Beyond space and blocks: Generating networks with arbitrary structure
Rémy Cazabet, Jacques Fize, Salvatore Citraro, Giulio Rossetti

To cite this version:
Rémy Cazabet, Jacques Fize, Salvatore Citraro, Giulio Rossetti. Beyond space and blocks: Generating networks with arbitrary structure. International School and Conference on Network Science 2023, Jul 2023, Vienne, Austria. hal-04441398

HAL Id: hal-04441398
https://hal.science/hal-04441398
Submitted on 6 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Random graph models are ubiquitous in network science. Often used as null models, they require to preserve the size of the observed networks, i.e., the number of nodes and edges. We focus on models in which the probability to observe an edge between each pair of node is independent, as in an ER or expected degree preserving (Chung-Lu) model. Such models are widespread for block structures (SBM) and spatial structure (e.g., Waxman, gravity model), but there is no simple solution to create random graphs with arbitrary —defined by the analyst— structures.

We propose a model allowing to generate graphs of a chosen number of nodes and expected number of edges, following an arbitrary structure, and with a controllable amount of randomness. The model requires:

- The number of nodes n
- The expected number of edges m
- A cost function describing the generated network's structure

Let’s take a simple example: a possible cost function for a spatial graph is simply the distance between the nodes —This distance can be attributed randomly, or kept from observed data.

We use a function allowing to tune the amount of randomness of the structure, ε ∈ [0,1]. When ε = 0, the structure is deterministically determined by the cost function. When ε = 1, the network is an ER random graph. The function interpolates between the two using a Bezier function, as illustrated in Fig. 2.

By convention, if

- If ε = 0, G = ER(n, m/n)
- If ε = 1, G = Rank(several functions on the binary tree, ranking node-pairs according to the distance).

As a default, we use the level-based function, which is the most efficient. In the simplest case, the probability to observe a graph in the resulting graph is proportional to the distance between the nodes in the tree.

To illustrate the flexibility of this approach, we first reproduce the famous Watts and Strogatz experiment, on which a regular circular network is progressively randomized, leading to the appearance of a small-world structure. (We use a normalized short path index increasing with shortest paths, in order to compare different network structures)

Following the same experimental process, we can observe how different structures behaves in term of clustering coefficient and short paths