COMBINATORICA
Bolyai Society – Springer-Verlag

COMBINATORICA **22** (2) (2002) 269–274

LARGE CLIQUES IN C_4-FREE GRAPHS

ANDRÁS GYÁRFÁS, ALICE HUBENKO*, JÓZSEF SOLYMOSI**

Dedicated to the memory of Paul Erdős

Received October 25, 1999

A graph is called C_4-free if it contains no cycle of length four as an induced subgraph. We prove that if a C_4-free graph has n vertices and at least c_1n^2 edges then it has a complete subgraph of c_2n vertices, where c_2 depends only on c_1. We also give estimates on c_2 and show that a similar result does not hold for H-free graphs—unless H is an induced subgraph of C_4. The best value of c_2 is determined for chordal graphs.

Graphs are understood to be simple, i.e. without loops or multiple edges and this is essential. The order of the largest complete subgraph of G is denoted by $\omega(G)$ and the order of the largest independent set of G is denoted by $\alpha(G)$. A graph is called C_4-free if it contains no cycle of length four as an induced subgraph. The following question has been asked by Paul Erdős: is it true that C_4-free graphs with n vertices and at least c_1n^2 edges must contain complete subgraphs of c_2n vertices, where c_2 depends only on c_1? We give the affirmative answer (Corollary 1) with $c_2=0.4c_1^2$. In fact, we shall prove a more general result, Theorem 1: A C_4-free graph with n vertices and average degree at least a must contain a complete subgraph of order at least $0.1a^2n^{-1}$. The role of C_4 is very important, similar results are not true for H-free graphs as the next proposition shows.

Proposition 1. *Suppose that H is a graph which is not an induced subgraph of C_4. There are H-free graphs G_n with n vertices and at least $n^2/4$ edges with $\omega(G)=o(n)$.*

Mathematics Subject Classification (2000): 05C35

* Supported by OTKA grant T029074.

** Supported by TKI grant stochastics@TUB and by OTKA grant T026203.

Proof. If H contains three independent vertices then define G_n as a 'Ramsey graph', i.e. a graph on n vertices in which no three vertices form an independent set and $\omega(G_n)$ is as small as possible. It is well known that G_n has $n^2/2 - o(n^2)$ edges and $\omega(G_n) \le c(nlogn)^{1/2} = o(n)$ ([4]). If $\alpha(H) \le 2$ then one can define G_n as a balanced complete bipartite graph, in this case it has $\lfloor n^2/4 \rfloor$ edges and $\omega(G_n) = 2$. Since the induced subgraphs of G_n satisfying $\alpha(H) \le 2$ are precisely the induced subgraphs of C_4, the proof is finished.

Theorem 1. *Suppose that G is a C_4-free graph on n vertices with average degree at least a. Then*

$$\omega(G) \ge 0.1 a^2 n^{-1}.$$

Applying Theorem 1 for graphs with at least $c_1 n^2$ edges we obtain the affirmative answer to the question of Erdős.

Corollary 1. *Let G be a C_4-free graph on n vertices with at least $c_1 n^2$ edges for some $0 < c_1 < \frac{1}{2}$. Then*

$$\omega(G) \ge 0.4 c_1^2 n.$$

The lower bound of $\omega(G)$ in Corollary 1 is paralleled by two upper bounds. The first construction is better if c_1 is close to zero and the second is better if c_1 is close to $1/2$.

Construction 1. It is easy to check that clique substitutions into vertices of a graph which does not contain C_4 (not necessarily induced) yields C_4-free graphs. It is well known that there are graphs on k vertices with approximately $(1/2)k^{3/2}$ edges which do not contain C_4. These graphs are the so called 'polarity graphs' ([1], [3]). Substituting cliques into vertices of polarity graphs yields graphs on n vertices, at least $c_1 n^2$ edges and with no cliques larger than $12 c_1^2 n$.

Construction 2. For two graphs, A and B, we denote by $A + B$ the graph C with vertex set $V(C) = V(A) \cup V(B)$ and edge set that consists of $E(A)$, $E(B)$ and all possible edges between $V(A)$ and $V(B)$.

Now consider a graph of form $A + B$ where A is a complete graph and B is a cycle with all chords of length less than $|B|/4$. When $1/4 \le c_1 < 1/2$, with suitable sizes of A, B one can get C_4-free graphs G with n vertices and at least $c_1 n^2$ edges satisfying

$$\omega(G) \le n\left(1 - \frac{3\sqrt{2}}{4}\sqrt{1 - 2c_1}\right).$$

Perhaps Construction 2 is best possible for $c_1 \geq 1/4$. However we could not even decide the case $c_1 = 1/4$: is it true that a C_4-free graph with n vertices and at least $n^2/4$ edges contains a clique of size $n/4$? (The best estimate we know is $n/6$, it comes by a special argument not shown in the paper.) A very special case we could not answer: assume that G is a $2k$-regular C_4-free graph with $4k+1$ vertices, is it true that $\omega(G) > k$? If true, it is best possible shown by the cycle C_{4k+1} with chords of length at most k.

Proof of Theorem 1. It is enough to prove Theorem 1 for the case when the minimum degree of G is at least $a/2$. Indeed, it is well known that every graph G with average degree a contains an induced subgraph G^* with average degree $a^* \geq a$ such that the minimum degree in G^* is at least $a^*/2$ (see for example [2], Proposition 1.2.2). So if we know the theorem for G^*, we get

$$\omega(G) \geq \omega(G^*) \geq \frac{0.1(a^*)^2}{|V(G^*)|} \geq \frac{0.1a^2}{n}.$$

Assuming that the minimum degree of G is at least $a/2$, fix an independent set S with $|S| = t$. Set $S = \{x_1, x_2, \dots, x_t\}$. Let A_i be the set of neighbors of x_i in G and set $m = \max_{i \neq j} |A_i \cap A_j|$. Since G is C_4-free, all the subgraphs $G[A_i \cap A_j]$ are complete graphs, and thus $m \leq \omega(G)$. Using that $|A_i| \geq a/2$ we get

$$ta/2 - \binom{t}{2} m \leq \sum_{1 \leq i \leq t} |A_i| - \sum_{1 \leq i < j \leq t} |A_i \cap A_j| \leq |\bigcup_{1 \leq i \leq t} A_i| \leq n$$

implying that

$$\omega(G) \geq m \geq \frac{ta/2 - n}{\binom{t}{2}}.$$

If $\alpha(G) \geq 4n/a$ then set $t = \lceil 4n/a \rceil$ and we get

$$\omega(G) \geq \frac{\lceil 4n/a \rceil a/2 - n}{\binom{\lceil 4n/a \rceil}{2}} \geq \frac{n}{\binom{\lfloor 4n/a \rfloor + 1}{2}}.$$

If $\alpha(G) \leq 4n/a$ then of course $\alpha(G) \leq \lfloor 4n/a \rfloor$ as well. Now we shall use the following claim: $\omega(G) \geq \frac{n}{\binom{\alpha(G)+1}{2}}$. This follows by selecting an independent set S with $|S| = \alpha(G) = \alpha$. Using the notation introduced above, the $\binom{\alpha}{2}$ sets $A_i \cap A_j$ and the α sets $\{x_i\} \cup B_i$ cover the vertex set of G where B_i denotes the set of vertices whose only neighbor in S is x_i. All of these sets span complete subgraphs because G is C_4-free and from the choice of S. Now we have

$$\omega(G) \geq \frac{n}{\binom{\alpha(G)+1}{2}} \geq \frac{n}{\binom{\lfloor 4n/a \rfloor+1}{2}}.$$

Therefore, in both cases we have

$$\omega(G) \geq \frac{n}{\binom{\lfloor 4n/a \rfloor+1}{2}} \geq \frac{n}{\binom{4n/a+1}{2}} = \frac{a^2}{8n+2a} \geq 0.1a^2/n$$

where in the last step, the trivial inequality $a \leq n$ was used for simple graphs. This finishes the proof.

Our next result shows that if c_1 tends to zero then Corollary 1 can be improved by a factor of ten.

Theorem 2. *Assume that G is a C_4-free graph with at least $c_1 n^2$ edges. Then $\omega(G) \geq c_2^2 n$ where c_2 tends to $4c_1^2$ if c_1 tends to zero.*

Proof. Let A denote the set of vertices in G whose degree is at least $10(c_1/c)^2 n$ where c will be fixed later. Then

$$2c_1 n^2 \leq 2e(G) = \sum_{x \in V} d(x) = \sum_{x \in A} d(x) + \sum_{x \in V-A} d(x) \leq \sum_{x \in A} d(x) + 10n^2(c_1/c)^2$$

which gives

$$\sum_{x \in A} d(x) \geq (2c_1 - 10(c_1/c)^2)n^2.$$

Case 1. There exists $x \in A$ whose neighborhood N in G span at least $c|N|^2$ edges. Now Corollary 1 can be applied to $G[N]$ which gives a complete subgraph of at least $0.4c^2|N| \geq 0.4c^2 10(c_1/c)^2 n = 4c_1^2 n$ vertices proving the theorem.

Case 2. For every $x \in A$, the neighborhood of x, N_x, span at most $c|N_x|^2$ edges. This means that at least $\binom{d(x)}{2} - cd(x)^2$ edges are missing from $G[N_x]$. Let us denote by e the number of edges in G. Then one can estimate p, the number of induced paths with three vertices in G as

$$p \geq \sum_{x \in A} \left(\binom{d(x)}{2} - cd(x)^2 \right) = \sum_{x \in A} \left(\frac{d(x)(d(x)-1)}{2} - cd(x)^2 \right) =$$

$$= (1/2 - c) \sum_{x \in A} d(x)^2 - e \geq (1/2 - c)n^{-1} \left(\sum_{x \in A} d(x) \right)^2 - e \geq$$

$$\geq (1/2 - c)n^3 \left(2c_1 - 10(c_1/c)^2 \right) - e.$$

Since the p paths connect at most $\binom{n}{2}$ pairs there are at least $p\binom{n}{2}^{-1}$ paths connecting the same vertex pair and the midpoints of these paths must form a complete graph since G is C_4-free. Therefore

$$\omega(G) \geq (1-2c)(2c_1 - 10(c_1/c)^2)^2 n - 1$$

which proves the theorem if c is selected so that both c and c_1/c tends to zero with c_1.

A graph is called *chordal* if every cycle of length at least four contains a chord. Chordal graphs are clearly induced C_4-free, thus Theorem 1 and Corollary 1 are valid for them. For chordal graphs we get (asymptotically) the best result.

Theorem 3. *Suppose that G is a chordal graph on n vertices with at least $\lceil c_1 n^2 \rceil$ edges, then*

$$\lfloor (1 - \sqrt{1-2c_1})n \rfloor \leq \omega(G).$$

Moreover, the the chordal graph which is the sum of an independent set of order $t = \lceil n\sqrt{1-2c_1} \rceil$ and a complete graph K_{n-t} shows that the inequality asymptotically is best possible.

Proof. It is well known that chordal graphs are perfect. See [2] Proposition 5.5.2. Thus the vertices of a chordal graph G can be partitioned into $\omega = \omega(G)$ independent sets $X_1, X_2, \ldots, X_\omega$. Since G is chordal, the edges between X_i and X_j form an acyclic subgraph of G. Thus, we obtain

$$c_1 n^2 \leq \lceil c_1 n^2 \rceil \leq \sum_{i<j} (|V(X_i)| + |V(X_j)| - 1) = (\omega - 1)\, n - \binom{\omega}{2} \leq \omega n - \omega^2/2.$$

Thus $\omega^2 - 2\omega n + 2c_1 n^2 \leq 0$ from which the stated lower bound follows.

To check that the inequality is asymptotically sharp for the chordal graph defined in the theorem is routine so it is omitted.

Acknowledgement. We appreciate the useful comments and suggestions of the referees.

References

[1] W. G. Brown: On graphs that do not contain a Thomsen graph, *Canad. Math. Bull.* **9** (1966), 281–285.

[2] R. Diestel: *Graph Theory,* Springer, 1997.

[3] P. Erdős, A. Rényi, V. T. Sós: On a problem of graph theory, *Studia Sci. Math. Hungar.,* **1** (1966), 51–57.

[4] J. H. Kim: The Ramsey Number $R(3,t)$ Has Order of Magnitude $t^2 log t$, *Random Structures and Algorithms,* **7** (1995), 173–207.

András Gyárfás

Computer and Automation Research Institute Hungarian Academy of Sciences
Gyarfas@luna.aszi.sztaki.hu

Alice Hubenko

Computer and Automation Research Institute Hungarian Academy of Sciences
hubenko@msci.memphis.edu

József Solymosi

Computer and Automation Research Institute Hungarian Academy of Sciences
solymosi@euclid.ucsd.edu