Lectures on Complex Networks

S. N. Dorogovtsev

University of Aveiro
and
Ioffe Institute, St Petersburg

CLARENDON PRESS • OXFORD
2010
Preface

This text is a very concise modern introduction to the science of networks, based on lectures which I gave at several universities to students and non-specialists. My aim is to introduce a reader without serious background in mathematics or physics to the world of networks.

The term ‘complex networks’ is young. It came to use in the late 1990s when researchers from very distinct sciences—computer scientists, biologists, sociologists, physicists, and mathematicians—started to intensively study diverse real-world networks and their models. This notion refers to networks with more complex architectures than, say, a uniformly random graph with given numbers of nodes and links. Usually, in these complex architectures, hubs—strongly connected nodes—play a pivotal role. In this sense, the great majority of real-world networks are complex.

The field of complex networks is currently a very hot and attractive research area. The reader may ask: why all the fuss around networks in fundamental sciences like physics? I prefer the question: why are networks so interesting? The answer is not only the tremendous importance of the Internet and cellular networks. The point is that the geometry and structural organization of these and many other networks are very different from those of other well-studied objects—lattices. Networks and their function cannot be understood based on theories developed for finite-dimensional lattices, and a new vision is needed.

On the other hand, random networks are objects of statistical mechanics. So the course is essentially based on the standard apparatus of classical statistical physics. There are already several excellent popular science books and serious reference volumes on complex networks, including books on particular types of networks. The introductory lectures for beginners fill the existing gap between these two kinds of literature. The intended audience is mostly undergraduate and postgraduate students in physics and other natural science disciplines. There is some risk that inevitable oversimplification will only create an illusion of understanding. I believe however that this illusion is not too dangerous and may even be stimulating. Moreover, I suggest that the strict selection of material and discussion of recent results and fresh ideas will make this thin book useful, even for many specialists in networks. The reader who needs more detailed information and rigorous derivations can afterwards refer to more difficult reference books and original papers.

I am deeply indebted to my friends and colleagues in Portugal for their encouragement and advice, first and foremost to Anna Rozhnova,

Aveiro

July 2009

S.N.D.
# Contents

1 First steps towards networks  
1.1 Euler’s graph  
1.2 Examples of graphs  
1.3 Shortest path length  
1.4 Lattices and fractals  
1.5 Milgram’s experiment  
1.6 Directed networks  
1.7 What are random networks?  
1.8 Degree distribution  
1.9 Clustering  
1.10 Adjacency matrix  

2 Classical random graphs  
2.1 Two classical models  
2.2 Loops in classical random graphs  
2.3 Diameter of classical random graphs  
2.4 The birth of a giant component  
2.5 Finite components  

3 Small and large worlds  
3.1 The world of Paul Erdős  
3.2 Diameter of the Web  
3.3 Small-world networks  
3.4 Equilibrium versus growing trees  
3.5 Giant connected component at birth is fractal  
3.6 Dimensionality of a brush  

4 From the Internet to cellular nets  
4.1 Levels of the Internet  
4.2 The WWW  
4.3 Cellular networks  
4.4 Co-occurrence networks  

5 Uncorrelated networks  
5.1 The configuration model  
5.2 Hidden variables  
5.3 Neighbour degree distribution  
5.4 Loops in uncorrelated networks  
5.5 Statistics of shortest paths
5.6 Uncorrelated bipartite networks 38

6 Percolation and epidemics 41
6.1 Connected components in uncorrelated networks 41
6.2 Ultra-resilience phenomenon 43
6.3 Finite-size effects 45
6.4 $k$-cores 46
6.5 Epidemics in networks 48

7 Self-organization of networks 51
7.1 Random recursive trees 51
7.2 The Barabási–Albert model 52
7.3 General preferential attachment 53
7.4 Condensation phenomena 57
7.5 Accelerated growth 58
7.6 The BKT transition 59
7.7 Deterministic graphs 60

8 Correlations in networks 61
8.1 Degree–degree correlations 61
8.2 How to measure correlations 62
8.3 Assortative and disassortative mixing 62
8.4 Why are networks correlated? 64
8.5 Degree correlations and clustering 66

9 Weighted networks 67
9.1 The strength of weak ties 67
9.2 World-wide airport network 69
9.3 Modelling weighted networks 70

10 Motifs, cliques, communities 73
10.1 Cliques in networks 73
10.2 Statistics of motifs 74
10.3 Modularity 76
10.4 Detecting communities 78
10.5 Hierarchical architectures 82

11 Navigation and search 83
11.1 Random walks on networks 83
11.2 Biased random walks 85
11.3 Kleinberg’s problem 86
11.4 Navigability 88
11.5 Google PageRank 90

12 Traffic 93
12.1 Traffic in the Internet 93
12.2 Congestion 95
12.3 Cascading failures 97
13 **Interacting systems on networks** 99  
13.1 The Ising model on networks 99  
13.2 Critical phenomena 101  
13.3 Synchronization 102  
13.4 Games on networks 108  
13.5 Avalanches as branching processes 109

14 **Optimization** 113  
14.1 Critique of preferential attachment 113  
14.2 Optimized trade-offs 114  
14.3 The power of choice 115

15 **Outlook** 119

Further reading 121

References 123

Index 133
First steps towards networks

A network (or a graph) is a set of nodes connected by links. In principle, any system with coupled elements can be represented as a network, so that our world is full of networks. Specific networks—regular and disordered lattices—were main objects of study in physics and other natural sciences up to the end of the 20th century. It is already clear, however, that most natural and artificial networks, from the Internet to biological and social nets, by no means resemble lattices. The path to understanding these networks began in St Petersburg in 1735 with a mathematical problem formulated on a very small graph.

1.1 Euler’s graph

This small undirected graph (Fig. 1.1) with multiple links was considered by legendary Swiss born mathematician Leonhard Euler (1707–1783). Young Euler was invited to St Petersburg in 1727 and worked there until his death, with a 25-year break (1741–1766) in Berlin. In 1735 Euler made what is now regarded as the birth point of graph theory: he solved the Königsberg bridge problem. The structure of all possible paths within Königsberg in Euler’s time is represented in the form of a graph in Fig. 1.1. The nodes of the graph are separate land masses in old Königsberg, and its links are the bridges between these pieces of land. Could a pedestrian walk around Königsberg, crossing each bridge only once? In other words, is it possible to walk this graph passing through each link only once? Euler proved that such a walk is impossible.

In graph theory the total number of connections of a node is called degree (it is sometimes called connectivity in physics). Consequently, Euler’s graph has three nodes of degree 3 and one of degree 5. According to another definition, a simple graph does not have multiple links and loops of length 1. Otherwise, the graph is a multi-graph. Thus Euler’s graph is a multi-graph. Degree is a local characteristic. Any description of the structure of an entire network or of its parts is essentially based on two notions: a path and a loop. A path is an alternating sequence of adjacent nodes and links with no repeated nodes. A cycle (in graph theory) or a loop (in physics) is a closed path where only the start and end nodes coincide. Note that it is the presence of loops in Euler’s graph that makes the Königsberg bridge problem fascinating and profound.

---

1 The terms ‘vertices’ and ‘edges’ are more standard in graph theory.

2 Much later, in 1873, Carl Hierholzer proved that a walk of this kind exists if and only if every node in a graph has an even number of links.

---

Fig. 1.1 Euler’s graph for the Königsberg bridge problem. The undirected links of this graph are seven bridges of old Königsberg connecting four separate land masses—nodes Kniephof island and the banks of the Pregel river. As Euler proved in 1735, there is no walk on this graph that passes through each link only once.
Graphs without loops—trees—are usually more simple to analyse. For example, a one-dimensional chain is a tree. The numbers of nodes $N$ (which we call the ‘network size’) and links $L$ of a tree satisfy a simple relation $L = N - 1$.

### 1.2 Examples of graphs

A few simple graphs are shown in Fig. 1.2. A complete graph, Fig. 1.2 (a), which is widely used in exactly solvable models in physics, has all nodes interconnected. In a star graph (which is the most compact tree), Fig. 1.2 (b), the maximum separation between nodes is 2. Combs and brushes, containing numerous chains, are shown in Figs. 1.2 (b) and (c), respectively. Due to the chains, random walks on these graphs essentially differ from those on lattices. The next example—the Petersen graph in Fig. 1.2 (e)—is one of the so-called cage graphs. These graphs are regular in the sense that each node in the graph has the same number of connections. A $(q, g)$-cage is a graph with the minimum possible number of nodes for a given node degree $q$ and a given length $g$ of the shortest cycle. In synchronized systems the cage graph architectures provide optimal synchronization. The notion of a graph can be generalized. In one of the direct generalizations—hypergraphs—generalized links (hyperedges) connect triples, quadruples, etc. of nodes, see Fig. 1.2 (f).

The next two important regular graphs, a Cayley tree and a Bethe lattice, shown in Fig. 1.3, will be extensively discussed in these lectures. These are very different networks. A Cayley tree has a boundary, which contains a finite fraction of all nodes—dead ends—and a centre (a root). A Bethe lattice is obtained from an infinite Cayley tree by formal exclusion of dead ends. As a result, all nodes in a Bethe lattice are equivalent, so there is neither a boundary nor a centre. To get rid of boundaries, physicists often treat these graphs as containing infinite loops.

Collaboration and many other networks may have not one but two types of nodes. For example a network of scientific coauthorships contains nodes—authors and nodes—papers. Each scientific paper in this graph is linked to all of its authors. As a result we have a bipartite graph, shown in Fig. 1.4. These networks are actually hypergraphs, where a ‘node–paper’ together with its connections plays the role of a ‘hyperedge’. A one-mode projection of a bipartite graph, explained in Fig. 1.4, is less informative. Many empirical maps of networks are only one-mode projections of as yet unexplored real multi-partite networks.
1.3 Shortest path length

A distance $\ell_{ij}$ between two nodes $i$ and $j$ in a network is the length of the shortest path between them through the network. Two characteristics describe the separation between nodes in an entire network: the mean internode distance in the network and its diameter. The mean internode distance $\bar{\ell}$ (also mean geodesic distance) is the average of $\ell_{ij}$ over all those pairs of nodes $(i,j)$ between which there is at least one connecting path. (Note that, in general, networks will contain disconnected parts.) The diameter $\ell_D$ is the maximum internode distance in a network. We will demonstrate that in many large networks, there is no great difference between these two quantities.

It is the dependence of $\bar{\ell}$ or $\ell_D$ on network size $N$ that is particularly important for characterization of network architectures. In networks with a compact structure, which we discuss, $\bar{\ell}(N)$ grows with $N$ slower than in more loose structures—lattices.

1.4 Lattices and fractals

In finite-dimensional regular and disordered lattices (these lattices are supposed to have no long-range bonds), the size dependence $\bar{\ell}(N)$ is power-law,

$$\bar{\ell} \sim N^{1/d_f}. \tag{1.1}$$

Here $d$ is the dimensionality of a lattice—an integer number. In contrast, fractals (Fig. 1.5 explains this notion) may have non-integer dimensionalities. In fractals, $\bar{\ell} \sim N^{1/d_f}$, where $d_f$ is called a fractal or Hausdorff dimension. Note that there is no great difference between finite-dimensional lattices and fractals in respect of the dependence $\bar{\ell}(N)$; they are both ‘large worlds’. For example, for a two-dimensional lattice of $10^{12}$ nodes, $\bar{\ell} \sim 10^6$. Only when $d$ or $d_f$ tend to infinity does this dependence become non-power-law. Note that the fractal dimension can be found even if (large) $N$ is fixed. Simply count down the number of nodes $n(\ell)$ within a distance $\ell$ from a given node. In fractals (and lattices) this number is $n \sim \ell^{d_f}$.

For the sake of comparison, let us estimate $n(\ell)$ for the $q$-regular Bethe lattice and Cayley tree, Fig. 1.3. As is common, instead of the node degree $q$, we use another number—branching $b = q - 1$. Then $n = 1 + q(1 + b + b^2 + \ldots + b^{n-1}) \equiv qb^{n-1}$, where we assume that $n$ is large. Thus, $n \sim b^\ell$, in contrast to lattices and fractals, and so, for a Cayley tree, we have

$$\bar{\ell} \sim \frac{\ln N}{\ln b}, \tag{1.2}$$

which grows with $N$ much slower than for any finite-dimensional lattice.

In this respect, one may say that Cayley trees and Bethe lattices are infinite-dimensional objects—‘small worlds’. If, for example, a Cayley tree has $10^{12}$ nodes of degree, say, 5, then $\bar{\ell} \sim 10$, which is dramatically smaller than in the previous example for a two-dimensional lattice.
Generally, the term *small-world phenomenon* refers to a slower growth of $\bar{k}(N)$ than any positive power of $N$ (and to a more rapidly growing $n(\ell)$ than any power of $\ell$). The networks showing this phenomenon are called *small worlds*. Most of the explored real-world networks, which we will discuss, have compact architectures of this kind.

### 1.5 Milgram’s experiment

The small-world phenomenon was first observed in a social network. In 1967 prominent social psychologist Stanley Milgram (1933–1984) performed a seminal experiment for measuring distances in a network of acquaintances in the United States. The question was: how many intermediate social links separate two randomly selected (and geographically separated) individuals?

The idea of the experiment was elegant, see Fig. 1.6. Milgram chose two locations: Omaha, Nebraska and Boston. A target person was chosen at random in Boston. A large enough number of randomly selected residents of Omaha received a letter with the following instructions:

(i) If you know the target person ‘on a personal basis’ (his/her name and address were enclosed), send the letter directly to him/her.

(ii) Otherwise mail a copy of this instruction to your ‘personal’ acquaintance (someone you know on a first name basis) who is more likely than you to know the target person.

An essential fraction of letters approached the target, after passing through only, on average, 5.5 social links; which is a surprisingly small number. This is what is known as the ‘six degrees of separation’. One may think that the real shortest path length should be even smaller, since the experiment revealed only a small fraction of all possible chains between starting persons and the target.

It is dangerous however to believe sociologists too much: (i) they have to work with poorly defined and subjective material, (ii) they have to use poor statistics. The details of the experiment and the resulting number, the ‘six degrees’, were criticised, but nobody denies the essence of Milgram’s observation—the impressive smallness of the world of social relations.

### 1.6 Directed networks

In directed networks, at least some fraction of connections are directed. It seems that the first extensively studied nets of this type were networks of citations in scientific papers. The nodes of a citation network are scientific papers, and the directed links are citations of one paper within another (Fig. 1.7). New links in the citation networks emerge only between new nodes and already existing ones; new connections between existing nodes are impossible (one cannot update an already published paper). In graph theory networks of this kind are called *recursive graphs.*
Furthermore, all links in a network of citations have the same direction—to older papers. This is valid, of course, for publications in paper form, that is in printed journals and in books. In contrast, papers in many electronic archives may be updated. I can update my old works in the http://arXiv.org electronic archive and change their lists of references to cite more recent papers. So, some links in the citation networks of these electronic archives may be oppositely directed.

1.7 What are random networks?

Even if we ignore the directedness of connections, the apparently random network in Fig. 1.7 differs from the graphs shown in Figs. 1.2 and 1.3. But then, what is a random network from the point of view of a physicist or a mathematician? Note that, strictly speaking, the notion of randomness is not applicable to a single finite graph. Indeed, by inspecting this finite graph, one cannot find whether it was generated by a deterministic algorithm or by a non-deterministic one. In the spirit of statistical physics, a random network is not a single graph but a statistical ensemble. This ensemble is defined as a set of its members—particular graphs—where each member has its own given probability of realization, that is its statistical weight. By this definition, a given random network is some graph with one probability, another graph with another probability, and so on. To obtain some quantity, characterizing a random network, in principle we should collect the full statistics for all members of the statistical ensemble. To obtain the mean value of some quantity for a random network, we average this quantity over all members of the ensemble—over all realizations—taking into account their statistical weights.

The first example of a random graph is a classical random graph model, shown in Fig. 1.8. This is the \( G_{N,p} \) or Gilbert model defined as follows. Take a given number \( N \) of labelled nodes, say \( i = 1, 2, 3, \ldots, N \), and interlink each pair of nodes with a given probability \( p \). If \( N = 3 \), this gives eight possible configurations with the realization probabili-

\[ p(1-p)^2 \]
\[ p^2(1-p) \]
\[ p(1-p)^2 \]
\[ p^2(1-p) \]
\[ (1-p)^3 \]
\[ p(1-p)^2 \]
\[ p^2(1-p) \]
\[ p^3 \]

The http://arXiv.org is one of the largest electronic archives, used mostly by physicists but also mathematicians and computer scientists. Most papers on the statistical mechanics of networks can be found in the cond-mat and physics sections of this archive.

More rigorously, the statistical weights are proportional to the realization probabilities. However, the proportionality coefficient is an arbitrary constant.

Note the difference between the three kinds of scientist. As a rule, empirical researchers and experimenters collect statistics for a single realization of a random network. Scientists using numerical simulations (computer experiments) investigate a few or a relatively small number of realizations. Theorists consider all, or at least all essential, members of the statistical ensemble of a random network.

Fig. 1.7 Network of citations in scientific papers.

Fig. 1.8 The Gilbert model of a random graph (the \( G_{N,p} \) model) for \( N=3 \) with realization probabilities represented for all configurations. All graphs in each column are isomorphic, that is they can be transformed into each other by relabelling their nodes.
ties shown in the figure. Note that this graph is ‘labelled’ (has labelled nodes). As in classical statistical mechanics, where particles are distinguishable (i.e., can be labelled), networks are usually considered to be labelled, which is important for the resulting ensemble.

Physicists divide statistical ensembles into two classes—equilibrium and non-equilibrium—which correspond to equilibrium and non-equilibrium systems. This division is also relevant for random networks. For example, the ensemble presented in Fig. 1.8 is equilibrium—its statistical weights do not evolve. In non-equilibrium (evolving) ensembles, statistical weights of configurations vary with time, and the set of configurations may also vary. Growing networks are obviously non-equilibrium. However, even among networks with a fixed number number of nodes, one can find non-equilibrium nets.

Suppose now that the number of nodes in a random network approaches infinity. Then, as a rule, the statistics collected for one member of the ensemble almost surely coincides with the statistics for the entire ensemble—self-averaging takes place. In other words, a relative number of ensemble members with non-typical properties is negligibly small. It turns out that the self-averaging property is very common in disordered systems. So the features of many large, but finite individual graphs can be accurately described in terms of statistical ensembles. It is technically easier for a theoretical physicist to analyse a statistical ensemble than a single graph, and so the self-averaging is really useful.

1.8 Degree distribution

The degree distribution $P(q)$ is the probability that a randomly chosen node in a random network has degree $q$:

$$P(q) = \frac{\langle N(q) \rangle}{N}. \quad (1.3)$$

Here $\langle N(q) \rangle$ is the average number of nodes of degree $q$ in the network, where the averaging is over the entire statistical ensemble. We assume that the total number of nodes in each member of the ensemble is the same, $N = \sum_{q} \langle N(q) \rangle$. An empirical researcher, who studies a single graph, say graph $g$, measures the frequency of occurrence of nodes with degree $q$ in this graph: $P_g(q) = N_g(q)/N$. Here $N_g(q)$ is the number of nodes of degree $q$ in graph $g$. This quantity is also usually called a degree distribution. $P_g(q)$ approaches $P(q)$ in the infinite network limit.

The degree distribution is the simplest statistical characteristic of a random network, and it is usually only the first step towards the description of a net. Remarkably, in many situations knowledge of the degree distribution is sufficient for the understanding of a network and the processes taking place on it. In principle, the entire degree distribution is significant: its low- and high-degree parts are important for different network properties and functions. In classical random graphs such as shown in Fig. 1.8, degree distributions decay quite rapidly, $P(q) \sim 1/q!$
for large $q$ (see the next lecture). All their moments $\sum_q q^n P(q)$ are finite even as the network size approaches infinity, and so the mean degree $\langle q \rangle = \sum_q q P(q)$ is a typical scale for degrees. There are practically no strongly connected hubs in these networks.

In contrast, numerous real-world networks, from the Internet to cellular nets, have slowly decaying degree distributions, where hubs occur with noticeable probability and play essential roles. Higher moments of the degree distributions of these networks diverge if we tend the size of the network to infinity. A dependence with power-law asymptotics $P(q) \sim q^{-\gamma}$ at large $q$ gives a standard example of a slowly decaying degree distribution.\(^{11}\) The power-law distributions are also called scale-free and networks with these distributions are called scale-free networks. This term implies the absence of a typical node degree in the network.\(^ {12}\)

### 1.9 Clustering

Clustering is about how the nearest neighbours of a node in a network are interconnected, so it is a non-local characteristic of a node. In this respect clustering goes one step further than degree. The clustering coefficient of a node is the probability that two nearest neighbours of a node are themselves nearest neighbours. In other words, if node $j$ has $q_j$ nearest neighbours with $t_j$ connections between them, the local clustering coefficient is

$$C_j(q_j) = \frac{t_j}{q_j(q_j - 1)/2},$$

see Fig. 1.9. When all the nearest neighbours of node $j$ are interconnected, $C_j = 1$; when there are no connections between them, as in trees, $C_j = 0$. The number $t_j$ is the total number of triangles—loops of length 3—attached to the node, and so the clustering refers to the statistics of small loops—triangles—in a network. Importantly, most real-world networks have strong clustering.

In general, the clustering coefficient of a node depends on its degree. Empirical researchers often present their data on degree-dependent clustering by using an averaged quantity—the mean clustering coefficient of a node of degree $q$—that is $\bar{C}(q) = \langle C_j(q) \rangle$. Two different less informative integral characteristics of network clustering are traditionally used. The first is the mean clustering of a network, which is the average of the local clustering coefficient, eqn (1.4), over all nodes,

$$\bar{C} = \langle t_j/[q_j(q_j - 1)/2] \rangle = \sum_q P(q) \bar{C}(q).$$

The second characteristic—the clustering coefficient $C$ of a network or transitivity—allows one to find the total number of loops of length 3 in the network.\(^ {13}\) The clustering coefficient of a network is defined as

$$C = \frac{\text{the number of loops of length 3 in a network}}{\text{the number of connected triples of nodes}}.$$  

A triple here is a node and two of its nearest neighbours.\(^ {14}\) A 3-loop

\(^{11}\) The value of the moment $\sum_q q^n P(q)$ here is determined by the upper limit of the sum. In an infinite network, this limit approaches infinity. So, if exponent $\gamma \leq n + 1$, then the $n$th and higher moments of the distribution diverge.

\(^{12}\) More strictly, the term ‘scale-free’ refers to the following property of a power-law distribution $q^{-\gamma}$. A rescaling of $q$ by a constant, $c \rightarrow cq$, only has the effect of multiplication by a constant factor: $(cq)^{-\gamma} = c^{-\gamma}q^{-\gamma}$.

\(^{13}\) The notion of clustering was adapted from sociology, where it is usually called transitivity.

\(^{14}\) One can easily find that the number of connected triples of nodes equals $\sum_j q_j(q_j - 1)/2 = N((q^2) - (q))/2$.  

---

Fig. 1.9 The clustering coefficient of the central node equals 2/3.
consists of three triples, which explains the coefficient 3. The denominator gives three times the maximum possible number of loops of length 3. One can easily see that \( C \) is also the ratio of the average numerator of expression (1.4) and its average denominator, \( C \equiv \langle t_j \rangle / \langle q_j (q_j - 1) / 2 \rangle \). Compare this with the definition of mean clustering. If \( \bar{C}(q) \) is independent of degree \( q \), then the mean clustering and the clustering coefficient coincide, \( \bar{C} = C \).

### 1.10 Adjacency matrix

Networks are naturally represented in matrix form. A graph of \( N \) nodes is described by an \( N \times N \) adjacency matrix \( \hat{a} \) whose non-zero elements indicate connections between nodes.\(^{15} \) For undirected networks, a non-diagonal element \( a_{ij} \) of an adjacency matrix is equal to the number of links between nodes \( i \) and \( j \), and so the matrix is symmetric. A diagonal element \( a_{ii} \) is twice the number of loops of length 1 attached to node \( i \). The factor 2 here is clear: each 1-loop plays the role of a double connection for a node. As a result, the degree of node \( i \) is \( q_i = \sum_j a_{ij} \).

Any structural characteristic of a network can be expressed in terms of the adjacency matrix. See, for example, the expression for the total number \( T \) of triangles in a graph without 1-loops:

\[
T = \frac{1}{6} \sum_i (\hat{a}^3)_{ii} = \frac{1}{6} \text{Tr} \hat{a}^3. \tag{1.6}
\]

Here Tr denotes the trace of a matrix—the sum of its diagonal elements.\(^{16} \) This formula leads to a compact expression for the clustering coefficient.

Numerical calculations with adjacency matrices of large networks require huge memory resources. Fortunately, one can often avoid using adjacency matrices. The point is that real-world networks and their models are typically sparse. That is, the numbers of connections in these networks are much smaller than in complete graphs: \( L \ll N^2 \), i.e. \( \langle q \rangle \ll N \). In 1999, in the WWW, for example, the average number of outgoing and incoming hyperlinks per web page was about eight. Therefore the great majority of matrix elements in the adjacency matrices of these networks are zeros. So, instead of an adjacency matrix \( N \times N \), it is better to use a set of \( N \) vectors, \( i = 1, 2, \ldots, N \), where the components of vector \( i \) are the labels of the nearest neighbours of node \( i \). This takes up much less memory, \( \langle q \rangle N \ll N^2 \).
Further reading

Here is a short, incomplete list of popular science books, reference books, and comprehensive reviews on complex networks.


References


References


References

Lett. 89, 208701.


Index

adaptive routing, 96
adjacency matrix, 8, 78, 104, 106
ARPANET, 25
avalanche, 98, 109
Barabási-Albert model, 52, 53, 65, 71, 105, 109
Berezinskii-Kosterlitz-Thouless transition, 59
Bethe lattice, 2, 3, 12, 45, 47
betweenness centrality, 37, 55, 80, 95
distribution, 37, 56
branching, 3, 12, 20, 22, 42, 85, 109
cascading failures, 97
Cayley tree, 2
Chung-Lu model, 34
CiteULike, 32
clique, 11, 73, 81
clustering, 7
coefficient, 7, 8, 11, 20, 35, 38, 53, 66
local, 7, 27, 66
mean, 7, 27
community, 69, 77, 78
condensation, 57
configuration model, 33, 41
congestion, 95
threshold, 96
connected component, 13
finite, 15
distribution, 15, 43
giant, 14, 22, 41, 59
in-, 29
out-, 29
strongly, 29
weakly, 29
copying process, 55
correlations, 61
assortative, 63
degree-degree, 61
disassortative, 63
Curie-Weiss law, 15
cycle, 1
decentralized search algorithms, 86
degree, 1
distribution, 6
cumulative, 26, 76
cut-off, 45, 56, 65
heavy-tailed, 7
exponent, 7
exponential, 51
nearest-neighbour, 35
Poisson, 10
power-law, 7
scale-free, 7
in-degree, 18
distribution, 29
out-degree, 18
distribution, 29
Deep Web, 28
diameter, 3, 29
dimension, 3
fractal, 3
Hausdorff, 3, 24
spectral, 23
disease genes, 32
duplication-divergence process, 55
Echenique-Gómez-Gardehes-Moreno model, 96
edge, 1
electric power grid, 97
epidemic threshold, 48
epidemic outbreaks, 49
Erdős number, 17
Erdős-Rényi graph, 9, 33
Fabricant-Koutsoupias-Papadimitriou model, 115
Flickr, 32
forest fires, 110
fractal, 3, 22
games, 108
genome, 30, 32
Genetic disorder, 32
Gilbert model, 5
Granovetter's hypothesis, 67
graph, 1
bipartite, 2, 17, 32
brush, 2, 23
cage, 2, 108
comb, 2
classical random, 5, 9, 33, 61
deterministic, 60
exponential random, 57
fully connected, 2, 99, 104
hypergraph, 2, 32
isomorphic, 5, 75
labeled, 5, 33
multi-graph, 1
random geometric, 89
random regular, 12
regular, 2
simple, 1
star, 2
greedy algorithm, 86
delivery time, 87
greedy routing, 87
Google, 28, 90
PageRank, 90
growth, 51
accelerated, 27, 58
linear, 58
hidden variables, 34, 66
hierarchical organization, 82
Hurst exponent, 94
hyperedge, 2, 32
hyperlink, 25
hypertext, 28
Ising model, 99
immunization, 49
Internet, 25
Autonomous System level, 26, 46
protocols, 25, 93
router level, 26, 46
Kac formula, 84
k-core, 46
Königsberg bridge problem, 1
Kuramoto model, 104
Laplacian matrix, 106
Laplacian spectrum, 107
lattice, 3, 19, 23, 87
link, 1
weight, 67
load, 37, 97
loop, 1, 11, 35
mean-field theory, 15, 45, 100
metapopulation models, 50
metric space, 89
Metropolis algorithm, 57
Milgram's experiment, 4, 86
mixing, 63
assortative, 63
Index

dissortative, 63
modularity, 76
Molloy-Reed criterion, 42
motif, 66, 73
Motter-Lai model, 97
network, 1
airport, 70, 71
annealed, 100
coeauthorship, 17
complex, v
directed, 4, 29, 75, 90
equilibrium, 6, 20, 33
Fortune 1000, 38
human disease, 32
locally tree-like, 11, 36
karate club, 78
metabolic reactions, 30
navigable, 87, 90
neuronal, 103
non-equilibrium, 20, 57
phone call, 68
protein, 30
physical interactions, 30
functional interactions, 55
random, 5
scale-free, 7
searchable, 88
social, 4, 32, 63, 67
sparse, 8
small-world, 19, 107
tagged social, 32
uncorrelated, 10, 33
undirected, 1
weighted, 31, 67
Newman-Girvan algorithm, 77, 80
node, 1
strength, 70
order parameter, 14, 100
complex, 104
congestion, 96
packet switching, 93
Pareto distribution, 53
path, 1
optimal, 71
shortest, 3
Pearson coefficient, 62
percolation, 14
bootstrap, 50
clique, 74
explosive, 117
threshold, 14, 44
Petersen graph, 2
phase transition, 14
continuous, 14
first order, 47
hybrid, 47
infinite order, 59
mixed, 47
second order, 100
Poisson process, 93
preferential attachment, 52
linear, 54
nonlinear, 56
proportional, 52
prevalence, 48
prisoner's dilemma, 108
spatial, 109
projection, 2, 38
random walk, 83
autocorrelation function, 85
biased, 85
first return time, 84
on directed networks, 90
recurrent, 84
simple, 83
transient, 84
randomization algorithm, 65
reproductive number, 48
rich-club phenomenon, 46, 65
routing, 93
adaptive, 96
greedy, 87
routing table, 89, 93
sandpile model, 111
self-averaging, 6, 98
self-organized criticality, 109
self-organization, 54
shortest path, 3
shortest-path length, 3, 13, 18, 37
Simon model, 54
SIR model, 48
SIS model, 48
small world, 3
static model, 35
statistical ensemble, 5
statistical weight, 5
Surface Web, 28
synchronization, 107
criterion, 107
synchronization, 102
tolerance parameter, 97
traffic, 93
autocorrelation function, 94
fluctuations, 95
self-similar, 94
transitivity, 7
tree, 2
random equilibrium, 22
random recursive, 21, 51
root, 2, 51
spanning, 22
uniform, 22
ultra-resilience, 43, 48
ultra-small world, 37
vertex, 1
Watts-Strogatz model, 19, 107
generalization, 87
weak ties, 67
World Wide Web (WWW), 8, 18, 25, 28
Yule processes, 53
Zipf law, 53
Zipf plot, 76