In Chapter 7 we have studied the configuration model with undirected edges. Many real-life networks however have naturally directed edges. Think, for example of Twitter and the World Wide Web. In this project we study the directed configuration model. What are the main differences and common properties compared to the undirected scenario?

What you minimally should do:

- Study the paper: Colin Cooper and Allan Frieze: "The Size of the Largest Strongly Connected Component of a Random Digraph with a Given Degree Sequence" Combinatorics, Probability and Computing, Volume 13 Issue 3, May 2004 Pages 319 - 337. Summarize the properties (structure) of the directed configuration models and compare them to the undirected scenario. What is the main difference between directed and undirected configuration model? Describe the 'bowtie' structure of the (strongly) connected component.

What you can do:

- You may analyse (by simulation and/or by theoretical investigations) the following related model in terms of the degree distribution and structure. There are $n$ vertices and each vertex has a fixed number of out-edges, say $d \geq 2$ many. These out-edges choose their end-vertex uniformly at random, thus the in-degree of any vertex $v$ is random and equals the number of out-edges (out of $dn$ many) that have chosen $v$ as their end-vertex.

  • Determine the asymptotic in-degree distribution of this model, as $n \to \infty$. To do this,

    ★ Calculate the probability that a vertex has in-degree $k$. Using the convergence of Binomial to Poisson random variables, show that this probability converges for each fixed $k$.

    ★ Apply first and second moment method on the proportion of vertices with in-degree $k$ to show the desired convergence.

  • Identify the proportion of vertices in each part of the bowtie structure of the connected components in this model. To do this,

    ★ Describe the offspring distribution of the approximating forward and backward branching processes, similarly as described in the Copper and Frieze paper, but now apply it for the branching process given by the degree distributions here.

    ★ Calculate the survival probability of the forward and backward branching processes.
★ Using these, identify the asymptotic proportion of vertices that will be in the in-part, and the out-part of the bowtie as well as in the strongly connected component.

• Confirm your calculations by simulation and statistical analysis.

What you can further do if you have time:
- You may numerically investigate more advanced properties of the above mentioned model, for example average graph distances and sizes of connected components.
- You may randomize the out-degree distribution. Instead of each vertex having out-degree $d$, the out-degrees can be independent and identically distributed from some distribution. The most interesting case is when the out-degree distribution follows a power-law, i.e.,

$$P(\text{deg}(v) > k) = \frac{C}{k^{\tau - 1}}$$

for some $k$ and $\tau > 2$. Do you see an abrupt change in e.g. distances in the graph at/around $\tau = 3$?