Critical Phenomena in Complex Networks

Term essay for Physics 563: Phase Transitions and the Renormalization Group
University of Illinois at Urbana-Champaign

Vikyath Deviprasad Rao

11 May 2012

Abstract

Networks arise in wide variety of contexts across disciplines, from social networks in sociology and transport networks in engineering, to networks for disease spreading in epidemiology and metabolic networks in biology. In this essay we provide an introduction to the systematic study of networks. We discuss a few important network models, focusing particularly on features that are similar to those that arise in statistical physics. We also discuss the application of renormalization group methods to examine the critical properties of a simple network model.

1 Introduction to networks

In its simplest form, a *network* consists of a set of points (called *vertices* or *nodes*) that are connected together in pairs by lines (called *edges* or *links*). The study of networks has garnered increasing interest amongst scientists across disciplines in the last two decades, not least because the class of systems that lend themselves to the network approach is very large. For example, in social networks, the vertices represent people, and various criteria may be used to link a pair of individuals together. In social networking websites such as Facebook, the link is obvious, but empirical studies in the past have also used telephone calls, e-mail exchanges, co-authorship of an academic paper, or sexual contact (to mention but a few) as their criterion.

Networks are also commonly used in the study of technological infrastructure, such as the internet and telephony backbones, the power grid, and train and airline routes. In biology, networks arise at the level of biochemistry (in mapping metabolic pathways or protein interactions, for instance) and at the level of ecology, in food webs. The brain is a complex network of neurons that are physically linked to one another, and the study of both artificial and biological neural networks is an active area of research.

The advantage of mapping a problem onto a network is that all the results derived for a general, abstract network can in principle be translated into the context of the specific problem to generate useful information. On the other hand, networks in general can be difficult to understand; Strogatz [7] lists the following common complications:

1. *Structural complexity:* the topology of the network may be difficult to characterize.
2. *Network evolution:* the presence or absence of particular edges may change with time.
3. *Connection diversity:* not all edges need to be identical, and they may have weights or directions associated with them.
4. *Node diversity:* there could be many different kinds of node.
5. *Dynamical complexity:* the state of each node could vary non-linearly in time.
6. The above complications can influence one another.

The typical approach in the study of complex networks is to ignore all but one of these complications. In this essay, we will focus mainly on static networks that are structurally complex: that is, we will examine the features of some networks with specific topologies.

In the remainder of this section, we will introduce some of the terminology and formalism used in the mathematical description of networks. In section 2, we will examine three networks with different structures: small-world networks, scale-free networks and random graphs, focusing particularly on the emergence of a giant component in the case of random graphs. In section 3, we will discuss the critical properties of a simple small-world network model, using renormalization group techniques.

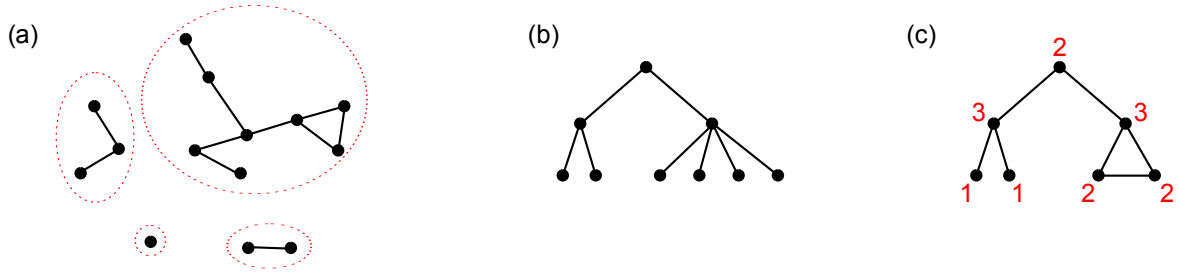

Figure 1: **A few basic networks**. (a) A disconnected network consisting of four components. (b) A tree, i.e. a connected network with no loops. (c) A connected network that contains one loop; the numbers indicate the degree of the vertices.

A network is mathematically represented by its *adjacency matrix* A, where each of the N vertices of the network are labelled $i = 1, 2, 3, \ldots, N$, and the matrix elements (in the simplest case) are

$$\mathsf{A}_{ij} = \begin{cases} 1 & \text{if vertices } i \text{ and } j \text{ are linked by an edge} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

In some applications, it is useful to weight edges differently; in this case, the elements of the adjacency matrix are allowed to take a range of values, rather than being binary. It is also sometimes useful to use a *directed network*, in which each edge has a direction associated with it; this results in a non-symmetric adjacency matrix.

A *path* in a network is a sequence of vertices such that every adjacent pair of vertices in the sequence is linked by an edge. A *connected network* is one in which every vertex may be reached from any other vertex by some path along the edges. By contrast, a disconnected network consists of distinct *components*, such that a vertex in one component cannot be reached from a vertex in another (see Fig. 1a). The adjacency matrix of a disconnected network can be written in block diagonal form (with appropriate labelling of the vertices). A *tree* is a connected network that contains no closed loops (see Fig. 1b–c).

The most important locally-defined quantity in a network is the *degree* (denoted k) of a given vertex, which is simply the number of edges connected to it. The degree of a vertex is one way of measuring its centrality, i.e. how important it is in the network.

2 Some important networks and their properties

The networks that we consider in this section all have complex structures: that is, the degrees of the vertices are distributed in a non-trivial manner. We first discuss random graphs, in which edges between a pair of vertices are placed at random in some way. Our exposition of random graphs is largely based on Newman's textbook [4]. We then discuss small-world networks, which are constructed from a regular network by a random re-wiring procedure. We also discuss scale-free networks, which have important practical applications. Finally,

we discuss the procedure of percolation, in which local modifications are made to a network.

2.1 Classical random graphs

A natural way two define a random graph is to take n vertices and place edges between m randomly chosen pairs of vertices. This type of random graph is denoted G_{nm}. The definition really selects for an ensemble of graphs with m edges, chosen from the set of all graphs with n vertices. Since each edge includes two vertices, it is easy to see that the mean degree $\langle k \rangle \equiv c$ in this model is $c = m/(n/2) = 2m/n$.

It turns out to be mathematically more feasible to deal with random graphs defined in a slightly different manner. Again, we start with n vertices, but now we fix the probability p of finding an edge between distinct pairs of vertices. This is known as the G_{np} model for random graphs, or the Erdős-Rényi random graph, named after two people who studied many of its properties [3].

Note that if we have a graph with n vertices, there are $\binom{n}{2}$ distinct pairs of vertices, and so the probability of choosing a graph with m edges is distributed binomially as follows:

$$\Pr(m) = \binom{\binom{n}{2}}{m} p^m (1-p)^{\binom{n}{2}-m}. \tag{2}$$

Thus the mean value of m (given p), and the mean degree of a vertex in the graph, are

$$\langle m \rangle = \binom{n}{2} p \qquad \text{and} \qquad c \equiv \langle k \rangle = \frac{\langle m \rangle}{n/2} = (n-1)p. \tag{3}$$

Finally, the probability of a given vertex being connected to k other vertices is

$$p_k = \binom{n-1}{k} p^k (1-p)^{n-1-k}. \tag{4}$$

This follows from the fact that we are dealing with independent probabilities, and there are a total of $(n-1)$ possible choices of other vertices to connect to. In the limit of large n, this simplifies to

$$p_k = e^{-c} \frac{c^k}{k!} \qquad (\text{large } n), \tag{5}$$

that is, we obtain a Poisson distribution.

Perhaps the most interesting feature of random graphs is the emergence of a giant component. If the $p = 0$, the graph has no edges and is therefore completely disconnected, i.e. each vertex is a component. If $p = 1$, then the graph is maximally connected: every vertex is directly linked to every other vertex, and the graph consists of a single component with n vertices. In particular, the size of this component is trivially extensive, i.e. it scales with n; for intermediate values of p, we focus on the largest component of the network, and examine how its size scales with n. It was shown [3] that the fraction S of vertices in the

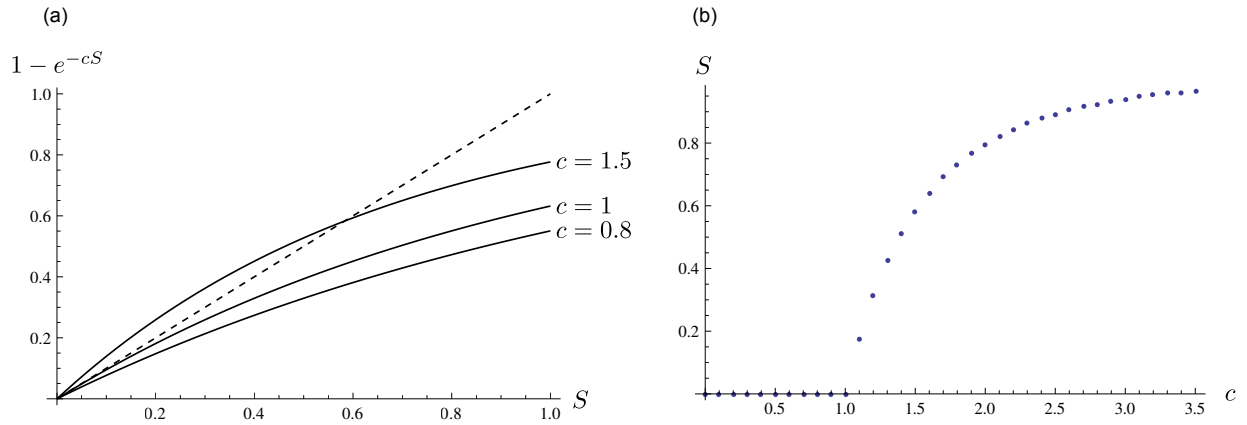

Figure 2: **Giant-component phase transition**. (a) Graphical solution of eq. (6) for three values of c. (b) Numerical solutions of eq. (6) for several values of c.

largest component obeys the equation

$$S = 1 - e^{-cS}, \tag{6}$$

where c again is the mean degree.

This equation cannot be solved analytically, but the graphical solution is depicted in Fig. 2a. For $c < 1$, no non-zero solutions exist, while for $c > 1$, there is a non-zero solution (in addition to $S = 0$). We solved eq. (6) numerically for several values of c, and plotted those in Fig. 2b. We observe a continuous phase transition at $c = 1$, where the largest component changes from scaling intensively to scaling extensively with n. When the largest component scales extensively, it is called a *giant component*. In the presence of a giant component, the remainder of the network consists of a number of small components whose average size is constant and does not scale with the network size.

Random graphs are perhaps the simplest model of structurally-complex networks, and the emergence of a giant component in an important feature that has interesting consequences for any system that is mapped onto a random graph. For example, the anecdotal "six degrees of separation" between any two people in the world might be explained by the presence of a giant component, which results in a short separation length between two vertices belonging to that component. In practice, however, most real-world networks do not exhibit the Poissonian distribution (5) of vertices characteristic of the G_{np} random graph. For this and other reasons, there are several other useful models of networks that are often used.

2.2 Small-world networks

Most real world networks are not completely random, nor are they completely regular. The small-world network model of Watts and Strogatz [8] aims to find the middle ground by beginning with a regular lattice network, and then replacing the original edges with random ones, with some probability p. They found that even a slight rewiring of the lattice leads to a

"small world" characterized by short paths between two vertices, as in the giant component of a random graph.

Quantitatively, this was examined using two quantities: first, the characteristic path length $L(p)$, which is the number of edges in the shortest path between two vertices in a lattice that has been re-wired with probability p, averaged over all pairs of vertices. And second, the clustering coefficient $C(p)$, which is defined as follows: given all possible edges between *neighbours* of a vertex i, let C_i be the fraction that are actually present; $C(p)$ is then defined as the average over all i of C_i. Thus, $C(p)$ is the measure of the "cliquishness" of a typical neighbourhood.

Watts and Strogatz found that for a broad range of p, the characteristic separation $L(p)$ is comparable to that of a random system (i.e. $L(p) \sim L(1)$), while the network remains highly clustered locally, like the lattice they started with (i.e. $C(p) \gg C(1)$). They tested their numerical predictions against three real-world networks: a social network of actors, linked by having acted in a film together; the power grid, with vertices representing generators or transformers, and edges representing high-voltage transmission lines; and finally, the neural network of the roundworm *C. elegans*. All these networks were seen to exhibit the small-world phenomenon.

2.3 Scale-free networks

Scale-free networks are characterized by a power-law degree distribution: $p_k \sim k^{-\gamma}$. This decays much more slowly than the Poisson distribution for random graphs, and this is typical of the tails of many real-world networks [7], including the Internet backbone, metabolic reaction networks, and the world-wide web. It has been shown theoretically [6, 2] that such a degree distribution arises in a model in which the network grows by the addition of vertices, but with new vertices being preferentially attached to vertices that are already well-connected.

Scale-free networks have the property that they are generally robust. This is seen in real networks such as the internet, where the failure of a single node typically does not affect the global properties of the network [1]. However, the price of this robustness is that a few particular vertices—the highly connected ones at the tail of the degree distribution—are particularly sensitive to failure, and so the network as a whole is highly vulnerable with respect to these vertices.

Fig. 3 illustrates the topology of a typical scale-free network, as well as those of the two other networks we have discussed so far. Notice that the random graph in Fig. 3a has a giant component that dominates the network, while the scale-free network (Fig. 3c) has a small number of highly-connected vertices.

2.4 Percolation

In this section we will discuss a process that has important practical applications. The process of percolation refers to the removal of some fraction of the vertices in a network along with attached edges (*site percolation*), or to the removal of some of the edges in a

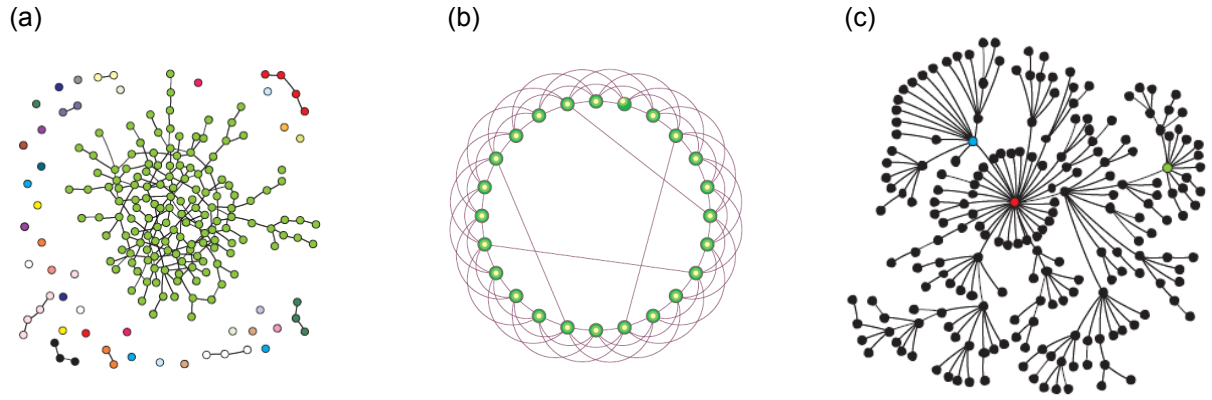

Figure 3: **Network topologies**. Typical diagrams for (a) a random graph, (b) a small-world network, and (c) a scale-free network. From Ref. [7].

network (*bond percolation*). This might (for example) correspond to the failure of particular hubs on the internet, or to the immunization of individuals in a disease-spreading network.

Site percolation is typically characterized by the probability ϕ that a randomly selected vertex in the original network is present after the percolation procedure. If $\phi = 1$, all the vertices are present, while for $\phi = 0$, all the vertices have been removed. For ϕ close to unity, a random graph that begins with a giant component continues to possess a giant component after percolation, but as ϕ is reduced, the giant component gradually disintegrates until the network consists only of small components. The transition from having a giant component to the absence of one is sharp in the limit as the total number of vertices approaches infinity [4]. This transition is called the *percolation transition*, and is similar to the emergence of the giant component discussed earlier. Again, this type of transition is important in, for example, models of epidemics: the disease can spread extensively if a giant component is present, but its effect is much smaller if the network only consists of a set of small components.

3 RG analysis of a small-world network

In this section, we discuss the application of renormalization group (RG) methods to deduce some critical properties of a small-world network [5]. We begin with a one-dimensional lattice of n sites with periodic boundary conditions. Each vertex is connected by edges to all neighbours within a range of K sites, so that the degree of each vertex is $2K$.

Recall that in order to construct a small-world network, we usually replace one of the Kn edges in the regular lattice with a random edge, with some probability p. For the purposes of this RG analysis, it is simpler to just add in an edge between Kn randomly chosen pairs of vertices, with probability p. For large n, this does not make a difference to the quantity of interest, namely the characteristic path length $L(p)$, which is the mean separation between a pair of vertices in the network. We will refer to the randomly added edges as "shortcuts" to distinguish them from the edges of the regular lattice.

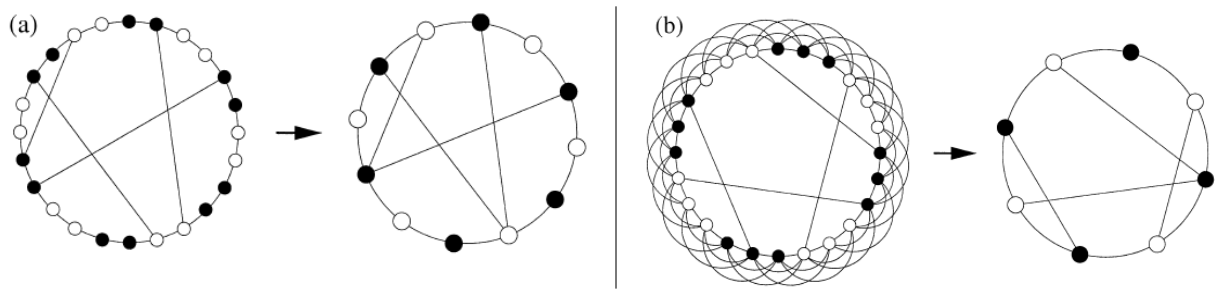

Figure 4: **RG transformations of small-world networks**. (a) $K = 1$: we begin with a lattice in which nearest neighbours are linked, and group together pairs of adjacent vertices. (b) From Ref. [5].

We are interested in the behaviour of the system in the limit as $p \to 0$, i.e. as the density of shortcuts approaches zero. Suppose we fix p and construct an ensemble of small-world networks by the procedure described above. The characteristic length L has two regimes of behaviour: first, if the system size n is small enough, the average number of shortcuts in this ensemble will be less than one. Then L is determined by the regular edges of the lattice, and in particular, we expect that L scales linearly with the size n of the system. The other possibility is that the system size n is large enough that the average number of shortcuts is greater than one. In this case, L will scale more slowly with n, and so we expect that $L \sim \log n$ in this regime. The transition between these regimes occurs at some system size $n = \xi$, and by the argument just outlined, we expect that at this system size,

$$p \cdot K\xi = p \cdot [\text{no. of trials}] = [\text{average no. of shortcuts}] \approx 1. \tag{7}$$

Thus we expect that ξ will diverge as $p \to 0$, since $\xi \sim p^{-1}$ by the equation above.

We will see that RG and finite-size scaling allow us to make the heuristic argument above mathematically rigorous. The divergence of the length scale ξ is analogous to that of the correlation length in a statistical mechanical system, and it corresponds to a one-sided phase transition at $p = 0$, analogous to that at $t = 0$ in the one-dimensional Ising model. We assume that $\xi \sim p^{-\tau}$, and we will derive the value of the critical exponent τ exactly.

First, consider the case $K = 1$, i.e. a one-dimensional lattice of n vertices (with the topology of a circle) in which each vertex is linked to its two nearest neighbours. Near the critical point, we expect that L obeys a finite-size scaling law of the form

$$L = nf(n/\xi), \tag{8}$$

and by the scaling arguments outlined above, we expect that the asymptotics of f are

$$f(z) \sim \begin{cases} \text{constant}, & z \ll 1 \\ z^{-1} \log z, & z \gg 1 \end{cases} \tag{9}$$

so that we recover the scaling forms $L \sim n$ and $L \sim \log n$ for the system size—measured

by the average number of shortcuts—being small and large respectively. We may therefore write $L = nf(np^\tau)$ near the critical region.

We now perform a real-space RG transformation in which we group together adjacent sites (see Fig. 4a). For sufficiently large n, the number of shortcuts is conserved under this transformation, while the system size halves, so we have the two renormalized variables

$$n' = \frac{n}{2} \qquad \text{and} \qquad p' = 2p. \tag{10}$$

Now consider the length of the shortest path between a given pair of vertices. In general, this consists of some regular edges and some shortcuts; the number of regular edges in this path halves under the RG transformation, while the number of shortcuts is unchanged. For $n \gg 1$ and $p \ll 1$, we may assume that in a typical path, there are many more regular edges than shortcuts, and so we expect that L scales like the regular edges, i.e.

$$L' = \frac{L}{2}. \tag{11}$$

Substituting the transformation equations (10) and (11) into the scaling form, we obtain $f(n'p'^\tau) = f(np^\tau)$, which gives

$$\tau = \frac{\log(n/n')}{\log(p'/p)} = 1, \tag{12}$$

in agreement with the result from the heuristic argument.

We now turn to the more general case where $K > 1$. This time, we group together adjacents site in groups of K sites in the RG transformation; Fig. 4b depicts this for $K = 3$. In the large n limit, the number of shortcuts is again preserved, and we have the following renormalized parameters:

$$n' = \frac{n}{K}, \quad p' = K^2 p, \quad K' = 1, \quad \text{and} \quad L' = L. \tag{13}$$

This transformation gives us a renormalized network with $K' = 1$, that is, a nearest-neighbour lattice model. Thus the transformation can be used to reduce a general small-world network (one with $K > 1$) to the simple $K = 1$ case considered above, and so $\tau = 1$ for all cases. Note that the scaling form for the transformed network is now

$$L = L' = n' f(n'p'^\tau) = \frac{n}{K} f(pKn), \tag{14}$$

and this holds asymptotically for $n/K = n' \gg 1$ and $K^2 p = p' \ll 1$. The first of these conditions ensures that we can neglect any inaccuracies that arise because we round of n to the nearest multiple of K in the RG transformation. The second condition may be re-written as $pKn \ll n/K$, which says that the average number of shortcuts must be less than the number of regular edges in the renormalized lattice.

4 Conclusion

We have attempted to present an overview of the complex networks that are increasingly used as models in various fields of study today. While we have mentioned a few applications of complex networks in passing, we have seen that an abstract complex network itself has interesting features, such as the emergence of a giant component or a particular form for the degree distribution. We have also seen that the methods of statistical physics can be brought to bear on complex networks, as illustrated by the application of RG methods to a simple small-world network.

References

[1] R. Albert, H. Jeong, and A.-L. Barabási. Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382, 2000.

[2] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

[3] P. Erdős and A. Rényi. On random graphs. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.

[4] M. Newman. *Networks: an introduction*. Oxford University Press, 2010.

[5] M.E.J. Newman and D.J. Watts. Renormalization group analysis of the small-world network model. *Physics Letters A*, 263(4-6):341–346, 1999.

[6] H.A. Simon. On a class of skew distribution functions. *Biometrika*, 42(3/4):425–440, 1955.

[7] S.H. Strogatz. Exploring complex networks. *Nature*, 410(6825):268–276, 2001.

[8] D.J. Watts and S.H. Strogatz. Collective dynamics of 'small-world' networks. *nature*, 393(6684):440–442, 1998.