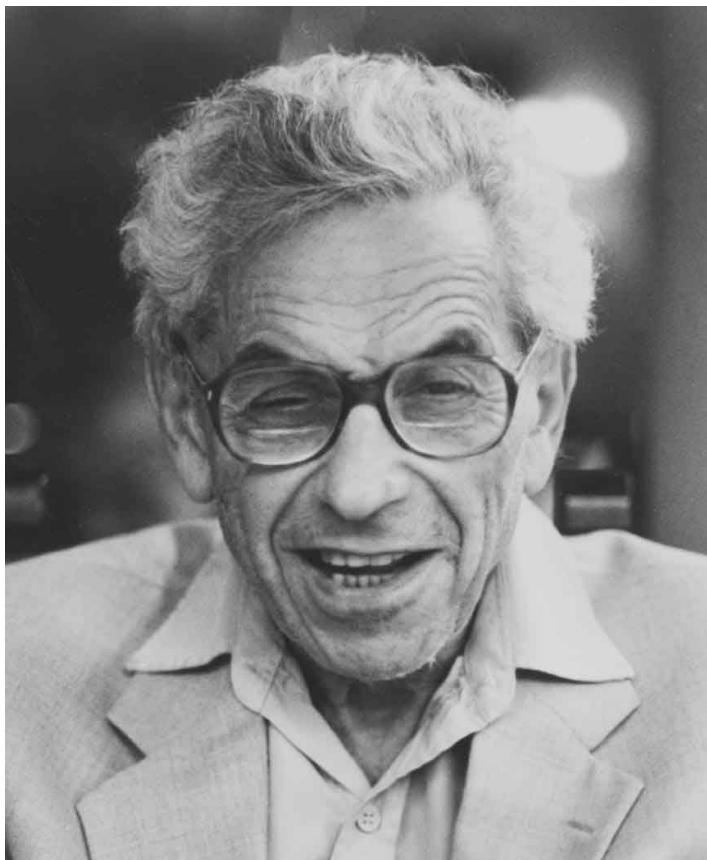


Image 3.2a
Pál Erdős (1913–1996)

Hungarian mathematician known for both his eccentricity and exceptional scientific output, having published more papers than any other mathematician in the history of mathematics. His productivity had its roots in his fondness for collaboration: he co-authored papers with over five hundred mathematicians, inspiring the concept of Erdős number. His legendarily personality and profound professional impact has inspired two biographies [19, 27] and a documentary [7].



Image 3.2b
Alfréd Rényi (1921–1970)

Hungarian mathematician with fundamental contributions to combinatorics, graph theory, and number theory. His impact goes beyond mathematics: the Rényi entropy is widely used in chaos theory and the random network model he co-developed is at the heart of network science. He is remembered through the hotbed of Hungarian mathematics, the Alfréd Rényi Institute of Mathematics in Budapest. He once said, "*A mathematician is a device for turning coffee into theorems*", a quote often attributed to Erdős.

THE NUMBER OF LINKS IS VARIABLE

Each random network we generate with the same parameters N, p will look slightly different ([Image 3.3](#)). Not only the detailed wiring diagram will vary between realizations, but so will the number of links L . It is useful, therefore, to determine how many links we expect for a particular realization of a random network with fixed N and p .

The probability that a random network has exactly L links is the product of three terms:

1. The probability that L of the attempts to connect the $N(N-1)/2$ pairs of nodes have resulted in a link, which is p^L .
2. The probability that the remaining $N(N-1)/2 - L$ attempts have not resulted in a link, which is $(1-p)^{N(N-1)/2-L}$

3. A combinational factor, $\binom{N}{2} \binom{N(N-1)}{L}$ counting the number of different ways we can place L links among $N(N-1)/2$ node pairs.

Hence the probability that a particular realization of a random graph has exactly L links is

$$p_L = \binom{N}{2} \binom{N(N-1)}{L} p^L (1-p)^{\frac{N(N-1)}{2}-L}. \quad (1)$$

As Eq. (1) is a binomial distribution ([Box 3.3](#)), the expected number of links in a random graph can be calculated as

$$\langle L \rangle = \sum_{L=0}^{N(N-1)} L p_L = p \frac{N(N-1)}{2}. \quad (2)$$

Hence $\langle L \rangle$ is the product of the probability p that two nodes are connected and the number of pairs we attempt to connect, which is $L_{\max} = N(N-1)/2$ ([Chapter 2](#)).

Using Eq. (2) we obtain the average degree of a random network as

$$\langle k \rangle = \frac{2\langle L \rangle}{N} = p(N-1). \quad (3)$$

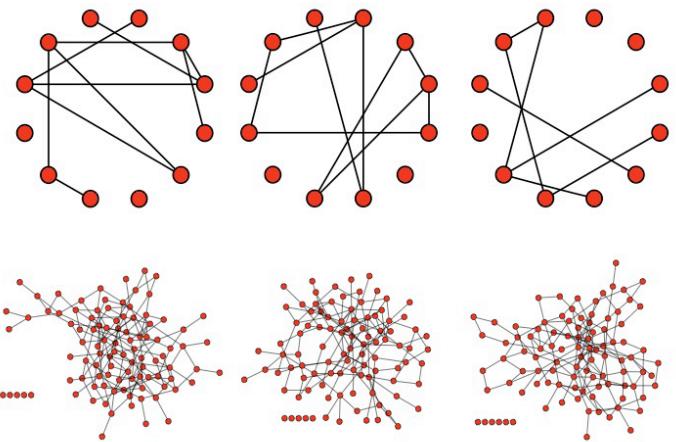


Image 3.3
Random networks are truly random.

Top row: Three realizations of a random network generated with the same parameters $N = 12$ and $p = 1/6$. Despite the identical parameters, the networks not only look different, but they differ in the number of links they have ($L = 8, 10, 7$) and in the degree of the individual nodes.

Bottom row: Three realizations of a random network with $N = 100$ and $p = 1/6$.

Hence $\langle k \rangle$ is the product of the probability p that two nodes are connected and $(N-1)$, representing the maximum number of links a node can have in a network of size N .

In summary the number of links in a random network is not fixed, but varies between realizations. Its expected value is determined by N and p . If we increase p from $p = 0$ to $p = 1$ the random network becomes denser and the average number of links increase linearly from $\langle L \rangle = 0$ to L_{\max} and the average degree of a node increases from $\langle k \rangle = 0$ to $\langle k \rangle = N-1$.

Binomial distribution: Mean and variance.

If we toss a fair coin N times, tails and heads should occur with the same probability $p = 1/2$. The binomial distribution provides the probability p_x that we obtain exactly x heads in a sequence of N throws. In general, the binomial distribution describes the number of successes in N independent experiments with two possible outcomes, in which the probability of one outcome is p , and of the other is $1-p$.

The binomial distribution has the form

$$p_x = \binom{N}{x} p^x (1-p)^{N-x}.$$

The mean of the distribution (first moment) is

$$\langle x \rangle = \sum_{x=0}^N x p_x = Np. \quad (4)$$

Its second moment is

$$\langle x^2 \rangle = \sum_{x=0}^N x^2 p_x = p(1-p)N + p^2 N^2, \quad (5)$$

providing its standard deviation as

$$\sigma_x = \left(\langle x^2 \rangle - \langle x \rangle^2 \right)^{\frac{1}{2}} = [p(1-p)N]^{\frac{1}{2}}. \quad (6)$$

Box 3.3

DEGREE DISTRIBUTION

As [Image 3.3](#) illustrates, in a given realization of a random network some nodes are lucky, gaining numerous links, while others have only a few or no links. These differences are captured by the degree distribution p_k providing the probability that a randomly chosen node has degree k .

In a random network the probability that node i has exactly k links is the product of three terms [5]:

- The probability that k of its links are present, or p^k .
- The probability that the remaining $(N - 1 - k)$ links are missing, or $(1-p)^{N-1-k}$.
- The number of ways we can select k links from $N - 1$ potential links a node can have, or $\binom{N-1}{k}$.

Hence the degree distribution of a random network follows the binomial distribution

$$p_k = \binom{N-1}{k} p^k (1-p)^{N-1-k}. \quad (7)$$

The shape of this distribution depends on the system size N and the probability p ([Image 3.4](#)). Using the properties of the binomial distribution ([Box 3.3](#)), from the degree distribution (7) we can calculate the network's average degree $\langle k \rangle$, recovering Eq. (3). We can also determine the second moment $\langle k^2 \rangle$ and the variance σ_k of the degree distribution ([Image 3.4](#)), quantities that will play an important role later.

Most real networks are sparse, hence $\langle k \rangle \ll N$ ([Table 3.1](#), [Image 3.4b](#)). In this limit the degree distribution (7) is well approximated by the Poisson distribution ([Advanced Topics 3. A](#))

$$p_k = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}, \quad (8)$$

which is often called, together with (7), the degree distribution of a random network.

The binomial and the Poisson distribution describe the

same quantity, hence they have several common properties ([Image 3.4a](#)):

- Both distributions have a peak around $\langle k \rangle$. If we keep N constant and increase p , the network becomes denser, increasing $\langle k \rangle$ and moving the peak to the right.
- The width of the distribution (dispersion) is also controlled by p or $\langle k \rangle$. The denser the network, the wider is the distribution, hence the larger are the differences in the degrees.

As we use the Poisson form in Eq. (8), we need to keep in mind that:

- The exact result for the degree distribution is the binomial form in Eq. (7), thus Eq. (8) represents only an approximation to (7) valid in the $\langle k \rangle \ll N$ limit. For most networks of practical importance this condition is easily satisfied.
- The advantage of the Poisson form is that key network characteristics, like $\langle k \rangle$, $\langle k^2 \rangle$ and σ_k , have a much simpler form ([Image 3.4a](#)), depending on a single parameter, $\langle k \rangle$.
- The Poisson distribution in Eq. (8) does not explicitly depend on the number of nodes N . Therefore, Eq. (8) predicts that the degree distributions of networks of different sizes but the same average degree $\langle k \rangle$ are indistinguishable from each other ([Image 3.4b](#)).

Despite the fact that the Poisson distribution is only an approximation to the degree distribution of a random network, thanks to its analytical simplicity, it is the preferred form for p_k . Hence throughout this book, unless noted otherwise, we will refer to the Poisson form in Eq. (8) as the degree distribution of a random network.

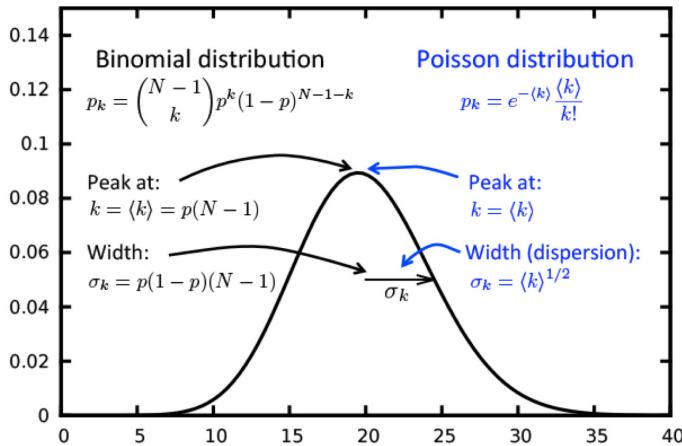


Image 3.4a

Anatomy of a binomial and a Poisson degree distribution.

The exact form of the degree distribution of a random network is the binomial distribution (left). For $N \gg \langle k \rangle$, the binomial can be well approximated by a Poisson distribution (right). As both distributions describe the same quantity, they have the same properties, which are expressed in terms of different parameters: the binomial distribution uses p and N as its fundamental parameters, while the Poisson distribution has only one parameter, $\langle k \rangle$.

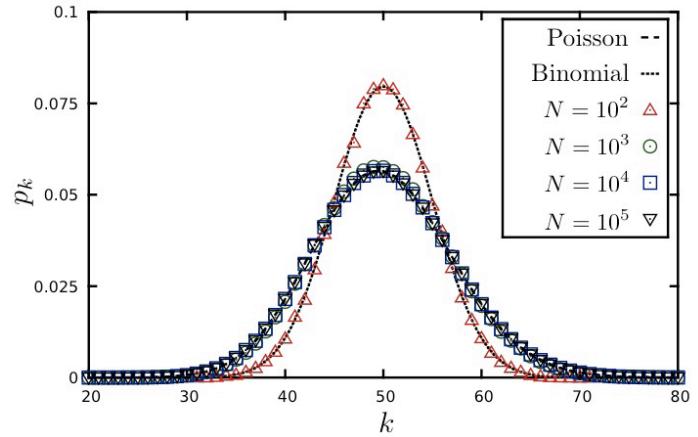


Image 3.4b

Degree distribution is independent of the network size.

The degree distribution of a random network with average degree $\langle k \rangle = 50$ and sizes $N = 10^2, 10^3, 10^4$. For $N = 10^2$ the degree distribution deviates significantly from the Poisson prediction (8), as the condition for the Poisson approximation, $N \gg \langle k \rangle$, is not satisfied. Hence for small networks one needs to use the exact binomial form of Eq. (7) (dotted line). For $N = 10^3$ and larger networks the degree distribution becomes indistinguishable from the Poisson prediction, (8), shown as a continuous line, illustrating that for large N the degree distribution is independent of the network size. In the figure we averaged over 1,000 independently generated random networks to decrease the noise in the degree distribution.

REAL NETWORKS DO NOT HAVE A POISSON DEGREE DISTRIBUTION

The degree of a node in a random network can vary between 0 and $N-1$, raising an important question: How big are the differences between the node degrees in a particular realization of a random network? That is, can highly connected nodes, or hubs, coexist with small degree nodes? We address answer these questions by estimating the size of the largest and the smallest node in a random network.

Let us assume that the world's social network is described by the random network model. This may not be as far fetched hypothesis as it first sounds: there is significant randomness in whom we meet and whom we choose to become acquainted with. Sociologists estimate that a typical person knows about 1,000 individuals on a first name basis, suggesting that $\langle k \rangle = 1,000$. Using the results obtained so far about random networks, we arrive to a number of surprising conclusions about a random society (see [Advanced Topics 3.B](#)):

- The most connected individual (the largest degree node) in a random society is expected to have degree $k_{\max} = 1,185$.
- The least connected individual is expected to have degree $k_{\min} = 816$.
- The dispersion of a random network is $\sigma_k = \langle k \rangle^{1/2}$, which for $\langle k \rangle = 1,000$ is $\sigma_k = 31.62$. This means that the number of friends of a typical individual should be mainly in the $\langle k \rangle \pm \sigma_k$ range, or between 970 and 1,030, a rather narrow range.

In other words, in a random society everyone would have a comparable number of friends. We would lack outliers, or highly popular individuals, and no one would be left behind, having only a few friends. This calculation illustrates that in a *large random network* the degree of most nodes is in the narrow vicinity of $\langle k \rangle$ ([Box 3.4](#)).

This prediction blatantly conflicts with reality. Indeed, there is extensive evidence of individuals who have considerably more than 1,018 acquaintances. For example, US president Franklin Delano Roosevelt's appointment book had about 22,000 names in it, individuals he met person-

Why hubs are absent in random network.

To understand why hubs are absent in random networks, we turn to the degree distribution (8). We first realize that the $1/k!$ term in (8) significantly decreases the chances of observing large degree nodes. Indeed, the Stirling approximation

$$k! \sim \left[\sqrt{2\pi k} \right] \left(\frac{k}{e} \right)^k$$

allows us rewrite Eq. (8) as

$$p_k = \frac{e^{-\langle k \rangle}}{\sqrt{2\pi \langle k \rangle}} \left(\frac{e \langle k \rangle}{\langle k \rangle} \right)^{\langle k \rangle}. \quad (9)$$

For degrees $k > e \langle k \rangle$ the term in the parenthesis is smaller than one, hence for large k both k -dependent terms in (9), i.e. $1/\sqrt{k}$ and $(e \langle k \rangle / k)^{\langle k \rangle}$ decrease rapidly with increasing k . Overall Eq. (9) predicts that in a random network *the chance of observing a hub decreases faster than exponentially*.

Box 3.4

ally [[17](#), [26](#)]. Similarly, a study of the social network behind Facebook has documented numerous individuals with 5,000 Facebook friends, the maximum allowed by the social networking platform [[4](#)]. The reason behind these systematic discrepancies can be understood by comparing the degree distribution of real and random networks.

In [Image 3.5](#) we show the degree distribution of three real networks, together with the corresponding Poisson fits. The figure documents considerable differences between the random network predictions and the real data:

- The Poisson form significantly underestimates the number of high degree nodes. For example, according to the random network model the maximum degree for the Internet is expected to be around 20, while the data indicates the existence of nodes with degrees close to 10^3 .

- The spread in the degrees of real networks is much wider than expected in a random network. This difference is captured by the dispersion σ_k (Image 3.4a). For example, if the Internet were to be random, we would expect $\sigma_k = 2.52$, while the measurements indicate $\sigma_{\text{internet}} = 14.14$, significantly higher than predicted.

These differences are not limited to the networks shown in Image 3.5, but all networks listed in Table 2.1 share this property. Hence the comparison with the real data indicates that the random network model does not capture the degree distribution of real networks. While in a random network most nodes have comparable degrees, forbidding hubs, in real networks we observe a significant number of highly connected nodes and large differences in node degrees. We will resolve these differences in Chapter 4.

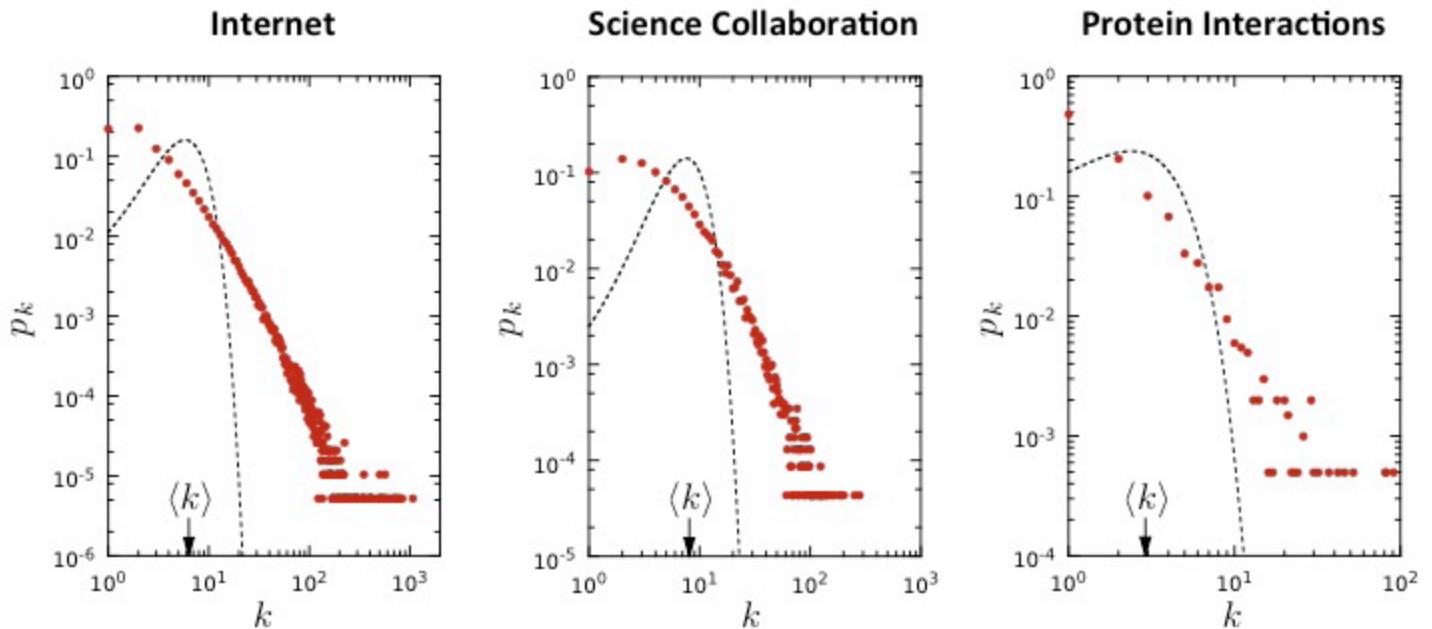
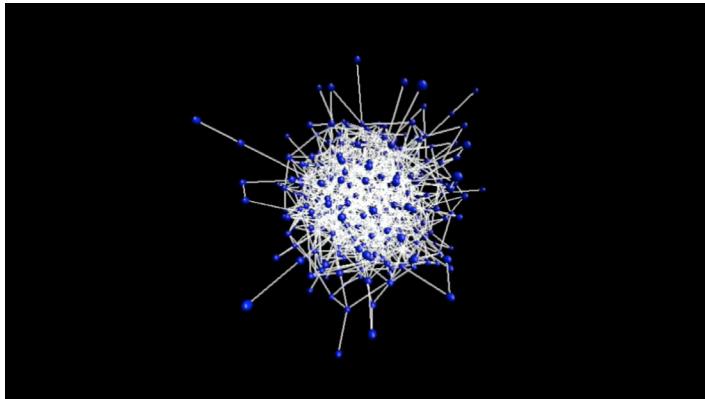


Image 3.5

Degree distribution of real networks.

The degree distribution of the Internet, science collaboration network, and the protein interaction network of yeast (Table 2.1). The dashed line corresponds to the Poisson prediction, obtained by measuring $\langle k \rangle$ for the real network and then plotting Eq. (8). The significant deviation between the data and the Poisson fit indicates that the random network model underestimates the size and the frequency of highly connected nodes, or hubs.

THE EVOLUTION OF A RANDOM NETWORK



Movie 3.1

Evolution of a random graph.

Changes in the structure of a random graph with increasing p , illustrating the absence of a giant component for small p and its sudden emergence once p exceeds a critical value.

The cocktail party we encountered at the beginning of the chapter captures a dynamical process: starting with N isolated nodes, the links are added gradually through random encounters between the guests. Within the random network model this corresponds to a gradual increase of p , with striking consequences on the network topology ([Movie 3.1](#)). To quantify this process, we first inspect how the size N_G of the *giant component*, which is the largest cluster within the network, varies with $\langle k \rangle$. The two extreme cases are easy to understand:

- For $p = 0$ we have $\langle k \rangle = 0$, hence we observe only isolated nodes. Therefore $N_G = 1$ and $N_G / N \rightarrow 0$ for large N .
- For $p = 1$ we have $\langle k \rangle = N-1$, hence the network is a complete graph and all nodes belong to a single cluster. Therefore $N_G = N$ and $N_G / N = 1$.

One would expect that the giant component will grow gradually from $N_G = 1$ to $N_G = N$ if we increase $\langle k \rangle$ from 0 to $N-1$. Yet, as [Image 3.6a](#) indicates, this is not the case: N_G / N remains zero for small $\langle k \rangle$, indicating the lack of a giant component for a range of $\langle k \rangle$ values. Once $\langle k \rangle$ ex-

ceeds a critical value, N_G / N increases rapidly, signaling the emergence of a giant component. Erdős and Rényi in their classical 1959 paper predicted that the *condition for the emergence of the giant component* is

$$\langle k \rangle = 1. \quad (10)$$

In other words, we have a giant component if and only if when *each node has on average one link* ([Advanced Topics 3.C](#)).

The fact that at least one link per node is *necessary* for a giant component is not unexpected. Indeed, for a giant component to exist, each of its nodes must be linked to at least one other node. It is somewhat counterintuitive, however that one link is *sufficient* for its emergence.

If we wish to express Eq. (10) in terms of p , using Eq. (3) we obtain

$$p_c = \frac{1}{N-1} \approx \frac{1}{N}, \quad (11)$$

indicating that the larger a network, the smaller p is sufficient for the giant component.

The emergence of the giant component is only one of the important transitions displayed by a random network. Changes in $\langle k \rangle$ allow us to distinguish four topologically distinct regimes ([Image 3.6](#)), each with its unique characteristics:

(a) Subcritical regime: $0 < \langle k \rangle < 1$, ($p < \frac{1}{N}$).

For $\langle k \rangle = 0$ the network consists of N isolated nodes. Increasing $\langle k \rangle$ is equivalent with adding $N\langle k \rangle = pN(N-1)/2$ links to the network. Yet, given the small number of links in the network in this regime, these links will mainly form clusters of size two ([Image 3.6b](#)). Upon increasing $\langle k \rangle$ further, some of the new links will join these pairs, forming tiny clusters. While we can designate at any moment the largest such cluster to be the giant component, in this regime the relative size of the largest cluster, N_G / N , remains zero. The reason is that for $\langle k \rangle < 1$ the largest cluster is a tree

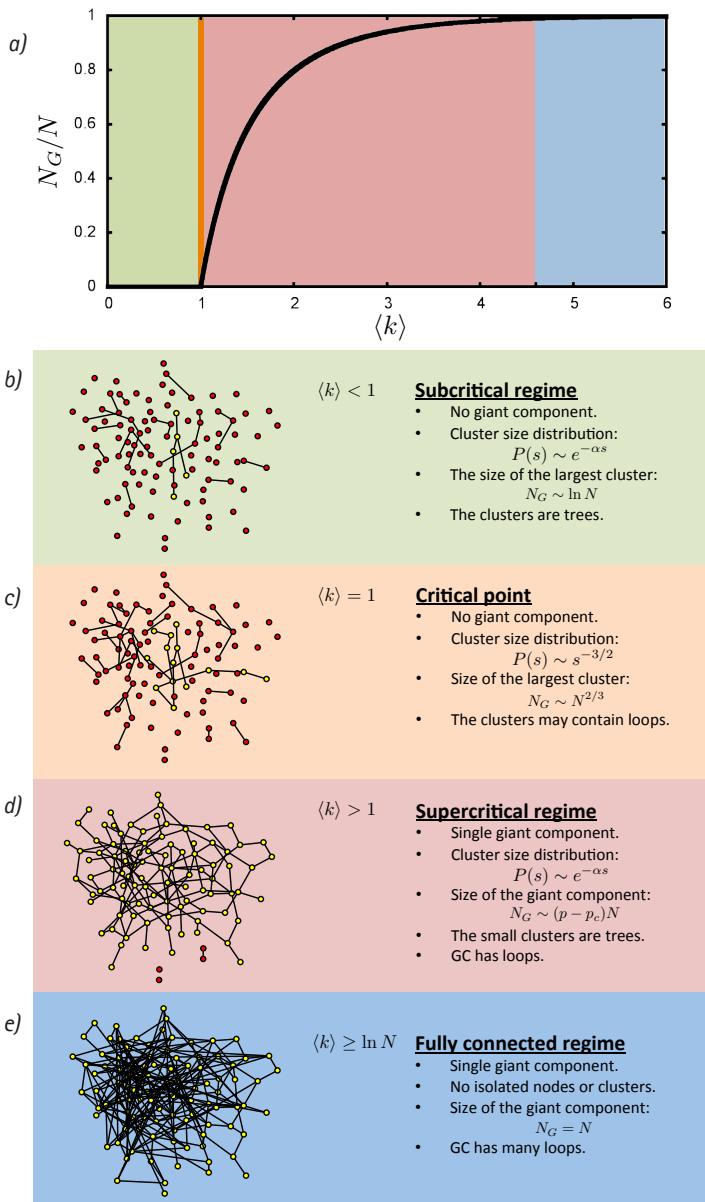


Image 3.6

Evolution of a random network.

- (a) The relative size of the giant component in function of the average degree $\langle k \rangle$ in the Erdős-Rényi model.
- (b)–(e) The main network characteristics in the four regimes that characterize a random network.

with size $N_G \sim \ln N$. Therefore $N_G/N \approx \ln N/N \rightarrow 0$ in the $N \rightarrow \infty$ limit, indicating that the largest component is tiny compared to the size of the network.

In summary, in the subcritical regime the network consists of numerous tiny components, whose size follows an exponential distribution. Hence these components have comparable sizes, lacking a clear winner that we could designate as a giant component ([Advanced Topics 3.D](#)).

(b) Critical Point: $\langle k \rangle = 1$, $(p = \frac{1}{N})$.

The critical point separates the regime where there is not yet a giant component ($\langle k \rangle < 1$) from the regime where there is one ($\langle k \rangle > 1$). While it signals the emergence of the giant component, the relative size of the largest component in this point is still zero ([Image 3.6c](#)). Indeed, the calculations indicate that the size of the largest component is $N_G \sim N^{2/3}$, so its relative size decreases as $N_G/N \sim N^{-1/3}$, indicating that N_G is still tiny compared to the network's size.

In absolute terms there is a significant increase in the size of the largest component at $\langle k \rangle = 1$. For example, for a random network of $N = 7 \times 10^9$ nodes, the size of the globe's social network, for $\langle k \rangle < 1$ the largest cluster is of the order of $N_G \approx \ln N = \ln(7 \times 10^9) \approx 22.7$. In contrast at $\langle k \rangle = 1$ we expect $N_G \sim N^{2/3} = (7 \times 10^9)^{2/3} \approx 3 \times 10^6$, a jump of about five orders of magnitude. Yet, both in the subcritical regime ($\langle k \rangle < 1$) and at the critical point ($\langle k \rangle = 1$) *the largest component contains a vanishing fraction of the total number of nodes in the network*.

Therefore most nodes are located in numerous small components, whose size distribution follows Eq. (36), a power law form indicating that components of rather different sizes coexist. These numerous small components are mainly trees, while the giant component may contain loops. Note that many properties of the network at the critical point resemble the properties of a physical system undergoing a phase transition ([Advanced Topics 3.F](#)).

(c) Supercritical regime: $\langle k \rangle > 1$, $(p > \frac{1}{N})$.

This regime has the most relevance to real systems, as for the first time we have a giant component that looks like a network. In the vicinity of the critical point the size of the giant component varies as

$$N_G/N \sim \langle k \rangle - 1, \quad (12)$$

or

$$N_G \sim (p - p_c)N, \quad (13)$$

where p_c is given by Eq. (11). In other words, the *giant component contains a finite fraction of all nodes in the network*. The further we move from the critical point, a larger fraction of nodes will belong to it. Note that Eq. (12) is valid only in the vicinity of $\langle k \rangle = 1$, and for large $\langle k \rangle$ the dependence between N_G and $\langle k \rangle$ is nonlinear ([Image 3.6d](#)).

In the supercritical regime there are still numerous isolated components that coexist with the giant component, their size distribution being given by Eq. (35). These small

components are trees, while the giant component contains numerous loops and cycles. The supercritical regime lasts until all nodes are absorbed by the giant component.

(d) Connected regime: $\langle k \rangle \geq \ln N$, ($p \geq \frac{\ln N}{N}$).

For sufficiently large p the giant component will absorb all nodes and components, hence $N_G \approx N$. In the absence of isolated nodes the network becomes connected. The average degree at which this happens depends on N as (Advanced Topic 3.E)

$$\langle k \rangle \sim \ln N. \quad (14)$$

Note that when we enter the connected regime the network is still relatively sparse, as $\ln N / N \rightarrow 0$ for large N . The network turns into a complete graph only at $\langle k \rangle = N - 1$.

In summary, the emergence of a network within the random network model is not a smooth process: the isolated nodes and tiny components observed for small $\langle k \rangle$ organize themselves into a giant component rather suddenly, through a process called phase transition (Advanced Topics 3.F). Along the way we encounter four topologically distinct regimes (Image 3.6). The discussion offered above follows an empirical perspective, fruitful if we wish to compare the observed networks to real systems. A different perspective, leading to its own rich behavior, is discussed in the mathematical literature (Box 3.5).

Network evolution in graph theory.

In the random graph literature it is often assumed that the connection probability $p(N)$ scales as N^z , where z is a tunable parameter between $-\infty$ and 0. The greatest discovery of Erdős and Rényi was that as we vary z , key properties of random graphs appear quite suddenly. To be precise, a graph has a given property Q if the probability of having Q approaches 1 as $N \rightarrow \infty$. That is, for a given probability either almost every graph has the property Q or, almost no graph has it. For example, for z less than $-3/2$ almost all graphs contain only isolated nodes and edges.

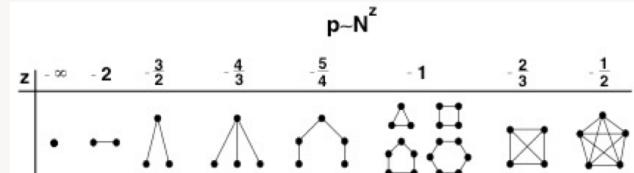


Image 3.7

Evolution of random graph.

The threshold probabilities at which different subgraphs appear in a random graph, as defined by exponent z in the $p(N) \sim N^z$ relationship. For $z < -3/2$ the graph consists of isolated nodes and edges. When z passes $-3/2$ trees of order 3 appear, while at $z = -4/3$ trees of order 4 appear. At $z = -1$ trees of all orders are present, together with cycles of all orders. Complete subgraphs of order 4 appear at $z = -2/3$, and as z increases further, complete subgraphs of larger and larger order emerge.

SECTION 7

REAL NETWORKS ARE SUPERCRITICAL

Two predictions of random network theory are of special importance for real networks:

- Once the average degree exceeds $\langle k \rangle = 1$, a giant component emerges that contains a finite fraction of all nodes. Hence only for $\langle k \rangle > 1$ the nodes organize themselves into a recognizable network.
- For $\langle k \rangle > \ln N$ all components are absorbed by the giant component, resulting in a single connected network.

But, do real networks satisfy the criteria for the existence of a giant component, i.e. $\langle k \rangle > 1$? And will this giant component contain all nodes, i.e. is $\langle k \rangle > \ln N$, or do we expect some nodes and components to remain disconnected? These questions can be answered by comparing the measured $\langle k \rangle$ with the theoretical thresholds uncovered above.

Network	N	L	$\langle k \rangle$	$\ln N$
Internet	192,244	609,066	6.34	12.17
Power Grid	4,941	6,594	2.67	8.51
Science Collaboration	23,133	186,936	8.08	10.04
Actor Network	212,250	3,054,278	28.78	12.27
Yeast Protein Interactions	2,018	2,930	2.90	7.61

Table 3.1

Are real networks connected?

The number of nodes N and links L for several undirected networks, together with $\langle k \rangle$ and $\ln N$. A giant component is expected for $\langle k \rangle > 1$ and all nodes should join the giant component for $\langle k \rangle \geq \ln N$. While for all networks $\langle k \rangle > 1$, for most $\langle k \rangle$ is under the $\ln N$ threshold.

The measurements indicate that real networks extravagantly exceed the $\langle k \rangle = 1$ threshold. Indeed, sociologists estimate that an average person has around 1,000 acquaintances; a typical neuron is connected to dozens of other neurons, some to thousands; in our cells, each molecule takes part in several chemical reactions, some, like water, in hundreds. This conclusion is supported by Table 3.1, listing the average degree of several undirected networks,

in each case finding $\langle k \rangle > 1$. Hence the average degree of real networks is well beyond the $\langle k \rangle = 1$ threshold, implying that they all have a giant component.

Let us now inspect if we have single component (if $\langle k \rangle > \ln N$), or we expect the network to be fragmented into multiple components (if $\langle k \rangle < \ln N$). For social networks this would mean that $\langle k \rangle \geq \ln(7 \times 10^9) \approx 22.7$. That is, if the average individual has more than two dozens acquaintances, then a random society would have a single component, leaving no node disconnected. With $\langle k \rangle \approx 1,000$ this is clearly satisfied. Yet, according to Table 3.1 most real networks do not satisfy this criteria, indicating that they should consist of several disconnected components. This is a disconcerting prediction for the Internet, as it suggests that we should have routers that, being disconnected from the giant component, are unable to communicate with other routers. This prediction is at odd with reality, as these routers would be of little utility.

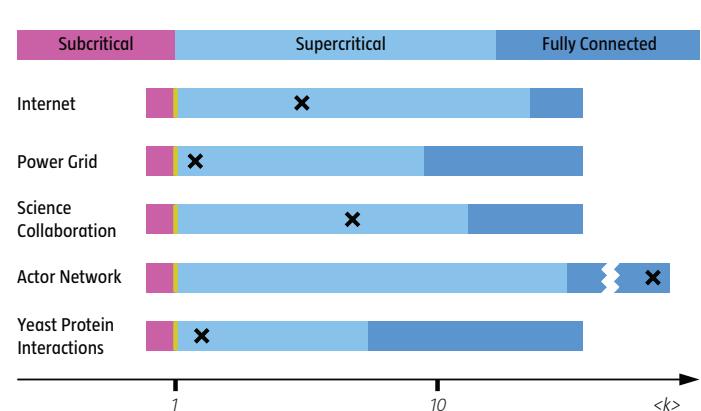


Image 3.8

Most real networks are supercritical.

The four regimes predicted by random network theory, marking with a cross the location of several real networks of Table 3.1. The diagram indicates that most networks are in the supercritical regime, hence they are expected to be broken into numerous isolated components. Only the actor network is in the connected regime, meaning that all nodes are expected to be part of a single giant component. Note that while the boundary between the subcritical and the supercritical regime is always at $\langle k \rangle = 1$, the boundary between the supercritical and the connected regimes is at $\ln N$, hence varies from system to system.

Taken together, we find that most real networks are in the supercritical regime ([Image 3.8](#)). This means that these networks have a giant component, but it coexists with many disconnected components and nodes. This is true, however, only if real networks are accurately described by the Erdős-Rényi model, i.e. if real networks are random. In the coming chapters, as we learn more about the structure of real networks, we will understand why real networks can stay connected despite failing the $k > \ln N$ criteria.

SMALL WORLD PROPERTY

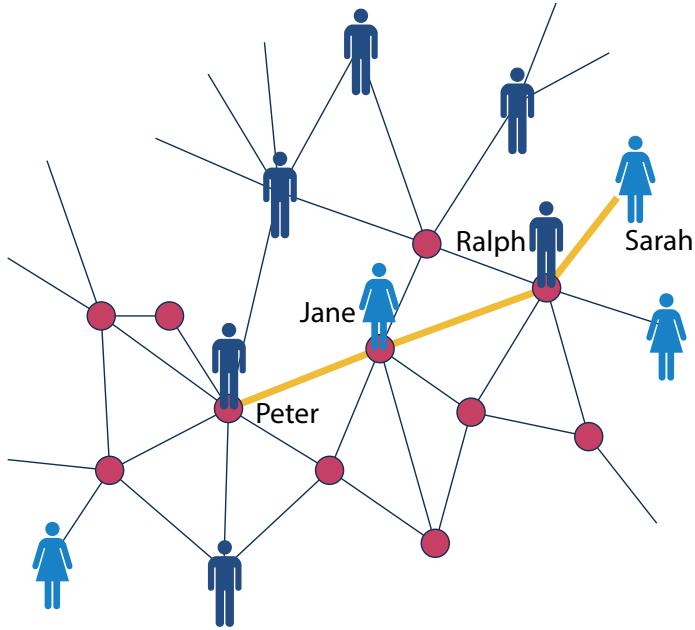


Image 3.9

Six degrees of separation.

According to six degrees of separation any two individuals, anywhere in the world, can be connected through a chain of six or fewer acquaintances. This means that while Sarah does not know Peter, she knows Ralph, who knows Jane and who in turn knows Peter. Hence Sarah is three degrees from Peter. In the language of network science six degrees, also called the small world property, states that the distance between any two nodes in a network is unexpectedly small.

Small world phenomena, also known as *six degrees of separation*, has long fascinated the general public. It states that if you choose any two individuals anywhere on earth, you will find a path of at most six acquaintances between them (Image 3.9). The fact that individuals who live in the same city are only a few handshakes from each other is by no means surprising. The small world concept goes further, however, stating that even individuals who are on the opposite side of the globe are six or fewer hand-shakes from us.

In the language of network science small world phenomena implies that *the distance between two randomly chosen nodes in a network is surprisingly short*. This statement raises two questions:

- What does short (or small) mean, i.e. short compared to what?
- How do we explain the existence of these short distances?

Both of these questions are answered by a simple calculation within the context of random networks. Consider a random network with average degree $\langle k \rangle$. A node in this network has on average:

- $\langle k \rangle$ nodes at distance one ($d=1$).
- $\langle k \rangle^2$ nodes at distance two ($d=2$).
- $\langle k \rangle^3$ nodes at distance three ($d=3$).
- ...
- $\langle k \rangle^d$ nodes at distance d .

For example, if $\langle k \rangle = 1,000$, we expect 10^6 individuals at distance two and about a billion individuals, i.e. almost the whole earth's population, at distance three from us.

To be precise, the expected number of nodes up to distance d from our starting node is

$$N(d) \approx 1 + \langle k \rangle + \langle k \rangle^2 + \dots + \langle k \rangle^d = \frac{\langle k \rangle^{d+1} - 1}{\langle k \rangle - 1}. \quad (15)$$

Yet, $N(d)$ must not exceed the total number of nodes, N , in the network. Therefore the distances cannot take up arbitrary values. We can identify a maximum distance d_{max} or the network's diameter at which $N(d)$ reaches N . By setting

$$N(d_{max}) \approx N, \quad (16)$$

and assuming that $\langle k \rangle \gg 1$, we can neglect the (-1) term in both the nominator and denominator of Eq. (15), obtaining

$$\langle k \rangle^{d_{max}} \approx N. \quad (17)$$

Therefore the diameter of a random network follows

$$d_{max} \propto \frac{\log N}{\log \langle k \rangle}, \quad (18)$$

which represents the *quantitative formulation of the small world phenomena*. The key, however is its interpretation:

- As derived, Eq. (18) predicts the scaling of the network diameter, d_{max} . Yet, for most networks Eq. (18) offers a better approximation to the average distance between two randomly chosen nodes, $\langle d \rangle$, than to d_{max} (Table 3.2). This is because d_{max} is often dominated by a few extreme paths, while $\langle d \rangle$ is averaged over all node pairs, a process that diminishes the fluctuations. Hence typically the small world property is defined by

$$\langle d \rangle \propto \frac{\log N}{\log \langle k \rangle}, \quad (19)$$

describing the dependence on N and $\langle k \rangle$ of the average distance in a network.

- In general $\log N \ll N$, hence the dependence of $\langle d \rangle$ on $\log N$ implies that the distances in a random network are *orders of magnitude smaller than the size of the network*. Consequently small world phenomena implies that the average path length or the diameter depends logarithmically on the system size. Hence, “small” means that $\langle d \rangle$ is proportional to $\log N$, rather than N or some power of N (Image 3.10).
- The $1/\log \langle k \rangle$ term implies that the denser the network, the smaller is the distance between the nodes.
- In real networks there are systematic corrections to Eq. (18), rooted in the fact that the number of nodes at distance $d > \langle d \rangle$ drops rapidly (Advanced Topics 3.F).

Network Name	N	L	$\langle k \rangle$	$\langle d \rangle$	d_{max}	$\frac{\log N}{\log \langle k \rangle}$
Internet	192,244	609,066	6.34	6.98	26	6.59
WWW	325,729	1,497,134	4.60	11.27	93	8.32
Power Grid	4,941	6,594	2.67	18.99	46	8.66
Mobile Phone Calls	36,595	91,826	2.51	11.72	39	11.42
Email	57,194	103,731	1.81	5.88	18	18.4
Science Collaboration	23,133	186,936	8.08	5.35	15	4.81
Actor Network	212,250	3,054,278	28.78	-	-	-
Citation Network	449,673	4,707,958	10.47	11.21	42	5.55
E Coli Metabolism	1,039	5,802	5.84	2.98	8	4.04
Yeast Protein Interactions	2,018	2,930	2.90	5.61	14	7.14

Table 3.2

Six degrees of separation.

The average distance $\langle d \rangle$ and the maximum distance d_{max} of the ten networks explored in this book. The last column provides $\langle d \rangle$ predicted by Eq. (19), indicating that it offers a reasonable approximation to $\langle d \rangle$. Yet, the agreement is not perfect – we will see in the next chapter that for many real networks Eq. (19) needs to be adjusted. For directed networks we list the average out-degree $\langle k_{out} \rangle$ and the path lengths are measured only along the direction of the links.

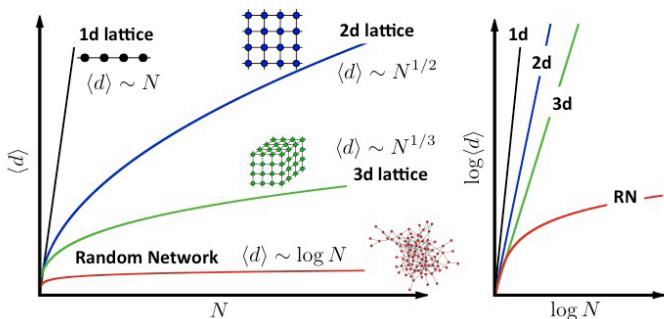


Image 3.10

Why are small worlds surprising?

Much of our intuition about distance is based on our experience with regular lattices, which do not display the small world phenomenon. Indeed,

- For a one-dimensional lattice (a line of length N) the diameter and the average path length scale linearly with N : $d_{\max} \sim \langle d \rangle \sim N$.
- For a square lattice $d_{\max} \sim \langle d \rangle \sim N^{1/2}$.
- For a cubic lattice $d_{\max} \sim \langle d \rangle \sim N^{1/3}$.
- In general, for a d -dimensional lattice we have $d_{\max} \sim \langle d \rangle \sim N^{1/d}$.

Such polynomial dependence predicts a much faster increase with N than Eq. (19), indicating that in regular lattices the path lengths are significantly longer than in a random network. The figure shows the predicted N -dependence of $\langle d \rangle$ for regular and random networks on a linear (left) and on a log-log (right) scale. If the social network would form a regular 2d lattice, where each individual knows only its nearest neighbors, the average distance between two individuals would be roughly $(7 \times 10^9)^{1/2} = 83,666$. Even if we correct for the fact that a person has about 1,000 acquaintances, not four, the average separation will be orders of magnitude larger than predicted by Eq. (19).

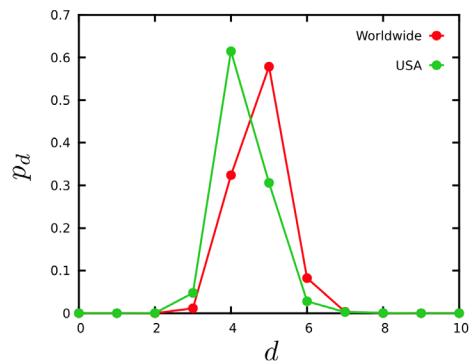


Image 3.11

Six degrees? Facebook finds only four.

Milgram's experiment could not detect the true distance between his study's participants, as he lacked an accurate map of the full social network. Today Facebook has the most extensive social network map ever assembled. Using Facebook's social graph of May 2011, consisting of 721 million active users and 68 billion symmetric friendship links, the average distance between the users was 4.74. The figure shows the distance distribution, p_d , for all pairs of Facebook users worldwide (full dataset) and within the US only. Therefore, instead of 'six degrees' researchers detected only 'four degrees of separation' [4], closer to the prediction of Eq. (20) than to Milgram's six degrees [23]. Using Facebook's N and L Eq. (19) predicts the average degree to be approximately 3.90, not far from the reported four degrees.

Let us illustrate the implications of Eq. (19) for social networks. Using $N = 7 \times 10^9$ and $\langle k \rangle \approx 10^3$, we obtain

$$\langle d \rangle = \frac{\ln 7 \times 10^9}{\ln(10^3)} = 3.28. \quad (20)$$

Therefore, all individuals on Earth should be within three to four handshakes of each other, about a half of "six degrees". The estimate (20) is probably closer to the real value given by Eq. (7) than the frequently quoted six degrees (Image 3.11).

While discovered in the context of social systems, the small world property applies beyond social networks. In Table 3.2 we compare the prediction of Eq. (19) with the average path length $\langle d \rangle$ for several real networks, finding that despite the diversity of these systems and the significant differences between them in terms of N and $\langle k \rangle$, Eq. (19) offers a reasonable approximation to the empirically observed $\langle d \rangle$.

The small world property has not only ignited the public's

imagination, but plays an important role in network science as well. It affects most network characteristics, from the spread of ideas in social networks to search on networks. The small world phenomena can be reasonably well understood in the context of the random network model: it is rooted in the fact that the number of nodes at distance d from a node increases exponentially with d . While in the coming chapters we will see that in real networks we encounter systematic deviations from Eq. (19), forcing us to replace it with more accurate predictions, the intuition offered by the random network model on the origin of the phenomenon remains valid.

A BRIEF HISTORY OF SIX DEGREES



Image 3.12
Frigyes Karinthy (1887–1938)

Hungarian writer, journalist and playwright, the first to describe the small world property. He remains one of the most popular writers in Hungary. English translation of *Chains*, the 1929 short story describing the small world phenomena, is available in [25].



Image 3.13
Stanley Milgram (1933–1984)

American social psychologist known for his experiments on obedience and authority. He designed and carried out the small world experiment in 1967 as part of his Harvard dissertation.

The first description of small world phenomena goes back to a 1929 story collection entitled *Minden másképpen van* (Everything is Different) by the Hungarian writer Frigyes Karinthy [21]. In *Láncszemek* (Chains), a short story in the volume, Karinthy suggests that one could name any person among earth's one and a half billion inhabitants (estimated population in 1929) and through at most five acquaintances, one of which he knew personally, he could link to him. To demonstrate his thesis Karinthy links a Nobel Prize winner to himself, noting that the Nobelist must know King Gustav, the Swedish monarch who hands out the Nobel Prize, who in turn is a consummate tennis player and occasionally plays with a tennis champion who is one of Karinthy's good friends. Remarking that finding a chain of acquaintances to celebrities, like a Nobelist, is easy, he next links a worker in Ford's factory to himself:

"The worker knows the manager in the shop, who knows Ford; Ford is on friendly terms with the general director of Hearst Publications, who last year became good friends with Árpád Pásztor, someone I not only know, but to the best of my knowledge a good friend of mine."

The first experimental study of small world phenomena took place four decades after Karinthy, in 1967, when Stanley Milgram turned the idea into an experiment probing the structure of social networks [23]. Milgram chose a stock broker in Boston and a divinity student in Sharon, Massachusetts as "targets". Randomly selected residents of Wichita, Kansas and Omaha, Nebraska received a letter containing a short summary of the study's purpose, a photograph, the name, address and information about the target person. They were asked to forward the letter to a friend, relative or acquaintance, who is more likely to know the target person. Milgram wrote in 1969: *"I asked a person of intelligence how many steps he thought it would take, and he said that it would require 100 intermediate persons, or more, to move from Nebraska to Sharon."* Yet, within a few days the first letter arrived, passing through only two links. Eventually 42 of the 160 letters made it back, some requiring close to a dozen intermediates. These completed chains allowed Milgram to determine the number of individuals required to get the letter to the target. He found that the median number of intermediates was 5.5, a relatively small number and remarkably close to Karinthy's 1929 insight.

A BRIEF HISTORY OF SIX DEGREES

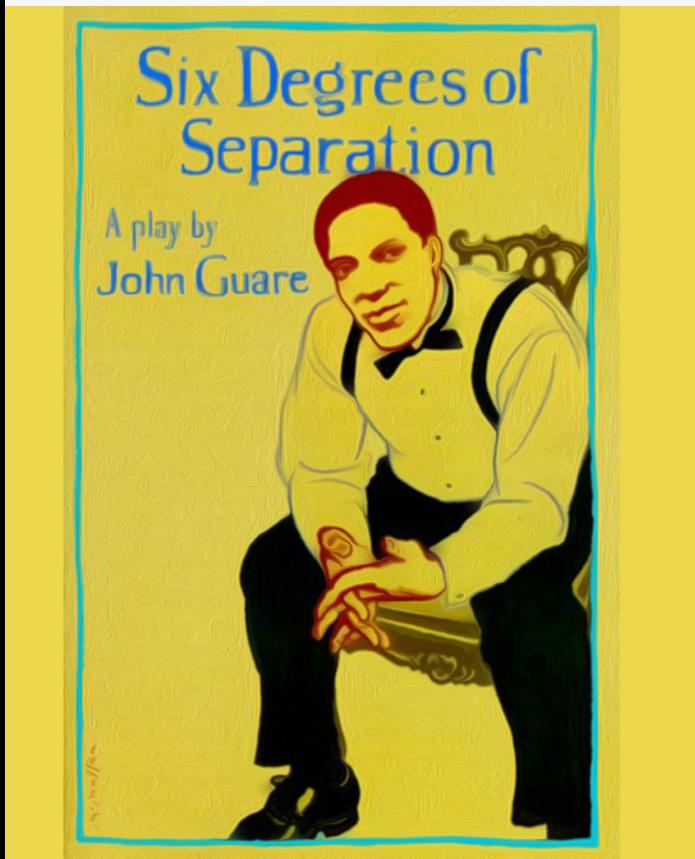


Image 3.14
Six Degrees of Separation.

Cover of John Guare's *Six Degrees of Separation* play, that helped turn six degrees into a catch phrase of popular culture.

The phrase "six degrees of separation" was introduced in 1991 by the playwright John Guare, who used it as the title of his Broadway play, later turned into a movie. The play's lead character, Ousa, musing about the world's interconnectedness, tells her daughter:

"Everybody on this planet is separated by only six other people. Six degrees of separation. Between us and everybody else on this planet. The president of the United States. A gondolier in Venice. It's not just the big names. It's anyone. A native in a rain forest. A Tierra del Fuegan. An Eskimo. I am bound to everyone on this planet by a trail of six people. It's a profound thought. How every person is a new door, opening up into other worlds."

Milgram's study was confined to the United States, linking individuals in Wichita and Omaha to Boston. Guare, however, with the sweep of a writer's imagination, generalized six degrees to the whole planet, bringing it closer in spirit to Karinthy's 1929 description. As more people watch movies than read sociology papers, Guare's version prevailed in popular thought.

Box 3.6c

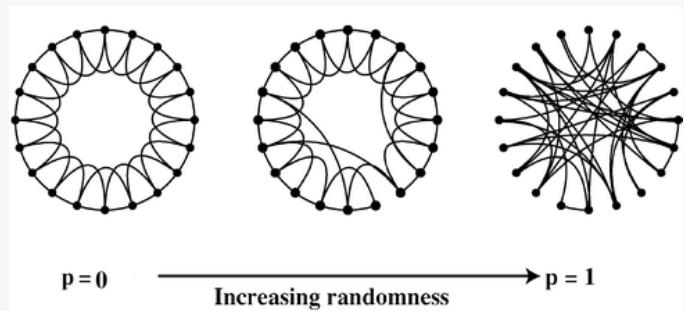


Image 3.15
Watts-Strogatz model.

The model starts from a ring of nodes, each node connected to their immediate and next neighbors, a configuration in which each node has clustering coefficient $C = 3/4$ ($p = 0$). With probability p each link is rewired to a randomly chosen node. For small p the network maintains a high average clustering coefficient but the random long-range links drastically decrease the distances between the nodes, inducing the small world effect (middle). For large p ($p = 1$) the network turns into a random network. (After [30]).

A new wave of interest in small worlds emerged following the 1998 study of Duncan Watts and Steven Strogatz, applied mathematicians working at Cornell [30]. They analyzed three real systems, the actor network of Hollywood, the neural network of the worm *C. elegans*, and the North American power grid, in each case finding that the average distance between the nodes is comparable to the random network prediction Eq. (19). Hence they found that the small world property applies to networks appearing in natural and technological systems as well. Watts and Strogatz also noted that these networks have a much higher clustering coefficient than expected for a random network, prompting them to propose a model to account for the coexistence of small path lengths and large clustering (Image 3.15). The model's properties are discussed in detail in the chapter devoted to social networks.

Box 3.6d

CLUSTERING COEFFICIENT

The local clustering coefficient C_i captures the density of links in node i 's immediate neighborhood: $C = 0$ means that there are no links between i 's neighbors; $C = 1$ implies that each of the i 's neighbors link to each other (Sect. 2.10). To calculate C_i for a node in a random network we need to estimate the expected number of links L_i between the node's k_i neighbors. In a random network the probability that two of i 's neighbors link to each other is p . As there are $k_i(k_i - 1)/2$ possible links between the k_i neighbors of node i , the expected value of L_i is

$$\langle L_i \rangle = p \frac{k_i(k_i - 1)}{2}.$$

Thus the local clustering coefficient of a random graph is

$$C_i = \frac{2\langle L_i \rangle}{k_i(k_i - 1)} = p = \frac{\langle k \rangle}{N}. \quad (21)$$

Equation (21) makes two predictions:

- (a) For fixed $\langle k \rangle$, the larger the network, the smaller is a node's clustering coefficient. Consequently the network's average clustering coefficient $\langle C \rangle$ is expected to decrease as $1/N$.
- (b) The local clustering coefficient of a node is independent of the node's degree.

To test the validity of Eq. (21) we plot $\langle C \rangle / \langle k \rangle$ in function of N for several undirected networks (Image 3.16a). We find that $\langle C \rangle / \langle k \rangle$ does not decrease as N^{-1} , but it is largely independent of N , in violation of Eq. (21). In Image 3.16b-d we also show the dependency of C on the node's degree k , for three real networks, finding that $C(k)$ systematically decreases with the degree, again in violation of Eq. (21).

Taken together, we find that the random network model does not capture the local clustering of real networks. Instead real networks have a much higher clustering coefficient than expected for a random network of similar N and L , and high-degree nodes tend to have a smaller clustering coefficient than low-degree nodes.

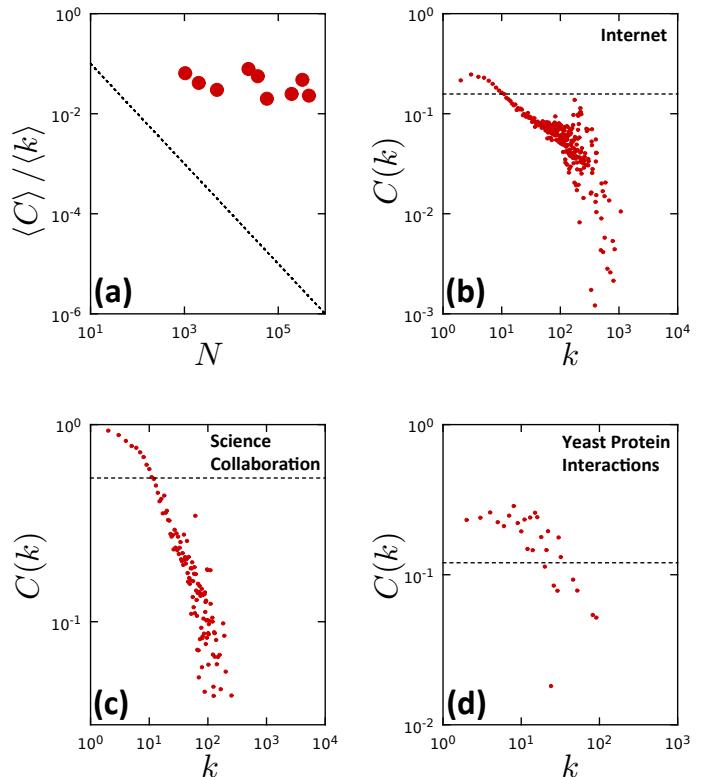


Image 3.16
Clustering in real networks.

(a) Comparison between the average clustering coefficient of real networks and the prediction Eq. (21) for random networks. Each circle corresponds to a network from Table 3.2. Directed network were made undirected to calculate C . The dashed line corresponds to Eq. (21), predicting that for random networks the average clustering coefficient should decrease as N^{-1} . In contrast, for real networks $\langle C \rangle$ has only a weak dependence on N .

(b)-(d) The dependence of the local clustering coefficient, $C(k)$, on the node's degree for (b) the Internet, (c) science collaboration network and (d) protein interaction network. $C(k)$ is measured by averaging the local clustering coefficient of all nodes with the same degree k . The dashed line corresponds to the prediction of Eq. (21) of the random network model, for which $C(k)$ is independent of k . In many real networks, the clustering coefficient decreases with k .

REAL NETWORKS ARE NOT RANDOM

For about four decades following its introduction in 1959 the random network model has dominated mathematical approaches to complex networks. The model suggests that if a network is not as regular as a square lattice, we should describe it as random. With that it equated complexity with randomness. We must therefore ask:

Do we really believe that real networks are random?

The answer is clearly no. The interactions between our proteins are governed by the strict laws of biochemistry so for the cell to function its chemical architecture can not be random. Similarly, in a random society an American student would be more likely to have among his friends Chinese factory workers than one of her classmates. In reality we suspect the existence of a deep order behind most complex systems. That order must be reflected in the structure of the network that describes their architecture, resulting in systematic deviations from a pure random configuration.

The degree to which random networks describe (or fail to describe) real systems must not be decided by epistemological arguments, but by a systematic quantitative comparison. This is possible because random network theory makes a number of quantitative predictions that can be tested on real networks:

Degree distribution: The degrees of a random network follow a binomial distribution, well approximated by a Poisson distribution in the $k \ll N$ limit. Yet, as shown in [Image 3.5](#), the Poisson distribution fails to capture the degree distribution of real networks. Instead in real systems we have more highly connected nodes than the random network model could account for.

Connectedness: Random network theory predicts that for $\langle k \rangle > 1$ we should observe a giant component, a condition satisfied by all networks we examined. Most networks do not satisfy the $\langle k \rangle > \ln N$ condition, which implies that these networks should be broken into isolated clusters ([Table 3.1](#)). Some networks are indeed fragmented, most are not.

Average path length: Random network theory predicts that the average path length scales as $\langle d \rangle \sim \log N / \log \langle k \rangle$, a prediction that captures the order of magnitude of the path lengths. Hence the random network model can account for the fundamental features of small world phenomena.

Clustering coefficient: In a random network the local clustering coefficient is independent of the node's degree and $\langle C \rangle$ depends on the system size as $1/N$. In contrast, measurements indicate that for real networks C decreases with the node degrees and is largely independent of the system size ([Image 3.16](#)).

Taken together, it appears that the small world phenomena is the only property reasonably explained by the random network model. All other network characteristics, from the degree distribution to the clustering coefficient, are significantly different in real and random networks. In fact, the more we learn about real networks, the more we will arrive at the startling conclusion that *we do not know of any real network that is accurately described by the random network model*.

This conclusion begs a legitimate question: If real networks are not random, why did we devote a full chapter to the random network model? The answer is simple: the model serves as a fundamental reference as we try to understand the properties of real networks. Each time we observe some network property we will have to ask if it could have emerged by chance. For this we turn to the random network model as a guide: if the property is present in the model, it means that randomness can account for it. If the property is absent in random networks, it may represent some signature of order, requiring a deeper explanation. So, the random network model may be the wrong model for most real systems, yet, *it remains quite relevant for network science* ([Box 3.8](#)).

Random networks and network science.

The lack of agreement between random and real networks raises an important question: how could a theory survive so long given its poor agreement with reality? The answer is simple: random network theory was never meant to serve as a model of real systems. True Erdős and Rényi did write in their first paper [9] that "This may be interesting not only from a purely mathematical point of view. In fact, the evolution of graphs may be considered as a rather simplified model of the evolution of certain communication nets (railways, road or electric network systems, etc.) of a country or some unit." Yet, this is the only mention of the potential practical value of their approach. The subsequent development of random graphs was driven by inherent mathematical challenges.

It is tempting to follow Thomas Kuhn and view network science as a paradigm change from random graphs to a theory of real networks [22]. In reality, there was no network paradigm before the end of 1990s. This period is characterized by a lack of interest in the problem, without systematic attempts to compare the properties of real networks with graph theoretical models. The work of Erdős and Rényi has gained prominence outside mathematics only after the emergence of network science (see [Image 3.17](#)).

Network theory does not lessen the contributions of Erdős and Rényi, but demonstrates the unintended importance of their work. When we point out the discrepancies between the predictions of the random network model and real networks, we do so only to offer a proper ground on which we can understand the properties of real systems.

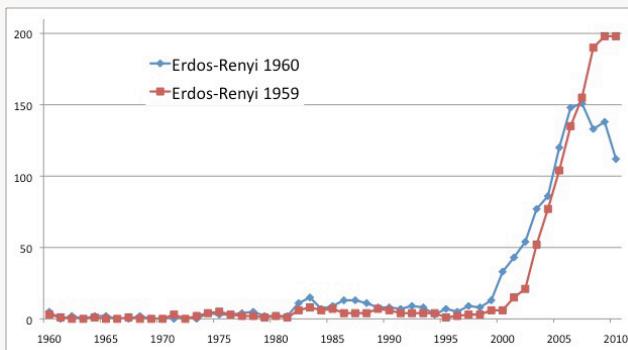


Image 3.17

Network science and random networks.

While today we perceive the Erdős-Rényi model as the cornerstone of network theory, the model was hardly known outside a small segment of mathematics. This is illustrated by the yearly citations of the first two papers by Erdős and Rényi, published in 1959 and 1960. For four decades after their publication the papers gathered less than 10 citations per year. The number of citations exploded after the first papers on scale-free networks [2, 3, 20] have turned Erdős and Rényi's work into the reference model of network theory.

Box 3.7

SUMMARY:

THE FIRST LAW OF NETWORKS

Network science has distilled a small number of fundamental organizing principles that govern the structure and evolution of real networks. We call these *network laws* as just like the laws of physics, they encode generic principles obeyed by many real networks. A network property quantifies as a law if

- (A) it has a unique quantitative, testable and falsifiable formulation;
- (B) it is obeyed by a large number of real networks;
- (C) it does not emerge by chance, hence it cannot be explained within the context of the random network model.

The results of this chapter allow us to formulate the first of these laws:

THE FIRST LAW: SMALL WORLD PROPERTY

In complex networks there are short distances between any pair of nodes.

Evidence for the first law is provided in Sect. 3.8. To recap in the context of the criteria A-C:

- A. Equation (19) offers the quantitative formulation of the First Law, predicting that the average distance between two randomly chosen nodes scales as a logarithm of the system size. Hence node-to-node distances are small compared to the network size.
- B. Table 3.2 offers evidence that most real networks obey the first law.
- C. As the small world property is present in random networks, the First Law apparently fails criterion C. Yet, we will see in the next chapter that in real networks distances are different from those expected in random networks, forcing us to modify Eq. (19).

At a glance: Random networks

- *Definition:* N nodes, where each node pair is connected with probability p .
- *Average degree:* $\langle k \rangle = p(N-1)$
- *Average number of links:* $\langle L \rangle = \frac{p(N-1)}{2}$
- *Degree distribution:* $P_k = \binom{N-1}{k} p^k (1-p)^{N-1-k}$.

For sparse networks ($k \ll N$), P_k has the Poisson form

$$P_k = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}.$$

- *Giant component (N_G):*

$\langle k \rangle < 1$: no giant component ($N_G \sim \ln N$)

$1 < \langle k \rangle < \ln N$: one giant component and disconnected clusters

$$\left(N_G \sim N^{\frac{2}{3}} \right)$$

$\langle k \rangle > \ln N$: all nodes join the giant component $N_G \sim (p - p_c)N$

- *Average distance:* $\langle d \rangle \propto \frac{\log N}{\log \langle k \rangle}$,
- *Clustering coefficient:* $C = \frac{\langle k \rangle}{N}$.

ADVANCED TOPICS 3.A: DERIVING THE POISSON DEGREE DISTRIBUTION

We start from the exact binomial distribution (7) or

$$p_k = \binom{N-l}{k} p^k (1-p)^{l-k} \quad (22)$$

$$p_x = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}. \quad (25)$$

that characterizes a random graph, and we rewrite the first term on the r.h.s. as

(23)

$$\binom{N-l}{k} = \frac{(N-l)(N-l-1)(N-l-2)\dots(N-l-k+1)(N-l-k)!}{k!(N-l-k)!} = \frac{(N-l)^k}{k!}$$

The last term of Eq. (22) can be simplified as

$$\ln[(1-p)^{(N-l)-k}] = (N-l-k) \ln(1 - \frac{\langle k \rangle}{N-l})$$

and using the series expansion

$$\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots, \forall |x| \leq 1$$

we obtain

$$\ln[(1-p)^{(N-l)-k}] \approx (N-l-k) \frac{\langle k \rangle}{N-l} = -\langle k \rangle (1 - \frac{k}{N-l}) \approx -\langle k \rangle,$$

which is valid if $N \gg k$, representing the small degree approximation at the heart of this derivation. Therefore the last term of Eq. (22) becomes

$$(1-p)^{(N-l)-k} = e^{-\langle k \rangle}. \quad (24)$$

Combining Eqs. (22), (23), and (24) we obtain the Poisson form of the degree distribution

$$p_k = \binom{N-l}{k} p^k (1-p)^{(N-l)-k} = \frac{(N-l)^k}{k!} p^k e^{-\langle k \rangle}$$

$$= \frac{(N-l)^k}{k!} \left(\frac{\langle k \rangle}{N-l} \right)^k e^{-\langle k \rangle},$$

ADVANCED TOPICS 3.B: MAXIMUM AND MINIMUM DEGREES

To determine the expected degree of the *largest node* in a random network, called the *network's upper cutoff*, we define the degree k_{\max} such that in a network of N nodes we have at most one node with degree higher than k_{\max} . Mathematically this means that the area behind the Poisson distribution p_k for $k \geq k_{\max}$ should be approximately one ([Image 3.18](#)). Since the area is given by $1 - P(k_{\max})$, where $P(k)$ is the cumulative degree distribution of p_k , the network's largest node satisfies:

$$N[1 - P(k_{\max})] \approx 1. \quad (26)$$

We write \approx instead of $=$, because k_{\max} is an integer, so in general the exact equation does not have a solution. For a Poisson distribution

$$1 - P(k_{\max}) = 1 - e^{-\langle k \rangle} \sum_{k=0}^{k_{\max}} \frac{\langle k \rangle^k}{k!} = e^{-\langle k \rangle} \sum_{k=k_{\max}+1}^{\infty} \frac{\langle k \rangle^k}{k!} \approx e^{-\langle k \rangle} \frac{\langle k \rangle^{k_{\max}+1}}{(k_{\max}+1)!}, \quad (27)$$

where in the last term we approximate the sum with its largest (leading) term.

For $N = 10^9$, and $\langle k \rangle = 1,000$ corresponding to roughly the size and average degree of the globe's social network, we obtain $k_{\max} = 1,185$, indicating that a random network lacks extremely popular individuals, or hubs.

We can use a similar argument to calculate the degree of the smallest node k_{\min} , or the *natural smallest cutoff*. Indeed, by requiring that there should be at most one node with degree smaller than k_{\min} we can write

$$NP(k_{\min}) \approx 1. \quad (28)$$

If $P(0) > 1$ the equation has no solution and $k_{\min} = 0$. For the *ER* network we have

$$P(k_{\min}) = e^{-\langle k \rangle} \sum_{k=0}^{k_{\min}} \frac{\langle k \rangle^k}{k!} \quad (29)$$

Solving Eq. (28) with $N = 10^9$ and $\langle k \rangle = 1,000$ we obtain $k_{\min} = 816$.

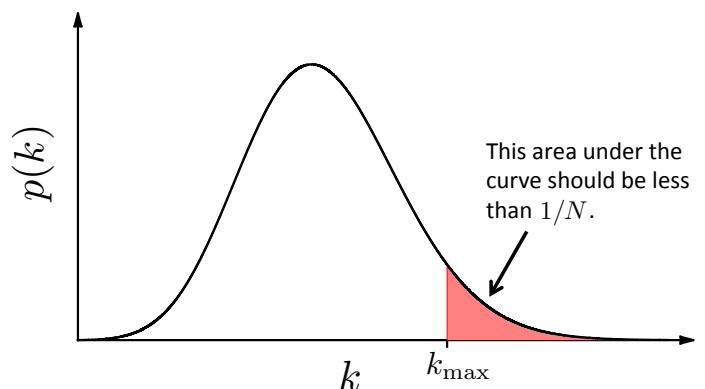


Image 3.18

Approximating the minimum and the maximum degree.

The maximum degree k_{\max} is chosen so that there is at most one node whose degree is higher than k_{\max} . This is often called the *natural upper cutoff* of a degree distribution. To calculate it, we need to set k_{\max} such that the area under the degree distribution p_k for $k \geq k_{\max}$ is exactly equal to $1/N$, hence this area multiplied by N , capturing the total number of nodes expected in the regime, is exactly one. We follow a similar argument to determine k_{\min} , or the expected smallest degree.

ADVANCED TOPICS 3.C: GIANT COMPONENT

Our aim here is to reproduce the argument, put forward independently by Solomonoff and Rapoport [28], and by Erdős and Rényi [8], on the emergence of giant component at $\langle k \rangle = 1$ (see also [24]).

Let us denote with $u = 1 - N_G/N$ the fraction of nodes that are not in the giant component (GC), whose size we take to be N_G . If node i is part of the GC, it must link to another node j , which is also part of the GC. Hence if i is not part of the GC, that could happen for two reasons:

- There is no link between i and j (probability for this is $1-p$).
- There is a link between i and j , but j is not part of the GC (probability for this is pu).

Therefore the total probability that i is not part of the GC via node j is $1-p+pu$. The probability that i is not linked to the GC via *any other node* is therefore $(1-p+pu)^{N-1}$, as there are $N-1$ nodes that could serve as a potential links to the GC for node i . As u is the fraction of nodes that do not belong to the GC, for any p and N the solution of the equation

$$u = (1 - p + pu)^{N-1} \quad (30)$$

provides the size of the giant component via $N_G = N(1-u)$. Using $p = \langle k \rangle / (N-1)$ and taking the log of both sides, for $\langle k \rangle \ll N$ we obtain

$$\ln u \approx (N-1) \ln \left[1 - \frac{\langle k \rangle}{N-1} (1-u) \right]. \quad (31)$$

Taking an exponential of both sides leads to $u = \exp[-\langle k \rangle(1-u)]$. If we denote with S the fraction of nodes in the giant component, $S = N_G/N$, then $S = 1-u$ and Eq. (31) provides

$$S = 1 - e^{-\langle k \rangle S}. \quad (32)$$

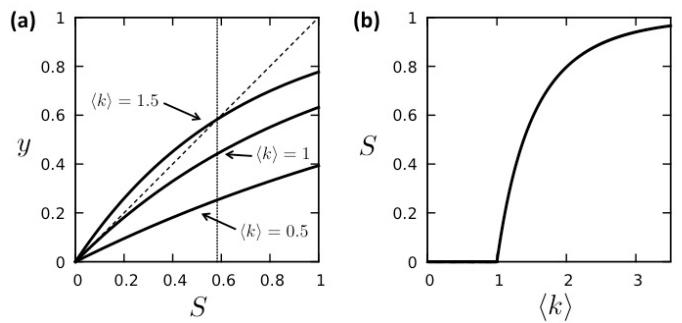
This simple looking equation provides the size of the giant component S in function of $\langle k \rangle$ ([Image 3.19](#)). Yet, Eq. (32)

does not have a closed solution. We can solve it graphically by plotting the right hand side of Eq. (32) as a function of S for various values of $\langle k \rangle$. To have a nonzero solution, the obtained curve must intersect with the dotted diagonal, representing the left hand side of Eq. (32). For small $\langle k \rangle$ the two curves intersect each other only for $S = 0$, indicating that for small $\langle k \rangle$, the size of the giant component is zero. Only when $\langle k \rangle$ exceeds a threshold value, does a non-zero solution emerge.

To determine the value of $\langle k \rangle$ at which we start having a nonzero solution we take a derivative of Eq. (32), as the phase transition point is when the r.h.s. of Eq. (32) has the same derivative as the l.h.s. of Eq. (32), i.e.

$$\begin{aligned} \frac{d}{dS} (1 - e^{-\langle k \rangle S}) &= 1, \\ \langle k \rangle e^{-\langle k \rangle S} &= 1. \end{aligned} \quad (33)$$

Setting $S = 0$, we obtain that the phase transition point is at $\langle k \rangle = 1$.



[Image 3.19](#)
Graphical solution for the size of the giant component.

- (a) The three curves in the left panel show $y = 1 - \exp[-\langle k \rangle S]$ for various $\langle k \rangle$. The diagonal dashed line corresponds $y = S$, and the intersection of the dotted and continuous lines provides the solution to Eq. (32), $S = 1 - \exp[-\langle k \rangle S]$. For the bottom curve there is only one intersection, at $S = 0$, indicating the absence of a giant component. The top curve is a solution at $S = 0.583\dots$ (vertical dashed line). The middle curve is precisely at the threshold between the regime where a non-zero solution for S exists and the regime where there is only the solution $S = 0$.
- (b) The size of the giant component in function of $\langle k \rangle$ as predicted by Eq. (32) [24].

ADVANCED TOPICS 3.D: COMPONENT SIZES

In [Image 3.5](#) we focused only on the size of the giant component, leaving an important question open: how many smaller components do we expect for a given $\langle k \rangle$, and what is their expected sizes? The aim of this section is to discuss these topics.

Component size distribution: For a random network the probability that a randomly chosen node belongs to a component of size s (different from the giant component G) is [24]

$$p_s \sim \frac{(s\langle k \rangle)^{s-1}}{s!} e^{-\langle k \rangle s}. \quad (34)$$

Replacing $\langle k \rangle^{s-1}$ with $\exp[(s-1) \ln \langle k \rangle]$ and using the

Stirling-formula $s! \approx \sqrt{2\pi s} \left(\frac{s}{e}\right)^s$

for large s we obtain

$$p_s \sim s^{-3/2} e^{-(\langle k \rangle - 1)s + (s-1)\ln \langle k \rangle}. \quad (35)$$

Therefore the component size distribution has two contributions: a slowly decreasing power law term $s^{-3/2}$ and a rapidly decreasing exponential term $e^{-(\langle k \rangle - 1)s + (s-1)\ln \langle k \rangle}$. Given that an exponential dominates for large s , Eq. (35) predicts that large components are prohibited. The only exception is at the critical point, $\langle k \rangle = 1$, where all terms in the exponential cancel, hence p_s follows the power law

$$p_s \sim s^{-3/2}. \quad (36)$$

As a power law decreases relatively slowly, at the critical point we expect to observe clusters of widely different sizes, a property consistent with the behavior of a system during a phase transition ([Advanced Topics 3.E](#)). These predictions are supported by numerical simulations in [Image 3.20](#), that shows p_s for three $\langle k \rangle$ values.

Average component size: The calculations also indicate that the average component size (once again, excluding

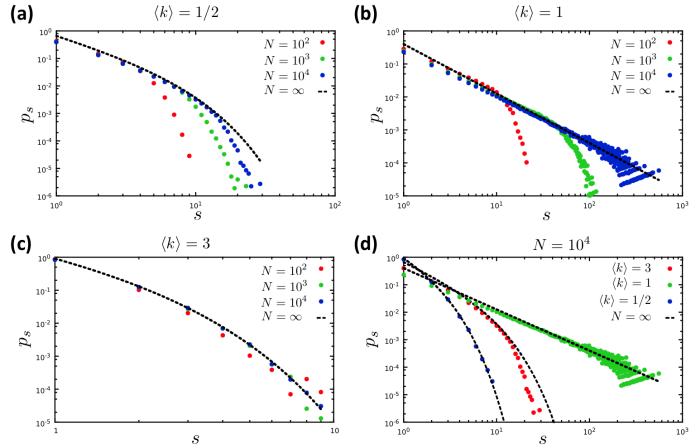


Image 3.20
Component size distribution.

Component size distribution in a random network, p_s excluding the giant component. (a)-(c) shows p_s for different $\langle k \rangle$ values and N , indicating that p_s converges for large N to the prediction (34). In (d) we show the results for $N = 10^4$, plotting together p_s for different $\langle k \rangle$. The plot clearly shows that while for $\langle k \rangle < 1$ and $\langle k \rangle > 1$ the p_s has a exponential form, right at the critical point $\langle k \rangle = 1$ the distribution follows the power law (36). The dotted line in each image correspond to the theoretical prediction (35). The first numerical study of the component size distribution in random networks was carried out in 1998, preceding the exploding interest in complex networks.

the giant component) follows [24]

$$\langle s \rangle = \frac{1}{1 - \langle k \rangle + \langle k \rangle N_G / N}. \quad (37)$$

For $\langle k \rangle < 1$ we lack a giant component ($N_G = 0$), hence Eq. (37) becomes

$$\langle s \rangle = \frac{1}{1 - \langle k \rangle}, \quad (38)$$

which diverges when the average degree approaches the critical point $\langle k \rangle = 1$. Therefore as we approach the critical point, the clusters are becoming bigger, signaling the

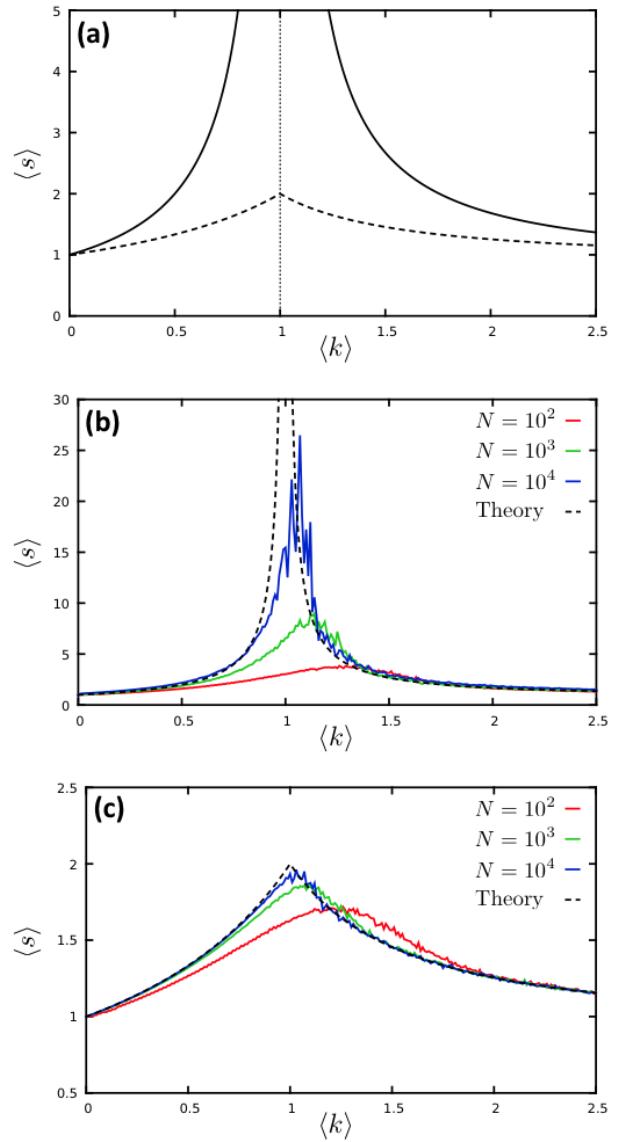
emergence of the giant component at $\langle k \rangle = 1$. Once again, numerical simulation support these predictions for large N ([Image 3.21](#)).

To determine the average component size for $\langle k \rangle > 1$ using Eq. (37), we need to first determine the size of the giant component. This can be done in a self-consistent manner, obtaining that the average cluster size decreases for $\langle k \rangle > 1$, as most of the clusters are gradually absorbed by the giant component.

Note that Eq. (37) predicts the size of the component to which a randomly chosen node belongs to. This is a biased measure, as the chance of belonging to a larger cluster is higher than the chance of belonging to a smaller one. The bias is linear in the cluster size, s . If we correct for this bias, we obtain the average size of the small components that we would get if we were to inspect each cluster one by one and measuring their average size [24]

$$\langle s' \rangle = \frac{2}{2 - \langle k \rangle + \langle k \rangle N_G / N}. \quad (39)$$

[Image 3.21](#) again offers numerical support for Eq. (39).



[Image 3.21](#)
Average component size.

- a. Upper curve: the average size $\langle s \rangle$ of a component to which a randomly chosen node belongs to as predicted by Eq. (39). Lower curve: the overall average size $\langle s' \rangle$ of a component as predicted by Eq. (37). The dotted vertical line marks $\langle k \rangle = 1$ (Redrawn after Newman, 2010).
- b. The average cluster size in a network measured in by numerical simulations, where we picked a node in the network and determined the size of the cluster it belongs to. This measure is biased, as each component of size s' will be counted s' times. The larger N becomes, the more closely the numerical data follows the prediction of Eq. (37). As predicted, $\langle s \rangle$ diverges at the $\langle k \rangle = 1$ critical point, supporting the existence of a phase transition in the system ([Advanced Topics 3.F](#)).
- c. The average cluster size in a network, where we corrected for the bias in (b) by selecting each component only once. The larger N becomes, the more closely the numerical data follows the prediction of Eq. (39).

ADVANCED TOPICS 3.E: SUPERCRITICAL REGIME.

To determine the value of $\langle k \rangle$ at which most nodes became part of the giant component, we calculate the probability that a randomly selected node does not have a link to the giant component, which is $(1-p)^{N_G} \approx (1-p)^N$, as in this regime $N_G \approx N$. The expected number of such isolated nodes is

$$I_N = N(1-p)^N = N \left(1 - \frac{N \cdot p}{N} \right)^N \approx Ne^{-Np}, \quad (40)$$

where we used $(1 - \frac{x}{n})^n \approx e^{-x}$, an approximation valid for

large n . If we make p sufficiently large, we arrive to the point where only one node remains disconnected from the giant component. At this point $I_N = 1$, hence according to Eq. (40) p needs to satisfy $Ne^{-Np} = 1$. Consequently, the value of p at which we are about to enter the fully connected regime is

$$p \sim \frac{\ln N}{N}, \quad (41)$$

which leads to Eq. (14) in terms of $\langle k \rangle$.

ADVANCED TOPICS 3.F: PHASE TRANSITIONS.

The emergence of the giant component at $\langle k \rangle = 1$ in the random network model is reminiscent of a *phase transition*, a much studied phenomenon in physics and chemistry [29]. Consider two examples:

- i. *Water-Ice Transition* ([Image 3.22a](#)): At high temperatures the H₂O molecules engage in a diffusive dance, forming small groups and then breaking apart to group up with other molecules. If cooled, at 0°C the molecules suddenly form a perfectly ordered ice crystal.
- ii. *Magnetism* ([Image 3.22b](#)): In ferromagnetic metals like iron at high temperatures the spins point in randomly chosen directions. Under some critical temperature T_c , however, all atoms orient their spins in the same direction and the metal becomes a magnet.

The freezing of a liquid and the emergence of magnetization are examples of phase transitions, representing transitions from disorder to order. Indeed, relative to the perfect order of the crystalline ice, liquid water is rather disordered. Similarly, the randomly oriented spins in a ferromagnetic take up the highly ordered common orientation under T_c .

Many properties of a system undergoing a phase transition are universal, that is, they are the same in a wide range of systems, from magma freezing into rock to a ceramic material turning into a superconductor. Furthermore, near the phase transition point, called the *critical point*, many quantities of interest follow power-laws. The phenomena observed near the critical point $\langle k \rangle = 1$ in a random network in many ways is similar to such a phase transition:

- The similarity between [Image 3.6a](#) and the magnetization diagram of [Image 3.22b](#) is not accidental: they both show transition from disorder to order, manifested as the emergence of a giant component as $\langle k \rangle$ exceeds $\langle k \rangle = 1$ in a random network.
- As we approach the freezing point, ice crystals of widely different sizes are observed, and so are domains of atoms with spins pointing in the same direction. The size distribution of the ice crystals or magnetic do-

- mains follows a power law. Similarly, while for $\langle k \rangle < 1$ and $\langle k \rangle > 1$ the cluster sizes follow an exponential distribution, in a random network right at the phase transition point, p_s follows a power law given by Eq.(36), implying the coexistence of components of widely different sizes.
- At the critical point the average size of the ice crystals or of the magnetic domains diverges, assuring that the whole system turns into a single frozen ice crystal or that all spins point in the same direction. Similarly in a random network the average cluster size $\langle s \rangle$ diverges as we approach $\langle k \rangle = 1$ ([Advanced Topics 3.D](#)).

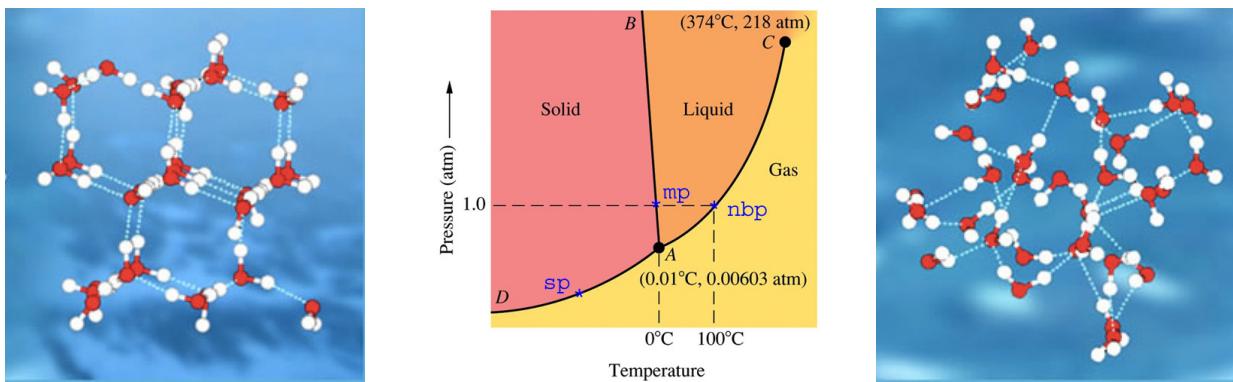


Image 3.22a
Water-Ice phase transition.

The hydrogen bonds that hold the water molecules together (dotted lines) are weak, constantly breaking up and re-forming, maintaining partially ordered local structures (left panel). The temperature-pressure phase diagram indicates (center panel) that by lowering the temperature, the water undergoes a phase transition, moving from a liquid (orange) to a frozen solid (red). In the solid phase each water molecule binds rigidly to four other molecules, forming an ice lattice (right panel). After <http://www.lbl.gov/Science-Articles/Archive/sabl/2005/February/water-solid.html>; phase diagram after <http://stevengoddard.wordpress.com/2010/09/02/the-ideal-world-phase-diagrams-part-deux/>

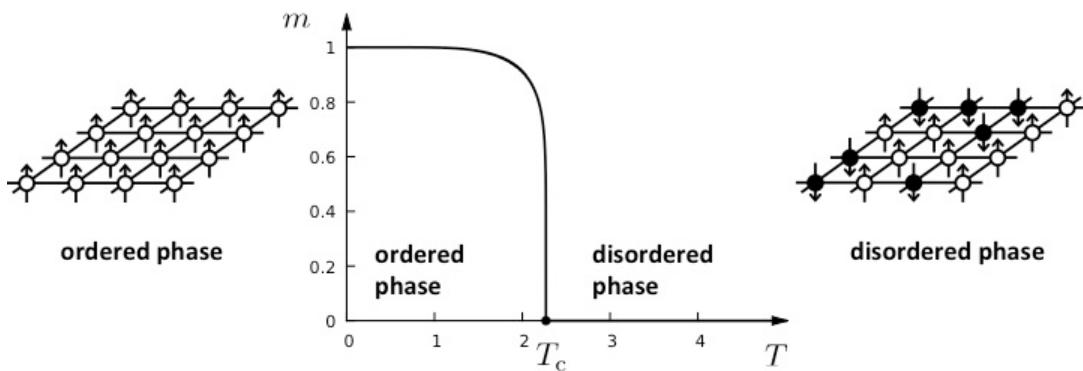


Image 3.22b
Magnetic phase transition.

In magnetic materials the magnetic moments of the individual atoms (spins) can point in two different directions. At high temperatures they choose randomly their direction (right panel), hence the system's total magnetization, $m = \Delta M / N$, where ΔM is the number of up spins minus the number of down spins, is zero. The phase diagram (middle panel) indicates that by lowering the temperature X , the system undergoes a phase transition at $T = T_c$ when a nonzero magnetization emerges, hence $m = M / N$ converges to one. In this ordered phase all spins point in the same direction (left panel).

ADVANCED TOPICS 3.G: CORRECTION TO SMALL WORLDS

Equation (18) offers only a rough approximation to the network diameter, valid for very large N and small d . Indeed, as soon as $\langle k \rangle^d$ approaches the system size N the $\langle k \rangle^d$ scaling must break down, as we are hitting the boundary of the network and there are not enough nodes to continue the $\langle k \rangle^d$ expansion. Such finite size effects result in corrections to Eq. (18).

For a random network with average degree $\langle k \rangle$, the network diameter is better approximated by (Fernholz & Ramachandran, 2007)

$$d_{\max} = \frac{\ln N}{\ln \langle k \rangle} + \frac{2 \ln N}{\ln [-W(\langle k \rangle \exp(-\langle k \rangle))]}, \quad (42)$$

where the Lambert W -function $W(z)$ is the principal inverse of $f(z) = z \exp(z)$. The first term on the r.h.s is Eq. (18), while the second is the correction that depends on the average degree. The correction increases the diameter, accounting for the fact that when we approach the network's diameter the number of modes must grow slower than $\langle k \rangle$. The magnitude of the correction becomes more obvious if we consider the various limits of Eq. (42).

In the $\langle k \rangle \rightarrow 0$ limit, i.e. when the network approaches the phase transition point, we can determine the Lambert W -function and the diameter becomes

$$d_{\max} = 3 \frac{\ln N}{\ln \langle k \rangle}. \quad (43)$$

Hence in the moment when the giant component emerges the network diameter is three times our prediction (18). This is due to the fact that at the critical point $\langle k \rangle = 1$ the network has a tree-like structure, consisting of long chains with hardly any loops, a configuration that significantly increases d_{\max} .

In the $\langle k \rangle \rightarrow \infty$ limit, corresponding to a very dense network, Eq. (42) becomes

$$d_{\max} = \frac{\ln N}{\ln \langle k \rangle} + \frac{2 \ln N}{\langle k \rangle} + \ln N \left(\frac{\ln \langle k \rangle}{\langle k \rangle^2} \right). \quad (44)$$

Hence if $\langle k \rangle$ increases, the second and the third terms vanish and the solution (42) converges to the result (18).

BIBLIOGRAPHY

- [1] Barabási, A.-L. (2003). *Linked: The new science of networks*. New York: Plume Books, 1 edition.
- [2] Barabási, A.-L. & Albert R. (1999). *Emergence of scaling in random networks*. Science, 286:509–512.
- [3] Barabási, A.-L., Albert, R., and Jeong, H. (1999). *Meanfield theory for scale-free random networks*. Physica A: Statistical Mechanics and its Applications, 272:173–187.
- [4] Backstrom, L., Boldi, P., Rosa, M., Ugander, J. & Vigna, S. (2011). *Four degrees of separation*. CoRR, abs/1111.4570.
- [5] Bollobás, B. (2001). *Random Graphs*. Cambridge University Press.
- [6] Christensen, K., Donangelo, R., Koiller, B., and Sneppen, K. (1998). *Evolution of Random Networks*. Physical Review Letters, 81:2380–2383.
- [7] Csicsery, G. P. (1993). *N is a Number: A Portait of Paul Erdős*.
- [8] Erdős, P. & Rényi, A. (1959). *On random graphs, I*. Publicationes Mathematicae (Debrecen), 6:290–297.
- [9] Erdős, P. & Rényi, A. (1960). *On the evolution of random graphs*. Publ. Math. Inst. Hung. Acad. Sci., 5:17–61.
- [10] Erdős, P. & Rényi, A. (1961a). *On the evolution of random graphs*. Bull. Inst. Internat. Statist., 38:343–347.
- [11] Erdős, P. & Rényi A. (1961b), *On the Strength of Connectedness of a Random Graph*, Acta Math. Acad. Sci. Hungary 12: 261–267.
- [12] Erdős, P. & Rényi, A. (1963). *Asymmetric graphs*. Acta Mathematica Acad. Sci. Hungarica, 14(3–4):295–315.
- [13] Erdős, P. & Rényi, A. (1966). *On random matrices*. Publ. Math. Inst. Hung. Acad. Sci., 8:455–461.
- [14] Erdős, P. & Rényi, A. (1966). *On the existence of a factor of degree one of a connected random graph*. Acta Math. Acad. Sci. Hungar., 17:359–368.
- [15] Erdős, P. & Rényi, A. (1968). *On random matrices II*. Studia Sci. Math. Hung., 13:459–464.
- [16] Fernholz, D. & Ramachandran, V. (2007). *The diameter of sparse random graphs*. Random Structures and Algorithms, 31(4):482–516.
- [17] Freeman, L. C. & Thompson, C. R. (1989). *Estimating Acquaintanceship*. Volume, pg. 147–158, in *The Small World*, Edited by Manfred Kochen (Ablex, Norwood, NJ)
- [18] Gilbert, E. N. (1959). *Random graphs*. The Annals of Mathematical Statistics, 30:1141–1144.
- [19] Hoffman, P. (1998). *The Man Who Loved Only Numbers: The Story of Paul Erdős and the Search for Mathematical Truth*. Hyperion Books.
- [20] Jeong, H., Albert, R. & Barabási, A. L. (1999). *Internet: Diameter of the world-wide web*. Nature, 401:130–131.
- [21] Frigyes K. , “Láncszemek,” in *Minden másképpen van* (Budapest: Atheneum Irodai es Nyomdai R.-T. Kiadása, 1929), 85–90. English translation is available in (Newman, Barabási, and Watts, 2006).
- [22] Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.
- [23] Milgram, S. (1967). *The Small World Problem*. Psychology Today, 2: 60–67.
- [24] Newman, M. (2010). *Networks: An Introduction*. Oxford University Press, 1 edition.
- [25] Newman, M., Barabási, A. L., and Watts, D. J. (2006). *The Structure and Dynamics of Networks*. Princeton University Press.

- [26] Rosenthal, H. (1960). *Acquaintances and contacts of Franklin Roosevelt*. Unpublished thesis. Massachusetts Institute of Technology.
- [27] Schechter, B. (1998). *My Brain is Open: The Mathematical Journeys of Paul Erdős*. Simon & Schuster.
- [28] Solomonoff, R. & Rapoport, A. (1951). *Connectivity of random nets*. Bulletin of Mathematical Biology, 13:107–117.
- [29] Stanley, H. E. (1987). *Introduction to Phase Transitions and Critical Phenomena*. Oxford University Press.
- [30] Watts, D. J. & Strogatz, S. H. (1998). *Collective dynamics of 'small-world' networks*. Nature 393: 409–10.