RANDOM NEAR-REGULAR GRAPHS AND THE NODE PACKING PROBLEM *

G.R. GRIMMETT

School of Mathematics, University of Bristol, Bristol B58 1TH, UK

W.R. PULLEYBLANK

Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ont. N2L 3G1, Canada

Received April 1985
Revised October 1985

Nemhauser and Trotter [12] proposed a certain easily-solved linear program as a relaxation of the node packing problem. They showed that any variables receiving integer values in an optimal solution to this linear program also take on the same values in an optimal solution to the (integer) node packing problem. Let π be the property of graphs defined as follows: a graph G has property π if and only if there is a unique optimal solution to the linear-relaxation problem, and this solution is completely fractional. If a graph G has property π then no information about the node packing problem on G is gained by solving the linear relaxation. We calculate the asymptotic probability that a certain type of 'sparse' random graph has property π, as the number of its nodes tends to infinity. Let m be a fixed positive integer, and consider the following random graph on the node set $\{1, 2, \ldots, n\}$. We join each node, j say, to exactly m other nodes chosen randomly with replacement, by edges oriented away from j; we denote by $G_n(m)$ the undirected graph obtained by deleting all orientations and allowing all parallel edges to coalesce. We show that, as $n \to \infty$,

$$P(G_n(m) \text{ has property } \pi) \to \begin{cases} 0 & \text{if} \quad m = 1, \\ 1 & \text{if} \quad m \geq 3, \end{cases}$$

and we conjecture that $P(G_n(2) \text{ has property } \pi) \to (1 - 2e^{-2})^{1/2}$.

node packing * random graphs

1. The node packing problem

The node packing problem is the following.

(NP) Given a graph $G = (V, E)$, find a set $S \subseteq V$ such that the following two conditions hold:
(1.1) no two nodes of S are adjacent,
(1.2) $|S|$ is maximized subject to (1.1).

Any set S satisfying (1.1) is called a (*node*) *packing* (or a *stable* or *independent* set). The related weighted problem (WNP) arises when each node v of the graph has a real weight $w(v)$ associated with it, and we seek a packing S for which

$$w(S) = \sum_{v \in S} w(v)$$

is maximized.

*Research supported in part by the Natural Sciences and Engineering Research Council of Canada.

The problem WNP may be formulated naturally as an integer linear program (IP) as follows:

(IP)

Maximize $\sum_{v \in V} w(v)\chi(v)$,

subject to $0 \leq \chi(v) \leq 1$, $\chi(v)$ is an integer, for all nodes $v \in V$, (1.3)

$\chi(u) + \chi(v) \leq 1$ for all edges $\{u, v\} \in E$. (1.4)

If we assume that evergy node is incident with at least one edge, then the upper bounds in (1.3) are implied by (1.4) and can be omitted. The feasible solutions to IP are precisely the incidence vectors of the packings of G, and solving IP is equivalent to solving WNP.

0167-6377/85/$3.30

If we remove the integrality condition from (1.3) we obtain a linear relaxation of the problem which has received some attention. It is well-known that if χ is a basic solution to this relaxation then $\chi(v) \in \{0, \frac{1}{2}, 1\}$ for all $v \in V$. Nemhauser and Trotter [12] showed that if an optimal solution to this relaxation has any integer components then these components may be 'fixed' at these values and the remaining smaller problem may be solved by some other means.

Theorem 1 (Nemhauser and Trotter [12]). *Let χ^* be a basic optimal solution to the problem*

$$\begin{aligned} &\textit{Maximize} && \sum_{v \in V} w(v)\chi(v), \\ &\textit{subject to} && 0 \le \chi(v) \le 1 \\ & && \quad \text{for all } v \in V, \quad \text{and} \\ & && \chi(u) + \chi(v) \le 1 \\ & && \quad \text{for all } \{u, v\} \in E. \end{aligned}$$

Let $W = \{v \in V : \chi^*(v)$ equals either 0 or 1$\}$. Then there exists an optimal solution $\tilde{\chi}$ to IP such that $\tilde{\chi}(v) = \chi^*(v)$ for all $v \in W$.

Consequently, it is of interest to obtain a solution to the relaxation having the minimum possible number of fractional components, and, in addition, to characterize those problems for which every optimal solution is completely fractional; it is for problems in the last class that the Nemhauser–Trotter approach provides no useful information.

The first question was considered by Picard and Quéyranne [13, 14] (as well as Pulleyblank [15]), where it was shown that a solution with a minimum cardinality set of fractional values can be obtained in polynominal time. The second question was considered in [15] for the special case when all weights equal 1, and so WNP reduces to NP. Before recalling the main result of [15] we need two definitions. A *2-matching* of a graph $G = (V, E)$ is an assignment of the integers 0,1 and 2 to the edgers of G in such a way that the sum of the integers over the edges incident with each node equals exactly 2. The graph G is *bicritical* if $|V| > 1$ and, for every node $v \in V$, $G \setminus v$ has a 2-matching (in [15] such graphs are called '2-bicritical').

Theorem 2 (Pulleyblank [15]). *The following a equivalent for a graph $G = (V, E)$.*

(i) *G is bicritical.*

(ii) *For each edge $e \in E$, there exists a 2-matchi μ of G for which $\mu(e) > 0$, and moreover ea component of G is non-bipartite.*

(iii) *The unique optimal solution to the linear pı gram*

$$\begin{aligned} &\textit{Maximize} && \sum_{v \in V} \chi(v), \\ &\textit{subject to} && 0 \le \chi(v) \le 1 \\ & && \quad \text{for all } v \in V, \quad \text{and} \\ & && \chi(u) + \chi(v) \le 1 \\ & && \quad \text{for all } \{u, v\} \in E, \end{aligned}$$

is obtained by setting $\chi(v) = \frac{1}{2}$ for all $v \in V$.

Henceforth we use the term 'bicritical' to de scribe any graph for which part (iii) of Theorem holds. That is, G is bicritical if it possesses th property π described in the abstract. Various othe characterizations of bicritical graphs are knowr For any $S \subseteq V$, we define the *neighbour set* $N(S$ by

$$N(S) = \{v \in V \setminus S : v \text{ is adjacent to a node of } S\}$$

Theorem 3 (Nemhauser and Trotter [12]). *G i bicritical if and only if $|N(I)| > |I|$ for all non empty node packings I of G.*

A graph G is said to be *regularizable* if it i: possible to replace each edge with a positive (in tegral) number of copies of itself and thereby obtain a regular graph.

Theorem 4 (Berge [1]). *$G = (V, E)$ is bicritical if and only if $|V| > 1$, G is regularizable, and eacl component of G is non-bipartite.*

Note that the question of determining whether or not a graph is bicritical is polynomially verifiable, since it reduces to solving a polynomial number of 2-matching problems, each of which is equivalent to a bipartite matching problem.

It is natural to pose the question of whether or not the property of bicriticality is common amongst graphs. We shall show in this paper that a graph G chosen randomly from a certain family of 'nearly regular' graphs is 'almost surely' bicritical so long as the expected node degrees of G are at least (about) 6; that is to say, the probability that such a graph G is bicritical approaches 1 as the number of vertices of G tends to infinity.

2. Random graphs

Let p satisfy $0 < p < 1$ and let $G_{n,p}$ be a random graph on the nodes $\{1, 2, \ldots, n\}$ in which each edge is present with probability p, independently of the presence or absence of all other edges. The probability that $G_{n,p}$ is bicritical satisfies (see [15])

$$P(G_{n,p} \text{ is bicritical}) - 1 \quad \text{as} \quad n \to \infty.$$

(If $n = 100$ and $p = \frac{1}{2}$ then the probability that $G_{n,p}$ is *not* bicritical is less than 1.4×10^{-8}.) Such random graphs are not attractive candidates as 'typical' optimization problems, since they contain about $\frac{1}{2}pn^2$ edges and have mean vertex degree $p(n-1)$; it is more interesting to consider another family of graphs which have much sparser edge sets and which are 'nearly regular'.

For any fixed positive integer m we construct the random graph $G_n(m)$ on the node set $V_n = \{1,2,\ldots,n\}$ as follows. For each node i, we perform m independent trials, in each case choosing one of the other $n-1$ nodes at random with equal probability $(n-1)^{-1}$. We construct edges joining i to each node thus chosen, and think of these edges as *emanating* from i. We carry out this process for each initial node i, and refer to the ensuing graph as the *original construction*. We obtain $G_n(m)$ by replacing each set of multiple edges by a single edge.

The following theorem is our main result.

Theorem 5. *As $n \to \infty$, we have that*
(a) *$P(G_n(1)$ is bicritical$) \to 0$,*
(b) *$\limsup_{n\to\infty} P(G_n(2)$ is bicritical$) < 1$,*
(c) *if $m \geq 3$ then $P(G_n(m)$ is bicritical$) \to 1$.*

The method of proof of part (b) of this theorem may be extended to show that

$$\limsup_{n\to\infty} P(G_n(2) \text{ is bicritical}) \leq (1 - 2e^{-2})^{1/2},$$

and we have found reason to conjecture that, as $n \to \infty$,

$$P(G_n(2) \text{ is bicritical}) \to (1 - 2e^{-2})^{1/2}.$$

Easy calculations show that, in $G_n(m)$,
(i) the expected number of edges is

$$\binom{n}{2}\left(1 - \left(1 - \frac{1}{n-1}\right)^{2m}\right) \sim mn \quad \text{as} \quad n \to \infty,$$

(ii) the expected degree of node i is

$$(n-1)\left(1 - \left(1 - \frac{1}{n-1}\right)^{2m}\right) \sim 2m$$
as $n \to \infty$,

(iii) the degree $D_n(i)$ of node i has asymptotic distribution, as $n \to \infty$,

$$P(D_n(i) = m + k) \to e^{-m}\frac{m^k}{k!}$$
$$\text{for } k = 0, 1, 2, \ldots \tag{2.1}$$

To see (iii), note that $D_n(i)$ comprises edges of the original construction which either emanate from i or which emanate from other nodes. The number of the latter is approximately binomially distributed with parameters n and m/n, which approaches the Poisson distribution of (2.1) as $n \to \infty$. Furthermore, with probability $1 - o(1)$, no edge emanating from i is parallel to an edge emanating from some other node.

The graph $G_n(m)$ is not regular, with probability $1 - o(1)$, but it is 'near-regular' in the sense that each node degree satisfies (2.1). Such random graphs are very similar to those of Fenner and Frieze [6, 7] and Shamir and Upfal [18], who considered a random graph $F_n(m)$ on V_n with edge set as follows. For each node i, a set of m nodes is chosen randomly, *without* replacement, from the remaining $n-1$ nodes, and edges are constructed between i and this set. $F_n(m)$ is obtained by allowing all parallel edges to coalesce as before. We believe that the global properties of $G_n(m)$ and $F_n(m)$ differ only in trivial respects (for example, the expected number of nodes in $G_n(2)$ and $F_n(2)$ which have degree 1 is approximately e^{-2} and 0, respectively). In most interesting respects we expect that the properties of $G_n(m)$ and $F_n(m)$, as $n \to \infty$, are very much the same. Of course, $G_n(1)$ and $F_n(1)$ have the same distribution, and are actually the graphs obtained in the *random mapping problem* in which a mapping f: $V_n \to V_n$ is chosen such that $\{f(i)\colon i \in V_n\}$ is a family of independent random nodes, $f_n(i)$ being uniformly distributed on $V_n \setminus i$. See Ross [16] for recent progress on this problem.

Fenner and Frieze [6, 7] have shown that, as $n \to \infty$,

$$P(F_n(m) \text{ is connected}) \to \begin{cases} 0 & \text{if} \quad m = 1, \\ 1 & \text{if} \quad m > 1, \end{cases}$$

and that $P(F_n(m)$ is Hamiltonian$) \to 1$ if $m \geq 23$. It is currently an open problem to find the minimal value of m such that $F_n(m)$ (or $G_n(m)$) is Hamiltonian with probability $1 - o(1)$. We note here that

$$P(G_n(2) \text{ is Hamiltonian}) \to 0 \quad \text{as} \quad n \to \infty, \tag{2.2}$$

since a standard application of Chebyshev's inequality (following a suggestion of A.M. Frieze (personal communication)) shows that both $F_n(2)$ and $G_n(2)$ contain, with probability $1 - o(1)$, a vertex with at least three neighbours having degree 2. It is natural to conjecture that $F_n(3)$ and $G_n(3)$ are Hamiltonian with probability $1 - o(1)$. See Bollobás [4] and Fenner and Frieze [8] for related work.

Shamir and Upfal [18] have shown that $P(F_{2n}(6)$ has a perfect matching$) \to 1$ as $n \to \infty$. (A perfect matching is a set of edges meeting each node exactly once.) It is currently an open problem to find the minimal m such that $F_n(m)$ has a perfect matching with probability $1 - o(1)$.

3. Proof of Theorem 5

Proof of (a). Let $N(n)$ be the number of nodes of $G_n(1)$ with degree 1. Standard calculations give that

$$E(N(n)) \sim ne^{-1},$$
$$E\left(N(n)^2\right) \sim n^2 e^{-2},$$

and hence, by Chebyshev's inequality,

$$P(N(n) \geq 1) \geq 2 - \frac{E\left(N(n)^2\right)}{E(N(n))^2} \to 1 \quad \text{as} \quad n \to \infty.$$

If we delete the neighbour of a degree 1 node, then clearly no 2-matching can exist, so graphs containing nodes of degree 1 cannot be bicritical.

Proof of (b). Let $N(n)$ be the number of nodes of $G_n(2)$ with degree 1. Then

$$E(N(n)) = n\frac{1}{n-1}\left(1 - \frac{1}{n-1}\right)^{2(n-1)} \to e^{-2} \quad \text{as} \quad n \to \infty.$$

Standard arguments may now be used to show that $N(n)$ is asymptotically distributed with a Poisson distribution in that, for $k = 0, 1, 2, \ldots$,

$$P(N(n) = k) \to \frac{1}{k!}\lambda^k e^{-\lambda} \quad \text{as} \quad n \to \infty,$$

where $\lambda = e^{-2}$. We do not prove this in detail, but refer the reader to exactly similar proofs in Erdös and Rényi [5, p. 27], Schürger [17, p. 50], Bollobás [2, p. 201] and Bollobás [3, p. 92]. Using the argument at the end of the proof of part (a), we find that

$$P(G_n(2) \text{ is bicritical}) \leq P(N(n) = 0) \to \exp(-e^{-2}) \quad \text{as} \quad n \to \infty.$$

A similar argument is easily seen to work for $F_n(2)$, by considering the number of pairs $A = \{i, j\}$ of nodes of $F_a(2)$ with the property that i and j are not adjacent in $F_n(2)$ but $|N(A)| \leq 2$.

Proof of (c). Suppose $m \geq 3$. By Theorem 3, if $G_n(m)$ is not bicritical then there exist sets S, T of nodes such that

$$S \cap T = \emptyset, \qquad 1 \leq |S| = |T|, \tag{3.1}$$

and such that no two nodes in T are adjacent and $N(T) \subseteq S$. Thus, the probability

$$\Pi_n(m) = P(G_n(m) \text{ is not bicritical})$$

satisfies

$$\Pi_n(m) \leq \sum_{S,T} P(E(S, T)),$$

where the summation is over all pairs S, T satisfying (3.1) and $E(S, T)$ is the event that, in the original construction, all edges emanating from T go into S and no edge emanating from $V_n \setminus (S \cup T)$ goes into T. Thus

$$\Pi_n(m) \leq \sum_{k=1}^{\frac{1}{2}n} \binom{n}{k}\binom{n-k}{k}\left(\frac{k}{n-1}\right)^{mk} \times \left(1 - \frac{k}{n-1}\right)^{m(n-2k)}. \tag{3.2}$$

Let the sequence $\{\alpha(n)\}$ of integers be given by

$$\alpha(n) = [\log n], \tag{3.3}$$

and divide the summation in (3.2) into three parts, dealing, respectively, with the ranges $1 \leq k \leq \alpha(n)$, $\alpha(n) < k \leq \frac{1}{2}n - \alpha(n)$, $\frac{1}{2}n - \alpha(n) < k \leq \frac{1}{2}n$. The first of these satisfies

$$\sum_{1}^{\alpha(n)} \binom{n}{k}\binom{n-k}{k}\left(\frac{k}{n-1}\right)^{mk}\left(1-\frac{k}{n-1}\right)^{m(n-2k)}$$

$$\le \sum_{1}^{\alpha(n)} \frac{n^{2k}}{(k!)^2}\left(\frac{\alpha(n)}{n-1}\right)^{mk}$$

$$\le \exp\left(n^2\left(\frac{\alpha(n)}{n-1}\right)^m\right)-1$$

$$\to 0 \quad \text{as} \quad n\to\infty \tag{3.4}$$

by (3.3). To deal with the second summation we use the following consequence of Stirling's formula: if $\{\gamma(n)\}$ is a sequence satisfying $\gamma(n)\to\infty$ as $n\to\infty$, then for all large n and all β such that βn is integer and $\gamma(n)\le\beta n\le n-\gamma(n)$ it is the case that

$$\binom{n}{\beta n}\le \frac{1}{\sqrt{(2n\beta(1-\beta))}}\left(\frac{1}{\beta^\beta(1-\beta)^{1-\beta}}\right)^n.$$

Make the change of variable $k=\beta n$ and approximate by an integral to find that, for all large n,

$$\sum_{\alpha(n)}^{\frac{1}{2}n-\alpha(n)} \binom{n}{k}\binom{n-k}{k}\left(\frac{k}{n-1}\right)^{mk}$$

$$\times\left(1-\frac{k}{n-1}\right)^{m(n-2k)}$$

$$\le \int_{\alpha(n)/n}^{(1-2\alpha(n)/n)/2} \frac{1}{\beta\sqrt{(1-2\beta)}}(g(\beta))^n \mathrm{d}\beta, \tag{3.5}$$

where

$$g(\beta)=\frac{\beta^{(m-2)\beta}(1-\beta)^{m(1-2\beta)}}{(1-2\beta)^{1-2\beta}} \quad \text{for } 0<\beta<\frac{1}{2},$$

$$\le \beta^\beta\left(\frac{(1-\beta)^3}{1-2\beta}\right)^{1/2\beta} \quad \text{since} \quad m\ge 3.$$

Note that

$$\frac{\beta^\beta}{(1-2\beta)^{1-2\beta}}<1 \quad \text{if} \quad \frac{1}{3}<\beta<\frac{1}{2}$$

and

$$\frac{(1-\beta)^3}{1-2\beta}<1 \quad \text{if} \quad 0<\beta<\frac{1}{2}(3-\sqrt{5}),$$

giving that for any γ satisfying $0<\gamma<\frac{1}{3}$ there exists $\delta=\delta(\gamma)$ such that $0<\delta<1$ and

$$g(\beta)\le\delta \quad \text{for} \quad \gamma\le\beta<\frac{1}{2}.$$

Furthermore, for values of β which are small and positive it is the case that

$$g(\beta)\le 1-\beta+o(\beta);$$

we choose $\gamma<\frac{1}{3}$ such that

$$g(\beta)\le 1-\tfrac{1}{2}\beta \quad \text{for} \quad 0<\beta\le\gamma.$$

For all large n, the integral in (3.5) is at most

$$A\int_{\alpha(n)/n}^{\gamma} \frac{n}{\alpha(n)}\left(1-\frac{1}{2}\beta\right)^n \mathrm{d}\beta$$

$$+B\int_{\gamma}^{1/2}\frac{\delta^n}{\sqrt{(1-2\beta)}}\mathrm{d}\beta$$

$$\le \frac{2A}{\alpha(n)}\left(1-\frac{\alpha(n)}{2n}\right)^{n+1}+B\delta^n$$

$$\to 0 \quad \text{as} \quad n\to\infty, \text{ by (3.3)},$$

where A and B are constants. Thus the summation in (3.5) tends to 0 as $n\to\infty$. Finally,

$$\sum_{\frac{1}{2}n-\alpha(n)+1}^{\frac{1}{2}n} \binom{n}{k}\binom{n-k}{k}\left(\frac{k}{n-1}\right)^{mk}$$

$$\times\left(1-\frac{k}{n\text{-}1}\right)^{m(n-2k)}$$

$$\le \alpha(n)2^n n^{2\alpha(n)}\left(\frac{1}{2}\right)^{m(n-2\alpha(n))/2}$$

$$\to 0 \quad \text{as} \quad n\to\infty$$

by (3.3) and the fact that $m\ge 3$. We have shown that the summation in (3.2) tends to 0 as $n\to\infty$, thus completing the proof of the theorem.

Acknowledgement

We thank Alan Frieze for his interest in this work and for finding an error in an earlier 'proof' of the conjecture concerning $G_n(2)$.

Note added in proof

Since the writing of this paper, two of the problems mentioned herein have been solved. Frieze [9] has proved that $P(F_{2n}(2)$ has a perfect matching$)\to 1$ as $n\to\infty$. One of the current authors (Grimmett [11]) has verified the conjecture mentioned in the abstract.

References

[1] C. Berge, "Regularizable graphs", in: *Proceedings of the Calcutta Conference on Graph Theory*, Indian Statistical Institute, Calcutta, 1976.

[2] B. Bollobás, "Threshold functions for small subgraphs", *Math. Proc. Cambridge Phil. Soc.* **90**, 197–206 (1981).

[3] B. Bollobás, "Random graphs", in: H.N.V. Temperley, ed., *Combinatorics*, London Math. Soc. Lect. Notes 52, 1981, 80–102.

[4] B. Bollobás, "Almost all regular graphs are Hamiltonian", *European Journal of Combinatorics* **4**, 97–106 (1983).

[5] P. Erdős and A. Rényi, "On the evolution of random graphs", *Publ. Math. Inst. Hung. Acad. Sci.* **5A**, 17–61 (1960).

[6] T.I. Fenner and A.M. Frieze, "On the connectivity of random *m*-orientable graphs and digraphs", *Combinatorica* **2**, 347–360 (1982).

[7] T.I. Fenner and A.M. Frieze, "On the existence of hamiltonian cycles in a class of random graphs", *Discrete Mathematics* **45**, 1–5 (1983).

[8] T.I. Fenner and A.M. Frieze, "Hamiltonian cycles in random regular graphs", *Journal of Combinatorial Theory* **B37**, 103–112 (1984).

[9] A.M. Frieze, "Maximum matchings in a class of random graphs", *Journal of Combinatorial Theory* **B**, forthcoming.

[10] G.R. Grimmett, "Random graphs", in: L. Beineke and R. Wilson, eds., *Selected Topics in Graph Theory 2*, Academic Press, London, 1983, 201–236.

[11] G.R. Grimmett, "An exact threshold theorem for random graphs and the node packing problem", *Journal of Combinatorial Theory* **B**, forthcoming.

[12] G.L. Nemhauser and L.E. Trotter, "Vertex packings: Structural properties and algorithms", *Mathematical Programming* **8**, 232–248 (1975).

[13] J.C. Picard and M. Quéyranne, "Vertex packings (VLP) – Reductions through alternate labelling", Tech. Rept. EP85-R-48, Ecole Polytechnique, University of Montreal, Montreal, 1975.

[14] J.C. Picard and M. Quéyranne, "On the integer valued variables in the linear vertex packing problem", *Mathematical Programming* **12**, 97–101 (1977).

[15] W.R. Pulleyblank, "Minimum node covers and 2-bicritical graphs", *Mathematical Programming* **17**, 91–103 (1979).

[16] S.M. Ross, "A random graph", *Journal of Applied Probability* **18**, 309–315 (1981).

[17] K. Schürger, "Limit theorems for complete subgraphs of random graphs", *Per. Math. Hungarica* **10**, 47–53 (1979).

[18] E. Shamir and E. Upfal, "One-factor in random graphs based on vertex choice", *Discrete Mathematics* **41**, 281–286 (1982).