Practical use of the configuration model

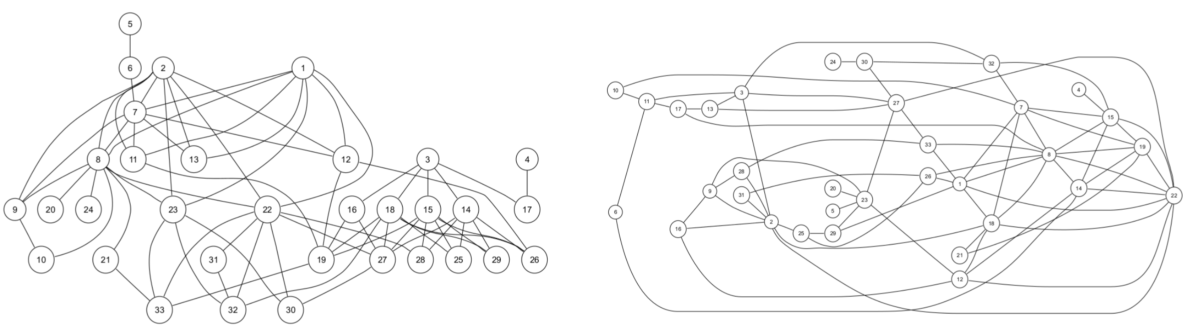
supervision: Lionel Tabourier, Matthieu Latapy (LIP6 – CNRS / SU)
location: Sorbonne Université – 4 place Jussieu, Paris

stages@complexnetworks.fr, http://complexnetworks.fr

Context:

Among models which represent real-world networks, the configuration model is among the most popular, because it is more realistic than random graphs with a fixed number of nodes and edges (Erdős-Rényi graphs) as it captures an essential element of complex networks, namely the fact that their degree is highly heterogeneously distributed.

Consequently, this model is often found for comparison to real networks, for example to understand some aspects of the structure of social networks [NP03], of the Internet [MKFV06], of ecological networks [MP04], among many others – see figure below.



Real structure of the food web for the Chesapeake bay ecosystem in summer (left) vs. one realization of its configuration model (right). Extracted from [BD11].

Goal:

However, generating a graph according to the configuration model properly can be problematic. Indeed, the comparison to real-world networks is meaningful only if the generation is uniform, i.e. if all graphs are generated with the same probability. The standard version of the configuration model (generation with rejection) guarantees uniformity, but it can be very slow with specific distributions. For instance, it is the case if the number of nodes with a large degree is important, which is quite usual for real networks.

Methods have been proposed to circumvent this problem. Most rely on much more elaborate algorithms that bring guarantees for some specific graph families [AGW19]. However, there is currently no method that guarantees a fast generation with any input. Our approach to the problem is then a bit different: we are looking for bounds on the generation time for different generation methods, in order to know what would be the most suitable method before starting the process itself. At least, we want relevant bounds that would indicate if using the configuration model with rejection can produce a sample within a reasonable time.

Requested profile:

This internship is directed at students with various background (complex networks, algorithmics, graph theory) but with a strong interest in graph algorithmics and/or theory and its applications.

The intern will be part of the Complex Networks team of the LIP6 (SU-CNRS), located in Paris on Jussieu Campus. The internship will be supervised by Lionel Tabourier and Matthieu Latapy.

References

[AGW19] Andrii Arman, Pu Gao, and Nicholas Wormald. Fast uniform generation of random graphs with given degree sequences. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1371–1379. IEEE, 2019.

[BD11] Joseph Blitzstein and Persi Diaconis. A sequential importance sampling algorithm for generating random graphs with prescribed degrees. *Internet mathematics*, 6(4):489–522, 2011.

[MKFV06] Priya Mahadevan, Dmitri Krioukov, Kevin Fall, and Amin Vahdat. Systematic topology analysis and generation using degree correlations. *ACM SIGCOMM Computer Communication Review*, 36(4):135–146, 2006.

[MP04] István Miklós and János Podani. Randomization of presence–absence matrices: comments and new algorithms. *Ecology*, 85(1):86–92, 2004.

[NP03] Mark EJ Newman and Juyong Park. Why social networks are different from other types of networks. *Physical review E*, 68(3):036122, 2003.