Counting loops in random graphs and real-world networks

E. Marinari [1], R. Monasson [2] and G. Semerjian [3]

[1] *Dipartimento di Fisica and INFN, Università di Roma La Sapienza, P. A. Moro 2, 00185 Roma, Italy.*

[2] *CNRS-Laboratoire de Physique Théorique de l'ENS, 24 rue Lhomond, 75005 Paris, France,*

[3] *Dipartimento di Fisica and SMC-INFM,*

Università di Roma La Sapienza, P. A. Moro 2, 00185 Roma, Italy.

(Dated: July 25, 2005)

Data gathering in fields as diverse as social sciences, biology or Internet measurements, has provided an impressive amount of knowledge on the topology of the underlying interaction networks (i.e. graphs) in these domains. A crucial direction of research is now to identify characteristic features of these networks, in particular in view of (in)validating proposed modelings.

These features can be roughly classified in two categories. "Local" ones, for instance the connectivity distribution and the clustering coefficient, can be efficiently computed on any graph, even if very large. However the most distinguishing features might be "global", involving large patterns of the network. Identifying and counting these patterns becomes a very demanding task when their sizes increase, prohibiting in general the use of exact counting algorithms. Among these global features, a very natural one is the distribution of the lengths of the circuits (closed loops) in the graphs. Whereas it is quite easy to measure the number of short loops (triangles for instance), this becomes very hard when the loops studied have a length of the order of the size of the network.

We have developed in [1] an alternative approach to this counting problem, based on its reformulation in terms of a statistical mechanics model, which is treated within the Bethe approximation. The outcomes of this method are of two types. On the one hand, it yields an efficient approximate counting algorithm based on a message passing procedure, with a computational cost linear in the size of the graph. On the other hand, we have also obtained results on the typical number of long circuits in ensembles of random graphs, in particular in situations where usual probabilistic methods fail because of large fluctuations in these numbers.

---

[1] E. Marinari, R. Monasson and G.Semerjian, `cond-mat/0507525`.