Distance–Related Molecular Descriptors

Bono Lučić,¹ Sonja Nikolić,¹ and Nenad Trinajstić¹

¹ The Rugjer Bošković Institute, P. O. Box 180, HR–10002 Zagreb, Croatia

Received: September 4, 2008; Accepted: September 12, 2008; Published: September 30, 2008

Citation of the article:
Internet Electronic Journal of Molecular Design

Distance–Related Molecular Descriptors

Bono Lučić,1 Sonja Nikolić,1 and Nenad Trinajstić1, *

1 The Rugjer Bošković Institute, P. O. Box 180, HR–10002 Zagreb, Croatia

Received: September 4, 2008; Accepted: September 12, 2008; Published: September 30, 2008

Abstract

A brief review of various molecular descriptors based on graph–theoretical distances is presented with a special emphasis on those distance–descriptors in whose development István Lukovits has been involved.

Keywords. Distance descriptors; Wiener index and its variants; Harary index and its variants; detour index and its variants; Vérhalom index; Pasaréti index; resistance–distance indices.

1 INTRODUCTION

István Lukovits was much interested in the distance–related molecular descriptors [e.g., 1–8]. He was especially interested in the mathematical and computational properties of the Wiener index and hyper–Wiener index and their applications in the structure–property–activity modeling [e.g., 1–4, 9–13]. We also collaborated with Lukovits in several project related to molecular descriptors based on graph–theoretical distances [e.g., 14–19]. Here we will review, in remembering Lukovits and his work in developing the theory of distance–related molecular descriptors, a number of distance descriptors and in particular those in whose derivation he was actively participating. Let us also mention that Lukovits has been attracted to chemical graph theory after lectures on graphs in chemistry, delivered by Živković and Trinajstić at the conference held in spring 1974 in Mátrafüred. These two authors also prepared a joint paper based on their lectures that has been published in Hungarian – the translation from English was expertly done by Professor Gábor Náray–Szabó [20].

But, before proceeding, let us first review the basic graph–theoretical concepts that we will use in the article. In presenting the needed (chemical) graph–theoretical concepts, we consulted the textbooks by Harary [21] and Wilson [22] and monographs by Trinajstić [23] and Gutman and

# Dedicated to the memory of Dr István Lukovits (1945–2007), dear friend, colleague and true gentleman, who helped us so much in our research. Collaboration with him was a real enjoyment. We deeply regret his passing so early in his life.

* Correspondence author; E–mail: trina@irb.hr.

BioChem Press

http://www.biochempress.com
Polansky [24]. A graph $G$ is formally defined to be a pair $[V(G), E(G)]$, where $V(G)$ is a non-empty finite set of elements called vertices and $E(G)$ is a finite set of unordered pairs of (not necessarily distinct) elements of $V(G)$ called edges. Molecular graphs represent the constitution of molecules. They are generated using the following rule: vertices stand for atoms and edges for bonds. Molecular graphs we consider are connected graphs. A graph is connected if every pair of vertices is joined by a path. The graph-theoretical distance is the length of the shortest path in terms of the minimum number of edges between any two vertices in $G$. The detour distance is the length of the longest path in terms of the largest number of edges between any two vertices in $G$. Obviously the graph-theoretical distance and detour distance are identical in acyclic structures. An invariant of a graph $G$ is a number associated with $G$ that has the same value for any graph isomorphic to $G$. If $G$ is a molecular graph, than the corresponding invariants are called molecular descriptors or topological indices by practitioners of chemical graph theory.

2 SELECTED DISTANCE DESCRIPTORS

2.1 Wiener index

The oldest graph-theoretical descriptor is the Wiener index, usually denoted by $W$, introduced in 1947 [25]. It is so named after his inventor Harry Wiener (1924–1998). His life and his achievements first in chemistry and later in medicine have been well-presented by Rouvray [26].

The Wiener index can be calculated in a number of ways [2,27]. The original method of Wiener was valid only for acyclic structures (trees) and may be summarized as “Multiply the number of vertices on each side of an edge and add all such contributions”. This statement may be formalized as follows. Let $T$ be a (molecular) tree with $N$ vertices and $e$ one of its edges. Let $N_1(e)$ and $N_2(e) = N - N_1(e)$ be the numbers of vertices of the two parts of $T - e$. Then the edge-decomposition formula is given by:

$$ W = \sum_{e} N_1(e)N_2(e) $$

(1)

The Wiener edge-decomposition formula which is valid only for acyclic graphs has been generalized by Lukovits and Gutman in 1994 to be applicable to cycle-containing systems [28]. The general edge-composition formula is given by:

$$ E = \sum_{e} \sum_{i\neq j} [p_{ij}(e)/p_{ij}] $$

(2)

where $p_{ij}$ is the total number of paths between vertices $i$ and $j$ that are of length $l(i,j)$. The number of such paths that contain the edge $e$ is denoted by $p_{ij}(e)$. Lukovits also devised an efficient algorithm for computing the edge-contributions of the Wiener index [4].
However, the graph–theoretical definition of the Wiener index that has been given by Hosoya in 1971 [29] is computationally much simpler. The Wiener index $W$ of a molecular graph $G$ is equal to the half–sum of the off–diagonal elements of the distance matrix $D$:

$$ W = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} [D]_{ij} $$

where $[D]_{ij}$ represents the length of a shortest path between vertices $i$ and $j$ in $G$.

The distance matrix $D$ of $G$ is a real symmetric $V \times V$ matrix whose elements are defined as [30]:

$$ [D]_{ij} = \left\{ \begin{array}{ll}
 l(i,j) & \text{if } i \neq j \\
 0 & \text{if } i = j 
\end{array} \right. $$

where $l(i,j)$ is the minimum number of edges between vertices $i$ and $j$ in $G$. The setting up the distance matrix for large systems is rather involved. However, there is also available a very efficient computer program for constructing the distance matrix and computing the Wiener index [31].

The Wiener index found considerable application in QSPR and QSAR [e.g., 9,10,32–34].

### 2.2 Partial Wiener indices

Lukovits [10] studied contributions of different types of bonds on the Wiener index using the following expression:

$$ W = W_S + W_D + W_T + W_A $$

where $W_S$ is obtained by adding the contributions of all single bonds, $W_D$ by adding the contributions of all double bonds, $W_T$ by adding the contributions of all triple bond and $W_A$ by adding the contributions of all aromatic bonds in the considered molecule.

### 2.3 Modified Wiener Index

The modified Wiener index, denoted $^mW$, can be calculated by modifying the original Wiener’s method: “Multiply the reciprocal number of vertices on each side of an edge and add all such contributions”. Therefore, the modification of original Wiener edge–decomposition formula, introduced by Nikolić et al. [35] in 2001, is given by:

$$ ^mW = \sum_{e} \frac{1}{N_1(e)} \frac{1}{N_2(e)} $$

where $^mW$ denotes the modified Wiener index. The above formula is applicable only to acyclic graphs. The modified Wiener number found so far rather limited application.
2.4 Hyper–Wiener Index

The hyper–Wiener index, denoted by $WW$, was introduced by Randić [36] in 1993. However, his algorithm for computing $WW$ could be applied only to acyclic structures. Lukovits et al. [37,38] showed that $WW$ can be computed for all structures (cyclic and acyclic) with the following formula:

$$WW = \frac{1}{4} \sum_{i=1}^{N} \sum_{j=1}^{N} \left( [D]_{ij} + [D]^{2}_{ij} \right)$$

This equation can be rewritten as:

$$WW = \frac{W}{2} + \frac{1}{4} \sum_{i=1}^{N} \sum_{j=1}^{N} [D]^{2}_{ij}$$

Lukovits was also involved in the study of trees with extremal hyper–Wiener index [39] and in designing an algorithm for the computation of $WW$ index and its variants for the characterization and discrimination of branched acyclic and cyclic molecules [3,12,40,41].

The hyper–Wiener index has been applied with some success to QSPR modeling [e.g., 16,32].

2.5 The Multiplicative Version of the Wiener Index

The multiplicative version of the Wiener index $\pi(W)$ of a molecular graph $G$ is equal to the product of shortest distances between all pairs of vertices in $G$ [42,43]:

$$\pi(W) = \prod_{i<j} [D]_{ij}$$

Since this index is, even for small molecular graphs, rather large number, e.g. 288 and 34560 for pentane and hexane graphs, the authors suggested $\ln \pi(W)$ to be used instead of $\pi(W)$ in the QSPR modeling.

2.6 Pasaréti Index

The Pasaréti index $P$ is an all–path version of the Wiener index [6,8,17]. It is defined as:

$$P = \sum_{i<j} \sum_{p_{ij}} | p_{ij} |$$

where $p_{ij}$ stands for a path between vertices $i$ and $j$ and $| p_{ij} |$ denotes the length of this path. The summation is between all pairs of vertices $i$ and $j$ and for all paths between vertices $i$ and $j$. The Pasaréti index was successfully tested in the structure–boiling point modeling of acyclic and cyclic hydrocarbons [17]. It is worth mentioning on this sad occasion the origin of the term Pasaréti index. The name Pasaréti index is given for this molecular descriptor because it was devised in the home of late István Lukovits that was located in a lovely part of Budapest called Pasaréti. A few years later he moved to another part of Budapest where he ended his days in this world.
2.7 Vérhalom Index

The Pasaréti index is an exponential function of graph size in terms of the number of vertices \( N \). This feature of the Pasaréti index might hinder its use in the structure–property–activity modeling because it becomes too large a number in the homologous series of molecules. Therefore, the Pasaréti index was transformed into a variant that is numerically much more user–friendly. It is called the Vérhalom index and denoted by \( V \). This descriptor is defined rather simply:

\[
V = \frac{P}{n}
\]  

(11)

where \( n \) is the total number of paths in a molecular graph divided by \( N(N–1)/2 \). The Vérhalom index was also successfully tested on the same set of molecules for the same property as the Pasaréti index [17].

What about the term Vérhalom index? It is again connected with the visit of one of us (NT) to Budapest. The Vérhalom index was devised in the office of late István Lukovits in the Chemical Research Center of the Hungarian Academy of Sciences located in the district Vérhalom in Budapest, hence the term.

2.8 Harary Index

The Harary index, denoted by \( H \), has independently been defined by Plavšić et al. [44] and Ivanciuc et al. [45] in 1993. It is defined as the half–sum of the off–diagonal elements of the reciprocal distance matrix \( D' \):

\[
H = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} [D']_{ij}
\]  

(12)

The reciprocal distance matrix \( D' \) can simply be obtained by replacing all elements in the distance matrix \( [D]_{ij} \) by their reciprocals [30]:

\[
[D']_{ij} = \frac{1}{[D]_{ij}}, \quad i \neq j
\]  

(13)

whereas diagonal elements \( [D']_{ii} \) are zero by definition. Harary index also found application in QSPR modeling [e.g., 17,33,44,45].

2.9 Hyper–Harary Index

The hyper–Harary, denoted by \( HH \), has been introduced and applied by Diudea [46] in 1997. This index is defined in a similar way as the hyper–Wiener index, that is, as the quarter–sum of the sum–matrix made up from the elements of the reciprocal distance matrix \( [D']_{ij} \) and their squares \( [D'_{ij}]^2 \):
Distance–Related Molecular Descriptors


\[ HH = \frac{1}{4} \sum_{i=1}^{N} \sum_{j=1}^{N} \left( D'_{ij} \right)_y + \left( D'^2 \right)_y \] (14)

This equation can also be rewritten as:

\[ HH = \frac{H}{2} + \frac{1}{4} \sum_{i=1}^{N} \sum_{j=1}^{N} \left[ D'_{ij} \right]^{2}_y \] (15)

The hyper–Harary found limited use in the structure–property modeling [e.g., 18,33].

2.10 Detour Index

The detour index has been named so by Lukovits [5] in 1996. However, Ivanciuc and Balaban defined this index [47] in 1994 as the maximum path sum index MPS, whereas Amić and Trinajstić introduced the same index [48] in 1995 as the Wiener–like index. The detour index, denoted by \( \omega \), is defined in the same way as the Wiener index \( W \), that is, the detour index is equal to the half–sum of the elements in the detour matrix \( \Delta \):

\[ \omega = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} [\Delta]_{ij} \] (16)

where \( [\Delta]_{ij} \) represents the length of a longest path between vertices \( i \) and \( j \) in \( G \).

The detour matrix \( \Delta \) of \( G \) is a real symmetric \( V \times V \) matrix whose elements are defined as [30]:

\[ [\Delta]_{ij} = \begin{cases} L(i,j) & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \] (17)

where \( L(i,j) \) is the maximum number of edges between vertices \( i \) and \( j \) in \( G \). The setting up the detour matrix for large systems is not easy. There are several methods available for setting up this matrix [e.g., 48,49]. Lukovits was also involved in computing the detour matrix. He and Razinger devised a computer program [50] for the paper–and–pencil method of path tracing proposed by Amić and Trinajstić [48]. The detour index found considerable application in QSPR [e.g., 5,14,18,33,48,51,52].

2.11 Hyper–Detour Index

The hyper–detour index, denoted by \( \omega_\omega \), has been proposed by Lukovits [5,53]. It is defined similarly as the hyper–Wiener index \( WW \) and hyper–Harary index \( HHI \), that is, as the quarter–sum of the sum–matrix made up from the off–diagonal elements of the detour matrix \( [\Delta]_{ij} \) and their squares \( [\Delta]_{ij}^2 \):

\[ \omega_\omega = \frac{1}{4} \sum_{i=1}^{N} \sum_{j=1}^{N} \left( [\Delta]_{ij} + [\Delta]_{ij}^2 \right) \] (18)
This equation can also be rewritten as:

\[ \omega = \frac{\omega}{2} + \frac{1}{4} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{i,j} [\Lambda]_{ij}^2 \]  \hspace{1cm} (19)

Since the Wiener index \( W \) and the detour index \( \omega \) are identical for acyclic structures, the same is also true for their hyper–counterparts.

The hyper–detour index found modest application in the structure–property modeling \([e.g., 17,18,33]\).

### 2.12 The Multiplicative Version of the Detour Index

In analogy with the multiplicative version of the Wiener index, Lukovits and Trinajstić [54] proposed the multiplicative version of the detour index \( \pi(\omega) \). The index \( \pi(\omega) \) of a molecular graph \( G \) is equal to the product of longest distances between all pairs of vertices in \( G \):

\[ \pi(\omega) = \prod_{i<j} [\Lambda]_{ij} \]  \hspace{1cm} (20)

In this case, the authors also suggested \( \ln \pi(\omega) \) to be used instead of \( \pi(\omega) \) in the QSPR modeling. Descriptors \( \ln \pi(W) \) and \( \ln \pi(\omega) \) are, of course, identical for acyclic structures. However, the authors never completed work on this descriptor due to Lukovits’ premature death.

### 2.13 Kirchhoff Index

The Kirchhoff index \( Kf \) is defined by means of the resistance–distance matrix [55]:

\[ Kf = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} [\Omega]_{ij} \]  \hspace{1cm} (21)

where \([\Omega]_{ij}\) represents the resistance–distance between vertices \( i \) and \( j \) of \( G \). With an aim to examine other possible metrics in molecular graph, Klein and Randić [56] introduced in 1993 the concept of the resistance distance. The roots of this concept may be found in an early report by Foster [57] as recently pointed out by Palacios [58].

The resistance–distance matrix \( \Omega \) is defined as [30]:

\[ [\Omega]_{ij} = \begin{cases} r(i,j) & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \]  \hspace{1cm} (22)

where \( r(i,j) \) is the resistance–distance between vertices \( i \) and \( j \). An algorithm in whose development has been involved Lukovits, based on the Laplacian matrix \([30,59,60]\), was proposed in 2002 for the efficient computing of the resistance–distance matrix [19]. The Laplacian matrix \( L \) is defined as the following difference matrix [30]:

\[ L = V - A \]  \hspace{1cm} (23)
where \( V \) is the degree matrix and \( A \) is the adjacency matrix of a molecular graph \( G \). The degree matrix \( V \) is a diagonal matrix with entries:

\[
[V]_i = d(i)
\]  

(24)

where \( d(i) \) is the degree of the vertex \( i \) in \( G \). The adjacency matrix \( A \) has entries either 1 when the corresponding two vertices are adjacent or 0 if these vertices are not adjacent.

### 2.14 Quotient Matrices and Wiener–Sum Indices

Quotient matrices have been introduced by Randić [61]. Several quotient matrices and the related Wiener–sum indices have been proposed and applied [e.g., 17, 61–64]. Here we will review only four Wiener–sum indices \( WS \) because Lukovits was involved in their development and applications to cyclicity of polycyclic graphs [19] and structure–boiling point modeling of acyclic and cyclic hydrocarbons [17]. However, the potential of \( WS \) indices in the structure–property modeling has not so far been sufficiently explored.

#### 2.14.1 Wiener–sum index of the \( D/\Delta \) matrix

The Wiener–sum index, denoted by \( WS(D/\Delta) \), is defined as the half–sum of the diagonal elements of the quotient matrix \( D/\Delta \) of a molecular graph \( G \):

\[
WS(D/\Delta) = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{[D]_{ij}}{[\Delta]_{ij}}
\]

(25)

The \( D/\Delta \) matrix is defined as [30]:

\[
[D/\Delta]_{ij} = \begin{cases} 
[D]_{ij}/[\Delta]_{ij} & \text{if } i \neq j \\
0 & \text{if } i = j
\end{cases}
\]

(26)

Randić [61] introduced the average row–sum index, denoted by \( RS \), as a measure of cyclicity. The \( RS \) index is trivially related to the \( WS \) index as \( RS = 2WS/N \).

Therefore, the \( RS \) index introduces only a modicum of information beyond the \( WS \) index. The \( WS(D/\Delta) \) has also been used in the structure–boiling point modeling of selected alkanes and cycloalkanes [17].

#### 2.14.2 Wiener–sum index of the \( \Delta/D \) matrix

The Wiener–sum index, denoted by \( WS(\Delta/D) \), is defined as the half–sum of the diagonal elements of the quotient matrix \( \Delta/D \) of a molecular graph \( G \):

\[
WS(\Delta/D) = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{[\Delta]_{ij}}{[D]_{ij}}
\]

(27)

The \( \Delta/D \) matrix can be simply be obtained from the \( D/\Delta \) matrix by inverting its off–diagonal
elements [30]:

\[
[D/D]_{ij} = \begin{cases} 
[A]_{ij}/[D]_{ij} & \text{if } i \neq j \\
0 & \text{if } i = j 
\end{cases}
\]  

(28)

The \(WS(\Delta/D)\) index found practical applications in structure–boiling point modeling for acyclic and cyclic hydrocarbons [16] and for benzenoids [62].

### 2.14.3 Wiener–sum index of the \(D/\Omega\) matrix

Another Wiener–sum index, denoted by \(WS(D/\Omega)\), is defined as the half–sum of the diagonal elements of the quotient matrix \(D/\Omega\) of a molecular graph \(G\) [19]:

\[
WS(D/\Omega) = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{[D]_{ij}}{[\Omega]_{ij}}
\]  

(29)

The \(D/\Omega\) matrix is defined [30]:

\[
[D/\Omega]_{ij} = \begin{cases} 
[D]_{ij}/[\Omega]_{ij} & \text{if } i \neq j \\
0 & \text{if } i = j 
\end{cases}
\]  

(30)

This index has so far been used only for studying cyclicity of five–vertex graphs, Platonic solids and \(C_{60}\) and \(C_{70}\) fullerenes [19].

### 2.14.4 Wiener–sum index of the \(\Omega/D\) matrix

The Wiener–sum index of the \(\Omega/D\) matrix, denoted by \(WS(\Omega/D)\), was also given the name Kirchhoff–sum index and was denoted by \(KfS\) [19]. It is defined as follows:

\[
KfS = WS(\Omega/D) = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{[\Omega]_{ij}}{[D]_{ij}}
\]  

(31)

The quotient matrix \(\Omega/D\) is simply obtainable from the \(D/\Omega\) matrix by inverting its non–diagonal elements [30]:

\[
[\Omega/D]_{ij} = \begin{cases} 
[\Omega]_{ij}/[D]_{ij} & \text{if } i \neq j \\
0 & \text{if } i = j 
\end{cases}
\]  

(32)

The \(WS(\Omega/D)\) index or the Kirchhoff–sum index \(KfS\) was used to discuss the cyclicity in four classes of polycyclic graphs: all five–vertex graphs, Schlegel graphs representing Platonic solids, \(C_{60}\) and \(C_{70}\) fullerenes.

### 3 CONCLUDING REMARKS

István Lukovits published at least 106 papers – 102 papers are listed under his name in the *Web of Science* (as seen on February 20, 2008). In this source four of his papers are not listed.
The first one it is here listed as ref. [1] and contains Lukovits’ report on the generalized Wiener index for molecules containing double bonds and its application to modeling the partition coefficients of selected alkanes, cycloalkanes, alkenes and cycloalkenes. It is possible that this paper is not listed because the journal *Reports in Molecular Theory* (Editors: Gábor Náray–Szabó and Harel Weinstein, published by CRC Press) has been very short–lived, just two issues in 1990. This paper is however discussed by Todeschini and Consonni in their *Handbook* [33].

The second paper missing is listed in this report as ref. [8]. In this paper, Lukovits reports on the Wiener–type graph invariants. This paper is not listed in the *Web of Science* because this source does not contain books.

The third paper missing is listed in this report as ref. [28]. In this paper, Lukovits and Gutman reported the edge–decomposition of the Wiener number. This paper is not listed in the *Web of Science* presumably because MATCH was not yet recorded in this source. This paper is however mentioned by Todeschini and Consonni in their *Handbook* [32].

The fourth paper missing is listed here as ref. [65]. In this paper, Lukovits discusses the generation formulae for isomers and it is based on his lecture at the International Conference on the Role of Topology in Chemistry (Athens, GA, March 20–24, 2001). Probably the same reasoning why this paper is not listed in the *Web of Science* applies here as for the paper above that also appeared in the book.

His last paper, published post–morten, is listed here as ref. [66]. This paper is listed in his 102 papers that the *Web of Science* contains. In this paper, the authors discussed aromaticity of carbon nanotubes and it was based on the invited lecture delivered by István Lukovits at the 21st Dubrovnik International Course and Conference MATH/CHEM/COMP 2006 (Dubrovnik, June 19–24, 2006).

His 106 papers may roughly be classified as follows: 26 have been concerned with development and computation of molecular descriptors, 23 with chemical graph–theoretical problems, 32 with uses of various descriptors in the structure–property–activity modeling, 15 with quantum–chemical studies, 4 with nanotubes, 4 with the corrosion inhibitors, one with the philosophy of science and one with chemical information.

5 REFERENCES


F. Harary, Graph Theory, Addison–Wesley, Reading, MA, 1971, 2nd printing.  
R. J. Wilson, Introduction to Graph Theory, Oliver & Boyd, Edinburgh, 1974, 2nd printing.  


