Optimal and Random Partitions of Random Graphs

Lenwood S. Heath and Joseph W. Lavinus

TR 93-24

Department of Computer Science
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

July 30, 1993

Submitted to COMBINATORICS, PROBABILITY, AND COMPUTING, 1993.

Optimal and Random Partitions of Random Graphs*

Lenwood S. Heath
Department of Computer Science
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24060

Joseph W. Lavinus
Department of Computer Science
University of Virginia
Charlottesville, Virginia 22903

Abstract

The behavior of random graphs with respect to graph partitioning is considered. Conditions are identified under which random graphs cannot be partitioned well, i.e., a random partition is likely to be almost as good as an optimal partition.

Graph algorithms are often tested using random graphs as input (see, e.g., Heath and Lavinus [5]). In the simplest and most common model [1], $R_{n,p}$ denotes the class of random graphs with n labeled vertices in which each edge is present with probability p. A natural question is whether such random graphs are useful for testing algorithms for various graph optimization problems. Turner [7] considers the problem of minimizing the *bandwidth* of a linear ordering of a graph, which is the length of the longest edge under that ordering. He shows ([7], Theorem 2.2) that the bandwidth of a random graph from $R_{n,p}$ almost certainly exceeds $(1-\varepsilon)n$ for any constant $\varepsilon > 0$;

*The authors gratefully acknowledge the support of National Science Foundation grant IRI–9116991.

that is, with respect to bandwidth minimization, a *random ordering* of a large random graph is almost as good as an *optimal ordering.* Thus, random graphs are not useful in testing bandwidth-minimization algorithms. McDiarmid and Miller [6] show analogous results for an extension of bandwidth to multi-dimensional lattices.

In this paper, we prove that a similar phenomenon occurs with respect to partitioning a random graph. The problem of partitioning a graph consists of dividing the vertices of the graph into subsets of cardinality not exceeding some bound K such that the number of edges whose endpoints lie in different subsets is minimized. More precisely, the problem of *graph partitioning* is the following:

> *Given an undirected graph $G = (V, E)$ and a set size K, find a partition of V into disjoint subsets $V_1, V_2, \ldots, V_m$ such that $|\{(u,v) : (u,v) \in E, u \in V_i, v \in V_j, i \neq j\}|$ is minimized, subject to $|V_i| \leq K$ for all i, $1 \leq i \leq m$.*

The corresponding decision problem is NP-complete (see Garey and Johnson [3], page 209). In this paper, we prove that a random graph cannot be partitioned well in the sense that a random partition is almost as good as an optimal partition.

For simplicity, assume that the number of vertices in the graph is an integer multiple of K. Suppose $G = (V, E)$ is a random graph from $R_{n,p}$. Let $\Pi = \{V_1, V_2, \ldots, V_{n/K}\}$ be a random partition of V in which every subset V_i contains exactly K vertices. An edge whose endpoints lie in different subsets is *external.* Let $\phi(\Pi)$ be the number of external edges in the partition Π. Let $\phi(G)$ be the number of external edges in an optimal partition of G.

The number of non-external edges (both of whose endpoints lie in the same subset) in Π falls between 0 and $\frac{n}{K}\binom{K}{2}$. The number of external edges falls between 0 and

$$\begin{aligned} N &= \binom{n}{2} - \frac{n}{K}\binom{K}{2} \\ &= n(n-K)/2. \end{aligned}$$

The expected number of external edges in a random partition Π is

$$\mathrm{E}[\phi(\Pi)] = pN.$$

We proceed to derive conditions under which a random partition is almost as good as an optimal partition of G (Corollaries 3 and 4). The following lemma bounds the probability that $\phi(\Pi)$ is much below its expected value.

Lemma 1 *Let G be a random graph from $R_{n,p}$. Constrain partitions of G to n/K subsets. Let $\zeta = \zeta(n)$ be a real-valued function of n. Then*

$$\ln \Pr[\phi(G) \le (1-\zeta)pN] \;\le\; n \ln n - n \ln K - \frac{\zeta^2 pN}{4(1-p)} + \frac{\ln n}{2} + O\left(\frac{n \ln K}{K}\right).$$

Proof: Let Π be a random partition of G into n/K subsets. The number of external edges $\phi(\Pi)$ follows a binomial distribution with parameters N and p; that is,

$$\Pr[\phi(\Pi) = k] = \binom{N}{k} p^k (1-p)^{N-k}.$$

The probability of interest is a tail of this distribution, to wit,

$$\Pr[\phi(\Pi) \leq (1-\zeta)pN] = \sum_{i=0}^{(1-\zeta)pN} \binom{N}{i} p^i (1-p)^{N-i}.$$

An upper bound on this probability can be obtained from Hoeffding's inequality (see [2], page 126), which asserts

$$\Pr[\mathrm{E}[\phi(\Pi)] - \phi(\Pi) \geq r] \leq e^{-r^2/4Np(1-p)}$$

for arbitrary r. Substituting $\mathrm{E}[\phi(\Pi)] = pN$ and selecting $r = \zeta pN$, we have

$$\Pr[\phi(\Pi) \leq (1-\zeta)pN] \leq e^{-\zeta^2 pN/4(1-p)}.$$

There are $n!/(K!)^{n/K}$ such partitions Π. For the optimal layout of G to have at most $(1-\zeta)pN$ external edges, at least one of these $n!/(K!)^{n/K}$ partitions must have $(1-\zeta)pN$ or fewer external edges. Hence,

$$\Pr[\phi(G) \leq (1-\zeta)pN] \leq \frac{n! e^{-\zeta^2 pN/4(1-p)}}{(K!)^{n/K}}.$$

Taking natural logarithms, we obtain

$$\ln \Pr[\phi(G) \le (1-\zeta)pN] \le \ln n! - \frac{n}{K} \ln K! - \frac{\zeta^2 pN}{4(1-p)}.$$

Applying Stirling's approximation for the factorial function (see [4], page 467), we have

$$\begin{aligned}\ln \Pr[\phi(G) \le (1-\zeta)pN] &\le n \ln n - n + \frac{\ln n}{2} - \frac{n}{K}\left(K \ln K - K + \frac{\ln K}{2} + O(1)\right) - \\ &\quad \frac{\zeta^2 pN}{4(1-p)} + O(1). \\ &= n \ln n - n \ln K - \frac{\zeta^2 pN}{4(1-p)} + \frac{\ln n}{2} + O\left(\frac{n \ln K}{K}\right),\end{aligned}$$

as required. □

Our central result is Theorem 2, which identifies conditions under which $\Pr[\phi(G) \le (1-\zeta)pN]$ is asymptotically 0.

Theorem 2 *Let G be a random graph from $R_{n,p}$. Constrain partitions of G to n/K subsets. Let $\zeta = \zeta(n)$ be a real-valued function of n. Suppose $K = o(n)$, and*

$$1 > \gamma \ge p \ge \frac{\tau \ln n}{n},$$

where γ and τ are constants such that, for sufficiently large n,

$$\frac{\zeta^2\tau}{8(1-\gamma)} > 1.$$

Then

$$\Pr[\phi(G) \leq (1-\zeta)pN] \to 0$$

as $n \to \infty$.

Proof: By Lemma 1,

$$\ln\Pr[\phi(G) \leq (1-\zeta)pN] \leq n\ln n - n\ln K - \frac{\zeta^2 pN}{4(1-p)} + \frac{\ln n}{2} + O\left(\frac{n\ln K}{K}\right).$$

Substituting the definition of N and the bounds on p, we obtain

$$\begin{aligned}
\ln\Pr[\phi(G) \leq (1-\zeta)pN] &\leq n\ln n - n\ln K - \frac{\zeta^2\frac{1}{2}n(n-K)\tau n\ln n/n}{4(1-\gamma)} + \frac{\ln n}{2} + \\
&\quad O\left(\frac{n\ln K}{K}\right) \\
&= n\ln n - n\ln K - \frac{\zeta^2\tau(n-K)\ln n}{8(1-\gamma)} + \frac{\ln n}{2} + O\left(\frac{n\ln K}{K}\right) \\
&= \left(1 - \frac{\zeta^2\tau}{8(1-\gamma)}\right) n\ln n + \frac{\zeta^2\tau}{8(1-\gamma)} K\ln n - n\ln K + \frac{\ln n}{2} + \\
&\quad O\left(\frac{n\ln K}{K}\right) \\
&= \left(1 - \frac{\zeta^2\tau}{8(1-\gamma)}\right) n\ln n + o(n\ln n)
\end{aligned}$$

which approaches $-\infty$ as $n \to \infty$. Hence

$$\Pr[\phi(G) \leq (1-\zeta)pN] \to 0$$

as $n \to \infty$, as desired. $\square$

As a corollary, we obtain this special case.

Corollary 3 *Let G be a random graph from $R_{n,p}$. Constrain partitions of G to n/K subsets, where $K = o(n)$. Suppose that ε is a positive constant and that γ is a constant satisfying $1 > \gamma \geq p$. If the expected degree of G is $\omega(\ln n)$, then*

$$\Pr[\phi(G) \leq (1-\varepsilon)pN] \to 0$$

as $n \to \infty$.

Proof: The expected degree of G is $p(n-1)$. If the expected degree is $\omega(\ln n)$, then $p = \omega(\ln n/n)$. Since $K = o(n)$, $N = \Theta(n^2)$. Hence we have $pN = \Omega(n \ln n)$. Choose τ sufficiently large that

$$\frac{\varepsilon^2 \tau}{8(1-\gamma)} > 1.$$

Choosing $\zeta = \varepsilon$ in Theorem 2, we conclude

$$\Pr[\phi(G) \leq (1-\varepsilon)pN] \to 0$$

as $n \to \infty$, as desired. □

In other words, if the degree of a random graph is slightly greater than logarithmic, then the number of external edges in an optimal partition is, asymptotically, almost the same as the number of external edges in a random partition.

If we want the degree to be *exactly* logarithmic, we obtain a slightly weaker corollary.

Corollary 4 *Let G be a random graph from $R_{n,p}$ where $p = \tau \ln n/n$. Constrain partitions of G to n/K subsets, where $K = o(n)$. Suppose that ε is a positive constant. If $\tau > 8/\varepsilon^2$, then*

$$\Pr[\phi(G) \leq (1-\varepsilon)pN] \to 0$$

as $n \to \infty$.

In particular, if $\tau > 32$, then the probability that an optimal partition of G has fewer than half the expected number of external edges in a random partition is asymptotically 0.

In the face of Corollaries 3 and 4, we conclude that random graphs, in the standard model we study here, are not useful for testing heuristics for the graph partitioning problem. Experimental evidence for this conclusion is presented in Heath and Lavinus [5]. In that work, some alternate random graph models based on geometry *do* prove useful in testing heuristics.

References

[1] B. Bollobás, *Random Graphs*, Academic Press, Orlando, Florida, 1985.

[2] T. H. Cormen, C. E. Leiserson, and R. L. Rivest, *Introduction to Algorithms*, MIT Press, Cambridge, Massachusetts, 1990.

[3] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W.H. Freeman and Company, San Francisco, 1979.

[4] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*, Addison-Wesley, Reading, Massachusetts, 1989.

[5] L. S. Heath and J. W. Lavinus, *Heuristics for laying out information graphs*. Submitted, 1993.

[6] C. McDiarmid and Z. Miller, *Lattice bandwidth of random graphs*, Discrete Applied Mathematics, 30 (1991), pp. 221–227.

[7] J. S. Turner, *On the probable performance of heuristics for bandwidth minimization*, SIAM Journal on Computing, 15 (1986), pp. 561–580.