The Annals of Probability
1983, Vol. 11, No. 1, 217–221

A METHOD OF INVESTIGATING THE LONGEST PATHS IN CERTAIN RANDOM GRAPHS

BY T. NEMETZ AND N. KUSOLITSCH

Mathematical Institute of the Hungarian Academy and Technische Universität Wien

In this paper we consider specific directed graphs called "ladders". The vertices of the graph are randomly colored by green or red. Deleting the edges with at least one red endpoint one gets a random graph. We give a method for finding the exact asymptotics of the longest path of this random graph if the "height" of the ladder goes to infinity. The result is a generalization of a celebrated theorem of Erdös-Rényi. An example is given, illustrating the method.

1. Introduction. The real line can be viewed as a labelled directed graph, where the vertices are the integers and the edges connect consecutive integers, always directed to the larger ones. Suppose we are given a sequence $\eta_1, \eta_2, \cdots$ of i.i.d. random variables with probability $P(\eta_1 = 0) = P(\eta_1 = 1) = ½$. Then we can color the integer point i of the "real-line graph" by green or red if $\eta_i = 0$ or $\eta_i = 1$, respectively. Any edge will be deleted if at least one of its endpoints is red. Then the paths in this random graph correspond to the 0-runs in $\eta_1, \eta_2, \cdots$. Therefore any problem concerning paths in this directed random graph has an equivalent one concerning 0-runs. This equivalence leads us to utilize run-theoretical methods in investigating longest-path problems in a class of directed graphs, called "ladders", which can be viewed as a generalization of the real-line graph.

DEFINITION. A *ladder* is a labelled directed graph $G = G(V, E)$ with the following properties:

The set V of vertices is partitioned into subsets $V^i := \{A_0^i, \cdots, A_{r-1}^i\}$, $i = 0, 1, \cdots$.

The subgraph with vertex-set $V^0 \cup V^1$ is a bipartite graph, all edges being directed from vertices in V^0 towards vertices in V^1. (A graph is called bipartite if the set of vertices can be divided into two subsets with no edge joining two vertices in the same subset.)

There is no edge between A_k^i, A_h^j if $|i - j| > 1$.

There is an edge going out of A_i^0 and also there is one ending in A_i^1 for all indices $i = 0, \cdots, r - 1$.

There is a directed edge (A_k^i, A_h^{i+1}) iff A_k^0, A_h^1 are connected by an edge.

The set V^i is called the ith rung of the ladder.

Obviously a ladder is completely characterized by the subgraph restricted to $V^0 \cup V^1$.

The subgraph restricted to the first $n + 1$ rungs is denoted by G_n.

EXAMPLE 1. Let $r = 3$, and let the bipartite graph $V^0 \cup V^1$ contain the six edges (A_j^0, A_j^1) $j = 0, 1, 2$, (A_0^0, A_1^1), (A_1^0, A_2^1), (A_2^0, A_0^1). The example is depicted in Figure 1.

Now suppose that the vertices A_k^i of G_n are independently randomly colored by green with probability $p_k > 0$ and by red with probability $1 - p_k$, where p_k does not depend on i. Then we get a random graph by deleting the edges having at least one red endpoint. The

Received December 1981; revised April 1982.
AMS 1980 *subject classifications.* Primary 60F15; secondary 60F10, 05C20.
Key words and phrases. Erdös-Rényi laws, random graph, longest runs, large deviation.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Probability.*
www.jstor.org

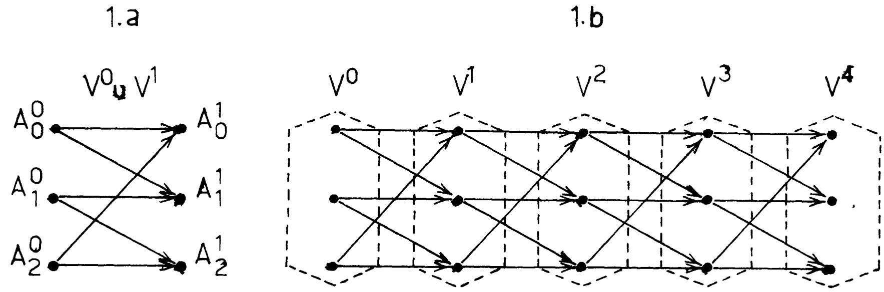

FIG. 1

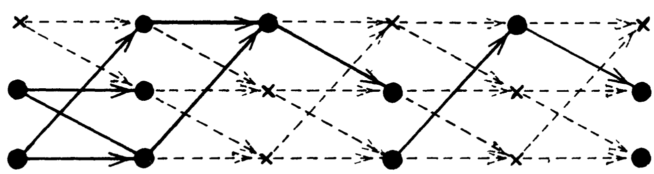

FIG. 2

resulting random graph will be denoted by $R_n = R_n(G; p_0, \cdots, p_{r-1})$. This is illustrated in Figure 2, where green vertices of the graph of Example 1 are indicated by circles and the red ones by crosses. Deleted edges are indicated by dashed lines.

We are interested in the asymptotic behaviour of the length $\nu_G(n)$ of the longest path in the random subgraph R_n of G_n. The length of a path is the number of rungs connected by it minus one. In the case of the real-line graph $\nu_G(n)$ is well known to have the order of magnitude of log n. See Erdös-Rényi (1970). In this paper we show that this statement is generally true, namely we will prove the following theorem.

THEOREM. *Let $G(V, E)$ be an arbitrary ladder with r vertices $A_0^i, \cdots A_{r-1}^i$ in the ith rung. The vertices A_k^i of G_n are randomly independently colored by green with probability $0 < p_k < 1$ and by red with probability $1 - p_k$, $k = 0, \cdots, r-1$. The edges of G_n having at least one red endpoint are deleted. Then the length $\nu_G(n)$ of the longest path in the resulting random graph R_n satisfies*

$$P\left(\lim_{n\to\infty} \frac{\nu_G(n)}{\log n} = -\frac{1}{\log \lambda}\right) = 1$$

with some positive constant $\lambda = \lambda(G; p_0, \cdots, p_{r-1}) < 1$.

A method is given for determining this constant as the largest eigenvalue of a quadratic matrix of size $2^r - 1$.

Preliminary results. In order to prove our theorem we have to establish some results concerning the calculation of the constant λ. For this reason we will first prove two lemmas in terms of the following cumbersome notations:

$B_i := \{j : 0 \le j \le r-1 \text{ and } (A_i^0, A_j^1) \text{ is an edge in } G(V, E)\}$ $i = 0, \cdots, r-1$.

We use the letter I (or $\tilde{I}$) for denoting a general set of indices, that is $I = \{i_1, \cdots, i_k\}$ $1 \le k \le r$, $i_t \ne i_s$ if $t \ne s$.

Let

$$B(I) := \cup_{h=1}^{k} B_{i_h} \quad \text{if} \quad I = \{i_1, \cdots, i_k\}$$

and

$$P(I) := \prod_{h=1}^{k} p_{i_h}.$$

Let $C^n(I)$ be the event that there are paths of length n in R_n and the set of their vertices in V^n is $\{A^n_{i_1} \cdots, A^n_{i_k}\}$, $I = \{i_1, \cdots, i_k\}$.

The independence and "stationarity" of the coloring allows us to determine the conditional probabilities $P(\tilde{I} \mid I) := P(C^{n+1}(\tilde{I}) \mid C^n(I))$ for all pairs $(I, \tilde{I})$.

LEMMA 1.

$$P(\tilde{I} \mid I) = \begin{cases} 0 & \text{if } \tilde{I} \not\subset B(I) \\ P(\tilde{I}) \cdot \prod_{j \in B(I) \setminus \tilde{I}} (1 - p_j) & \text{otherwise.} \end{cases}$$

PROOF. If $\tilde{I} \not\subset B(I)$, then there is at least one index $\tilde{\imath}$, such that $A^{n+1}_{\tilde{\imath}}$ cannot be reached from any of the vertices $A^n_{i_1}, \cdots, A^n_{i_k}$. In this case the conditional probability must be zero. In the remaining case all vertices A^{n+1}_i, $i \in \tilde{I}$ must be green and all vertices A^{n+1}_j, $j \in B(I) \setminus \tilde{I}$ must be red. For the remaining vertices there is no restriction at all. Since all the vertices are colored independently this settles the second case $\tilde{I} \subset B(I)$. □

Let $I_1, \cdots, I_{2^r-1}$ be the non-empty index-sets in a given order. We introduce the "truncated" transition probability matrix

$$\Pi_G = (p_{t,s}), \text{ with } p_{t,s} := P(I_s \mid I_t) \quad s, t = 1, \cdots, 2^r - 1.$$

The row sums of Π_G are strictly less than one, i.e. $\sum_s p_{t,s} \leq 1 - \prod_{j=0}^{r-1} (1 - p_j)$, because all vertices of the subsequent rung may be colored red. The Frobenius theorem tells us that the matrix Π_G has a positive eigenvalue $\lambda = \lambda(G; p_0, \cdots, p_{r-1})$ majorizing all other eigenvalues in moduli and

(1) $$\lambda < 1.$$

LEMMA G. $\Pi^n_G = \sum_{i=1}^q \sum_{j=0}^{v_i-1} \frac{n!}{(n-j)!} \lambda_i^{n-j} D_{i,j}$, *where* $\lambda_1, \cdots, \lambda_q$ *are the eigenvalues of* Π_G *with multiplicities* $v_1, \cdots, v_q$ *and* $D_{i,j}$ *are matrices not depending on* n. (*cf. Gantmacher*, 1966).

This lemma allows us to compute the probability $P(\nu_G(n) = n)$, i.e. the probability that there is a path of length n in R_n.

LEMMA 2. *There are constants* $d_{i,j}$, $i = 1, \cdots, q, j = 1, \cdots, v_i$ *such that for every integer*

$$P(\nu_G(n) = n) = \sum_{i=1}^q \sum_{j=1}^{v_i-1} \frac{n!}{(n-j)!} \lambda_i^{n-j} d_{i,j}.$$

PROOF. The idea is to find a recursion formula for the $2^r - 1$ probabilities $P(C^n(I))$, instead of giving an explicit form or a recursion for $P(\nu_G(n) = n)$. Obviously $P(\nu_G(n) = n)$ is the sum of these probabilities.

Recalling the notations preceding Lemma 1 we get

(2) $$P(C^n(I_s)) = \sum_{t=1}^{2^r-1} p_{t,s} P(C^{n-1}(I_t)), \quad s = 1, \cdots, 2^r - 1; r \geq 2.$$

(Here the nonempty index sets I_s are ordered in the same way as before.)

Defining $C^o(I)$ as the event that in the 0-th -rung exactly the vertices $A^o_{i_1}, \cdots A^o_{i_k}$, $I = \{i_1, \cdots, i_k\}$ are colored by green, the recursion extends to $n = 1$. On the other hand obviously,

$$P^1(I) := P(C^0(I)) = P(I) \prod_{j \notin I} (1 - p_j).$$

With the notation $P^1 := (P^1(I_1), \cdots, P^1(I_{2^r-1}))$, (2) implies an explicit form for the vector

$P^n := (P(C^n(I_1)), \cdots, P(C^n(I_{2^r-1})))$, namely

$$P^n = P^1 \Pi_G^n \qquad n = 1, 2, \cdots.$$

Let $\bar{1}^T$ denote the $2^r - 1$ dimensional vector with unit components, i.e. $\bar{1}^T = (1, \cdots, 1)$. Since

$$P(\nu_G(n) = n) = \sum_{t=1}^{2^r-1} P(C^n(I_t)) = P^n \bar{1}$$

we get by applying Lemma G:

$$P(\nu_G(n) = n) = P^1 \Pi_G^n \bar{1} = \sum_{i=1}^{q} \sum_{j=0}^{v_i-1} \frac{n!}{(n-j)!} \lambda_i^{n-j} d_{i,j} \text{ with } d_{i,j} = P^1 D_{i,j} \bar{1}. \qquad \square$$

PROOF OF THE THEOREM. Let us denote by $E_{i,N}$ the event that there is a path from the ith rung to the $i + N$th rung in R_n. Obviously $P(E_{i,N}) = P(\nu_G(N) = N)$ for all integers i and N. Therefore we get from Lemma 2

(3) $$P(\nu_G(n) \geq N) \leq \sum_{i=0}^{n-N} P(E_{i,N}) \leq cn\lambda_G^N$$

for some positive constant c.

On the other hand

$$P(\nu_G(n) < N) \leq P(\cap_{i=0}^{[n/(N+1)]-1} E_{i(N+1),N}^c.$$

But the events $E_{i,N}$, $E_{j,N}$ are independent for $|i - j| > N$, so

$$P(\nu_G(n) < N) \leq (1 - P(\nu_G(N) = N))^{n/(N+1)-2}.$$

Taking into account, that

$$P(\nu_G(N) = N) = P^1 \Pi_G{}^N \bar{1} = \sum_{s,t} P^1(I_s) p_{s,t}^{(N)} \geq \min_{1 \leq s \leq 2^r-1} P^1(I_s) \max_s \sum_t p_{s,t}^{(N)} \geq \tilde{c}\lambda_G^N,$$

where $\tilde{c} = \min_{1 \leq s \leq 2^r-1} P^1(I_s) > 0$, one gets finally

(4) $$P(\nu_G(n) < N) \leq (1 - \tilde{c}\lambda_G^N)^{n/(N+1)-2}.$$

Inequality (4) implies

$$P\left(\nu_G(n) < \left[\frac{(1-\varepsilon)\log n}{-\log \lambda_G}\right]\right) \leq \exp(-cn^{\varepsilon/2}),$$

for every $\varepsilon > 0$. (Here and in what follows $[x]$ always denotes the largest integer not greater than x.) Since $\sum_{n \in N} \exp(-cn^{\varepsilon/2}) < \infty$ an application of the Borel-Cantelli-Lemma yields

$$P\left(\lim_n \inf \frac{\nu_{G(n)}}{\log n} \geq (1-\varepsilon) \frac{1}{-\log \lambda_G}\right) = 1.$$

But inequality (3) implies

$$P\left(\nu(n) \geq \left[\frac{(1+\varepsilon)\log n}{-\log \lambda_G}\right]\right) \leq \frac{cn^{-\varepsilon}}{\lambda_G}.$$

This, together with the Borel-Cantelli-Lemma, implies

$$P\left(\lim_n \sup \frac{\nu_G(2^n)}{\log 2^n} \leq \frac{1+\varepsilon}{-\log \lambda_G}\right) = 1.$$

Since for an integer m with $2^{n-1} < m \leq 2^n$,

$$\frac{\nu_G(m)}{\log m} \leq \frac{\nu_G(2^n)}{\log 2^n - \log 2}$$

holds true, we get

$$P\left(\lim_n \sup \frac{\nu_G(n)}{\log n} \leq \frac{1+\varepsilon}{-\log \lambda_G}\right) = 1.$$

TABLE 1

s \ t	1	2	3	4	5	6	7
1	2	2	0	2	0	0	0
2	0	2	2	0	0	2	0
3	2	0	2	0	2	0	0
4	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1

Since $\varepsilon > 0$ is chosen arbitrarily, one gets

$$P\left(\lim_n \frac{\nu_G(n)}{\log n} = \frac{1}{-\log \lambda_G}\right) = 1. \qquad \square$$

EXAMPLE 1 (*continued*). We illustrate the theorem by carrying out the computations in the example above. Let $p_0 = p_1 = p_2 = 1/2$. First we must compute the "truncated" transition probability matrix Π_G: we have 7 non-empty index-sets. Let $I_1 = \{0\}$, $I_2 = \{1\}$, $I_3 = \{2\}$, $I_4 = \{0, 1\}$, $I_5 = \{0, 2\}$, $I_6 = \{1, 2\}$, $I_7 = \{0, 1, 2\}$ be their ordering. The sets B_i are then $B_i = \{i, i \oplus 1 \bmod 3\}$ $i = 0, 1, 2$, so that $B(I_t) = B_t$ for $t = 1, 2, 3$ and $B(I_t) = I_7$ for $t \geq 4$. The "truncated" transition probability matrix is given by Table 1, where $p_{s,t}$ is the s, t-entry of the table divided by 8.

The characteristic polynomial of this matrix is given by $\lambda^3(6 - (4 - 8\lambda)^2)((8\lambda - 1)^2 + 3)$, so its eigenvalues are $\lambda_{1,2} = (4 \pm \sqrt{6})/8$, $\lambda_{3,4} = (1 \pm i\sqrt{3})/8$, $\lambda_5 = \lambda_6 = \lambda_7 = 0$, and $\lambda_G = (4 + \sqrt{6})/8$. Therefore the "exact" asymptotic is given by

$$P\left(\lim_n \frac{\lambda_G(n)}{\log n} = \frac{1}{\log 8 - \log(4 + \sqrt{6})}\right) = 1. \qquad \square$$

Concluding remark. As was mentioned, our result is a generalization of the Erdös-Rényi law of large numbers. J. Komlós pointed out to us that the very method could be applied to extend the validity of the more general results of Komlós and Tusnády (1975), Erdös and Révész (1975), Révész (1982).

REFERENCES

[1] ERDÖS, P. and RÉNYI, A. (1970). On a new law of large numbers. *J. Análys e Math.* **23** 103–111.
[2] ERDÖS, P. and RÉVÉSZ, P. (1975). On the length of the longest head-run. Topics in Information Theory, Keszthely (Hungary), *Coll. Math. Soc. János Bolyai* **16** 219–228.
[3] GANTMACHER, F. R. (1966). *Matrizenrechnung.* DVW, Berlin Bd II.
[4] KOMLÓS, J. and TUSNÁDY, G. (1975). On sequences of "pure heads". *Ann. Probability* **3** 608–617.
[5] RÉVÉSZ, P. (1982). A characterization of the asymptotic properties of a stochastic process by four classes of deterministic curves. Unpublished manuscript.

MATHEMATICAL INSTITUTE
OF THE HUNGARIAN ACADEMY
OF SCIENCES
REALTANODA UT 13-15
H-1053 BUDAPEST V
HUNGARY

INSTITUT FÜR STATISTIK UND
WAHRSCHEINLICHKEITSTHEORIE
TECHNISCHE UNIVERSITÄT WIEN
KARLSPLATZ 13
A-1040 WIEN
AUSTRIA