Discrete Mathematics 14 (1976) 359–364.

HAMILTONIAN CIRCUITS IN RANDOM GRAPHS

L. PÓSA
Eötvös Lorand University, Budapest, Hungary

Received 26 October 1974

The probability that a random graph with n vertices and $cn \log n$ edges contains a Hamiltonian circuit tends to 1 as $n \to \infty$ (if c is sufficiently large).

By a graph we mean a graph without loops and multiple edges. We denote by (p, q) the edge between the vertices p and q. The edges (p_1, p_2), (p_2, p_3), ..., (p_{n-1}, p_n) form a *path* if $p_i \neq p_j$. We denote this path by $U(p_1, p_2, ..., p_n)$ By the length of a path we mean the number of its edges.

The edges (p_1, p_2), (p_2, p_3), ..., (p_{n-1}, p_n), (p_n, p_1) form a *circuit* if $p_i \neq p_j$ and $n \geq 3$. By the *length* of a circuit we mean the number of its edges.

A path (respectively a circuit) is in a graph G if every edge of the path (circuit) occurs in G.

We call a path passing through every vertex (i.e., having the length $n - 1$) a *Hamiltonian line*, a circuit passing through every vertex (i.e., having the length n) a *Hamiltonian circuit*.

One more notation: $|X|$ denotes the number of elements of the set X.

Erdős and Rényi raised the following problem: For what function $f(n)$ does the probability that a random graph with n vertices and $f(n)$ edges contains a Hamiltonian circuit tend to 1 as $n \to \infty$?

Erdős and Rényi showed that $f(n) = \frac{1}{2} n \log n$ guarantees neither the connectivity of the graph, nor the existence of a 1-factor, with probability tending to 1. (We do not define these notions because we shall not need them. It is enough to know that if a graph G contains a Hamiltonian circuit, then G is connected and – in the case of an even number of points – contains also a 1-factor.)

The best result in the other direction is due to Komlós and Szemerédi

[3]. They pointed out that $f(n) = cne^{\sqrt{\log n}}$ edges already guarantee the existence of a Hamiltonian circuit with probability tending to 1.

In this paper, we shall show that, for a sufficiently large c, as few as $cn \log n$ suffice.

Let G be an arbitrary graph, and let $U(x_1, x_2, ..., x_k)$ be a path of maximum length in G. If G contains the edge (x_1, x_j) $(1 < j < k)$, then, of course, G contains the path $U'(x_{j-1}, x_{j-2}, ..., x_1, x_j, x_{j+1}, ..., x_k)$, too. U' consists of the same vertices as U and, furthermore, has one end point (x_k) in common with U. We call the transformation $U \to U'$ just described an *allowable transformation.* We may perform allowable transformations several times successively ($U \to U'$, $U' \to U''$, etc.), but we have to be careful that x_k always remains an end point, i.e., only the other end point may be changed. Let us consider the set H of the "other end points" of the paths constructed in this way. x_1, the "other end point" of the path U with which we started, is also an element of H.

Let us consider the original path $U(x_1, x_2, ..., x_k)$. Let the set X consist of those vertices, differing from x_k, which do not belong to H and which are not even adjacent on the path U to a point belonging to H. (x_j is adjacent to x_{j-1} and x_{j+1} on the path U.) Thus all points of G not occurring in U are elements of X.

Lemma 1. *A vertex of H and a vertex of X cannot be joined by an edge.*

Proof. (1) A point p of H and a point q not occurring in U cannot be joined by an edge. In fact, because p is in H, we can form a path U^* having p and x_k as end points by means of allowable transformations. By adding the edge (p, q) to U^*, we would obtain a path in G that is longer than U, and this is impossible.

(2) Assume that the vertices x_i and x_j are joined by an edge, and $x_i \in H$, $x_j \in X$ $(1 \leqslant i < k,\ 1 < j < k)$. By the definition of H, there exists a path $U^*(x_i, ..., x_j, ..., x_k)$ which can be obtained from U by means of allowable transformations.

If the vertices adjacent to x_j in U^* are the same as in U, then the new end point of the path U^{**} that can be formed from U^* with the aid of the edge (x_i, x_j) will be adjacent to x_j on U^* and hence on U. Thus x_j is adjacent to some element of H, which is impossible, since $x_j \in X$.

But if the vertices adjacent to x_j in U^* are not the same as in U, then this means that one of the edges (x_j, x_{j-1}) and (x_j, x_{j+1}) must have been erased during one of the transformations which led from U to U^*. But when we erase an edge, one of its vertices becomes the new end point of the path. (For instance, we get the path $U_2(y_{i-1}, y_{i-2}, ..., y_1, y_i, y_{i+1}, ..., y_{k-1}, x_k)$ from $U_1(y_1, y_2, ..., y_{k-1}, x_k)$ with the aid of (y_1, y_i). Here we erased from U_1 only the edge (y_{i-1}, y_i) and, indeed, y_{i-1} has become the new end point.) We have shown, therefore, that one of the points x_{j-1}, x_j, x_{j+1} belongs to H. But this is impossible, since $x_j \in X$. This proves Lemma 1.

Remark. If we assume that the number of the vertices of G is n and $|H| = p$, then $|X| \geqslant n - 3p$.

Now we pass on to the examination of random graphs.

Lemma 2. *Assume that the edges of the graph G with n vertices are drawn in, mutually independently, with probability $(c \log n)/n$. (In other words, we consider n vertices, and we join the pairs of vertices by an edge, mutually independently, with probability $(c \log n)/n$. We denote the graph that arises in this way — and thus depends on chance — by G.) If c is sufficiently large, then the probability that, for some $p \leqslant \frac{1}{4}n$, there exists a set A of p vertices and a set B — disjoint from A — of $n - 3p - 1$ vertices such that no edge joins a vertex of A to a vertex of B, tends to 0 as $n \to \infty$.*

Proof. The probability in question can be estimated from above as follows:

$$\sum_{p=1}^{[n/4]} \binom{n}{p} \binom{n}{n-3p-1} \left(1 - \frac{c \log n}{n}\right)^{p(n-3p-1)} \leqslant$$

$$\leqslant \sum_{p=1}^{[n/4]} n^{4p+1}\, \mathrm{e}^{(-c \log n)p(n-3p-1)/n} \leqslant \sum_{p=1}^{[n/4]} n^{4p+1-cp/5} \to 0.$$

(We have employed $n - 3p - 1 \geqslant \frac{1}{5}n$ and $c \geqslant 30$.)

Theorem 1. *Assume that the edges of the graph G with n vertices are drawn in mutually independently with probability $(c \log n)/n$. Then, for a sufficiently large c, the probability that G contains a Hamiltonian line tends to 1 as $n \to \infty$.*

Proof. (Due to L. Lovász). Let us introduce notations for the following events.

K: For some $p \leqslant \frac{1}{4}n$, there exists a set A of p vertices and a set B – disjoint from A – of $n - 3p - 1$ vertices, such that no edge joins a vertex of A to a vertex of B.

$L(x)$: Any path of maximum length in G passes through x. (x is an arbitrary vertex of G.)

M: G contains a Hamiltonian line.

$\bar{A}$ denotes the complement of the event A.

We shall deal with the estimation of $\mathbf{P}(L(x))$, where x is a fixed vertex of G. Let us denote by $G(x)$ the graph with $n-1$ vertices obtained from G by erasing x.

$L(x)$: A path of maximum length in G.

Let us choose arbitrarily one of the paths of maximum length in $G(x)$ Denoting this path by U, we define the sets H and X *in* $G(x)$ (!) (see Lemma 1.)

We consider two cases.

(1) $|H| \leqslant \frac{1}{4}n$.

(2) $|H| > \frac{1}{4}n$.

The first case involves the occurrence of the event K, since $|H| = p$, $p \leqslant \frac{1}{4}n$, $|X| \geqslant n - 1 - 3p$. (The number of vertices of $G(x)$ is equal to $n-1$.)

In the second case, the occurrence of $\overline{L(x)}$ would imply that x could not be joined in G with any element of H (otherwise, by the definition of H, some path U^* that is obtainable from U could be elongated by x). The probability of this is

$$\leqslant \left(1 - \frac{c \log n}{n}\right)^{n/4} \leqslant e^{(n/4)(-c \log n/n)} = n^{-c/4}.$$

(Since the definition of H depends only on $G(x)$ and U, the situation of the edges between x and $G(x)$ is independent of the choice of H.)

Consequently, $\mathbf{P}(\overline{L(x)} \text{ and } \bar{K}) \leqslant n^{-c/4}$. Hence

$$\mathbf{P}(\text{there exists an } x \text{ such that } \overline{L(x)}; \text{ and } \bar{K}) \leqslant n^{1-c/4}.$$

By virtue of Lemma 2, $\mathbf{P}(K) \to 0$ ($n \to \infty$) for a sufficiently large c, thus

$$\mathbf{P}(\text{there exists an } x \text{ such that } \overline{L(x)}) \leqslant n^{1-c/4} + \mathbf{P}(K) \to 0,$$

i.e.

$$\mathbf{P}(\text{for every } x, L(x)) \to 1.$$

This means that, with probability tending to 1, every path of maximum length in G passes through all points of G. Thus Theorem 1 is proved.

In the following, let c denote a number for which the conclusions of Theorem 1 and Lemma 2 hold.

Theorem 2. *Suppose that the edges of the graph G with n vertices are drawn in, mutually independently, with probability $(c_1 \log n)/n$. Then, for a sufficiently large c_1, the probability that G contains a Hamiltonian circuit tends to 1 as $n \to \infty$.*

Proof. Let us consider n vertices and construct two random graphs G_1 and G_2 on them, such that a pair of vertices is joined by an edge with probability $(c \log n)/n$ in G_1, with probability $(\log n)/n$ in G_2, and all the events mentioned are independent. Let G be the union of G_1 and G_2 (i.e., the edges of G are the edges occurring in G_1 or G_2). G itself is also a random graph in which any two points are joined by an edge, mutually independently, with probability

$$\frac{c \log n}{n} + \frac{\log n}{n} - \frac{c \log n}{n} \frac{\log n}{n}.$$

By Theorem 1, the probability that G_1 contains a Hamiltonian line tends to 1. Consider (if there is one) a Hamiltonian line $U(x_1, x_2, ..., x_n)$ in G_1. With the aid of U we define the sets H and X (see Lemma 1). We shall achieve our purpose by distinguishing the cases $|H| \leqslant \frac{1}{4} n$ and $|H| > \frac{1}{4} n$.

If G contains no Hamiltonian circuit, then one out of the following three events of small probability must occur:

(1) There is no Hamiltonian line in G_1.

(2) (U, H defined as above) $|H| \leqslant \frac{1}{4} n$ (by Lemmas 1 and 2, the probability of this event tends to 0).

(3) $|H| > \frac{1}{4} n$. In this case, x_n cannot be joined by edges in G_2 to elements of H. (For, suppose that x_n is joined to an $h \in H$. By the definition of H, there exists a path U^*, with end points x_n and h, which can be obtained by allowable transformations from U. But U^*, together with the edge (x_n, h), is then a Hamiltonian circuit.) The probability of the event that there is no edge in G_2 between x_n and the elements of H is $\leqslant (1 - \log n/n)^{n/4} \to 0$.

Thus the exclusion of these three events with probabilities tending to 0 guarantees the existence of a Hamiltonian circuit. This completes the proof of Theorem 2 ($c_1 = c + 1$).

Theorem 3. *Let us consider n vertices and place* $[c_1 n \log n]$ *edges between them at random. The graph G so arising contains a Hamiltonian circuit with probability tending to* 1. *(*c_1 *is a number for which Theorem 2 holds.)*

Proof. Let us perform the following experiment.

We consider n vertices. We join by an edge every pair of vertices, mutually independently, with probability $(c_1 \log n)/n$. G_1 denotes the resulting graph. Now, if the number of edges in G_1 is less than $[c_1 n \log n]$, then we place additional edges in G_1 at random until the number of edges is equal to $[c_1 n \log n]$. If the number of edges of G_1 is at least $[c_1 n \log n]$, then we have nothing else to do.

We denote by G_2 the graph so constructed (whether identical with G_1 or not).

Let us introduce notations for the following events:

R: G_2 contains a Hamiltonian circuit.

S: The number of edges of G_1 is less than $[c_1 n \log n]$.

$\mathbf{P}(S) \to 1$ can be easily obtained from the Chebysheff inequality. Also, it follows from Theorem 2 that $\mathbf{P}(R) \to 1$ (since even G_1 contains a Hamiltonian circuit with probability tending to 1). Hence $\mathbf{P}(R \mid S) \to 1$. On the other hand, $\mathbf{P}(R \mid S)$ is just the probability we wanted, that is, the probability that a random graph with $[c_1 n \log n]$ edges contains a Hamiltonian circuit. In fact, the graphs G_2 that arise when S holds have precisely $[c_1 n \log n]$ edges, and each one has the same probability.

Thus we have finished the proof of Theorem 3.

I wish to thank L. Lovász for his ingenious simplification of the original proof of this theorem.

References

[1] P. Erdös and A. Rényi, On the evolution of random graphs, Mat. Kut. Int. Kozl. 5. (1960) 17–60.

[2] P. Erdös and A. Rényi, On the existence of a factor of degree one of connected random graphs, Acta Math. Acad. Sci. Hungar. 17 (1966) 359–379.

[3] J. Komlós and E. Szemerédi, Hamiltonian cycles in random graphs, in: A. Hajnal, R. Rado and V.T. Sós (eds.), Infinite and Finite Sets, Colloq. Math. Soc. János Bolyai 10 (North-Holland, Amsterdam 1975) 1003–1011.