Discrete Mathematics 91 (1991) 141–148 North-Holland 141

Pancyclic Hamilton cycles in random graphs

C. Cooper

Department of Computing Mathematics and Statistics, Polytechnic of North London, Holloway Rd, London, UK, N78DB

Received 28 September 1988 Revised 11 October 1989

Abstract

Cooper, C., Pancyclic Hamilton cycles in random graphs, Discrete Mathematics 91 (1991) 141-148.

Let $\mathscr{G}(n, p)$ denote the probability space of the set \mathscr{G} of graphs $G = (V_n, E)$ with vertex set $V_n = \{1, 2, \ldots, n\}$ and edges E chosen independently with probability p from $\mathscr{C} = \{\{u, v\}: u, v \in V_n, u \neq v\}$.

A graph $G \in \mathcal{G}(n, p)$ is defined to be pancyclic if, for all $s, 3 \le s \le n$ there is a cycle of size s on the edges of G. We show that the threshold probability $p = (\log n + \log \log n + c_n)/n$ for the property that G contains a Hamilton cycle is also the threshold probability for the existence of a 2-pancyclic Hamilton cycle, which is defined as follows. Given a Hamilton cycle H, we will say that H is k-pancyclic if for each s ($3 \le s \le n - 1$) we can find a cycle C of length s using only the edges of H and at most k other edges.

1. Introduction

Let $\mathscr{G}(n, p)$ denote the probability space of the set \mathscr{G} of graphs $G = (V_n, E)$ with vertex set $V_n = \{1, 2, ..., n\}$ and edges E chosen independently with probability p from $\mathscr{C} = \{\{u, v\}: u, v \in V_n, u \neq v\}$.

A graph $G \in \mathscr{G}(n, p)$ is defined to be pancyclic if, for all $s, 3 \le s \le n$ there is a cycle of size s on the edges of G. We show that the threshold probability $p = (\log n + \log \log n + c_n)/n$ for the property that G contains a Hamilton cycle is also the threshold probability for the existence of a 2-pancyclic Hamilton cycle, which is defined as follows. Given a Hamilton cycle H, we will say that H is k-pancyclic if for each s ($3 \le s \le n - 1$) we can find a cycle C of length s using only the edges of H and at most k other edges $e_1, e_2, \ldots, e_k \notin E(H)$. We shall refer to the edges e_i as chords of H and such a cycle C as a k-cycle.

The threshold for the existence of Hamilton cycles in $\mathcal{G}(n, p)$ was established by Komlós and Szemerédi [5] and is the same as that for minimum vertex degree at least 2 (Erdős and Rényi [3]). The question of the threshold for pancyclic graphs was raised by Korshunov [6] and has been studied by Luczak [7] and Cooper and Frieze [2], and was found to be the same as the threshold for the existence of Hamilton cycles. We now show this is also the threshold for the 2-pancyclic property.

Theorem 1.1. Let

$$p = \frac{\log n + \log \log n + c_n}{n}$$

then

$$\lim_{n \to \infty} \Pr(G \in \mathscr{G}(n, p) \text{ is 2-pancyclic}) = \begin{cases} 0 & c_n \to -\infty, \\ e^{-e^{-c}} & c_n \to c, \\ 1 & c_n \to \infty \end{cases}$$
(1.1)

 $(= \lim_{n \to \infty} \Pr(G \in \mathscr{G}(n, p) \text{ has minimum degree } \geq 2)).$

We conjecture that with p as given in Theorem 1.1, then p is the threshold probability for G to be 1-pancyclic.

2. Proof of Theorem 1.1

2.0. Construction of a multigraph

Our proof is based on constructing a set of multigraphs for which the set of underlying graphs is $\mathcal{G}(n, p)$.

We start with $\mathscr{G}(n, p_r)$ where $p_r = (\log n + 6 \log \log n)/2n$. In order to map $\mathscr{G}(n, p_r)$ onto $\mathscr{G}(n, p)$ we go from $\mathscr{G}(n, p_r)$ into a space of multigraphs by generating a second set of edges from \mathscr{E} with independent edge probability $\pi = 1 - (1 - p)/(1 - p_r)$, and then fuse multiple edges to give the underlying set of graphs. For some edges we do this directly, whilst for others we generate a set of edges with probability equivalent to π .

Step 1. Given $G \in \mathcal{G}(n, p_r)$ with edge set E_r let $A = \{v: d(v) < 2\}$ be the set of vertices of G of degree less than 2. Let $\mathscr{C} = \mathscr{C}_1 \cup \mathscr{C}_2$ be the set of potential edges from which selections are to be made independently with probability π , and let $\mathscr{C}_1 = \{\{u, v\}: u \in A, v \in V_n\}$ and $\mathscr{C}_2 = \{\{u, v\}: u, v \notin A\}$. Let each edge in \mathscr{C}_1 be chosen independently with probability π and call the set of edges so selected E'. Let Red = $E_r \cup E'$ be the edges of the resulting multigraph M(Red) where we regard the edges as coloured red. We will write $GM(\text{Red}: \delta \ge 2)$ if the underlying graph GM(Red) has minimum degree at least 2.

Step 2. For each $e = \{u, v\} \in \mathcal{E}_2$ we choose $\lambda = 8 \log 2(n/\log n)$ independent copies of e, $(e(1), \ldots, e(\lambda))$ with probability $p_b = 100(\log n/n^2)$, and call the set of selected edges Blue. The multigraph at this stage will be referred to as M(Red, Blue).

Step 3. We now choose independently a green edge for each $e \in \mathscr{E}_2$ with probability p_e and call the set of chosen edges Green, where

$$p_{g} = 1 - \frac{1 - p}{(1 - p_{r})(1 - p_{b})^{\lambda}} \ge \frac{\log n - 6 \log \log n - 1600 \log 2 - 2c_{n}}{2n}.$$

The proof of the next lemma follows standard lines and is given in the Appendix, A.

Lemma 2.1. Almost every (a.e.) $GM = GM(\text{Red}: \delta \ge 2, \text{Blue})$ contains a Hamilton cycle.

We now prove that a.e. multigraph found to have the Hamilton cycle property has the pancyclic Hamilton cycle property. In order to do this we fix our red-blue Hamilton cycle and relabel the vertices clockwise round the cycle as v_1, v_2, \ldots, v_n starting from vertex 1. Provided $v_i, v_{i+k} \in V_n - A$ then $\{v_i, v_{i+k}\} \in \mathcal{C}_2$ and is thus potentially an edge of the green subgraph $(\{v_i, v_{i+k}\} \in$ Green). We will call such an edge a k-chord with initial vertex v_i and terminal vertex v_{i+k} .

There are various constructions available for finding 2-cycles of a suitable size. We use two; the first based on 'triangles' works well for the smaller cycles, and the second based on 'rectangles' for the larger ones.

Type A cycles (triangles). For fixed v_i ($i \le n - 2s + 3$) and fixed cycle size s, partition the vertices v_{i+2}, \ldots, v_n on the Hamilton cycle into $\lfloor (n - i - 1)/(2s - 4) \rfloor$ sets of size 2s - 4 running sequentially from vertex v_{i+2} . Thus the first such set is $\{v_{i+2}, \ldots, v_{i+2s-3}\}$. On this set we can construct s - 2 potential triangular cycles using the chord pairs

$$(\{v_i, v_{i+2}\}, \{v_i, v_{i+s}\}), (\{v_i, v_{i+3}\}, \{v_i, v_{i+s+1}\}), \ldots, \\ (\{v_i, v_{i+s-1}\}, \{v_i, v_{i+2s-3}\}).$$

The existence of any pair $(\{v_i, v_{i+j}\}, \{v_i, v_{i+j+s-2}\})$ forms a cycle

 $(v_i, v_{i+j}, v_{i+j+1}, \ldots, v_{i+j+s-3}, v_{i+j+s-2}, v_i).$

For fixed *i* each pair of edges is examined only once, and as *i* runs from 1 to n - 2s + 3 no edge is ever reused for fixed *s*. Thus for fixed *s* we examine

$$\sum_{i=1}^{n-2s+3} \left\lfloor \frac{n-i+1}{2s-4} \right\rfloor (s-2) \ge \frac{1}{4}(n-2s)^2$$

independent pairs of edges, of which at most 3na will have either an initial or terminal vertex in A, (|A| = a).

Type B cycles (rectangles). Let $j \in \{1, 2, ..., \lfloor (k-1)/2 \rfloor\}$. For fixed k and fixed j, consider the pairs of potential green j- and (k-j)-chords $(\{v_i, v_{i+j}\}, \{v_i, v_{i+k-j}\})$ where the index i + j is understood to be i + j - n whenever i + j > n and $t \in \{i + j, ..., n + i - k + j\}$. The existence of such chords forms the follow-

C. Cooper

ing cycle of length n - k + 2

$$(v_i v_{i+j} v_{i+j+1}, \ldots, v_i v_{i+k-j} v_{i+k-j+1}, \ldots, v_{i-1} v_i).$$

Let $X_{k,j}$ be the number of cycles of length s = n - k + 2 we can form for fixed j, k. We note that for $j \neq j' X_{k,j}$ and $X_{k,j'}$ are independent random variables. For fixed j, k we have n(n-k) potential chord pairs of which at most 2a(n-k) + 2an will have a vertex in A.

Conditional on $|A| < \sqrt{n}/\log n$ which by Lemma A.1 holds for a.e. $GM(\text{Red}: \delta \ge 2, \text{ Blue, Green})$ we have

Lemma 2.2. For Type B cycles where $10n/\log n \le s \le n-1$ and s = n-k+2

$$\Pr(X_{k,j}=0) \leq \frac{5n}{(n-k)\log n}$$

Proof. Let $X_{k,j} = X$ for fixed k, j. We first note that

$$(n-\sqrt{n})(n-k)p_g^2 \leq E(X) \leq n(n-k)p_g^2.$$

To calculate E(X(X-1)) consider all ordered pairs of chord pairs indexed by the initial vertices of the *j*, (k-j)-chords, and written $((v_{\alpha}, v_{\beta}), (v_{\alpha'}, v_{\beta'}))$. Let *Y* count pairs where $\alpha \neq \alpha', \beta \neq \beta'$ and *Z* all others. Then

$$(n - \sqrt{n})_2(n - k)_2(p_g^2)^2 \leq E(Y) \leq (n)_2(n - k)_2(p_g^2)^2$$

allowing for A and also for the case when $v_{\beta} \in \{v_{\alpha'}, \ldots, v_{\alpha'+j}\}$ and thus does not restrict the choice of $v_{\beta'}$. In the case where not all the chords are distinct we have

$$E(Z) \leq n(n-k)p_g^2 + n(n-k)_2 p_g^3 + (n)_2(n-k)p_g^3$$

and thus $\operatorname{Var}(X) \leq (2 + o(1))n^2(n - k)p_g^3$. Applying $P(X = 0) \leq \operatorname{Var}(X)/E(X)^2$ the lemma follows. \Box

Proof of Theorem 1.1. We prove that a.e. $G \in \mathcal{G}(n, p)$ of minimum degree 2 contains a 2-pancyclic Hamilton cycle.

We first note that the equivalence classes of GM(Red, Blue, Green), where the equivalence relation is 'the same uncoloured graph' are exactly a $G \in \mathcal{G}(n, p)$ as each edge is chosen independently with probability p.

Case 1: $3 \le s \le n/3$.

Using type A cycles we have

Pr(∃s: G does not contain a type A cycle of length s,
$$3 \le s \le n/3$$
)
≤ $n/3(1-p_g^2)^{\frac{1}{2}(n-2s)^2(1-o(1))} \le ne^{-(\log n/13)^2} = o(1)$

144

Case 2: $n/3 \le s \le n - n^{\frac{1}{4}}$.

We use type B cycles for $n^{\frac{1}{4}} \le k \le 2n/3$ and s = n - k + 2. Let Y_i be defined as follows:

 $Y_j = 1$ if $X_{k,j} > 0$, $Y_j = 0$ otherwise.

Thus $\sum Y_j$ is a lower bound for the number of cycles of size s. The inequality of Hoeffding [4] gives the following probability bound for sums of independent random variables taking values in the interval [0, 1],

$$\Pr(\sum Y_j \leq (1 - \varepsilon)E(\sum Y_j)) \leq e^{-(\varepsilon^2/2)E(\sum Y_j)} \quad \varepsilon \in (0, 1)$$

Thus

$$\Pr\left(\sum_{j=1}^{\lfloor \frac{k-1}{2} \rfloor} Y_j \le (1-\varepsilon) \left\lfloor \frac{k-1}{2} \right\rfloor \left(1 - \frac{15}{2\log n}\right)\right) \le e^{-(\varepsilon^{2/4})k(1-o(1))}$$

which holds simultaneously over the range of k.

Case 3: $n - n^{\frac{1}{4}} \le s \le n - 1$.

A simple construction using single (s-1)-chords is adequate here. Any such chord divides the Hamilton cycle into a cycle of length s, and one of length n-s. Pr(No(s-1)-chord exists for some s,

$$n - n^{\frac{1}{4}} \le s \le n - 1) \le n^{\frac{1}{4}} (1 - p_g)^{n - 2\sqrt{n}} = o(1).$$

Appendix A. Proof of Lemma 2.1

The proofs in this appendix follow standard lines, (see for example Bollobás [1]) save for slight differences in the edge probabilities used.

Lemma A.1. (a) Almost every (a.e.) $G \in \mathcal{G}(n, p_r)$ satisfies the following:

(i) G contains a path of length $n(1 - (8 \log 2)/\log n)$.

(ii) If $A_i = \{v: d(v) = i\}$ i = 0, 1 then $|A_0 \cup A_1| = |A| \le \sqrt{n/\log n}$.

(iii) $G = C \cup A_0$ where C is a connected 'giant' component.

- (b) $\lim_{n\to\infty} \Pr(\delta(GM(\text{Red})) \ge 2)$ is given by (1.1).
- (c) A.e. $GM(\text{Red}: \delta \ge 2)$ is connected.

Proof. (a)(i) follows from Fernandez de la Vega's path algorithm [9] which states that if $\pi = \theta/n$, $0 < \theta < \log n - 3 \log \log n$, then a.e. $G \in \mathcal{G}(n, \pi)$ contains a path of a least $n(1 - (4 \log 2)/\theta)$.

(a)(ii) follows from $E(|A_0|) = O(\sqrt{n} \log^{-3} n)$, $E(|A_1|) = O(\sqrt{n} \log^{-2} n)$ and the Markov inequality.

(a)(iii) follows from:

(1) If T_k is the number of trees of order k, then $T_k = 0$ in a.e. $G \in \mathcal{G}(n, \pi)$ provided

$$\pi = \frac{\log n + (k-1) \log \log n + \omega(n)}{kn}, \quad \omega(n) \to \infty.$$

(2) For $\pi = \omega(n)/n$, $\omega(n) \to \infty$, every component of G with the exception of the giant component is a tree, a.e. $G \in \mathcal{G}(n, \pi)$.

See for example Bollobás [1, p. 94 Theorem 4(v) and p. 136 Theorem 10(iii)].

(b) Given $G \in \mathcal{G}(n, p_r)$, let X_1 count the vertices of degree 0, and let X_2 count the vertices of degree 0 in the corresponding GM(Red).

$$E(X_2) = E(E(X_2 \mid X_1 = x_1)) = E((1 - \pi)^{n-1}X_1)$$

= $(1 - \pi)^{n-1}n(1 - p_r)^{n-1} = O(\log^{-1} n)$

Similarly, let $Y_2^{(j)}$ (j = 0, 1) count the vertices of degree 1 in GM(Red) coming from vertices of degree j in G and let Y_1 count the vertices of degree 1 in G.

$$E(Y_2^{(1)} + Y_2^{(0)}) = E(E(Y_2^{(1)} | Y_1 = y_1)) + E(E(Y_2^{(0)} | X_1 = x_1))$$

= $E((1 - \pi)^{n-2}Y_1) + E((n - 1)\pi(1 - \pi)^{n-2}X_1)$
= $(1 - \pi)^{n-2}n(n - 1)p_r(1 - p_r)^{n-2}$
+ $(n - 1)\pi(1 - \pi)^{n-2}n(1 - p_r)^{n-1}$
= $n(n - 1)p(1 - p)^{n-2} = (1 + o(1))e^{-c}$.

Similar calculations show that the *t*th factorial moment of $Y_2^{(1)} + Y_2^{(0)}$ is asymptotic to $(e^{-c})^t$ thus providing the required Poisson parameter of e^{-c} for (1.1),

(c) Let
$$B \subseteq A_0$$
, $|B| = b$.

 $\phi(b) = E(E(\text{number of sets } B \text{ not connected to } V_n - B \mid X_1 = x_1))$

$$= E\left(\binom{X_1}{b}(1-\pi)^{b(n-b)}\right) = (1-\pi)^{b(n-b)}\binom{n}{b}(1-p_r)^{b(n+b)+\binom{k}{2}}$$

where the final part of the above term is the expected number of b subsets of vertices of degree zero in G. Thus $\sum_{b} \phi(b) = O(\log^{-1} n)$. \Box

Lemma A.2. For a.e. $GM \in \{GM(\text{Red}: \delta \ge 2)\}$, for all $U \subset V_n$, if $|U| \le n/4$ then $|U \cup \Gamma(U)| \ge 3 |U|$ where $\Gamma(U)$ is the disjoint neighbour set of U.

Proof. We prove the lemma using p_r on the condition that $\delta \ge 2$ and note that increasing some of the edge probabilities to p will not affect the results as the property is monotone increasing. We assume that $|\Gamma(U)| < 2|U|$ and show the result follows for a.e. GM by contradiction.

146

Case 1: $n^{\frac{3}{5}} \le u \le n/4$. The expected number of such sets |U| = u, $|\Gamma(U)| = w$ for fixed u, w is

$$E(u, w) = {\binom{n}{u}} {\binom{n-u}{w}} (1-(1-p_r)^u)^w (1-p_r)^{u(n-(u+w))}$$
$$\leq \left(\frac{ne^{-\theta+1+\theta u/n}}{u}\right)^u \left(\frac{u\theta e^{1+u\theta/n}}{w}\right)^w,$$

where $p_r = \theta/n$ and $\theta = (\log n + 6 \log \log n)/2$ and $(1 - (1 - p_r)^u)^w \le (p_r u)^w$. Thus

$$\sum_{w=1}^{2u-1} E(u, w) \leq (2u) \left(\frac{n\theta^2}{4u} e^{-\theta + 3 + 3u\theta/n}\right)^u \text{ as } \left(\frac{u\theta e^{1+u\theta/n}}{w}\right)^w$$

is monotone increasing for $w \in \{1, 2, ..., 2u - 1\}$. Finally

$$\sum_{u=n^{3/5}}^{n/4} \sum_{w=1}^{2u-1} E(u, w) \leq \sum_{u=n^{3/5}}^{n/4} (2u) \left(\frac{2n^{\frac{1}{2}+3u/2n} (\log n)^{-1+9u/n}}{u}\right)^{u} \leq O(n^{-\sqrt{n}}).$$

Case 2: $t \ge 7$, $u \le n^{\frac{3}{5}}$.

Separate the induced subgraph $G[U \cup \Gamma(U)]$ into connected components $(U_i, \Gamma(U_i))$ where at least one component T satisfies $|\Gamma(U_i)| < 2 |U_i|$. Let |T| = t and by hypothesis $|U_i| = u$, where $u \ge \lfloor t/3 \rfloor$. As T is connected, it must contain at least a tree, so the number of edges h in G[T] is at least t-1. For fixed t, the expected number E(u, t) of such components is at most

$$\sum_{h \ge t-1} \binom{n}{t} \binom{t}{u} \binom{\binom{t}{2}}{h} p_r^h (1-p_r)^{u(n-t)}.$$
(A.2.1)

Replacing u by $\lfloor t/3 \rfloor$ and summing over t we have

$$\sum_{t=7}^{3n^{3/5}} \sum_{h \ge t-1} \binom{n}{t} \binom{t}{\lceil t/3 \rceil} \binom{\binom{t}{2}}{h} p_r^h (1-p_r)^{\lceil t/3 \rceil (n-t)} \\ \le \sum_{t=7}^{3n^{3/5}} \binom{n}{t} (1-p_r)^{t/3(n-t)} \binom{t}{\lceil t/3 \rceil} \sum_{h \ge t-1} \binom{t^2 e p_r}{2h}^h \\ \le \sum_{t=7}^{3n^{3/5}} \frac{n}{\log n} (3e^2 (\log n)^{t/n} n^{-\frac{1}{6}(1-t/n)})^t = o(1).$$

For if

$$A_{h} = \left(\frac{t^{2}ep_{r}}{2h}\right)^{h}$$
 then $A_{h+1} \le \left(\frac{t^{2}p_{r}}{2h}\right)A_{h}$ as $\left(1 + \frac{1}{h}\right)^{-(h+1)} \le e^{-1}$.

Thus $A_{h+1} < ((t \log n)/n)A_h$ for $h \ge t-1$.

Case 3: $t \leq 6$

As $\delta \ge 2$ the only possible pairs (u, t-u) contradicting t < 3u are $\{(i, j): 2 \le i \le 5, i+j \le 6, 1 \le j < 2i\}$. Using expression (A.2.1) with these feasible values of

u and t gives a term of order

 $O(n^{t}p_{r}^{t-1}(1-p_{r})^{u(n-t)}) = O(n^{1-u/2}(\log n)^{t-1-3u})$

which is at most $O(\log^{-1} n)$. \Box

Lemma 2.1. A.e. $GM = GM(\text{Red}: \delta \ge 2, \text{ Blue})$ contains a Hamilton cycle.

Proof. We assume all longest paths of length l in GM cannot be extended and apply Pósa [8] type path rotations to obtain a contradiction. Suppose that the vertices of the longest path $x_0 P x_1$ we have selected are labelled $v_1 v_2, \ldots, v_l$ and $v_1 = x_0$. By $\delta \ge 2$ there is an edge $\{v_i, w\}$ and w must be a path vertex $(v_i \text{ say})$ or the path will extend by at least one vertex. We rotate on $\{v_i, v_i\}$ giving a new path $v_1v_2, \ldots, v_iv_iv_{i-1}, \ldots, v_{i+1}$. Fix $x_0 = v_1$ and let $U = \{v: \text{ rotation endpoints}\}$ of $x_0 P x_1$. Let $\Gamma(U)$ be the disjoint neighbour set of U in the graph, then by maximal path length each $v \in \Gamma(U)$ is adjacent to some $u \in U$ on the path. The set of path adjacent vertices is at most 2|U| - 1, so $|U| \ge n/4$ by Lemma A.2. Thus $|U \cap A^{c}| \ge n/5$ say, where $A^{c} = V_{n} - A$. A similar argument fixing x_{1} gives us at least $n^2/50$ potential edges, any one of which would close the path into a cycle and allow path extension by the connectivity hypothesis. We have to extend the path at most $\lambda = (8 \log 2)n/\log n$ times. Using the *i*th set of blue edges $\{e(i)\}$ for the *i*th extension of the path, the probability there is no edge between the endpoint sets is $(1 - p_b)^{n^2/50} = O(n^{-2})$. We conclude GM contains the required Hamilton cycle. \Box

Acknowledgement

We would like to thank Trevor Fenner and Alan Frieze for their helpful criticisms and suggestions on the contents of this paper.

References

- [1] B. Bollobás, Random Graphs (Academic Press, New York, 1985).
- [2] C. Cooper and A. Frieze, Pancyclic random graphs, Proc. 1987 Conf. on Random Graphs, Poznan.
- [3] P. Erdős and A. Rényi, On the strength of connectedness of a random graph, Acta. Math. Acad. Sci. Hungar. 12 (1961) 261–267.
- [4] W. Hoeffding, Probability inequalities for sums of bounded random variables, J. Amer. Statist. Assoc. 58 (1963) 13-30.
- [5] J. Komlós and E. Szemerédi, Limit distribution for the existence of Hamilton cycles in a random graph, Discrete Math. 49 (1983) 55-63.
- [6] A. Korshunov, Solution of a problem of Erdős and Rényi on Hamilton cycles in non-oriented graphs, Soviet Math. Dokl. 17 (1976) 760-764.
- [7] T. Luczak, Cycles in random graphs (1987) to appear.
- [8] L. Pósa, Hamilton cycles in random graphs, Discrete Math. 14 (1976) 359-364.
- [9] W. Fernadez de la Vega, Long paths in random graphs, Studia. Sci. Math. Hungar. 14 (1979) 335-340.