J Math Chem (2009) 46:1369–1376
DOI 10.1007/s10910-009-9520-x

ORIGINAL PAPER

Topological index based on the ratios of geometrical and arithmetical means of end-vertex degrees of edges

Damir Vukičević · Boris Furtula

Received: 7 December 2008 / Accepted: 6 January 2009 / Published online: 22 January 2009

Abstract Research on the topological indices based on end-vertex degrees of edges has been intensively rising recently. Randić index, one of the best-known topological indices in chemical graph theory, is belonging to this class of indices. In this paper, we introduce a novel topological index based on the end-vertex degrees of edges and its basic features are presented here. We named it as *geometrical-arithmetic index* (GA).

Keywords Tree · Degree of vertex · Arithmetical mean · Geometrical mean · End-vertex degrees of edges

1 Introduction

Molecular descriptors are playing significant role in chemistry, pharmacology, etc. Among them, topological indices have a prominent place [1]. There are numerous of topological descriptors that have found some applications in theoretical chemistry, especially in QSPR/QSAR research. Within all topological indices ones of the most investigated are the descriptors based on the valences of atoms in molecules (in graph-theoretical notions *degrees of vertices of graph*).

Let us consider following formula:

$$\sum_{uv \in E(G)} (d_u d_v)^{\alpha}$$

D. Vukičević
Faculty of Natural Sciences and Mathematics, University of Split, 21000 Split, Croatia

B. Furtula (✉)
Faculty of Science, University of Kragujevac, P.O. Box 60, 34000 Kragujevac, Serbia
e-mail: furtula@kg.ac.rs

where summation goes over all edges of the edge set $E(G)$ of graph G, and d_u, d_v are the degrees of end-vertices of an edge uv.

Then, if $\alpha = 1$ one obtains second Zagreb index, M_2 [2]; for $\alpha = -1/2$ we get Randić connectivity index, χ [3]which is one of mostly used topological descriptors today; for $\alpha = -1$ one obtains modified Zagreb indices [4], etc.

Beside above-mentioned there are other topological descriptors based on end-vertex degrees of edges of graph that have found some applications in QSPR/QSAR research (see for example [5,6]. In a past few years a number of papers have appeared considering these topological indices (for example see [7–15] and references cited therein).

Here, we are proposing a new index which belonging to this class of topological indices. It is defined as follows:

$$GA(G) = \sum_{uv \in E(G)} \frac{\sqrt{d_u d_v}}{(d_u + d_v)/2} = \sum_{uv \in E(G)} \frac{2\sqrt{d_u d_v}}{d_u + d_v},$$

where summation goes over all edges of graph G, and d_u, d_v are the degrees of vertices that are connected with edge uv.

This index is named as *geometrical-arithmetic index* (*GA*) because, as it can be seen from the definition, it consists from geometrical mean of end-vertex degrees of an edge uv ($\sqrt{d_u d_v}$) as numerator and arithmetic mean of end-vertex degrees of the edge uv ($(d_u + d_v)/2$) as denominator.

In the following sections, predictive power of *GA* index will be discussed as well as its some basic mathematical properties.

2 *GA* index as possible tool for QSPR/QSAR research

In order to investigate predictive power of *GA* index we used octanes and some of their physico-chemical properties as resource. We found experimental results at the www.moleculardescriptors.eu. The following physico-chemical properties have been modeled:

- Boiling point (*BP*)
- Entropy (*S*)
- Enthalpy of vaporization (*HVAP*)
- Standard enthalpy of vaporization (*DHVAP*)
- Enthalpy of formation (*HFORM*)
- Acentric factor (*AcenFac*)

and results are compared with those obtained using the well-known Randić index. We chose those physico-chemical properties for which *GA* index and Randić index give reasonably good correlations, i.e. correlation coefficients are larger than 0.8.

In the Table 1 are depicted graphs that show correlations between *GA* and Randić index on one hand and above-mentioned properties on the other. From depicted graphs, it is not obvious which index gives better results. Therefore, we conducted a simple statistical analysis to compare *GA* and Randić index. Results are presented in Table 2.

Table 1 Graphs showing correlation between some phyisico-chemical properties and *GA* and Randić index respectively

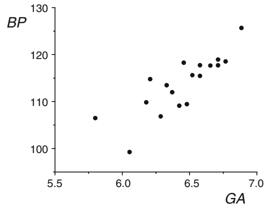
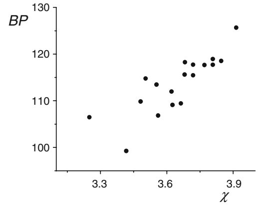
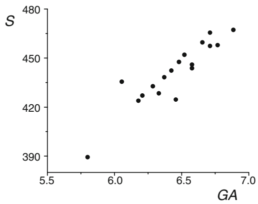
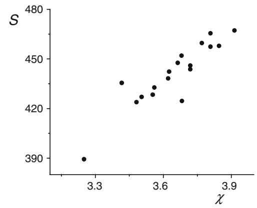
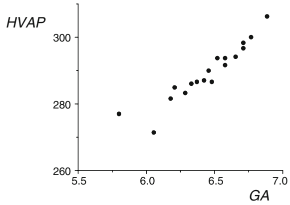
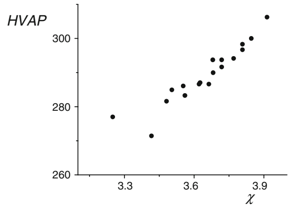
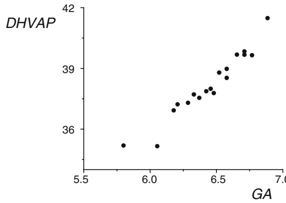
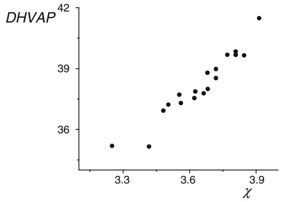
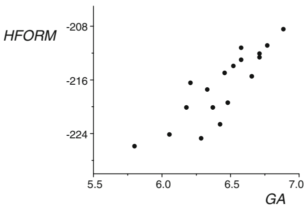
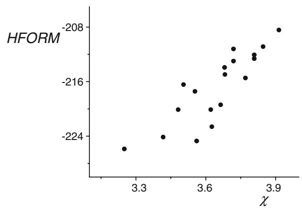
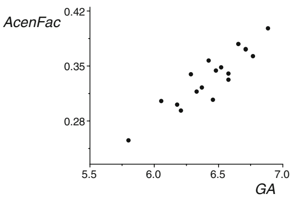
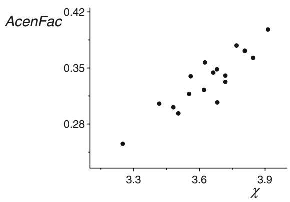
	GA index	Randić index
Boiling point (*BP*)		
Entropy (*S*)		
Enthalpy of vaporization (*HVAP*)		
Standard enthalpy of vaporization (*DHVAP*)		
Enthalpy of formation (*HFORM*)		
Acentric factor (*AcenFac*)		

Table 2 Correlation coefficients and ratio of quadratic mean of residuals for graphs depicted in Table 1

	Correlation coefficient (R)		1 - RQR (%)
	GA index	Randić index	
Boiling point (*BP*)	0.823	0.821	0.562
Entropy (*S*)	0.912	0.906	2.942
Enthalpy of vaporization (*HVAP*)	0.941	0.936	4.152
Standard enthalpy of vaporization (*DHVAP*)	0.966	0.958	9.005
Enthalpy of formation (*HFORM*)	0.858	0.850	2.494
Acentric factor (*AcenFac*)	0.912	0.904	4.051

It can be seen from data for correlation coefficient (R) (Table 2) that in all cases GA index gives somewhat better results. Apparently, a superficial glance on the correlation coefficients do not show strong justification for introducing a new index, because correlation coefficients that we obtain in the case of GA index are not significantly better than those obtained using Randić index. However, predicting power of a new index is reasonable and that can be seen from the ratio of quadratic mean of residuals (RQR):

$$RQR = \frac{\sqrt{\frac{\sum_{i=1}^{n}(a \cdot GA_i + b - Exp_i)^2}{n}}}{\sqrt{\frac{\sum_{i=1}^{n}(a' \cdot \chi_i + b' - Exp_i)^2}{n}}} = \sqrt{\frac{\sum_{i=1}^{n}(a \cdot GA_i + b - Exp_i)^2}{\sum_{i=1}^{n}(a' \cdot \chi_i + b' - Exp_i)^2}}$$

One should observe that only for boiling point (BP) GA index does not give better results than Randić. In all other cases, the prediction power of GA index is at least for 2.5% better than prediction power of Randić index. The greatest improvement in prediction with GA index comparing to Randić index is obtained in the case of standard enthalpy of vaporization (more than 9%). That is why we believe that GA index should be considered in the future QSPR/QSAR researches.

3 Lower and upper bounds of *GA* index for general graphs and chemical graphs

In this section are given some basic mathematical features of geometrical-arithmetic index (GA).

Theorem 1 *Let G be a simple graph with n vertices, then*

$$0 \leq GA(G) \leq \binom{n}{2}.$$

Lower bound is achieved if and only if G is an empty graph and upper bound is achieved if and only if G is a complete graph.

Proof Note that contribution of each edge is positive. Hence, $GA(G) \geq 0$ and $GA(G) = 0$ only if G is an empty graph. Now, let us prove the upper bound. It is well known that geometric mean is less or equal to arithmetic mean. Hence,

$$GA(G) \leq 1 \cdot m \leq \binom{n}{2},$$

where m is the number of edges. Moreover, the equality is obtained if and only if $G$is regular graph with $\binom{n}{2}$ edges. The only such graph is the complete graph. □

Theorem 2 *Let G be a simple connected graph with n vertices, then*

$$\frac{2(n-1)^{3/2}}{n} \leq GA(G) \leq \binom{n}{2}.$$

Lower bound is achieved if and only if G is a star and upper bound is achieved if and only if G is a complete graph.

Proof Assume that $n \geq 2$, because otherwise the claim is trivial. Upper bound follows from Theorem 1. Now, let us prove that for contribution of each edge uv holds:

$$\frac{2\sqrt{d_u d_v}}{d_u + d_v} \leq \frac{2\sqrt{n-1}}{n}.$$

Without loss of generality, we may assume that $d_u \leq d_v$. Denote $x = \frac{d_u}{d_v}$ and note that $\frac{1}{n-1} \leq x \leq 1$. It holds:

$$\frac{2\sqrt{d_u d_v}}{d_u + d_v} = \frac{2\sqrt{x}}{1+x}.$$

Simple differential calculation shows that the right hand-side is ascending on the interval $\left[\frac{1}{n-1}, 1\right)$, hence it achieves minimum at $x = \frac{1}{n-1}$. Hence, indeed $\frac{2\sqrt{d_u d_v}}{d_u+d_v} \leq \frac{2\sqrt{n-1}}{n}$. Since, graph is connected, it has at least $n-1$ edges. Therefore,

$$GA(G) \geq (n-1) \cdot \frac{2\sqrt{n-1}}{n} = \frac{2(n-1)^{3/2}}{n}.$$

Moreover, the equality is obtained if and only if graph has $n-1$ edges each connecting vertices of degrees 1 and n. The only such graph is a star. □

Theorem 3 *Let T be a tree with n vertices, then*

$$\frac{2(n-1)^{3/2}}{n} \leq GA(T) \leq \begin{cases} 0, & n = 1 \\ 1, & n = 2 \\ \frac{4\sqrt{2}}{3} + (n-3), & n \geq 3 \end{cases}$$

Lower bound is achieved if and only if T is a star and upper bound is achieved if and only if T is a path.

Proof The lower bound follows from Theorem 2. Let us prove the upper bound. Assume that $n \geq 3$, because otherwise the claim is trivial. Let u be a pendant vertex in T and v its only neighbor. Then

$$\frac{2\sqrt{d_u d_v}}{d_u + d_v} = \frac{2\sqrt{d_v}}{1 + d_v}.$$

Simple calculation shows that right hand-side is descending on the segment $[2, n-1]$. Hence,

$$\frac{2\sqrt{d_u d_v}}{d_u + d_v} = \frac{2\sqrt{d_v}}{1 + d_v} \leq \frac{2\sqrt{2}}{3}.$$

Taking into account that contribution of every edge with end-degrees ≥ 2 is at most one, it holds:

$$GA\,(T) \leq \frac{2\sqrt{2}}{3} \cdot l + (n - 1 - l) \cdot 1,$$

where l is the number of leaves in tree T. Since, every tree has at least two edges, it holds:

$$GA\,(T) \leq \frac{2\sqrt{2}}{3} \cdot l + (n - 1 - l) \cdot 1 \leq \frac{4\sqrt{2}}{3} \cdot l + (n - 3)\,.$$

Moreover, the equality holds if and only if tree has exactly two edges connecting vertices of degrees 1 and 2; and all other edges connect vertices of the same degree. The only such graph is path. □

Theorem 4 *Let T be a chemical tree with n vertices, then*

$$\frac{13}{15}n - \frac{17}{15} \leq GA\,(T) \leq \begin{cases} 0, & n = 1 \\ 1, & n = 2 \\ \frac{4\sqrt{2}}{3} + (n - 3)\,, & n \geq 3 \end{cases}$$

Lower bound is achieved for trees that contains only vertices of degrees 1 and 4; and upper bound is achieved if and only if T is a path.

Proof The upper bound follows from Theorem 3. Let us prove the lower bound. If $n \leq 2$, the claim is trivial, hence assume that $n \geq 3$. Suppose to contrary that $GA\,(T) < \frac{13}{15}n - \frac{17}{15}$ and that T is such tree with the smallest number of vertices. First, let us prove that T has no vertices of degree 2. Suppose to the contrary that T has at least one vertex u of degree 2 and let its neighbors be v and w. Let T' be a tree obtained by replacing path vw with the single edge, i.e. $T' = (T - u) \cup \{vw\}$. Since, T' has less vertices then T, it follows that

$$GA\left(T'\right) \geq \frac{13}{15}\,(n - 1) - \frac{17}{15},$$

but then

$$\begin{aligned}GA\left(T\right) &= GA\left(T'\right) + \frac{2\sqrt{d_u d_v}}{d_u + d_v} + \frac{2\sqrt{d_u d_w}}{d_u + d_w} - \frac{2\sqrt{d_v d_w}}{d_w + d_w} \\ &\geq \frac{13}{15}\left(n-1\right) - \frac{17}{15} + \frac{2\sqrt{2d_v}}{2+d_v} + \frac{2\sqrt{2d_w}}{2+d_w} - \frac{2\sqrt{d_v d_w}}{d_w + d_w} \\ &\geq \frac{13}{15}\left(n-1\right) - \frac{17}{15} + \max_{1\leq x\leq y\leq 4}\left(\frac{2\sqrt{2x}}{2+x} + \frac{2\sqrt{2y}}{2+y} - \frac{2\sqrt{xy}}{x+y}\right) \\ &\geq \{\text{simple check of all 15 possibilities}\} \\ &> \frac{13}{15}\left(n-1\right) - \frac{17}{15} + 0.885 > \frac{13}{15}n - \frac{17}{15},\end{aligned}$$

which is a contradiction. Denote by n_i number of vertices of degree i and denote by m_{ij} number of edges that connect vertices of degrees i and j. Put $x_{ij} = \frac{2\sqrt{ij}}{i+j}$. It holds:

$$\begin{aligned}GA\left(T\right) &= m_{13}x_{13} + m_{14}x_{14} + m_{33}x_{33} + m_{34}x_{34} + m_{44}x_{44} \\ &\geq \left(m_{13} + m_{14}\right)\min\left\{x_{13}, x_{14}\right\} + m_{33}x_{33} + m_{34}x_{34} + m_{44}x_{44} \\ &= n_1 x_{14} + m_{33}x_{33} + m_{34}x_{34} + m_{44}x_{44}\end{aligned}$$

Note that $m_{33}+m_{34}+m_{44} = n-n_1-1$ and that $m_{33}+m_{34} \leq \min\left\{3n_3, n-n_1-1\right\}$. Therefore,

$$\begin{aligned}GA\left(T\right) &\geq n_1 x_{14} + m_{33}x_{33} + m_{34}x_{34} + m_{44}x_{44} \\ &\geq n_1 x_{14} + \min\left\{3n_3, n-n_1-1\right\}\cdot\min\left\{x_{33}, x_{34}, x_{44}\right\} \\ &\quad + \left(n-n_1-1-\min\left\{3n_3, n-n_1-1\right\}\right)\cdot x_{44} \\ &= n_1 x_{14} + \min\left\{3n_3, n-n_1-1\right\}\cdot x_{34} \\ &\quad + \left(n-n_1-1-\min\left\{3n_3, n-n_1-1\right\}\right)\cdot x_{44} \\ &= n_1 x_{14} + \left(n-n_1-1\right)x_{44} - \left(x_{44}-x_{34}\right)\cdot\min\left\{3n_3, n-n_1-1\right\} \\ &\geq n_1 x_{14} + \left(n-n_1-1\right)x_{44} - 3n_3\left(x_{44}-x_{34}\right).\end{aligned}$$

From $n = n_1 + n_3 + n_3$ and $n_1 = n_3 + 2n_4 + 2$, it follows that $n_1 = \frac{2}{3} + \frac{2n}{3} - \frac{n_3}{3}$, hence

$$\begin{aligned}GA\left(T\right) &\geq \left(\frac{2}{3} + \frac{2n}{3} - \frac{n_3}{3}\right)x_{14} + \left(\frac{n}{3} + \frac{n_3}{3} - \frac{5}{3}\right)x_{44} - 3n_3\left(x_{44} - x_{34}\right) \\ &= \left(\frac{2}{3}x_{14} + \frac{1}{3}x_{44}\right)n + \left(-\frac{x_{14}}{3} + \frac{x_{44}}{3} - 3x_{44} + 3x_{34}\right)n_3 + \left(\frac{2}{3}x_{14} - \frac{5}{3}x_{44}\right).\end{aligned}$$

Simple calculation shows that factor standing by n_3 is positive, hence

$$GA(T) \geq \left(\frac{2}{3}x_{14} + \frac{1}{3}x_{44}\right)n + \left(\frac{2}{3}x_{14} - \frac{5}{3}x_{44}\right) = \frac{13n-17}{15}.$$

Finally, let us tree q vertices of degree and $2q+2$ vertices of degree one. In this tree, it holds: $m_{44} = q-1$ and $m_{14} = 2q+2$, hence

$$\begin{aligned} GA(T) &\geq (q-1)\cdot x_{44} + (2q+2)\,x_{14} \\ &= (q-1) + (2q+2)\cdot\frac{4}{5} = \frac{13(3q+2)-17}{15} = \frac{13n-17}{15}. \end{aligned}$$

This proves the Theorem. □

4 Conclusion

We proposed a new topological index based on end-vertex degrees of edges. It has been shown that this index can be used as predictive tool in QSPR/QSAR researches. Predictive power of this index has been tested on some physico-chemical properties of octanes. Obtained results show that it gives somewhat better results comparing with well-known Randić index. In addition, we analyzed some of its basic mathematical properties. It has been found a lower and upper bounds in the case of simple connected graphs and trees as well as in the case of chemical graphs and chemical trees.

Acknowledgements This work is supported by Ministry of Science and Technological Development of the Republic of Serbia through the project No. 144015G (B. Furtula), and Ministry of Science, Education and Sports of the Republic of Croatia through Grants Nos. 177-0000000-0884 (D. Vukičević), 037-0000000-2779 (D. Vukičević).

References

1. R. Todeschini, V. Consonni, *Handbook of Molecular Descriptors* (Wiley-VCH, Weinheim, 2000)
2. I. Gutman, N. Trinajstić, Chem. Phys. Lett. **17**, 535 (1972)
3. M. Randić, J. Am. Chem. Soc. **97**, 6609 (1975)
4. S. Nikolić, G. Kovačević, A. Miličević, N. Trinajstić, Croat. Chem. Acta **76**, 113 (2003)
5. E. Estrada, L. Torres, L. Rodríguez, I. Gutman, Indian J. Chem. **37A**, 849 (1998)
6. E. Estrada, Chem. Phys. Lett. **463**, 422 (2008)
7. D. Vukičević, A. Graovac, MATCH Commun. Math. Comput. Chem. **60**, 37 (2008)
8. H. Hua, MATCH Commun. Math. Comput. Chem. **60**, 45 (2008)
9. L. Sun, R.S. Chen, MATCH Commun. Math. Comput. Chem. **60**, 57 (2008)
10. D. Vukičević, S.M. Rajtmajer, N. Trinajstić, MATCH Commun. Math. Comput. Chem. **60**, 65 (2008)
11. H. Liu, M. Lu, F. Tian, J. Math. Chem. **44**, 301 (2008)
12. Y. Jiang, M. Lu, J. Math. Chem. **43**, 955 (2008)
13. H. Hua, M. Wang, H. Wang, J. Math. Chem. **43**, 737 (2008)
14. J. Zhang, H. Deng, S. Chen, J. Math. Chem. **42**, 941 (2007)
15. R. García-Domenech, J. Gálvez, J.V. de Julián-Ortiz, L. Pogliani, Chem. Rev. **108**, 1127 (2008)