On connectivity related extremal problems *

Peter Hajnal

Bolyai Institute, University of Szeged
Aradi Vértanúk tere 1., Szeged, Hungary 6720
hajnal@math.u-szeged.hu

Let $f_k(n)$ be the maximal number of edges of a simple graph on n vertices without k connected subgraph. W. Mader started to investigate the order of magnitude of this function. The first results on $f_k(n)$ are due to W. Mader who proved that $(3k-4)/2(n-(k-1)) \le f_k(n) \le (1+1/\sqrt{2})(k-1)(n-(k-1))$, assuming that n is large enough. He also conjectured that the lower bound is the right order of magnitude of $f_k(n)$. Further improvement is due to Matula, who proved that $f_k(n) \le 5/3(k-1)(n-(k-1))$. In this paper we improve Matula's upper bound by proving that $f_k(n) \le (1+\sqrt{6}/4)kn \approx 1.612kn$.

The improvement is not a major breakthrough but we think that the problem deserves more attention. We also want to popularize other related questions. We present applications of this results to Ramsey theory on connectivity and vertex partition of graphs with conditions on connectivity. These applications shed light on other connectivity related open problems.

0. Introduction

Extremal graph theory is a major research direction in graph theory with various applications (see combinatorial geometry for many excellent examples).

We want to shed light on few extremal questions related to graph connectivity. (For other problems in this direction see [5] and [6].) The most natural question (following Turán's theorem's lead) is: How many edges guarantee a k connected subgraph in a simple graph on n nodes? An equivalent formulation is: What is the maximal number of edges in a simple graph on n vertices with no k connected subgraph? First W. Mader exhibited an example. Let $k-1$ be a divisor of n. Our vertex set will be divided into $n/(k-1)$ many $k-1$ element sets. The induced subgraphs of these $k-1$ element sets will be cliques with one exception when the corresponding subgraph is an empty graph. The additional edges are all the edges connecting the independent $k-1$ set to the vertices of the cliques. Easy to check that the graph does not contain a k connected subgraph and it

* Partially supported by OTKA F021271

has $(3k-4)/2 \cdot (n-(k-1))$ edges. In terms of the function introduced in the abstract it means that $(3k-4)/2 \cdot (n-(k-1)) \leq f_k(n)$ at least for certain values of n. It is not so hard to construct graph with more edges than $(3k-4)/2 \cdot (n-(k-1))$ and without k connected subgraph. All the known examples are small in terms of the number of vertices. For the author there is no example known with more than $(3k-4)/2 \cdot (n-(k-1))$ edges, without k connected subgraph and with more than $(k-1)^2/2$ vertices. It is plausible to conjecture [4] that $(3k-4)/2(n-(k-1)) \leq f_k(n)$ for large enough n (with a lower bound condition on n, that is a function of k). The conjecture is verified in the case of $k \leq 7$ ([4]), but the general case is still open. To underline the difficulty we mention that various counterexamples exist for small values of n, with completely different structures. Even the large examples, showing the sharpness of the conjecture, are showing diversity.

Next we state the current best upper bound on $f_k(n)$ due to D. Matula.

Theorem. *(Matula [7]) Let G be a simple graph with $|V(G)| \geq 2(k-1)$ and $|E(G)| > \left(1+\sqrt{2}/2\right)(k-1)\left(|V(G)|-(k-1)\right)$ then G has a k connected subgraph.*

Our main contribution is to improve the upper bound.

Theorem. *Let $k \geq 3$. Let G be a simple graph assuming that*

$$n = |V(G)| \geq (k-1) + \frac{1}{2}\sqrt{6k^2-18k+16}$$

and

$$|E(G)| > \left(\frac{1}{4}\sqrt{6k^2-18k+16} + k - \frac{3}{2}\right)(n-(k-1)).$$

Then G has a k connected subgraph.

The order of magnitude of our bound is $(1+\sqrt{6}/4)kn \approx 1.612kn$.

Finally we mention few applications and related problems.

1. Notations

We use standard notation (for example see [3]). All graphs are supposed to be simple undirected graphs. $V(G)$ denotes the vertex set of the graph G and $E(G)$ denotes its edge set. $C \subset V(G)$ is a cutset of a connected graph G if after deleting the vertices in C the resulting graph $(G-C)$ is not connected. G is k connected iff it has more than k vertices and it has no cutset of size smaller than k.

If G has more than k vertices and it is not k connected, then it must have a cutset C of size $k-1$. In this case we think about G as a graph obtained by gluing together two graphs (G_1 and G_2) along $G|_C$ (the subgraph of G induced by C, consisting of the elements of C as vertices, and all the edges of G, connecting two elements of C as edges). Both G_1 and G_2 have at least k elements, and $V(G_1) \cap V(G_2) = C$.

If G has at least k vertices and it does not have a k connected subgraph then G itself is not k connected, so G can be thought as a graph built up from G_1 and G_2 by gluing them along a $k-1$ element set. In this case of course both G_1 and G_2 do not have a k connected subgraph. Hence both of them can be thought as a graph on k vertices or a graph obtained from two graphs by gluing them along a $k-1$ element set. To summarize the ideas above G can be built up from graphs with k vertices by a *gluing procedure*: in each step of the procedure we glue two already built up graph along a set of size $k-1$.

2. Proofs

First we prove a lemma which is only interesting for small graphs, but in that case the given bound is sharp. The lemma is present in [7] but we state it with proof for the sake of completeness.

Lemma. *Let G be a simple graph on $n(>k)$ vertices and*

$$|E(G)| > \frac{1}{6}\left(n^2 + (4k-7)n + (4k-2k^2)\right).$$

Then G has a k connected subgraph.

Proof. We prove the claim by induction on n.

If $n = k+1$, then the assumption on the number of edges gives us that $|E(G)| > \binom{k+1}{2} - 1$, hence G is a complete graph on $k+1$ vertices, itself a k connected graph.

Let us assume that we know the claim for graphs on fewer that $n = |V(G)|$ vertices. If G is k connected we are done. If not then G is obtained by gluing G_1 and G_2 along a $k-1$ element set, C. Let $n_1 = |V(G_1)|$ and $n_2 = |V(G_2)|$ $((n_1-(k-1)) + (n_2 - (k-1)) = n-(k-1)$, $k-1 < n_1, n_2 < n)$. W.l.o.g. we assume that $n_1 \geq n_2$, hence $n_1 \geq (n+(k-1))/2 > k$.

If $|E(G_1)| > 1/6 \cdot \left(n_1^2 + (4k-7)n_1 + (4k-2k^2)\right)$, then the induction hypothesis can be applied and we are done.

If $|E(G_1)| \leq 1/6 \cdot \left(n_1^2 + (4k-7)n_1 + (4k-2k^2)\right)$, then we can bound the number of edges in G by estimating the edges of G outside G_1 by the number of edges of the complete graph on $V(G_2) - C$ plus the number of edges of the complete bipartite graph between the color classes C and $V(G_2) - C$:

$$|E(G)| \leq \frac{1}{6}\left(n_1^2 + (4k-7)n_1 + (4k-2k^2)\right) + \binom{n_2-(k-1)}{2} + (k-1)\left(n_2-(k-1)\right).$$

Using the assumption on the number of edges of G we get

$$\frac{1}{6}\left(n^2 + (4k-7)n + (4k-2k^2)\right) < \frac{1}{6}\left(n_1^2 + (4k-7)n_1 + (4k-2k^2)\right) + \binom{n_2-(k-1)}{2} + (k-1)\left(n_2-(k-1)\right).$$

After rearranging the inequality we obtain $n_2 > n_1$ that contradicts our assumption. This completes the proof of the lemma. ∎

The above lemma is sharp when $k+1 \le n \le 2k-2$, at least if $n = k-1+2^l$ then there exists simple graph on n vertices with $1/6 \cdot \big(n^2 + (4k-7)n + (4k-2k^2)\big)$ many edges and with no k connected subgraph:

We define G by describing its complement. $\bar{G}$ will have components as follows: $K_{1,1}, K_{2,2}, K_{4,4}, \ldots, K_{2^{l-1},2^{l-1}}$ and $(k-1) - (2^l - 2)$ many isolated nodes. The number of edges of G can be calculated easily and it turns out to be the promised value. Now we are going to prove that G has no k connected subgraph.

The vertices not in the $K_{2^{l-1},2^{l-1}}$ component of $\bar{G}$ give us a cutset C_0 of size $k-1$ in G. Hence any k connected subgraph of G must be inside this cutset with one component of $G - C_0$. Either way the assumed k connected subgraph must lie in a graph G_1 that the complement of the graph with components: $K_{1,1}, K_{2,2}, K_{4,4}, \ldots, K_{2^{l-2},2^{l-2}}$ and $(k-1) - (2^l-2) + 2^{l-1}$ many isolated nodes. The vertices not in the $K_{2^{l-2},2^{l-2}}$ component of $\bar{G}_1$ give us a cutset C_1 of size $k-1$. Hence any k connected subgraph of G must be inside this cutset with one component of $G_1 - C_1$. Either way the assumed k connected subgraph must lie in a graph G_2 that is the complement of the graph with components: $K_{1,1}, K_{2,2}, K_{4,4}, \ldots, K_{2^{l-3},2^{l-3}}$ and $(k-1) - (2^l-2) + 2^{l-1} + 2^{l-2}$ many isolated nodes. We can continue this procedure till we force the assumed k connected subgraph into a k element subset of $V(G)$, where "there is no enough room".

After the preliminary lemma we can prove the main theorem.

Theorem. *Let $k \ge 3$. Let G be a simple graph assuming that*

$$n = |V(G)| \ge (k-1) + \frac{1}{2}\sqrt{6k^2 - 18k + 16}$$

and

$$|E(G)| > \left(\frac{1}{4}\sqrt{6k^2-18k+16} + k - \frac{3}{2}\right)(n - (k-1)).$$

Then G has a k connected subgraph.

Proof. We prove the theorem by induction on n.

1. case: $(k-1) + \frac{1}{2}\sqrt{6k^2-18k+16} \le n \le (k-1) + \sqrt{6k^2-18k+16}$.

Then it is easy to check that

$$\left(\frac{1}{4}\sqrt{6k^2-18k+16} + k - \frac{3}{2}\right)(n-(k-1)) \ge 1/6\left(n^2 + (4k-7)n + (4k-2k^2)\right),$$

hence the lemma is applicable, providing the claim.

2. case: $(k-1) + \sqrt{6k^2-18k+16} \le n$.

If G itself is k connected we are done.

If not then G is obtained from G_1 and G_2 by gluing them together along a $k-1$ element set, C. Let $n_1 = |V(G_1)|$ and $n_2 = |V(G_2)|$ $((n_1 - (k-1)) + (n_2 - (k-1)) = n - (k-1),$

$k-1 < n_1, n_2 < n$). W.l.o.g. we assume that $n_1 \geq n_2$, hence $n_1 \geq (n + (k-1))/2 \geq (k-1) + \frac{1}{2}\sqrt{6k^2 - 18k + 16}$.

If $|E(G_1)| > \left(\frac{1}{4}\sqrt{6k^2 - 18k + 16} + k - \frac{3}{2}\right)(n_1 - (k-1))$, then we can apply the induction hypothesis and obtain a k connected subgraph of G_1.

If $|E(G_1)| \leq \left(\frac{1}{4}\sqrt{6k^2 - 18k + 16} + k - \frac{3}{2}\right)(n_1 - (k-1))$, then we consider two subcases.

1. subcase: $n_2 \geq (k-1) + \frac{1}{2}\sqrt{6k^2 - 18k + 16}$.

As above we can assume that $|E(G_2)| \leq \left(\frac{1}{4}\sqrt{6k^2 - 18k + 16} + k - \frac{3}{2}\right)(n_2 - (k-1))$ and we obtain an upper bound on the number of edges in G:

$$\begin{aligned} |E(G)| \leq & |E(G_1)| + |E(G_2)| \\ \leq & \left(\frac{1}{4}\sqrt{6k^2 - 18k + 16} + k - \frac{3}{2}\right)(n_1 - (k-1)) \\ & + \left(\frac{1}{4}\sqrt{6k^2 - 18k + 16} + k - \frac{3}{2}\right)(n_2 - (k-1)) \\ = & \left(\frac{1}{4}\sqrt{6k^2 - 18k + 16} + k - \frac{3}{2}\right)(n - (k-1)). \end{aligned}$$

This contradicts the conditions of our theorem.

2. subcase: $n_2 < (k-1) + \frac{1}{2}\sqrt{6k^2 - 18k + 16}$.

Now we can bound the number of edges in G by estimating the edges of G outside G_1 by the number of edges of the complete graph on $V(G_2) - C$ plus the number of edges of the complete bipartite graph between the color classes C and $V(G_2) - C$. Hence

$$\begin{aligned} |E(G)| \leq & \left(\frac{1}{4}\sqrt{6k^2 - 18k + 16} + k - \frac{3}{2}\right)(n_1 - (k-1)) \\ & + \binom{n_2 - (k-1)}{2} + (k-1)(n_2 - (k-1)). \end{aligned}$$

Using the assumption on the number of edges of G we get

$$\begin{aligned} & \left(\frac{1}{4}\sqrt{6k^2 - 18k + 16} + k - \frac{3}{2}\right)(n - (k-1)) \\ & < \left(\frac{1}{4}\sqrt{6k^2 - 18k + 16} + k - \frac{3}{2}\right)(n_1 - (k-1)) \\ & \quad + \binom{n_2 - (k-1)}{2} + (k-1)(n_2 - (k-1)). \end{aligned}$$

After rearranging the inequality we obtain that $n_2 > (k-1) + \frac{1}{2}\sqrt{6k^2 - 18k + 16}$, that contradicts our assumption. This completes the proof of the theorem. ∎

We needed the complicated formulas to make the induction to work. The following corollary makes the claim a little bit weaker but the order of our upper bound on $f_k(n)$ is more transparent.

Corollary. *Let n be at least $k+1$. Then*

$$(3k-4)/2 \cdot (n-(k-1)) \leq f_k(n) \leq (1+\sqrt{6}/4)kn \approx 1.612kn.$$

Proof. The lower bound comes from [4]. Easy to check that $(1+\sqrt{6}/4)kn$ is greater than the lemma's bound on the number of edges if $k+1 \leq n \leq (k-1)+\frac{1}{2}\sqrt{6k^2-18k+16}$ and it is greater than the theorem's bound on the number of edges if $(k-1)+\frac{1}{2}\sqrt{6k^2-18k+16} \leq n$. ∎

3. Applications

We mention two simple applications of the above result. Both of them is just plugging our result into existing proofs.

The first application is vertex partition problem of E. Győri [1]. He asked whether there exists a function $f(s,t)$ such that the vertices of any $f(s,t)$ connected graph can be partitioned into two sets S and T such a way that $G|_S$ is an s connected graph and $G|_T$ is a t connected graph.

The question was answered affirmatively by C. Thomassen [9], M. Szegedy [8] and P. Hajnal [2]. Further on $f(s,t)$ denotes the minimal possible value that is allowed. The proofs use the $f_k(n)$ function. If one plugs our new bound into the best proof ([2] Theorem 4.3.) obtains the following theorem.

Corollary. *If $s \geq 3$, $t \geq 2$ and G is a $(2+\sqrt{6}/2)(s+t)$ connected graph, then there exists an $\{S,T\}$ partition of its vertex set such that $G|_S$ is s connected and $G|_T$ is t connected.*

The second application is Ramsey theory for connectivity. The classical Ramsey theorem says that there exists a function $R_c(k)$ that for arbitrary c coloring of the edges of a complete graph on $R_c(k)$ vertices there must be monochromatic clique of size k. Determining the minimal value of $R_c(k)$ is one of the major open question of graph theory.

D. Matula asked what happens if we look for k connected monochromatic subgraph. The problem turned out to be significantly simpler than the case of complete graphs. It is easy to see that there exists a function $F_c(k)$ such that for arbitrary c coloring of the edges of a complete graph on $F_c(k)$ vertices there must be monochromatic clique of size k. Further on $F_c(k)$ denotes the minimal possible value.

D. Matula gave upper and lower bounds (they are constant factor apart) for $F_c(k)$. The upper bound uses the $f_k(n)$ function. Hence our improved bound immediately gives the following result.

Corollary. $F_c(k) < (2+\sqrt{6}/2)c \cdot k$.

4. Open problems

The major question is W. Mader's conjecture: Is it true that $f_k(n) = (3k-4)/4 \cdot (n - (k-1))$ for large enough n?

One can also consider other classes of graph, like graphs without k connected minor. What is the maximal number of edges in a simple graph on n vertices without k connected minor?

The two applications of our result also hide two nice conjectures.

C. Thomassen conjectures that $f(s,t) = s + t + 1$.

D. Matula conjectures that $F_c(k) = 2c \cdot (k-1) + 1$.

The later two conjectures has relation to the $f_k(n)$ function through existing proof techniques. Settling Mader's conjecture does not resolve the later two problems. Their complete solutions require new ideas.

References

[1] E. Győry, Problem presented at the Sixth Hungarian Colloqim on Combinatorics, Eger, Hungary 1981.

[2] P. Hajnal, Partition of graphs with condition on the connectivity and minimum degree, *Combinatorica*, **3**(1983), 95-99.

[3] L. Lovász, *Combinatorial problems and exercises*, second edition, North Holland Publishing Co., Amsterdam, 1993.

[4] W. Mader, Existenz n-fash zusammenanhagender Teilgraphen in Graphen genügend grossen Kantendichte, *Abh. Sem. Hamburg Univ.*, **37**(1972), 86—97.

[5] W. Mader, Extremal connectivity problems, in: *Infinite and finite sets, Keszthely, Hungary 1973*, Colloquia Mathematica Societatis János Bolyai, 1973, 1089–1093.

[6] W. Mader, Connectivity and edge-connectivity in finite graphs, in: *Surveys in combinatorics*, ed: B. Bollobás, Cambridge University Press, London (1979), 66–95.

[7] D.W. Matula, Ramsey theory for graph connectivity, *Journal of Graph Theory*, **7**(1983), 95–103.

[8] M. Szegedy, personal communication.

[9] C. Thomassen, Graph decompositions with constrains on connectivity and minimal degree, *Journal of Graph Theory*, **7**(1983), 165–167.