Merrifield-Simmons index and minimum Number of Independent Sets in Short Trees

Allan Frendrup (frendrup@math.aau.dk)
Dept. of Mathematics, Aalborg University,
Fredrik Bajers Vej 7 G, 9220 Aalborg øst, Denmark

Anders Sune Pedersen (asp@imada.sdu.dk)
Dept. of Mathematics & Computer Science, University of Southern Denmark,
Campusvej 55, 5230 Odense M, Denmark

Alexander A. Sapozhenko (sapozhenko@mail.ru)
Lomonosov University of Moscow
Faculty of Computational Mathematics and Cybernetics
Leninskie Gory, 119992 Moscow, Russia

Preben Dahl Vestergaard (pdv@math.aau.dk)
Dept. of Mathematics, Aalborg University,
Fredrik Bajers Vej 7 G, 9220 Aalborg øst, Denmark

MR Subject Classification: 05C69, 05C05

April 1, 2008

Abstract

In Ars Comb. 84 (2007), 85–96, Pedersen and Vestergaard posed the problem of determining a lower bound for the number of independent sets in a tree of fixed order and diameter d. Asymptotically, we give here a complete solution for trees of diameter $d \leq 5$. The lower bound is $5^{n/3}$ and we give the structure of the extremal trees. A generalization to connected graphs is stated.

1 Introduction

Half a century ago authors counted maximal independent sets in a graph ([7, 8]) and the first results on the number of independent subsets of a graph appeared in [11, 2, 3], here $i(G)$ was called the Fibonacci number of G. In chemical literature $i(G)$ is called the Merrifield-Simmons index. It is treated in a monograph ([6]) and in a wealth of later papers ([1, 17, 16, 15, 14, 13, 12]).

In [10] several upper and lower bounds for $i(G)$ were presented in terms of order, size or independence number and also bounds for $i(G)$ in trees and in unicyclic graphs were obtained.

Denoting n-order trees with diameter d by $T(n,d)$, we have that

$$i(T(n,d)) \leq \text{fib}(d) + 2^{n-d}\,\text{fib}(d+1) \tag{1}$$

[9, Th. 3.1], [5, Th. 1].

Formula (1) gives a tight upper bound for the number of independent sets in a tree in terms of its diameter and order, in [9] we also determined the trees for which that upper bound is attained. In the same paper we posed the problem of determining the corresponding lower bound in terms of diameter and order, and asked for a characterization of the trees for which the lower bound is attained. This is for sufficiently large orders done here for diameters four and five. Asymptotically the number of independent sets in n-order trees of diameter five turns out to be $5^{n/3}$ (Corollary 3). The results for diameter three and four are also given in a recent paper [4].

2 Notation

All graphs will be assumed simple and finite. A vertex of degree one is called a *leaf* and its unique neighbour is called a *stem*. In a graph G the set of vertices which are neigbours to a vertex $v \in V(G)$ is denoted by $N_G(v)$ and by $N(v)$ if the graph G is obvious from context. The set of vertices consisting of the vertex v and all its neighbours is denoted by $N[v]$, i.e. $N[v] = \{v\} \cup N(v)$. Let $S \subseteq V(G)$, then $N(S)$ denotes the set of vertices in $V(G)$ having a neighbour in S and $N[S] = S \cup N(S)$. For a set S of vertices, $S \subseteq V(G)$ we let $G - S$ denote the graph obtained from G by deleting from G all vertices of S and all edges incident with a vertex of S.

Given a graph G, a subset S of $V(G)$ is said to be independent, if no two vertices of S are adjacent in G, in particular, the empty set is considered to be an independent set of any graph. The number of independent sets in a graph G is denoted by $i(G)$.

We shall often consider some tree T of a given diameter d and order n such that $i(T)$ is minimum. By this we mean that T is a tree of diameter d and order n such that no other tree T' of diameter d and order n contains fewer independent sets than T does.

3 Helpful results

In this section we state some basic observations and results, which will be helpful for characterizing trees of a given diameter and order which contain the fewest possible number of independent sets.

Observation 1. Let G be a graph and let $v \in V(G)$ and $e = uz \in E(G)$. Then
(i) $i(G) = i(G-v) + i(G-N[v])$
(ii) $i(G-e) = i(G) + i(G-N[\{u,z\}])$
(iii) $i(G) = i(G-\{u,z\}) + i(G-N[u]) + i(G-N[z])$

Observation 2. If H is a induced subgraph of G then $i(H) \leq i(G)$ and equality holds if and only if $G \cong H$. If H is a spanning subgraph of G then $i(H) \geq i(G)$ and equality holds if and only if $H \cong G$.

Lemma 1. Let G be a graph containing two leaves l_1 and l_2 such that $d(l_1, l_2) \leq 3$ and let s_i denote the stem adjacent to l_i for $i \in \{1, 2\}$. If $G' := G - s_2l_2 + l_1l_2$ then $i(G') \leq i(G)$ and if equality holds then either
(i) $d(l_1, l_2) = 2, s_1 = s_2$ and $N_G(s_1) = \{l_1, l_2\}$, i.e., in G the three vertices s_1, l_1, l_2 span a P_3 as a component or
(ii) $d(l_1, l_2) = 3, s_1 \neq s_2$ and $N_G(s_2) = \{s_1, l_2\}$.

Proof. Observe that $G - l_2s_2 = G' - l_1l_2$ so by Observation 1(ii)

$$\begin{aligned} i(G) &= i(G') - i(G - N_G[\{l_2, s_2\}]) + i(G' - N_{G'}[\{l_1, l_2\}]) \qquad (2) \\ &= i(G') - i(G - N_G[s_2]) + i(G' - N_{G'}[l_1]). \end{aligned}$$

Since $G - N_G[s_2] \cong (G' - N_{G'}[l_1]) - N_{G'}(s_2)$ the graph $G - N_G[s_2]$ is an induced subgraph of $G' - N_{G'}[l_1]$. Therefore $i(G' - N_{G'}[l_1]) - i(G - N_G[s_2]) \geq 0$ and hence we have that $i(G) \geq i(G')$. If $i(G) = i(G')$ then $i(G' - N_{G'}[l_1]) - i(G - N_G[s_2]) = 0$ and for $d(l_1, l_2) = 2$, i.e., $s_1 = s_2$, we have $N_G(s_1) = \{l_1, l_2\}$ while for $d(l_1, l_2) = 3$, i.e., $s_1 \neq s_2$, we have that $N_G(s_2) = \{s_1, l_2\}$. This proves Lemma 1. □

Lemma 2. Let T be a tree of diameter $d \geq 4$ and order n such that $i(T)$ is minimum. Then no vertex in T is adjacent to more than two leaves, and if a vertex is adjacent to two leaves then it is penultimate on a diametrical path of G.

Proof. Assume that a vertex v is adjacent to two leaves l_1, l_2. By Lemma 1 it follows that v is the second vertex on a diametrical path $v_1, v_2 = v, v_3, v_4, \ldots, v_{d+1}$. Otherwise the graph $G' = G - vl_2 + l_1l_2$ would have $i(G') < i(G)$ and the diameter of G' would not be larger than that of G. So only the penultimate vertex of a diametrical path can support multiple leaves. We shall prove that v can support at most two leaves. Let $L := \{l_1, \ldots, l_k\}$, $k \geq 3$, be the leaves adjacent to v and consider the tree $T' := T - \{l_1v, l_2v\} + \{v_3l_1, l_1l_2\}$. Let C be the component of $T - v_2v_3$ containing v_3 then

$$\begin{aligned} i(T) &= i(C - v_3)(2^{|L|} + 1) + i(C - N[v_3])2^{|L|} \\ &> 3i(C - v_3)(2^{|L|-2} + 1) + i(C - N[v_3])2^{|L|-1} = i(T'). \end{aligned}$$

Since T' is a tree with diameter d and order n, we have a contradiction with the minimality of $i(T)$. Thus, v is not adjacent to more than two leaves. $\square$

Lemma 3. Let H be a graph with a vertex v. Let $G_1, \ldots, G_k$, $k \geq 7$, be copies of K_2 and let $v_i \in G_i$. If $G = H \cup G_1 \cup \cdots \cup G_k + \{vv_1, \ldots, vv_k\}$ and $G' = H \cup G_1 \cup \cdots \cup G_{k-1} + \{x, y\} + \{vv_1, \ldots, vv_{k-1}, v_1x, v_2y\}$, then $i(G') < i(G)$.

Proof. By considering G and G' we observe that $i(G) = 3^k i(H - v) + 2^k i(H - N[v])$ and $i(G') = 25 \cdot 3^{k-3} i(H - v) + 2^{k+1} i(H - N[v])$. Thus, since $k \geq 7$, we obtain

$$i(G) - i(G') = 2 \cdot 3^{k-3} i(H - v) - 2^k i(H - N[v]) \geq (2 \cdot 3^{k-3} - 2^k) i(H - N[v]) > 0.$$

$\square$

4 Trees of diameter three

For trees of diameter three the problem is straightforward. For completeness we describe the trees T of diameter three for which $i(T)$ is minimum.

Proposition 1. Given any fixed $n \geq 4$, let T denote a tree of diameter three and order n for which the number of independent sets is minimum. Let $P_4 : x_0x_1x_2x_3$ denote a diametrical path of T. Then

$$\{\deg(x_1), \deg(x_2)\} = \left\{ \left\lfloor \frac{n-2}{2} \right\rfloor, \left\lceil \frac{n-2}{2} \right\rceil \right\}.$$

5 Trees of diameter four

Let G_{2k+2}, $k \geq 2$, be the graph obtained from $K_{1,k+1}$ by subdividing k of its edges. Consider a tree T with diameter 4 and order n such that $i(T)$ is minimum. Let $v_1, \ldots, v_5$ be a diametrical path in T. If $n \geq 7$ it follows from Lemma 1 and Lemma 2 that $T \cong G_n$ or that each component of $T - v_3$ is isomorphic to K_2 or P_3. If $T \not\cong G_n$ then let $s(T)$ and $t(T)$ denote the number of components from $T - v_3$ isomorphic to K_2 and P_3, respectively. Then $n = 1 + 2s + 3t$ and $i(T) = 2^{s(T)} 4^{t(T)} + 3^{s(T)} 5^{t(T)}$.

Theorem 1. Let T_n be a tree of diameter four and order n for which the parameter $i(T_n)$ attains its minimum value and let $v_1, \ldots, v_5$ be a diametrical path in T_n. Then $T_5 = P_5$, $T_6 = G_6$ and if $n \geq 7$ then each component of $T_n - v_3$ is isomorphic to K_2 or P_3 and

- $s(T_n)$ is as indicated in the following table when $7 \leq n \leq 25$.
- $s(T_n) = 2n + 1 \mod 3$ for $n \geq 26$.

n	7	8	9	10	11	12	13	14	15	16
$s(T_n)$	3	2	4	3	5	4	6	5	4	3

n	17	18	19	20	21	22	23	24	25
$s(T_n)$	5	4	3	2	4	3	2	1	3

Proof. The theorem is easily verified for $n \leq 6$. Thus, we may assume $n \geq 7$. By considering G_n (if n is even) it easily follows that the graph T' obtained from G_n by removing the leaf adjacent to the center vertex and attaching a second leaf to another stem satisfies $i(T') < i(G_n)$. Thus we may assume that $T_n \not\cong G_n$ and only $s(T_n)$ has to be determined.

Now consider trees T' and T'' with the same structure as T, i.e., having diameter four and such that all components obtained by deletion of the central vertex are K_2's or P_3's. Assume further that that $s(T'') = s(T') - 3 \geq 0$ and $t(T'') = t(T') + 2$. If $s' := s(T')$ and $t' := t(T'')$ then

$$i(T') - i(T'') = \frac{2}{27} 3^{s'} 5^{t'} - 2^{s'+2t'}.$$

It follows that

$$i(T') - i(T'') \geq 0 \Leftrightarrow \left(\frac{3}{2}\right)^{s'} \left(\frac{5}{4}\right)^{t'} \geq \frac{27}{2} \Leftrightarrow s' \log 3/2 + t' \log 5/4 \geq \log 27/2.$$

Since $n' := |V(T')| = 1 + 2s' + 3t'$ we may obtain that $i(T') - i(T'') \geq 0$ if and only if $s' \geq a - bn'$ for real numbers a and b, $a = \frac{\log(27/2)+(1/3)\log(5/4)}{\log(3/2)-(2/3)\log(5/4)}$, $b = \frac{(1/3)\log(5/4)}{\log(3/2)-(2/3)\log(5/4)}$, ($a \approx 10,429$ and $b \approx 0,2898$).

It follows that if k is the largest integer such that $k \leq a - bn$ and $n = 1 + 2k + 3t$ for some integer $t \geq 0$ then $s(T_n) = k$ if and only if $k \geq 3$. Using these observations, it is straightforward to derive the values of $s(T_n)$ for $n \leq 25$. For $n \geq 26$ the inequality $s' \leq a - bn$ implies that $s' < 3$ and therefore $s(T_n) \leq 2$. By the equation $n = 1+2s+3t$ we obtain that $2s(T_n) \equiv n - 1 (\text{mod } 3)$ and the statement is obtained since this implies that $s(T_n) \equiv -2s(T_n) \equiv 1 - n \equiv 2n + 1 (\text{mod } 3)$. □

6 Trees of diameter five

In order to describe trees of diameter five with minimum number of independent sets we introduce the following terminology.

Let T denote a tree of diameter five with a diametrical path $P_6 : x_0x_1x_2x_3x_4x_5$. If there is exactly one leaf attached to $\{x_2, x_3\}$, then we refer to T as a *center-leaf tree*, and if there is no leaf attached to $\{x_2, x_3\}$, then we refer to T as a *center-leaf-free tree*.

Let T denote a center-leaf-free tree. If every component of $T - \{x_2, x_3\}$ is a $K_{1,1}$ then T is referred to as a center-leaf-free $K_{1,1}$-tree. If every component of $T - \{x_2, x_3\}$ is a $K_{1,2}$, then T is referred to as a center-leaf-free $K_{1,2}$-tree. If every component of $T - \{x_2, x_3\}$ is a $K_{1,1}$ or a $K_{1,2}$, then T is referred to as a center-leaf-free mixed-$K_{1,1}$-$K_{1,2}$-tree.

6.1 Some lemmas concerning trees of diameter five

In the following we prove some results needed for the characterization of trees of diameter five with minimum number of independent sets.

Lemma 4. Let T be a tree of diameter five for which $i(T)$ is minimum, and let $P_6 : x_0x_1x_2x_3x_4x_5$ denote a diametrical path of T. Then

(1) The neighbourhood $N[x_2, x_3]$ contains at most one leaf,

(2) if there is a leaf l attached to either x_2 or x_3, then every component of $T - \{x_2, x_3, l\}$ is a $K_{1,1}$, and

(3) if neither x_2 nor x_3 has a leaf attached, then every component of $T - \{x_2, x_3\}$ is a $K_{1,1}$ or a $K_{1,2}$.

Proof. Statement (1) follows from Lemma 1, while statement (3) follows from Lemma 2. To prove statement (2), we may assume that a leaf l is adjacent to x_2. From Lemma 1 it follows that all vertices from $N(x_2)\backslash\{x_3\}$ have degree at most two in T. Thus we may assume that a vertex $y \in N(x_3)\backslash\{x_2\}$ has degree at least three in T. By Lemma 2 y has degree exactly three.

Let l', x be a the two leaves adjacent to y and consider the tree $T' := T - yl' + ll'$. Observe that an independent set S in T' is independent in T unless $\{l', y\} \subsetneq S$. An independent set in T containing both l and l' is not independent in T'. Therefore

$$i(T) = i(T') - i(T' - N_{T'}[l', y]) + i(T - N_T[l, l']). \qquad (3)$$

One component of $T' - N_{T'}[l', y]$ is A, the subdivided star with center x_2 and the other components are a collection B of $K_{1,1}$'s and $K_{1,2}$'s, each joined by a (deleted) edge to x_3. Compare this to $T - N_T[\{l, l'\}]$ which as components has $A - x_2$ and the isolated vertex x in one group (corresponding to A) and one further component $B \cup \{x_3\}$. We see by Observation 2 that $i(T' - N_{T'}[l', y]) < i(T - N_T[l, l'])$. Thus (3) implies $i(T) > i(T')$ which contradicts the choice of T. □

Lemma 4 states that given an integer $n \geq 6$, the minimum value of $i(T)$ over all trees of order n and diameter five is attained for a center-leaf $K_{1,1}$-tree or a center-leaf-free mixed-$K_{1,1}$-$K_{1,2}$-tree. Given any center-leaf-free mixed-$K_{1,1}$-$K_{1,2}$-tree T, we let $p(T)$ and $q(T)$ denote the number of $K_{1,1}$'s attached to x_2 and x_3, respectively, and let $r(T)$ and $s(T)$ denote the number of $K_{1,2}$'s attached, by their center vertex, to x_2 and x_3, respectively. Whenever the context is clear we will simply write p, q, r and s for $p(T), q(T), r(T)$ and $s(T)$. For the number of independent sets in a center-leaf-free mixed-$K_{1,1}$-$K_{1,2}$-tree T we can use Observation 1(iii) to obtain Proposition 2 below.

Proposition 2. The number of independent sets in any center-leaf-free mixed-$K_{1,1}$-$K_{1,2}$-tree T is

$$3^{p+q}5^{r+s} + 2^p3^q4^r5^s + 2^q3^p5^r4^s. \qquad (4)$$

Lemma 5. Let T be a tree of diameter five for which $i(T)$ is minimum. If there is a leaf attached to x_2 or x_3, then $n(T) \leq 27$.

Proof. Suppose there is a leaf attached to x_2 or x_3. Then it follows from Lemma 4, that T is a center-leaf $K_{1,1}$-tree. If p and q denote the number of K_2's attached to x_2 and x_3, respectively, then $p, q \leq 6$, according to Lemma 3. Since $n(T) = 3 + 2(p+q)$, the desired bound on $n(T)$ follows. □

Corollary 1. If $n \geq 28$, then a tree T of diameter five and order n for which $i(T)$ is minimum is a center-leaf-free mixed-$K_{1,1}$-$K_{1,2}$-tree.

Proof. As noted above in Lemma 4, the tree T is either a center-leaf $K_{1,1}$-tree or a center-leaf-free mixed-$K_{1,1}$-$K_{1,2}$-tree. Since $n \geq 28$, the claim follows from Lemma 5. □

Lemma 6. Let T be a tree of diameter five for which $i(T)$ is minimum, and let $P_6 : x_0x_1x_2x_3x_4x_5$ denote a diametrical path of T. Let r and s denote the number of $K_{1,2}$'s attached by their center vertex to x_2 and x_3, respectively. By symmetry, we may assume $r = s + c$ for some non-negative integer c. If $c \geq 1$ then $p \leq q$ and given values of p and q we have that c is the largest possible integer such that

$$c \leq \left\lfloor \frac{\log(5/4) + (q-p)\log(3/2)}{\log(5/4)} \right\rfloor.$$

Proof. Suppose that $c \geq 1$. Let T' denote the center-leaf-free mixed $K_{1,1}$-$K_{1,2}$-tree with $p(T') = p$, $q(T') = q$, $r(T') = r - 1 = s + c - 1$ and $s(T') = s + 1$. According to Observation 1(iii),

$$\begin{aligned} i(T) &= 3^{p+q}5^{r+s} + 2^p3^q4^r5^s + 2^q3^p5^r4^s \quad \text{and} \\ i(T') &= 3^{p+q}5^{r+s} + 2^p3^q4^{r-1}5^{s+1} + 2^q3^p5^{r-1}4^{s+1}. \end{aligned}$$

Thus $i(T) - i(T') = \frac{1}{5}2^q3^p5^r4^s - \frac{1}{4}2^p3^q4^r5^s$. Now consider the logarithm of the ratio of these two terms

$$\log\left(\frac{\frac{1}{5}2^q3^p5^r4^s}{\frac{1}{4}2^p3^q4^r5^s}\right) = \log\left(\frac{4}{5}\left(\frac{2}{3}\right)^{q-p}\left(\frac{5}{4}\right)^c\right) = \log\left(\frac{4}{5}\right) + (q-p)\log\left(\frac{2}{3}\right) + c\log\left(\frac{5}{4}\right).$$

This term is at most zero since $i(T) \leq i(T')$. This implies that $p \leq q$ and since $i(T)$ is minimum and we by hypothesis create a tree T' with larger $i(T')$ each time we move a $K_{1,2}$-component attached to x_2 over to x_3 we want c to be the largest possible integer such that

$$c \leq \frac{\log(5/4) + (q-p)\log(3/2)}{\log(5/4)}.$$

□

Analogously to Lemma 3 we can obtain Lemma 7.

Lemma 7. Let H be a graph with a vertex v. Let $G_1, \ldots, G_9$ be copies of $K_{1,2}$ and let v_i be the center vertex of G_i, $1 \leq i \leq 9$. Let F_1, F_2, F_3 be copies of $K_{1,1}$ and let f_i be a vertex in F_i, $1 \leq i \leq 3$. If $G = H \cup G_1 \cup \cdots \cup G_7 \cup F_1 \cup F_2 \cup F_3 + \{vv_1, \ldots, vv_7, vf_1, vf_2, vf_3\}$ and $G' = H \cup G_1 \cup \cdots \cup G_9 + \{vv_1, \ldots, vv_7, vv_8, vv_9\}$, then $i(G') < i(G)$.

Proof. $i(G)-i(G') = 2\cdot 5^7\cdot i(H-v)-2^{17}i(H-N[v]) > 0$ because $i(H-v) \geq i(H-N[v])$ and $5^7 > 2^{16}$. □

From Lemma 7 we obtain the following corollary.

Corollary 2. Let T be a tree of diameter five for which $i(T)$ is minimum. If $n \geq 88$ then $p \leq 2, q \leq 2$ and $|r-s| \leq 4$.

Proof. Assume that $p \geq 3$ or $q \geq 3$. Since $p \leq 6, q \leq 6$ by Lemma 3 and thus $|r-s| \leq 11$ by Lemma 6 the assumption $n \geq 88$ implies that either $p \geq 3$ and $r \geq 7$ or $q \geq 3$ and $s \geq 7$. Now Lemma 7 implies that $i(T)$ is not minimum.

By using Lemma 6 and $p \leq 2$, $q \leq 2$ we obtain that $|r-s| \leq 4$ and if $r > s$ then $p \leq q$. □

6.2 Main result for trees of diameter five

By using the results from Section 6.1 we obtain the main results for trees of diameter five.

Theorem 2. For any $n \geq 28$, a tree T of diameter five and order n for which $i(T)$ is minimum is a center-leaf-free mixed-$K_{1,1}$-$K_{1,2}$-tree with $r(T) = s(T) + c$ and $q(T) = p(T) + d$ for non-negative integers c and d. Moreover, $c \leq 11$ and $p(T), q(T) \leq 6$. If $n \geq 88$ then $c \leq 4$ and $p(T), q(T) \leq 2$.

Proof. The proof relies on the results of Section 6.1. According to Corollary 1, the tree T as described in the theorem is a center-leaf-free mixed-$K_{1,1}$-$K_{1,2}$-tree. The bounds on the parameters $p(T), q(T), r(T), s(T)$ follow from Lemma 3, Lemma 6 and Corollary 2. □

If T_n is a tree of diameter five and order n for which $i(T_n)$ is minimum, then it follows from the above theorem that as n increases the tree T_n will be an increasingly 'well-balanced' center-leaf-free mixed-$K_{1,1}$-$K_{1,2}$-tree, that is, the ratio of $r(T_n)$ and $s(T_n)$ will tend to one, and the ratio of $(p(T_n)+q(T_n))/n$ will be small.

Lemma 8. There is an integer n' such that if T_n is a tree of diameter five and order $n \geq n'$ for which $i(T_n)$ is minimum then $p(T_n)+q(T_n) \leq 2$.

Proof. Let T^n be any center-leaf-free mixed-$K_{1,1}$-$K_{1,2}$-tree of order n such that $p(T^n)+q(T^n) \geq 3$ and $p(T^n), q(T^n) \leq 6$. Consider a center-leaf-free mixed-$K_{1,1}$-$K_{1,2}$-tree T_2^n of order n with $p(T_2^n)+q(T_2^n) = p(T^n)+q(T^n)-3$ and $r(T_2^n)+s(T_2^n) = r(T^n)+s(T^n)+2$. From equation (4) it follows that

$$\lim_{n\to\infty} \frac{i(T^n)}{3^{p(T^n)+q(T^n)}5^{r(T^n)+s(T^n)}} = 1 \quad \text{and} \quad \lim_{n\to\infty} \frac{i(T_2^n)}{3^{p(T^n)+q(T^n)}5^{r(T^n)+s(T^n)}} = \frac{25}{27}.$$

Thus there must exist an integer n' such that $i(T^n) > i(T_2^n)$ when $n \geq n'$.This implies that T_n can not be the graph T^n for $n \geq n'$ and the statement follows. □

By using the result from Lemma 8 we can obtain the following characterization of T_n when $n \geq n'$.

Theorem 3. There is an integer n' such that if T_n is a tree of diameter five and order $n \geq n'$ for which $i(T_n)$ is minimum. Then T_n is a center-leaf-free mixed-$K_{1,1}$-$K_{1,2}$-tree and p, q, r and s is as indicated in the following table (it is assumed that $c := r - s \geq 0$ and $p \leq q$):

n mod 6	q	p	c
0	1	1	0
1	1	0	1
2	0	0	0
3	2	0	3
4	1	0	2
5	0	0	1

Proof. Let n' be the integer from Lemma 8 and consider T_n when $n \geq n'$. Then $p + q \leq 2$ and Lemma 6 implies that $c \leq 1$ if $p = q$, $c \in \{1, 2\}$ if $q - p = 1$ and $c \in \{3, 4\}$ if $q - p = 2$. By considering cases depending on $n \mod 6$ it can be observed that this determines the parameters p, q and c when $n \mod 3 \neq 0$. Further in the case $n \mod 6 = 0$ either $p = 1, q = 1$ and $c = 0$ or $p = 0, q = 2$ and $c = 4$ and in the case $n \mod 6 = 3$ either $p = 1, q = 1$ and $c = 1$ or $p = 0, q = 2$ and $c = 3$. In both cases we only have to compare the number of independent sets in the two trees that might be isomorphic to T_n and the result is as indicated in the table. □

It can be shown that the integer n' form Lemma 8 and Theorem 3 can be choosen to be smaller than one hundred.

From Theorem 3 we immediately obtain

Corollary 3. Asymptotically the minimum number of independent sets in n-order trees of diameter five is $5^{n/3}$.

7 The lower bound of i on graphs of fixed order and diameter

The following theorem gives an optimal bound for i for connected graphs of fixed order and diameter. The graph obtained by attaching a path P to a vertex v in a graph G is the graph $P \cup G + uv$ where u is a vertex of P with minimum degree. The Fibonacci numbers $fib(0), fib(1), \ldots$ is defined by the equations $fib(0) := 0$, $fib(1) := 1$ and $fib(n) := fib(n-1) + fib(n-2)$ for $n \geq 2$.

Theorem 4. If G is a connected graph of order n and diameter $d \geq 2$, then

$$2fib(d+1) + (n-d)fib(d) \leq i(G), \tag{5}$$

where equality occurs if and only if G is isomorphic to the graph obtained from K_{n-d+2} by removing an edge wv and attaching a path P_{d-2} at v (if $d \geq 3$).

Proof. If $G \cong P_{d+1}$ then the statement is true for G since $i(G) = fib(d+3) = 2fib(d+1) + (n-d)fib(d)$. Let G be a connected graph of order n and diameter d, $G \not\cong P_{d+1}$. Assume that the statement is true for each graph of order less than

n. Consider a diametrical path $P : v_1, \ldots, v_{d+1}$ in G. Since $G \ncong P_{d+1}$ there must be a vertex $u \notin V(P)$ such that $G - u$ is connected and since P is a diametrical path u can at most be adjacent to three vertices of P. Thus $G - u$ is a graph with diameter at least d and $G - N[u]$ has at least $d - 2$ vertices. By assumption we have that $i(G - u) \geq 2fib(d + 1) + (n - 1 - d)fib(d)$ and Observation 2 implies that $i(G - N[u]) \geq i(P_{d-2}) = fib(d)$. If equality holds in both inequalities then $G - N[u] \cong P_{d-2}$ and $G - u$ has diameter d and can be constructed as one of the graphs described in the statement. Thus if equality holds in both inequalities G must be one of the graphs described in the statement. By applying Observation 1 we obtain that

$$i(G) = i(G - u) + i(G - N[u]) \geq 2fib(d + 1) + (n - d)fib(d)$$

and equality occurs if and only if G is one of the graphs described in the statement. □

References

[1] Ivan Gutman. Extremal hexagonal chains. *J. Math. Chem.*, 12(1-4):197–210, 1993. ISSN 0259-9791. Applied graph theory and discrete mathematics in chemistry (Saskatoon, SK, 1991).

[2] Peter Kirschenhofer, Helmut Prodinger, and Robert F. Tichy. Fibonacci numbers of graphs. II. *Fibonacci Quart.*, 21(3):219–229, 1983. ISSN 0015-0517.

[3] Peter Kirschenhofer, Helmut Prodinger, and Robert F. Tichy. Fibonacci numbers of graphs. III. Planted plane trees. In *Fibonacci numbers and their applications (Patras, 1984)*, volume 28 of *Math. Appl.*, pages 105–120. Reidel, Dordrecht, 1986.

[4] Arnold Knopfmacher, Robert F. Tichy, Stephan Wagner, and Volker Ziegler. Graphs, partitions and Fibonacci numbers. *Discrete Appl. Math.*, 155(10):1175–1187, 2007. ISSN 0166-218X.

[5] Xueliang Li, Haixing Zhao, and Ivan Gutman. On the Merrifield-Simmons index of trees. *MATCH Commun. Math. Comput. Chem.*, 54(2):389–402, 2005. ISSN 0340-6253.

[6] Richard E. Merrifield and Howard E. Simmons. *Topological Methods in Chemistry.* Wiley-Interscience, 1989.

[7] R. E. Miller and D.E. Muller. A problem of maximum consistent subsets. Technical report, IBM research report RC-240, J.T. Watson Research Center, 1960.

[8] J. W. Moon and L. Moser. On cliques in graphs. *Israel J. Math.*, 3:23–28, 1965. ISSN 0021-2172.

[9] Anders Sune Pedersen and Preben Dahl Vestergaard. An upper bound on the number of independent sets in a tree. *Ars Comb.*, 84:85–96, 2007.

[10] Anders Sune Pedersen and Preben Dahl Vestergaard. Bounds on the number of vertex independent sets in a graph. *Taiwanese J. Math.*, 10(6):1575–1587, 2006. ISSN 1027-5487.

[11] Helmut Prodinger and Robert F. Tichy. Fibonacci numbers of graphs. *Fibonacci Quart.*, 20(1):16–21, 1982. ISSN 0015-0517.

[12] H. Wang and H. Hua. Unicycle graphs with extremal merrifield-simmons index. *J. of Math. Chem.*, 43(1):202–209, 2008.

[13] M. Wang, H. Hua, and D: Wang. The first and second largest merrifield-simmons indices of trees with prescribed pendent vertices. *Journal of Mathematical Chemistry*, 43(2):727–736, 2008.

[14] A. Yu and X. Lv. The merrifield-simmons indices and hosoya indices of trees with k pendent vertices. *J. Math. Chem.*, 41(1), 2007.

[15] Aimei Yu and Feng Tian. A kind of graphs with minimal Hosoya indices and maximal Merrifield-Simmons indices. *MATCH Commun. Math. Comput. Chem.*, 55(1):103–118, 2006. ISSN 0340-6253.

[16] Lian-Zhu Zhang and Feng Tian. Extremal catacondensed benzenoids. *J. Math. Chem.*, 34(1-2):111–122, 2003. ISSN 0259-9791.

[17] Lianzhu Zhang. The proof of gutman's conjectures concerning extremal hexagonal chains. *Journal of Systems Science and Mathematical Science*, 18(4):460–465, 1998.