The domain model for protein-protein interaction networks.

Chris Cannings, U of Sheffield, Nov., 2008

There has been much recent interest in the growth of networks, particularly with reference to the WWW, but also in various social and in biological contexts. These are often described as having a power law distribution for the degrees of the nodes. Such a power law arises from the preferential attachment model (due to Simon(1954) but usually accredited to Barabasi and Albert(1999) in which new nodes are added to an existing network one at a time, and each new node is then linked to some m pre-existing nodes which are chosen with probabilities proportional to their degree. Such a process leads to a power law distribution for the degree distribution (Barabasi and Riorden,2001, Jordan,2006). This simple model is attractive since it envisages nodes which have already proved of *value* as acquiring new links more readily than those of lesser *value*.

The simplicity of this model has led some to claim there is some universality for some a process, though often there is little justification either statistically or logically. For example, in a Y2H (yeast-two-hybrid) experiment pairs of proteins are assayed to see if they bind, the outcome can then be represented as a graph in which each protein is represented as a node, and each edge corresponds to binding between the corresponding proteins. The structure of the network is of considerable interest in the attempt to understand organic evolution. It has been reported that the degree distribution (degree of a node is the number of edges adjacent to that node) satisfies a power law. However this claim is hardly substantiated by the data since there is little attempt to fit the data properly, and the experimental design is completely ignored.

Here I shall discuss the inadequacy of the fitting of the power law and introduce a "new" model, the domain model which attempts to model the process of protein binding (albeit in a highly abstracted form). We shall present results regarding the fit of this model to the basic data (allowing for the sampling), and results on the probabilities of certain motifs.

Refs.

Barabasi AL & Albert R (1999) Emergence of scaling in random networks. Science, 286, 509

Bollobas B, Riordan J, Spenser J & Tusnady G (2001) The degree sequence of a scale-free random graph process. Random Struct.Alg. 16, 279.

Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y., A comprehensive two-hybrid analysis to explore the yeast protein interactome, *Proc. Natl. Acad. Sci.U S A*, 98:4569–4574, 2001.

Jordan J (2006) The degree sequence and spectra of scale-free random graphs. Random Struct.Alg. 29,226.

Simon HA (1955) On a class of skew distribution functions. Biometrika, 42, 425

Thomas A, Cannings RC, Monk NAM and Cannings C (2003) On the structure of protein-protein interaction networks. Biochem Soc Trans., 31 1491-1496

Uetz,P. et al. (2000) **A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae***. *Nature* **403**: 623-627