The Wiener Index and the Szeged Index of Benzenoid Systems in Linear Time

Victor Chepoi*

Laboratoire de Biomathématiques, Faculté de Medecine, Université d'Aix Marseille II, 27 Bd Jean Moulin, F-13385 Marseille Cedex 5, France

Sandi Klavžar

Department of Mathematics, PEF, University of Maribor, Koroška cesta 160, 2000 Maribor, Slovenia

Received February 21, 1997[⊗]

A linear time algorithm is presented which, for a given benzenoid system G, computes the Wiener index of G. The algorithm is based on an isometric embedding of G into the Cartesian product of three trees, combined with the notion of the Wiener index of vertex-weighted graphs. An analogous approach yields also a linear algorithm for computing the Szeged index of benzenoid systems.

1. INTRODUCTION

Distance properties of molecular graphs form an important topic in chemical graph theory.[1] To justify this statement just recall the famous Wiener index which is also known as the Wiener number. This index is the first[2] but also one of the most important topological indices of chemical graphs. Its research is still very active; see recent reviews[3,4] and several new results in a volume[5] dedicated to the 50th anniversary of Wiener's paper.[2]

Many methods and algorithms for computing the Wiener index of a graph were proposed in the chemical literature; papers in refs 4, 6−10 present just a sample of these studies. The fastest general algorithm[7] for computing the Wiener index of a graph is of complexity $O(mn)$, where n and m denote the number of vertices and the number of edges, respectively, of a considered graph. Linear algorithms were previously proposed only for trees[7] and for fasciagraphs and rotagraphs.[9] We wish to add, however, that in the latter case the linearity is a more theoretical than practical term. This is due to the fact that the size of the problem is supposed to be the number of copies of a given graph G from which a fasciagraph (rotagraph) is build up and not the size of G itself.

The Szeged index (Sz) of a (molecular) graph is a recently proposed[11,12] topological index which has already received considerable attention.[13−20] Among others, a method for its computation was given,[14] and an $O(mn)$ algorithm for its computation has been proposed.[17]

A new approach to the study of distance properties of molecular graphs was proposed by Klavžar, Gutman, and Mohar.[21] It was shown that benzenoid systems provide so-called isometric embeddings into hypercubes, and based on this fact a simple formula for the Wiener index of these graphs has been obtained. The approach was further developed in the subsequent papers.[10,14] In particular, combinatorial expressions for the Wiener index of compact pericondensed benzenoid hydrocarbons were computed.[10] Along these lines it was observed[22] that benzenoid systems can also be isometrically embedded into the Cartesian product of three trees. As an application it was demonstrated how to compute the diameter of a benzenoid system in linear time.

Here we continue to develop this line of research. Our main result asserts that the Wiener index of a benzenoid system with n vertices can be computed in $O(n)$ time. In the next section we present the algorithm and prove its correctness. We also give an example from which it can be seen that this linear algorithm for the Wiener index not only is fast on computers but also is conceptually simple and easy to follow on the paper. In the last section we demonstrate that our approach can also be used to compute the Szeged index of a benzenoid system on n vertices in $O(n)$ time.

The *Wiener index* or *Wiener number* of a (molecular) graph $G = (V,E)$ is defined as follows:

$$W(G) = \frac{1}{2}\sum_{u\in V}\sum_{v\in V} d_G(u,v)$$

Here $d_G(u,v)$ stands for the usual shortest path distance of the graph G, i.e., the length of a shortest path connecting u with v in G.

Given two connected graphs G and H, we say that G admits an *isometric embedding* (alias distance-preserving embedding) into H if there exists a mapping

$$\beta : V(G) \rightarrow V(H)$$

such that

$$d_H(\beta(u),\beta(v)) = d_G(u,v)$$

for all vertices $u,v \in V(G)$.

The *Cartesian product* $H = H_1 \times H_2 \times \ldots \times H_m$ of connected graphs $H_1, H_2, \ldots, H_m$ is the graph on the vertex set

$$V(H) = \{u = (u_1,u_2,\ldots,u_m) : u_i \in V(H_i),\ i = 1,\ldots,m\}$$

Two vertices $u = (u_1,u_2,\ldots,u_m)$ and $v = (v_1,v_2,\ldots,v_m)$ of H are adjacent if and only if the vectors u and v coincide in all but one position i, where u_i and v_i are adjacent in H_i. The most important distance property for the Cartesian product is the following. If $x = (x_1,\ldots,x_m)$ and $y = (y_1,\ldots,y_m)$ are vertices of H then we have

$$d_H(x,y) = \sum_{i=1}^{m} d_{H_i}(x_i,y_i)$$

[⊗] Abstract published in *Advance ACS Abstracts,* June 1, 1997.

S0095-2338(97)00007-3 CCC: $14.00

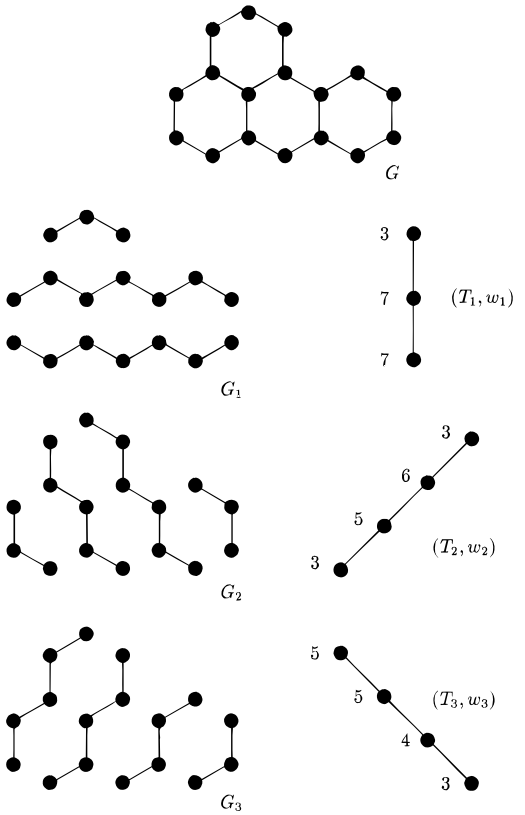

Figure 1. Steps in computing the Wiener index of the benzenoid system *G*.

Benzenoid systems (or *benzenoid graphs*) are graphs constructed in the following manner.[23] Let H be the hexagonal (graphite) lattice and let *Z* be a cricuit on it. Then a benzenoid system is formed by the vertices and edges of H, lying on *Z* and in the interior of *Z*. Note that the edge set of a benzenoid system can be partitioned into three classes of parallel edges.

2. WIENER INDEX OF BENZENOID SYSTEMS IN LINEAR TIME

In order to obtain a linear algorithm for computing the Wiener index of benzenoid systems we need to recall the embedding of a benzenoid system into the Cartesian product of three trees.[22]

Let G be a benzenoid system and let E_1, E_2, and E_3 denote the edges of G of a given direction. For $i = 1, 2, 3$ let G_i be the graph which is obtained from G be deleting all the edges of E_i. Note that the connected components of the graph G_i are paths. Define a graph T_i whose vertices are the connected components of G_i and two such components P' and P'' are adjacent in T_i if and only if there are vertices $u \in P'$ and $v \in P''$ which are endvertices of an edge from E_i. (See Figure 1 as well as figures from ref 22 for examples.) Then every T_i is a tree and moreover, the mapping

$$\alpha = V(G) \rightarrow V(T_1 \times T_2 \times T_3)$$

defined as $\alpha(v) = (v_1, v_2, v_3)$, is an isometric embedding. Here v_i is the vertex of T_i corresponding to the connected component of G_i containing v. We call the mapping α the *canonical embedding* of the benzenoid system G. In fact, the embedding α can be extended in a certain way such that it is an isometric embedding of an (arbitrary) graph into the Cartesian product of graphs. This general result is due to Graham and Winkler.[24] We can now state the following result.[22]

Theorem 1. *The canonical embedding* $\alpha : V(G) \rightarrow V(T_1 \times T_2 \times T_3)$ *of a benzenoid system G with n vertices is an isometric embedding. Moreover, the trees* T_1, T_2, and T_3 *and the corresponding labels of the vertices of G can be computed in O(n) time.*

We next invoke a concept of the Wiener index on weighted graphs.[25] A (*vertex*)-*weighted graph* (G,w) is a graph G together with a function w: $V(G) \rightarrow |\mathbf{N}^+$. The Wiener number $W(G,w)$ of a weighted graph (G,w) is defined as

$$W(G,w) = \frac{1}{2} \sum_{u,v \in V(G)} w(u)w(v)d_G(u,v)$$

Let G be a benzenoid system a let T_1, T_2, and T_3 be the trees from the canonical embedding. For $i = 1, 2, 3$ we introduce weighted trees (T_i,w_i) as follows: for $u \in T_i$ let $w_i(u)$ be the number of vertices x of G such that the ith component of $\alpha(x)$ is u. In other words, $w_i(u)$ is just the number of vertices in the connected component of G_i which corresponds to the vertex u.

Proposition 2. *Let G be a benzenoid system and let* α *be the canonical embedding of G into* $T_1 \times T_2 \times T_3$. *Then*

$$W(G) = W(T_1,w_1) + W(T_2,w_2) + w(T_3,w_3)$$

Proof. Let $V = V(G)$ and let $H = T_1 \times T_2 \times T_3$. For $u \in V$ let $\alpha(u) = (u_1,u_2,u_3)$. Then we have

$$\begin{aligned}
W(G) &= \frac{1}{2}\sum_{u\in V}\sum_{v\in V} d_G(u,v) \\
&= \frac{1}{2}\sum_{u\in V}\sum_{v\in V} d_H(\alpha(u),\alpha(v)) \\
&= \frac{1}{2}\sum_{u\in V}\sum_{v\in V}\sum_{i=1}^{3} d_{T_i}(u_i,v_i) \\
&= \sum_{i=1}^{3}\left(\frac{1}{2}\sum_{u\in V}\sum_{v\in V} d_{T_i}(u_i,v_i)\right) \\
&= \sum_{i=1}^{3}\left(\frac{1}{2}\sum_{u_i\in T_i}\sum_{v_i\in T_i} w_i(u)w_i(v)d_{T_i}(u_i,v_i)\right) \\
&= \sum_{i=1}^{3} W(T_i,w_i)
\end{aligned}$$

By Proposition 2, a linear algorithm for computing $W(G)$ will be provided by a linear algorithm for computing the Wiener index of a weighted tree. For this sake we state the following.

Lemma 3. *Let (T,w) be a weighted tree. For an edge e of T, let* T_1 *and* T_2 *be the connected components of T\e and for* $i = 1, 2$ *set*

$$n_i(e) = \sum_{u\in T_i} w(u)$$

Then we have

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

Lemma 3 was previously[25] proved in a more general setting and represents a generalization of the classical Wiener's result[2] which asserts the above formula in the case when $w(u) = 1$ for any vertex u of G. Moreover, this result was used by Mohar and Pisanski[7] to obtain a simple linear algorithm for computing the Wiener index of a tree. By a simple modification of their algorithm we can thus also compute the Wiener index of a weighted tree in linear time. (This is, in fact, implicit in the paper.[7]) Briefly, this algorithm can be described as follows. We order the vertices of a given weighted tree (T,w) so that the next vertex v is a leaf in the subtree induced by the vertices with a larger index. Let u be the neighbor of v in this subtree. Then add the factor $w(v)(n - w(v))$ to the current sum and update $w(u)$ by letting $w(u) := w(u) + w(v)$. Note finally that using Theorem 1 it is easy to obtain the weighted trees T_i in linear time. Combining these observations with Proposition 2 we have arrived to the main result of this paper:

Theorem 4. *The Wiener index of a benzenoid system on n vertices can be computed in O(n) time.*

We illustrate the above procedure with the following example. It should be clear from it that the method proposed not only is simple and fast for a computer but also efficient for a "hand" manipulation. The steps of the example are illustrated in Figure 1.

Let G be a benzenoid system from Figure 1. Then the graphs G_1, G_2, and G_3 are obtained from G be deleting all edges of a given direction, respectively. (Observe that the connected components of the graphs G_i are paths.) Then each component of G_i is a vertex of the factor graph T_i, for $i = 1, 2, 3$. The weight of a given vertex in the weighted tree (T_i,w_i) is simply the number of vertices in the corresponding component of G_i.

Then by Lemma 3 we have

$$W(T_1,w_1) = 3\cdot(7+7) + (3+7)\cdot 7 = 112$$

$$W(T_2,w_2) = 3\cdot(5+6+3) + (3+5)\cdot(6+3) + (3+5+6)\cdot 3 = 156$$

$$W(T_3,w_3) = 5\cdot(5+4+3) + (5+5)\cdot(4+3) + (5+5+4)\cdot 3 = 172$$

and from here using Proposition 2 we conclude

$$W(G) = 112 + 156 + 172 = 440$$

3. SZEGED INDEX OF BENZENOID SYSTEMS IN LINEAR TIME

The Szeged index is defined in the following way. Let e be an edge of the molecular graph G, connecting the vertices u and v. Let $n_1(e|G)$ be the number of vertices x of the graph G, having the property $d_G(x,u) < d_G(x,v)$ and let $n_2(e|G)$ be the number of vertices x on the graph G with $d_G(x,u) > d_G(x,v)$. The Szeged index of the graph G is

$$Sz(G) = \sum_{e \in E(G)} n_1(e|G) n_2(e|G)$$

Let G be a benzenoid system. Recall that a straight line segment c with end points p and q is called a *cut segment* if c is orthogonal to one of the three edge directions, each p and q is the center of an edge, and the graph obtained from G by removing all edges intersected by c has exactly two connected components. The *cut* is the set of edges intersected by a cut segment.

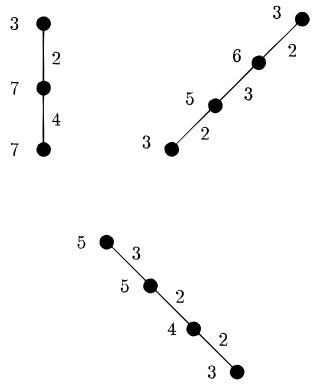

Figure 2. Edge labelings in the process of computing the Szeged index of the benzenoid system G from Figure 1.

Proposition 5. *Let G be a benzenoid system which is canonically embedded into $T_1 \times T_2 \times T_3$. For $i = 1, 2$, and 3 let (T_i,w_i) be weighted trees as in Proposition 2 and let n_i be as in Lemma 3. For an edge e of T_i let f(e) be the number of edges in the cut of G separating the path corresponding to its end-vertices. Then we have*

$$Sz(G) = \sum_{i=1}^{3} \sum_{e \in T_i} f(e) n_1(e) n_2(e)$$

Proof. We have

$$Sz(G) = \sum_{e \in E(G)} n_1(e|G) n_2(e|G) = \sum_{C} f(e) n_1(e|G) n_2(e|G) = \sum_{i=1}^{3} \sum_{e \in T_i} f(e) n_1(e) n_2(e)$$

The first equality is the definition of the Szeged index. In the second equality we sum over all cuts of G, i.e., in every cut C we select an arbitrary representative of C. That the equality indeed holds follows from the fact[21] that if two edges e and e' belong to the same cut C, then $n_1(e|G) = n_1(e'|G)$. Finally, for the third line of the above computations observe that every cut corresponds to a unique vertex in one of the trees T_i and that $n_1(e)$ sums up to $n_1(e|G)$.

From Proposition 5 and the results of section 2 we immediately obtain theorem 6.

Theorem 6. *The Szeged index of a benzenoid system on n vertices can be computed in O(n) time.*

Consider, for instance, the benzenoid from Figure 1. Then as for the Wiener index we construct the weighted trees (T_i,w_i). In addition, for all edges e of these trees we compute $f(e)$, i.e., the number of edges in the cut corresponding to e, see Figure 2.

Using Proposition 5 we have

$$Sz(G) = 2\cdot3\cdot(7+7) + 4\cdot(3+7)\cdot7 + 2\cdot3\cdot(5+6+3) + 3\cdot(3+5)\cdot(6+3) + 2\cdot(3+5+6)\cdot3 + 3\cdot5\cdot(5+4+3) + 2\cdot(5+5)\cdot(4+3) + 2\cdot(5+5+4)\cdot3$$
$$= 364 + 384 + 404 = 1152$$

ACKNOWLEDGMENT

The work of one author (S.K.) was supported in part by the Ministry of Science and Technology of Slovenia under the grant J1-7036.

REFERENCES AND NOTES

(1) Trinajstić, N. *Chemical Graph Theory*; CRC Press: Boca Raton, FL, 1993.
(2) Wiener, H. Structural determination of paraffin boiling points, *J. Am. Chem. Soc.* **1947**, *69*, 17−20.
(3) Gutman, I.; Yeh, Y. N.; Lee, S. L.; Luo, Y. L. Some recent results in the theory of the Wiener number. *Indian J. Chem.* **1993**, *32A*, 651−661.
(4) Nikolić, S.; Trinajstić, N.; Mihalić, Z. The Wiener index: Development and applications. *Croat. Chem. Acta* **1995**, *68*, 105−129.
(5) Fifty Years of the Wiener Index; Gutman, I., Klavžar, S., Mohar, B., Eds.; MATCH **1997**, *35*, in press.
(6) Canfield, E. R.; Robinson, W. R.; Rouvray, D. H. Determination of the Wiener molecular branching index for the molecular tree. *J. Comput. Chem.* **1985**, *6*, 598.
(7) Mohar, B.; Pisanski, T. How to compute the Wiener index of a graph. *J. Math. Chem.* **1988**, *2*, 267−277.
(8) Gutman, I. Calculating the Wiener number: The Doyle-Graver method. *J. Serb. Chem. Soc.* **1993**, *58*, 745−750.
(9) Juvan, M.; Mohar, B.; Graovac, A.; Klavžar, S.; Žerovnik, J. Fast computation of the Wiener index of fasciagraphs and rotagraphs. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 834−840.
(10) Gutman, I.; Klavžar, S. A Method for calculating Wiener numbers of benzenoid hydrocarbons. *Models in Chemistry (Acta Chim. Hung.)* **1996**, *133*, 389−399.
(11) Gutman, I. A formula for the Wiener number of trees and its extension to graphs containing cycles. *Graph Theory Notes New York* **1994**, *27*, 9−15.
(12) Khadikar, P. V.; Deshpande, N. V.; Kale, P. P.; Dobrynin, A.; Gutman, I. The Szeged index and an analogy with the Wiener index. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 547−550.
(13) Dobrynin, A.; Gutman, I. On a graph invariant related to the sum of all distances in a graph. *Publ. Inst. Math. (Beograd)* **1994**, *56(70)*, 18−22.
(14) Gutman, I.; Klavžar, S. An algorithm for the calculation of the Szeged index of benzenoid hydrocarbons. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 1011−1014.
(15) Dobrynin, A. A.; Gutman, I.; Dömötör, G. A Wiener-type graph invariant for some bipartite graphs. *Appl. Math. Lett.* **1995**, *8(5)*, 57−62.
(16) Klavžar, S.; Rajapakse, A.; Gutman, I. The Szeged and the Wiener index of graphs. *Appl. Math. Lett.* **1996**, *9(5)*, 45−49.
(17) Žerovnik, J. Computing the Szeged index. *Croat. Chem. Acta* **1996**, *69*, 837−843.
(18) Gutman, I.; Popović, L.; Khadikar, P. V.; Karmarkar, S.; Joshi, S.; Mandloi, M. Relations between Wiener and Szeged indices of monocyclic molecules. *Comm. Math. Comp. Chem. (MATCH)* **1997**, *35*, in press.
(19) Dobrynin, A. A.; Gutman, I. Szeged index of some polycyclic bipartite graphs with circuits of different size. *Comm. Math. Comp. Chem. (MATCH)* **1997**, *35*, in press.
(20) Dobrynin, A. A. The Szeged index for complements of hexagonal chains. *Comm. Math. Comp. Chem. (MATCH)* **1997**, *35*, in press.
(21) Klavžar, S.; Gutman, I.; Mohar, B. Labeling of benzenoid systems which reflects the vertex-distance relations. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 590−593.
(22) Chepoi, V. On distances in benzenoid systems. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1169−1172.
(23) Gutman, I.; Cyvin, S. J. *Introduction to the Theory of Benzenoid Hydrocarbons*; Springer-Verlag: Berlin, 1989.
(24) Graham, R. L.; Winkler, P. M. On isometric embeddings of graphs. *Trans. Amer. Math. Soc.* **1985**, *288*, 527−536.
(25) Klavžar, S.; Gutman, I. Wiener number of vertex-weighted graphs and a chemical application. *Discrete Appl. Math.* Accepted for publication.

CI9700079