k-CONNECTIVITY IN RANDOM UNDIRECTED GRAPHS*

John H. REIF

Aiken Computation Laboratory, Harvard University, Cambridge, MA 02138, U.S.A.

Paul G. SPIRAKIS

Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, U.S.A.

This paper concerns vertex connectivity in random graphs. We present results bounding the cardinality of the biggest k-block in random graphs of the $G_{n,p}$ model, for any constant value of k. Our results extend the work of Erdös and Rényi and Karp and Tarjan. We prove here that $G_{n,p}$, with $p \geq c/n$, has a giant k-block almost surely, for any constant $k > 0$. The distribution of the size of the giant k-block is examined. We provide bounds on this distribution which are very nearly tight. We furthermore prove here that the cardinality of the biggest k-block is greater than n-log n, with probability at least $1 - 1/(n^2 \log n)$, for $p \geq c/n$ and $c > k+3$.

1. Introduction

A graph $G = (V, E)$ consists of a finite nonempty set V of vertices together with a prescribed set E of unordered pairs of distinct elements of V (set of edges). (We allow no loops nor multiple edges.) The vertex connectivity $k(G)$ of an undirected graph G is the minimum number of vertices whose removal results in a disconnected graph or a trivial graph (consisting of just one vertex). Note that we follow here [6] in defining k-connectivity, which we find to be most natural. MacLane [8] gives a (somewhat different) definition of triconnectivity so that he can have the theorem that a graph is planar if its triconnected components are. MacLane [8] shows that his triconnected components are homeomorphic to 3-blocks. Vertex k-connectivity seems to be a fundamental property of a graph and has numerous applications to other graph problems (such as planarity testing, routing problems etc.) It is relevant to questions concerning vulnerability of a graph to separation. Cluster analysis methods considering the nature and inherent reliability of proximity data use the theory of k-connectivity to find groups of likes and dislikes in object pair association graphs ([6, 7] also [4]).

A k-block of an undirected graph G is a maximal k-connected subgraph. A k-block is trivial if it has only one vertex [6]. Clearly, each k-block consists of $\geq k$ vertices or it is trivial.

* This work was supported in part by the National Science Foundation Grants NSF-MCS 82-00269 and NSF-MCS 83-00630 and the Office of Naval Research Contract N00014-80-C-0674.

Preprint of paper appearing in Discrete Mathematics, Volume 54, Issue 2, April 1985, Pages 181-191.

Matula [6] examined certain properties of k-blocks in graphs (number of them, separation lemma) and Erdös and Renyi [2] and Karp and Tarjan [5] examined the distribution of the size of the biggest 1 and 2-blocks in random graphs $G_{n,p}$ with $p \geq c/n$ and $G_{n,N}$ with $N \geq cn$. They proved that there is a giant k-block for $k = 1, 2$, with exponentially decaying probability of error. For $p > (\log n)/n$ Erdös, Rényi [3] showed that $G_{n,p}$ becomes almost surely k-connected, for $k > 0$.

In our pap . we examine k-connectivity in the model $G_{n,p}$, defined precisely as follows: For $0 \leq p \leq 1$ and $n \geq 0$ let $G_{n,p}$ be a random variable whose values are graphs on the vertex set $\{1, 2, \ldots, n\}$. If $e = \{u, v\}$ and $u, v \in \{1, 2, \ldots, n\}$ then Prob$\{e$ is an edge$\} = p$ and these probabilities are independent for different e.

We prove that for each constant $k > 0$ and each ε $(0 \leq \varepsilon \leq 1)$ there is a $c(k, \varepsilon) > k$ such that, for $p \geq c/n$, $G_{n,p}$ has a k-block of cardinality at least εn, with probability tending to 1 as n tends to infinity. Furthermore, for any $\alpha > 1$, $k > 0$ and ε on $(0, 1]$ there is a $c(k, \varepsilon, \alpha) > \alpha k$ such that, for $p \geq c/n$, the above probability is at least $1 - e^{-\alpha n}$.

We also prove that for any constant $k > 0$ and any $m < n/2k$, there are constants $c(k) > k$ and $d(k) > 0$ such that the cardinality of the biggest k-block of $G_{n,p}$, with $p \geq (c(k) \log n)/n$ is equal to $n - m$ with probability at least $1 - n^{-md(k)}$. A corollary of this result and the giant k-block result is the fact that, for such p, $G_{n,p}$ becomes almost surely k-connected (this was proved in [3]). For p above the threshold value, we get tight bounds on the probability estimates.

Finally, we prove that for any $m \leq \sqrt{n}$ and any constant $k > 0$, there is a constant $c_1(k) > 1 + k$ and a function $t(n) > (c_1(k) \log n)/m$, such that, if $p > t(n)/n$, then the cardinality of the biggest k-block of $G_{n,p}$ is greater than $n - m$ with probability at least $1 - 1/mn^{c_1(k)-1-k}$. A corollary is that if $p \geq c_1(k)/n$ with $c_1(k) > k + 3$, then $G_{n,p}$ has a k-block of cardinality greater than $n - \log n$ with probability at least $1 - 1/(n^2 \log n)$.

2. Properties of k-blocks

Proposition 1 (Matula [6]). *For each $k \geq 0$, any two k-blocks have no more than $k - 1$ vertices in common.*

Definition (Matula [6]). A separation set S of G is a vertex subset $S \subseteq V(G)$ such that $G - S$ is disconnected. A minimum separating set $S \subset V(G)$ has $|S| = k(G)$.

Definition. Let G be a graph (V, E) and let $S \subseteq V$ be a set of vertices. Then by $\langle S \rangle$ we denote the subgraph induced by S on G.

Lemma 1 (Matula [6], Block separation lemma). *Let $S \subseteq V(G)$ be a minimum separating set of the noncomplete graph G with $\langle A_1 \rangle, \langle A_2 \rangle, \ldots, \langle A_m \rangle$, $m \geq 2$ the components of $G - \langle S \rangle$ and let $k \geq k(G) + 1$. Then each k-block of G is a k-block of*

$\langle A_i \cup S\rangle$ *for precisely one value of* i, *and each* k*-block of* $\langle A_i \cup S\rangle$ *for every* i *is a* k*-block of* G.

For a proof, see [6].

Remark. Matula [6] shows that for each $k \geq 1$ the total number of nontrivial k'-blocks for $1 \leq k' \leq k$, is $\leq \lfloor (2n-1)/3 \rfloor$ for any graph G with n vertices.

3. Giant k-blocks in random graphs

In the following we introduce special notation for very large subgraphs. For each ε, $0 \leq \varepsilon \leq 1$, a subgraph H of a graph G of n vertices is called an ε*-giant of* G if the cardinality of the vertex set of H is $\geq \varepsilon n$.

Definition. Given a vertex set $S \subseteq V$ in the graph $G=(V, E)$, the *boundary vertices* of S are the set $B(S)=\{u \in S \mid \exists v \in V-S \text{ such that } \{u, v\} \in E\}$.

Definition. Let X be a random variable whose values are the cardinality of the maximum k-block of instances of $G_{n,p}$. Let $F_{n,p,k}(a)=\text{Prob}\{X \leq a\}$ be the distribution function of X.

Definition. If $G=(V, E)$ and A, B are subsets of V, then

$$\text{CROSS}(A, B)=\{e=\{u, v\} \in E \mid u \in A \text{ and } v \in B\}.$$

Lemma 2. *For any* $\alpha_1, \varepsilon_1, \varepsilon_2>0$ *where* $\varepsilon_1+\varepsilon_2 \leq 1$ *and* $\alpha_1 \geq 1$ *there are constants* $c, \varepsilon_3, \varepsilon_4>0$ *such that a random graph* $G_{n,p}$ *with* $p \geq c/n$ *has the property* (*) *with probability* $\geq 1-e^{-\alpha_1 n}$.

If A, B *are any two vertex subsets of* V *such that* $|A| \geq \lfloor \varepsilon_1 n \rfloor$, $|B| \geq \lfloor \varepsilon_2 n \rfloor$ *and* $A \cap B=\emptyset$ *then* $|\text{CROSS}(A, B)|>0$. (*)

Proof. The complement of (*) is the event I: "There are two vertex subsets A, B such that $|A| \geq \lfloor \varepsilon_1 n \rfloor$, $|B| \geq \lfloor \varepsilon_2 n \rfloor$ and $A \cap B=\emptyset$ and $|\text{CROSS}(A, B)|=0$."

Clearly

$$\text{Prob}\{\text{CROSS}(A, B)=\emptyset\} \leq (1-p)^{\varepsilon_1 n \varepsilon_2 n} \leq \left(\left(1-\frac{c}{n}\right)^n\right)^{\varepsilon_1 \varepsilon_2 n} \leq e^{-c \varepsilon_1 \varepsilon_2 n}.$$

The number of ways to select these A, B is

$$n_{AB} \leq \binom{n}{\varepsilon_1 n}\binom{n-\varepsilon_1 n}{\varepsilon_2 n}$$

(assume for simplicity here that $\varepsilon_1 n,\ \varepsilon_2 n \in N$). Note that [9]

$$\binom{n}{k} \sim \left(\frac{n}{2\pi k(n-k)}\right)^{1/2} \left(\frac{n}{k}\right)^k \left(\frac{n}{n-k}\right)^{n-k}$$

leading to

$$n_{AB} \leqslant \frac{1}{(4\pi^2 \varepsilon_1 \varepsilon_2 (1-\varepsilon_1)(1-\varepsilon_2) n^2)^{1/2}} \left(\frac{1}{\varepsilon_1}\right)^{\varepsilon_1 n} \left(\frac{1}{\varepsilon_2}\right)^{\varepsilon_2 n} \left(\frac{1}{1-\varepsilon_1}\right)^{(1-\varepsilon_1)n} \left(\frac{1}{1-\varepsilon_2}\right)^{(1-\varepsilon_2)n}$$

Let

$$\gamma(\varepsilon_1, \varepsilon_2) = \frac{1}{\varepsilon_1^{\varepsilon_1} \varepsilon_2^{\varepsilon_2} (1-\varepsilon_1)^{1-\varepsilon_1} (1-\varepsilon_2)^{1-\varepsilon_2}}.$$

Clearly $\gamma(\varepsilon_1, \varepsilon_2) \leqslant 4$ and $\gamma(\varepsilon_1, \varepsilon_2) = 4$ when $\varepsilon_1 = \varepsilon_2 = \frac{1}{2}$. So

$$n_{AB} \leqslant \frac{1}{2\pi n} \frac{1}{\sqrt{\beta(\varepsilon_1, \varepsilon_2)}} \gamma^n(\varepsilon_1, \varepsilon_2) \tag{1}$$

where

$$\beta(\varepsilon_1, \varepsilon_2) = \varepsilon_1 \varepsilon_2 (1-\varepsilon_1)(1-\varepsilon_2) \leqslant 4. \tag{2}$$

Clearly, since $\varepsilon_1, \varepsilon_2 \geqslant 1/n$, we have

$$\beta(\varepsilon_1, \varepsilon_2) \geqslant \frac{1}{n^2} + \frac{1}{n^4} - \frac{2}{n^3}. \tag{3}$$

It is straightforward to show that

$$\frac{1}{\sqrt{\beta(\varepsilon_1, \varepsilon_2)}} \leqslant n \frac{1}{\sqrt{1-2/n}} \leqslant n\sqrt{2} \quad \text{for } n \geqslant 4.$$

We then get

$$\begin{aligned}
\text{Prob}\{I\} &\leqslant \sum_{\text{all } A,B} \text{Prob}\{\text{CROSS}(A, B) = \emptyset\} \\
&\leqslant \frac{\sqrt{2}}{2\pi} 4^n e^{-c\varepsilon_1 \varepsilon_2 n} \quad \text{(from (1), (2), (3))} \\
&\leqslant \frac{\sqrt{2}}{2\pi} (4e^{-c\varepsilon_1 \varepsilon_2})^n \leqslant \frac{\sqrt{2}}{2\pi} e^{-\alpha_1 n}
\end{aligned}$$

for

$$c \geqslant \frac{\alpha_1 + \log_e 4}{\varepsilon_1 \varepsilon_2}. \quad \square$$

Now we can prove

Theorem 1. *For every ε on $(0, 1)$ and $k > 0$ there is a $c = c(\varepsilon, k) > 0$ such that, for $p > c/n$, $\lim_{n \to +\infty} F_{n,p,k}(\varepsilon n) = 0$. In other words, the random graph $G_{n,p}$ with $p \geqslant c/n$ has an ε-giant k-block with probability tending to 1 as n tends to $+\infty$.*

Proof. Let $G=(V,E)$ be an instance of the random graph $G_{n,p}$. Let E_1 be the event "G has no ε-giant k-block". Assume event E_1 be true in the instance G of $G_{n,p}$. Let initially the set $A=\emptyset$. Do the following construction just until A has cardinality $\geqslant \varepsilon'(n/2)$, where $\varepsilon'=\min(\varepsilon, 1-\varepsilon)$.

(a) Find a minimum separating set S of G. Let $\langle A_1\rangle,\ldots,\langle A_m\rangle$, $m\geqslant 2$ be the components of $G-S$. Let $\langle A_i\rangle$ be the smallest of them. Let A be $(A_i\cup S)\cup A$. Let B be the union of the rest of the components and let G be the graph induced by $B\cup S$. If $|A|<\varepsilon'\cdot\frac{1}{2}n$, then go to (a).

By the above method of constructing A, each addition of a component in A adds at most $k-1$ vertices to $B(A)$ (i.e. the vertices of the cut) and at least one vertex to $A-B(A)$ (by the block separation lemma and by the fact that k-blocks have $\geqslant k$ vertices if they are nontrivial) or causes the transformation of a boundary to a nonboundary vertex. Thus, at least $1/k$ of the vertices of A are not in $B(A)$.

By this construction, finally the k-blocks of G are going to be separated. Because all k-blocks have been assumed to have cardinality $\leqslant \varepsilon n$, we will finally have

$$\varepsilon'\frac{n}{2}\leqslant |A|\leqslant \min\left[\varepsilon'\frac{n}{2}+\varepsilon n, \frac{3n\varepsilon'}{4}\right].$$

So

$$|A-B(A)|\geqslant \frac{\min(\varepsilon, 1-\varepsilon)}{2k}\cdot n$$

and

$$|V-A|\geqslant n\left(1-\min\left[\left(\varepsilon+\frac{\varepsilon'}{2}\right), \left(\frac{3\varepsilon'}{4}\right)\right]\right).$$

Let $Y=A-B(A)$ and $Z=V-A$. Then $|Y|\geqslant \varepsilon_1 n$ and $|Z|\geqslant \varepsilon_2 n$, where $\varepsilon_1=\varepsilon'/2k$ and for ε_2 it is straightforward to show that

$$1-\varepsilon_2=\begin{cases}\dfrac{3\varepsilon}{4} & \text{if } \varepsilon\leqslant\frac{1}{2},\\[2ex] \dfrac{3-3\varepsilon}{4} & \text{if } \varepsilon>\frac{1}{2}.\end{cases}$$

Finally

$$\text{if } \varepsilon\leqslant\tfrac{1}{2} \text{ then } \varepsilon_1=\frac{\varepsilon}{2k},\ \varepsilon_2=1-\frac{3\varepsilon}{4}$$

$$\text{else } \varepsilon_1=\frac{1-\varepsilon}{2k},\ \varepsilon_2=\frac{1}{4}+\frac{3\varepsilon}{4}.$$

Also, $\text{CROSS}(Y,Z)=\emptyset$, by construction. Let E_2 be the above event. We have just shown that E_1 implies E_2. So,

$$\text{Prob}\{E_1\}\leqslant \text{Prob}\{E_2\}$$

and, by using the proof of Lemma 1

$$\text{Prob}\{E_2\} \leqslant \frac{\sqrt{2}}{2\pi}(\gamma(\varepsilon_1, \varepsilon_2)e^{-c\varepsilon_1\varepsilon_2})^n. \tag{4}$$

Clearly, when $\ln \gamma(\varepsilon_1, \varepsilon_2) < c\varepsilon_1\varepsilon_2$ then the right-hand side of (4) tends to zero as $n \to +\infty$. So, for $c > \ln \gamma(\varepsilon_1, \varepsilon_2)/(\varepsilon_1\varepsilon_2)$ we get $\text{Prob}\{E_1\} \to 0$ as $n \to +\infty$. □

Note. The constant c must be greater than

$$c_{\text{th}} = \frac{\ln \gamma(\varepsilon_1, \varepsilon_2)}{\varepsilon_1\varepsilon_2} = -\frac{\ln \varepsilon_1}{\varepsilon_2} - \frac{\ln \varepsilon_2}{\varepsilon_1} - \frac{(1-\varepsilon_1)}{\varepsilon_1\varepsilon_2}\ln(1-\varepsilon_1) - \frac{(1-\varepsilon_2)}{\varepsilon_1\varepsilon_2}\ln(1-\varepsilon_2).$$

Clearly $c_{\text{th}}(\varepsilon_1, \varepsilon_2) = c_{\text{th}}(1-\varepsilon_1, 1-\varepsilon_2)$. From [1] it is known that the threshold for a giant connected component ($k=1$) is $1/n$, i.e. c must be >1. Let us find $\min c_{\text{th}}(\varepsilon_1, \varepsilon_2)$ for $\varepsilon_1 + \varepsilon_2 \leqslant 1$, $\varepsilon_1, \varepsilon_2 \geqslant 1/n$, and $k=1$. Clearly $c_{\text{th}}(\varepsilon_1, \varepsilon_2) > 1$ and $\varepsilon_1 \to 0$, $\varepsilon_2 \to 1$ gives us $c_{\text{th}}(\varepsilon_1, \varepsilon_2) \to 1$. So, our threshold matches the threshold of [1] for $k=1$. In general, $(k \geqslant 1)$, $c > k$.

Corollary 1. *For every ε on $(0, 1)$, $\alpha > 1$ and $k > 0$ there is a $c = c(k, \varepsilon, \alpha) > 0$ such that, for $p \geqslant c/n$, $F_{n,p,k}(\varepsilon n) \leqslant e^{-\alpha n}$.*

Proof. Just take $c \geqslant (\alpha + \ln \gamma(\varepsilon_1, \varepsilon_2))/\varepsilon_1\varepsilon_2$ in (4) of the proof of Theorem 1. We get $\text{Prob}\{E_1\} \leqslant \text{Prob}\{E_2\} \leqslant e^{-\alpha n}$. □

Remark. In the proof of Theorem 1, note that $|Y| < (1-\varepsilon_2)n$ and $|Z| < (1-\varepsilon_1)n$. Then, the probability that $\text{CROSS}(Y, Z) = \emptyset$ is greater than $(1-p)^{(1-\varepsilon_1)(1-\varepsilon_2)n^2}$. Also, the number of ways to select such Y, Z is

$$n_{yz} \geqslant \binom{n}{(1-\varepsilon_2)n}\binom{n-(1-\varepsilon_2)n}{(1-\varepsilon_1)n}.$$

By the proof of Lemma 2 then, the probability that there are two such sets is at least

$$\frac{\sqrt{2}}{2\pi}(\gamma(1-\varepsilon_1, 1-\varepsilon_2)e^{-c(1-\varepsilon_1)(1-\varepsilon_2)})^n$$

leading to the conclusion that such sets almost surely exist when $c < \ln \gamma(1-\varepsilon_1, 1-\varepsilon_2)/(1-\varepsilon_1)(1-\varepsilon_2)$. But $\gamma(1-\varepsilon_1, 1-\varepsilon_2) = \gamma(\varepsilon_1, \varepsilon_2)$ and, when $\varepsilon_1 + \varepsilon_2 = 1$ then $(1-\varepsilon_1)(1-\varepsilon_2) = \varepsilon_1\varepsilon_2$. So, we conclude:

Corollary 2. *The constant $c_{\text{th}} = \ln \gamma(\varepsilon_1, \varepsilon_2)/(\varepsilon_1\varepsilon_2)$ is optimal.*

4. *k*-blocks of dense random graphs

This section considers random graphs with edge probability $p \geqslant c \log n/n$. Erdös and Rényi [3] gave the threshold probability for a random graph of the $G_{n,N}$

model to be k-connected. That threshold, translated to $G_{n,p}$ implies that

$$p_{\text{th}} = \frac{1}{n}\left(\log n + \frac{k}{2}\log\log n + \alpha + o\left(\frac{1}{n}\right)\right) \quad \text{(where } \alpha \text{ is a constant)}$$

is the threshold for $G_{n,p}$ to be k-connected. We examine here a related question, namely the distribution of the size of the biggest k-block of dense random graphs. This question has not been sufficiently answered in the past. Erdös and Rényi [3] used a powerful technique for proving tight bounds on the probability that a random graph is k-connected, namely, they locate suitable subgraphs not joined by an edge, and derive tight bounds on the probability of the existence of such subgraphs. We use essentially the same ideas here.

Theorem 2. *For any constant integer $k > 0$ and any n and $m < n/2k$ there are constants $c(k) > k$ and $d(k) > 0$ such that the cardinality X of the biggest k-block of the graph $G_{n,p}$ with $p \geq c(k)\log n/n$ satisfies*

$$\text{Prob}\{X = n - m\} \leq n^{-md(k)}.$$

Proof. Let G be an instance of $G_{n,p}$ and let the event $X = n - m$ be true in that instance. Let A be a k-block with $|A| = X$. For every $u \in V - A$ we have that

$$|\{\{u, v\} \in E(G) : v \in A\}| \leq k - 1$$

(else, u would belong to A). Let

$$A_1 = \{v \in A : \exists u \in V - A : \{u, v\} \in E(G)\}$$

then

$$|A_1| \leq (k-1)\,|V - A| = (k-1)m.$$

Let $A_2 = A - A_1$. We get

$$|A_2| \geq n - m - (k-1)m = n - km.$$

Furthermore, there is no edge from $V - A$ to A_2. Let E be this event. The probability of E is bounded above by

$$u(m, n) = \binom{n}{m}\binom{n-m}{n-km}(1-p)^{(n-km)m} \tag{5}$$

But

$$(1-p)^{(n-km)m} \leq \left(1 - \frac{c\log n}{n}\right)^{(n-km)m} \leq \left(1 - \frac{c\log n}{n}\right)^{(m(n-km)/\log n)\cdot\log n}$$
$$\leq e^{-cm\log n} = n^{-cm}. \tag{6}$$

Also, note that $\binom{n-m}{n-km} = \binom{n-m}{(k-1)m}$. Since $(k-1)m < \frac{1}{2}(n-m)$ we have $\binom{n-m}{(k-1)m} < (n-m)^{(k-1)m} < n^{(k-1)m}$. Also, since $m < \frac{1}{2}n$ we get $\binom{n}{m} < n^m$. We conclude

$$u(m, n) < n^{km} n^{-cm}. \tag{7}$$

Let $d(k) = c - k$. Clearly $d(k) > 0$ iff $c > k$. We thus have $\text{Prob}\{E\} \leq n^{-m\cdot d(k)}$ □

Theorem 3. *For any constant integer $k>0$ and any sufficiently large n, there are positive constants $c(k)$ and $d(k)$ such that the cardinality X of the biggest k-block of the graph $G_{n,p}$ with $p \geq c(k)\log n/n$ satisfies*

$$\text{Prob}\{X \leq n - \log n\} < 2n^{(1-d(k)\log n)}.$$

Proof. By using Theorem 2, we get

$$\text{Prob}\left\{\log n \leq n - X < \frac{n}{2k}\right\} = \sum_{m=\log n}^{n/2k} n^{-md(k)}$$

with $d(k) = c - k$, $c > k$. So,

$$\text{Prob}\left\{\log n \leq n - X < \frac{n}{2k}\right\} < n \cdot n^{-\log n \,.\, d(k)} < n^{1-d(k)\log n}.$$

Also, by Corollary 1 and using $\varepsilon = 1/2k$ we get $\text{Prob}\{n - X > n/2k\} < e^{-\alpha \cdot n}$ for any $\alpha > 1$ and $c(k) \geq (\alpha + \log_e \gamma(\varepsilon_1\varepsilon_2))/\varepsilon_1\varepsilon_2$ with $\varepsilon_1\varepsilon_2 = (1/(2k))(1-3/(8k))$. So, for $c(k) < \max(k, (\alpha + \log_e 4)/\varepsilon_1\varepsilon_2)$ or

$$c(k) > 16(\alpha + \log_e 4)k^2 \tag{8}$$

we get

$$\text{Prob}\{\log n \leq n - X\} < e^{-\alpha \cdot n} + n^{1-\log n \cdot d(k)}$$

or

$$\text{Prob}\{X \leq n - \log n\} < 2n^{(1-d(k)\log n)} \quad \text{for any sufficiently large } n. \quad \square$$

Note. Theorem 3 says that for $p \geq c(k)\log n/n$ the graph $G_{n,p}$ has a k-block of size $\geq n - \log n$ with probability tending to 1 as $n \to \infty$.

Theorem 4. *For any constant integer $k>0$ and n sufficiently large, there are constants $c(k) > k^2$ and $d'(k) = c(k) - 1 - k$ such that the random graph $G_{n,p}$ with $p \geq c(k)\log n/n$ is k-connected with probability at least $1 - 2n^{-d'(k)}$.*

Proof. Let $R = n - X$ where X is the cardinality of the biggest k-block of $G_{n,p}$. By using the previous two theorems, with

$$c(k) > 1 + \max\left(k, \frac{\alpha + \log_e \gamma(\varepsilon_1\varepsilon_2)}{\varepsilon_1\varepsilon_2}\right) \quad \text{with } \varepsilon_1\varepsilon_2 = \frac{1}{2k}\left(1 - \frac{3}{8k}\right),$$

we get that

$$\text{Prob}\{1 \leq R\} < e^{-\alpha \cdot n} + n^{1-(c-k)}.$$

Let $d'(k) = c - k - 1$. Then $d'(k) > 0$ for

$$c(k) > 2 + \max\left(k, \frac{\alpha + \log_e \gamma(\varepsilon_1, \varepsilon_2)}{\varepsilon_1\varepsilon_2}\right)$$

and

$$\text{Prob}\{1 \leq R\} \leq e^{-\alpha \cdot n} + n^{-d'(k)} < 2n^{-d'(k)}$$

for large n. Hence, $\text{Prob}\{R = 0\} > 1 - 2n^{-d'(k)}$. $\square$

Note. Erdös and Rényi [3] prove that (for the $G_{n,N}$ model) the threshold probability that $G_{n,N}$ is k-connected is asymptotically the same as the threshold probability that the mindegree of $G_{n,N}$ is k. This result can be easily shown to hold for the $G_{n,p}$ model too. Our proof of Theorem 3 (of which Theorems 4, 5 are corollaries) is based on the examination of vertices whose degree is $\leq k-1$. The optimality of the result of [3] for $G_{n,N}$ then, allows us to conjecture that the estimate of Theorem 2 is close to optimal.

5. k-blocks for intermediate edge densities

Let $c/n \leq p \leq c' \log n/n$. We wish to study the k-connectivity for random graphs of the $G_{n,p}$ model, with p as above. Note that the result of Theorem 5 below, does not translate easily to $G_{n,N}$. Below, $c' \leq 1$.

Theorem 5. *For any constant $k \geq 0$ and any $m \leq \sqrt{n}$ there is a constant $c_1(k) > 1 + c'k$ and a function $t(n) > (c_1(k) \log n)/m$ such that, if $p > t(n)/n$ then if X is the cardinality of the biggest k-block of $G_{n,p}$ then*

$$\text{Prob}\{X \leq n-m\} \leq \frac{1}{m}\frac{1}{n^{c_1-1-c'k}} \to 0 \quad \textit{as } n \to \infty.$$

Proof. Assume that in the instance G of $G_{n,p}$ the cardinality X of the biggest k-block satisfies the inequality $X \leq n-m$. Then, we can find two sets Y, Z (as in the proof of Theorem 3) such that $|Y| = m$, $|Z| = n-km$ and no edge between them. This event is above bounded by the probability $1-q$ where

$q = \text{Prob}\{$for every pair of disjoint sets Y, Z of vertices of the above sizes, there is at least one edge between Y, $Z\}$.

We shall show $q \to 1$ as $n \to \infty$. Let us enumerate all possible pairs of sets of vertices of the above sizes. Call them

$$(Y_1, Z_1), (Y_2, Z_2), \ldots, (Y_g, Z_g)$$

where

$$g = \binom{n}{m}\binom{n-m}{n-km} = \binom{n}{m}\binom{n-m}{(k-1)m}.$$

We have

$$q = \text{Prob}\{\text{CROSS}(Y_1, Z_1) \neq \emptyset \wedge \cdots \wedge \text{CROSS}(Y_g, Z_g) \neq \emptyset\}.$$

By Bayes's formula,

$$q = \text{Prob}\{\text{CROSS}(Y_1, Z_1) \neq \emptyset\}$$

$$\cdot \text{Prob}\left\{\frac{\text{CROSS}(Y_2, Z_2) \neq \emptyset}{\text{CROSS}(Y_1, Z_1) \neq \emptyset}\right\} \cdot \ldots \cdot \text{Prob}\left\{\frac{\text{CROSS}(Y_g, Z_g) = \emptyset}{\bigcap_{1 \leq i \leq g-1} \text{CROSS}(Y_i, Z_i) \neq \emptyset}\right\}.$$

We need the following enumeration lemma:

Lemma 3. *For every two sets Y_i, Z_i having at least one edge e between them, there are at least*

$$g_1 = \binom{n-2}{m-1}\binom{n-2-(m-1)}{(k-1)m-1}$$

pairs of sets of sizes m, $n-km$ which also have this edge between them.

This lemma can be proved easily by taking out the two vertices of e and enumerating.

Corollary 3. *There is a suitable enumeration of the sets in the q product such that for every term i not equal to 1 the next g_1 or more terms (conditioned on the existence of an edge from A_i to B_i) will be equal to 1.*

Hence, the value of q is

$$q \geq [\text{Prob}\{E(Y_1, Z_1) \neq \emptyset\}]^{g/g_1}$$

but

$$\frac{g}{g_1} = \left(\frac{n-1}{n-km}\right)\left(\frac{n}{m}\right) \quad \text{and} \quad \lim_{n\to+\infty} \frac{g}{g_1} = \frac{n}{m},$$

hence

$$q \geq (1-(1-p)^{m(n-km)})^{g/g_1}$$

or

$$q \geq (1-((1-p)^{1/p})^{pm(n-km)})^{g/g_1}$$

or

$$q \geq (1-e^{-pm(n-km)})^{g/g_1}$$

or

$$q \geq 1 - \frac{g}{g_1} e^{-m\cdot t(n)+pkm^2}$$

or

$$q \geq 1 - \left(\frac{n-1}{n-km}\right)\left(\frac{n}{m}\right) e^{-m\cdot t(n)+pkm^2}.$$

Since $p < c' \log n/n$ we have $pkm^2 < (c'/n)km^2 \log n$. For $m \leq \sqrt{n}$ we get $pkm^2 < c'k \log n$ and as n grows large, we get

$$q \geq 1 - \left(\frac{n}{m}\right) e^{-t(n)\cdot m + c'k \log n}$$

or

$$q \geq 1 - e^{\log n - \log m - mt(n) + c'k \log n}$$

or

$$q \geq 1 - \frac{1}{m} n^{-(c_1 - 1 - c'k)} \tag{9}$$

(because $mt(n) > c_1 \log n$). So, if $c_1 > c'k + 1$ we have

$$\text{Prob}\{X < n - m\} < \frac{1}{m} \cdot \frac{1}{n^{c_1 - 1 - c'k}} \to 0 \quad \text{as } n \to +\infty. \quad \square$$

Corollary 4. *For $m = \log n$ we get: For each $k > 0$, the graph $G_{n,p}$ with $c_1(k)/n \leq p \leq c' \log n/n$ and $c_1(k) > c'k + 3$, has a k-block of cardinality $> n - \log n$ with probability $\geq 1 - 1/(mn^2)$.*

6. Conclusion

We provided here nearly tight bounds on the probability distribution of the size of the biggest k-block in random graphs. We leave as an open problem the problem of determining k-blocks of a graph efficiently when the graph is random. This problem has been successfully answered, for $k = 1$ and 2, by Karp and Tarjan [5]. Algorithms for finding the k-blocks (for any k) when the graph is not necessarily random, have been discovered by Matula [7]. We believe that considerable improvements in the time efficiency of these algorithms can be made if the algorithms are adjusted to make a use of the structural properties of random graphs.

References

[1] P. Erdös and A. Rényi, On random graphs, Publ. Math. 6 (1959) 290–297.
[2] P. Erdös and A. Rényi, On the evolution of random graphs, Publ. Math. Inst. Hung. Acad. Sci. 5A (1960) 17–61.
[3] P. Erdös and A. Rényi, On the strength of connectedness of a random graph, Acta Math. Acad. Sci. Hung. 12 (1961), also in: J. Spenser, ed., The Art of Counting—Selected Writings (MIT Press, Cambridge, 1973).
[4] N. Jardine and R. Sibson, Mathematical Taxonomy (Wiley, London, 1971).
[5] R.M. Karp and R.E. Tarjan, Linear expected time algorithms for connectivity problems, Proc. 12th Annual ACM Symposium on Theory of Computing, Los Angeles, CA (1980) 368–377.
[6] D. Matula, k-blocks and ultrablocks in graphs, J. Combin. Theory Ser. B. 24 (1978) 1–13.
[7] D. Matula, Graph theoretic techniques for cluster analysis algorithms, in: F. Van Ryzon, ed., Classification and Clustering (Academic Press, New York, 1977) 95–129.
[8] S. MacLane, A structural characterization of planar combinatorial graphs, Duke Math. J. 3 (1937) 340–472.
[8] W. Feller, An Introduction to Probability Theory and its Applications, Vol. 1 (Wiley, New York, 1957).