Discrete Mathematics 341 (2018) 3165–3173

Contents lists available at ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

On the diameter of Kronecker graphs

Justyna Banaszak, Tomasz Łuczak *

Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań, Poland

ARTICLE INFO

Article history:
Received 4 May 2017
Received in revised form 5 June 2018
Accepted 26 July 2018

Keywords:
Random graphs
Kronecker graph
Diameter
Connectivity

ABSTRACT

We prove that a.a.s. as soon as a Kronecker graph becomes connected it has a finite diameter.

1. Introduction

A Kronecker graph is a random graph with vertex set $V = \{0, 1\}^n$, where the probability that two vertices $u, v \in V$ are adjacent strongly depends on the structure of the vectors $u = (u_1, \ldots, u_n)$, and $v = (v_1, \ldots, v_n)$. More specifically, let $\mathbf{P}$ be a symmetric matrix

$$\mathbf{P} = \begin{array}{c} \\ 1 \\ 0 \end{array} \begin{array}{c} \begin{array}{cc} 1 & 0 \end{array} \\ \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} \end{array},$$

where zeros and ones are labels of rows and columns of $\mathbf{P}$, $\alpha, \beta, \gamma \in [0, 1]$, and $\alpha \geq \gamma$. In the Kronecker graph $\mathcal{K}(n, \mathbf{P})$ two vertices $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n) \in V = \{0, 1\}^n$ are adjacent with probability

$$p_{u,v} = \prod_{i=1}^{n} \mathbf{P}[u_i, v_i],$$

independently for each such pair.

Kronecker graphs were introduced by Leskovec, Chakrabarti, Kleinberg and Faloutsos in [2] to model some real world networks (see also [1,3,7]). Since then they have been studied by several authors but their properties are still far from being well understood (see [4] and references therein). In particular, Radcliffe and Young [9] determined the exact threshold for the property that $\mathcal{K}(n, \mathbf{P})$ is connected, supplementing a slightly weaker result of Mahdian and Xu [8].

* Corresponding author.
E-mail addresses: tabor@amu.edu.pl (J. Banaszak), tomasz@amu.edu.pl (T. Łuczak).

https://doi.org/10.1016/j.disc.2018.07.026

Theorem 1.

$$\lim_{n\to\infty} \mathbb{P}(\mathcal{K}(n,\mathbf{P}) \text{ is connected}) = \begin{cases} 0 & \text{if } \beta+\gamma=1,\ \beta\neq 1 \\ 0 & \text{if } \beta=1,\ \alpha=\gamma=0 \\ 1 & \text{if } \beta=1,\ \alpha>0 \text{ and } \gamma=0 \\ 1 & \text{if } \beta+\gamma>1. \end{cases}$$

The main result of this work states that as soon as $\mathcal{K}(n, \mathbf{P})$ becomes connected its diameter is bounded by a constant.

Theorem 2. *If either $\beta + \gamma > 1$, or $\beta = 1, \alpha > 0$ and $\gamma = 0$, then there exists a constant $a = a(\alpha, \beta, \gamma)$ such that a.a.s. $\mathrm{diam}(\mathcal{K}(n, \mathbf{P})) \le a$.*

2. The idea of the proof

In order to sketch our argument let us recall how one shows that the diameter is bounded from above for the binomial random graph model $G(N, p)$, and for many other random graph models. Typically, since random graphs are good expanders, it is proven first that for some small k the k-neighbourhood of each vertex is much larger than $\sqrt{N}$. Then, in the second part of the proof, one argues that since two random subsets of vertices of size larger than $\sqrt{N}$ intersect with large probability, each pair of vertices is a.a.s. connected by a path of length at most $2k$. However, in our case this procedure fails completely. The main reason is that most neighbours of a given vertex v have a similar structure, and so the events '$x \sim v$' and '$y \sim v$' are strongly correlated. Thus, the k-neighbourhood of a given vertex is very far from being a random subset, which is crucial for the second step of the procedure. Even more importantly, we do not understand expanding properties of $\mathcal{K}(n, \mathbf{P})$ and it is hard to control how fast the k-neighbourhoods of a vertex $\mathcal{K}(n, \mathbf{P})$ grow, which in most of the other random graph models is quite easy to investigate.

In [8], the diameter of $\mathcal{K}(n, \mathbf{P})$ is studied for $\gamma \le \beta \le \alpha$. For this specific range of parameters the probability of appearance of an edge of $\mathcal{K}(n, \mathbf{P})$ grows with the weights of its ends, i.e. for every two vertices u, v the probability that there exists an edge uv is always greater than the probability of an edge uv', whenever v has greater weight than v'. Using this fact the authors of [8] bounded from above the diameter of $\mathcal{K}(n, \mathbf{P})$ using well known bounds for the diameter of binomial random graphs.

To handle the difficulties related to the dependence of edges in $\mathcal{K}(n, \mathbf{P})$ in the general case we use the following approach. We consider two vertices, v and u which are 'similar' to each other (more specifically, we choose both of them from the middle layer of the n-cube and assume that they have small Hamming distance from each other). Then we generate their neighbourhoods at the same time until, for some k, we observe that the k-neighbourhood of v does not expand according to its expected rate. This is because many, in fact most, candidates for $(k + 1)$-neighbours of v have already been placed in the i-neighbourhood of v for some $i \le k$. However, the chance that a vertex x is in the i-neighbourhood of v is roughly the same as the probability that x is in the i-neighbourhood of u so, if most potential $(k + 1)$-neighbours of v are already in its kth neighbourhood, many of them are also in the kth neighbourhood of u. Consequently, there is a path of length at most $2k$ joining v and u.

The structure of the paper is the following. First we treat a special 'pathological' case $\beta = 1$. Then we present the crucial part of our argument showing that the subgraph induced in $\mathcal{K}(n, \mathbf{P})$ by its middle layer has a.a.s. a small diameter. Finally, we complete the proof showing that a.a.s. each vertex of $\mathcal{K}(n, \mathbf{P})$ is connected to the middle layer by a short path.

3. Case $\beta = 1$

In this section we show that if $\beta = 1, \alpha > 0$, and $\gamma = 0$, then the diameter of $\mathcal{K}(n, \mathbf{P})$ is a.a.s. bounded by a constant. This set of parameters α, β, γ is somewhat special as it is the only case, when $\beta + \gamma = 1$ and still $\mathcal{K}(n, \mathbf{P})$ is a.a.s. connected.

We introduce some notation, which we shall use throughout the paper. By $d(v, u)$ we denote the Hamming distance between two vertices v and u and $w(v)$ stands for the weight of a vertex $v = (v_1, \dots, v_n)$, i.e. the number of ones in its label that is

$$w(v) = \sum_{i=1}^{n} v_i.$$

For a vertex $v = (v_1, \dots, v_n)$, we set $\bar{v} = (1 - v_1, \dots, 1 - v_n)$.

Now let us go back to the case $\beta = 1$. Note that

$$\mathbb{P}(v \sim \bar{v}) = \beta^n = 1,$$

and observe that either v or $\bar{v}$ has weight at least $n/2$. Thus, to show the assertion it is enough to verify that a.a.s. there exists a path of bounded length between every pair of vertices in R defined as

$$R = \{v \in V : w(v) \ge n/2\}.$$