Mathematical Statistics
Stockholm University
Research Report **2017:19**,
http://www.math.su.se

The tail does not determine the size of the giant

Maria Deijfen Sebastian Rosengren Pieter Trapman

October 2017

Abstract

The size of the giant component in the configuration model is given by a well-known expression involving the generating function of the degree distribution. In this note, we argue that the size of the giant is not determined by the tail behavior of the degree distribution but rather by the distribution over small degrees. Upper and lower bounds for the component size are derived for an arbitrary given distribution over small degrees $d \leq L$ and given expected degree, and numerical implementations show that these bounds are very close already for small values of L. On the other hand, examples illustrate that, for a fixed degree tail, the component size can vary substantially depending on the distribution over small degrees. Hence the degree tail does not play the same crucial role for the size of the giant as it does for many other properties of the graph.

Keywords: Configuration model, component size, degree distribution.

AMS 2010 Subject Classification: 05C80.

1 Introduction and results

The configuration model is one of the simplest and most well-known models for generating a random graph with a prescribed degree distribution. It takes a probability distribution with support on the non-negative integers as input and gives a graph with this degree distribution as output. The model is very well studied and there are precise answers to most questions concerning properties of the model such as the threshold for the occurrence of a giant component [13, 11], the size of the largest component [14, 11], diameter and distances in the supercritical regime [7, 8, 9], criteria for the graph to be simple [10] etc; see [5, Chapter 7] and [6, Chapters 4-5] for detailed overviews. Empirical networks often exhibit power law distributions, that is, the number of vertices with degree d decays as an inverse power of d for large degrees. For this reason, there has been a lot of attention on properties of the configuration model with this type of degree distribution. Here we focus on the size of the largest component in the supercritical regime as a functional of the degree distribution. Our main message is that the size of the largest supercritical component is not determined by the tail behavior of the degree distribution, but by the distribution over small degrees. While this is not surprising, in view of the general focus on degree tails in the literature, we think it deserves to be pointed out and elaborated on.

The model and its phase transition

To define the model, fix the number n of vertices in the graph and let $F = \{p_d\}_{d\geq 0}$ be a probability distribution with support on the non-negative integers. Assign a random number D_i of half-edges independently to each vertex $i = 1, \dots, n$, with $D_i \sim F$. If the total number of half-edges is odd, one extra half-edge is added to a uniformly chosen vertex. Then pair half-edges uniformly at random to create edges, that is, first pick two half-edges uniformly at random and