COMBINATORICA 2(3) (1982) 223—228

LONG PATHS IN SPARSE RANDOM GRAPHS

Béla BOLLOBÁS

Dedicated to Tibor Gallai on his seventieth birthday

Received 8 April 1982

We consider random graphs with n labelled vertices in which edges are chosen independently and with probability c/n. We prove that almost every random graph of this kind contains a path of length $\geqq (1-\alpha(c))n$ where $\alpha(c)$ is an exponentially decreasing function of c.

Given $n\in\mathbf{N}$ and $0<p=p(n)<1$, the probability space $\mathcal{G}(n, P(\text{edge})=p)$ consists of all graphs with a fixed set of n labelled vertices, in which the edges are chosen independently and with probability p. As customary, we say that *almost every* (a.e.) *random graph* $G_{n,p}$ has a property Q if the probability that a graph $G\in\mathcal{G}(n, P(\text{edge})=p)$ has Q tends to 1 as $n\to\infty$.

Some important recent results in the theory of random graphs concern the existence of long paths in a.e. random graph $G_{n,c/n}$, where c is a constant. (For basic properties of graphs and random graphs see [2].) In order to formulate these results, for $c>0$ set

$$1-\alpha(c) = \sup\{\alpha\in\mathbf{R}\colon \text{ a.e. } G_{n,c/n} \text{ contains a path of length at least } \alpha n\},$$

$$1-\beta(c) = \sup\{\beta\in\mathbf{R}\colon \text{ a.e. } G_{n,c/n} \text{ contains a cycle of length at least } \beta n\}.$$

Clearly $0\leqq\alpha(c)\leqq\beta(c)\leqq 1$ for every c, and $\alpha(c)$, $\beta(c)$ are decreasing functions of c.

Erdős and Rényi [3] proved that in a.e. $G_{n,1/n}$ the largest component has $O(n^{2/3})$ vertices, so $\alpha(c)=1$ for $c\leqq 1$. Ajtai, Komlós and Szemerédi [1] proved the somewhat surprising result that $\alpha(c)<1$ for every $c>1$, and de la Véga [5] showed independently that $\alpha(c)\leqq c_0/c$ for every $c>0$ and some absolute constant c_0. Our aim is to show that this bound on $\alpha(c)$ can be replaced by an exponential function of c.

It is easily seen that $\alpha(c)$ cannot decay faster than an exponential function of c. Indeed, Erdős and Rényi [3] showed that for every $c>1$ a.e. $G_{n,c/n}$ is such that the order of its largest component is asymptotically $\left\{1-\frac{1}{c}\sum_{k=1}^{\infty} k^{k-1}(ce^{-c})^k/k!\right\}n$. Hence

$$\alpha(c) \geqq \frac{1}{c}\sum_{k=1}^{\infty} k^{k-1}(ce^{-c})^k/k! \geqq e^{-c}.$$

AMS subject classification (1980): 05 C 38, 60 C 05.

Theorem. *For* $c>2$ *a.e.* $G_{n,c/n}$ *contains a cycle of length at least* $(1-c^{24}e^{-c/2})n$. *In particular,*

$$\alpha(c) \leqq \beta(c) \leqq c^{24}e^{-c/2}.$$

Proof. We may assume that $c>200$ since if $c^{24}<e^{c/2}$, there is nothing to prove. Let $k, n\in \mathbf{N}$, $k<n$, and let V be a set of n labelled vertices. Denote by $\vec{\mathcal{G}}_{k\text{-out}}(n)$ the set of all directed graphs with vertex set V in which the outdegree of every vertex is k: $d^+(x)=k$ for every $x\in V$. Turn $\vec{\mathcal{G}}_{k\text{-out}}(n)$ into a probability space by giving all elements the same probability.

Given a directed graph $\vec{G}$, denote by $\theta(\vec{G})$ the graph having the same vertex set as $\vec{G}$, in which xy is an edge iff at least one of $\vec{xy}$ and $\vec{yx}$ is an edge of $\vec{G}$. Set $\mathcal{G}_{k\text{-out}}(n)=\{\theta(\vec{G})\colon \vec{G}\in\vec{\mathcal{G}}_{k\text{-out}}(n)\}$ and turn $\mathcal{G}_{k\text{-out}}(n)$ into a probability space by giving $G\in\mathcal{G}_{k\text{-out}}(n)$ the probability of the set $\theta^{-1}(G)\subset\vec{\mathcal{G}}_{k\text{-out}}(n)$.

Our proof is based on a recent deep theorem of Fenner and Frieze [4] stating that if $k\geqq 23$ then a.e. $G\in\mathcal{G}_{k\text{-out}}(n)$ is Hamiltonian. In order to make use of this result, we shall represent $\mathcal{G}(n, P(\text{edge})=c/n)$ as the image of a space of random directed graphs.

Define $r=d/n$, $0<r<1$, by

$$c/n = 2r-r^2.$$

Then clearly

$$c/2+c^2/8n < d < c/2+c^2/7n$$

if n is sufficiently large. Let $\vec{\mathcal{G}}(n, r)$ be the probability space of directed graphs with a fixed set V of n labelled vertices in which the edges are chosen independently and with probability r. Then in $\theta(G)$ the edges are chosen independently and with probability

$$1-(1-r)^2 = 2r-r^2=c/n,$$

so $\{\theta(\vec{G})\colon \vec{G}\in\vec{\mathcal{G}}(n, r)\}$, with probability inherited from $\vec{\mathcal{G}}(n, r)$, is just another representation of $\mathcal{G}(n, P(\text{edge})=c/n)$. Hence it suffices to show that a.e. $\vec{G}\in\vec{\mathcal{G}}(n, r)$ is such that $\theta(\vec{G})$ contains a cycle of length at least $(1-c^{24}e^{-c/2})n$.

We shall do this by showing that a.e. $\vec{G}\in\vec{\mathcal{G}}(n, r)$ is such that every vertex of a large set W dominates at least 23 vertices of W. Then the Fenner—Frieze theorem applied to the random graph $\vec{G}[W]$ will imply that a.e. $\vec{G}\in\vec{\mathcal{G}}(n, r)$ is such that $\theta(\vec{G})$ contains a cycle of length at least W. We shall look for an appropriate set W by omitting from V disjoint sets $U_0, U_1, \ldots, U_t$. The set U_0 will consist of all vertices of small outdegree (together with the vertices of too large outdegree, but this is only for the sake of convenience), then U_1 will be the set of vertices of small outdegree in $G-U_0$, and so on.

Given $x\in V$, the outdegree $d^+(x)$ of x in a random directed graph $\vec{G}_{n,r}$ has binomial distribution with parameters $n-1$ and $r=d/n$. Let us estimate $P(d^+(x)\leqq 24)$ and $P(d^+(x)\geqq 6d)$. Since $c>104$, if n is sufficiently large then

$$(1-d/n)^{n-1-i} < \exp\{-d+25\,d/n\} < \exp\{-c/2-c^2/8n+13c/n\} < e^{-c/2}$$

for $i<25$. Consequently

$$P(d^+(x) \leqq 24) = \sum_{i=0}^{24} \binom{n-1}{i} \left(\frac{d}{n}\right)^i (1-d/n)^{n-1-i} < e^{-c/2} \sum_{i=0}^{24} (c/2)^i/i!.$$

Also,

$$P(d^+(x) \geqq 6d) < \sum_{i=\lceil 6d \rceil}^{n-1} \binom{n-1}{i} \left(\frac{d}{n}\right)^i < \sum_{i=\lceil 6d \rceil}^{\infty} d^i/i! < \sum_{i=\lceil 6d \rceil}^{\infty} (e/6)^i < e^{-c}.$$

For $\vec{G} \in \vec{\mathscr{G}}(n, r)$ let

$$U_0 = U_0(\vec{G}) = \{x \in V:\ d^+(x) \leqq 24 \quad \text{or} \quad d^+(x) \geqq 6d\}.$$

Then $|U_0|$ has binomial distribution with parameters n and p, where $p=p(n)$ satisfies

$$p < e^{-c} + e^{-c/2} \sum_{i=0}^{24} (c/2)^i/i! = p_0.$$

Let us note the following simple property of the binomial distribution $S_{n,p}$ with parameters n and p: if $0<p<1/2$ and $(pn)^{-1/2}<\varepsilon \leqq 1/2$ then

$$P(|S_{n,p} - pn| \geqq \varepsilon pn) < e^{-\varepsilon^2 pn/3}. \tag{1}$$

By applying (1) we find that

$$P\left(|U_0| \geqq \frac{3}{2} p_0 n\right) < e^{-p_0 n/12} = P_0.$$

Now let us define disjoint sets $U_1, U_2, \ldots$, all functions of a random directed graph $\vec{G} \in \vec{\mathscr{G}}(n, r)$, as follows. Suppose $i \geqq 1$ and we have defined $U_0, U_1, \ldots, U_{i-1}$. Set

$$U_i = \left\{x \in V - \bigcup_{j=0}^{i-1} U_j:\ \left|\Gamma_+(x) \cap \bigcup_{j=0}^{i-1} U_j\right| \geqq 2\right\},$$

where

$$\Gamma_+(x) = \{y \in V:\ \vec{xy} \in \vec{E}(\vec{G})\}.$$

Let $D=9d^2$, $t=\left\lfloor \frac{2}{3} \log n / \log(1/Dp_0) \right\rfloor$ and

$$p_j = D^j p_0^{j+1}, \quad j = 1, 2, \ldots, t.$$

It can be checked that the assumptions we have made imply that $Dp_0 < 1/4$ and so $p_j < 4^{-j-1}$, $j=0, 1, \ldots, t$. We shall show that the probability that $|U_j| \geqq \frac{3}{2} p_j n$, conditional on $|U_i| \leqq \frac{3}{2} p_i n$, $i=0, 1, \ldots, j-1$, is rather small. To be precise,

(2) $$P\left(|U_j| \geqq \frac{3}{2} p_j n \,\middle|\, |U_i| \leqq \frac{3}{2} p_i n, \quad i=0, 1, \ldots, j-1\right) \leqq e^{-p_j n/12} = P_j,$$

$$j = 1, 2, \ldots, t.$$

In order to prove (2), consider disjoint sets $V_0, V_1, \ldots, V_{j-1}$ of V satisfying

$$v_i = |V_i| \leqq \frac{3}{2} p_i n, \quad i = 0, 1, \ldots, j-1.$$

Let $x_0 \in V - \bigcup_{i=0}^{j-1} V_i$. We claim that

(3) $$P\left(\left|\Gamma_+(x_0) \cap \bigcup_{i=0}^{j-1} V_i\right| \geqq 2 \,\middle|\, U_i = V_i, \quad i = 0, 1, \ldots, j-1\right) \leqq p_j,$$

$j=1, 2, \ldots, t$. Note that, by (1), inequality (3) implies (2), so it suffices to prove (3).

First consider $j=1$. Since $P(x \in U_0) < 1/2$,

$$P(|\Gamma_+(x_0) \cap V_0| \geqq 2 \,|\, U_0 = V_0) \leqq 2 \sum_{i=25}^{[6d]} \sum_{k=2}^{i} \binom{v_0}{k} \binom{n-1-v_0}{i-k} r^i (1-r)^{n-1-i}$$

$$\leqq 3 \sum_{i=25}^{[6d]} \sum_{k=2}^{i} \left(\frac{3}{2} p_0\right)^k \frac{d^i}{k!(i-k)!} e^{-d} \leqq 3e^{-d} \sum_{k=2}^{[6d]} \frac{(3p_0 d/2)^k}{k!} \sum_{i=k}^{\infty} d^{i-k}/(i-k)! \leqq 5d^2 p_0^2 < p_1.$$

The last but one inequality above holds because $p_0 d < 1/5$.

Now assume that $2 \leqq j \leqq t$ and we have proved (3) and (2) for smaller values of j. If a vertex x belongs to U_j then x dominates at least one vertex of U_{j-1} and at least two vertices of $\bigcup_{i=0}^{j-1} U_i$. Taking into account that $\frac{3}{2} \sum_{i=0}^{j-1} p_i < 2p_0 < 1/2$, we find that for $x_0 \notin \bigcup_{i=0}^{j-1} V_i$

$$P\left(\left|\Gamma_+(x_0) \cap \bigcup_{i=0}^{j-1} U_i\right| \geqq 2 \,\middle|\, U_i = V_i, \quad i = 0, 1, \ldots, j-1\right) \leqq$$

$$\leqq 2 \sum_{k=25}^{[6d]} v_{j-1} \sum_{i=0}^{j-1} v_i \binom{n-2}{k-2} r^k (1-r)^{n-1-k} \leqq$$

$$\leqq 3 \sum_{k=25}^{[6d]} \left(\frac{3}{2} p_{j-1}\right) (2p_0) \frac{d^k}{(k-2)!} e^{-d} \leqq 9d^2 p_0 p_{j-1} = p_j.$$

This proves (3) and so (2).

The set $\bigcup_{i=0}^{t} U_i$ will not always do for the set $U=V-W$ of exceptional vertices. Set

$$\tilde{U} = \left\{x\in V - \bigcup_{j=1}^{t} U_j \colon\ |\Gamma_+(x)\cap U_t| \geqq 2\right\}.$$

Let us estimate the probability of $\tilde{U}$ being large, conditional on $U_0, U_1, \ldots, U_t$ being not too large. As before, consider disjoint sets $V_0, V_1, \ldots, V_t$ of V satisfying

$$v_j = |V_j| \leqq \frac{3}{2} p_j n, \quad j = 0, 1, \ldots, t.$$

Then for $x_0\in V - \bigcup_{j=0}^{t} V_j$

$$P(|\Gamma_+(x_0)\cap V_t| \geqq 2 \,|\, U_i = V_i, \quad i = 0, 1, \ldots, t) \leqq$$

$$\leqq 2 \sum_{i=25}^{\lfloor 6d \rfloor} \binom{v_t}{2}\binom{n-2}{i-2} r^i (1-r)^{n-1-i} \leqq 2d^2 p_t^2 \leqq (Dp_0)^{2(t+1)} \leqq n^{-4/3}.$$

Hence

$$P\left(\tilde{U} \neq \emptyset \,\middle|\, |U_i| \leqq \frac{3}{2} p_i n, \quad i = 0, 1, \ldots, t\right) \leqq n^{-1/3}. \tag{4}$$

Putting together (2) and (4) we find that

$$P\left(|U_j| \leqq \frac{3}{2} p_j n, \quad j = 0, 1, \ldots t, \quad \text{and} \quad \tilde{U} = \emptyset\right) \geqq 1 - \sum_{i=0}^{t} P_j - n^{-1/3} = 1 - o(1).$$

Now let W be a set of maximum cardinality for which

$$|\Gamma_+(x)\cap W| \geqq 23$$

for every $x\in V-W$. If $\tilde{U}=\emptyset$ then the set $V-\bigcup_{j=0}^{t} U_j$ will do for W since a vertex not in $\bigcup_{j=0}^{t} U_j$ dominates at most one vertex in U_t and at most one vertex in $\bigcup_{j=0}^{t-1} U_j$, so it dominates at least $25-2=23$ vertices in $V-\bigcup_{j=0}^{t} U_j$.

Consequently

$$P(|W| \geqq (1-2p_0)n) \geqq P\left(|U_j| \leqq \frac{3}{2} p_j n, \quad j = 0, \ldots, t, \quad \text{and} \quad \tilde{U} = \emptyset\right) = 1 - o(1).$$

Having found this large set W the proof is almost complete. Indeed, if W_0 is a fixed set of vertices with $|W_0| \geqq (1-2p_0)n$ then

$$P(\theta(\vec{G})[W] \quad \text{is Hamiltonian} \,|\, W = W_0) = 1 - o(1).$$

Consequently a.e. $\vec{G}\in\vec{\mathcal{G}}(n, r)$ is such that $\theta(\vec{G})$ contains a cycle of length at least $(1-2p_0)n$.

Note that in the proof we were very generous in our estimate of p_0. This is because the correct upper bound for $\beta(c)$ will be c_1ce^{-c} where c_1 is an absolute constant. We shall return to this in a later note.

References

[1] M. Ajtai, J. Komlós and E. Szemerédi, The longest path in a random graph, *Combinatorica* **1** (1981) 1—12.
[2] B. Bollobás, *Graph Theory — An Introductory Course,* Graduate Texts in Mathematics, Springer-Verlag, New York, Heidelberg and Berlin, 1979.
[3] P. Erdős and A. Rényi, On the evolution of random graphs, *Publ. Math. Inst. Hungar. Acad. Sci.* **5** (1960) 17—61.
[4] T. I. Fenner and A. M. Frieze, On the existence of hamiltonian cycles in a class of random graphs, *to appear.*
[5] W. F. de la Véga, Long paths in random graphs, *Combinatorica* **3** (1983)

Béla Bollobás

Department of Pure Mathematics and Mathematical Statistics
University of Cambridge
Cambridge CB2 1SB, U.K.

Department of Mathematics
Louisiana State University
Baton Rouge, LA 70803, U.S.A.