On the number of independent subsets in trees with restricted degrees

Eric Ould Dadah Andriantiana and Stephan Wagner

Department of Mathematical Sciences
Mathematics Division
Stellenbosch University
Private Bag X1, Matieland 7602
South Africa

Abstract

We study the number of independent vertex subsets (known as the Merrifield-Simmons index in mathematical chemistry) and the number of independent edge subsets (called the Hosoya index) for trees whose vertex degrees are restricted to 1 or d (for some $d \geq 3$), a natural restriction in the chemical context. We find that the minimum of the Merrifield-Simmons index and the maximum of the Hosoya index are both attained for path-like trees; furthermore, one obtains the second-smallest value of the Merrifield-Simmons index and the second-largest value of the Hosoya index for generalized tripods. Analogous results are also found for a closely related parameter, the graph energy, that also plays an important rôle in mathematical chemistry.

Key words: independent sets, topological index, extremal trees, restricted degree
2000 MSC: 05C05, 05C35, 92E10

1 Introduction

The *Merrifield-Simmons index* $\sigma(G)$ [1], defined as the number of independent vertex subsets of a graph, and the *Hosoya index* $Z(G)$ [2], defined analogously as the number of matchings (independent edge subsets), certainly belong to the most popular topological indices in mathematical chemistry, whose main purpose is to predict physico-chemical properties of compounds from their

Email address: ericoda@sun.ac.za,swagner@sun.ac.za (Eric Ould Dadah Andriantiana and Stephan Wagner).

Preprint submitted to Mathematical and Computer Modelling 21 February 2010

structure (that can be modeled as a graph). A wealth of theoretical results on these two parameters has been obtained in recent years, in particular regarding trees and tree-like structures (such as unicyclic graphs [3–5] or bicyclic graphs [6,7]). Upper and lower bounds are known under various restrictions, such as diameter [8], number of leaves [9] or number of cut edges [10]. The interested reader is referred to [11] for a survey on this topic.

The Hosoya index is intimately related to another important parameter, namely the *energy*, defined as the sum of the absolute values of all eigenvalues of a graph [12]. For trees, this strong relation becomes clear from the representation as a *Coulson integral* (see [13]):

$$E(G) = \frac{2}{\pi} \int_0^\infty x^{-2} \log \left(\sum_{k \geq 0} m(G,k) x^{2k} \right) dx, \tag{1}$$

where $m(G,k)$ denotes the number of matchings of size k.

Degree restrictions are particularly natural in the chemical context; trees whose maximum degree is at most 4 are also known as *chemical trees* [14]. Maximum and minimum values of our two indices for trees with given maximum degree were determined in [15,16]. See also [17] for a related extremal problem concerning trees with given maximum degree. The present paper is devoted to another natural type of degree restriction: we consider trees whose vertex degrees are all either 1 or d; the set of all such trees will be denoted by $\mathbb{T}_{1,d}$. Note in particular that for $d = 4$, we obtain trees that represent saturated hydrocarbons (alkanes); in chemical practice, one usually removes all hydrogen atoms (corresponding to leaves in a tree) of a compound to obtain the associated chemical graph, but the results of this paper show that one actually obtains very similar behavior if these vertices are kept. The same class of trees has been investigated, for instance, in [18]. There, the trees in $\mathbb{T}_{1,d}$ that maximize the Wiener index are characterized.

Our main result is the characterization of those trees in $\mathbb{T}_{1,d}$ that maximize the Hosoya index and energy and minimize the Merrifield-Simmons index. It turns out that they coincide with those trees that have been found to maximize the Wiener index in the aforementioned paper [18]. The dual problem (maximum Merrifield-Simmons index, minimum Hosoya index and energy) has essentially been solved in [15]. The second-largest resp. second-smallest values and associated graphs are characterized as well; there is a striking similarity to the behavior observed for trees without any restrictions (compare [16,19]).

2 Preliminaries

As mentioned in the introduction, we will be concerned with the class $\mathbb{T}_{1,d}$ of trees whose vertex degrees are either 1 or d (where $d \geq 3$; the case $d = 2$ is essentially trivial, since $\mathbb{T}_{1,2}$ coincides with the set of all paths). Our aim is to determine the trees in this class that maximize the Hosoya index and energy and minimize the Merrifield-Simmons index. Write $\mu(G, x) = \sum_{k\geq 0} m(G, k)x^{2k}$ for the polynomial that occurs in (1). Clearly, the Hosoya index is just $Z(G) = \mu(G, 1)$. The following lemma is well known (see for instance [13]) and frequently used as a tool in this context.

Lemma 1. *Let v be a vertex in a graph G. Then we have*

$$\mu(G, x) = \mu(G - v, x) + x^2 \sum_{w\in N_G(v)} \mu(G - \{v, w\}, x)$$

and

$$\sigma(G) = \sigma(G - v) + \sigma(G - (\{v\} \cup N_G(v))).$$

If G and G' are two disjoint graphs, then we have

$$\mu(G \cup G', x) = \mu(G, x)\mu(G', x) \text{ and } \sigma(G \cup G') = \sigma(G)\sigma(G').$$

The trees described in the following definitions will be of particular interest.

Definition 1. For $n \geq 1$, let $c_1, \ldots, c_n$ be nonnegative integers. The tree which is obtained from a path $v_1, \ldots, v_n$ of length $n - 1$ by attaching c_i new leaves to v_i, for $1 \leq i \leq n$, is called $(c_1, \ldots, c_n)$-*caterpillar* (Fig. 1). In particular, we write C_n for the caterpillar with $c_1 = c_n = d - 1$ and $c_2 = \cdots = c_{n-1} = d - 2$, and C'_n for the caterpillar with $c_1 = d - 1$ and $c_2 = \cdots = c_n = d - 2$.

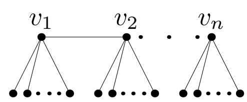

Fig. 1. A caterpillar

For convenience we define C_0 to be the path of length one and C'_0 the isolated vertex.

Note that C_n is an element of $\mathbb{T}_{1,d}$, but C'_n is not (since v_n has degree $d - 1$). Elements of $\mathbb{T}_{1,d}$ can be regarded as extensions of trees whose degrees are at most d (by adding an appropriate number of leaves to each vertex). In this regard, C_n corresponds to the path. Analogously, we define d-tripods as extensions of tripods, which were found to play an essential rôle in this context, see [16].

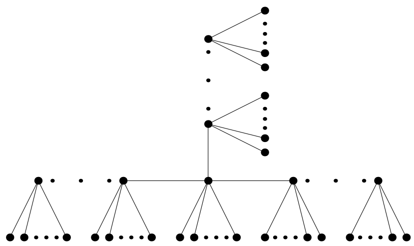

Fig. 2. A d-tripod

Definition 2. A *d-tripod*, denoted by $T_d(l, m, n)$, is an element of $\mathbb{T}_{1,d}$ that is constructed by attaching three branches C'_l, C'_m, C'_n (where $n \geq m \geq l \geq 1$) to a common vertex, see Fig. 2.

Remark 2. From Lemma 1, one can easily deduce the relations

$$\mu(C'_n, x) = (1 + (d-2)x^2)\mu(C'_{n-1}, x) + x^2\mu(C'_{n-2}, x), \tag{2}$$
$$\sigma(C'_n) = 2^{d-2}(\sigma(C'_{n-1}) + \sigma(C'_{n-1})). \tag{3}$$

Equation (2) shows that the sequence $\mu(C'_n, x)$ $(n \geq 0)$ satisfies a linear recurrence relation whose explicit solution is given by

$$\mu(C'_n, x) = \frac{X + x^2}{X - Y}X^n - \frac{Y + x^2}{X - Y}Y^n, \tag{4}$$

where

$$X = \frac{1 + (d-2)x^2 + \sqrt{1 + 2dx^2 + (d-2)^2x^4}}{2}$$

and

$$Y = \frac{1 + (d-2)x^2 - \sqrt{1 + 2dx^2 + (d-2)^2x^4}}{2}.$$

In a similar way, using the recurrence (3), one also obtains an explicit expression for $\sigma(C'_n)$:

$$\sigma(C'_n) = \frac{2X' + 1}{X' - Y'}X'^n - \frac{2Y' + 1}{X' - Y'}Y'^n, \tag{5}$$

where $X' = 2^{d-3} + \sqrt{2^{2d-6} + 2^{d-2}}$ and $Y' = 2^{d-3} - \sqrt{2^{2d-6} + 2^{d-2}}$. Next we need a simple yet crucial lemma:

Lemma 3. *For all nonnegative integers k and n such that $n \geq 3$ and real numbers $a \in (0, 1)$, the function defined on $I_n = \{0, 1, \ldots, \lfloor \frac{n-1}{2} \rfloor\}$ by*

$$\begin{aligned} f_{a,k} : & I_n \longrightarrow \mathbb{R} \\ & i \longmapsto a^i + (-1)^k a^{n-1-i} \end{aligned}$$

is decreasing on I_n.

Proof. For all $i \in I_n \backslash \{\lfloor \frac{n-1}{2} \rfloor\}$ we have

$$f_{a,k}(i+1) - f_{a,k}(i) = (a-1)(a^i - (-1)^k a^{n-i-2}) \text{ and } a^i > \left|(-1)^k a^{n-i-2}\right|.$$

Therefore $f_{a,k}(i+1) - f_{a,k}(i) < 0$ meaning that $f_{a,k}$ is decreasing on I_n. □

3 Intermediary results

Let G be a connected graph with a vertex v of degree $d-2$ that is not isomorphic to the star S_{d-1}. Let us label the vertices of degree d in C_n by $v_1, \ldots, v_n$. We denote by $C(n,k,G,v)$ the graph obtained by removing the $d-2$ leaves attached to v_k in C_n and identifying v_k with the vertex v of G (see Fig. 3). The next lemma provides information on the behavior of the number of independent vertex subsets and the energy of $C(n,k,G,v)$ as k varies. Similar results, considering a path P_n instead of C_n, can be found in [19] for the Merrifield-Simmons index and in [16] for the Hosoya index.

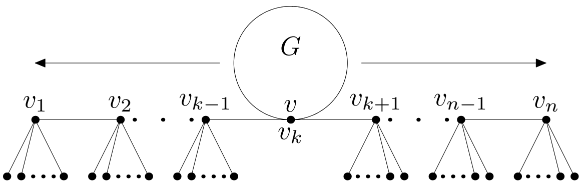

Fig. 3. $C(n,k,G,v)$

Lemma 4. *Let m be a nonnegative integer, $n = 4m+i, i \in \{1,2,3,4\}$, and let l be the integer part of $\frac{i-1}{2}$. Then for all real numbers $x > 0$ we have*

$$\mu(C(n,2,G,v),x) < \mu(C(n,4,G,v),x) < \cdots < \mu(C(n,2m+2l,G,v),x) < \mu(C(n,2m+1,G,v),x) < \cdots < \mu(C(n,3,G,v),x) < \mu(C(n,1,G,v),x)$$

and

$$\sigma(C(n,2,G,v)) > \sigma(C(n,4,G,v)) > \cdots > \sigma(C(n,2m+2l,G,v)) > \sigma(C(n,2m+1,G,v)) > \cdots > \sigma(C(n,3,G,v)) > \sigma(C(n,1,G,v)).$$

Proof. Using the first equation in Lemma 1 for $C(n,i+1,G,v)$, we find

$$\mu(C(n,i+1,G,v),x) = (\mu(C'_{i-1},x)\mu(C'_j,x) + \mu(C'_i,x)\mu(C'_{j-1},x))x^2\mu(G-v,x) + \mu(C'_i,x)\mu(C'_j,x)\left(\mu(G-v,x) + x^2\sum_{w\in N_G(v)}\mu(G-\{v,w\},x)\right).$$

where $j = n-1-i$.

Using the notations of Section 2, we have $XY = -x^2$. If we take $A = (X+x^2)/(X-Y)$ and $B = -(Y+x^2)/(X-Y)$, then $AB = (1-d)x^4/(X-Y)^2$ and Equation (4) yields

$$\begin{aligned}\mu(C'_i,x)\mu(C'_j,x) &= A^2X^{n-1} + B^2Y^{n-1} + (-1)^i\frac{(1-d)x^{2i}}{(X-Y)^2}x^4(X^{n-1-2i}+Y^{n-1-2i})\\ &= A^2X^{n-1} + B^2Y^{n-1}\\ &\quad + (-1)^i\frac{(1-d)X^{n-1}}{(X-Y)^2}x^4\left(\left(\frac{x^2}{X^2}\right)^i + (-1)^{n-1}\left(\frac{x^2}{X^2}\right)^j\right)\end{aligned}$$

$$\begin{aligned}\mu(C'_i,x)\mu(C'_{j-1},x) &= A^2X^{n-2} + B^2Y^{n-2}\\ &\quad + (-1)^i\frac{(1-d)x^2X^n}{(X-Y)^2}\left(\left(\frac{x^2}{X^2}\right)^{i+1} + (-1)^n\left(\frac{x^2}{X^2}\right)^j\right)\end{aligned}$$

$$\begin{aligned}\mu(C'_{i-1},x)\mu(C'_j,x) &= A^2X^{n-2} + B^2Y^{n-2}\\ &\quad + (-1)^{i-1}\frac{(1-d)x^2X^n}{(X-Y)^2}\left(\left(\frac{x^2}{X^2}\right)^{i} + (-1)^n\left(\frac{x^2}{X^2}\right)^{j+1}\right)\end{aligned}$$

$$\begin{aligned}\mu(C'_{i-1},x)\mu(C'_j,x) + \mu(C'_i,x)\mu(C'_{j-1},x) &= 2A^2X^{n-2} + 2B^2Y^{n-2}\\ &\quad + (-1)^i\frac{(1-d)x^2X^n(x^2-X^2)}{X^2(X-Y)^2}\left(\left(\frac{x^2}{X^2}\right)^{i} + (-1)^{n-1}\left(\frac{x^2}{X^2}\right)^{j}\right).\end{aligned}$$

Therefore, with $f_{\frac{x^2}{X^2},n-1}$ as in Lemma 3,

$$\mu(C(n,i+1,G,v),x) = D + E(-1)^i f_{\frac{x^2}{X^2},n-1}(i) \tag{6}$$

where

$$\begin{aligned}D &= (2A^2X^{n-2} + 2B^2Y^{n-2})x^2\mu(G-v,x)\\ &\quad + (A^2X^{n-1} + B^2Y^{n-1})\left(\mu(G-v,x) + x^2\sum_{w\in N_G(v)}\mu(G-\{v,w\},x)\right)\\ E &= \frac{(1-d)x^4X^n}{X(X-Y)^2}\left(\frac{x^2+X-X^2}{X}\mu(G-v,x) + x^2\sum_{w\in N_G(v)}\mu(G-\{v,w\},x)\right).\end{aligned}$$

As $X > |Y|$ and $A > |B|$, D is clearly positive. Since $|N_G(v)| = d-2$ by our assumptions, and $\mu(G-\{v,w\},x) \le \mu(G-v,x)$ for all $w \in N_G(v)$, we have

$$\sum_{w\in N_G(v)}\mu(G-\{v,w\},x) \le (d-2)\mu(G-v,x).$$

Equality can only hold if all neighbors of v are leaves, which has been excluded. Therefore we obtain

$$E > \frac{(1-d)x^4X^n}{X(X-Y)^2}\left(\frac{x^2+(1+(d-2)x^2)X-X^2}{X}\mu(G-v,x)\right) = 0.$$

Finally it becomes clear that Equation (6), together with Lemma 3 (it is easy to see that $\frac{x^2}{X^2} < 1$) implies the first part of the lemma.

For the case of the Merrifield-Simmons index, we can follow the same way to end up with

$$\sigma(C(n,i+1,G,v)) = F + K(-1)^{i-1}f_{\frac{2^{d-2}}{X'^2},n-1}(i), \tag{7}$$

where, with $X' = 2^{d-3} + \sqrt{2^{2d-6}+2^{d-2}}$, $Y' = 2^{d-3} - \sqrt{2^{2d-6}+2^{d-2}}$, $A' = (2X'+1)/(X'-Y')$ and $B' = -(2Y'+1)/(X'-Y')$,

$$\begin{aligned} F &= (A'^2X'^{n-1} + B'^2Y'^{n-1})\sigma(G-v) \\ &\quad + 2^{2d-4}(A'^2X'^{n-3} + B'^2Y'^{n-3})\sigma(G-(\{v\}\cup N_{G'}(v))), \\ K &= \frac{X'^{n-1}(2^{d-1}-1)}{(X'-Y')^2}\left(2^{d-2}\sigma(G-(\{v\}\cup N_G(v))) - \sigma(G-v)\right), \end{aligned}$$

are two positive constants. This proves the second part of the lemma. $\square$

4 Main Results

Lemma 4 can now be used to determine the extremal trees in $\mathbb{T}_{1,d}$, as follows:

Theorem 1. Let $d \geq 3$ and n be two nonnegative integers. The tree that maximizes the energy and minimizes the Merrifield-Simmons index among all elements of $\mathbb{T}_{1,d}$ of order $(d-1)n+2$ is the caterpillar C_n. For all real $x > 0$ we have

$$\begin{aligned} \mu(C_n,x) &= \frac{(X+x^2)^2}{X-Y}X^{n-1} - \frac{(Y+x^2)^2}{X-Y}Y^{n-1}, \\ \sigma(C_n) &= \frac{3X'+2^{d-2}+1}{X'-Y'}X'^n - \frac{3Y'+2^{d-2}+1}{X'-Y'}Y'^n, \end{aligned}$$

with X, Y, X', Y' as in equations (4) and (5).

Proof. Let T be an element of $\mathbb{T}_{1,d}$ whose order is $(d-1)n+2$. If T is not the caterpillar, then it has a branching vertex v with at least three branches that are not just single vertices and such that two of them are C'_i and C'_j for certain integers $i \geq j \geq 1$.

Now we apply Lemma 4, where G is taken to be the graph consisting of all branches of v other than C'_i and C'_j. Moving G to the end of the caterpillar formed by these two branches, we obtain a new tree T_1 that has strictly smaller Merrifield-Simmons index than T and satisfies $\mu(T, x) < \mu(T_1, x)$ for arbitrary $x > 0$. Together with the Coulson integral (1), this shows that T cannot be maximal with respect to the energy or minimal with respect to the Merrifield-Simmons index.

The two explicit expressions follow from Equations (4) and (5) using the relations $\mu(C_n, x) = \mu(C'_n, x) + x^2\mu(C'_{n-1}, x)$ and $\sigma(C_n) = \sigma(C'_n) + 2^{d-2}\sigma(C'_{n-1})$. □

Lemma 4 also allows us to go one step further and determine the extremal trees if C_n is excluded.

Theorem 2. For all integers $d \geq 3$ and for a given number of vertices $(d-1)n + 2$, the element of $\mathbb{T}_{1,d} \backslash \{C_n, n \in \mathbb{N}\}$ with minimum Merrifield-Simmons index and maximum energy is

- the d-tripod $T_d(1, 1, 1)$ if $n = 4$,
- the d-tripod $T_d(2, 1, 1)$ if $n = 5$,
- the d-tripod $T_d(2, 2, n-5)$ if $n > 5$.

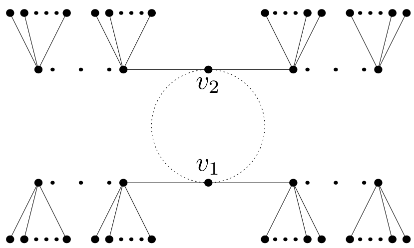

Fig. 4. A non d-tripod in $\mathbb{T}'_{1,d}$

Proof. The cases where $n < 4$ are not interesting because C_n is the only element of $\mathbb{T}_{1,d}$ in these cases, and the cases $n = 4$ and $n = 5$ are trivial as well.

Now we assume that $n > 5$. We start by showing that the minimizer of the σ-index or the maximizer of the energy in $\mathbb{T}'_{1,d} = \mathbb{T}_{1,d} \backslash \{C_n, n \in \mathbb{N}\}$ has to be a d-tripod. Let T be an element of $\mathbb{T}'_{1,d}$ that is not a d-tripod, then there are two possible cases:

- Case a: T consists of a vertex with more than three caterpillar branches that are not just single vertices. Then, applying Lemma 4 as in the proof of

Theorem 1, we can merge two branches to form a single caterpillar branch and get an element T_1 of $\mathbb{T}'_{1,d}$ such that $\sigma(T_1) < \sigma(T)$ and $\mu(T,x) < \mu(T_1,x)$ for all $x > 0$.

- Case b: T has at least two branching vertices with at least three branches each, as pictured in Fig. 4. Once again we can use Lemma 4 to merge two branches of v_1, thereby constructing an element of $\mathbb{T}'_{1,d}$ with $\sigma(T_1) < \sigma(T)$ and $\mu(T,x) < \mu(T_1,x)$ for all $x > 0$, as in the previous case.

Therefore such a tree T cannot minimize the Merrifield-Simmons index or maximize the energy, which leaves us with the set of d-tripods. Applying Lemma 4 once again, we find that $T_{2,2,n-5}$ is the d-tripod that minimizes the σ-index and maximizes the energy. □

Remark 5. In particular, as $\mu(G,1)$ coincides with the Hosoya index of G, the maximal trees in $\mathbb{T}_{1,d}$ with respect to the number of matchings are the same as the maximal trees that we found with respect to the energy.

5 Conclusion

The results of this paper show an interesting analogy between trees in the class $\mathbb{T}_{1,d}$ and general trees without degree restrictions. With a little bit more effort, one can certainly also determine the third-largest/smallest values and perhaps more, as in [16,19]. We also believe that our crucial Lemma 4 can be applied to the study of unicyclic graphs or other tree-like structures with similar degree restrictions. A natural problem for further study would be to consider trees with more general degree restrictions, such as prescribing the entire degree sequence. This might be a very hard problem, but partial results can possibly be obtained along the lines of this paper.

Acknowledgment

This material is based upon work supported financially by the National Research Foundation of South Africa under grant number 70560.

References

[1] R. E. Merrifield, H. E. Simmons, Topological Methods in Chemistry, Wiley, New York, 1989.

[2] H. Hosoya, Topological Index. A Newly Proposed Quantity Characterizing the Topological Nature of Structural Isomers of Saturated Hydrocarbons, Bull. Chem. Soc. Jpn. 44 (1971) 2332–2339.

[3] H. Hua, Minimizing a class of unicyclic graphs by means of Hosoya index, Math. Comput. Modelling 48 (5-6) (2008) 940–948.

[4] J. Ou, On extremal unicyclic molecular graphs with maximal Hosoya index, Discrete Appl. Math. 157 (2) (2009) 391–397.

[5] A. S. Pedersen, P. D. Vestergaard, The number of independent sets in unicyclic graphs, Discrete Appl. Math. 152 (1-3) (2005) 246–256.

[6] H. Deng, The smallest Merrifield-Simmons index of $(n, n+1)$-graphs, Math. Comput. Modelling 49 (1-2) (2009) 320–326.

[7] M. Startek, A. Włoch, I. Włoch, Fibonacci numbers and Lucas numbers in graphs, Discrete Appl. Math. 157 (4) (2009) 864–868.

[8] W. Yan, L. Ye, On the minimal energy of trees with a given diameter, Appl. Math. Lett. 18 (9) (2005) 1046–1052.

[9] A. Yu, X. Lv, The Merrifield-Simmons indices and Hosoya indices of trees with k pendant vertices, J. Math. Chem. 41 (1) (2007) 33–43.

[10] H. Hua, A sharp upper bound for the number of stable sets in graphs with given number of cut edges, Appl. Math. Lett. 22 (9) (2009) 1380–1385.

[11] S. Wagner, I. Gutman, Maxima and minima of the Hosoya index and the Merrifield-Simmons index: A survey of results and techniques, submitted (2009).

[12] I. Gutman, The energy of a graph: old and new results, in: Algebraic combinatorics and applications (Gößweinstein, 1999), Springer, Berlin, 2001, pp. 196–211.

[13] I. Gutman, O. E. Polansky, Mathematical concepts in organic chemistry, Springer-Verlag, Berlin, 1986.

[14] M. Fischermann, I. Gutman, A. Hoffmann, D. Rautenbach, D. Vidović, L. Volkmann, Extremal Chemical Trees, Z. Naturforsch. 57a (2002) 49–52.

[15] C. Heuberger, S. Wagner, Maximizing the number of independent subsets over trees with bounded degree, J. Graph Theory 58 (1) (2008) 49–68.

[16] S. Wagner, Extremal trees with respect to Hosoya index and Merrifield-Simmons index, MATCH Commun. Math. Comput. Chem. 57 (1) (2007) 221–233.

[17] R. Kirk, H. Wang, Largest number of subtrees of trees with a given maximum degree, SIAM J. Discrete Math. 22 (3) (2008) 985–995.

[18] M. Fischermann, A. Hoffmann, D. Rautenbach, L. Székely, L. Volkmann, Wiener index versus maximum degree in trees, Discrete Appl. Math. 122 (1-3) (2002) 127–137.

[19] H. Zhao, X. Li, On the Fibonacci numbers of trees, Fibonacci Quart. 44 (1) (2006) 32–38.