Wiener-Number-Related Sequences

Douglas J. Klein† and Ivan Gutman*,‡

Texas A & M University at Galveston, Galveston, Texas 77553-1675, and Faculty of Science, University of Kragujevac, P.O. Box 60, YU-34000 Kragujevac, Yugoslavia

Received July 24, 1998

It is shown that several graph-distance-based structure-descriptors (among which are the recently introduced hyper-Wiener and Tratch−Stankevitch−Zefirov-type indices) are all mutually related and can be expressed in terms of distance distribution moments (i.e., moments for the distribution of shortest path intersite distances in the molecular graph).

1. INTRODUCTION

Recently various generalizations or extensions of the now "classical" Wiener number[1] have become of interest, both as regards new individual numbers[2−7] and as regards sequences of numbers.[8−12] Of particular interest here are those numbers of Tratch et al.[2] and of Randić,[3] their manner of interrelation, their formulas for trees as extended to general cycle-containing graphs, and their extension to a series of numbers, which turn out to be related to the moments of the distribution of distances.

The numbers of interest for the case of a tree graph T (with N vertices) may be obtained via the following algorithm.

0. For each pair of vertices i, j of T identify the unique path $\pi \equiv \pi(i, j)$ of length d_{ij} between i and j.

1. Remove from T all edges that belong to π and delete all vertices of π except i and j. Delete any side branches attached to these internal vertices of π. The remaining subgraph T_π is disconnected, possessing two components, with vertex counts of (say) a_π and b_π.

2. Form the sums (over the $N(N-1)/2$ pairs of vertices $\{i, j\}$).

$$W_n(T) \equiv \sum_{i<j} a_\pi \cdot (d_{ij})^n \cdot b_\pi$$

For example, for the (hydrogen-depleted) graph T_0 of 2-methylbutane, depicted in Figure 1, we have

vertex pair i, j	vertices of $\pi = \pi(i, j)$	$d(i, j)$	a_π	b_π
1, 2	1, 2	1	1	4
1, 3	1, 2, 3	2	1	2
1, 4	1, 2, 3, 4	3	1	1
1, 5	1, 2, 5	2	1	1
2, 3	2, 3	1	3	2
2, 4	2, 3, 4	2	3	1
2, 5	2, 5	1	4	1
3, 4	3, 4	1	4	1
3, 5	3, 2, 5	2	2	1
4, 5	4, 3, 2, 5	3	1	1

* To whom all correspondence should be addressed. E-mail: gutman@knez.uis.kg.ac.yu.
† A & M University.
‡ Faculty of Science.

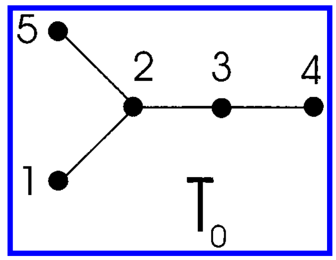

Figure 1. The hydrogen-depleted molecular graph of 2-methylbutane.

and therefore

$$W_n(T_0) = [1\times4+3\times2+4\times1+4\times1]1^n + [1\times2+1\times1+3\times1+2\times1]2^n + [1\times1+1\times1]3^n$$

i.e.,

$$W_n(T_0) = 18 + 8 \times 2^n + 2 \times 3^n$$

For T being a tree, the number $W_0(T)$ is that of Randić[3] (usually referred to as the hyper-Wiener index) while $W_1(T)$ is that of Tratch, Stankevitch, and Zefirov.[2] As a motivation for the relation to the Wiener number[1] one may note that if d_{ij} in the above formula is replaced by the (i, j)th element of the adjacency matrix of T, then the sum becomes restricted to edges and the Wiener number is obtained, independently of n, so that both indices of Randić[3] and of Tratch−Stankevitch−Zefirov[2] can be seen as individual number extensions, at least for trees. But, in fact, the whole sequence for $n = 0, 1, 2, ...$ is equally an extension.

Formally, the Wiener number of T corresponds to the limit value of $W_n(T)$ when $n \to -\infty$.

A third, less obvious, relation of the quantity $W_n(T)$ to the Wiener number will become evident in the subsequent discussion.

For cycle-containing connected graphs G Randić was silent about the definition of his $W_0(G)$ alternative, but Tratch et al. made a natural extension for $W_1(G)$ in terms of *geodesics* (which are shortest length paths between a pair of vertices). In the defining formula they simply replace the product $a_\pi \cdot b_\pi$ by a number $\#_{ij}(G)$, which they took to be the number of geodesics of G which contain both i and j. But here we choose $\#_{ij}(G)$ to be the number of pairs of vertices $\{p, q\}$ of

10.1021/ci980133s CCC: $18.00
Published on Web 04/09/1999

G such that there is a geodesic between p and q containing both i and j, this amounting to the same thing as Tratch et al. for the special case of trees. This extension is quite natural, and we imagine it applied for any n, so that

$$W_n(G) \equiv \sum_{i<j} \#_{ij}(G)\cdot(d_{ij})^n$$

This formula evidently includes an answer to Randić's implicit question as to the definition for general connected graphs, as has been a subject of some consideration.[9,13–15]

2. AN INTERRELATING FORMULA

The indicated definition is to be reformulated solely in terms of the distances d_{ij}. We first note that

$$W_n(G) = \sum_{i<j} \sum_{\{p,q\}}^{i,j} (d_{ij})^n$$

where the second sum is over all pairs of points which have a geodesic containing both i and j. Then upon reversal of the two sums one obtains

$$W_n(G) = \sum_{p<q} \sum_{i<j}^{\{p,q\}} (d_{ij})^n$$

where now the first sum is over all pairs of points $\{p, q\}$ of G and the second sum is over all pairs $\{i, j\}$ in some geodesic between p and q. It should be noted that the value taken by this second sum is independent of which geodesic one chooses between p and q.

Now, with the length of a geodesic g between p and q being d_{pq}, and noting that the number of subpaths of length d between two points of such a g is $d_{pq} - d + 1$, one can re-express the inner summation as

$$\sum_{i<j}^{\{p,q\}} (d_{ij})^n = \sum_{d=1}^{d_{pq}} (d_{pq} - d + 1)\cdot d^n$$

Thence

$$W_n(G) = \sum_{p<q} [(d_{pq} + 1)\cdot S_n(d_{pq}) - S_{n+1}(d_{pq})]$$

where we have introduced standard sum functions which are polynomials in their arguments

$$S_n(D) \equiv \sum_{d=1}^{D} d^n = \sum_{m=0}^{n} \gamma_{m,n}\cdot D^{n+1-m}$$

Here the coefficients $\gamma_{m,n}$ are given as follows[16]

$$\gamma_{0,n} = \frac{1}{n+1}$$

for $n \geq 0$

$$\gamma_{1,n} = \frac{1}{2}$$

for $n \geq 1$ whereas $\gamma_{1,0} = 0$

$$\gamma_{2k,n} = \frac{1}{2k}\binom{n}{2k-1} B_{2k}$$

and

$$\gamma_{2k+1,n} = 0$$

for $k \geq 1$.

Notice that $\gamma_{m,n} = 0$ whenever $m > n$. Besides, it is assumed that $\gamma_{m,n} = 0$ also for $m < 0$.

Hence,

$$S_0(D) = D$$

$$S_1(D) = \frac{1}{2}D^2 + \frac{1}{2}D$$

and for $n > 2$,

$$S_n(D) = \frac{1}{n+1}D^{n+1} + \frac{1}{2}D^n + \sum_{k=1}^{\lfloor n/2 \rfloor} \frac{1}{2k}\binom{n}{2k-1} B_{2k} D^{n+1-2k}$$

In the above formulas B_m stands for the mth Bernoulli number. Recall[16] that $B_0 = 1$ and that for $m \geq 1$ these numbers can be computed recursively by means of the relation

$$\sum_{k=0}^{n-1} \binom{n}{k} B_k = 0; \quad n \geq 2$$

Thus

$$B_1 = -\frac{1}{2} \quad B_2 = \frac{1}{6} \quad B_3 = 0$$

$$B_4 = -\frac{1}{30} \quad B_5 = 0 \quad B_6 = \frac{1}{42}$$

$$B_7 = 0 \quad B_8 = -\frac{1}{30} \quad B_9 = 0$$

$$B_{10} = \frac{5}{66} \quad B_{11} = 0 \quad B_{12} = -\frac{691}{2730}$$

etc. In the general case, $B_m = 0$ for odd m, $m \geq 3$, whereas for even m the B_m values alternate in sign.

Now defining *distance distribution moments* (i.e., moments for the distribution of shortest path intersite distances in G) as

$$M_m = \sum_{p<q} (d_{pq})^m$$

one obtains the Wiener number extensions $W_n(G)$ in terms of these moments

$$W_n(G) = \sum_{m=0}^{n+1} \Gamma_{m,n} M_m(G)$$

where

$$\Gamma_{m,n} = \gamma_{m-1,n} + \gamma_{m,n} - \gamma_{m,n+1}$$

The above m-summation then has no more than $n + 2$

(nonzero) terms. For the case $n = 0$ the suggested extension[9] of Randić's tree formula[3] to general connected graphs G is recovered, namely

$$W_0(G) = \frac{1}{2}M_2 + \frac{1}{2}M_1 \tag{1}$$

The respective expression for the Tratch−Stankevitch−Zefirov-type index reads

$$W_1(G) = \frac{1}{6}M_3 + \frac{1}{2}M_2 + \frac{1}{3}M_1 \tag{2}$$

For $n \geq 2$ a lengthy, but elementary calculation gives

$$\Gamma_{0,n} = \frac{1}{(n+1)(n+2)}$$

$$\Gamma_{1,n} = \frac{1}{n+1}$$

$$\Gamma_{2,n} = \frac{5}{12}$$

$$\Gamma_{m,n} = \frac{1}{m-1}\binom{n}{m-2}B_{m-1}$$

if m is odd, $m \geq 3$, and

$$\Gamma_{m,n} = -\frac{m-1}{n+2-m}\binom{n}{m-1}B_m$$

if m is even, $m \geq 4$. From this we obtain our final result, namely the explicit expression 3 for the generalization of the hyper-Wiener ($n = 0$) and Tratch−Stankevitch−Zefirov index ($n = 1$) to the case of $n \geq 2$:

$$W_n(G) = \frac{1}{(n+1)(n+2)}M_{n+2} + \frac{1}{n+1}M_{n+1} + \frac{5}{12}M_n + \sum_{k=1}^{\lfloor (n+1)/2 \rfloor} \frac{1}{2k}\binom{n}{2k-1}B_{2k}M_{n+1-2k} - \sum_{k=2}^{\lfloor (n+1)/2 \rfloor} \frac{2k-1}{n+2-2k}\binom{n}{2k-1}B_{2k}M_{n+2-2k} \tag{3}$$

The previously known[9] formula 1 for the hyper-Wiener index, the formula 2 for the previously introduced Tratch−Stankevitch−Zefirov-type index,[2] and their generalization formula 3 are the main results of this work. The general formula 3 relates the index $W_n(G)$ to the moment distributions, which themselves (in a "normalized" form) have been advocated by Randić[8] as relevant descriptors. Applications of the formulas 1−3 were recently reported elsewhere.[17]

ACKNOWLEDGMENT

D.J.K. acknowledges support of the Welch Foundation of Houston, Texas. I.G. thanks the Institute for Advanced Studies of the Hebrew University of Jerusalem and the Mathematical Institute in Belgrade for financial support.

REFERENCES AND NOTES

(1) Wiener, H. Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc.* **1947**, *69*, 17−20.
(2) Tratch, S. S.; Stankevitch, M. I.; Zefirov, J. Combinatorial Models and Algorithms in Chemistry. The Expanded Wiener Number−A Novel Topological Index. *J. Comput. Chem.* **1990**, *11*, 899−908.
(3) Randić, M. Novel Molecular Descriptor for Structure−Property Studies. *Chem. Phys. Lett.* **1993**, *211*, 478−483.
(4) Klein, D. J.; Randić, M. Resistance Distance. *J. Math. Chem.* **1993**, *12*, 81−95.
(5) Khadikar, P. V.; Deshpande, N. V.; Kale, P. P.; Dobrynin, A.; Gutman, I.; Dömötör, G. The Szeged Index and an Analogy with the Wiener Index. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 547−550.
(6) Gutman, I.; Mohar, B. The Quasi-Wiener and the Kirchhoff Indices Coincide. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 982−985.
(7) Lukovits, I. An All-Path Version of the Wiener Index. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 125−129.
(8) Randić, M. Molecular Shape Profiles. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 373−382.
(9) Klein, D. J.; Lukovits, I.; Gutman, I. On the Definition of the Hyper-Wiener Index for Cycle-Containing Structures. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 50−52.
(10) Zhu, H. Y.; Klein, D. J.; Lukovits, I. Extensions of the Wiener Number. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 420−428.
(11) Gutman, I. Buckley-Type Relations for Wiener-Type Structure-Descriptors. *J. Serb. Chem. Soc.* **1998**, *63*, 487−492.
(12) Estrada, E.; Ivanciuc, O.; Gutman, I.; Gutierrez, A.; Rodríguez, L. Extended Wiener Indices. A New Set of Descriptors for Quantitative Structure−Property Studies. *New J. Chem.* In press.
(13) Randić, M.; Guo, X.; Oxley, T.; Krishnapriyan, H. Wiener Matrix: Source of Novel Graph Invariants. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 709−716.
(14) Randić, M.; Guo, X.; Oxley, T.; Krishnapriyan, H.; Naylor, L. Wiener Matrix Invariants. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 361−367.
(15) Lukovits, I.; Linert, W. A Novel Definition of the Hyper-Wiener Index for Cycles. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 899−902.
(16) The expansion of $S_n(D)$ into a polynomial in the variable D is a classical mathematical result, due to Jacob Bernoulli (early XVIII century). For further details see, for instance: *Encyclopaedia of Mathematics*; Reidel: Dordrecht, 1987; Vol. 1, p 360.
(17) Gutman, I.; Vidović, D.; Popović, L. On Graph Representation of Organic Molecules−Cayley's Plerograms vs His Kenograms. *J. Chem. Soc., Faraday Trans.* **1998**, *94*, 857−860.

CI980133S