

ON THE STRENGTH OF CONNECTEDNESS OF A RANDOM GRAPH

By

P. ERDŐS (Budapest), corresponding member of the Academy,
and A. RÉNYI (Budapest), member of the Academy

Let G be a non-oriented graph without parallel edges and without slings, with vertices V_1, V_2, \dots, V_n . Let us denote by $d(V_k)$ the *valency* (or degree) of a point V_k in G , i. e. the number of edges starting from V_k . Let us put

$$(1) \quad c(G) = \min_{1 \leq k \leq n} d(V_k).$$

If G is an arbitrary non-complete graph, let $c_p(G)$ denote the least number k such that by deleting k appropriately chosen vertices from G (i. e. deleting the k points in question and all edges starting from these points) the resulting graph is not connected. If G is a complete graph of order n , we put $c_p(G) = n - 1$. Let $c_e(G)$ denote the least number l such that by deleting l appropriately chosen edges from G the resulting graph is not connected. We may measure the strength of connectedness of G by any of the numbers $c_p(G)$, $c_e(G)$ and in a certain sense (if G is known to be connected) also by $c(G)$. Evidently one has

$$(2) \quad c(G) \geq c_e(G) \geq c_p(G).$$

It is known further that any two points of G are connected by at least $c_p(G)$ paths having no point in common, except the two endpoints (theorem of MENGER—WHITNEY, see [1] and [2]) and by at least $c_e(G)$ paths having no edge in common (theorem of FORD and FULKERSON, see [3]).

We shall denote by $\nu_r(G)$ the number of vertices of G which have the valency r ($r = 0, 1, 2, \dots$).

As in two previous papers ([4], [5]) we consider the random graph $\Gamma_{n, N}$ defined as follows: Let there be given n labelled points V_1, V_2, \dots, V_n . Let us choose at random N edges among the $\binom{n}{2}$ possible edges connecting these n points, so that each of the $\binom{\binom{n}{2}}{N}$ possible choices of these edges should be equiprobable. We denote by $\Gamma_{n, N}$ the random graph thus obtained. We shall denote by $\mathbf{P}(\cdot)$ the probability of the event in the brackets.

The aim of this note is to investigate the strength of connectedness of the random graph $\Gamma_{n,N}$ when n and N both tend to $+\infty$, $N=N(n)$ being a function of n . As it has been shown in [4], the following theorem holds:

THEOREM 1. *If we have $N(n) = \frac{1}{2}n \log n + \alpha n + o(n)$ where α is a real constant, then the probability of $\Gamma_{n,N(n)}$ being connected tends to $\exp(-e^{-2\alpha})$ for $n \rightarrow +\infty$.*

In this paper we shall prove the following theorem:

THEOREM 2. *If we have $N(n) = \frac{1}{2}n \log n + \frac{r}{2}n \log \log n + \alpha n + o(n)$ where α is a real constant and r a non-negative integer, then*

$$(3) \quad \lim_{n \rightarrow +\infty} \mathbf{P}(c_p(\Gamma_{n,N(n)}) = r) = 1 - \exp\left(-\frac{e^{-2\alpha}}{r!}\right),$$

further

$$(4) \quad \lim_{n \rightarrow +\infty} \mathbf{P}(c_e(\Gamma_{n,N(n)}) = r) = 1 - \exp\left(-\frac{e^{-2\alpha}}{r!}\right)$$

and

$$(5) \quad \lim_{n \rightarrow +\infty} \mathbf{P}(c(\Gamma_{n,N(n)}) = r) = 1 - \exp\left(-\frac{e^{-2\alpha}}{r!}\right).$$

REMARK. Clearly Theorem 2 can be considered as a generalization of Theorem 1. As a matter of fact, any of the statements $c_p(G) = 0$ or $c_e(G) = 0$ is equivalent to G not being connected and thus for $r = 0$ (3) and (4) reduce to the statement of Theorem 1. It has been shown further in [4] that if $N(n) = \frac{n}{2} \log n + \alpha n + o(n)$ and $\Gamma_{n,N(n)}$ is not connected, then it consists almost surely of a connected component and of a few isolated points. Therefore (5) is for $r = 0$ also equivalent to the statement of Theorem 1. Thus in proving Theorem 2 we may restrict ourselves to the case $r \geq 1$.

The statement (5) of Theorem 2 gives information about the *minimal* valency of points of $\Gamma_{n,N}$. In a forthcoming note we shall deal with the same question for larger ranges of N (when $c(\Gamma_{n,N})$ tends to infinity with n), further with the related question about the *maximal* valency of points of $\Gamma_{n,N}$.

We shall prove further the following

THEOREM 3. *If we have $N(n) = \frac{1}{2}n \log n + \frac{r}{2}n \log \log n + \alpha n + o(n)$ where α is a real constant and r a non-negative integer, then we have*

$$(6) \quad \lim_{n \rightarrow +\infty} \mathbf{P}(\nu_r(\Gamma_{n,N(n)}) = k) = \frac{\lambda^k e^{-\lambda}}{k!} \quad \text{for } k = 0, 1, \dots$$

where $\lambda = \frac{e^{-2\alpha}}{r!}$; in other words, the distribution of $\nu_r(\Gamma_{n, N(n)})$ tends to a Poisson distribution.

PROOF OF THEOREMS 2 AND 3. Let $r \geq 1$ be an integer and $-\infty < \alpha < +\infty$. Let us suppose that

$$(7) \quad N(n) = \frac{1}{2}n \log n + \frac{r}{2}n \log \log n + \alpha n + o(n).$$

Let $\Gamma_{n, N}$ be a random graph with the n vertices V_1, V_2, \dots, V_n and having N edges. Let $P_k(n, N, r)$ denote the probability that by removing r suitably chosen points from $\Gamma_{n, N}$ there remain two disjoint graphs, consisting of k and $n - k - r$ points, respectively. We may suppose $k < \left\lfloor \frac{n-r}{2} \right\rfloor$. First we have clearly

$$P_k(n, N, r) \leq \binom{n}{r} \binom{n-r}{k} \frac{\binom{n}{2} - k(n-k-r)}{\binom{n}{N}}.$$

It follows by some obvious estimations that

$$(8) \quad \sum_{\substack{(r+3) \frac{\log n}{\log \log n} < k \leq \left\lfloor \frac{n-r}{2} \right\rfloor}} P_k(n, N(n), r) = O\left(\frac{1}{n}\right).$$

Now we consider the case $k \leq (r+3) \frac{\log n}{\log \log n}$. Let $P_k^*(n, N, r)$ denote the probability that by removing r suitably chosen points (the set of which will be denoted by \mathcal{A}) $\Gamma_{n, N}$ can be split into two disjoint subgraphs Γ' and Γ'' consisting of k and $n - k - r$ points, respectively, but that $\Gamma_{n, N}$ can not be made disconnected by removing only $r - 1$ points. If $\Gamma_{n, N}$ has these properties and if s denotes the number of edges of $\Gamma_{n, N}$ connecting a point of \mathcal{A} with a point of Γ' , then we have clearly $s \geq r$. Otherwise, by definition, $s \leq rk$. Thus we have

$$(9) \quad P_k^*(n, N, r) \leq \sum_{s=r}^{rk} \binom{n}{r} \binom{n-r}{k} \binom{rk}{s} \frac{\binom{n}{2} - k(n-k)}{\binom{n}{N-s}}.$$

It follows that

$$(10) \quad P_k^*(n, N(n), r) = O\left(\frac{1}{\log n}\right).$$

From (8) and (10) it follows that for $n \rightarrow +\infty$

$$(11) \quad \mathbf{P}(c_p(I_{n, N(n)}) = r) \sim \mathbf{P}(c(I_{n, N(n)}) = r).$$

As a matter of fact, (8) and (10) imply that if by removing r suitably chosen points (but not by removing less than r points) $I_{n, N(n)}$ can be split into two disjoint subgraphs I' and I'' consisting of k and $n-k-r$ points, respectively, where $k \leq \left\lfloor \frac{n-r}{2} \right\rfloor$, then only the case $k=1$ has to be considered, the probability of $k > 1$ being negligibly small. It remains to prove (5). This can be done as follows. First we prove that

$$(12) \quad \lim_{n \rightarrow +\infty} \mathbf{P}(c(I_{n, N(n)}) \leq r-1) = 0.$$

For $r=1$ this follows already from Theorem 1. Thus we may suppose here $r \geq 2$. We have

$$\mathbf{P}(c(I_{n, N}) \leq r-1) \leq \sum_{h=1}^{r-1} n \binom{n-1}{h} \frac{\binom{n}{2} - (n-1)}{\binom{n}{N}},$$

and thus

$$(13) \quad \mathbf{P}(c(I_{n, N(n)}) \leq r-1) = O\left(\frac{1}{\log n}\right)$$

which proves (12).

Now let $\nu_r(I_{n, N})$ denote the number of vertices of $I_{n, N}$ which have the valency r . Then we have clearly by (12)

$$(14) \quad \mathbf{P}(c(I_{n, N(n)}) = r) \sim \mathbf{P}(\nu_r(I_{n, N(n)}) \neq 0).$$

Now evidently

$$(15) \quad \mathbf{P}(\nu_r(I_{n, N(n)}) \neq 0) = \sum_{j=1}^n (-1)^{j-1} S_j$$

where

$$(16) \quad S_j = \sum_{1 \leq k_1 < k_2 < \dots < k_j \leq n} \mathbf{P}(d(V_{k_1}) = r, d(V_{k_2}) = r, \dots, d(V_{k_j}) = r).$$

Evidently, if we stop after taking an even or odd number of terms of the

sum on the right-hand side of (15), we obtain a quantity which is greater or smaller, respectively, than the left-hand side of (15). Now clearly

$$\mathbf{P}(d(V_k) = r) = \binom{n-1}{r} \frac{\binom{\binom{n}{2} - (n-1)}{N(n) - r}}{\binom{\binom{n}{2}}{N(n)}} \sim \frac{e^{-2\alpha}}{nr!},$$

and thus

$$(17) \quad \lim_{n \rightarrow +\infty} S_1 = \frac{e^{-2\alpha}}{r!}.$$

Now let us consider $\mathbf{P}(d(V_{k_1}) = r, d(V_{k_2}) = r)$ where $k_1 \neq k_2$. If both V_{k_1} and V_{k_2} have valency r , three cases have to be considered: a) either V_{k_1} and V_{k_2} are not connected, and there is no point which is connected with both V_{k_1} and V_{k_2} ; b) or V_{k_1} and V_{k_2} are not connected, but there is a point connected with both; c) V_{k_1} and V_{k_2} are connected. We denote the probabilities of the corresponding subcases by $\mathbf{P}_a(d(V_{k_1}) = r, d(V_{k_2}) = r)$, $\mathbf{P}_b(d(V_{k_1}) = r, d(V_{k_2}) = r)$ and $\mathbf{P}_c(d(V_{k_1}) = r, d(V_{k_2}) = r)$, respectively. We evidently have

$$\mathbf{P}_a(d(V_{k_1}) = r, d(V_{k_2}) = r) = \frac{(n-2)!}{r!^2(n-2r-2)!} \frac{\binom{\binom{n}{2} - (2n-3)}{N(n) - 2r}}{\binom{\binom{n}{2}}{N(n)}} \sim \left(\frac{e^{-2\alpha}}{n \cdot r!}\right)^2,$$

and thus

$$(18) \quad \sum_{1 \leq k_1 < k_2 \leq n} \mathbf{P}_a(d(V_{k_1}) = r, d(V_{k_2}) = r) \sim \frac{1}{2} \left(\frac{e^{-2\alpha}}{r!}\right)^2.$$

On the other hand (denoting by l the number of points which are connected with both V_{k_1} and V_{k_2}), we have

$$(19) \quad \begin{aligned} &\mathbf{P}_b(d(V_{k_1}) = r, d(V_{k_2}) = r) = \\ &= \sum_{l=1}^r \frac{(n-2)!}{l!(r-l)!(n-2r+l-2)!} \frac{\binom{\binom{n}{2} - (2n-3)}{N(n) - 2r}}{\binom{\binom{n}{2}}{N(n)}} = O\left(\frac{1}{n^3}\right). \end{aligned}$$

Similarly one has

$$(20) \quad \mathbf{P}_c(d(V_{k_1})=r, d(V_{k_2})=r) = \sum_{l=0}^{r-1} \frac{(n-2)!}{l!(r-l-1)!(n-2r+l)!} \frac{\binom{n}{2} - (2n-3)}{\binom{n}{2}} \frac{N(n)-2r}{N(n)} = O\left(\frac{1}{n^4}\right).$$

Thus we obtain

$$\lim_{n \rightarrow +\infty} S_2 = \frac{1}{2} \left(\frac{e^{-2\alpha}}{r!} \right)^2.$$

The cases $j > 2$ can be dealt with similarly. Thus we obtain

$$(21) \quad \lim_{n \rightarrow +\infty} S_j = \frac{1}{j!} \left(\frac{e^{-2\alpha}}{r!} \right)^j \quad (j = 1, 2, 3, 4, \dots).$$

It follows from (16) and (21) that

$$(22) \quad \lim_{n \rightarrow +\infty} \mathbf{P}(\nu_r(\Gamma_{n, N(n)}) \neq 0) = 1 - \exp\left(-\frac{e^{-2\alpha}}{r!}\right).$$

In view of (2), (11) and (14) Theorem 2 follows.

To prove Theorem 3 it is sufficient to remark that by the well-known formula of CH. JORDAN

$$(23) \quad \mathbf{P}(\nu_r(\Gamma_{n, N(n)}) = k) = \sum_{j=0}^{n-k} (-1)^j \binom{j+k}{j} S_{j+k},$$

and thus by (21), putting $\lambda = \frac{e^{-2\alpha}}{r!}$, we obtain for $k = 0, 1, \dots$

$$(24) \quad \lim_{n \rightarrow +\infty} \mathbf{P}(\nu_r(\Gamma_{n, N(n)}) = k) = \frac{\lambda^k}{k!} \sum_{j=0}^{\infty} \frac{(-1)^j \lambda^j}{j!} = \frac{\lambda^k e^{-\lambda}}{k!}.$$

Thus Theorem 3 is proved.

Our thanks are due to T. GALLAI for his valuable remarks.

(Received 12 October 1960)

References

- [1] D. KÖNIG, *Theorie der endlichen und unendlichen Graphen* (Leipzig, 1936).
- [2] C. BERGE, *Théorie des graphes et ses applications* (Paris, 1958).
- [3] L. R. FORD and D. R. FULKERSON, Maximal flow through a network, *Canadian Journal of Math.*, **8** (1956), p. 399.
- [4] P. ERDŐS and A. RÉNYI, On random graphs. I, *Publ. Math. Debrecen*, **6** (1959), pp. 290—297.
- [5] P. ERDŐS and A. RÉNYI, On the evolution of random graphs, *Publ. Math. Inst. Hung. Acad. Sci.*, **5** (1960), pp. 17—61.