Todo list

Degree distribution in GCC

Benoît Richard Guiyuan Shi

June 18, 2018

Abstract

Write abstract

1 Introduction

Studying the fundamental properties of networks require to be able to abstract from the particular examples found in nature. This is usually done [?] by using a random model for the network generation and averaging the properties of interest over the set of possible networks. One common model is the configuration model [?] that allows to uniformly sample the space of all networks with a given degree distribution [?]. However, many real examples of networks intrinsically need to be connected, as for example the World Wide Web or railroad networks, but no model known to us allows to sample the space of all connected networks of a given degree distribution.

Missing ref

Missing ref

Missing ref

Find other examples

A way to still study connected networks is to only consider the Giant Connected Components (GCC) of networks generated using the configuration model [?]. We study here how this method implies bias on the degree distribution of the GCC.

Missing ref

2 Degree distribution in GCC

Per Bayes theorem we have for two event A and B

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = P(B|A)\frac{P(A)}{P(B)}. \tag{1}$$

{Bayes theorem}

This allows us to compute the degree distribution of the vertices in the giant connected component

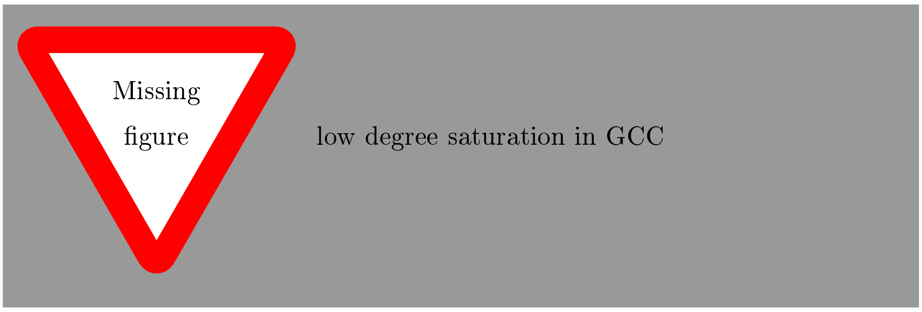
$$\begin{aligned} P(deg(v) = k | v \in GCC) &= P(v \in GCC | deg(v) = k)\frac{P(deg(v) = k)}{P(v \in GCC)} \qquad (2) \\ &= (1 - P(v \notin GCC | deg(v) = k))\frac{p_k}{S} \qquad (3) \\ &= \frac{p_k}{S}(1 - u^k)., \qquad (4) \end{aligned}$$

{Degree distribution in GCC}

where S is the probability that a random node is part of the GCC, p_k is the probability that a node has degree k and u is the probability that a node reached by following an edge of the network is not part of the GCC.

More details for the last step

In the context of the configuration model we choose the probabilities p_k. We now introduce the generating function for the degree distribution



g_0 and the generating function for excess degree distribution g_1, defined as

$$g_0(z) = \sum_{k=0}^{\infty} p_k z^k \tag{5}$$

$$g_1(z) = \frac{g_0'(z)}{g_0'(1)}, \tag{6}$$

where the prime notation is used for derivative with respect to z. They allow to determine the quantities S and u as [?]

Missing ref

$$u = g_1(u) \tag{7}$$

$$S = 1 - g_0(u). \tag{8}$$

By multiplying this expression by z^k for each k and summing, we find the generating function $G_0(z)$ of the degree distribution in the giant connected component

$$G_0(z) = \frac{1}{S}(g_0(z) - g_0(uz)). \tag{9}$$

3 Discussion