Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica A 355 (2005) 641–656
PHYSICA A
www.elsevier.com/locate/physa

A stochastic evolutionary model exhibiting power-law behaviour with an exponential cutoff

Trevor Fenner*, Mark Levene, George Loizou

School of Computer Science and Information Systems, Birkbeck College, University of London, London WC1E 7HX, UK

Available online 26 January 2005

Abstract

Recently several authors have proposed stochastic evolutionary models for the growth of complex networks that give rise to power-law distributions. These models are based on the notion of preferential attachment leading to the "rich get richer" phenomenon. Despite the generality of the proposed stochastic models, there are still some unexplained phenomena, which may arise due to the limited size of networks such as protein and e-mail networks. Such networks may in fact exhibit an exponential cutoff in the power-law scaling, although this cutoff may only be observable in the tail of the distribution for extremely large networks. We propose a modification of the basic stochastic evolutionary model, so that after a node is chosen preferentially, say according to the number of its inlinks, there is a small probability that this node will be discarded. We show that as a result of this modification, by viewing the stochastic process in terms of an urn transfer model, we obtain a power-law distribution with an exponential cutoff. Unlike many other models, the current model can capture instances where the exponent of the distribution is less than or equal to two. As a proof of concept, we demonstrate the consistency of our model by analysing a yeast protein interaction network, the distribution of which is known to follow a power law with an exponential cutoff.

*Corresponding author.
E-mail addresses: trevor@dcs.bbk.ac.uk (T. Fenner), mark@dcs.bbk.ac.uk (M. Levene), george@dcs.bbk.ac.uk (G. Loizou).

0378-4371/$ - see front matter
doi:10.1016/j.physa.2005.01.007

1. Introduction

Power-law distributions taking the form

$$f(i) = Ci^{-\tau} , \tag{1}$$

where C and τ are positive constants, are abundant in nature [1]. The constant τ is called the *exponent* of the distribution. Examples of such distributions are: *Zipf's law*, which states that the relative frequency of words in a text is inversely proportional to their rank, *Pareto's law*, which states that the number of people whose personal income is above a certain level follows a power-law distribution with an exponent between 1.5 and 2 (Pareto's law is also known as the *80:20 law*, stating that about 20% of the population earn 80% of the income) and *Gutenberg–Richter's law*, which states that, over a period of time, the number of earthquakes of a certain magnitude is roughly inversely proportional to the magnitude. Recently, several researchers have detected power-law distributions in the topology of various networks such as the World-Wide-Web [2,3] and author citation graphs [4].

The motivation for the current research is two-fold. First, from a complex network perspective, we would like to understand the stochastic mechanisms that govern the growth of a network. This has lead to fruitful interdisciplinary research by a mixture of Computer Scientists, Mathematicians, Statisticians, Physicists, and Social Scientists [5–9], who are actively involved in investigating various characteristics of complex networks, such as the degree distribution of the nodes, the diameter, and the relative sizes of various components. These researchers have proposed several stochastic models for the evolution of complex networks; all of these have the common theme of *preferential attachment*—which results in the "rich get richer" phenomenon—for example, where new links to existing nodes are added in proportion to the number of links to these nodes currently present. Considering the web as an example of a complex network, one of the challenges in this line of research is to explain the empirically discovered power-law distributions [10]. It turns out that the evolutionary model of preferential attachment fails to explain several of the empirical results, due to the fact that the exponents predicted are inconsistent with the observations. To address this problem, we proposed in [11] an extension of the stochastic model for the web's evolution in which the addition of links utilises a mixture of preferential and non-preferential mechanisms. We introduced a general stochastic model involving the transfer of balls between urns that also naturally models quantities such as the numbers of web pages in and visitors to a web site, which are not naturally described in graph-theoretic terms.

Another extension of the preferential attachment model, proposed in [6], takes into account the ageing of nodes, so that a link is connected to an old node, not only preferentially, but also depending on the age of the node: the older the node is, the less likely it is that other nodes will be connected to it. It was shown in [6] that if the ageing function is a power law then the degree distribution has a phase transition from a power-law distribution, when the exponent of the ageing function is less than one, to an exponential distribution, when the exponent is greater than one. A different model of node ageing was proposed in [12] with two alternative ageing