Iranian J. Math. Chem. 9 (3) September (2018) 179 − 186

Iranian Journal of Mathematical Chemistry

Journal homepage: ijmc.kashanu.ac.ir

One–Alpha Descriptor

DAMIR VUKIČEVIĆ[1] AND ZAHRA YARAHMADI[2,•]

[1]Department of Mathematics, University of Split, Nikole Tesle 12, HR-21000 Split, Croatia
[2]Department of Mathematics, Faculty of Science, Khorramabad Branch, Islamic Azad University, Khorramabad, I. R. Iran

ARTICLE INFO

Article History:

Received 5 February 2018
Accepted 11 March 2018
Published online 23 June 2018
Academic Editor: Gholam Hossein Fath-Tabar

Keywords:

One-alpha descriptor
Extermal graph
Tree

ABSTRACT

Recently, one-two descriptor has been defined and it has been shown that it is a good predictor of the heat capacity at *P* constant (CP) and of the total surface area (TSA). In this paper, we analyze its generalizations by replacing the value 2 by arbitrary positive value α. We show that these analyses may be on interest, because even good predictions of CP and TSA can be slightly improved. Furthermore, it can be expected that this more general descriptor can find a wider range of application than the original one. The extremal values of trees have been found for all values of α.

1 INTRODUCTION

The molecular descriptor is the final result of a logical and mathematical procedure which transforms chemical information encoded within a symbolic representation of a molecule into a useful number or the result of some standardized experiment [2]. Molecular descriptors have been shown to be useful in modeling many physico-chemical properties in numerous QSAR and QSPR studies [3-5].

In this paper, we introduce one-alpha descriptor. It is defined as the sum of the vertex contributions is such a way that each pendent vertex contributes 1, each vertex of degree two adjacent to pendent vertex contributes α, and also each vertex of degree higher than two also contributes α and another vertex contributes 0. If we take α=2, we get the previously defined [1] one-two descriptor. As in [1], we illustrate this definition for 3-ethyl-hexane in Figure 1.

• Corresponding Author (Email: z.yarahmadi@khoiau.ac.ir)
DOI: 10.22052/ijmc.2018.118091.1342

One-alpha descriptor of graph G will be denoted by OA(G). For instance, if G is 3-ethyl-hexane, then OA(G) = 3 + 4α. We show that one-alpha descriptor as generalization of one-two descriptor may be of interest in chemistry, since it can slightly improve predictions of the heat capacity at *P* constant (CP) and of the total surface area (TSA) for octane isomers. Further, we analyze mathematical properties of this descriptor. Namely, we find tight upper and lower bounds in the families of the trees with n vertices and the chemical trees with n vertices.

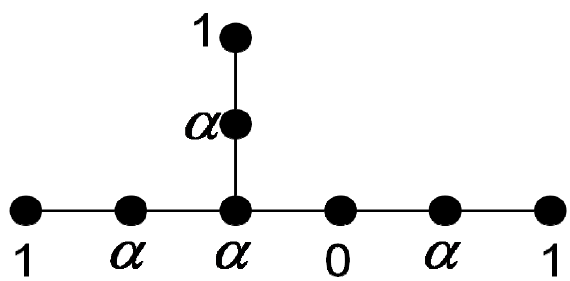

Figure 1. Vertex contributions of 3-ethyl-hexane. Each pendent vertex contributes 1, each vertex of degree two adjacent to pendent vertex contributes α, and also each vertex of degree higher than two also contributes α and another vertex contributes 0.

2. QSAR RESULTS

International Academy of Mathematical Chemistry [6] proposed four benchmark sets [7] as sets for testing the molecular descriptors. Also, recently Adriatic descriptors [8−12] have been proposed and in many cases they have provided better results than benchmark descriptors [8,11]. One-two descriptor outperformed both sets of descriptors in the linear modeling of TSA, and it was of comparable quality (slightly better) than benchmark descriptors in the linear modeling of CP, but not as good as the best Adriatic index [8−12].

However, here we show that α =2 does not give the best results in the set of all one-alpha descriptors. Namely in the linear modeling of TSA, the best results are obtained for α ≈1.8 and in the linear modeling of CP, the best results are obtained for α ≈ 2.18. The histograms that illustrate the changes of r^2 in the dependence of the values of α are presented on the Figures 2 and 3:

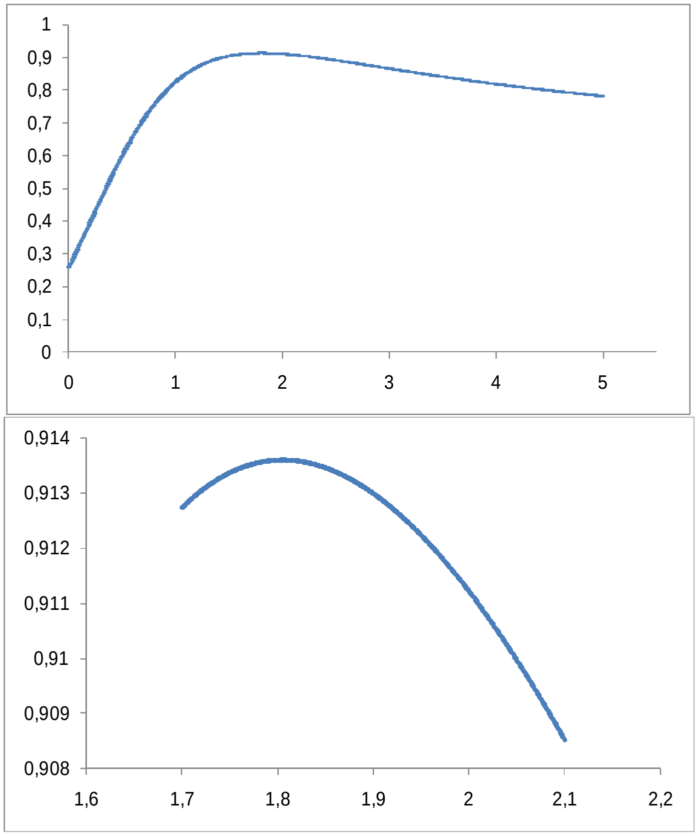

Figure 2. r^2 values of the linear models for estimation of the total surface area by one-alpha descriptors. On the left hand-side diagram α ∈ [0,5] and on the right hand-side diagram α ∈ [1.7, 2.1].

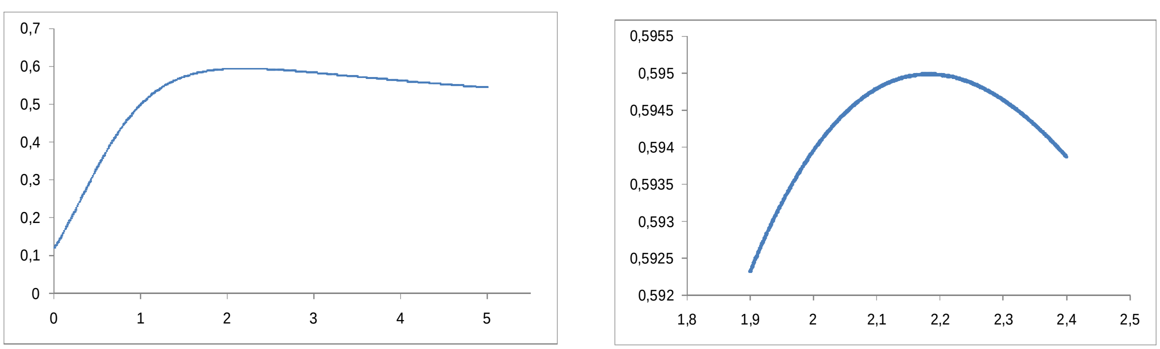

Figure 3. r^2 values of the linear models for estimation of the heat capacity at P constant by one-alpha descriptors. On the left hand-side diagram α ∈ [0,5] and on the right hand-side diagram α ∈ [1.9, 2.4].

3. MATHEMATICAL PROPERTIES

Before proving the main theorems, let us introduce some notations. By $n_i(G)$ we denote the number of vertices of degree i in G and by $d_G(u)$ we denote the degree of vertex u in graph G. Let x be any real number. By [x] we denote the greatest integer not greater than x. In the proofs of the main theorems, we shall use the following well known lemmas:

Lemma 1. Let G be a tree with at least 2 vertices. Then it holds:

$$n_1(G) = \sum_{i\geq 3}(i-2)\,n_i(G) + 2.$$

Proof. It is easy to see that:

$$\begin{aligned}\sum_{i\geq 3}(i-2)\,n_i(G) + 2 &= \sum_{i\geq 3} i n_i(G) - 2\sum_{i\geq 3} n_i(G) + 2\\ &= \left(2\,|V(G)| - n_1(G) - 2n_2(G) - 2\right)\\ &\quad - 2\left(|V(G)| - n_1(G) - n_2(G)\right) + 2\\ &= n_1(G).\end{aligned}$$

Lemma 2. Let G be a tree with maximal upper bound for one-alpha descriptor. Then G does not contain any vertices of contribution 0 to OA index.

Proof. Supposed to the contrary that there exists a vertex u of contribution 0, and adjacent to vertices v_1 and v_2 such that $d_G(v_1)$, $d_G(v_2) \geq 2$. Let $G' = G - uv_2 + v_1v_2$. It can be easily seen that contributions of all the vertices except u and v_1 to OA index remained the same, the contribution of v_1 did not decrease and the contribution of u increased from 0 to 1. Hence, $OA(G') > OA(G)$, which is contradiction. Hence, indeed there are no vertices of contribution 0. □

Lemma 3. Let G be a tree with maximal upper bound for one-alpha descriptor and $\alpha > 1$. Then G has at least a vertex of degree 2.

Proof. By Lemma 1,

$$n_1(G) = \sum_{i\geq 3}(i-2)\,n_i(G) + 2 \geq \sum_{i\geq 3} n_i(G) + 2.$$

By the above inequality we conclude that, $n_1(G) - 2 \geq \sum_{i\geq 3} n_i(G)$. By contrary we assume that $n_2(G) = 0$. Since $n_1(G) + \sum_{i\geq 3} n_i(G) = n$, $n_1(G) + n_1(G) - 2 \geq n_1(G) + \sum_{i\geq 3} n_i(G) = n$ and so $n_1(G) \geq n/2 + 1$, where $|V(G)| = n$.

Also, since $n_2(G) = 0$, $OA(G) = n_1(G) + \alpha\sum_{i\geq 3} n_i(G) = n_1(G) + \alpha(n - n_1(G))$. By assumption OA(G) is maximal and $\alpha > 1$, then $n_1(G)$ must have minimum value. Now we construct a graph G' such that $|V(G')| = n$ and $n_1(G) < n/2+1$ and $OA(G') > OA(G)$, which makes contradiction. There are three cases for $n \geq 4$, $n = 3k + 1$, $n = 3k + 2$ or $n = 3k + 3$.

For each case we construct the graph G' as Figure 4, such that $OA(G') = n_1(G') + \alpha(n - n_1(G'))$, $n_1(G') < n/2+1$ and $OA(G') > OA(G)$, which is contradiction. □

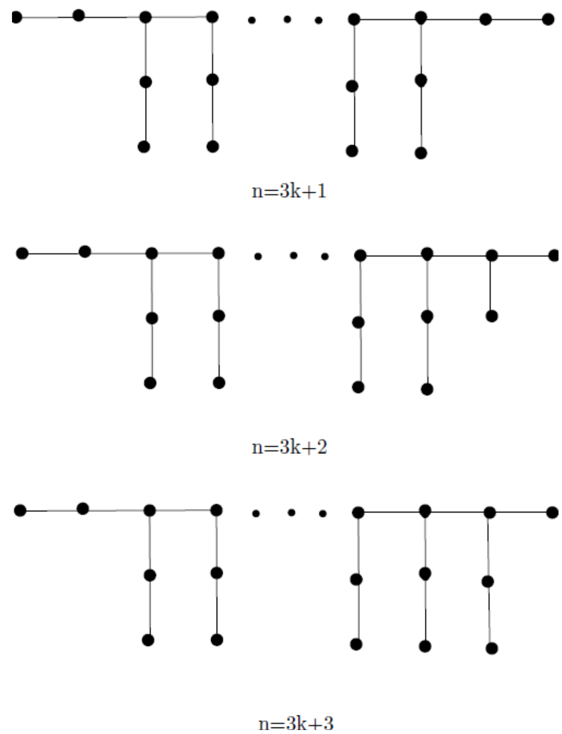

Figure 4. The constructed graph related Lemma 3.

Now, we can obtain lower and upper bounds for trees to different values of α.

Theorem 4. Let *G* be a tree with n vertices. It holds

$$OA(G) \geq \begin{cases} 0 & n=1 \\ 2 & n=2 \\ 2+\alpha & n=3 \\ \begin{cases} 2+2\alpha & \alpha < n-3 \\ 2n-4 & \alpha = n-3 \\ (n-1)+\alpha & \alpha > n-3 \end{cases} & n \geq 4 \end{cases}.$$

Proof. To prove the lower bound for OA(G), it can be easily checked for $n \leq 3$. For n=1, 2, 3 the lower bound can be obtained immediately. Hence, let us assume that $n \geq 4$. We use this well known fact that each tree with at least two vertices has at least two leaves. If G is a star, then it is easy to see that $OA(G) = (n-1) + \alpha$. If G is not a star, than there are at least two vertices adjacent to leaves hence $OA(G) \geq 2 + 2\alpha$. Now if $\alpha = n - 3$, then $2 + 2\alpha = (n-1) + \alpha = 2n - 4$ and hence for $\alpha < n-3$, we have $2 + 2\alpha < (n - 1) + \alpha$. Also for $\alpha > n - 3$, $2 + 2\alpha > (n - 1) + \alpha$. Examples of the extremal graphs obtaining the lower bounds for $\alpha < n - 3$ are paths P_n, for $n \geq 4$ and for $\alpha > n - 3$ are the stars S_n, for $n \geq 4$. Also for $\alpha = n - 3$, $OA(P_n) = OA(S_n) = 2n - 4$. □

Theorem 5. Let *G* be a tree with *n* vertices. Then

$$OA(G) \leq \begin{cases} (n-1)+\alpha & \alpha \leq 1 \\ \begin{cases} \frac{n+2}{3} + \alpha(n - \frac{n+2}{3}) & \alpha > 1,\ \frac{n+2}{3} \in Z \\ \left\lfloor \frac{n+2}{3} \right\rfloor + 1 + \alpha(n - \left\lfloor \frac{n+2}{3} \right\rfloor - 1) & \alpha > 1,\ \frac{n+2}{3} \notin Z \end{cases} \end{cases}.$$

Proof. First assume that $\alpha \leq 1$. By Lemma 2, $OA(G) = n_1(G) + \alpha(n - n_1(G))$, then $OA(G)$ is maximum if and only if $n_1(G)$ is maximum. A tree with maximum number of leaves is a star and $OA(S_n) = (n - 1) + \alpha$. Therefore, $OA(G) \leq (n - 1) + \alpha$.

Now we assume that $\alpha > 1$. Let us prove that for each n, there exists an n-vertices tree G with maximum OA such that $\Delta(G) \leq 3$. Suppose that G is a tree with n vertices such that OA(G) is maximum. If $\Delta(G) \leq 3$, then there is nothing to prove. Let $\Delta(G) > 4$, by Lemma 3, there exists a vertex v in V(G) such that $deg_G v = 2$. Let $u \in V(G)$, $deg_G u = \Delta(G)$ and $T_1, T_2, ..., T_{\Delta(G)}$ be branches from u, see Figure 5. Without loss of generality, we can assume that v is in T_1.

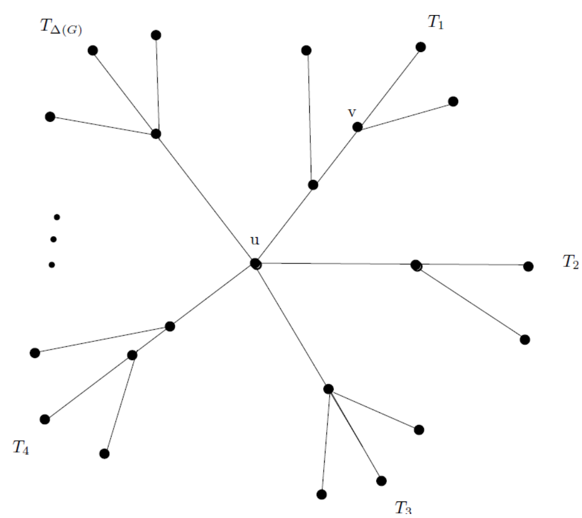

Figure 5. The configuration of graph G in Theorem 5.

Now we instruct a graph G_1 as follow: We omit the branches $T_4, T_5, ..., T_{\Delta(G)}$ and join them to vertex v. By this transformation we obtain the tree G_1, such that $OA(G_1)$ =

OA(G) and since OA(G) is maximum then $OA(G_1)$ will be maximum, then by Lemma 3 there exists a vertex v_1 in $V(G_1)$, such that $d_{G1}(v_1)=2$. It is clear that $\Delta(G_1) \leq \Delta(G)$ and $d_{G1}(v) = \Delta(G) - 1$. By continuing the above process we can obtain the graph G_2, such that $\Delta(G_2) \leq \Delta(G_1) \leq \Delta(G)$ and $OA(G_2) = OA(G_1) = OA(G)$. Finally by continuing this process we can obtain the graph G_s from G_{s-1}, such that $\Delta(G_s) \leq 3$ and $OA(G_s) = ... = OA(G_1) = OA(G)$. Hence from beginning we can assume that G is a tree with maximum OA and $\Delta(G) \leq 3$. Now by Lemma 1, $n_1(G) = \sum_{i\geq 3}(i - 2)\, n_i(G) + 2 = n_3(G) + 2$. We have $n_1(G) + n_2(G) + n_3(G) = 2$, and so

$$2\, n_1(G) + n_2(G) = n + 2. \quad (1)$$

By Lemma 2, G does not contain any vertices of contribution 0 to OA index then $OA(G) = n_1(G) + \alpha(n - n_1(G))$. Since $\alpha > 1$, OA(G) is maximum if and only if $n_1(G)$ is minimum. Again, by Lemma 2, we have $n_2(G) \leq n_1(G)$. From Equation (1), we conclude that $3n_1(G) \geq 2n_1(G) + n_2(G) = n + 2$ and then $n_1(G) \geq (n + 2)/3$. Hence if

$$(n + 2)/3 = k \in Z \ (n \equiv 1 \bmod 3),$$

then $n_1(G)=(n+2)/3$ is minimum value for $n_1(G)$ and if

$$(n + 2)/3 = k \notin Z \ (n \equiv 0 \text{ or } 2 \bmod 3),$$

then $n_1(G)=[(n+2)/3]+1$ is minimum value. The examples of the extremal graphs obtaining the upper bounds are presented in the Figure 6.

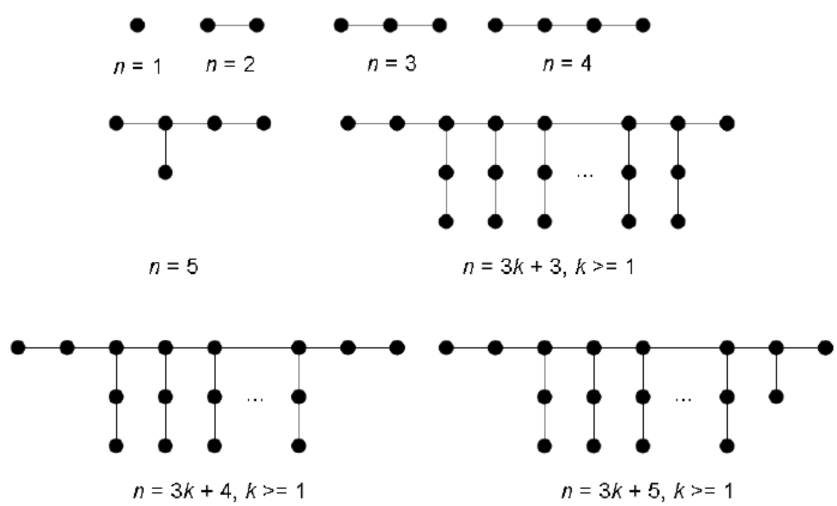

Figure 6. Extremal graphs obtaining the upper bounds.

This proves the Theorem. □

ACKNOWLEDGMENT. The partial support of Croatian Ministry of Science, Education and Sport (grants no. 177-0000000-0884 and 037-0000000-2779) is gratefully acknowledged.

REFERENCES

1. D. Vukičević, M. Bralo, A. Klarić, A. Markovina, D. Spahija, A. Tadić and A. Žilić, One-Two descriptor, *J. Math. Chem.* **48** (2010) 395–400.
2. R. Todeschini and V. Consonni, *Handbook of Molecular Descriptors,* Wiley−VCH, Weinheim, 2000.
3. N. Trinajstić, *Chemical Graph Theory*, CRC Press, Boca Raton, 1992.
4. J. Devillers and A. T. Balaban (Eds.), Topological indices and related descriptors in QSAR and QSPR, Gordon and Breach, Amsterdam, 1999.
5. M. Karelson, Molecular *Descriptors in QSAR/QSPR*, Wiley−Interscience, New York, 2000.
6. http://www.iamc-online.org/
7. http://www.moleculardescriptors.eu/dataset/dataset.htm
8. D. Vukičević and M. Gašperov, Bond additive modeling 1. Adriatic indices, *Croat. Chem. Acta* **83** (3) (2010) 243–260.
9. D. Vukičević, Bond additive modeling 2. Mathematical properties of max-min rodeg index, *Croat. Chem. Acta* **83** (3) (2010) 261–273.
10. D. Vukičević, Bond additive modeling 3. Comparison between the product-connectivity index and sum-connectivity index, *Croat. Chem. Acta* **83** (3) (2011) 349−351.
11. D. Vukičević, Bond additive modeling 4. QSPR and QSAR studies of variable adriatic indices, *Croat. Chem. Acta* **84** (1) (2011) 87−91.
12. D. Vukičević, Bond additive modeling 5. Mathematical properties of variable sum exdeg index, *Croat. Chem. Acta* **84** (1) (2011) 93–101.