Molecular Topological Index: A Relation with the Wiener Index

Douglas J. Klein

Department of Marine Sciences, Texas A&M University, Galveston, Texas 77553-1675

Zlatko Mihalić

Faculty of Science and Mathematics, The University of Zagreb, 41000 Zagreb, The Republic of Croatia

Dejan Plavšić and Nenad Trinajstić*

The Rugjer Bošković Institute, P.O. Box 1016, 41001 Zagreb, The Republic of Croatia

Received January 31, 1992

It is shown analytically that the recently introduced molecular topological index and the Wiener index, known since 1947, are closely related graph-theoretical invariants for acyclic structures.

Schultz[1] has recently introduced an index for characterization of alkanes by an integer, which he named the molecular topological index (MTI). Since this is not a particularly distinctive label, the MTI was also called the Schultz index[2] after its originator. The MTI appears to be an attractive graph-theoretical index that is easy to compute and has structural significance, although it is not an especially discriminative descriptor.[3]

The Schultz index is defined as

$$\mathrm{MTI} = \sum_{j=1}^{N} e_j \tag{1}$$

where

$$e_j = \sum_{i=1}^{N} v_i (A_{ij} + D_{ij}) \tag{2}$$

In eq 2 v_i is the valence of vertex i in a molecular graph[4] **G**, A_{ij} is the element of the adjacency matrix[5,6] for **G** (which is equal to unity if vertices i and j are adjacent and otherwise is zero), D_{ij} is the element of the distance matrix[5,6] for **G** (which is equal to the length of the shortest path between vertices i and j), and N is the number of vertices in **G**.

In a recent comparative study[7] of molecular descriptors derived from the distance matrix, it has been found that the MTI and the Wiener index[8] (W) are strongly ($r = 0.999$) linearly correlated distance indices for the set of trees (depicting carbon skeletons of alkanes) with up to 10 vertices. This result suggests the existence of a formal relation between MTI and W for trees. In the present report we wish to show that there is indeed a close connection between these two distance descriptors.

The Wiener index was introduced by Wiener[8] in 1947, but its graph-theoretical definition was first given by Hosoya[9] in 1971. The Wiener index W is defined as

$$W = \frac{1}{2} \sum_{i,j=1}^{N} D_{ij} \tag{3}$$

or if

$$W_i = \sum_{j=1}^{N} D_{ij} \tag{4}$$

then

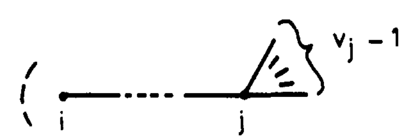
$$W = \frac{1}{2} \sum_{i=1}^{N} W_i \tag{5}$$

We now propose the following theorem: *If G is a tree, then*

$$e_i = 2W_i + (v_i + v'_i) - N + 1 \tag{6}$$

where v'_i is the number of next-nearest neighbors to i.

The proof of this theorem is as follows. For a pair of vertices i and j define the sum S_{ji} over all D_{ki} such that k is adjacent to j but more distant from i than j is. Clearly $S_{ii} = v_i$. For $j \neq i$ we have in a tree the circumstance depicted below:



That is, there is a single path of length D_{ji} from i to j, followed by single steps to $v_j - 1$ more distant neighbors of j. Thence

$$S_{ji} = (v_j - 1)(D_{ji} + 1),\ i \neq j \tag{7}$$

With the addition of the appropriate correction for $i = j$, we then have

$$S_{ji} = (v_j - 1)(D_{ji} + 1) + \delta_{ij} \tag{8}$$

Now returning to the definition of S_{ji}, one sees that summation over all j results in a sum over all D_{ki} each a single-time since each k ($\neq i$) is adjacent (in a tree) to exactly one other vertex (say j) that is closer to i than k is. That is

$$\sum_{j=1}^{N} S_{ji} = \sum_{k=1}^{N} D_{ki} = W_i \tag{9}$$

But with the substitution from eq 8, we have

$$\begin{aligned} W_i &= \sum_{j=1}^{N} [(v_j - 1)(D_{ji} + 1) + \delta_{ij}] \\ &= \sum_{j=1}^{N} v_j D_{ji} - \sum_{j=1}^{N} D_{ji} + \sum_{j=1}^{N} (v_j - 1) + \sum_{j=1}^{N} \delta_{ij} \end{aligned} \tag{10}$$

Here the last j-sum clearly is equal to 1, while the next to the last j-sum gives

$$\sum_{j=1}^{N} v_j - N = 2(N-1) - N = N - 2 \tag{11}$$

The above follows from the handshaking lemma[10]

$$\sum_{j=1}^{N} v_j = 2M \tag{12}$$

(where M is the number of edges in G) and the relationship between M and N in a tree

$$M = N - 1 \tag{13}$$

Thence, further recalling the definitions of e_i and W_i we have

$$W_i = e_i - \sum_{j=1}^{N} v_j A_{ij} - W_i + N - 1 \tag{14}$$

The j-sum in eq 14 can be given in a more convenient form

$$\sum_{j=1}^{N} v_j A_{ij} = \sum_{j=1}^{N} (v_j - 1) A_{ji} + \sum_{j=1}^{N} A_{ji} = v'_i + v_i \tag{15}$$

Combining eqs 14 and 15 we obtain

$$W_i = e_i - (v'_i + v_i) - W_i + N - 1 \tag{16}$$

and the theorem follows immediately.

We may now readily note

$$\begin{aligned} \mathrm{MTI} &= \sum_{i=1}^{N} e_i \\ &= \sum_{i=1}^{N} [2W_i + (v_i + v'_i) - N + 1] \\ &= 4W + \sum_{i=1}^{N} (v_i + v'_i) - N^2 + N \end{aligned} \tag{17}$$

Since the remnant i-sum in our last equation can be given, utilizing again the (generalized) handshaking lemma, in terms of the number of vertices and the number p_2 of paths of length two in a tree

$$\sum_{i=1}^{N} (v_i + v'_i) = 2[(N-1) + p_2] \tag{18}$$

we have the following corollary: *If G is a tree, then*

$$\mathrm{MTI} = 4W + 2p_2 - (N-1)(N-2) \tag{19}$$

Therefore, if we know the Wiener index of a tree, we immediately also know its molecular topological index, because the information on p_2-number is also contained in the distance matrix used to generate W. We note here that the p_2-number is identical to the Gordon–Scantlebury index,[11] which in turn is equal to half of the Platt index.[12,13] It is also interesting to mention that there exist several relationships among topological indices as one discussed here, or similar ones, which all belong to the so-called "quadratic index family".[14] Combinations of the Wiener and Platt indices, like the Schultz index, had been used already successfully by Platt[15] for correlations with thermodynamic properties of alkanes.

ACKNOWLEDGMENT

This work was supported in part by NSF Grant JPN 001, by the Welch Foundation of Houston, TX, and by the Ministry of Science, Technology and Informatics of the Republic of Croatia through Grants 1-07-159, 1-07-165, and 1-07-185.

REFERENCES AND NOTES

(1) Schultz, H. P. Topological Organic Chemistry. 1. Graph Theory and Topological Indices of Alkanes. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 227–228.

(2) Trinajstić, N. *Chemical* Graph *Theory*, 2nd ed.; CRC Press: Boca Raton, FL, 1992; Chapter 10.

(3) Müller, W. R.; Szymanski, K.; Knop, J. V.; Trinajstić, N. Molecular Topological Index. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 160–163.

(4) Trinajstić, N. *Chemical Graph Theory*; CRC Press: Boca Raton, FL, 1983; Vol. I, Chapter 3.

(5) Harary, F. *Graph Theory*, 2nd ed.; Addison-Wesley: Reading, MA, 1971; Chapter 16.

(6) Cvetković, D. M.; Doob, M.; Sachs, H. *Spectra of Graphs: Theory and Application*; Academic Press: New York, 1980; Chapter 8.

(7) Mihalić, Z.; Nikolić, S.; Trinajstić, N. A Comparative Study of Molecular Descriptions Derived from the Distance Matrix. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 28–37.

(8) Wiener, H. Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc.* **1947**, *69*, 17–20.

(9) Hosoya, H. Topological Index: A Newly Proposed Quantity Characterizing the Topological Nature of Structured Isomers of Saturated Hydrocarbons. *Bull. Chem. Soc. Jpn.* **1971**, *44*, 2332–2339.

(10) Wilson, R. J. *Introduction to Graph Theory*; Oliver & Boyd: Ediburgh, 1972; Chapter 2.

(11) Gordon, M.; Scantlebury, G. R. Non-Random Polycondensation: Statistical Theory of the Substitution Effect. *Trans. Faraday Soc.* **1964**, *60*, 604.

(12) Platt, J. R. Influence of Neighbor Bonds on Additive Bond Properties in Paraffins. *J. Chem. Phys.* **1947**, *15*, 419.

(13) Balabn, A. T.; Motoc, I.; Bonchev, D.; Mekenyan, O. Topological Indices for Structure–Activity Correlations. *Top. Curr. Chem.* **1983**, *114*, 21–55.

(14) Balaban, A. T. Chemical Graph: XXXIV. Five New Topological Indices for the Branching of Tree-Like Graphs. *Theor. Chim. Acta* **1979**, *53*, 355.

(15) Platt, J. R. Prediction of Isomeric Differences in Paraffin Properties. *J. Phys. Chem.* **1952**, *56*, 328.