MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **76** (2016) 69-80

ISSN 0340 - 6253

Entropy of Weighted Graphs with the Degree–Based Topological Indices as Weights

Ramin Kazemi

Department of Statistics, Imam Khomeini International University, Qazvin, Iran

r.kazemi@SCI.ikiu.ac.ir

(Received December 6, 2015)

Abstract

Topological indices are numerical parameters of a graph which characterize its topology and are usually graph invariant. Among the large number of existing topological indices, an important class of such measures relies on Shannon's entropy to characterize graphs by determining their structural information content. In this paper, we study the entropy of weighted graphs with the degree–based topological indices (specially, the first and second Zagreb indices, the general Randić index, the harmonic and sum–connectivity index) as weights.

1 Introduction

A graph G is an ordered pair of sets $V(G)$ and $E(G)$ such that the elements $uv \in E(G)$ are a sub-collection of the unordered pairs of elements of $V(G)$. For convenience, we denote a graph by $G = (V, E)$ sometimes. The elements of $V(G)$ are called vertices and the elements of $E(G)$ are called edges. If $e = uv$ is an edge, then we say vertices u and v are adjacent, and u, v are two endpoints (or ends) of e. If G is a graph with n vertices and m edges, then we say the order of G is n and the size of G is m. A graph of order n is addressed as an n-vertex graph. A graph is connected if, for every partition of its vertex set into two nonempty sets X and Y, there is an edge with one end in X and one end in Y. Otherwise, the graph is disconnected. In other words, a graph is disconnected if its vertex set can be partitioned into two nonempty subsets X and Y so that no edge has one end in X and one end in Y. All vertices adjacent to vertex u are called neighbors of

u. The neighborhood of u is the set of the neighbors of u. The number of edges adjacent to vertex u is the degree of u, denoted by d_u. Vertices of degrees 0 and 1 are said to be isolated and pendent vertices, respectively. A pendent vertex is also referred to as a leaf of the underlying graph. A connected graph without any cycle is a *tree*. A weighted graph is a graph in which a number (the weight) is assigned to each edge. Such weights might represent for example costs, lengths or capacities, depending on the problem at hand. Some authors call such a graph a network.

A topological index is a numerical parameter mathematically derived from the graph structure. It is a graph invariant thus it does not depend on the labelling or pictorial representation of the graph. The topological indices of molecular graphs are widely used for establishing correlations between the structure of a molecular compound and its physico-chemical properties or biological activity (e.g., pharmacology). Gutman [3] introduced the following general form for degree–based topological indices:

$$TI = TI(G) = \sum_{uv \in E(G)} F(d_u, d_v), \tag{1}$$

where the summation goes over all pairs of adjacent nodes u, v of graph G, and where $F = F(x, y)$ is an appropriately chosen function. In particular, $F(x, y) = x+y$ for the first Zagreb index, $F(x, y) = xy$ for the second Zagreb index, $F(x, y) = (xy)^\lambda$ $(\lambda \in \mathbb{R})$ for the general Randić index, $F(x, y) = 2(x+y)^{-1}$ for the harmonic index and $F(x, y) = (x+y)^{-\frac{1}{2}}$ for the sum–connectivity index [6].

The plan of the paper is as follows. In Section 2, the entropy of weighted graphs is defined. The main results are given in Section 3. In this section, we start with the constant weights. Then we give the main theorem (Theorem 4) with the essential Equation (5) for degree–based topological indices. In passing, we study some well-known graphs.

2 Entropy of Weighted Graphs

Studies of the information content of complex networks and graphs have been initiated in the late 1950s based on the seminal work due to Shannon. Numerous measures for analyzing complex networks quantitatively have been contributed. A variety of problems in, e.g., discrete mathematics, computer science, information theory, statistics, chemistry, biology, etc., deal with investigating entropies for relational structures. For example, graph entropy measures have been used extensively to characterize the structure of graph–

based systems in mathematical chemistry, biology and in computer science–related areas [1].

Rashevsky is the first who introduced the so-called structural information content based on partitions of vertex orbits [12]. Mowshowitz used the the same measure and proved some properties for graph operations (sum, join, etc.) [11]. Moreover, Rashevsky used the concept of graph entropy to measure the structural complexity of graphs. Mowshowitz introduced the entropy of a graph as an information–theoretic quantity, and he interpreted it as the structural information content of a graph. Mowshowitz later studied mathematical properties of graph entropies measures thoroughly and also discussed special applications thereof.

For a given graph G and vertex v_i, let d_i be the degree of v_i. For an edge v_iv_j, one defines:

$$p_{ij} = \frac{w(v_iv_j)}{\sum_{j=1}^{d_i} w(v_iv_j)}, \tag{2}$$

where $w(v_iv_j)$ is the weight of the edge v_iv_j and $w(v_iv_j) > 0$. The node entropy has been defined by:

$$H(v_i) = -\sum_{j=1}^{d_i} p_{ij} \log(p_{ij}). \tag{3}$$

Motivated by this method, Chen *et al.* [2] introduced the definition of the entropy of edge-weighted graphs, which also can be interpreted as multiple graphs. For an edge–weighted graph, $G = (V, E, w)$, where V, E and w denote the vertex set, the edge set and the edge weight of G, respectively.

Definition 1. *For an edge weighted graph $G = (V, E, w)$, the entropy of G is defined by:*

$$I(G, w) = -\sum_{uv \in E} p_{u,v} \log p_{u,v}, \tag{4}$$

where

$$p_{u,v} = \frac{w(uv)}{\sum_{uv \in E} w(uv)}.$$

The above definition of the entropy for edge–weighted graphs is based on the probability function (2).

3 Main Results

As our first result, we prove the following theorem.

Theorem 1. *Let $G = (V, E, w)$ be a graph with n vertices, m edges and $w(e) = c > 0$ for each edge e, where c is a constant. Then*

$$\log(n-1) \leq I(G, w) \leq \log \binom{n}{2}.$$

Proof. By definition,

$$I(G, w) = -\sum_{uv \in E} \frac{1}{m} \log \frac{1}{m} = \log m.$$

Proof is completed since $n - 1 \leq m \leq \binom{n}{2}$. ■

As a corollary, for a tree T with n vertices, $I(T, w) = \log(n-1)$.

Example 1. *Let $w(e) = w(uv) = d_u + d_v$. Then $\sum_{uv \in E(G)} w(e) = M_1(G)$ where M_1 is the first Zagreb index.*

1- The path P_n is a tree of order n with exactly two pendent vertices. Thus

$$M_1(P_n) = 4n - 6.$$

2- The star of order n, denoted by S_n, is the tree with $n-1$ pendent vertices. Thus

$$M_1(S_n) = n(n-1).$$

3- A complete graph K_n is a simple undirected graph in which every pair of distinct vertices is connected by a unique edge. Thus

$$M_1(K_n) = n(n-1)^2.$$

Then

$$\begin{aligned} I(P_n, w) &= -\frac{4(n-3)}{4n-6} \log \frac{4}{4n-6} - \frac{6}{4n-6} \log \frac{3}{4n-6}, \\ I(S_n, w) &= \log(n-1), \\ I(K_n, w) &= -\log \frac{2}{n(n-1)}. \end{aligned}$$

Theorem 2. *Let $G = (V, E, w)$ be a regular graph with $n \geq 3$ vertices, m edges and $w(e) = F(d_u, d_v)$. Then*

$$\log n \leq I(G, w) \leq \log \binom{n}{2}.$$

Proof. Suppose G is k-regular. Then, $k \geq 2$, since G is connected and $n \geq 3$. By definition,

$$I(G, w) = -\sum_{uv \in E} \frac{F(k,k)}{\sum_{uv \in E} F(k,k)} \log \frac{F(k,k)}{\sum_{uv \in E} F(k,k)} = \log \frac{nk}{2}.$$

Proof is completed since $2 \leq k \leq n - 1$. ■

Theorem 3. *Let $G = (V, E, w)$ be a complete bipartite graph with n vertices and $w(e) = F(d_u, d_v)$. Then*

$$\log(n-1) \leq I(G,w) \leq \log\left(\left\lfloor \frac{n}{2} \right\rfloor \left\lceil \frac{n}{2} \right\rceil\right).$$

Proof. Suppose G is a complete bipartite graph with n vertices, and the two parts have n_1 and n_2 vertices, respectively. Therefore, $n = n_1 + n_2$. By definition,

$$I(G,w) = -\sum_{uv\in E} \frac{F(n_1,n_2)}{\sum_{uv\in E} F(n_1,n_2)} \log \frac{F(n_1,n_2)}{\sum_{uv\in E} F((n_1,n_2)} = \log(n_1 n_2).$$

■

Theorems 2 and 3 show that the entropy of weighted regular and complete bipartite graphs do not depend on the weights.

Theorem 4. *Let $G = (V, E, F(d_u, d_v))$ be a connected graph with n vertices. Also, let $T_m = \min TI(G)$ and $T_M = \max TI(G)$.*

1- For the first and second Zagreb indices,

$$\log \frac{T_m}{F(n-1,n-1)} \leq I(G,w) \leq \log \frac{T_M}{F(1,2)}.$$

2- For the harmonic and sum–connectivity indices,

$$\log \frac{T_m}{F(1,2)} \leq I(G,w) \leq \log \frac{T_M}{F(n-1,n-1)}.$$

Proof. We have

$$\begin{aligned}
I(G,w) &= -\sum_{uv\in E(G)} p_{u,v} \log p_{u,v} \\
&= -\sum_{uv\in E(G)} \frac{F(d_u,d_v)}{TI(G)} \log \frac{F(d_u,d_v)}{TI(G)} \\
&= -\frac{1}{TI(G)} \sum_{uv\in E(G)} F(d_u,d_v) \log \frac{F(d_u,d_v)}{TI(G)} \\
&= -\frac{1}{TI(G)} \sum_{uv\in E(G)} F(d_u,d_v)(\log F(d_u,d_v) - \log TI(G)) \\
&= \log TI(G) - \frac{1}{TI(G)} \sum_{uv\in E(G)} F(d_u,d_v) \log F(d_u,d_v).
\end{aligned}$$

Thus $I(G,w)$ can be expressed as:

$$I(G,w) = \log TI(G) - \frac{1}{TI(G)} \sum_{uv\in E(G)} F(d_u,d_v) \log F(d_u,d_v). \tag{5}$$

1- From (5),

$$\begin{aligned} I(G,w) &\leq \log T_M(G) - \frac{1}{TI(G)} \log F(1,2) \sum_{uv \in E(G)} F(d_u,d_v) \\ &= \log T_M(G) - \log F(1,2). \end{aligned}$$

Also,

$$\begin{aligned} I(G,w) &\geq \log T_m(G) - \frac{1}{TI(G)} \log F(n-1,n-1) \sum_{uv \in E(G)} F(d_u,d_v) \\ &= \log T_m(G) - \log F(n-1,n-1). \end{aligned}$$

Part 2 is proved similarly. ■

Corollary 1. *Let $G = (V, E, F(d_u, d_v))$ be a connected graph with n vertices and $w(e) = (d_u d_v)^\lambda$ where $\lambda \in \mathbb{R}$ (the general Randić index as weight). Then*

1- for $\lambda > 0$,

$$\log \frac{T_m}{(n-1)^{2\lambda}} \leq I(G,w) \leq \log T_M - \lambda.$$

2- for $\lambda < 0$,

$$\log T_M - \lambda \leq I(G,w) \leq \log \frac{T_m}{(n-1)^{2\lambda}}.$$

The following corollary is an immediate consequence of reference [8].

Corollary 2. *1- Let G be a graph with n vertices and no isolated vertices and $w(e) = (d_u d_v)^\lambda$. Then*

i) For $\lambda \in (-1/2, 0)$:

$$\log\Big(\min\{(n-1)^{1+\lambda}, \frac{n}{2}(even\ n), \frac{n-3}{2} + 2^{1+\lambda}(odd\ n)\}\Big) - \lambda \leq I(G,w) \leq \log(n(n-1)) - 1.$$

ii) For $\lambda \in (-\infty, -1)$, when n is even:

$$\log(n(n-1)^{1+2\lambda}) - \lambda - 1 \leq I(G,w) \leq \log n - 2\lambda \log(n-1) - 1$$

and when n is odd:

$$\log(n(n-1)^{1+2\lambda}) - \lambda - 1 \leq I(G,w) \leq \log(n - 3 + 2^{2+\lambda}) - 2\lambda \log(n-1) - 1.$$

2- Let T be a tree with n vertices and $w(e) = (d_u d_v)^\lambda$. Then

i) For $\lambda \in [-1/2, 0]$:

$$(\lambda + 1)\log(n-1) - \lambda \leq I(G,w) \leq \log(1 + (n-3)2^{\lambda-1}) - 2\lambda \log(n-1) + \lambda + 1.$$

ii) For $\lambda \in [-\infty, -2]$, *when* n *is odd:*

$$(\lambda+1)\log(n-1)-\lambda \le I(G,w) \le \log((n-1)^\lambda+2^\lambda)+(1-2\lambda)\log(n-1)-1$$

and when n *is even:*

$$(\lambda+1)\log(n-1)-\lambda \le I(G,w) \le \log(\frac{n-2}{2}((n-2)^\lambda+2^\lambda))+4^\lambda)-2\lambda\log(n-1).$$

The reader see reference [10] for evaluating entropy in simple connected graphs and [13] for trees.

Corollary 3. *Let* G *be a graph with* n *vertices. Let* δ *and* Δ *be the minimum degree and the maximum degree of* G*, respectively.*

1- For the first and second Zagreb indices, the general Randić index ($\lambda > 0$)*:*

$$\log\frac{T_m}{F(\Delta,\Delta)} \le I(G,w) \le \log\frac{T_M}{F(\delta,\delta)}.$$

2- For the harmonic and sum–connectivity indices, the general Randić index ($\lambda < 0$)*:*

$$\log\frac{T_m}{F(\delta,\delta)} \le I(G,w) \le \log\frac{T_M}{F(\Delta,\Delta)}.$$

Let $w(e)=(d_ud_v)^\lambda$ *where* $\lambda \in \mathbb{R}$*. Then*

3- for $\lambda > 0$,

$$\log\frac{T_m}{\Delta^{2\lambda}} \le I(G,w) \le \log\frac{T_m}{\delta^{2\lambda}}$$

and for $\lambda < 0$,

$$\log\frac{T_m}{\delta^{2\lambda}} \le I(G,w) \le \log\frac{T_m}{\Delta^{2\lambda}}.$$

Example 2. *Let* $w(e)=d_u+d_v$ *and* T_n *be a tree with maximum degree* Δ*. Then the first Zagreb index* $\le n^2-3n+2(\Delta+1)$ *[7]. Thus*

$$I(T_n,w) \le \log\frac{n^2-3n+2(\Delta+1)}{2\delta}.$$

Also, $4n-6 \le$ *the first Zagreb index of* $T_n \le n(n-1)$ *[4]. Then*

$$\log\frac{4n-6}{2\Delta} \le I(G,w) \le \log\frac{n(n-1)}{2\delta}.$$

Dendrimers are large and complex molecules with very well–defined chemical structures. They are nearly perfect monodisperse macromolecules with a regular and highly branched three-dimensional architecture [5]. A *dendrimer* is a tree with two additional parameters; the progressive degree t and the radius r. Every internal node of the tree has degree $t+1$. As in every tree, a dendrimer has one (monocentric dendrimer) or two (dicentric dendrimer) central nodes; the radius r denotes the (largest) distance from an external node to the (closer) center. If all external nodes are at a distance r from the center, then the dendrimer is called homogeneous. Internal nodes different from the central nodes are called branching nodes and are said to be on the i-th orbit if their distance to the (nearer) center is r. Every branching vertex has one incoming edge, as well as t outgoing edges. Let $D(t;r)$ denote the dendrimer graph with parameters t and r. If $D(t;r)$ has only one center, then we have $n = 1 + \frac{(t+1)(t^r-1)}{t-1}$. As an example, we show a dendrimer with one center such that $t = 3$ and $r = 3$ in Figure 1.

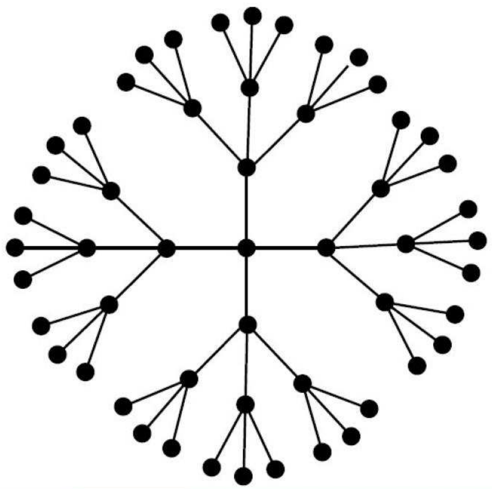

Figure 1: A dendrimer (one center) with $t = 3$ and $r = 3$.

Theorem 5. *Let $D(t;r)$ be a dendrimer with n vertices with only one center and $w(e) = F(d_u, d_v)$. Then*

$$\begin{aligned} I(D(t;r),w) &= -(t+1)t^{r-1}\frac{F(t+1,1)}{TI(D)}\log\frac{F(t+1,1)}{TI(D)} \\ &- (t+1)(1-t^{r-1})\frac{F(t+1,t+1)}{TI(D)}\log\frac{F(t+1,t+1)}{TI(D)}, \end{aligned}$$

where

$$TI(D) = (t+1)t^{r-1}F(t+1,1) + (t+1)(1-t^{r-1})F(t+1,t+1).$$

Proof. If $r = 1$, then $D(t;1)$ is a star. Now, by Example 1 (Part 2), $I(D(t;1),w) = \log(t+1)$ since

$$n = 1 + \frac{(t+1)(t-1)}{t-1} = t+2.$$

If $t = 1$, then $D(1;r)$ is a path. Now, similar to Example 1 (Part 1),

$$I(D(t;1),w) = -(n-3)\frac{F(2,2)}{TI(D)}\log\frac{F(2,2)}{TI(D)} - 2\frac{F(1,2)}{TI(D)}\log\frac{F(1,2)}{TI(D)}.$$

In a dendrimer $D(t;r)$ with one center, there are $(t+1)t^{r-1}$ leaves and both end vertices of any edge have degree $t+1$. Then

$$\begin{aligned} TI(D) &= (t+1)t^{r-1}F(t+1,1) + (n-1-(t+1)t^{r-1})F(t+1,t+1) \\ &= (n-1)(n-2)^{r-1}F(n-1,1) \\ &+ (n-1)(1-(n-2)^{r-1})F(n-1,n-1). \end{aligned}$$

Proof is completed since $n = t+2$. ■

In Theorem 5, we can find the lower and upper bounds for the entropy with evaluating $I(D(t;r),w)$ as a function on t. For example, in the case of Randić weights ($\lambda < 0$), $I(D(t;r),w)$ is an increasing function on t with a maximum at point $t = n-2$ and a minimum at point $t = 1$.

Corollary 4. *If $D(t;r)$ has only two centers, then we have $n = \frac{2(t^{r+1}-1)}{t-1}$. Then we can obtain a result for this structure with the same manner.*

Example 3. *Let $w(e) = d_u + d_v$. Then*

$$TI(D(t;2)) = n(n-1)(n-2) + 2(n-1)^2(3-n).$$

Then

$$\begin{aligned} I(D(t;r),w) &= -(n-1)(n-2)\frac{n}{TI(D(t;2))}\log\frac{n}{TI(D(t;2))} \\ &- (n-1)(3-n)\frac{2n-2}{TI(D(t;2))}\log\frac{2n-2}{TI(D(t;2))}. \end{aligned}$$

A *comet* is a tree composed of a star and a pendent path [9]. For any numbers n and $2 \le t \le n-1$, we denote by $C(n;t)$ the comet of order n with t pendent vertices, i.e., a tree formed by a path P_{n-t} of which one end vertex coincides with a pendent vertex of a star S_{t+1} of order $t+1$. Observe that $C(n;t)$ is the path graph if $t = 2$ and is the star graph if $t = n-1$.

Theorem 6. *Let $C(n;t)$ be a comet and $w(e) = F(d_u, d_v)$. Then*

$$\begin{aligned} I(C(n;t),w) &= -\frac{F(1,2)}{TI(C(n;t))}\log\frac{F(1,2)}{TI(C(n;t))} - \frac{F(2,t)}{TI(C(n;t))}\log\frac{F(2,t)}{TI(C(n;t))} \\ &- (t-1)\frac{F(1,t)}{TI(C(n;t))}\log\frac{F(1,t)}{TI(C(n;t))} \\ &- (n-t-2)\frac{F(2,2)}{TI(C(n;t))}\log\frac{F(2,2)}{TI(C(n;t))}, \end{aligned}$$

where

$$TI(C(n;t)) = F(1,2) + F(2,t) + (t-1)F(1,t) + (n-t-2)F(2,2).$$

Proof. It is enough to note that by definition of a comet,

$$TI(C(n;t)) = F(1,2) + F(2,t) + (t-1)F(1,t) + (n-t-2)F(2,2).$$

■

Example 4. *If $w(e) = d_u + d_v$, then*

$$TI(C(n;t)) = t^2 - 3t + 4n - 4.$$

Thus

$$\begin{aligned} I(C(n;t),w) &= -\frac{3}{TI(C(n;t))}\log\frac{3}{TI(C(n;t))} - \frac{2+t}{TI(C(n;t))}\log\frac{2+t}{TI(C(n;t))} \\ &- (t-1)\frac{t+1}{TI(C(n;t))}\log\frac{t+1}{TI(C(n;t))} \\ &- (n-t-2)\frac{4}{TI(C(n;t))}\log\frac{4}{TI(C(n;t))}, \end{aligned}$$

We can extend our results to the multiplicative version of the topological indices. Let us define

$$p_{u,v} = \frac{w(uv)}{\prod_{uv\in E(G)} w(uv)}.$$

In this case, $\prod_{uv\in E(G)} p_{uv} = 1$ but p_{uv} is not a probability function. Also if $w(e) = F(d_u, d_v)$, then $TI_{mu}(G) = \prod_{uv\in E(G)} w(e)$ is a multiplicative version of a topological index. Similar to (5),

$$I(G,w) = -\frac{TI(G)}{TI_{mu}(G)}\log TI_{mu}(G) - \frac{1}{TI_{mu}(G)}\sum_{uv\in E(G)} F(d_u,d_v)\log F(d_u,d_v).$$

Thus for example,

$$I(G,w) \geq -\log TI_{mu}(G) - \frac{\log F(\Delta,\Delta)}{TI_{mu}(G)}TI(G) \geq -\log(F(\Delta,\Delta)TI_{mu}(G)).$$

4 Conclusion

This paper mainly considered edge weights defined by some degree–based topological indices. For future work, it would be interesting to consider other degree–based topological indices, such as the atom-bond connectivity (ABC) index, which is well studied with applications in chemistry. The entropy for vertex–weighted graphs can be defined similarly, which has already been studied extensively. Also, we can study the entropy of weighted graphs with other version of the topological indices such as reformulated version. Studying the entropy of weighted graphs with an information on degrees or edges is also possible. In the presented results, equality holds if and only if the graph G is a specified graph. For example, In Theorem 2, the left equality holds if and only if G is the cycle graph, and the right equality holds if and only if G is the complete graph.

Acknowledgments: This work was supported by the Imam Khomeini International University in Iran.

References

[1] D. Bonchev, *Information Theoretic Indices for Characterization of Chemical Structures*, Res. Studies Press, Chichester, 1983.

[2] Z. Chen, M. Dehmer, Y. Shi, A note on distance–based graph entropies, *Entropy* **16** (2014) 5416–5427.

[3] I. Gutman, Degree–based topological indices, *Croat. Chem. Acta.* **86** (2013) 351–361.

[4] I. Gutman, K. C. Das, The first Zagreb index 30 years after, *MATCH Commun. Math. Comput. Chem.* **50** (2004) 83–92.

[5] A. Iranmanesh, N. A. Gholami, Computing the Szeged index of styrylbenzene dendrimer and triarylamine dendrimer of generation 1–3, *MATCH Commun. Math. Comput. Chem.* **62** (2009) 371–379.

[6] R. Kazemi, The second Zagreb index of molecular graphs with tree structure, *MATCH Commun. Math. Comput. Chem.* **72** (2014) 753–760.

[7] K. C. Das, I. Gutman, B. Horoldagva, Comparing Zagreb indices and coindices of trees, *MATCH Commun. Math. Comput. Chem.* **98** (2012) 189–198.

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

[9] H. Lu, B. Zhou, Lower bounds for the Randić index R_{-1} of trees, *MATCH Commun. Math. Comput. Chem.* **54** (2005) 435–440.

[10] M. Lu, H. Liu, F. Tian, The connectivity index, *MATCH Commun. Math. Comput. Chem.* **51** (2004) 149–154.

[11] A. Mowshowitz, Entropy and the complexity of the graphs I: An index of the relative complexity of a graph, *Bull. Math. Biophys.* **30** (1968) 175–204.

[12] N. Rashevsky, Life, information theory, and topology, *Bull. Math. Biophys.* **17** (1955) 229–235.

[13] I. Tomescu, M. Jamil, Maximum general sum–connectivity index for trees with given independence number, *MATCH Commun. Math. Comput. Chem.* **72** (2014) 715–722.