Network Complexity

D J Raine[1] & V Norris[2]

[1]Department of Physics and Astronomy, University of Leicester, Leicester, LE1 7RH, UK
[2]Laboratoire des processus ioniques cellulaires UPRES A CNRS 6037, Faculté des Sciences et Techniques de Rouen , F76821 Mont Saint Aignan, France

Abstract

We discuss the background to the recent explosion of interest in the structure of homogeneous networks. We include a brief review of 'small worlds' and of the scale-free models that exhibit a power law distribution of nodes as a function of the number of links to or from them. The so-called standard canonical law of Mandelbrot is also considered and some examples from biology given. We then present the various ideas of the complexity of a network and introduce a measure that we call ?-complexity. Finally we show how this notion might arise from a thermodynamics of stationary systems maintained far from equilibrium by constraints.

1 Prehistory

1.1 To show subsection numbering

Until relatively recently it was widely believed that naturally occurring, homogeneous networks could be at least adequately modeled as random. Relational networks, such as food webs, which show a trophic structure [1], are consequently not homogeneous so are not counterexamples. Other relational graphs do however turn out to have interesting statistical properties and their structure may provide insights in many areas of science.

To set the scene we begin briefly with random graphs of N nodes, often called ER graphs [2]. These graphs are obtained by connecting randomly chosen pairs of nodes with probability p. Interesting questions that can be addressed usually involve the way in which properties scale with N or with p. For example, the smallest number of edges connecting randomly chosen nodes averaged over a random graph scales as logN. On the other hand, the appearance of a dominant connected subset sets in around p=1/N. In the following we shall refer to networks and graphs virtually interchangeably, except that we shall regard networks always as connected.

Of most interest to us is the distribution of the degrees of the nodes, that is the probability p(k) of picking a node having k edges, or, equivalently, the expected number of nodes of degree k, Np(k). For the ER graphs this is given by a Poisson distribution

$$p(k) = \frac{\mu^k}{k!}e^{-\mu}$$

where $\mu = 2Np$ is the mean number of edges per node. (Strictly this is the distribution obtained by averaging over many realisations of the graph, but will hold to a good approximation for any particular realisation.) It is a simple matter to compare this with the distribution of node degrees in naturally occurring networks. Perhaps the most interesting feature of the distribution is how many networks in nature do not actually follow it. In fact, many recent investigations have shown power law distributions over much of the range of degrees. Examples range from metabolic networks [3, 4] and protein networks [5] to the internet, co-authorship, citations and the language structure (see the table in [6]). These networks have come to be called 'scale-free' [7].