Character of Graphs with Extremal Balaban Index[*]

Hawei Dong [a,b]   **and**   **Xiaofeng Guo** [a]

[a] _School of Mathematical Sciences, Xiamen University,_
Xiamen, Fujian 361005, China

[b] _Department of Mathematics, Minjiang University,_
Fuzhou, Fujian 350108, China

e-mail: **donghawei@126.com** , **xfguo@xmu.edu.cn**

(Received April 9, 2009)

Abstract

The Balaban index (also called J index) of a connected graph G is defined as $J = J(G) = \frac{|E(G)|}{\mu+1} \sum_{uv \in E(G)} \frac{1}{\sqrt{\sigma_G(u)\sigma_G(v)}}$, where $\sigma_G(u) = \sum_{w \in V(G)} d_G(u, w)$ and μ is the cyclomatic number. Balaban index has been used in various QSAR and QSPR studies. In this paper, we characterize the graphs with both some parameters (such as the number of vertices, connectivity, diameter) and extreme Balaban indices.

INTRODUCTION

Molecular topology can be expressed numerically in term of molecular descriptors, and among these descriptors topological indices (TIs) occupy a special place because they are more complex than counts of atoms, groups or bonds, but less complicated than quantum-chemical parameters. Consequently, they can be computed in a very short time from various types of input data on atom connectivities, and be used for quantitative structure-property relationship (QSPR) and quantitative structure activity relationship (QSAR) [1] .

The oldest TI, the Wiener index introduced by Wiener [2] in 1947, is defined as the sum of distances between vertices in a graph G:

$$W(G) = \sum_{\{u,v\} \subseteq V(G)} d_G(u, v),$$

where $V(G)$ is the vertex set of G and $d_G(u, v)$ is the distance between vertices u and v of G.

The molecular connectivity index χ was defined by Randić [3] in terms of vertex degree:

$$\chi = \chi(G) = \sum_{uv \in E(G)} \frac{1}{\sqrt{d_G(u)d_G(v)}},$$

where $d_G(u)$ and $d_G(v)$ are the degree of vertices u and v in G, respectively.

Balaban index was proposed by Balaban [4, 5] which also is called J index. It was defined analogously as Randić index but vertex degrees were replaced by distance sums, and each distance sum multiplies the cyclomatic number plus one and divides the number of edges :

$$J = J(G) = \frac{|E(G)|}{\mu + 1} \sum_{uv \in E(G)} \frac{1}{\sqrt{\sigma_G(u)\sigma_G(v)}}, \tag{1}$$

where $\sigma_G(u) = \sum_{w \in V(G)} d_G(u, w)$ and $\mu = |E(G)| - |V(G)| + 1$ is the cyclomatic number. $\sigma_G(u)$ (resp. $\sigma_G(v)$) is also called the distance of vertex u (v) and is equal to the row sum of distance matrix of G corresponding to u (v).

There are other indices constructed in similar manner, such as JJ index. Instead of using the row sums of the adjacency matrix or distance matrix, it uses the row sums of the Wiener matrix, which has a good discrimination [6].

In [7], Balaban et al compared the ordering of constitutional isomers of alkanes with 6 through 9 carbon atoms. It was shown that the ordering induced by Balaban index parallels the ordering induced by Wiener index, but reduces the degeneracy of the latter index and provides a much higher discriminating ability. Therefore Balaban index also is called 'sharpened Wiener index' sometimes.

The behavior of Balaban index mimics the behavior of the melting temperatures and glass transition temperatures for linear macromolecules, which possess an asymptotic limit for these physical properties. The asymptotic value of Balaban index for an infinite path is the number $\pi = 3.14159$ in [8] and the asymptotic properties for Fibonacci trees are analyzed in [9]. An approximate hyperbolic dependence is presented in [10] between the asymptotic value of Balaban index and normalized Wiener index of infinite polymers.

Moreover, Balaban index was used subsequently in various QSAR and QSPR studies [11]–[42].

In [38], QSAR study on benzensulphonamide carbonic anhydrase inhibitors has been made by using Balaban index which was found better than the widely used Wiener index and Randić connectivity index.

Balaban index has some applications in biology. In [39], Grassy et al applied Balaban index to the optimization of a biologically active lead compound using 13 molecular descriptors. Bermudez et al [40] showed that Balaban index is useful in establishing relationships among RNAs. And Shu et al [41] indicated that the topological indices can characterize the details of RNA structures and may have a protential role in identifying and classifying ncRNAs.

For chemical applications, it may be interesting to identify the graph with the maximum and minimum TIs in given class of graphs, especially trees. Numerous related results on graphs with extreme Tls, such as Wiener index [43]–[47] and Randić index [48], have been gotten and some tight lower and upper bounds of Balaban indices of some graphs are also given [49].

In this paper, we study extreme Balaban indices of graphs with various paraments, and prove that among all the graphs with n vertices, path P_n has the minimum Balaban index, complete graph K_n has the maximum Balaban index, and star S_n is the tree with the maximum Balaban index. In addition, we characterize the k-connected (k-edge-connected) graphs with n vertices and the maximum Balaban indices, and prove that cycle C_n is the only 2-connected graph with n vertices and the minimum Balaban index. Finally, we characterize the graph (resp. tree) with n vertices, diameter d, and the maximum Balaban index.

PRELIMINARIES

In this article we consider only finite, undirected graphs without loops or multiple edges. Throughout the paper, if S is a set, $|S|$ will denote its cardinality. For convenience, we denote the weight of an edge uv in a graph G about Balaban index as: $J_G(uv) = \frac{1}{\sqrt{\sigma_G(u)\sigma_G(v)}}$, and denote the weight of a subgraph G' of G about Balaban index as $J_G(G') = \sum\limits_{uv \in E(G')} \frac{1}{\sqrt{\sigma_G(u)\sigma_G(v)}}$. Thus the Balaban index of G can be expressed as: $J(G) = \frac{|E(G)|}{\mu+1} \sum\limits_{e \in E(G)} J_G(e) = \frac{|E(G)|}{\mu+1} \sum\limits_{i=1}^{p} J_G(G_i)$, where G_i, $i = 1, 2, \cdots, p$, are edge disjoint subgraphs of G and $\cup_{i=1}^{p} E(G_i) = E(G)$.

Now we give some useful inequalities.

Lemma 1 *If a, a', b, b', w, x, y, $z \in R^+$ such that $\frac{b}{x} \geq \frac{a}{w}$, $\frac{b'}{y} \geq \frac{a'}{z}$, $w \geq x$ and $z \geq y$, then*

$$\frac{1}{\sqrt{(w+a)(z+a')}} + \frac{1}{\sqrt{xy}} \geq \frac{1}{\sqrt{wz}} + \frac{1}{\sqrt{(x+b)(y+b')}}. \text{ And the equation holds if and only if } b = a, \ b' = a',$$

$w = x$ and $z = y$.

Proof. Let $\frac{w}{x} = c_1$ and $\frac{z}{y} = c_2$, then $c_1 \geq 1$ and $c_2 \geq 1$.

$$\frac{1}{\sqrt{xy}} - \frac{1}{\sqrt{(x+b)(y+b')}} \geq \frac{1}{\sqrt{c_1 c_2}} \left(\frac{1}{\sqrt{xy}} - \frac{1}{\sqrt{(x+b)(y+b')}} \right)$$

that is,

$$\frac{1}{\sqrt{xy}} + \frac{1}{\sqrt{c_1 c_2}} \frac{1}{\sqrt{(x+b)(y+b')}} \geq \frac{1}{\sqrt{c_1 c_2}} \frac{1}{\sqrt{xy}} + \frac{1}{\sqrt{(x+b)(y+b')}}.$$

Hence

$$\frac{1}{\sqrt{xy}} + \frac{1}{\sqrt{(w+a)(z+a)}} \geq \frac{1}{\sqrt{xy}} + \frac{1}{\sqrt{(c_1 x + c_1 b)(c_2 y + c_2 b')}}$$

$$\geq \frac{1}{\sqrt{c_1 c_2}} \frac{1}{\sqrt{xy}} + \frac{1}{\sqrt{(x+b)(y+b')}} = \frac{1}{\sqrt{wz}} + \frac{1}{\sqrt{(x+b)(y+b')}}.$$

∎

The proof of two lemmas in the following are easy and are omitted.

Lemma 2 *If $x, y, a \in R^+$ such that $x \geq y + a$, then $\frac{1}{\sqrt{(x-a)(y+a)}} \leq \frac{1}{\sqrt{xy}}$. And the equation holds if and only if $x = y + a$.*

Lemma 3 *Let G be a graph. Then $J(G - e) < J(G)$ for any $e \in E(G)$.*

In [50], a maximal subtree containing a vertex v of a tree T as an end vertex is called a branch of T at v, the weight of a branch B of T, denoted by $bw(B)$, is the number of edges in B. The centroid of a tree T, denoted by $C(T)$, is the set of vertices v of T for which the maximum branch weight at v is minimized.

Theorem 4 *[50] If $C = C(T)$ be the centroid of a tree T of order n, then one of the following holds:*

(1) $C := \{c\}$ and $bw(c) \leq (n-1)/2$;

(2) $C := \{c_1, c_2\}$ and $bw(c_1) = bw(c_2) = n/2$.

In both cases, if $v \in V(T) \setminus C$ then $bw(v) \geq n/2$.

Theorem 5 [51] *If $p_0 p_1 \cdots p_k$ is a path in a tree T, $p_0 \in C(T)$ and $p_1 \notin C(T)$, then $\sigma_T(p_0) < \sigma_T(p_1) < \cdots < \sigma_T(p_k)$.*

Lemma 6 *Let $P = p_1 \cdots p_l$ be a path in a graph G_0, and let G_j $(1 \le j \le l)$ be the component of $G_0 - E(P)$ containing p_j. Let H_1, H_2, H_1', H_2' be trees such that $|V(H_1)| + |V(H_2)| = |V(H_1')| + |V(H_2')|$ and $|V(H_1')| < |V(H_1)| \le |V(H_2)| < |V(H_2')|$. Construct the graph G (resp. G') by identifying p_1 to a vertex in H_1 (H_1') and p_l to a vertex in H_2 (H_2'). If $|V(G_i)| \le |V(G_{l+1-i})|$ $(1 \le i \le \lceil \frac{l}{2} \rceil - 1)$, then $J_G(P) < J_{G'}(P)$.*

Proof. Since $|V(H_1')| < |V(H_1)| \le |V(H_2)| < |V(H_2')|$ and $|V(G_i)| \le |V(G_{l+1-i})|$ $(1 \le i \le \lceil \frac{l}{2} \rceil - 1)$, $\sigma_G(p_i) > \sigma_{G'}(p_{l+1-i})$. For $1 \le i \le \lceil \frac{l}{2} \rceil - 1$, we have $\sigma_{G'}(p_i) - \sigma_G(p_i) = \sigma_G(p_{l+1-i}) - \sigma_{G'}(p_{l+1-i}) > 0$. Let $x = \sigma_{G'}(p_{l+1-i})$, $y = \sigma_{G'}(p_{l-i})$, $w = \sigma_G(p_i)$, $z = \sigma_G(p_{i+1})$, $a = \sigma_{G'}(p_i) - \sigma_G(p_i)$, $a' = \sigma_{G'}(p_{i+1}) - \sigma_G(p_{i+1})$, $b = \sigma_G(p_{l+1-i}) - \sigma_{G'}(p_{l+1-i})$ and $b' = \sigma_G(p_{l-i}) - \sigma_{G'}(p_{l-i})$, then $\frac{b}{x} \ge \frac{a}{w}$, $\frac{b'}{y} \ge \frac{a'}{z}$, $w > x$ and $z \ge y$. So, by Lemma 1,

$$J_G(p_i p_{i+1}) + J_G(p_{l+1-i} p_{l-i}) < J_{G'}(p_i p_{i+1}) + J_{G'}(p_{l+1-i} p_{l-i}), \ 1 \le i \le \lceil \frac{l}{2} \rceil - 1.$$

If l is odd, $J_G(P) < J_{G'}(P)$. If l is even, then $J_G(p_{\frac{l}{2}} p_{\frac{l}{2}+1}) < J_{G'}(p_{\frac{l}{2}} p_{\frac{l}{2}+1})$ by Lemma 2, and so $J_G(P) < J_{G'}(P)$. ∎

Lemma 7 *Let $P = p_1 \cdots p_l$ be a path in a graph G, and B the union of some branches at p_1 not containing p_2. Let G_i be the component of $G - E(P) - E(B)$ containing p_i for $1 \le i \le l$. Define the graph $G' := G - \{p_1 x | x \in N_G(p_1) \cap V(B)\} + \{p_l x | x \in N_G(p_1) \cap V(B)\}$ (see Fig.1). Suppose $|V(G_i)| \le |V(G_{l+1-i})|$ and there is an injection ϕ_i from $E(G_i)$ to $E(G_{l+1-i})$ for $2 \le i \le \lceil \frac{l}{2} \rceil - 1$ (resp. ϕ_1 from $E(G_1)$ to $E(G_l \cup B)$) such that, for any edge $ab \in E(G_i)$ and $\phi_i(ab) = a'b'$, either $\sigma_G(a) \ge \sigma_{G'}(a')$ and $\sigma_G(b) \ge \sigma_{G'}(b')$ or $\sigma_G(a) \ge \sigma_{G'}(b')$ and $\sigma_G(b) \ge \sigma_{G'}(a')$. Furthermore if $\phi_1(E(G_1)) \cap B \ne \emptyset$, assume $|V(G_l)| \ge |V(G_1)| + |V(B)|$. Then $J(G) \le J(G')$ and $J(G) < J(G')$ if one of the above inequations is strict.*

Proof. Let t_j be any vertex in G_j $(1 \le j \le l)$. For $1 \le i \le \lfloor \frac{l}{2} \rfloor$, $\sigma_{G'}(t_i) - \sigma_G(t_i) = (l + 1 - 2i)|V(B)| = \sigma_G(t_{l+1-i}) - \sigma_{G'}(t_{l+1-i}) > 0$. Hence, for any edge $e \in E(G_i)$ $(1 \le i \le \lfloor \frac{l}{2} \rfloor)$, if $\phi_i(e) \in E(G_{l+1-i})$, then we have $J_G(e) + J_G(\phi_i(e)) \le J_{G'}(e) + J_{G'}(\phi_i(e))$ by Lemma 1. If $e \in E(G_1)$ and $\phi_1(e) \in E(B)$, for any end vertex t_1 of e and any end vertex b of $\phi_1(e)$, $\sigma_G(b) - \sigma_{G'}(b) = (l-1)(|V(G_l)| - |V(G_1)|) + (l-3)(|V(G_{l-1})| - |V(G_2)|) + \cdots + (\lceil \frac{l}{2} \rceil - \lfloor \frac{l}{2} \rfloor)(|V(G_{\lceil \frac{l}{2} \rceil})| - |V(G_{\lfloor \frac{l}{2} \rfloor})|) \ge 0$.
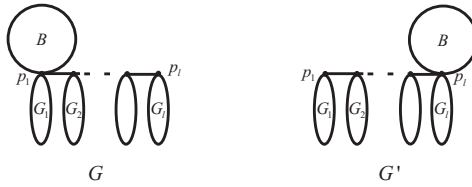
Figure 1: G and G' in Lemma 7.

Combining that $|V(G_l)| \geq |V(G_1)| + |V(B)|$, so $\sigma_G(b) - \sigma_{G'}(b) \geq \sigma_{G'}(t_1) - \sigma_G(t_1)$. Then we also have $J_G(e) + J_G(\phi_1(e)) \leq J_{G'}(e) + J_{G'}(\phi_1(e))$ by Lemma 1.

For the weight of P in G and G', $J_G(P) \leq J_{G'}(P)$ by Lemma 6.

For any edge e not considered above, $J_G(e) \leq J_{G'}(e)$ obviously.

So, $J(G) \leq J(G')$, and $J(G) < J(G')$ if one of the above inequations is strict. ∎

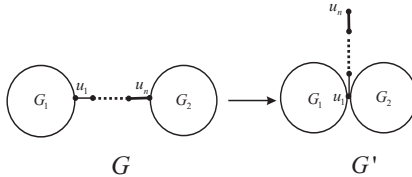By Lemma 7, we can easily get two corollaries in the following.

Figure 2: The path-lifting transformation from graph G to G'.

Corollary 8 *Let $P = u_1 u_2 \cdots u_n$ be a path in a graph G with $d(u_2) = d(u_3) = \cdots = d(u_{n-1}) = 2$ and $d(u_1), d(u_n) \geq 2$, let G_1 be the component of $G - u_2$ containing u_1, and let G_2 be the component of $G - u_{n-1}$ containing u_n. Construct $G' = G - \{u_n v | v \in V(G_2) \cap N_G(u_n)\} + \{u_1 v | v \in V(G_2) \cap N_G(u_n)\}$ (See Fig. 2), then $J(G') > J(G)$.*

Definition 9 *The transformation from G to G' in Corollary 8 is called the path-lifting transformation (on path P) of G.*

Corollary 10 *Let G_0 be a graph, and $P = v_0 v_1 \cdots v_t$ a path of length t. Let G (resp. G') be the graph obtained from G_0 and P by identifying a vertex v_{k+1} (v_k), $k < \lfloor t/2 \rfloor$, to the same vertex of G_0 (See Fig. 3). Then $J(G) > J(G')$.*
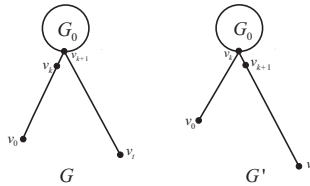
Figure 3: the path-moving transformation from G to G'.

Definition 11 *We call the transformation in Corollary 10 from G to G' the path-moving transformation (on path P) of G.*

GRAPHS (TREES) WITH EXTREME BALABAN INDEX AMONG GRAPHS (TREES) WITH n VERTICES

For any tree T with n vertices but star S_n or path P_n, T can be transformed into S_n by carrying out path-lifting transformation repeatedly and P_n by carrying out via path-lifting transformation repeatedly. So we can make the following conclusion by Corollary 8:

Theorem 12 *If T is a tree with n vertices, $n > 1$, then $J(S_n) \geq J(T) \geq J(P_n)$. The lower bound is realized if and only if $T \cong P_n$ and the upper bound is realized if and only if $T \cong S_n$.*

Theorem 13 *If G is a connected graph with n vertices, then $J(K_n) \geq J(G) \geq J(P_n)$. The lower bound is realized if and only if $G \cong P_n$ and the upper bound is realized if and only if $G \cong K_n$.*

Proof. In the set of connected graphs with n vertices, K_n is the only graph with the maximum edge number and then it has the maximum Balaban index by Lemma 3. Note that trees have the minimum edge number in the set of connected graphs with n vertices and Lemma 12 tells us that among all the trees with n vertices, the tree with the minimum Balaban index is P_n. Hence P_n has the minimum Balaban index of connected graphs with n vertices. ∎

GRAPHS WITH EXTREME BALABAN INDEX AMONG k-CONNECTED (k-EDGE-CONNECTED) GRAPHS WITH n VERTICES

Let G and H be two graphs with $V(G) \cap V(H) = \emptyset$. Denote the disjoint union of G and H by $G \cup H$. The join of G and H, denoted $G+H$, is the graph with vertex set $V(G+H) = V(G) \cup V(H)$ and edge set $E(G + H) = E(G) \cup E(H) \cup \{uv | u \in V(G), v \in V(H)\}$.

Theorem 14 *If G is a k-connected (k-edge-connected) graph with n vertices, then $J(G) \leq J(K_k + (K_1 \cup K_{n-k-1}))$. The bound is realized if and only if $G \cong K_k + (K_1 \cup K_{n-k-1})$.*

Proof. Let graph H have the maximum Balaban index among all k-connected graphs with n vertices and X be a vertex cut of H with $|X| = k$. Denote the components of $H - X$ by $H_1, H_2, \cdots, H_\omega$. In fact, each of the subgraphs $H[X], H_1, H_2, \cdots, H_\omega$ must be complete. Otherwise, by adding an edge between two nonadjacent vertices in a subgraph not complete, we would arrive at a graph with the same number of vertices and the same connectivity, but larger Balaban index. It is a contradiction.

If $\omega > 2$, by adding an edge between a vertex in H_1 and a vertex in H_ω the resulting graph would still have connectivity k, but its Balaban index would increase, a contradiction. So, $H - X$ has exactly two components H_1 and H_2. By a similar argument, we conclude that any vertex in H_1 and H_2 is adjacent to any vertex in X.

Denote the number of vertices of H_1 by n_1 and that of H_2 by n_2. Without loss of generality, suppose $n_1 \leq n_2$. If $n_1 = 1$, $H \cong K_k + (K_1 \cup K_{n-k-1})$. Hence suppose $n_1 > 1$. Let $v \in V(H_1)$ and $H' = H - \{vu | u \in V(H_1 - v)\} + \{uw | u \in V(H_1 - v), \ w \in V(H_2)\}$. Then for any vertex $u \in V(H_1) \cup V(H_2) - v$ and $w \in X$, $\sigma_{H'}(u) < \sigma_H(u)$ and $\sigma_{H'}(w) = \sigma_H(w)$. Let $v' \in V(H_1) - v$. Clearly $J_{H'}(H' - \bigcup_{w \in X}\{vw\} - \bigcup_{w \in X}\{v'w\}) > J_H(H - \bigcup_{w \in X}\{vw\} - \bigcup_{w \in X}\{v'w\})$, $\sigma_{H'}(v) > \sigma_H(v')$ and $\sigma_H(v) - \sigma_{H'}(v) \leq \sigma_{H'}(v') - \sigma_H(v') < 0$. By Lemma 1, $J_{H'}(e) + J_{H'}(e') > J_H(e) + J_H(e')$ for any edge $e \in \bigcup_{w \in X}\{vw\}$ and $e' \in \bigcup_{w \in X}\{v'w\}$. So $J(H') > J(H)$ but H' is also k-connected, a contradiction. It follows $n_1 = 1$, that is, $H \cong K_k + (K_1 \cup K_{n-k-1})$.

Let Q be a graph with the maximum Balaban index among k-edge-connected graphs with n-vertices. Now we show that $Q \cong K_k + (K_1 \cup K_{n-k-1})$. If Q is also k-vertex-connected, the conclusion is true; otherwise, the connectivity of Q is less than k. By adding some edges, Q can be transformed into a new k-connected graph Q'. By Lemma 3, $J(Q) < J(Q')$. Since $J(Q') \leq J(K_k + (K_1 \cup K_{n-k-1}))$, $J(Q) < J(K_k + (K_1 \cup K_{n-k-1}))$. Note that $K_k + (K_1 \cup K_{n-k-1})$ is also k-edge-connected. This contradicts the assumption. ∎

The natural question is to ask for the graph having the minimum Balaban index of k-connected (k-edge-connected) graphs with n vertices. This problem seems to be much more difficult and, at this moment, we cannot offer any solution except the graph having the minimum Balaban index of 2-connected graphs with n vertices.

Lemma 15 [*46*] *Let G be a 2-connected graph with n vertices. Then for any vertex $v \in V(G)$ we have $\sigma(v) \leq \lfloor \frac{1}{4} n^2 \rfloor$. Moreover, this bound is achieved for every vertex $v \in V(G)$ iff G is a cycle.*

Since C_n is a minimally 2-connected graph and the distance of each vertex reaches the upper bound given in Lemma 15, the following theorem is clear. By simply calculating, we get that $J(C_n)$ depends on the parity of n. Moreover, for even n, all Balaban indices are equal to 2.

Theorem 16 *If G is a 2-connected graph with n vertices, then $J(G) \geq J(C_n)$. The bound is realized if and only if $G \cong C_n$.*

THE GRAPH WITH THE MAXIMUM BALABAN INDEX AMONG GRAPHS WITH n VERTICES AND DIAMETER d

The diameter of a graph G, denoted by $diam(G)$, is the maximum distance between any two vertices of G. A graph G is said to be maximal with respect to its diameter if $diam(G + e) < diam(G)$ for every edge e in the complement of G. The following Lemma characterizes such graphs.

Lemma 17 [*52, 53*] *A graph G of diameter d is maximal with respect to the diameter if and only if there exists a vertex v_0 such that the distance layers V_i, where $V_i = \{x | d_G(v_0, x) = i\}$ $(i = 0, 1, \cdots, d)$, fulfills the condition that the induced subgraphs $G_i = G(V_{i-1} \cup V_i)$ are complete whenever $1 \leq i \leq d$ and $E(G) = E(G_1) \cup \cdots \cup E(G_d)$. Further, if $d \geq 2$ then $|V_d| = 1$.*

In [46], Plesnik has shown that among graphs with n vertices and diameter d the minimum Wiener index is achieved by a maximal graph of diameter d each of all layers of which, except middle layer, has exactly one vertex, where the middle layer refers to $\frac{1}{2}d$ layer if d is even, and $\lfloor \frac{1}{2}d \rfloor$ layer and $\lceil \frac{1}{2}d \rceil$ layer if d is odd. We will show, whatever d is odd or even, the graph having the maximum Balaban index among graphs with n vertices and diameter d is also isomorphic to a maximal graph of diameter d, denoted by G_d, where each of all layers, except exactly one of $\lfloor \frac{1}{2}d \rfloor$ layer and $\lceil \frac{1}{2}d \rceil$ layer, has exactly one vertex.

Theorem 18 *Let G be a graph with n vertices and diameter d, then $J(G) \leq J(G_d)$ and the bound is achieved iff $G \cong G_d$.*

Proof. First we suppose that G is a graph with the maximum Balaban index among graphs with n vertices and diameter d. Clearly, G is a maximal graph of diameter d and thus it has the form given in Lemma 17. Suppose that $|V_{\lceil \frac{1}{2}d \rceil}| > 1$ and there exists a layer V_k $(k \neq \lceil \frac{1}{2}d \rceil)$ with $|V_k| > 1$. There are two cases to be considered.

Case 1. $k < \lceil \frac{1}{2}d \rceil$. We can assume that k is the least number with $|V_k| > 1$. Choose a vertex $x \in V_k$ and form a new maximal graph G^* of diameter d with layers $V_0, \cdots, V_{k-1}, \{x\}, (V_k - \{x\}) \cup V_{k+1}, V_{k+2}, \cdots, V_d$. One can easily count that:

$$\sigma_{G^*}(v) - \sigma_G(v) = \begin{cases} |V_k| - 1, & v \in V_0 \cup \cdots \cup V_{k-1}; \\ 1 - |V_k|, & v \in V_{k+2} \cup \cdots \cup V_d; \\ 0, & v \in V_{k+1} \cup \{x\}; \\ \sum_{i=0}^{k-1} |V_i| - \sum_{i=k+2}^{d} |V_i| & v \in V_k - \{x\} \end{cases}$$

For any vertices $u \in V_i$ and $v \in V_{2k+1-i}$ $(0 \leq i \leq k-1)$, $\sigma_G(u) - \sigma_G(v) = (2k-2i)(|V_{2k+1-i}| + \cdots + |V_d| - |V_i| - \cdots - |V_1|) + \sum_{j=1}^{\lfloor \frac{k-i-1}{2} \rfloor} (2k + 1 - 2i - 2j)(|V_{2k+1-i-j}| - |V_{i+j}|) \geq 0$. By Lemma 7, we can easily show $J(G^*) > J(G)$, a contradiction.

Case 2. $k > \lceil \frac{1}{2}d \rceil$. We can assume that k is the maximal possible number with $|V_k| > 1$ and prove similarly as Case 1, again a contradiction.

Thus, we get the result. ∎

Now we have charactered the graph with the maximum Balaban index among graphs with n vertices and diameter d. Furthermore, we will show that the tree with the maximum Balaban index among trees with n vertices and diameter d is constructed by coinciding a centroid of path P_{d+1} with the nonpendent vertex of star S_{n-d}, denoted by $PS(n,d)$.

Recall a tree is called a caterpillar if the deletion of its endvertices produces a path.

Theorem 19 *Let T be a tree with n vertices and diameter d, then $J(T) \leq J(PS(n,d))$, and the bound is achieved iff $T \cong PS(n,d)$.*

Proof. Suppose T is the tree with the maximum Balaban index among trees with n vertices and diameter d. Let path $P = p_1 p_2 \cdots p_{d+1}$. If there is a nonpendent edge e not in P, we do a path-lifting transformation on e to obtain a new tree T'. Then T' also is a tree with n vertices and diameter d but $J(T') > J(T)$ by Corollary 8, a contradiction. So T is a caterpillar. Denote $i = min\{k|d(p_k) > 2, \ 1 \leq k \leq d+1\}$ and $j = max\{k|d(p_k) > 2, \ 1 \leq k \leq d+1\}$.

If $i \neq j$, suppose, without loss of generality, $i \leq (d+1)/2$ and construct a tree T'' such that $V(T) = V(T'')$ and

$$E(T'') = E(T) - \bigcup_{w \in N(p_i) - p_{i-1} - p_{i+1}} p_i\, w + \bigcup_{w \in N(p_i) - p_{i-1} - p_{i+1}} p_{i+1}\, w\ .$$

Then $J(T'') > J(T)$ by Lemma 7, a contradiction. So $i = j$. And if $i \neq \lfloor d/2 \rfloor$ or $\lceil d/2 \rceil$, then do a path-moving transformation on P and get a contradiction by Corollary 10. Then we get the result. ∎

References

[1] A. T. Balaban, O. Ivanciuc, Historical development of topological indices, in: J. Devillers, A. T. Balaban (Eds.), *Topological Indices and Related Descriotors in QSAR and QSPR*, Gordon and Breach, 1999, pp. 21–57.

[2] H. Wiener, Structural determination of paraffin boiling points, *J. Am. Chem. Soc.* **69** (1947) 17–20.

[3] M. Randić, On characterization of molecular branching, *J. Am. Chem. Soc.* **97** (1975) 6609–6615.

[4] A. T. Balaban, Highly discriminating distance–based topological index, *Chem. Phys. Lett.* **89** (1982) 399–404.

[5] A. T. Balaban, Topological indices based on topological distances in molecular graphs, *Pure Appl. Chem.* **55** (1983) 199–206.

[6] X. Guo, M. Randić, Trees with the same topological index JJ, *SAR and QSAR Environ. Res.* **10** (1999) 381–394.

[7] A. T. Balaban, A comparison between various topological indices, particularly between the index J and Wiener's index W, in: D. H. Rouvray, R. B. King (Eds.), *Topology in Chemistry — Discrete Mathematics of Molecules*, Horwood, Chichester, 2002, pp. 89–112.

[8] A. T. Balaban, N. Ionescu–Pallas, T. S. Balaban, Asymptotic values of topological indices J and J' (average distance sum connectivities) for infinite cyclic and acyclic graphs, *MATCH Commun. Math. Comput. Chem.* **17** (1985) 121–146.

[9] N. Jia, K. W. McLaughlin, Fibonacci trees: A study of the asymptotic behavior of Balaban's index, *MATCH Commun. Math. Comput. Chem.* **51** (2004) 79–95.

[10] T. S. Balaban, A. T. Balaban, D. Bonchev, A topological approach to predicting properties of infiite polymers. Part VI. Rational formulas for the normalized Wiener index and a comparison with index J, *J. Mol. Struct. (Theochem)* **535** (2001) 81–92.

[11] J. Devillers, A. T. Balaban (Eds.), *Topological Indices and Related Descriptors in QSAR and QSPR*, Gordon & Breach, Amsterdam, 1999.

[12] M. P. Hanson, D. H. Rouvray, Novel applications of topological indices. 2. Prediction of the threshold soot index for hydrocarbon fuels, *J. Phys. Chem.* **91** (1987) 2981–2985.

[13] M. Vedrina, S. Nikolić, M. Medić-Šarić, N. Trinajstić, TAM: a program for the calculation of topological indices in QSPR and QSAR studies, *Computers Chem.* **6** (1997) 355–361.

[14] J. F. Caputo, K. J. Cook, A graph–theoretical approach to the prediction of physical properties of alkanes based on the distance matrix, *Pharm. Res.* **6** (1989) 809–812.

[15] S. Gupta, M. Singh, A. K. Madan, Connective eccentricity index: a novel topological descriptor for predicting biological activity, *J. Mol. Graphics Model.* **18** (2000) 18–25.

[16] M. Jalali–Heravi, M. H. Fatemi, Prediction of thermal conductivity detection response factors using an artificial neural network, *J. Chromatogr. A* **897** (2000) 227–235.

[17] M. Jalali–Heravi, M. Asadollahi–Baboli, QSAR analysis of platelet–derived growth inhibitors using GA-ANN and shuffling crossvalidation, *QSAR Comb. Sci.* **27** (2008) 750–757.

[18] M. Randić, M. Pompe, The variable connectivity index $^1\chi^f$ versus the traditional molecular descriptors: a comparative study of $^1\chi^f$ against descriptors of codessa, *J. Chem. Inf. Model.* **41** (2001) 631–638.

[19] E. Estrada, Generalization of topological indices, *Chem. Phys. Lett.* **336** (2001) 248–252.

[20] E. Estrada, Generalized graph matrix, graph geometry, quantum chemistry, and optimal description of physicochemical properties, *J. Phys. Chem. A* **107** (2003) 7482–7489.

[21] E. Estrada, A. R. Matamala, GTI-space: the space of generalized topological indices, *J. Math. Chem.* **43** (2008) 508–517.

[22] J. Zheng, Y. Polyakova, K. H. Row, Prediction of retention of nucleic compounds based on a QSPR model, *Chromatographia* **64** (2006) 1–9.

[23] A. Krawczuk, A. Voelkel, J. Lulek, R. Urbaniak, K. Szyrwiska, Use of topological indices of polychlorinated biphenyls in structure–retention relationships, *J. Chromatogr. A* **1018** (2003) 63–71.

[24] A. Ratkiewicz, T. N. Truong, Application of chemical graph theory for automated mechanism generation, *J. Chem. Inf. Comput. Sci.* **43** (2003) 36–44.

[25] S. N. Rao, T. P. Stockfisch, Partially unified multiple property recursive partitioning (PUMP-RP) analyses of cyclooxygenase (COX) inhibitors, *J. Chem. Inf. Comput. Sci.* **43** (2003) 1614–1622.

[26] J. M. Santiuste, Relationships between retention data of benzene and chlorobenzenes with their physico–chemical properties and topological indices, J. M. Takacs, *Chromatographia* **58** (2003) 87–96.

[27] K. Roy , G. Ghosh, QSTR with extended topochemical atom indices. 4. Modeling of the acute toxicity of phenylsulfonyl carboxylates to vibrio fischeri using principal component factor analysis and principal component regression analysis, *QSAR Comb. Sci.* **23** (2004) 526–535.

[28] C. Rücker, G. Rücker, M. Meringer, Exploring the limits of graph invariant– and spectrum–based discrimination of (sub) structures, *J. Chem. Inf. Comput. Sci.* **42** (2002) 640–650.

[29] A. Jezierska, Counter–propagation artificial neural network as a tool for the independent variable selection: structure–mutagenicity study on aromatic amines, *Mol. Divers.* **8** (2004) 371–377.

[30] H. Gonzalez–Diaz, G. Aguero, M. A. Cabrera, R. Molina, L. Santana, E. Uriarte, G. Delogu, N. Castanedo, Unified Markov thermodynamics based on stochastic forms to classify drugs considering molecular structure, partition system, and biological species: distribution of the antimicrobial G1 on rat tissues. *Bioorg. Med. Chem. Lett.* **15** (2005) 551–557.

[31] M. R. McGillen, C. J. Perciva, T. Raventos–Duran, G. Sanchez–Reyn, D. E. Shallcross, Can topological indices be used to predict gas-phase rate coefficients of importance to tropospheric chemistry? Free radical abstraction reactions of alkanes, *Atmospheric Env.* **40** (2006) 2488–2500.

[32] A. Panaye, J. P. Doucet, J. Devillers, N. Marchand–Geneste, J. M. Porcher, Decision trees versus support vector machine for classification of androgen receptor ligands, *SAR QSAR Env. Res.* **19** (2008) 129–151.

[33] V. Rastija, M. Medric–Saric, QSAR Study of antioxidant activity of wine polyphenols, *Eur. J. Med. Chem.* **44** (2009) 400–408.

[34] Y. Polyakova, M. Long, K. H. Row, Linear regression based QSPR models for the prediction of the retention mechanism of some nitrogen containing heterocycles, *J. Liq. Chromatogr. Rel. Techn.* **29** (2006) 533–552.

[35] C. R. Munteanu, H. Gonzalez–Diaz, F. Borges, A.L. Magalhaes, Natural random protein classification models based on star network topological indices, *J. Theo. Bio.* **254** (2008) 775–783.

[36] H. Dureja, A. K. Madan, Superaugmented eccentric connectivity indices: A new generation of highly discriminating topological descriptors for QSAR/QSPR modeling, *Med. Chem. Res.* **16** (2007) 331–341.

[37] S. Prasanna, E. Manivannan, S. C. Chaturvedi, QSAR analyses of conformationally restricted 1,5-diaryl pyrazoles as selective COX-2 inhibitors: application of connection table representation of ligands, *Bioorg. Med. Chem.* **15** (2005) 2097–2102.

[38] A. Thakur, M. Thakur, P. V. Khadikar, C. T. Supuran, P. Sudele, QSAR study on benzenesulphonamide carbonic anhydrase inhibitors: topological approach using Balaban index, *Bioorg. Med. Chem.* **12** (2004) 789–793.

[39] G. Grassy, B. Calas, A. Yasri, R. Lahana, J. Woo, S. Iyer, M. Kaczorek, R. Floc'h, R. Buelow, Computer–assisted rational design of immunosuppressive compounds, *Nature Biotechnol.* **16** (1988) 748–752.

[40] C. Bermudex, E. E. Daza, E. Andrade, Characterization and comparison of Escherichia coli transfer RNA by graph theory based on secondary structure, *J. Theor. Biol.* **197** (1999) 193–205.

[41] W. Shu, X. Bo, Z. Zheng, and S. Wang, A novel representation of RNA secondary structure based on element-contact graphs, *BMC Bioinformatics* **9** (2008) 188–196.

[42] E. Estrada, Y. Gutierrez, The Balaban *J* index in the multidimensional space of generalized topological indices. Generalizations and QSPR improvements, *MATCH Commun. Math. Comput. Chem.* **44** (2001) 155–16.

[43] A. A. Dobrynin, I. Gutman, S. Klavžar, P. žigert, Wiener index of hexagonal systems, *Acta Appl. Math.* **72** (2002) 247–294.

[44] A. A. Dobrynin, R. Entringer, I. Gutman, Wiener index of trees: theory and applications, *Acta Appl. Math.* **66** (2001) 211–249.

[45] R. C. Entringer, D. E. Jackson, D. A. Snyder, Distance in graphs, *Czech. Math. J.* **26** (1976) 283–296.

[46] J. Plesnik, On the sum of all distances in a graph or digraph, *J. Graph Theory* **8** (1984) 1–21.

[47] I. Gutman, S. Zhang, Graph connectivity and Wiener index, *Bull. Acad. Serbe Sci. Arts (Cl. Sci. Math. Natur.)* **133** (2006) 1–5.

[48] X. Li, Y. Shi, A survey on the Randić index, *MATCH Commun. Math. Comput. Chem.* **59** (2008) 127–156.

[49] B. Zhou, N. Trinajstić, Bounds on the Balaban index, *Croat. Chem. Acta.* **81** (2008) 319–323.

[50] C. Jordan, Sur les assemblages de lignes, *J. Reine Angew. Math.* **70** (1869) 185–190.

[51] C. A. Barefoot, R. C. Entringer, L. A. Szely, External values for ratios of distance in trees, *Discr. Appl. Math.* **80** (1997) 37–56.

[52] O. Ore, Diameters in graphs, *J. Comb. Theory* **5** (1968) 75–81.

[53] R. D. Luce, Commectivity and generalized cliques in sociometric group structure, *Psychometrika* **15** (1950) 169–190.