Journal of Mathematical Chemistry Vol. 33, No. 2, February 2003 (© 2003)

The overall hyper-Wiener index

X.H. Li and J.J. Lin

Department of Chemistry, Wenzhou Normal College, Wenzhou, Zhejiang 325027, P.R. China

Received 21 March 2002

The concept of overall connectivity of a graph G was extended here to the definition of the overall hyper-Wiener index $OWW(G)$ of a graph G, defined as the sum of the hyper-Wiener indexes in all subgraphs of G, as well as the sum of eth-order terms, ${}^eOWW(G)$, with e being the number of edges in the subgraph. The potential usefulness of the overall hyper-Wiener index in QSAR/QSPR is evaluated by its correlation with a number of properties of C3-C8 alkanes and cycloalkanes.

KEY WORDS: overall hyper-Wiener index, hyper-Wiener index, QSAR/QSPR

1. Introduction

The hyper-Wiener index (WW) was proposed by Randic et al. [1], as a generalization of the much studied Wiener index (W) of graph invariant [2–4]. Recall that W is defined as the sum of distances between pairs of vertices of the graph under study. Randic et al. generalized W to WW by considering paths p instead of edges e in the Wiener index. In his definition, WW is the sum of contributions K_{ij}, where subscripts i and j denote a pair of vertices. K_{ij} may be obtained [5,6] by using the following two-step algorithm:

(1) Remove the path (for trees there is only one path) between vertices i and j, whence two nonconnected subgraphs are obtained.

(2) Multiply the number of vertices of the first subgraph by the number of vertices of the second one.

There are $N(N-1)/2$ terms K_{ij} altogether, where N denotes the number of vertices (i.e., the number of carbons in H-suppressed graphs).

Klein et al. [7] provide an interpretation of WW extending beyond trees. They show that

$$WW = \frac{1}{2}\left[\sum_{i<j} d_{ij}^2 + \sum_{i<j} d_{ij}\right] = \frac{1}{2}\sum_{i<j} d_{ij}^2 + \frac{1}{2}W \tag{1}$$

agrees with Randic's definition in the case of trees. Thus, the hyper-Wiener index is proportional to the Wiener index plus the sum of squared distances between pairs of vertices. Roberto et al. [8] propose an algorithm with a complexity linear in the number of

0259-9791/03/0200-0081/0

vertices for the computation of the hyper-Wiener index of chemical trees. Lukovits et al. [9] define a novel hyper-Wiener index for cycles. As the hyper-Wiener index measures the extensiveness of a graph, their definition produces values for cyclic structures that are lower than those for corresponding chains. But in fact, some properties of cyclic structures are sometimes higher than their n-alkane counterparts, for example, boiling points. So they said that there should be a further factor that is not accounted for by hyper-Wiener index WW. This study aims to make up the pitfalls of the hyper-Wiener number while preserving the best of its features.

2. The overall hyper-Wiener index

Recently, Bonchev proposes a new approach to the topological characterization of molecules [10–12]. It proceeds from the total number of subgraphs $K(G)$ of a molecular graph G, and their presentation uses complete series of classes of various orders e, which is the number of edges in each subgraph of a certain class. Here $e = 0, 1, 2, \ldots E$, with E being the number of edges in the entire graph. He used the Wiener numbers of each subgraph, extending thus that seminal molecular descriptor to its most complete version, the overall Wiener index [13]. In this study, we use the hyper-Wiener indices of each subgraph, extending this molecular descriptor to overall hyper-Wiener index.

Definition 1. The overall hyper-Wiener index $OWW(G)$ of any graph G is defined as the sum of the hyper-Wiener indices $WW_i(G_i)$ of all K connected subgraphs of G:

$$OWW(G) = \sum_{i=1}^{K} W_i(G_i \subset G). \tag{2}$$

For the hyper-Wiener index of cycles we adopt the definition of Lukovits [9].

Definition 2. The eth-order overall hyper-Wiener index ${}^eOWW(G)$ of any graph G is defined as the sum of the hyper-Wiener indices $WW_j({}^eG_j)$ of all ${}^eK$ subgraphs ${}^eG_j \subset G$ that have e edges:

$${}^eOWW(G) = \sum_{j=1}^{{}^eK} WW_j\big({}^eG_j \subset G\big). \tag{3}$$

Definition 3. The eth-order overall hyper-Wiener index ${}^eOWW(G)$ can be presented as a sum of terms, ${}^eOWW_k(G)$, representing the sum of the hyper-Wiener indices in the subgraphs of specified type. For acyclic graphs these are the path ($k = \mathrm{p}$), cluster ($k = \mathrm{c}$), and pathcluster ($k = \mathrm{pc}$) type, as defined by Kier and Hall [14]:

$$\begin{aligned} {}^eOWW(G) &= {}^eOWW_{\mathrm{p}}(G) + {}^eOWW_{\mathrm{c}}(G) + {}^eOWW_{\mathrm{pc}}(G) \\ &= \sum_{j=1}^{{}^eK_{\mathrm{p}}} WW_{\mathrm{p}j} + \sum_{l=1}^{{}^eK_{\mathrm{p}}} WW_{\mathrm{c}l} + \sum_{m=1}^{{}^eK_{\mathrm{p}}} WW_{\mathrm{pc}m}. \end{aligned} \tag{4}$$

Definition 4. The overall hyper-Wiener index vector $OWW'(G)$ of any graph G is the sequence of all ${}^{e}OWW(G)$'s listed in an ascending order of the number of edges e:

$$OWW'(G) = OWW\{{}^{0}OWW,\ {}^{1}OWW,\ {}^{2}OWW,\ \ldots,\ {}^{E}OWW\}, \tag{5}$$

or in more detail for acyclic graphs:

$$OWW'(G) = OWW\{{}^{0}OWW,\ {}^{1}OWW,\ {}^{2}OWW,\ {}^{3}OWW_{\mathrm{p}},\ {}^{3}OWW_{\mathrm{c}}, \ldots, {}^{E}OWW_{\mathrm{p}},\ {}^{E}OWW_{\mathrm{c}},\ {}^{E}OWW_{\mathrm{pc}}\}. \tag{6}$$

The Eth-order overall hyper-Wiener index, ${}^{E}OWW(G)$, is the hyper-Wiener index $WW(G)$ itself:

$${}^{E}OWW(G) = WW(G). \tag{7}$$

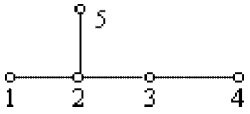
The zero-order overall hyper-Wiener index, ${}^{0}WW(G)$, is equal to zero:

$${}^{0}OWW(G) = 0. \tag{8}$$

The first-order overall hyper-Wiener index, ${}^{1}OWW(G)$, is equal to the number of graph edges E:

$${}^{1}OWW(G) = E. \tag{9}$$

Example.



$e = 1$:	1–2, 2–3, 3–4, 2–5;	${}^{1}OWW = 4 \times 1 = 4$,
$e = 2$:	1–2–3, 2–3–4, 1–2–5, 5–2–3;	${}^{2}OWW = 4 \times 5 = 20$,
$e = 3$:	1–2–3–4, 5–2–3–4, 1–2–3–5;	${}^{3}OWW = 2 \times 15 + 12 = 42$,
$e = 3$:	the entire graph;	${}^{4}OWW = 28$,
$OWW = 4 + 20 + 42 + 28 = 94$;		$OWW' = 94(4, 20, 42, 28)$.

3. Quantitative structure–property relationships with overall hyper-Wiener index

In this study, we apply multiple regression analysis (MRA) in a QSAR modeling of seven physical properties of 44 alkanes and cycloalkanes for C3–C8 alkanes [12]. The properties are: boiling point, BP; molecular volume, MV; molecular refraction, MR; heat of vaporization, HV; surface tension, ST; critical temperature, TC; and critical pressure, PC (table 1) [15,16]. Two different sets of overall hyper-Wiener indices were used in the modeling, the first one includes the OWW index and its eth-order components, ${}^{e}OWW$, for $e = 1, 2, \ldots, 7$ (table 2). The extended set of overall hyper-Wiener indices includes also the path (p), cluster (c), and pathcluster (pc) terms of order 3–7 (table 3).

Table 1
Experimental values for the physical properties of the 44 alkanes.

No.	Molecule[a]	BP	MV	MR	HV	TC	PC	ST
1	C3	−42.070				96.80	42.01	
2	nC4	−0.500				152.01	37.47	
3	2MC3	−11.730				134.98	36.00	
4	nC5	36.074	115.205	25.2656	26.42	196.62	33.31	16.00
5	2MC4	27.852	116.426	25.2923	24.59	187.80	32.90	15.00
6	22MMC3	9.503	122.074	25.7243	21.78	160.60	31.57	
7	nC6	68.740	130.688	29.9066	31.55	234.70	29.92	18.42
8	2MC5	60.271	131.933	29.9459	29.86	224.90	29.95	17.38
9	3MC5	63.282	129.717	29.8016	30.27	231.20	30.83	18.12
10	23MMC4	57.988	130.240	29.8104	29.12	227.10	30.99	17.37
11	22MMC4	49.741	132.744	29.9347	27.69	216.20	30.67	16.30
12	nC7	98.427	146.540	34.5504	36.55	267.01	27.01	20.26
13	2MC6	90.052	147.656	34.5908	34.80	257.90	27.20	19.29
14	3MC6	91.850	145.821	34.4597	35.08	262.40	28.10	19.79
15	3EC5	93.475	143.517	34.2827	35.22	267.60	28.60	20.44
16	24MMC5	80.500	148.949	34.6192	32.88	247.10	27.40	18.15
17	23MMC5	89.784	144.153	34.3237	34.24	264.60	29.20	19.96
18	22MMC5	79.197	148.695	34.6166	32.43	247.70	28.40	18.02
19	33MMC5	86.064	144.530	34.3323	33.02	263.00	30.00	19.59
20	223MMMC4	80.882	145.191	34.3736	32.04	258.30	29.75	18.76
21	nC8	125.665	162.592	39.1922	41.48	296.20	24.64	21.76
22	2MC7	117.647	163.663	39.2316	39.68	288.00	24.80	20.60
23	3MC7	118.925	161.832	39.1001	39.83	292.00	25.60	21.17
24	4MC7	117.709	162.105	39.1174	39.67	290.00	25.60	21.17
25	3EC6	118.534	160.072	38.9441	39.40	292.00	25.74	21.51
26	25MMC6	109.103	164.697	39.2596	37.86	279.00	25.00	19.73
27	24MMC6	109.429	163.093	39.1300	37.76	282.00	25.80	20.05
28	22MMC6	106.840	164.285	39.2525	37.29	279.00	25.60	19.60
29	23MMC6	115.607	160.395	38.9808	38.79	293.00	26.60	20.99
30	34MMC6	117.725	158.814	38.8453	39.02	298.00	27.40	21.64
31	2M3EC5	115.650	158.794	38.8362	38.52	295.00	27.40	21.52
32	33MMC6	111.969	160.879	39.0087	37.93	290.84	27.20	20.63
33	3M3EC5	118.259	157.026	38.7171	37.99	305.00	28.90	21.99
34	224MMMC5	99.238	165.083	39.2617	35.13	271.15	25.50	18.77
35	234MMMC5	113.467	158.852	38.8681	37.61	295.00	27.60	21.14
36	223MMMC5	109.841	159.526	38.9249	36.91	294.00	28.20	20.67
37	233MMMC5	114.760	157.292	38.7617	37.22	303.00	29.00	21.56
38	2233MMMMC4	106.470				270.87	24.50	
39	Cyclopropane	−32.800						
40	Cyclobutane	12.600						
41	Cyclopentane	49.300						
42	Cyclohexane	80.700						
43	Cycloheptane	118.400						
44	Cyclooctane	149.000						

[a] M and E stand for methyl and ethyl, respectively; the numbers denote the position of the methyl and ethyl branches.

Table 2
Overall hyper-Wiener index *OWW*, and its *e*th-order components, calculated for the 44 alkanes.

No.	Molecule[a]	^1OWW	^2OWW	^3OWW	^4OWW	^5OWW	^6OWW	^7OWW	*WW*	*OWW*
1	C3	2	5						5	7
2	nC4	3	10	15					15	28
3	2MC3	3	15	12					12	30
4	nC5	4	15	30	35				35	84
5	2MC4	4	20	42	28				28	94
6	22MMC3	4	30	48	22				22	104
7	nC6	5	20	45	70	70			70	210
8	2MC5	5	25	57	98	58			58	243
9	3MC5	5	25	72	91	54			54	247
10	23MMC4	5	30	84	112	47			47	278
11	22MMC4	5	35	93	106	44			44	283
12	nC7	6	25	60	105	140	126		126	462
13	2MC6	6	30	72	133	198	108		108	547
14	3MC6	6	30	87	161	182	99		99	565
15	3EC5	6	30	102	189	162	90		90	579
16	24MMC5	6	35	84	196	232	91		91	644
17	23MMC5	6	35	114	210	213	83		83	661
18	22MMC5	6	40	108	211	218	84		84	667
19	33MMC5	6	40	138	225	196	76		76	681
20	223MMMC4	6	45	150	274	229	69		69	773
21	nC8	7	30	75	140	210	252	210	210	924
22	2MC7	7	35	87	168	268	360	185	185	1110
23	3MC7	7	35	102	196	322	333	170	170	1165
24	4MC7	7	35	102	231	310	324	165	165	1174
25	3EC6	7	35	117	259	360	288	150	150	1216
26	25MMC6	7	40	99	196	396	432	161	161	1331
27	24MMC6	7	40	114	259	426	397	147	147	1390
28	22MMC6	7	45	123	246	428	408	149	149	1406
29	23MMC6	7	40	129	280	411	389	143	143	1399
30	34MMC6	7	40	144	308	449	364	134	134	1446
31	2M3EC5	7	40	144	343	433	346	129	129	1442
32	33MMC6	7	45	153	330	440	358	131	131	1464
33	3M3EC5	7	45	183	379	456	318	118	118	1506
34	224MMMC5	7	50	135	344	566	441	127	127	1670
35	234MMMC5	7	45	156	364	542	423	122	122	1659
36	223MMMC5	7	50	180	407	565	402	115	115	1726
37	233MMMC5	7	50	195	428	547	387	111	111	1725
38	2233MMMMC4	7	60	231	548	687	414	97	97	2044
39	Cyclopropane	3	15	3					3	21
40	Cyclobutane	4	20	60	10				10	94
41	Cyclopentane	5	25	75	175	25			25	305
42	Cyclohexane	6	30	90	210	420	51		51	807
43	Cycloheptane	7	35	105	245	490	882	94.5	94.5	1858.5
44	Cyclooctane	8	40	120	280	560	1008	1680	172	3868

[a] M and E stand for methyl and ethyl, respectively; the numbers denote the position of the methyl and ethyl branches.

Table 3
Path (p), cluster (c), and pathcluster (pc) terms of overall hyper-Wiener index *OWW* of the 44 alkanes.

No.	${}^3OWW_p$	${}^3OWW_c$	${}^4OWW_p$	${}^4OWW_c$	${}^4OWW_{pc}$	${}^5OWW_p$	${}^5OWW_c$	${}^5OWW_{pc}$	${}^6OWW_p$	${}^6OWW_c$	${}^6OWW_{pc}$	${}^7OWW_p$	${}^7OWW_c$	${}^7OWW_{pc}$
1														
2	15	0												
3	0	12												
4	30	0	35	0	0									
5	30	12	0	0	28									
6	0	48	0	22	0									
7	45	0	70	0	0	70	0	0						
8	45	12	70	0	28	0	0	58						
9	60	12	35	0	56	0	0	54						
10	60	24	0	0	112	0	47	0						
11	45	48	0	22	84	0	0	44						
12	60	0	105	0	0	140	0	0	126	0	0			
13	60	12	105	0	28	140	0	58	0	0	108			
14	75	12	105	0	56	70	0	112	0	0	99			
15	90	12	105	0	84	0	0	162	0	0	90			
16	60	24	140	0	56	0	0	232	0	0	91			
17	90	24	70	0	140	0	47	166	0	0	83			
18	60	48	105	22	84	0	0	218	0	0	84			
19	90	48	35	22	168	0	0	196	0	0	76			
20	90	60	0	22	252	0	141	88	0	69	0			
21	75	0	140	0	0	210	0	0	252	0	0	210	0	0
22	75	12	140	0	28	210	0	58	252	0	108	0	0	185
23	90	12	140	0	56	210	0	112	126	0	207	0	0	170
24	90	12	175	0	56	140	0	170	126	0	198	0	0	165
25	105	12	175	0	84	140	0	220	0	0	288	0	0	150
26	75	24	140	0	56	280	0	116	0	0	432	0	0	161
27	90	24	175	0	84	140	0	286	0	0	397	0	0	147
28	75	48	140	22	84	210	0	218	0	0	408	0	0	149
29	105	24	140	0	140	140	47	224	0	0	389	0	0	143

Table 3
(Continued)

No.	$^{3}OWW_{p}$	$^{3}OWW_{c}$	$^{4}OWW_{p}$	$^{4}OWW_{c}$	$^{4}OWW_{pc}$	$^{5}OWW_{p}$	$^{5}OWW_{c}$	$^{5}OWW_{pc}$	$^{6}OWW_{p}$	$^{6}OWW_{c}$	$^{6}OWW_{pc}$	$^{7}OWW_{p}$	$^{7}OWW_{c}$	$^{7}OWW_{pc}$
30	120	24	140	0	168	70	47	332	0	0	364	0	0	134
31	120	24	175	0	168	0	47	386	0	0	346	0	0	129
32	105	48	140	22	168	70	0	370	0	0	358	0	0	131
33	135	48	105	22	252	0	0	456	0	0	318	0	0	118
34	75	60	210	22	112	0	0	566	0	0	441	0	0	127
35	120	36	140	0	224	0	94	448	0	0	423	0	0	122
36	120	60	105	22	280	0	141	424	0	69	333	0	0	115
37	135	60	70	22	336	0	141	406	0	69	318	0	0	111
38	135	96	0	44	504	0	423	264	0	414	0	0	97	0
39	15													
40	60	0	35											
41	75	0	175	0	0	70								
42	90	0	210	0	0	420	0	0	126					
43	105	0	245	0	0	490	0	0	882	0	0	210	0	0
44	120	0	280	0	0	560	0	0	1008	0	0	1680	0	0

The following correlations have been obtained:

$$\begin{aligned}
\mathrm{BP} &= -135.3764 + 19.140 \ln {}^{1}OWW + 37.8555\,{}^{1}OWW - 4.2774\,{}^{2}OWW \\
&\quad + 0.6996\,{}^{3}OWW_{\mathrm{c}} - 0.0364\,{}^{4}OWW_{\mathrm{p}} + 0.0465\,{}^{4}OWW_{\mathrm{pc}}, \qquad (10)\\
&\quad n = 44, \quad r = 0.9984, \quad s = 2.74, \quad F = 1944,\\
\mathrm{MV} &= 49.1343 + 16.3498\,{}^{1}OWW + 0.3156\,{}^{2}OWW \\
&\quad - 0.1359\,{}^{3}OWW_{\mathrm{p}} - 0.0604\,{}^{3}OWW_{\mathrm{c}}, \qquad (11)\\
&\quad n = 34, \quad r = 0.9997, \quad s = 0.39, \quad F = 11762,\\
\mathrm{MR} &= 6.4476 + 4.7567\,{}^{1}OWW + 0.0053\,{}^{2}OWW - 0.0094\,{}^{3}OWW_{\mathrm{p}}, \qquad (12)\\
&\quad n = 34, \quad r = 0.9999, \quad s = 0.024, \quad F = 371152,\\
\mathrm{HV} &= 3.9965 + 6.9941\,{}^{1}OWW - 0.3551\,{}^{2}OWW + 0.0273\,{}^{4}OWW \\
&\quad - 0.0340\,{}^{4}OWW_{\mathrm{p}} - 0.0213\,{}^{4}OWW_{\mathrm{pc}}, \qquad (13)\\
&\quad n = 34, \quad r = 0.9991, \quad s = 0.22, \quad F = 3096,\\
\mathrm{PC} &= 49.8616 - 4.4920\,{}^{1}OWW + 0.0055\,{}^{5}OWW \\
&\quad + 0.0584\,{}^{3}OWW_{\mathrm{p}} - 0.0124\,{}^{6}OWW_{\mathrm{c}}, \qquad (14)\\
&\quad n = 38, \quad r = 0.9904, \quad s = 0.54, \quad F = 422,\\
\mathrm{TC} &= 30.9442 + 35.4980\,{}^{1}OWW - 0.1802\,{}^{4}OWW \\
&\quad + 0.6848\,{}^{3}OWW_{\mathrm{p}} - 0.0516\,{}^{5}OWW_{\mathrm{p}}, \qquad (15)\\
&\quad n = 38, \quad r = 0.9960, \quad s = 4.80, \quad F = 1018,\\
\mathrm{ST} &= 8.3673 + 2.1675\,{}^{1}OWW - 0.14745\,{}^{2}OWW - 0.0026\,{}^{6}OWW \\
&\quad + 0.0414\,{}^{3}OWW_{\mathrm{p}} + 0.0228\,{}^{4}OWW_{\mathrm{p}}, \qquad (16)\\
&\quad n = 33, \quad r = 0.9930, \quad s = 0.23, \quad F = 379.
\end{aligned}$$

4. Results and discussion

All the correlations are better than those that only using the hyper-Wiener index. The boiling points of cyclohexane and cycloheptane predicted by this paper are 83.9°C and 115.1°C, which are closer to the experimental values 80.7° and 118.4°C, because the overall hyper-Wiener index accounts for more factors than the hyper-Wiener index. For example, although the hyper-Wiener indices of cycloalkanes are lower than their n-alkane counterparts, but as the cycloalkanes have more edges than their n-alkane counterparts, the *e*th-order components of the overall hyper-Wiener indices are higher than their n-alkane counterparts. So the properties predicted by the overall hyper-Wiener indices may be higher or lower than their n-alkane counterparts.

As the hyper-Wiener indices are included in the overall hyper-Wiener indices, the basic features of hyper-Wiener indices are preserved by the overall hyper-Wiener indices. So the overall hyper-Wiener indices preserve the features of the hyper-Wiener indices and eliminate their pitfalls, it may be useful for QSAR and QSPR studies.

References

[1] M. Randic, X. Gou, T. Oxley and H. Krishnapriyan, Wiener matrix: Source of novel matrix invariants, J. Chem. Inf. Comput. Sci. 33 (1993) 709.
[2] H. Wiener, Correlation of heats of isomerization and differences in heats of vaporization of isomers among the paraffin hydrocarbons, J. Am. Chem. Soc. 69 (1947) 2636.
[3] H. Wiener, Structural determination of paraffin boiling points, J. Am. Chem. Soc. 69 (1947) 17.
[4] S. Nikolic and Z. Trinajstic, The Wiener index: Developments and applications, Croat. Chim. Acta 68 (1995) 105.
[5] M. Randic, Novel molecular descriptor for structure–property studies, Chem. Phys. Lett. 211 (1993) 478.
[6] M. Randic, X. Gou, T. Oxley, H. Krishnapriyan and L. Naylor, Wiener matrix invariants, J. Chem. Inf. Comput. Sci. 34 (1994) 261.
[7] D.J. Klein, I. Lukovits and I. Gutman, On the definition of the hyper-Wiener index for cycle-containing structures, J. Chem. Inf. Comput. Sci. 35 (1995) 50.
[8] A. Roberto and H. Pierre, A linear algorithm for the hyper-Wiener index of chemical trees, J. Chem. Inf. Comput. Sci. 41 (2001) 958.
[9] I. Lukovits and W. Linert, A novel definition of the hyper-Wiener index for cycles, J. Chem. Inf. Comput. Sci. 34 (1994) 899.
[10] D. Bonchev, Kolmogorov's information, Shannon's entropy, and topological complexity of molecules, Bulg. Chem. Commun. 28 (1995) 567.
[11] D. Bonchev, Novel indices for the topological complexity of molecules, ASR/QSAR Envir. Res. 7 (1997) 23.
[12] D. Bonchev, Overall connectivity/topological complexities: A new tool for QSAR/QSPR, J. Chem. Inf. Comput. Sci. 40 (2000) 934.
[13] D. Bonchev, The overall Wiener index – A new tool for characterization of molecular topology, J. Chem. Inf. Comput. Sci. 41 (2001) 582.
[14] L.B. Kier and L.H. Hall, *Molecular Connectivity in Chemistry and Drug Research* (Academic Press, New York, 1976).
[15] D.E. Needham, I.C. Wei and P.G. Seybold, Molecular modeling of the physical properties of the alkanes, J. Am. Chem. Soc. 110 (1998) 4186.
[16] C.Z. Cao and H. Yuan, Topological indices based on vertex, distance, and rings: On the boiling points of paraffins and cycloalkanes, J. Chem. Inf. Comput. Sci. 41 (2001) 867.