A New Hyper-Wiener Index*

Ivan Gutman

Faculty of Science, University of Kragujevac, 34000 Kragujevac, Serbia and Montenegro
(E-mail: gutman@knez.uis.kg.ac.yu)

RECEIVED DECEMBER 19, 2002; REVISED MARCH 1, 2003; ACCEPTED MARCH 18, 2003

If two edges $e$ and $f$ are deleted from a tree $T$, then it decomposes into three components, possessing, $n_0(e, f)$, $n_1(e, f)$, and $n_2(e, f)$ vertices. Let $n_0(e, f)$ count the vertices lying between the edges $e$ and $f$. It is shown that the Wiener index $W$ of the tree $T$ is equal to the sum over all edges $e$ of the products $n_1(e, e) / n_2(e, e)$, and that the hyper-Wiener index $WW$ of $T$ is the sum over all pairs of edges $e, f$ of the products $n_1(e, f) / n_2(e, f)$. We now consider another structure-descriptor, denoted by $WWW$, equal to the sum over all pairs of edges of the products $n_0(e, f) / n_1(e, f) / n_2(e, f)$. We establish some basic properties of $WWW$ and show how it is related to $W$.

Key words
- Wiener index, $W$
- hyper-Wiener index, $WW$
- $WWW$ index

INTRODUCTION

In his seminal paper1 Wiener not only introduced the topological index $W$ that nowadays carries his name,2,3 but also gave an efficient method for its calculation, Eq. (1). Recall that the Wiener index $W$ is, by definition, equal to the sum of distances between all pairs of vertices of the respective (molecular) graph. Hence, for a graph on $n$ vertices, for the direct calculation of $W$ a total of $n(n–1)/2$ distances needs to be determined.4,5 In the case of trees, Wiener showed1 that his index can be computed by means of the formula

$$W(T) = \sum_{e \in E(T)} n_1(e, e) \cdot n_2(e, e)$$

where $T$ is a tree, $E(T)$ is its edge set and the summation goes over all edges of $T$. The number of vertices of $T$ lying on the two sides of the edge $e$ are denoted by $n_1(e, e)$ and $n_2(e, e)$. Thus the right hand side of formula (1) consists of only $n–1$ summands, each of which is easily evaluated.

Motivated by formula (1) Randić5 introduced the hyper-Wiener index $WW$. In the case of trees this structure descriptor is defined as follows. Let $u$ and $v$ be two vertices of a tree $T$ and let $\pi_{uv}$ be the unique path connecting $u$ and $v$. Let $n_1(\pi_{uv})$ and $n_2(\pi_{uv})$ be the number of vertices of $T$ lying on the two sides of the path $\pi_{uv}$. Then

$$WW(T) = \sum_{u,v} n_1(\pi_{uv}) \cdot n_2(\pi_{uv})$$

with summation going over all pairs of vertices of $T$. Note that if the summation in Eq. (2) would be restricted to pairs of adjacent vertices, then Eq. (2) would reduce to Eq. (1).

Formula (2) is not applicable to graphs that contain cycles. For such graphs the problem of defining $WW$ was solved by Klein et al.6

Formula (2) can be put in a somewhat more convenient form. To do this we need some preparations.

Let $T$ be a tree and $u$ and $v$ two of its vertices, see Figure 1(a). There is a unique path $\pi_{uv}$ connecting these

* Dedicated to Professor Nenad Trinajstić, my former teacher, from whom I learned much.
two vertices, see Figure 1(b). This path uniquely determines two edges e and f; we denote by e the edge of T, incident to vertex u and belonging to \( \pi_{uv} \), analogously, f is the edge incident to vertex v and belonging to \( \pi_{uv} \), see Figure 1(c). By deleting the edges e and f from T it decomposes into three components, \( T_0, T_1, \) and \( T_2 \), see Figure 1(d). These possess \( n_0(e, f), n_1(e, f), \) and \( n_2(e, f) \) vertices, respectively. Note that if u and v are adjacent vertices, then the edges e and f coincide, \( e = f \). Then \( T_0 \) does not exist, i.e., \( n_0(e, e) = 0 \).

The general structure of a tree T and the labeling of its components is shown in Figure 2.

If the tree T has \( n \) vertices, then evidently
\[
 n_0(e, f) + n_1(e, f) + n_2(e, f) = n
\]
holds for any choice of the edges e and f, i.e., for any two vertices u and v of T.

Bearing the above in mind, we see that the quantities \( n_1(\pi_{uv}) \) and \( n_2(\pi_{uv}) \), occurring in Eq. (2), coincide with \( n_1(e, f) \) and \( n_2(e, f) \), and that Eq. (2) may be rewritten as
\[
 \text{WW}(T) = \sum_{e,f} n_1(e, f) \cdot n_2(e, f) \tag{3}
\]
where the summation goes over all pairs of edges, including those in which \( e = f \).

In view of Eq. (1), formula (3) is transformed into
\[
 \text{WW}(T) = \text{W}(T) + \sum_{e,f} n_1(e, f) \cdot n_2(e, f) \tag{4}
\]
with \( \sum_{e,f} \) indicating summation only over pairs of different edges.

**A MODIFICATION OF THE HYPER-WIENER INDEX**

Already Eq. (4) implies that there must be some relation between the Wiener and hyper-Wiener indices. However, only a short time ago\(^7\) it was discovered that \( \text{WW} \) is bounded from both below and from above by linear functions of \( \text{W} \):
\[
 \frac{3n}{4} \frac{W(T)}{W(T)} - \frac{1}{12} n(n-1)^2(n+1) \leq \frac{W(T)}{W(T)} \leq \frac{n+1}{2} \frac{W(T)}{W(T)} - \frac{1}{2} (n^2 - 2n + 2) (n-1).
\]

Quite recently these estimates were improved as\(^8\)
\[
 \frac{3}{2} \frac{W(T)}{W(T)} - \frac{1}{2} (n-1) \leq \text{WW}(T) \leq \frac{n+2}{4} \frac{W(T)}{W(T)} - \frac{1}{4} n(n-1)(n+1).
\]

The above relations suggest that \( \text{W} \) and \( \text{WW} \) must be highly correlated quantities, what eventually was confirmed by extensive numerical testing.\(^8\)

In order to eliminate the effect on the hyper-Wiener index, Eq. (3), of terms pertaining to \( e = f \), we arrived at a new structure-descriptor \( \text{WWW} \), defined as follows.

Let \( T \) be a tree whose structure is depicted in Figure 2. Then the new hyper-Wiener index is equal to:
\[
 \text{WWW}(T) = \sum_{e,f} n_0(e, f) \cdot n_1(e, f) \cdot n_2(e, f) \tag{5}
\]
Because of \( n_0(e, e) = 0 \), formula (5) is tantamount to
\[
 \text{WWW}(T) = \sum_{e,f} n_0(e, f) \cdot n_1(e, f) \cdot n_2(e, f) \tag{6}
\]

Consequently, the terms that contribute to the Wiener index, Eq. (1), have no effect on WWW. As a result, the correlation between WWW and W is significantly weaker than the correlation between WW and W. An illustrative example is given in Table I and Figure 3.

The consequence of the good correlation between WW and W is that the structure-property and structure-activity relations, based on the ordinary hyper-Wiener index, are of exactly same quality as the corresponding relations based on the Wiener index. Because of the much weaker correlation between WWW and W (and also between WWW and WW), the efficiency of the new hyper-Wiener index in QSPR and QSAR studies is different, sometimes better, than the efficiency of W and WW. Details on this matter will be reported elsewhere.9

**A RELATION BETWEEN THE NEW HYPER-WIENER INDEX AND THE WIENER INDEX**

In this section we report a relation between the WWW-index of a tree T and the Wiener indices of some fragments of T. In order to formulate and prove this relation we shall use the following notation (illustrated in Figure 4).

Let T be a tree and e its edge. The subtrees obtained by deleting e from T are denoted by A and B. Clearly, A and B depend on the choice of the edge e.

The edge sets of T, A, and B are denoted by \(E(T)\), \(E(A)\) and \(E(B)\), respectively. The trees A and B have \(n_A\) and \(n_B\) vertices. Evidently, \(n_A + n_B = \) number of vertices of T and \(E(T) = E(A) \cup E(B) \cup \{e\}\).

With this notation, Wiener’s formula (1) reads

\[
W(T) = \sum_{e \in E(T)} n_A \cdot n_B .
\]

**TABLE I.** Wiener indices (W), hyper-Wiener indices (WW) and the new hyper-Wiener indices (WWW) for the isomeric octanes, cf. Figure 3

<table>
<thead>
<tr>
<th>Compound</th>
<th>W</th>
<th>WW</th>
<th>WWW</th>
</tr>
</thead>
<tbody>
<tr>
<td>octane</td>
<td>84</td>
<td>210</td>
<td>252</td>
</tr>
<tr>
<td>2-methylheptane</td>
<td>79</td>
<td>185</td>
<td>232</td>
</tr>
<tr>
<td>3-methylheptane</td>
<td>76</td>
<td>170</td>
<td>228</td>
</tr>
<tr>
<td>4-methylheptane</td>
<td>75</td>
<td>165</td>
<td>228</td>
</tr>
<tr>
<td>3-ethylhexane</td>
<td>72</td>
<td>150</td>
<td>224</td>
</tr>
<tr>
<td>2,2-dimethylhexane</td>
<td>71</td>
<td>149</td>
<td>204</td>
</tr>
<tr>
<td>2,3-dimethylhexane</td>
<td>70</td>
<td>143</td>
<td>208</td>
</tr>
<tr>
<td>2,4-dimethylhexane</td>
<td>71</td>
<td>147</td>
<td>210</td>
</tr>
<tr>
<td>2,5-dimethylhexane</td>
<td>74</td>
<td>161</td>
<td>214</td>
</tr>
<tr>
<td>3,3-dimethylhexane</td>
<td>67</td>
<td>131</td>
<td>198</td>
</tr>
<tr>
<td>3,4-dimethylhexane</td>
<td>68</td>
<td>134</td>
<td>204</td>
</tr>
<tr>
<td>3-ethyl-2-methylpentane</td>
<td>67</td>
<td>129</td>
<td>204</td>
</tr>
<tr>
<td>3-ethyl-3-methylpentane</td>
<td>64</td>
<td>118</td>
<td>192</td>
</tr>
<tr>
<td>2,2,3-trimethylpentane</td>
<td>63</td>
<td>115</td>
<td>182</td>
</tr>
<tr>
<td>2,2,4-trimethylpentane</td>
<td>66</td>
<td>127</td>
<td>188</td>
</tr>
<tr>
<td>2,3,3-trimethylpentane</td>
<td>62</td>
<td>111</td>
<td>180</td>
</tr>
<tr>
<td>2,3,4-trimethylpentane</td>
<td>65</td>
<td>122</td>
<td>190</td>
</tr>
<tr>
<td>2,2,3,3-tetramethylbutane</td>
<td>58</td>
<td>97</td>
<td>162</td>
</tr>
</tbody>
</table>

**Figure 3.** Correlation between the Wiener index (W), the hyper-Wiener index (WW) and the new hyper-Wiener index (WWW) for isomeric octanes, cf. Table I. The respective correlation coefficients are 0.997 (for WW vs. W) and 0.976 (for WWW vs. W).

**Figure 4.** The notation used in the proof of Eq. (8). It is in agreement with, and refers to, the notation specified in Figure 2.
From the definition of the WWW-index, Eq. (5), and the fact that \( n_0(e, e) = 0 \) we have

\[
\text{WWW}(T) = \sum_{e,f} n_0(e,f) \cdot n_1(e,f) \cdot n_2(e,f)
\]

\[
= \frac{1}{2} \sum_{e \in E(T)} \sum_{f \in E(T)} n_0(e,f) \cdot n_1(e,f) \cdot n_2(e,f)
\]

\[
= \frac{1}{2} \sum_{e \in E(T)} \left( \sum_{f \in E(A)} n_0(e,f) \cdot n_1(e,f) \cdot n_2(e,f) + \sum_{f \in E(B)} n_0(e,f) \cdot n_1(e,f) \cdot n_2(e,f) \right)
\]

because the edge \( f \) lies either in A or in B, see Figure 4.

Now, using the notation explained in Figure 4, if \( f \in E(A) \), then \( n_1(e,f) = n_B \), whereas if \( f \in E(B) \), then \( n_1(e,f) = n_A \). Therefore,

\[
\text{WWW}(T) = \frac{1}{2} \sum_{e \in E(T)} \left( n_B \sum_{f \in E(A)} n_0(e,f) \cdot n_1(e,f) \cdot n_2(e,f) + n_B \sum_{f \in E(B)} n_0(e,f) \cdot n_1(e,f) \cdot n_2(e,f) \right)
\]

In accordance with Eq. (1), for any fixed edge \( e \),

\[
\sum_{f \in E(A)} n_0(e,f) \cdot n_2(e,f) = W(A)
\]

\[
\sum_{f \in E(B)} n_0(e,f) \cdot n_2(e,f) = W(B)
\]

which results in the following remarkable identity:

\[
\text{WWW}(T) = \frac{1}{2} \sum_{e \in E(T)} \left( n_B W(A) + n_B W(B) \right) . \tag{8}
\]

It is worth comparing relation (8) with Wiener's formula (7) as well as with an analogous, recently obtained, expression for the ordinary hyper-Wiener index:

\[
\text{WW}(T) = \frac{n}{2} W(T) - \frac{1}{2} \sum_{e \in E(T)} \left( W(A) + W(B) \right) .
\]

REFERENCES


SAŽETAK

Novi hiper-Wiener indeks

Ivan Gutman

Ako se iz stabla \( T \) izbace dva bridja \( e \) i \( f \), onda se ono razlaže na tri komponente, koje posjeduju \( n_0(e,f) \), \( n_1(e,f) \), i \( n_2(e,f) \) vrhova. Neka \( n_0(e,f) \) broji one vrhove koji leže između \( e \) i \( f \). Pokazuje se da je Wienerov indeks \( W \) stabla \( T \) jednak zbroju preko svih bridova \( e \) produkta \( n_1(e,e) \cdot n_2(e,e) \), a da je hiper-Wienerov indeks \( WW \) istog stabla jednak zbroju preko svih parova bridova \( e,f \) produkta \( n_1(e,f) \cdot n_2(e,f) \). U radu se razmatra i jedan novi strukturni deskriptor, označen s \( WWW \), jednak zbroju preko svih parova bridova produkta \( n_0(e,f) \cdot n_1(e,f) \) \( n_2(e,f) \). Utvrđena su neka temeljna svojstva indeksa \( WWW \) i pokazano je kako je on povezan s \( W \).