

The Detour Index

István Lukovits

Central Research Institute for Chemistry, Hungarian Academy of Sciences,
H-1525 Budapest, P.O.Box 17, Hungary

Received October 20, 1995; revised December 14, 1995; accepted January 17, 1996

The detour index w is the sum of all entries in the upper triangle of the detour matrix Δ , where the ij -th entry Δ_{ij} denotes the length of the longest path between vertices i and j of the underlying graph. The Wiener index W and w are equivalent for acyclic structures, but are different in cycle-containing graphs. Similarly the hyper-Wiener index Γ and its analogue the hyper-detour index γ are equivalent in acyclic structures and are different in cyclic structures. Several formulas were derived for w and γ . Excellent correlations between boiling points and the composite indices $(Ww)^{1/8}$ and $(\Gamma\gamma)^{1/10}$ were found in a series consisting of 77 cyclic and acyclic alkanes.

INTRODUCTION

The detour index w was defined by Amić and Trinajstić.¹ w is the sum of the upper triangle of the detour matrix² Δ , the i,j -th entry Δ_{ij} denotes the longest path between vertices i and j of the underlying graph ($i,j = 1, 2, \dots, N$, where N denotes the number of vertices):

$$w = \sum \Delta_{ij} \quad (1)$$

and the summation has to be performed for all distinct pairs of vertices. Example: cyclobutane. Carbons are considered only, and the atoms are numbered consecutively. Then $\Delta_{12} = 3$, $\Delta_{13} = 2$, $\Delta_{14} = 3$, $\Delta_{23} = 3$, $\Delta_{24} = 2$, $\Delta_{34} = 3$, and $w = 3 + 2 + 3 + 3 + 2 + 3 = 16$. Note that the definition of w is similar to that of well known Wiener index³ W . Note that in order to obtain W , Δ_{ij} has to be replaced by distances d_{ij} in Eq. (1). W and w are equivalent in acyclic structures, because here $d_{ij} = \Delta_{ij}$ for all i and j , but they are necessarily

different in cyclic structures – except for non relevant cases involving cycles with just two vertices. w may account for structural characteristics of alkanes and may be used to derive quantitative structure-property relationships (QSPR), but this feature of w was not investigated by Amić and Trinajstić.¹

W has been investigated many times in QSPRs – in fact it is one of the most »popular« graph theoretical invariants. Two recent reviews^{4,5} summarize important results obtained thus far. Very high correlation between W and boiling points of acyclic was demonstrated by Trinajstić *et al.*⁶ In 1976 Rouvray and Crafford suggested an index that is practically equivalent – except for a factor of two – with the Wiener index.⁷ A typical weakness of the Wiener index – and in fact of most topological indices proposed by now – is that no significant correlation can be detected between W and any physical property of the compounds in series consisting of acyclic and cycle-containing structures. Partition coefficients⁸ and retention⁹ indices seem to be exceptions. A variant of the Wiener index, the hyper-Wiener index Γ has been proposed for QSPR studies by Randić,¹⁰ and this index has also been found to be a good parameter in a series of acyclic hydrocarbons.¹¹

The concept of the hyper-Wiener index was extended to define the hyper-detour index γ . The purpose of this paper was to investigate QSPRs involving w and γ in a series of acyclic and cycle-containing hydrocarbons. It was found that composite indices involving W and w , or Γ and γ , are the best parameters. The structure of this paper is as follows. In the next section several formulas for w and γ will be derived, the third section summarizes the QSPR equations, and the last section is the conclusion.

DERIVATION OF FORMULAS

Terms »graph« and »structural formula«, »edge« and »chemical bond« and »vertex« and »atom« will be used interchangeably in this paper. Hydrogen suppressed graphs will be considered hereafter. Those who are not familiar with the concepts of chemical graph theory may consult a textbook¹² on this topics. The review on the application of distance matrix in chemistry may also be helpful.¹³

In a simple cycle for each entry Δ_{ij} we can write:

$$\Delta_{ij} = N - d_{ij} \quad (2)$$

where N denotes the number of vertices and d_{ij} is the distance between vertices i and j . Using the sum rules for finite series and the formulas¹⁴ used to obtain W in simple cycles, we obtain from Eq. (2) that:

$$w = (3N^3 - 4N^2)/8 \quad (N \text{ is even}) \tag{3}$$

and

$$w = (3N^3 - 4N^2 + N)/8 \quad (N \text{ is odd}) . \tag{4}$$

Analogously with the definition of the connectedness index s_i by Seybold¹⁵ we can define the »detour-connectedness« index S_i ($i = 1, 2, \dots, N$):

$$S_i = 2w/N = (3N^2 - 4N)/4 \quad (N \text{ is even}) \tag{5}$$

and

$$S_i = 2w/N = (3N^2 - 4N + 1)/4 \quad (N \text{ is odd}) . \tag{6}$$

Eqs. (5) and (6) may be used to obtain formulas for cycles with a single side-chain (Figure 1.). The cycle contains k vertices, and the side chain contains

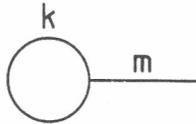


Figure 1. A scheme depicting a cycle with single side chain. $N = K + m - 1$.

m vertices and the branching vertex is considered twice, therefore $N = k + m - 1$. For any cutpoint graph, W and similarly w may be obtained from values W_k, s_k, W_m, s_m, k and m by the following formula:¹⁴

$$W = W_k + W_m + (k - 1)s_m + (m - 1)s_k \tag{7}$$

and

$$w = w_k + w_m + (k - 1)S_m + (m - 1)S_k . \tag{8}$$

Therefore for graphs depicted in Figure 1. By using Eqs. (5), (6) and (8) and the formula¹⁴ of the connectedness of a chain $S_m = s_m = (m^2 - m)/2$, we obtain:

$$w = (9k^3 + 4m^3 + 18k^2m + 12km^2 - 30k^2 - 12m^2 - 36km + 24k + 8m)/24 \quad (k \text{ is even}) \tag{9}$$

and

$$w = (9k^3 + 4m^3 + 18k^2m + 12km^2 - 30k^2 - 12m^2 - 36km + 27k + 14m - 6)/24 \quad (k \text{ is odd}) . \quad (10)$$

Eqs. (9) and (10) are equivalent with Eqs. (3) and (4), respectively, if $m = 1$.

The definition of the hyper-Wiener index Γ , proposed by Randić¹⁰ can also be extended to include the detour matrix Δ . For this purpose the formula derived for Γ has to be used:¹⁶

$$\Gamma = \Sigma (d_{ij}^2 + d_{ij})/2 \quad (11)$$

where the summation has to be performed for all pairs of vertices i and j . Replacing d_{ij} by Δ_{ij} in Eq. (11), we immediately obtain:

$$\gamma = \Sigma (\Delta_{ij}^2 + \Delta_{ij})/2 \quad (12)$$

γ will be denoted by the term »hyper-detour index« hereafter. The hyper-detour index of simple cycles is:

$$\gamma = (7N^4 - 3N^3 - 10N^2)/48 \quad (N \text{ is even}) \quad (13)$$

and

$$\gamma = (7N^4 - 3N^3 - 7N^2 + 3N)/48 \quad (N \text{ is odd}) . \quad (14)$$

Note that Eqs. (4) and (14) yield for methane ($N = 1$) $w_{\text{methane}} = 0$ and $\gamma_{\text{methane}} = 0$, respectively. It has to be mentioned that the derivation of numerical formulas for the detour index w and the hyper-detour index γ has yet to be done.

APPLICATIONS

A series consisting of acyclic and cyclic alkanes were considered. The total number of compounds considered (n) was 77. The experimental boiling points (B.P.) of acyclic hydrocarbons were the same considered by Trinajstić *et al.*,⁶ Balaban *et al.*¹⁷ Lukovits and Linert.¹¹ All acyclic structures up to $N = 8$ (*i.e.* octanes) were considered including methane.¹⁸ Acyclic nonanes and decanes were not considered in order to ensure a well balanced sample of acyclic and cyclic compounds. The boiling points of cycle-containing alkanes were collected from a Handbook¹⁸ up to $N = 10$. The data are listed in Table I. Acyclic n -alkanes are denoted by NX , where X denotes the number of carbons in the chain. Therefore $N2$ denotes ethane, $N4$ denotes butane, *etc.*

TABLE I

Boiling Points (°C) of 77 Acyclic and Cyclic Alkanes and Topological Indices

Molecule	B.P.	W	w	Γ	γ
N1	-164.0	0	0	0	0
N2	-88.5	1	1	1	1
N3	-44.5	4	4	5	5
N4	-0.5	10	10	15	15
2M-N3	-10.5	9	9	12	12
N5	36.5	20	20	35	35
2M-N4	27.9	18	18	28	28
2,2MM-N3	9.5	16	16	22	22
N6	68.7	35	35	70	70
3M-N5	63.2	31	31	54	54
2M-N5	60.2	32	32	58	58
2,3MM-N4	58.1	29	29	47	47
2,2MM-N4	49.7	28	28	44	44
N7	98.4	56	56	126	126
3E-N5	93.5	48	48	90	90
3M-N6	91.8	50	50	99	99
2M-N6	90.0	52	52	108	108
2,2MM-N5	89.8	46	46	84	84
3,3MM-N5	86.0	44	44	76	76
2,4MM-N5	80.5	48	48	91	91
2,2MM-N5	79.2	46	46	84	84
2,2,3MMM-N4	80.9	42	42	69	69
C4	12.0	8	16	10	30
E-C4	70.7	29	45	49	101
M-C4	36.3	16	28	23	57
C7	118.5	42	105	70	322
C6	80.7	27	63	42	168
B-C6	181.6	133	217	323	597
<i>i</i> -P-C6	171.3	88	156	176	464
<i>s</i> -B-C6	179.3	121	205	263	583
<i>t</i> -B-C6	171.5	114	198	233	611
1M-4E-C6	150.0	90	162	187	487
P-C6	156.7	94	166	203	457
1,1,3MMM-C6	139.0	82	158	152	462
1,2,5MMM-C6	139.5	84	162	160	481
C8	148.5	64	160	120	552
C5	49.2	15	35	20	80
1,2EE-C5	153.6 ⁺⁺	87	151	171	412
1,1MM-C5	87.5	39	75	61	185
1,2MM-C5	95.5 ⁺	40	79	64	203
E-C7	103.5	43	79	75	207
M-C5	71.8	26	54	39	131
1M-2P-C5	152.6 ⁺	90	151	185	416
C3	28.7	3	6	3	9

TABLE I
 (continuing)

Molecule	B.P.	W	w	Γ	γ
1,1MM-C3	20.6	15	22	20	38
1,2MM-C3	33.0	16	24	23	45
E-C3	34.5	17	24	26	46
M-C3	4.5	8	13	10	22
1,1,2MMM-C3	52.6	26	37	39	71
N8	125.8	84	84	210	210
3M-N7	118.8	76	76	170	170
3E-N6	118.9	72	72	150	150
3,4MM-N6	118.7	68	68	134	134
3E-3M-N5	118.2	64	64	118	118
4M-N7	117.7	75	75	165	165
2M-N7	117.6	79	79	185	185
3E-2M-N5	115.6	67	67	129	129
2,3MM-N6	115.3	70	70	143	143
2,4MM-N6	109.4	71	71	147	147
3,3MM-N6	112.0	67	67	131	131
2,5MM-N6	108.4	74	74	161	161
2,3,4MMM-N5	113.4	65	65	122	122
2,3,3MMM-N5	114.6	62	62	111	111
2,2MM-N6	107.0	71	71	149	149
2,2,3MMM-N5	110.5	63	63	115	115
2,2,4MMM-N5	99.3	66	66	127	127
2,2,3,3MMMM-N4	106.0	58	58	97	97
1,1MM-C6	119.5	59	119	103	337
1,2MM-C6	126.6 ⁺	60	126	106	362
1,3MM-C6	122.3 ⁺	61	123	110	355
1,4MM-C6	121.8 ⁺	62	122	115	349
cyclohexane, 1,2-dimethylene	124.0	58	160	101	552
E-C6	131.8	64	124	122	368
P-C5	131.0	67	111	135	315
<i>i</i> -P-C5	126.4	62	106	114	286
1,1,2MMM-C5	113.5	56	106	92	278
1,1,3MMM-C5	115.5	58	104	100	266

⁺ Mean value measured for *cis* and *trans* isomers

⁺⁺ *trans* isomer.

Similarly cycloalkanes are denoted by symbols CX, C3 denotes cyclopropane, C4 denotes cyclobutane. Alkyl substituents are denoted in accordance with the notation used by Balaban *et al.*¹⁷ by letters M, E, P and B for methyl, ethyl, propyl and butyl, respectively.

TABLE II

Correlations between Topological Indices and Boiling Points of 77 Acyclic and Cycle-Containing Alkanes and 29 Octanes

No.	Index	R Alkanes	R Octanes
1.	N	0.977	—
2.	$N^{1/2}$	0.986	—
3.	$W^{1/3}$	0.977	0.060
4.	$W^{1/4}$	0.979	0.060
5.	$w^{1/3}$	0.961	0.747
6.	$w^{1/4}$	0.977	0.744
7.	$\Gamma^{-1/4}$	0.967	0.314
8.	$\Gamma^{-1/5}$	0.969	0.316
9.	$\gamma^{1/4}$	0.945	0.755
10.	$\gamma^{1/5}$	0.957	0.754
11.	$\gamma^{1/6}$	0.960	0.753
12.	$(Ww)^{1/6}$	0.990	0.792
13.	$(Ww)^{1/7}$	0.994	0.791
14.	$(Ww)^{1/8}$	0.994	0.790
15.	$(\Gamma\gamma)^{1/8}$	0.989	0.776
16.	$(\Gamma\gamma)^{1/10}$	0.991	0.774
17.	$(Ww\Gamma\gamma)^{1/14}$	0.986	0.784
18.	$(Ww\Gamma\gamma)^{1/16}$	0.988	0.782

Table II. lists several correlation coefficients between B.P.-s and various topological indices. The variation of the boiling points of alkanes are explained to a great extent by the number of carbons (N) in each molecule.^{6,17} Often it is quite difficult to find an index with a higher correlation coefficient than obtained by using N . N was used determine the value of the lowest acceptable correlation coefficient in the series of molecules (Table I). Table II indicates that $N^{1/2}$ – more exactly $N^{0.498\pm 0.063}$ – was the best of all B.P. = $AN^B + C$ type equations with $R = 0.98604$, were A, B and C are regression parameters. The correlation coefficient calculated between B.P. and $w^{1/4}$ is significant at the level p 0.01, but $w^{1/4}$ alone could not be used for making predictions. None of the single indices $W^{1/4}$, $w^{1/4}$, $\Gamma^{-1/5}$ and $\gamma^{1/5}$ did surpass $N^{1/2}$ in this respect. Higher values of R were obtained with composite indices (entries 12–18 in Table II), and the respective regression equations are listed in Table III. The best results were obtained by using composite indices $(Ww)^{1/8}$ and $(\Gamma\gamma)^{1/10}$ (Table III). Linear combinations of $W^{1/4}$ and $w^{1/4}$ and also $G^{1/5}$ and $g^{1/5}$ also yielded significant higher correlation coefficients than $N^{1/2}$ did:

TABLE III
Selected Regression Equations ⁺ of Type B.P. = AX + B (*n* = 77)

Eq.	X	A	B	s	R	F
1.	$N^{1/2}$	151(±3)	-306	10.1	0.986	2630
2.	$(Ww)^{1/6}$	58.6(1.0)	-129.3	8.6	0.990	3594
3.	$(Ww)^{1/8}$	100.0(1.2)	-176.8	6.4	0.994	6553
4.	$(\Gamma\gamma)^{1/8}$	73.1(1.3)	-142.8	8.9	0.989	3352
5.	$(\Gamma\gamma)^{1/10}$	107.9(1.6)	-180.9	7.9	0.991	4338

⁺ The numbers in parentheses denote the standard errors of the regression coefficients.

$$\text{B.P.} = 55.8(3.7)W^{1/4} + 44.7(3.1)w^{1/4} - 177.4 \quad (15)$$

$$n = 77, s = 6.4, R = 0.994, F = 3266$$

and

$$\text{B.P.} = 66.4(3.7)\Gamma^{1/5} + 43.5(2.9)w^{1/5} - 183.5 \quad (16)$$

$$n = 77, s = 7.6, R = 0.992, F = 2321 .$$

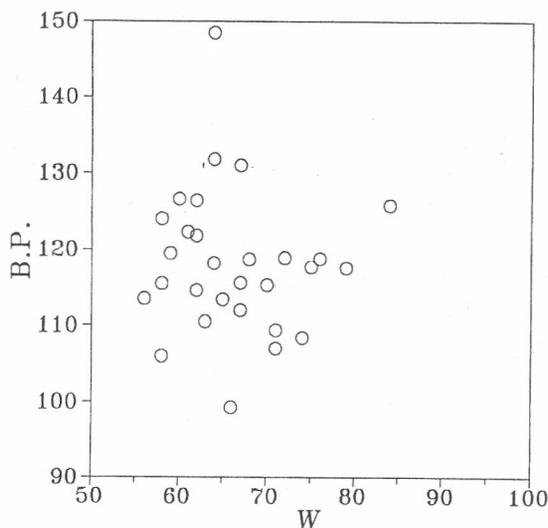


Figure 2. Correlation between boiling points and the Wiener index *W* of acyclic and cyclic octane isomers.

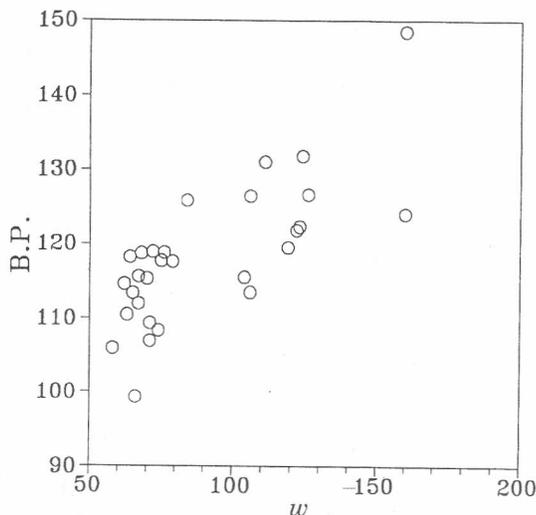


Figure 3. Correlation between boiling points and the detour index w of acyclic and cyclic octane isomers.

It seems that W and w (and similarly Γ and γ) together are adequate descriptors of the topological property of alkanes explaining the variation of the boiling points in the series considered.

The correlation coefficients obtained for the subseries of (acyclic and cyclic) octanes ($n = 29$) are also listed in Table II. It can be seen that the correlation coefficient between B.P. and W is practically equal to zero ($R = 0.057$) in the octane series (Figure 2.). Note that in the series of acyclic octanes¹⁹ R is equal to 0.5 only. However, there is a moderate, although significant ($p < 0.01$) correlation coefficient ($R = 0.747$) between w and B.P. in the same series (Figure 3.). The corresponding correlation coefficients obtained by using Γ and γ were 0.025 and 0.759, respectively.

CONCLUSIONS

The combined indices $(Ww)^{1/8}$ and $(\Gamma\gamma)^{1/10}$ seem to be useful descriptors of molecular properties of series consisting of acyclic and cyclic structures. The detour index w and the hyper-detour index γ alone are less efficient in QSPRs. Moderate correlation coefficients were detected between w and γ and the B.P.-s, whereas W and Γ failed show any correlation in a series composed of acyclic and cyclic octane isomers.

The method indicated in this paper allows to derive w for cutpoint graphs. No numerical method is available to calculate w and Γ numerically thus far. This problem is highly important which has to be solved in the future.

REFERENCES

1. D. Amić and N. Trinajstić, *Croat. Chem. Acta* **68** (1995) 53.
2. F. Buckley and F. Harary, *Distance in Graphs*, Addison-Wesley, Redwood City, CA 1990, p. 213.
3. (a) H. Wiener, *J. Am. Chem. Soc.* **69** (1947) 17. (b) H. Wiener, *J. Am. Chem. Soc.* **69** (1947) 2636.
4. I. Gutman, Y. N. Yeh, S. L. Lee and Y. L. Luo, *Indian J. Chem.* **32A** (1993) 651.
5. S. Nikolić, N. Trinajstić, and Z. Mihalić, *Croat. Chem. Acta* **68** (1995) 105.
6. N. Trinajstić, D. Babić, S. Nikolić, D. Plavšić, D. Amić, and Z. Mihalić, *J. Chem. Inf. Comput. Sci.* **34** (1994) 368.
7. D. Rouvray and B. C. Crafford, *S. Afr. J. Sci.* **72** (1976) 47.
8. I. Lukovits, *Reports Mol. Theor.* **1** (1990) 127.
9. N. Bošnjak, N. Aldler, M. Perić, and N. Trinajstić, in: Z. B. Maksić (Ed.), »*Modelling of Structure and Properties of Molecules*«, Ellis Horwood, Chichester, 1987, p. 103.
10. M. Randić, *Chem. Phys. Lett.* **211** (1993) 478.
11. I. Lukovits and W. Linert, *J. Chem. Inf. Comput. Sci.* **34** (1994) 899.
12. N. Trinajstić, *Chemical Graph Theory*, 2nd Ed., CRC, Boca Raton, 1992, p. 1.
13. Z. Mihalić, D. Veljan, D. Amić, S. Nikolić, D. Plavšić, and N. Trinajstić, *J. Math. Chem.* **11** (1992) 223.
14. I. Lukovits, *J. Chem. Inf. Comput. Sci.* **31** (1991) 503.
15. P. G. Seybold, *Int. J. Quantum Chem.: Quantum Biol. Symp.* **10** (1983) 103.
16. D. J. Klein, I. Lukovits, and I. Gutman, *J. Chem. Inf. Comput. Sci.* **35** (1995) 50.
17. A. T. Balaban, D. Ciubotariu, M. Medeleanu, *J. Chem. Inf. Comput. Sci.* **31** (1991) 517.
18. *Handbook of Chemistry and Physics*, 70th Ed., CRC-Press, Boca Raton, 1990, p. C-222.
19. E. C. Kirby, *J. Chem. Inf. Comput. Sci.* **34** (1994) 1030.

SAŽETAK

Indeks zaobilaznih udaljenosti

István Lukovits

Indeks zaobilaznih udaljenosti, w , jednak je sumi elemenata gornjeg trokuta matrice zaobilaznih udaljenosti, čiji element d_{ij} označuje duljinu najdulje staze između čvorova i i j pripadnog grafa. Wienerov indeks, W , i w ekvivalentni su za acikličke strukture, ali su različiti u grafovima s prstenom. Slično, hiper-Wienerov indeks i njegov analog, hiperindeks zaobilaznih udaljenosti, ekvivalentni su za acikličke strukture a različiti za cikličke strukture. Izvedeno je nekoliko formula za w i γ . Za niz oktana cikličke i acikličke strukture pokazana je umjerena korelacija njihovih vrelišta sa w i γ . Odlične korelacije između vrelišta i kompozitnih indeksa $(Ww)^{1/8}$ i $(\Gamma\gamma)^{1/10}$ nađene su za niz od 77 alkana.