ELSEVIER

Discrete Mathematics 150 (1996) 31–48

DISCRETE MATHEMATICS

Cycle-saturated graphs of minimum size

C.A. Barefoot[a], L.H. Clark[b, 1], R.C. Entringer[b], T.D. Porter[b, 1], L.A. Székely[b, *], Zs. Tuza[c]

[a] *New Mexico Institute of Mining & Technology, Socorro, NM 87801, USA*
[b] *University of New Mexico, Albuquerque, NM 87131, USA*
[c] *Computer and Automation Institute of the Hungarian, Academy of Sciences, H-1111 Budapest, Kendu u. 13-17, Hungary*

Received 13 September 1993; revised 22 September 1994

Abstract

A graph G is called C_k*-saturated* if G contains no cycles of length k but does contain such a cycle after the addition of any new edge. Bounds are obtained for the minimum number of edges in C_k-saturated graphs for all $k \neq 8$ or 10 and n sufficiently large. In general, it is shown that the minimum is between $n + c_1 n/k$ and $n + c_2 n/k$ for some positive constants c_1 and c_2. Our results provide an asymptotic solution to a 15-year-old problem of Bollobás.

1. Introduction

Given a graph G, we denote the vertex set, edge set and complement of G by $V(G)$, $E(G)$ and $\bar{G}$, respectively; the *order* and *size* of G are the cardinalities of $V(G)$ and $E(G)$. The path, cycle and complete graph of order k will be denoted by P_k, C_k and K_k, respectively. The distance between two vertices u and v of the graph G is denoted by $d_G(u, v)$. By *leaf* we will mean a copy of K_2 and by *triangle*, a copy of K_3. (Our definition of leaf differs from the usual, attaching a leaf to the graph does produce an endvertex however.) If H is a subgraph of G, we write $H \subseteq G$. A graph G_1 is said to be *attached* to the graph G at the vertex $\mathbf{0}$ iff $V(G_1) \cap V(G) = \{\mathbf{0}\}$. Of course, if G_1 is vertex-symmetric the vertex of attachment of G_1 need not be specified.

Given the graph F, the graph G is said to be F*-saturated* if $F \nsubseteq G$ but $F \subseteq G + e$ for every $e \in E(\bar{G})$. We note that if F has order greater than the order, n, of G then K_n is the only F-saturated graph (and vacuously so) of order n. Thus, we restrict our attention

* Corresponding author. Department of Computer Science, Eötvös Lorand University, H-1088 Budapest, Hungary. Supported in part by OTKA T 014105.
[1] Permanent address: Department of Mathematics, Southern Illinois University, Carbondale, IL, 62901, USA.

SSDI 0012-365X(95)00173-5

to $n \geqslant |V(F)|$ and for such n we define $sat(n, F)$ to be the minimum size of an F-saturated graph of order n.

Apparently, this notion occurred first with Erdős et al. [9], who determined the value of $sat(n, K_k)$ more than 25 years ago. These authors posed a related problem in which all the graphs involved are required to be bipartite. This problem was solved independently by Bollobás [1, 2] and Wessel [17, 18]. Since the paper of Erdős et al. several related results, both special and general, have been obtained. As special results, Kászonyi and Tuza [11] determined exact values of $sat(n, K_{1,k})$, $sat(n, kK_2)$ and $sat(n, P_k)$. In general, they proved $sat(n, F) = O(n)$ for fixed F; Tuza has conjectured, further, that for every graph F the limit $\lim_{n\to\infty}(sat(n, F))/n$ exists (see [15]). Truszczynski and Tuza [14] characterized the graphs F for which this limit exists and is less than 1. For results on hypergraphs see [8, 15]. Our point of departure is the class of problems with $F = C_k$. As pointed out by Bollobás [3, p. 103], these form a 'rather neglected set of unsolved problems'.

One of us (unpublished) had asked if $sat(n, C_k) = 3n/2 + o(n)$ for every fixed $k > 3$ and infinitely many n. Three exact values of $sat(n, C_k)$ and bounds, which we will determine, are listed in Table 1 in summary form. These show that there can be at

Table 1
Bounds and exact values for $sat(n, C_k)$, n sufficiently large

k	$sat(n, C_k)$	$n \geqslant$	Reference
3	$= n - 1$	3	
4	$= \left\lfloor \frac{3n-5}{2} \right\rfloor$	5	Theorem A [12, 16]
5	$\leqslant \frac{10n-4}{7}$	8	Proposition 1
$\geqslant 5$	$\left(1 + \frac{1}{2k+8}\right)n$	k	Theorem 1
6	$\leqslant \frac{3n}{2}$	11	Proposition 2
7	$\leqslant \frac{7n+12}{5}$	10	Proposition 3
9	$\leqslant \frac{43n}{22} + O(1)$	9	[11]
$\geqslant 9$ and $\equiv 1 \bmod 2$	$\leqslant \left(1 + \frac{6}{k-3}\right)n + O(k^2)$	$3k$	Theorem 2
12	$\leqslant \frac{29n+99}{22}$	12	Proposition 4
$\geqslant 14$ and $\equiv 0 \bmod 2$	$\leqslant \left(1 + \frac{4}{k-2}\right)n + O(k^3)$	$\frac{3k^2}{4} + \frac{5k}{2}$	Theorem 3
$\geqslant 20$ and $\equiv 4 \bmod 8$	$\leqslant \left(\frac{5}{4} + \frac{3}{4k-4}\right)n + \frac{k}{2}$	k	Theorem 4
n	$= \left\lfloor \frac{3n+1}{2} \right\rfloor$	53	Theorem B [5–7]

most eight values of k for which $sat(n, C_k) = 3n/2 + o(n)$, namely: 4, 6, 8, 9, 10, 11, 13 and 15. On the other hand, $sat(n, C_k) < 3n/2 + o(n)$ has not yet been proved for k tending to infinity faster than $n^{1/2}$, $k < n$.

One may also ask for $sat(n, C_k)$ when n and k both tend to ∞. Here different results may be expected over different regions. Our results give asymptotic formulas for $sat(n, C_k)$ up to the second term when k is allowed to tend to infinity slowly.

For odd $k \geqslant 9$,

$$\left(1 + \frac{1}{2k+8}\right)n \leqslant sat(n, C_k) \leqslant \left(1 + \frac{6}{k-3}\right)n + O(k^2)$$

so that, for n and odd k both tending to ∞ with $k = o(n^{1/3})$,

$$sat(n, C_k) = n + \Theta(n/k)$$

and

$$sat(n, C_k) = (1 + o(1))n$$

for $k = o(n^{1/2})$.

For even $k \geqslant 14$,

$$\left(1 + \frac{1}{2k+8}\right)n \leqslant sat(n, C_k) \leqslant \left(1 + \frac{4}{k-2}\right)n + O(k^3)$$

so that for n and even k tending to ∞ with $k = o(n^{1/4})$,

$$sat(n, C_k) = n + \Theta(n/k)$$

and

$$sat(n, C_k) = (1 + o(1))n$$

for $k = o(n^{1/3})$.

For $k = 4m$, m odd and at least 3 we have a reasonably good upper bound over a much wider region:

$$sat(n, C_k) \leqslant \left(\frac{5}{4} + \frac{3}{4k-4}\right)n + \frac{k}{2}.$$

2. C_k-saturated graphs, small k

As pointed out by Ollman, a few minutes reflection shows that the unique C_3-saturated graph of minimum size is the star $K_{1,n-1}$ so that $sat(n, C_3) = n - 1$. Already for the next entry in the table, $sat(n, C_4) = \lfloor (3n-5)/2 \rfloor$, the original proof [12] was 20 pages long; a later proof, [16], is still half that length. (It should be noted that the value of $sat(n, C_4)$ is misstated in [3, p. 167]).

Theorem A (Ollmann [12] and Tuza [16]). $sat(n, C_4) = \lfloor (3n-5)/2 \rfloor$, $n \geqslant 5$. *The only C_4-saturated graphs of order $n \geqslant 5$ and minimum size are pictured in Fig.* 1.

In reading Fig. 1 it is to be understood that in (i) there is a total of $(n-6)/2$ triangles attached to the base triangle at the bold vertices and in (ii) there is a total of $(n-5)/2$ triangles attached to the base triangle or the pentagon at the bold vertices.

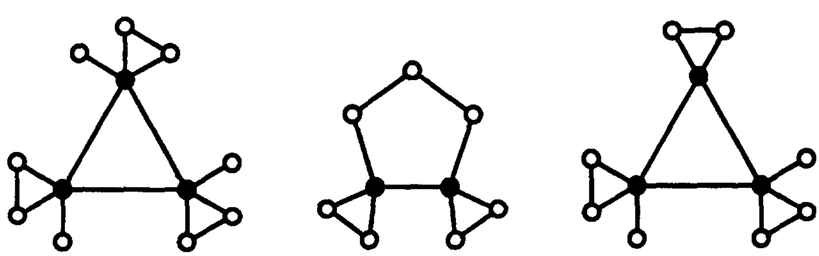
The graphs of Fig. 1 suggest a general construction that we will exploit on several occasions. Call a graph G a *C_k-builder* iff G is C_k-saturated and has a vertex, labelled **0**, such that if the **0** vertices of two copies are identified then the resulting graph is C_k-saturated. Clearly, if G is a C_k-builder and we identify the **0** vertices of $s \geqslant 1$ copies of G, then the resulting graph is C_k-saturated. We use this technique to obtain upper bonds for $sat(n, C_k)$, $k = 5$, 6, 7 and 12.

Some obvious properties of C_k-saturated graphs will be used without further reference. Certainly, all blocks of such graphs must themselves be C_k-saturated graphs. Furthermore, if one of those blocks is a copy of K_r for some r, $1 \leqslant r \leqslant k-1$, then it may be replaced by K_s for arbitrary s, $r \leqslant s \leqslant k-1$ and the resulting graph is still C_k-saturated.

Throughout this section we have attempted to strike a reasonable balance with respect to the inclusion of details. We have learned from experience that carelessness in verifying the existence of paths and nonexistence of cycles generally causes them to not exist, and exist, respectively.

Proposition 1. $sat(n, C_5) \leqslant (10n-4)/7$, $n \geqslant 8$.

Proof. It is easily verified that the graph G pictured in Fig. 2 is C_5-saturated. Since there is a path of length 2 from vertex **0** to every vertex of G different from **0**, it immediately follows that G is a C_5-builder. Let G_s, $s \geqslant 1$, be the graph obtained from



(i) (ii)

Fig. 1. The minimal C_4-saturated graphs; (i) n even, (ii) n odd.

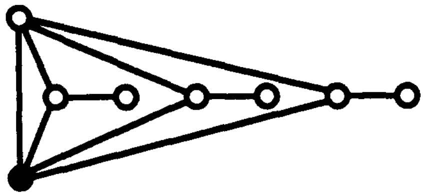


Fig. 2. A C_5-builder.

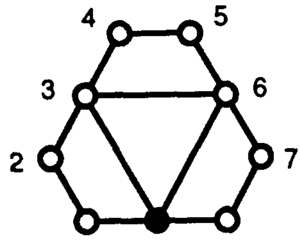


Fig. 3. A C_6-builder.

s copies of G by identifying the **0** vertices. Then G_s has order $n = 7s + 1$ and size $10s$. When $n = 7s + r$, $2 \leqslant r \leqslant 7$, a C_5-saturated graph with order n and size $10s + \lfloor 3r/2 \rfloor - 1$ can be constructed from G_s or G_{s+1} as follows.

If $r = 3$ or 5, attach one or two triangles, respectively, at **0** in G_s.

If $r = 2$, 4 or 6, delete one endvertex from G_s and attach one, two or three, respectively, triangles at **0**.

If $r = 7$ (and alternatively, for $r = 6$), delete one (two) endvertex (vertices) from G_{s+1}.

It is a simple matter to verify that all the constructions described above give C_5-saturated graphs and have the required numbers of edges. We leave it to the reader to do this. □

Proposition 2. $sat(n, C_6) \leqslant 3n/2$, $n \geqslant 11$.

Proof. Any 6-cycle in the graph G pictured in Fig. 3 must contain two of the paths **0**123, 3456 and 678**0** and so cannot exist. In view of the symmetry of G, the existence of the paths **0**65432, **0**87654, 123654, 1**0**8765, and 123678 shows that G is C_6-saturated. There are paths of length 3,1 and 4 from **0** to each of the vertices 1 and 8. Furthermore, there are paths of length 2 and 3 from **0** to each of the vertices 2, 3, 4, 5, 6 and 7. Thus, for any two (not necessarily distinct) vertices of G both different from **0**, there are paths from these vertices to **0** the sum of whose lengths is 5. We conclude that G is a C_6-builder.

Let G_s, $s \geqslant 1$, be the graph obtained from s copies of G by identifying the **0** vertices. Then G_s has order $n = 8s + 1$ and size $12s$. When $n = 8s + r$, $2 \leqslant r \leqslant 8$, a C_6-saturated graph with order n and size $12s + 3\lfloor r/2 \rfloor$ can be constructed from G_s or G_{s-1} as follows.

If $r = 3$, 5 or 7, to G_s attach one triangle at each of the vertices **0**, **0** and 3, **0**, 3 and 6, respectively, of one copy of G in G_s.

If $r = 4$, 6 or 8, attach a K_4 to G_s at **0**, also attach no triangles, one triangle at 3 and one triangle at 6, respectively, of one copy of G in G_s.

If $r = 2$ and $s \geqslant 2$ attach one K_4 at **0** and attach one triangle at each of the vertices **0**, 3 and 6 of one copy of G in G_{s-1}.

In this case also, it is simple matter to verify that all the constructions described above give C_6-saturated graphs and have the required number of edges. We again leave this to the reader. □

Proposition 3. $sat(n, C_7) \leqslant (7n + 12)/5$, $n \geqslant 10$.

Proof. Let G be the Petersen graph with seven leaves attached as indicated in Fig. 4. In proving that G is a C_7-builder we make heavy use of the fact that the Petersen graph is vertex-transitive.

The existence of the path **0**327891 implies the existence of paths of length 6 between any two distinct, nonadjacent vertices of the Petersen graph.

The existence of the paths **0**67891 and **0**67219 implies the existence of paths of length 5 between any two distinct vertices of the Petersen graph. Consequently, we have paths of length 6 between any two nonadjacent vertices in G, one in the Petersen graph, the other not.

The existence of the paths **0**6789 and **0**6721 implies the existence of paths of length 4 between any two distinct vertices of the Petersen graph. Consequently, we have paths of length 6 between any two distinct vertices in G, not in the Petersen graph.

Since it is easily verified that the Petersen graph has no 7-cycles (see [4], for example) we conclude that G is C_7-saturated.

Now consider the graph H formed by taking two copies of G, deleting the endvertex adjacent to **0** from one of the copies, and identifying the **0**-vertices. Suppose vertex u is in one copy of G, vertex v is in the other copy of G and $u \neq \mathbf{0} \neq v$.

The existence of the path **0**912 and **0**67219 implies the existence of paths of length 6 between u and v in the case u and v are in the Petersen graphs.

The existence of the paths **0**91, **0**321 and **0**6721 implies the existence of paths of length 6 between u and v in the case u is in the Petersen graph and v is not.

The existence of the path **0**91 implies the existence of paths of length 6 between u and v in the case neither u nor v is in the Petersen graph.

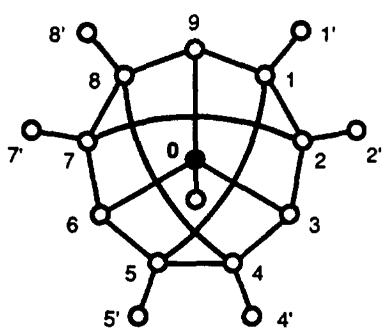


Fig. 4. A C_7-builder.

Since H obviously has no 7-cycles we conclude that H is C_7-saturated.

Consequently, if we identify all of the **0** vertices of $s \geqslant 1$ copies of G and delete all but one of the end vertices adjacent to **0**, the resulting graph, G_s, is C_7-saturated, has order $n = 15s + 2$ and size $21s + 1$. For $1 \leqslant r \leqslant 7$ a C_7-saturated graph with order $n = 15s - r + 2$ and size $21s - r + 1$ can be constructed from G_s by deleting r leaves. For $1 \leqslant r \leqslant 7$ a C_7-saturated graph with order $n = 15s + r + 2$ and size $21s + 2r + 1$ can be constructed from G_s by replacing each of r leaves by a triangle. □

Proposition 4. $sat(n, C_{12}) \leqslant (29n + 99)/22$, $n \geqslant 12$.

Proof. We denote by J_3 the Tietze graph. If the triangle of this graph is contracted to a vertex the resulting graph is the Petersen graph. Since the Petersen graph is nonhamiltonian so is the graph J_3 of Fig. 5. We have verified, by computer, that every two nonadjacent vertices of the Tietze graph are joined by paths of lengths 9, 10 and 11 and that every two nonadjacent vertices are joined by paths of lengths 9 and 10. Thus, the graph G obtained from J_3 by attaching a leaf at every vertex is C_{12}-saturated since every two nonadjacent vertices are the end vertices of a path of length 11. It was also verified by computer that for every pair of vertices u and v of J_3, distinct from **0** but not necessarily from each other, there are paths from u to **0** and from v to **0** the sum of whose lengths is 9 and 10 and 11. We conclude that G is a C_{12}-builder.

Consequently, if we identify all of the **0** vertices of $s \geqslant 1$ copies of G and delete all but 1 of the end vertices attached at **0**, the resulting graph, G_s, is C_{12}-saturated, has order $n = 22s + 2$ and size $29s + 1$. For $1 \leqslant r \leqslant 12$ a C_{12}-saturated graph with order $n = 22s - r + 2$ and size $29s - r + 1$ can be constructed from G_s by deleting r leaves. For $0 \leqslant r \leqslant 9$ a C_{12}-saturated graph with order $n = 22s + r + 2$ and size $29s + 2r + 1$ can be constructed from G_s by replacing each of r leaves by a triangle. □

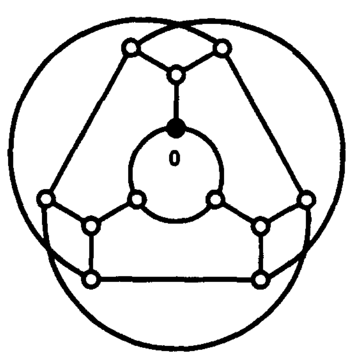


Fig. 5. The graph J_3; a C_{12}-builder.

3. A lower bound

We will need some additional terminology and notation for this section. If A and B are two disjoint subsets of vertices of a graph G we will denote by $e(A, B)$ the number of edges of G with one endvertex in A and the other in B and by $G[A]$ the subgraph of G induced by A. Given the vertex v of G we denote by $N_G(v)$ the set of vertices of G that are adjacent to v. We will not require all the subsets of a partition to be nonempty. We assume the implicit duplication of vertices when we write 'the path $abcPdef$ where P is the path $cxyzd$' will cause no confusion.

Theorem 1. *For $n \geqslant k \geqslant 5$, $sat(n, C_k) \geqslant n(1 + 1/(2k + 8))$.*

Proof. Assume, to the contrary, that there are counterexamples and let G be one of smallest order n. Necessarily, G is connected. Let

$$L = \{v \in V(G) \mid d_G(v) = 1\}.$$

(i) If $v_1, v_2 \in L$ then $d(v_1, v_2) \geqslant 3$. Since G is connected and $n \geqslant 3$, $v_1v_2 \notin E(G)$. Suppose v_1wv_2 is a path in G. Then $G + v_1v_2$ contains a k-cycle which must be $v_1wv_2v_1$ and $k = 3$.

(ii) If $v \in L$ and $vw \in E(G)$ then $d_G(w) \geqslant 3$. From (i) we have $d_G(w) \geqslant 2$. Suppose $d_G(w) = 2$ and let vwx be a path in G. Then $G + vx$ contains a k-cycle which must be $vxwv$ and $k = 3$.

(iii) If $v \in L$ and $vwxy$ is a path in G then $d_G(w) = 3$ implies $d_G(x) \geqslant 3$. Suppose $d_G(x) = 2$. Since $k \geqslant 5$, $G + vy$ contains a k-cycle which must be $vwPyv$ where P is a $w - y$ path in $G - v - x$ of length $k - 2$. But then $xwPyx$ is a k-cycle in G.

(iv) If $v_1, v_2 \in L$ and $v_1w_1w_2v_2$ is a path in G then $d_G(w_1), d_G(w_2) \geqslant 4$. From (ii) we have $d_G(w_1)$, $d_G(w_2) \geqslant 3$. Suppose $N_G(w_1) = \{v_1, w_2, x_1\}$. Since $v_2 \neq x_1$ and $k \geqslant 5$, $G + v_2x_1$ contains a k-cycle which must be $v_2x_1Pw_2v_2$ where P is a $x_1 - w_2$ path in $G - v_2 - w_1$ of length $k - 2$, so that $w_1x_1Pw_2w_1$ is a k-cycle in G.

(v) No endblock of G is a cycle. Let C_l be an l-cycle containing the cutvertex v of G. For $l = 3$, $H = G - (V(C_l) - \{v\})$ is C_k-saturated with

$$e(G) = e(H) + 3 \geqslant (n - 2)\left(1 + \frac{1}{2k + 8}\right) + 3 \geqslant n\left(1 + \frac{1}{2k + 8}\right)$$

while for $l \geqslant 4$, some chord may be added to C_l without forming a k-cycle.

Let

$$A = \{v \in V(G) \mid d_G(v) = 2\},$$

$$B = \{v \in V(G) \mid d_G(v) = 3,\ N_G(v) \cap L \neq \emptyset\},$$

$$C = \{v \in V(G) \mid d_G(v) \geqslant 3,\ N_G(v) \cap L = \emptyset\},$$

$$D = \{v \in V(G) \mid d_G(v) \geqslant 4,\ N_G(v) \cap L \neq \emptyset\}.$$

Then $\{L, A, B, C, D\}$ partitions $V(G)$ so that, by (i), each vertex of $B \cup D$ is adjacent to precisely one vertex of L, and, by (ii), each vertex of L is adjacent to a vertex of $B \cup D$ (see Fig. 6). Thus, $|B| + |D| = |L|$. From (iii) and (iv) we have

$$e(A, B) = e(A, D) = e(B, B) = e(B, D) = 0.$$

Since G is connected, $G[A]$ is the union of disjoint paths $P_1, \dots, P_t$.

(vi) For all i, $1 \leqslant i \leqslant t$, we have $|V(P_i)| \leqslant k - 1$. Let $P = v_1 v_2 \dots v_l$ be a path in $G[A]$ with $l \geqslant k$ and $v_1 x$, $v_l y \in E(G)$ where $x, y \in C$. By (v), $x \neq y$. Then $G + v_1 v_3$ contains a k-cycle C_k which must be $v_1 v_3 \dots v_l y Q x v_1$ where Q is a $y - x$ path of length at least 1. But then C_k must have length at least $l + 1 \geqslant k + 1$ which is impossible.

Let

$$S = \{c \in C \mid d_G(c) = e(A \cup B, \{c\})\},$$

$$S_1 = \{c \in S \mid e(B, \{c\}) = 0\},$$

$$S_2 = \{c \in S \mid e(B, \{c\}) \neq 0\}.$$

Obviously, $\{S_1, S_2\}$ is a partition of the independent set S and S_1 is the set of vertices with degree at least 3 that are adjacent only to vertices with degree 2.

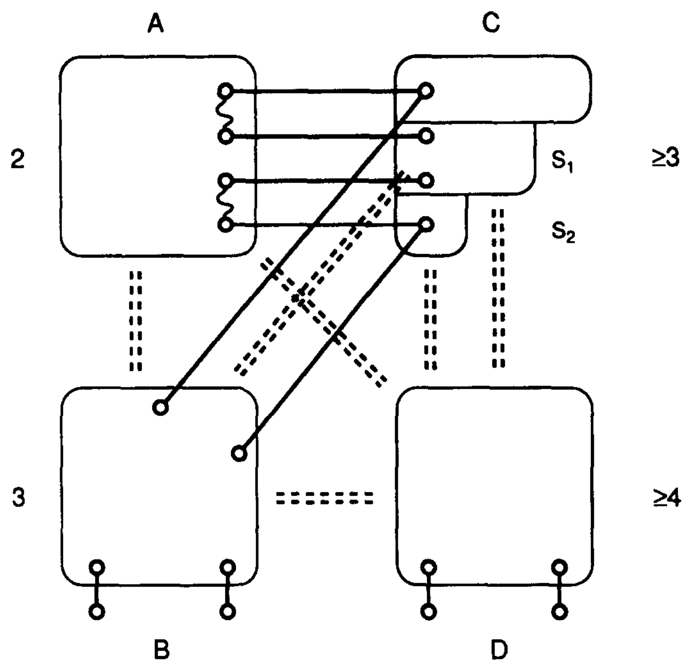


Fig. 6. The partition of $V(G)$.

For $s \in S$ let

$$P(s) = \{P \mid \text{ is a (component) path in } G[A],\ e(P, \{s\}) \neq 0\},$$

$$O(s) = \{P \in P(s) \mid P \notin P(r), \text{ for all } r \in S - \{s\}\},$$

$$T(s) = \{P \in P(s) \mid P \in P(r), \text{ for some } r \in S - \{s\}\}.$$

Obviously, $\{O(s), T(s)\}$ is a partition of $P(s)$, a path P in $P[A]$ is in $P(s)$ for at most two s in S and $O(s) \cap O(r) = \emptyset$ for distinct s, r in S.

(vii) For all $s \in S_1$ we have

$$\sum_{P \in P(s)} |V(P)| \leqslant (k-1)(d_G(s) - 1).$$

If $|V(P)| \leqslant 2$ for all $P \in P(s)$ then

$$\sum_{P \in P(s)} |V(P)| \leqslant 2d_G(s) \leqslant (k-1)(d_G(s) - 1)$$

since $k \geqslant 5$ and $d_G(s) \geqslant 3$. Thus, we may suppose $P = v_1 v_2 v_3 \ldots v_l \in P(s)$ with $l \geqslant 3$ and $v_1 s, v_l s' \in E(G)$. By (v), $s \neq s'$. Consequently, $G + v_1 v_3$ contains a k-cycle which must be $v_1 v_3 \cdots v_l s' Q s v_1$, where Q is an $s' - s$ path in $G - \{v_1, \ldots, v_l\}$ and Q must contain some $P' \in P(s)$ as a subpath since $s \in S_1$. Then

$$|V(P)| + |V(P')| \leqslant k - 1$$

and, by (vi),

$$\begin{aligned}\sum_{P \in P(s)} |V(R)| &= |V(P)| + |V(P)'| + \sum_{R \in P(s) - \{P, P'\}} |V(R)| \\ &\leqslant k - 1 + (k-1)(d_G(s) - 2) \\ &\leqslant (k-1)(d_G(s) - 1).\end{aligned}$$

Let

$$\mathscr{P} = \bigcup_{s \in S_1} P(s),$$

$$\mathscr{O} = \bigcup_{s \in S_1} O(s),$$

$$\mathscr{T} = \bigcup_{s \in S_1} T(s).$$

Note that $\{O(s) \mid s \in S_1\}$ partitions $\mathscr{O}$.

(viii) We have $|S_1| \leqslant \frac{2}{3}|\mathscr{P}| \leqslant \frac{2}{3}t$. Here

$$3|S_1| \leqslant \sum_{s \in S_1} d_G(s) = e\left(\bigcup_{P \in \mathscr{P}} V(P), S_1\right) = 2|\mathscr{P}| - |\mathscr{O}| \leqslant 2|\mathscr{P}|.$$

(ix) We have $t \geqslant |A|/(k-1) + |S_1|/2$.
By (vii),

$$\begin{aligned}\sum_{P \in \mathscr{P}} |V(P)| &= \frac{1}{2}\sum_{s \in S_1}\sum_{P \in P(s)} |V(P)| + \frac{1}{2}\sum_{s \in S_1}\sum_{P \in O(s)} |V(P)| \\ &\leqslant \frac{1}{2}\sum_{s \in S_1}(k-1)(d_G(s)-1) + \frac{1}{2}\sum_{s \in S_1}(k-1)|O(s)| \\ &= \frac{k-1}{2}(2|\mathscr{P}| - |\mathscr{O}| - |S_1|) + \frac{k-1}{2}|\mathscr{O}| \\ &= (k-1)(|\mathscr{P}| - |S_1|/2),\end{aligned}$$

while, by (vi),

$$\sum_{P \in \{P_1, \ldots, P_{\mathscr{P}}\}} |V(P)| \leqslant (k-1)(t - |\mathscr{P}|)$$

so that

$$\begin{aligned}|A| = \sum_{i=1}^{t} |V(P_i)| &\leqslant (k-1)(|\mathscr{P}| - |S_1|/2) + (k-1)(t - |\mathscr{P}|) \\ &\leqslant (k-1)(t - |S_1|/2).\end{aligned}$$

Partition S_2 as follows:

$$T_1 = \{s \in S_2 \mid e(B, \{s\}) = 1\},$$

$$T_2 = \{s \in S_2 \mid e(B, \{s\}) \geqslant 2\}.$$

(x) We have $2|S_2| \leqslant 2|B| + |T_1|$. Here

$$2|S_2| - |T_1| = |T_1| + 2|T_2| \leqslant e(B, s_2) \leqslant 2|B|.$$

Let

$$\mathscr{P}^* = \bigcup_{s \in T_1} P(s),$$

$$\mathscr{O}^* = \bigcup_{s \in T_1} O(s),$$

$$\mathscr{T}^* = \bigcup_{s \in T_1} T(s).$$

(xi) We have $|T_1| \leqslant |\mathscr{P}^*| \leqslant t$. Here

$$2|T_1| \leqslant \sum_{s \in T_1} (d_G(s) - 1) = e(A, T_1) = 2|\mathscr{P}^*| - |\mathscr{O}^*| \leqslant 2|\mathscr{P}^*|.$$

(xii) Finally, we have $|S_2| \leqslant 2|B|$. Here

$$|S_2| \leqslant e(B, S_2) \leqslant 2|B|.$$

Now, by counting degrees,

$$e \geqslant |A| + 2|B| + 3|C|/2 + 5|D|/2 = n + |D|/2 + |C|/2, \tag{1}$$

while, by counting edges,

$$e \geqslant \sum_{1}^{t}(|V(P_i)| + 1) + 3|B| + (|C| - |S|)/2 + 3|D|/2 + |D|$$

$$= |A| + t + 3|B| + |C|/2 + 5|D|/2 - |S|/2$$

$$= n + |D|/2 - |C|/2 + t + |B| - |S|/2, \tag{2}$$

since $e(C \cup D, \{v\}) \geqslant 1$ for all $v \in C - S$, so that

$$e(G[C \cup D]) \geqslant (|C| - |S|)/2 + 3|D/2.$$

First, by (1) and (2),

$$e \geqslant \max\{n + |D|/2 + |C|/2, n + |D|/2 - |C|/2 + t + |B| - |S|/2\}.$$

Now fix n, $|B|$, $|D|$, $|S|$, and t and define

$$f(|A|, |C|) = \max\{n + |D|/2 + |C|/2, n + |D|/2 - |C|/2 + t + |B| - |S|/2\}.$$

Then f minimizes at $|C| = |B| + t - |S|/2$ so that

$$e \geqslant \min_{|A|,|C|} f(|A|, |C|) = n + |D|/2 + t/2 + |B|/2 - |S|/4. \tag{3}$$

Next, using (ix), (xii), (1) and (2) we argue in a similar manner and obtain

$$e \geqslant \min_{|A|,|C|} \max\{n + |D|/2 + |C|/2, n + |D|/2 - |C|/2 + |A|/(k-1)\}$$

$$= n + |D|/2 + |A|/2(k-1)$$

$$= n\left(1 + \frac{1}{2k}\right) + \frac{k-2}{2k}|D| - \frac{1}{k}|B|. \tag{4}$$

For $|B| \leqslant (n/2)(1 - \bar{c})$, ($\bar{c}$ to be determined later),

$$e \geqslant n\left(1 + \frac{1}{2k}\right) - \frac{|B|}{k} \geqslant n\left(1 + \frac{\bar{c}}{2k}\right),$$

by (4), while

$$|S| = |S_1| + |S_2| \leqslant 7t/6 + |B|,$$

by (viii), (x), and (xi) so that, for $|B| \geqslant (n/2)(1 - \bar{c})$,

$$e \geqslant n + t/2 + |B|/2 - |S|/4 \geqslant n + 5t/24 + |B|/4 \geqslant n(9/8 - \bar{c}/8) \geqslant n\left(1 + \frac{\bar{c}}{2k}\right),$$

by (3) provided

$$\bar{c} \leqslant \frac{k}{k+4}.$$

Consequently,

$$e \geqslant n\left(1 + \frac{1}{2k+8}\right). \qquad \square$$

4. Upper bounds

A special case of a construction in [11] implies that

$$sat(n, C_k) \leqslant sat(n, P_{k-1}) + n - 1$$

for all k and n, $3 \leqslant k \leqslant n$. Combining this inequality with the value of $sat(n, P_{k-1})$ obtained in [11] gives the upper bound, for $k \geqslant 6$,

$$sat(n, C_k) \leqslant c_k n + O(1),$$

where

$$c_k = \begin{cases} 2 - \dfrac{1}{2^{k/2} - 2}, & k \text{ even}, \\ 2 - \dfrac{1}{3 \cdot 2^{(k-3)/2} - 2}, & \text{odd}. \end{cases}$$

Until now this has been the only general upper bound for $sat(n, C_k)$.

We present four constructions in this section all of which give upper bounds for $sat(n, C_k)$. Each of the first three contains a complete or nearly complete subgraph of order at least k and, consequently, these bounds are bad as long as n is small. The fourth construction works for small n/k ratios as well.

Theorem 2.

$$sat(n, C_{2s+1}) \leqslant \frac{s+2}{s-1}(n-r) + \binom{2s-3}{2} + \binom{r-1}{2} - 10,$$

where $n = (m+2)(s-1) + r$, $1 \leqslant r \leqslant s-1$, $s \geqslant 4$ *and* $m \geqslant 3$.

Consider the graph $G_{s,m}$ pictured in Fig. 7. It consists of a subgraph H, a complete subgraph K_{2s-3} with the edge OO' deleted and a complete subgraph K_r attached to an arbitrary vertex $O'' \notin \{O, O'\}$ of the subgraph $K_{2s-3} - OO'$. The subgraph H consists of two vertices A and B joined by m internally disjoint paths P_i, $1 \leqslant i \leqslant m$, each of length s. Vertex B is adjacent to all vertices of the subgraph $K_{2s-3} - OO'$. Both O and O' are adjacent to all vertices of H that are adjacent to B.

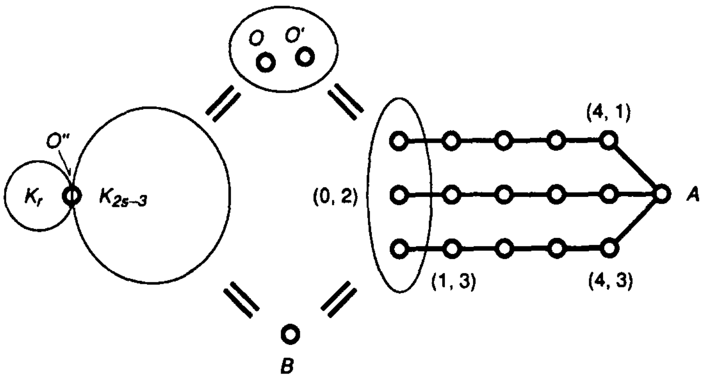


Fig. 7. The graph $G_{s,m}$; a C_k-saturated graph, $k = 2s + 1 \geqslant 9$.

It is easily seen that $G_{s,m}$ has order $n = 2 + m(s-1) + 2s - 3 + r - 1 = (m+2)(s-1) + r$ and size

$$ms + \binom{2s-3}{2} - 1 + 2m + 2s - 5 + \binom{r}{2}$$

$$= (m+2)(s+2) + \binom{2s-3}{2} + \binom{r}{2} - 10$$

$$= \frac{s+2}{s-1}(n-r) + \binom{2s-3}{2} + \binom{r}{2} - 10.$$

It remains to show the existence of paths of length $2s$ between all pairs of nonadjacent vertices of $G_{s,m}$. This is done in Part I. In Part II we show that $G_{s,m}$ contains no cycles of length $(2s+1)$. The details of Part I have not been included since they consist of the analysis of 16 cases. A copy of this analysis is available from the third author.

Part II. To show that $G_{s,m}$ contains no $(2s+1)$-cycles we assume, to the contrary, that C is such a cycle and note that it cannot have the form ABA ($= AP_iBP_jA$). Thus, C contains a vertex of $K_{2s-3} - OO'$. Furthermore, C must contain at least one of the vertices A, B for otherwise it could have length $2s - 2$ at most.

If A and B are both in C, then necessarily one of the paths BAO or BAO' would be a subgraph of C. Now each of these paths is of length $2s$. But then, since (O, B) and (O', B) are not edges of $G_{s,m}$, we would have $|E(C)| > 2s + 1$, a contradiction.

Suppose, then, that A is a vertex of C but B is not. Now the path OAO', which has length $2s$, is a subgraph of C. But then, since (O, O') is not an edge of $G_{s,m}$, we would again have $|E(C)| > 2s + 1$. We conclude that A is not in C. But any cycle of length greater than 6 in $G_{s,m} - A$ contains at most two vertices of the type $(0, q)$, none of the type (p, q) with $p \geqslant 1$ and none from $K_r - O''$. Thus, C has order at most $1 + 2s - 3 + 2 = 2s$. □

Theorem 3. *For $k \geqslant 14$ and even,*

$$sat(n, C_k) \leqslant \left(1 + \frac{4}{k-2}\right)n + O(k^3).$$

Proof. Throughout the proof, s will be even and at least 6 and $n \geqslant 103$.

Define $m \geqslant 2$ and r by $n = ms^2 + s + r + 1$ and $2s \leqslant r < s^2 + 2s$. The graph $G_{s,m}$, pictured in Fig. 8 with $m = 2$, is constructed as follows. Let $H = \bar{C}_s$ (which is connected since $s \geqslant 6$), with vertices labeled $(s,1), (s,2), \ldots, (s,s)$ so that (s,i) is not adjacent to $(s, i+1)$, $1 \leqslant i \leqslant s$ (indices are read modulo s). To each vertex (s,i) of H attach a complete graph K_{r_i} where the r_i's are as nearly equal as possible subject to $3 \leqslant r_i \leqslant s+3$ and $\sum_{i=1}^{s}(r_i - 1) = r$ (the K_{r_i} are not shown in Fig. 8). For $1 \leqslant j \leqslant m$, let $C_{s,j}$ be a cycle of length s with vertices labeled $(0,i)_j$ so that $(0,i)_j$ is adjacent to $(0, i+1)_j$, $1 \leqslant i \leqslant s$. Now construct disjoint $(0,i)_j - (s,i)$ paths of length s. Finally, append a vertex O and all edges $O(s/2, i)_j$, $1 \leqslant i \leqslant s$ and $1 \leqslant j \leqslant m$. It is easy to see that this graph has order $n = ms^2 + s + r + 1$ and size

$$m(s^2 + 2s) + \binom{s}{2} - s + \sum_{i=1}^{s}\binom{r_i - 1}{2} = \left(1 + \frac{2}{s}\right)n + O(s^3).$$

Now define $m \geqslant 2$ and r by $n = ms^2 + (m+1)s + r + 1$ and $2s \leqslant r < s^2 + 2s$. The graph $G'_{s,m}$ is the same as $G_{s,m}$ except that the vertices of H are labeled $(s+1, 1)$, $(s+1, 2), \ldots, (s+1, s)$ and the $(0,i)_j - (s,i)$ paths of length s are replaced by $(0,i)_j - (s+1, i)$ paths of length $s+1$. Thus, $G'_{s,m}$ has order

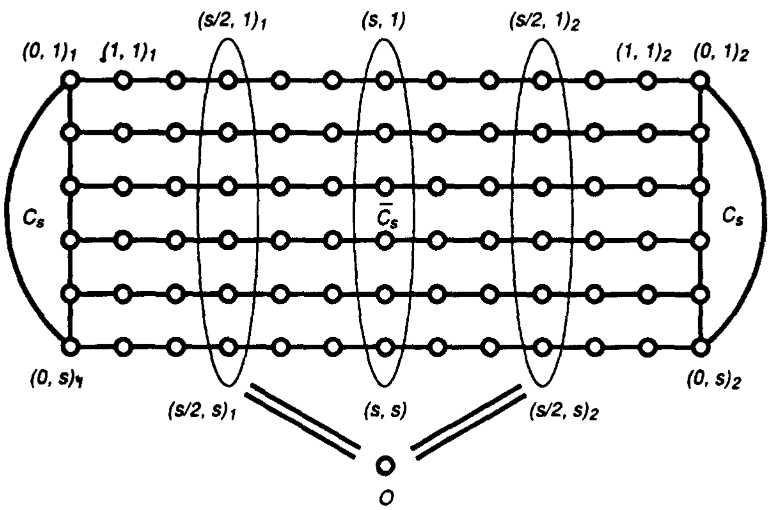


Fig. 8. The graph $G_{s,m}$; a C_k-saturated graph, $k = 2s + 2 \geqslant 14$.

$n = ms^2 + (m+1)s + r + 1$ and size

$$m(s^2 + 2s) + \binom{s}{2} - s + \sum_{i=1}^{s} \binom{r_i - 1}{2} = \left(1 + \frac{2}{s+1}\right)n + O(s^3).$$

In Part I we construct paths of length $2s+1$, $(2s+3)$ between all pairs of nonadjacent vertices in $G_{s,m}$, ($G'_{s,m}$, respectively). In Part II we will show that $G_{s,m}$, ($G'_{s,m}$) contains no $2s+2$, ($2s+4$, respectively)-cycles. As with Theorem 2, the details of Part I are not included but are available from the third author.

Part II. To show that there are no $2s+2$, $(2s+4)$-cycles in $G_{s,m}$, ($G'_{s,m}$, respectively), we note that any such cycle in either graph would necessarily contain vertices u, v where $u \in C_{s,a}$ for some a and $v \in \bar{C}_s$. But any cycle containing a vertex $u \in C_{s,a}$ necessarily contains an edge e in $C_{s,a}$ and, since $d_{G_{s,a}}(u,v) \geqslant s$ and $d_{G'_{s,a}}(u,v) \geqslant s+1$ there are at least $2s+1$, $(2s+3)$ edges of $E(G_{s,m}) - E(\bar{C}_s)$, ($E(G'_{s,m}) - E(\bar{C}_s)$ respectively) that are also in the cycle. Hence, every cycle of length $2s+2$, $(2s+4)$ in $G_{s,m}$, ($G'_{s,m}$, respectively) that contains vertices u and v must contain exactly one edge from each of $G_{s,a}$ and $\bar{C}_s$. But clearly this is impossible. Thus, any cycle containing u and v has length strictly greater than $2s+2$, $(2s+4)$ in $G_{s,m}$ ($G'_{s,m}$), respectively). □

Theorem 4. *For* $n \geqslant 4m \geqslant 12$ *and* m *odd* $sat(n, C_{4m}) \leqslant ((10m-1)/(8m-2))n + 2m$.

Proof. Recall the graph J_3 pictured in Fig. 5. In proving Proposition 4 it was shown that the graph G_s obtained by identifying the **0** vertices of $s \geqslant 1$ copies of J_3 and attaching one leaf to each vertex was C_{12}-saturated, had order $22s+1$ and size $29s$. There is a natural generalization, J_m, m odd, of J_3 where the 'inner triangle' now is a cycle C of length m each of whose vertices lies in one of the m copies of the $K_{1,3}$ subgraphs of J_m (J_5 is pictured in [6]). The structure of the cycle on the 'outer' two vertices of each of the J_3's is the same as in J_3: two consecutive, skip two, two consecutive, skip two, etc. These are the 'snarks' of Isaacs [10] who had shown them to be nonhamiltonian, i.e., they contain no $4m$-cycles.

Let $G_{s,m}$, $s \geqslant 1$, $m \geqslant 5$ and odd, be the graph obtained by identifying the **0** vertices of $s \geqslant 1$ copies of J_m and attaching one leaf to each vertex of $G_{s,m}$ different from **0**. Then $G_{s,m}$ has order $n = 2(4m-1)s + 1$ and size $(6m + (4m-1))s = ((10m-1)/(8m-2))(n-1)$. By deleting r leaves, $0 \leqslant r \leqslant 8m-3$, we obtain a graph of order $n = 2(4m-1)s + 1 - r$ and size

$$(10m-1)s - r = \frac{10m-1}{8m-2}(n-1) + \frac{2m+1}{8m-2}r < \frac{10m-1}{8m-2}n + 2m.$$

It remains to show that each two nonadjacent vertices of $G_{s,m}$ are joined by a path of length $4m-1$. Our proof of this fact is long (12 pages) and so will not be included here; full details can be found in [13]. We will, however, outline the proof.

The proof is inductive on m beginning at $m = 5$. In [6] it was shown that every two nonadjacent vertices of J_m are joined by a path of length $4m - 1$; in [13] it is shown, through consideration of seven cases, that this is true also for paths of lengths $4m - 2$ and $4m - 3$. To complete the proof that each two nonadjacent vertices of $G_{s,m}$ are joined by a path of length $4m - 1$ it now suffices to show that if the $\mathbf{0}$ vertices of two copies of J_m are identified then each of two nonadjacent vertices of the resulting graph are joined by paths of lengths $2m$, $2m - 1$ and $2m - 2$. This was accomplished by analysis of 24 cases. □

Finally, we consider the situation in which $k = n$. Then we are considering graphs that are not hamiltonian but become so upon the addition of any new edge. Minimum graphs of order n with this property were shown, [6], to have size $3n/2$ for all even $n \geqslant 36$ and, [5, 7], to have size $(3n + 1)/2$ for all odd $n \geqslant 53$. Combining these results we obtain

Theorem B. Clarke et al. [5–7]. $sat(n, C_n) = \lfloor (3n + 1)/2 \rfloor$, $n \geqslant 53$.

5. Concluding remarks

Although it may not be evident from the results, we made extensive use of computers, mainly to check our constructions. The computer also assisted in the proof of Theorem 1: linear programming showed that using the inequalities developed in the proof, the lower bound $n(1 + 1/(2k + 8))$ is close to the optimal value for various values of n and k. Actually, the experimental optimum seems to be $n(1 + 1/(2k + 3))$ from our computational evidence.

The computer also helped to discard some of our constructions. It is interesting to note that the computer check of our constructions for 14- and 16-cycle saturated graphs took hours. To find these graphs — the smallest members in their infinite families — with a computer search is impossible with any computer.

We were not satisfied with our construction for 6-cycle saturated graphs with $3\lfloor n/2 \rfloor$ edges. Therefore, we conducted a computer search for C_6-builders with up to 8 vertices but did not find anything better.

We think that some of our graphs are the extremal graphs (for $k = 5, 6, 7$) and that our general constructions, if not optimal, still are very close to being optimal and show some structural features of the extremal graphs.

The values of $sat(n, C_4) = \lfloor (3n - 5)/2 \rfloor$ and $sat(n, C_n) = \lfloor (3n + 1)/2 \rfloor$ together with the constructions showing $sat(n, C_k) < 3n/2$ for other values of k suggest the following problems.

Problem 1. Determine if $sat(n, C_k)$ is a convex function of k, $k > 3$, for fixed n or is convex at least when the parity of k is fixed.

If this is the case then

Problem 2. Determine the value of k which minimizes $sat(n, C_k)$ for fixed n.

On the other hand, if the answer to Problem 1 is in the negative, then the dual question arises: which k maximizes $sat(n, C_k)$? Finally,

Problem 3. Is $\limsup_n sat(n, C_k)/n$ a decreasing function of k, at least for odd k and even k, respectively?

References

[1] B. Bollobás, On a conjecture of Erdős, Hajnal and Moon, Amer. Math. Monthly 74 (1967) 178–179.
[2] B. Bollobás, Determination of extremal graphs by using weights, Wiss. Z. Techn. Hochsch. Ilmenau 13 (1967) 419–421.
[3] B. Bollobás, Extremal Graph Theory (Academic Press, London, 1978).
[4] F.C. Bussemaker, S. Cobeljic, D.M. Cvetkovic and J.J. Seidel, Computer investigation of cubic graphs, Technological University Eindhoven, Eindhoven, 1976.
[5] L.H. Clark, R.P. Crane, R.C. Entringer and H.D. Shapiro, On smallest maximally nonhamiltonian graphs, Congress. Numer. 53 (1986) 215–220.
[6] L.H. Clark and R.C. Entringer, Smallest maximally nonhamiltonian graphs, Period. Math. Hungar. 15 (1983) 57–68.
[7] L.H. Clark, R.C. Entringer and H.D. Shapiro, Smallest maximally nonhamiltonian graphs II, Graphs Combin. 8 (1992) 225–231.
[8] P. Erdős, Z. Füredi and Z. Tuza, Saturated r-uniform hypergraphs, Discrete Math., to appear.
[9] P. Erdős, A. Hajnal and J.W. Moon, A problem in graph theory, Amer. Math. Monthly 71 (1964) 1107–1110.
[10] R. Isaacs, Infinite families of nontrivial trivalent graphs which are not Tait colorable, Amer. Math. Monthly 82 (1975) 221–239.
[11] L. Kászonyi and Z. Tuza, Saturated graphs with minimal number of edges, J. Graph Theory 10 (1986) 203–210.
[12] L.T. Ollmann, $K_{2,2}$ saturated graphs with a minimal number of edges, In: F. Hoffman, R.B. Levow and R.S.D. Thomas, eds., Proc. 3rd Southeastern Conference on Combinatorics, Graph Theory and Computing, Boca Raton, Florida (1972) 367–392.
[13] T.D. Porter, Some results in combinatorial analysis and graph theory, Ph.D. Dissertation, University of New Mexico, July 1990.
[14] M. Trusczynski and Z. Tuza, Asymptotic results on saturated graphs, Discrete Math. 87 (1991) 309–314.
[15] Z. Tuza, Extremal problems on saturated graphs and hypergraphs, Ars Combin. 25B (1988) 105–113.
[16] Z. Tuza, C_4-saturated graphs of minimum size, Acta Univ. Carolin. — Math. Phys. 30 (1989) 161–167.
[17] W. Wessel, Über eine Klasse paarer Graphen. I. Beweis einer Vermutung von Erdős Hajnal und Moon, Wiss. Z. Techn. Hochsch. Ilmenau 12 (1966) 253–256.
[18] W. Wessel, Über eine Klasse paarer Graphen. II. Bestimmung der Minimalgraphen, Wiss. Z. Techn. Hochsch. Ilmenau 13 (1967) 423–426.