Large Subgraphs of Minimal Density or Degree

Paul Erdős
Hungarian Academy of Sciences

Ralph Faudree
University of Memphis

Arun Jagota
University of Memphis

Tomasz Luczak
Adam Mickiewicz University

Abstract

This paper addresses the following questions. In any graph G with at least $\alpha\binom{n}{2}$ edges, how large of an induced subgraph H can we guarantee the existence of with minimum degree $\delta(H) \geq \lfloor \alpha |V(H)| \rfloor$? In any graph G with at least $\alpha\binom{n}{2} - f(n)$ edges, where $f(n)$ is an increasing function of n, how large of an induced subgraph H can we guarantee the existence of containing at least $\alpha\binom{|V(H)|}{2}$ edges? In any graph G with at least αn^2 edges, how large of an induced subgraph H can we guarantee the existence of with at least $\alpha|V(H)|^2 + \Omega(n)$ edges? For $\alpha = 1 - 1/r$ for $r = 2, 3, \ldots$, the answer is zero since if G is a complete r-partite graph, no subgraph H of G has more than $\alpha|V(H)|^2$ edges. However, we show that for all admissible α except these, the answer is $\Omega(n)$. In any graph G with minimum degree $\delta(G) \geq \alpha n - f(n)$, where $f(n) = o(n)$, how large of an induced subgraph H can we guarantee the existence of with minimum degree $\delta(H) \geq \alpha|V(H)|$?

1 Introduction

Let G be an arbitrary graph without loops or multiple edges. As usual, $V(G)$, $E(G)$, $\delta(G)$, and n denote the vertex set, the edge set, the minimum degree of the vertices in G, and the number of vertices in G. For $S \subset V(G)$, let $G[S]$ denote the subgraph of G induced by S. In this paper the expression "subgraph of G" will always refer to one induced by a set of vertices. If H is a subgraph of G then $|H|$ will denote its order (the number of vertices).

For positive integer n and functions $f(n)$ and $g(n)$ we say that $f = O(g)$ if, for some positive constants c and n_0, $f(n) \leq cg(n)$ for all $n \geq n_0$. We

JCMCC 22 (1996), pp. 87-96

say that $f = o(g)$ if $\lim_{n\to\infty} f(n)/g(n) = 0$. We say that $f = \Omega(g)$ if $g = O(f)$ and $f = \Theta(g)$ if $f = O(g)$ and $g = O(f)$.

This paper addresses the following questions. How large a subgraph of a graph G can we find so that its density is sufficiently higher than that of G? How large a subgraph of a graph G can we find so that its minimum degree is sufficiently large?

We address the second question first. The following theorem is obtained by an elementary generalization of the lower bound results and proof of Erdős, Luczak, and Spencer [2] (Theorem 5.2). We think that this general version provides some additional useful information, therefore we include it here. We include the proof in full, to make the paper self-contained and clear.

Theorem 1 *Let $\alpha(n)$ and $\beta(n)$ be functions of n such that $0 < \alpha(n) < 1$ and $0 < \beta(n) \le \beta_0 < 1$ for some constant β_0. Furthermore, let H be a subgraph of maximum order such that $\delta(H) \ge \lfloor \beta(n)|H| \rfloor$ in a graph G containing at least m edges, and let $f(n)$ be a nonnegative nondecreasing function of n.*

1. *If $m = \alpha(n)\binom{n}{2} + f(n)$ then*
   (a) *for $\beta(n) = \alpha(n)$ we have $|H| = \Omega(\sqrt{f(n)}) + \Omega(\sqrt{n})$,*
   (b) *for $\beta(n) = \alpha(n) - 1/g(n)$, where $g(n)$ is a positive constant or increases with n,*
   $$|H| = \Omega(n/\sqrt{g(n)}) + \Omega(\sqrt{f(n)}) + \Omega(\sqrt{n}) .$$
2. *If $m = \alpha(n)\binom{n}{2} - f(n)$ then*
   (a) *for $\beta(n) = \alpha(n)$ and $f(n) = o(n)$ we have $|H| = \Omega(\sqrt{n})$,*
   (b) *for $\beta(n) = \alpha(n) - 1/g(n)$, where $g(n)$ is a positive constant or increases with n and $f(n) = o(n^2/g(n)) + o(n)$, we have $|H| = \Omega(n/\sqrt{g(n)}) + \Omega(\sqrt{n})$.*

The case when $\alpha = \beta = 1/2$ was studied in detail in [2], where, in particular, the results identical with (1a) and (2a) were shown. In fact the version of (2a) for $\alpha = \beta = 1/2$ in [2] is more general as it does not require that $f(n) = o(n)$. However, we may also remove this restriction by first using Theorem 2, which appears later, to find a "large" subgraph H of G containing at least $\alpha\binom{|H|}{2}$ edges, and then using case (2a) of Theorem 1 to find a "large" subgraph I of H with $\delta(I) \ge \lfloor \beta|I| \rfloor$. More importantly, in [2], upper bound results for the above questions are also obtained. For one important case, $f(n) = 0$ in (1a), it is shown that $|H| = O(n^{2/3})$. Whereas this is much better than $O(n)$, a significant gap still remains between the upper and lower bound, which deserves exploration.

To picture how the minimum size of H, given by Theorem 1, depends on the difference $\alpha(n) - \beta(n)$ let us look more closely at the case when $f(n) = 0$. If $\beta(n)$ is bounded from above by a constant smaller than $\alpha(n)$ then H is of the order of n. When $\beta(n)$ approaches $\alpha(n)$ as n increases, the minimum size of H drops down (and depends on the rate of convergence $\beta(n) \rightarrow \alpha(n)$), and, finally, for $\alpha(n) = \beta(n)$ the lower bound on $|H|$ is of the order of $\sqrt{n}$. For constant α, β and $\alpha < \beta$ our proof breaks down as it relies on the greedy method (see the proof). In this case, by the probabilistic method it may be shown that $|H| = O(\log n)$ (see [2], Theorem 5.1* case (i) and set $\epsilon = -0.01, \beta = 1/2$ for example).

Now consider $\alpha(n) = \beta(n)$ in Theorem 1. As noted earlier in [2], with $f(n) = O(n)$ the lower bound on $|H|$ is of the order of $\sqrt{n}$, with $f(n) = \Omega(n^2)$ it improves to order of n, and with $n = o(f(n))$ and $f(n) = o(n^2)$ it is in between.

Theorem 1 gives also some estimates for the following extremal problem: for constant k estimate the value of $F(n,k)$ defined as the smallest integer e so that every graph with n vertices and e edges has a subgraph H of order $m = \Omega(\sqrt{n})$ with $\delta(H) \geq k$. Setting $f(n) = 0$, $\alpha(n) = \beta(n) = k/m$ in Theorem 1 gives $F(n,k) \leq \alpha(n)\binom{n}{2} = O(n^{3/2})$.

A k-multigraph is one in which a pair of vertices is connected by at most k edges. With $0 < \alpha, \beta(n) < k$, Theorem 1 also holds for a k-multigraph in which k is a constant. If $k = k(n)$ is an increasing function of n then the analog of Theorem 1 holds for a k-multigraph in which every result is divided by $k(n)$. See the proof.

Theorem 1 represents "extremal" cases. We may also ask the question for "typical" cases. Specifically, we may ask, In a random graph in $G_{n,p}$, what is the expected order of the largest subgraph H with $\delta(H) \geq \lfloor \beta |H| \rfloor$. This question is addressed in [2] for $\beta = 1/2$. It is shown there, for example, that for $p = 1/2$ this expected order is $\Theta(n)$.

Another question that we may ask is of the Ramsey type. Namely, instead of asking whether G has a subgraph of a certain order satisfying certain conditions, we ask if either G or its complement $\bar{G}$ has a subgraph of a certain order satisfying those conditions. This question has been investigated in [3], in which the following results are obtained. For any graph G, let H be a maximum order subgraph of G or its complement $\bar{G}$ satisfying $\delta(H) \geq \alpha|H|$. If $0 < \alpha < 1/2$, then $|H| = \Omega(n)$. If $1/2 < \alpha < 1$, then $|H| = O(\log n)$. Thus the behavior changes suddenly at $\alpha = 1/2$. Further, at $\alpha = 1/2$, it is shown that $|H| = \Omega(n/\log n)$ and $|H| = O(n \log\log n/\log n)$.

We now address the first question. Let $\alpha(n) = a(n)/b(n)$ denote a rational function of n such that $0 < \alpha(n) \leq 1/2$. Let $f(n)$ be a nonnegative nondecreasing function of n.

Theorem 2 *In a graph G containing at least $\alpha(n)\binom{n}{2} - f(n)$ edges, let H be a subgraph of maximum order containing at least $\alpha(n)\binom{|H|}{2}$ edges. Then*

$$|H| = \begin{cases} \Omega(\frac{n}{(b(n)f(n)/n)^{\log(f(n)/(\alpha(n)n))}}) & \textit{if } f(n) = \Omega(\alpha(n)n) \\ \Omega(\frac{n}{(b(n)f(n)/n)} & \textit{if } f(n) = o(\alpha(n)n) + \Omega(n/b(n)) \\ \Theta(n) & \textit{if } f(n) = o(n/b(n)) \end{cases}$$

With $f(n) = \Theta(n)$ and $\alpha(n)$ a constant, Theorem 2 gives $|H| = \Theta(n)$.

It may be verified from the proof that Theorem 2 also holds, for $0 < \alpha(n) \leq 1/2$, if G is a multigraph. In contrast to Theorem 1, no restriction is placed on the number of edges connecting a pair of vertices.

The next question we ask is as follows. If $\delta(G) \geq \lfloor n/2 \rfloor - k$, where k is a constant, can we find a large subgraph H of G with $\delta(H) \geq \lfloor |H|/2 \rfloor$? Using Theorems 1 and 2, we are able to show that $|H| = \Omega(\sqrt{n})$. More generally, the following result holds.

Proposition 1 *Let α be a constant with $0 < \alpha \leq 1/2$ and let $f(n)$ be a nonnegative nondecreasing function of n. Then each graph G with $\delta(G) \geq \alpha n - f(n)$ contains a subgraph H with $\delta(H) \geq \alpha|H|$ such that*

$$|H| = \Omega(\sqrt{\frac{n}{(f(n)/n)^{\log(f(n)/n)}}})$$

We now ask the following question related to Theorem 2. The problem is whether, roughly, in the hypothesis of Theorem 2 one can replace a "deficit" term $f(n)$ for the number of edges of G by a "surplus" term $g(|H|, n)$ for the number of edges of H, i.e. whether every graph G with at least αn^2 edges contains a large subgraph H with at least $\alpha|H|^2 + g(|H|, n)$ edges. One can easily see that if G is a complete r-partite graph the answer is negative since no subgraph H of G has more than $\alpha|H|^2$ edges (note that in this case $\alpha = 1 - 1/r$ for $r = 2, 3, \ldots$). However, it turns out that these graphs are the only exceptions.

Theorem 3 *Let $\alpha > 0$ be a constant such that $\alpha \neq 1 - 1/r$ for $r = 2, 3, \ldots$. Then there exist constants $\delta = \delta(\alpha) > 0$, $\eta = \eta(\alpha) > 0$ and $N = N(\alpha)$ such that for every $n \geq N$ and $k = k(n)$, where $|k| \leq \delta n \log n$, every graph G with n vertices and more than $\alpha n^2 - k$ edges contains a subgraph on $n' \geq n - \eta k/\log n$ vertices with at least $\alpha(n')^2 + kn'/n$ edges.*

In particular, for every α as above, every graph G with n vertices and more than $\alpha n^2 - O(n)$ edges contains a subgraph on $n' \geq n - O(n/\log n)$ vertices with at least $\alpha(n')^2 + \Omega(n')$ edges.

One is tempted to conjecture that the above result is sharp in a very strong sense, i.e. that for every constant α there exists a constant $B = B(\alpha)$ and a graph G with αn^2 edges such that G contains no subgraphs on m vertices, $1 \leq m \leq n$, with more than $\alpha m^2 + Bm \log n$ edges. Although the existence of such a graph would be rather surprising at the present moment we are not able to exclude such a possibility.

2 Proof of Theorem 1

The proof employs the greedy method, as in [2], and works for a k-multigraph ($k = 1$ gives a graph).

```
while δ(G) < ⌊β|V(G)|⌋ do
      Delete a minimum degree vertex from G
```

First assume that G has at least $\alpha(n)\binom{n}{2}$ edges. Let $|V(G)| = n$ at the start of a non-terminating step of the algorithm. Then the deleted vertex has degree $\leq \beta(n)n - 1$. After termination, let $|V(G)| = t$. We want largest t such that

$$\alpha(n)\binom{n}{2} - [(\beta(n)n-1)+(\beta(n)(n-1)-1)+\ldots+(\beta(n)(t+1)-1)] \geq k\binom{t}{2}$$

which may be simplified to

$$[\alpha(n) - \beta(n)]\binom{n}{2} + [1-\beta(n)]n \geq \frac{k-\beta(n)}{2}t^2 + (1 - \frac{k+\beta(n)}{2})t \quad (1)$$

Addition or subtraction of $f(n)$ from the left hand side of (1) and a straightforward examination of the cases listed in Theorem 1 completes the proof. □

3 Proof of Theorem 2

The proof employs an interesting variant of the greedy method.

```
for k ← f(n)/n + α(n)/2; k ≥ α(n)/2; k ← k/2 do
    while |E(G)| < α(n)/2 |V(G)|^2 - k/2 |V(G)| do
        Delete a minimum degree vertex from G
```

At the start of any iteration of the *for* loop,

$$|E(G)| \geq \frac{\alpha(n)}{2} |V(G)|^2 - k\,|V(G)|$$

This is also true when the algorithm begins, since

$$\frac{\alpha(n)}{2}\,n^2 \;-\; (f(n)/n \;+\; \alpha(n)/2)n \;=\; \alpha(n)\binom{n}{2} \;-\; f(n)$$

At the end of any one iteration of the *for* loop

$$|E(G)| \geq \frac{\alpha(n)}{2}\,|V(G)|^2 \;-\; \frac{k}{2}\,|V(G)|$$

Thus when the algorithm terminates,

$$|E(G)| \geq \frac{\alpha(n)}{2}\,|V(G)|^2 \;-\; \frac{\alpha(n)}{2}\,|V(G)| \;=\; \alpha(n)\binom{|V(G)|}{2}$$

Let $|V(G)| = m$ at the start of iteration $k = k_0$ of the *for* loop and let $|V(G)| = t$ at the end of that iteration. In the step when $|V(G)| = s$, the deleted vertex has degree smaller than $\alpha(n)s - k_0$. It turns out that we need to bound it better. Recall that $\alpha(n) = a(n)/b(n) \leq 1/2$. Let $\Delta_d(x, y)$ denote the sum $\sum_{i=y}^{x} \alpha(n)i$ minus the sum of the degrees of the deleted vertices in the sequence of steps in which $|V(G)|$ equals $x, x-1, \ldots, y$. Then for any positive integer q,

$$\Delta_d(qb(n) + b(n) - 1, qb(n)) \geq 1 \;+\; \sum_{j=1}^{b(n)-1} \frac{ja(n) \bmod b(n)}{b(n)}$$

Note that $a(n)/b(n) \leq 1/2$. Suppose that

$$ka(n) \bmod b(n) < a(n).$$

It is easy to check that

$$(k+1)a(n) \bmod b(n) \geq a(n).$$

Hence

$$\begin{aligned}\Delta_d(qb(n) + b(n) - 1, qb(n)) \;&\geq\; a(n)/2 \;+\; (b(n) - a(n))/b(n) \\ &=\; b(n)\alpha(n)/2 + (1 - \alpha(n)) \qquad (2)\end{aligned}$$

We want largest t such that

$$\alpha(n)m^2/2 - k_0m - \Big(\sum_{i=t+1}^{m} (\alpha(n)i - k_0) \;-\; \Delta_d(m, t+1)\Big) \geq \alpha(n)t^2/2 \;-\; k_0t/2$$

which may be simplified to

$$-\alpha(n)m/2 + \Delta_d(m,t+1) \geq -\alpha(n)t/2 + k_0 t/2 \qquad (3)$$

Using (2),

$$\Delta_d(m,t+1) \geq \left(\frac{b(n)\alpha(n)}{2} + 1 - \alpha(n)\right)\frac{m-t}{b(n)}$$

Substituting in (3) gives

$$\frac{(1-\alpha(n))m}{b(n)} \geq \left(\frac{1-\alpha(n)}{b(n)} + \frac{k_0}{2}\right)t$$

from which

$$t = \frac{(1-\alpha(n))/b(n)}{(1-\alpha(n))/b(n) \; + \; k_0/2}m \geq \frac{m/2}{1/2 \; + \; b(n)f(m)/m}$$

which is

$$\begin{cases} \Theta(\frac{m}{b(n)f(m)/m}) & \text{if } f(m) = \Omega(m/b(n)) \\ \Theta(m) & \text{if } f(m) = o(m/b(n)) \end{cases}$$

Since we start with $k = f(n)/n \; + \; \alpha(n)/2$ and stop when $k < \alpha(n)/2$, the number of iterations of the *for* loop is $i = \log_2(f(n)/n+\alpha(n)/2)/\alpha(n)/2$ which equals

$$\begin{cases} \Theta(\log f(n)/(\alpha(n)n)) & \text{if } f(n) = \Omega(n\alpha(n)) \\ 1 & \text{if } f(n) = o(n\alpha(n)) \end{cases}$$

Noting that $\alpha(n)n \; = \; \Omega(n/b(n))$, we finally get, after the algorithm terminates

$$|V(G)| = \begin{cases} \Omega(\frac{n}{(b(n)f(n)/n)^{\log(f(n)/(\alpha(n)n))}}) & \text{if } f(n) = \Omega(\alpha(n)n) \\ \Omega(\frac{n}{(b(n)f(n)/n)}) & \text{if } f(n) = o(\alpha(n)n) \; + \; \Omega(n/b(n)) \\ \Theta(n) & \text{if } f(n) = o(n/b(n)) \end{cases}$$

□

Finally, we may observe that the above proof works also if G is replaced by a multigraph. Indeed the governing inequality remains (3) and the $\Delta_d(m,t+1)$ calculation remains unchanged.

4 Proof of Proposition 1

First, it may be seen that Theorem 2 also holds if we start with G containing at least $\alpha(n)n^2/2 \; - \; g(n)$ edges, $0 < \alpha(n) \leq 1/2$, and we want a subgraph H of G containing at least $\alpha(n)|H|^2/2$ edges.

Consider G with $\delta(G) \geq \alpha n - f(n)$. Then $|E(G)| \geq \alpha n^2/2 - nf(n)/2$. By the above observations on Theorem 2, G contains a subgraph H with $|E(H)| \geq \alpha|H|^2/2$ and $|H| \geq \Omega(n/((f(n)/n)^{\log(f(n)/n)}))$. By Theorem 1, case (1 a), H contains a subgraph I with $\delta(I) \geq \lfloor \alpha|I| \rfloor$ and $|I| = \Omega(\sqrt{|H|})$. The result follows. □

5 Proof of Theorem 3

Our simple argument is based on the following result of Bollobás, Erdős and Simonovits (see, for example, [1, Thm.VI.3.1i]).

Theorem 4 *There exists an absolute constant $\eta > 0$ such that if $0 < \epsilon < 1/(r-1)$ and*

$$M > (1 - 1/(r-1) + \epsilon)n^2/2$$

then every graph with n vertices and M edges contains a complete r-partite graph $K_r(t)$, such that each set of the r-partition has

$$t = \left\lfloor \frac{\eta \log n}{r \log(1/\epsilon)} \right\rfloor$$

vertices.

We shall show the following result.

Lemma 1 *Let $\alpha = (1 - 1/(r-1) + \epsilon)/2$, where $0 < \epsilon < 1/(r(r-1))$ and $M = \alpha n^2 + cn \log n$, with*

$$|c| = |c(n)| \leq A = \frac{1}{50} \frac{\eta\epsilon}{\log(2/\epsilon)} \left(\frac{1}{r(r-1)} - \epsilon \right) .$$

(Here η is the constant which appears in the theorem above.) Then, there exists a constant N such that every graph G with $n > N$ vertices and M edges contains a set S of s vertices,

$$A \log n \leq s \leq \frac{50Ar(r-1)}{\epsilon(1 - r(r-1)\epsilon)} \log n ,$$

such that a graph obtaining from G by removing vertices from S has $n' = n - s$ vertices and at least $\alpha(n')^2 + cn' \log n + sA \log n$ edges.

Remark. We have made no effort to make constants in the above statement best possible.

Proof of Lemma 1.

Case 1. G contains at least $\epsilon n/4$ vertices of degree at least $d = 2\alpha n + 2c\log n + (20A/\epsilon)\log n$.

In this case, it is enough to remove from the graph $s = \lceil A\log n\rceil$ vertices of smallest degree. Indeed, let $d_1 \le \ldots \le d_n$ be a degree sequence of G, and let $d_{n-k+1} \ge d$ for some $k \ge \epsilon n/4$. Then

$$\begin{aligned}\textstyle\sum_{i=1}^{n-k} d_i &\le 2\alpha n^2 + cn\log n - k(2\alpha n + c\log n + (20A/\epsilon)\log n)\\ &\le (n-k)(2\alpha n + 2c\log n - \tfrac{k}{n-k}\tfrac{20A}{\epsilon}\log n)\\ &\le (n-k)(2\alpha n + 2c\log n - 5A\log n)\end{aligned}$$

Hence, s vertices of the smallest degree are incident to at most $s(2\alpha n + 2c\log n - 5A\log n)$ edges. Thus, deleting these vertices from G results in a graph with $n' = n - s$ vertices and at least

$$\begin{aligned}\alpha n^2 &\; -2\alpha sn + cn\log n - 2cs\log n + 5sA\log n\\ &= \alpha(n-s)^2 + c(n-s)\log n + s(5A\log n - c\log n - \alpha s)\\ &\ge \alpha(n')^2 + cn'\log n + sA\log n\end{aligned}$$

edges.

Case 2. G contains less than $\epsilon n/4$ vertices of degree at least $d_0 = 2\alpha n + 2c\log n + (20A/\epsilon)\log n$.

Note that the number of edges of G incident to vertices of degree at least d_0 is less than $\epsilon n^2/4$, so the graph $\bar{G}$, obtained from G by removing all such vertices, has $\bar{n} \ge 0.8n$ vertices and at least

$$\left(1 - \frac{1}{r-1} + \frac{\epsilon}{2}\right)\frac{n^2}{2} \ge \left(1 - \frac{1}{r-1} + \frac{\epsilon}{2}\right)\frac{\bar{n}^2}{2}$$

edges. Thus, according to Theorem 4, $\bar{G}$ contains a complete r-partite graph $K_r(t)$ with

$$t = \left\lfloor \frac{\eta\log\bar{n}}{r\log(2/\epsilon)}\right\rfloor \ge \frac{0.9\eta\log n}{r\log(2/\epsilon)}\,.$$

Now delete all vertices of this r-partite subgraph from G. The resulting graph has $n' = n - rt$ vertices and at least

$$\begin{aligned}\alpha n^2 &\; -2rt\alpha n + cn\log n - 2crt\log n - (20Art/\epsilon)\log n + \tbinom{r}{2}t^2\\ &= \alpha(n-rt)^2 + c(n-rt)\log n + rt(t((r-1)/2 - \alpha r) - (20A/\epsilon)\log n\\ &\quad -c\log n)\\ &\ge \alpha(n')^2 + cn'\log n + rt(t(1/(r-1) - \epsilon r)/2 - (21A/\epsilon)\log n)\\ &\ge \alpha(n')^2 + cn'\log n + rt(A/\epsilon)\log n\\ &\ge \alpha(n')^2 + cn'\log n + rtA\log n\,.\end{aligned}$$

This completes the proof of Lemma 1. Theorem 3 follows easily as a corollary. □

References

[1] B. Bollobás. *Extremal Graph Theory*. Academic Press, New York, 1978.

[2] P. Erdős, T.Luczak, and J. Spencer. Subgraphs of large minimal degree. In A. Frieze and T. Łuczak, editors, *Random Graphs*, volume 2, chapter 5, pages 59–66. John Wiley & Sons, 1992.

[3] P. Erdős and J. Pach. On a quasi-Ramsey problem. *Journal of Graph Theory*, 7:137–147, 1983.