Applied Mathematics and Computation 431 (2022) 127325

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Szeged-like entropies of graphs

Matthias Dehmer[a,b,c,d], Frank Emmert-Streib[e], Niko Tratnik[f,g,*], Petra Žigert Pleteršek[f,h]

[a] Department for Biomedical Computer Science and Mechatronics, UMIT—Private University for Health Sciences, Medical Informatics and Technology, Eduard-Wallnöfer-Zentrum 1, 6060, Hall in Tyrol, Tyrol, Austria
[b] Swiss Distance University of Applied Sciences, Brig, Switzerland
[c] College of Artificial Intelligence, Nankai University, Tianjin, China
[d] School of Science, Xian Technological University, Xian, Shaanxi 710021, China
[e] Predictive Society and Data Analytic Lab, Faculty of Information Technology and Communication Sciences, Tampere University, Tampere 33100, Finland
[f] Faculty of Natural Sciences and Mathematics, University of Maribor, Slovenia
[g] Institute of Mathematics, Physics, and Mechanics, Ljubljana, Slovenia
[h] Faculty of Chemistry and Chemical Engineering, University of Maribor, Slovenia

ARTICLE INFO

Article history:
Received 6 January 2022
Revised 9 June 2022
Accepted 11 June 2022
Available online 4 July 2022

2020 MSC:
05C09
94A17
05C12
05C92
05C90

Keywords:
Szeged entropy
Mostar entropy
PI entropy
Cut method
Quotient graphs
Sensitivity of a topological descriptor

ABSTRACT

In this paper, we define some new distance-based graph measures and explore various properties. In particular, we develop a novel family of Szeged-like entropies and investigate several features thereof. We deduce a cut method for computing these entropies from quotient graphs, apply the obtained result to a molecular graph, and underpin the usefulness of the novel measures by numerical results.

1. Introduction

For a probability vector $p = (p_1, \ldots, p_n)$, where $\sum_{i=1}^{n} p_i = 1$ and $p_i \geq 0$, the *Shannon's entropy* [17] of p was defined as

$$I(p) = -\sum_{i=1}^{n} p_i \log p_i.$$

* Corresponding author at: Faculty of Natural Sciences and Mathematics, University of Maribor, Slovenia.
E-mail addresses: Matthias.Dehmer@umit.at (M. Dehmer), frank.emmert-streib@tuni.fi (F. Emmert-Streib), niko.tratnik@um.si (N. Tratnik).

https://doi.org/10.1016/j.amc.2022.127325

In the above definition, log denotes the logarithm of the base 2. Moreover, we assume that $0 \cdot \log 0 = 0$, which coincides with $\lim_{x \downarrow 0} x \log x = 0$.

Various entropy measures on graphs have been defined based on the partition of elements of a graph obtained from some equivalence relation. The pioneers in this field of research were Rashevsky, Trucco, and Mowshowitz. For more information and references see a survey paper on graph entropies [7].

Then, Dehmer [4] used a different approach which is not based on the problem of determining partitions of elements. His definition uses so-called information functionals which map graph elements (for example vertices) to the non-negative real numbers. More precisely, if $f : V(G) \to \mathbb{R}_0^+$ such that $\sum_{x \in V(G)} f(x) > 0$, then the *entropy* I_f of graph G is defined by

$$I_f(G) = -\sum_{x \in V(G)} p(x) \log p(x),$$

where for any $x \in V(G)$ we have

$$p(x) = \frac{f(x)}{\sum_{x \in V(G)} f(x)}.$$

In particular, there exist entropies related to degree-based topological indices (for example, see [9]). Various applications of these measures in computer science, chemistry, bioinformatics, complex networks and related disciplines are known (see [5–7,16] and references therein).

On the other hand, an entropy of a graph can be defined also by using edge contributions, i.e. information functionals that map the graph edges to the non-negative real numbers [2,12,15]. For example, an information functional can be a function used in the bond-additive topological index (TI). However, since Szeged-like TI are also well-known bond-additive topological descriptors, we focus in this present paper on entropies that can be defined by information functionals based on Szeged-like TI.

Consequently, we next introduce the most important Szeged-like TI. The well-known *Szeged index* [10] of a connected graph G is defined as

$$Sz(G) = \sum_{f=xy \in E(G)} n_x(f) n_y(f),$$

where $n_x(f)$ is the number of vertices in the graph G that are closer to x than to y and $n_y(f)$ is defined analogously, for any edge $f = xy \in E(G)$. Similarly, the *PI index* [13] of G is

$$PI(G) = \sum_{f=xy \in E(G)} (m_x(f) + m_y(f)),$$

where for an edge $f = xy \in E(G)$, $m_x(f)$ represents the number of edges closer to x than to y and $m_y(f)$ is defined analogously. Recently, another bond-additive topological index, named as the Mostar index, has been introduced [8]. For a connected graph G the *Mostar index* of G is defined as

$$Mo(G) = \sum_{f=xy \in E(G)} |n_x(f) - n_y(f)|.$$

However, there exist numerous variations of these topological indices, which are known as Szeged-like TI [3].

The aim of the present paper is to introduce the general form of a Szeged-like entropy of a graph. Next, we develop a cut method for the calculation of such graph measures based on the weighted quotient graphs (see [3] for more information on the cut method). As an example, we apply the obtained method to a molecular graph called phenylene. Finally, we statistically test selected Szeged-like entropies and discuss the obtained numerical results.

2. Preliminaries

If G is a graph, we denote the vertex set by $V(G)$ and the edge set by $E(G)$. Further, let the distance $d_G(x, y)$ be the length of any shortest path from x to y, where $x, y \in V(G)$. Let $x \in V(G)$ be a vertex and $e = uv \in E(G)$ be an edge of a graph G. Then, the distance from x to e equals

$$d_G(x, e) = \min\{d_G(x, u), d_G(x, v)\}.$$

If $x \in V(G)$, then $N(x) = \{y \in V(G) \mid xy \in E(G)\}$ is the *open neighbourhood* of x and $\deg(x) = |N(x)|$ is the *degree* of x.

Let $f_1 = x_1y_1$ and $f_2 = x_2y_2$ be two edges of a connected graph G. Then f_1 and f_2 are in relation Θ, denoted as $f_1 \Theta f_2$, if

$$d(x_1, x_2) + d(y_1, y_2) \neq d(x_1, y_2) + d(y_1, x_2).$$

This relation is the well-known Djoković-Winkler relation. In general, relation Θ is not an equivalence relation due to the transitivity. Therefore, we denote its transitive closure by Θ^*. For more information on relation Θ see [11].

For the edge set $E(G)$, we denote by $\mathcal{M} = \{M_1, \dots, M_r\}$ its Θ^*-partition. If $\mathcal{C} = \{C_1, \dots, C_k\}$ is a partition of $E(G)$ such that for any $i \in \{1, \dots, k\}$ the set C_i is the union of some sets from $\mathcal{M}$, we say that $\mathcal{C}$ is coarser than $\mathcal{M}$ and refer to $\mathcal{C}$ as a *c-partition*.

Let M be a subset of $E(G)$. The connected components of $G - M$ represent the vertices of the *quotient graph* G/M. We insert an edge between two vertices U_1 and U_2 of G/M if there exist $u_1 \in V(U_1)$ and $u_2 \in V(U_2)$ such that $u_1u_2 \in E(G)$. For any edge $F = U_1U_2$ in the quotient graph G/M, we define the set $\widehat{F} = \{u_1u_2 \in E(G) \mid u_1 \in V(U_1), u_2 \in V(U_2)\}$.

3. Szeged-like TI of strength-weighted graphs

In this section, we describe an efficient cut method, deduced in [3], which enables the computation of Szeged-like TI for any strength-weighted graph.

For a graph G, the *strength-weighted graph* G_{sw} is here defined as $G_{sw} = (G, SW_V, SW_E)$, where SW_V is a pair (w_v, s_v) of weights on vertices and SW_E is a pair (w_e, s_e) of weights on edges; i.e. w_v, s_v are functions from $V(G_{sw})$ to $\mathbb{R}$ and w_e, s_e are functions from $E(G_{sw})$ to $\mathbb{R}$ (see [1]).

Note that in [3] it was required that the weights w_v, s_v, w_e, s_e are non-negative. However, for some technical reasons (needed in Section 4) we allow that the weight of a vertex or an edge is any real number. It is easy to see that Theorem 3.1 remains valid also under such conditions.

Let us define the following sets for any $f = xy \in E(G)$:

$$\begin{aligned}
N_x(f|G) &= \{u \in V(G) \mid d_G(x,u) < d_G(y,u)\},\\
N_y(f|G) &= \{u \in V(G) \mid d_G(y,u) < d_G(x,u)\},\\
N_0(f|G) &= \{u \in V(G) \mid d_G(x,u) = d_G(y,u)\},\\
M_x(f|G) &= \{e \in E(G) \mid d_G(x,e) < d_G(y,e)\},\\
M_y(f|G) &= \{e \in E(G) \mid d_G(y,e) < d_G(x,e)\},\\
M_0(f|G) &= \{e \in E(G) \mid d_G(x,e) = d_G(y,e)\}.
\end{aligned}$$

Now we can define the following numbers for $f = xy$ of G_{sw}:

$$\begin{aligned}
n_x(f|G_{sw}) &= \sum_{u \in N_x(f|G_{sw})} w_v(u),\\
m_x(f|G_{sw}) &= \sum_{u \in N_x(f|G_{sw})} s_v(u) + \sum_{e \in M_x(f|G_{sw})} s_e(e),\\
n_0(f|G_{sw}) &= \sum_{u \in N_0(f|G_{sw})} w_v(u),\\
m_0(f|G_{sw}) &= \sum_{u \in N_0(f|G_{sw})} s_v(u) + \sum_{e \in M_0(f|G_{sw})} s_e(e).
\end{aligned}$$

Analogously, we set $n_y(f|G_{sw})$ and $m_y(f|G_{sw})$.

In addition, for $f = xy$ of G_{sw} we define the next value:

$$R(f|G_{sw}) = R\big(n_x(f|G_{sw}), n_y(f|G_{sw}), m_x(f|G_{sw}), m_y(f|G_{sw}), n_0(f|G_{sw}), m_0(f|G_{sw})\big),$$

where R is a function of six variables such that for all $(r_1, r_2, r_3, r_4, r_5, r_6)$ it holds

$$R(r_1, r_2, r_3, r_4, r_5, r_6) = R(r_2, r_1, r_4, r_3, r_5, r_6).$$

The last condition implies that R is symmetric in the first two pairs of coordinates. We say that such a function R is a *regular function* of G_{sw}. Then, for G_{sw}, the *Szeged-like TI* is defined as (see [3])

$$TI_R(G_{sw}) = \sum_{f \in E(G)} w_e(f) R(f|G_{sw}). \tag{1}$$

It is easy to see that all the mentioned topological indices in Section 1 belong to the family of Szeged-like TI.

Let $\{C_1, \dots, C_k\}$ be any c-partition of $E(G_{sw})$. For $j \in \{1, \dots, k\}$, we define the quotient graph $G_{sw}/C_j = (G/C_j, SW_v^j, SW_e^j)$ as the strength-weighted graph with the following weights (see [1]):

- $w_v^j : V(G_{sw}/C_j) \to \mathbb{R}$, $w_v^j(U) = \sum_{u \in V(U)} w_v(u)$, $\forall\ U \in V(G_{sw}/C_j)$,
- $s_v^j : V(G_{sw}/C_j) \to \mathbb{R}$, $s_v^j(U) = \sum_{f \in E(U)} s_e(f) + \sum_{u \in V(U)} s_v(u)$, $\forall\ U \in V(G_{sw}/C_j)$,
- $w_e^j : E(G_{sw}/C_j) \to \mathbb{R}$, $w_e^j(F) = \sum_{f \in \widehat{F}} w_e(f)$, $\forall\ F \in E(G_{sw}/C_j)$,
- $s_e^j : E(G_{sw}/C_j) \to \mathbb{R}$, $s_e^j(F) = \sum_{f \in \widehat{F}} s_e(f)$, $\forall\ F \in E(G_{sw}/C_j)$.

For any $j \in \{1, \ldots, k\}$, we usually write G_j instead of G_{sw}/C_j. The following theorem represents the main result of [3].

Theorem 3.1. *[3] If $\{C_1, \ldots, C_k\}$ is a c-partition and R a regular function of a connected G_{sw}, then*

$$TI_R(G_{sw}) = \sum_{j=1}^{k} TI_R(G_j).$$

4. Szeged-like entropies of strength-weighted graphs and the cut method

As already mentioned, an entropy of a graph can be defined by using an information functional that maps vertices to the non-negative real numbers. Similarly, if G is a connected graph and $h : E(G) \to \mathbb{R}_0^+$ with $\sum_{f \in E(G)} h(f) > 0$ is an information functional that maps the edges of G to the non-negative real numbers, then the *edge-entropy* I_h of graph G is defined as

$$I_h(G) = -\sum_{f \in E(G)} p(f) \log p(f), \tag{2}$$

where for any edge $f \in E(G)$ we have

$$p(f) = \frac{h(f)}{\sum_{f \in E(G)} h(f)}. \tag{3}$$

As special cases, we can obtain entropies defined by $h(f) = n_x(f)n_y(f)$, $h(f) = m_x(f) + m_y(f)$, and $h(f) = |n_x(f) - n_y(f)|$, which are called the *Szeged entropy*, the *PI entropy*, and the *Mostar entropy*, respectively [15]. Note that the Mostar entropy is defined only for graphs in which $n_x(f) \neq n_y(f)$ for at least one edge $f \in E(G)$ (i.e. for graphs with $Mo(G) \neq 0$).

In this section, we firstly formally define Szeged-like entropies by using Eq. (2) and then develop methods for calculating such entropies.

4.1. Szeged-like entropies

Let G_{sw} be a connected strength-weighted graph and R a regular function of G_{sw} such that $w_e(f) \geq 0$ and $R(f|G_{sw}) \geq 0$ for any $f \in E(G_{sw})$, but for at least one edge $f \in E(G_{sw})$ it holds $w_e(f)R(f|G_{sw}) > 0$. We now consider the edge-entropy of G_{sw} based on the information functional $h : E(G) \to \mathbb{R}_0^+$ defined as $h(f) = w_e(f)R(f|G_{sw})$ for every $f \in E(G_{sw})$. Such entropy $I_h(G_{sw})$ will be called a *Szeged-like entropy* of G_{sw}.

Obviously, by Eqs. (1) and (3) we obtain

$$p(f) = \frac{w_e(f)R(f|G_{sw})}{TI_R(G_{sw})} \tag{4}$$

for any $f \in E(G_{sw})$.

If we define $w_v(u) = 1$, $s_v(u) = 0$ for any $u \in V(G_{sw})$, $s_e(f) = 1$ for any $f \in E(G_{sw})$ together with the weight $w_e(f)$ and the regular function R as in Table 1, we obtain the corresponding selected Szeged-like entropies.

Of course there exist infinitely many other Szeged-like entropies, which are obtained by using different regular function or weight w_e. Several other examples of Szeged-like TI can be found in [3], for which the corresponding entropies can be defined.

Next, let R' be a function of six variables defined by

$$R' = R \log R. \tag{5}$$

Obviously, R' is also a regular function for G_{sw}.

Furthermore, we define the strength-weighted graph $G'_{sw} = (G, SW_V, (w'_e, s_e))$ to be the strength-weighted graph where for every $f \in E(G'_{sw}) = E(G_{sw})$ we have

$$w'_e(f) = w_e(f) \log(w_e(f)). \tag{6}$$

Obviously, strength-weighted graphs G_{sw} and G'_{sw} differ only in the weight w_e. The following proposition can now be stated.

Table 1
Some Szeged-like entropies, corresponding functions, and weights w_e.

Szeged-like entropy	regular function R	$w_e(f) = w_e(xy)$
Szeged entropy (I_{Sz})	$r_1 r_2$	1
PI entropy (I_{PI})	$r_3 + r_4$	1
Mostar entropy (I_{Mo})	$\lvert r_1 - r_2 \rvert$	1
revised edge-Szeged entropy ($I_{Sz_e^*}$)	$(r_3 + r_6/2)(r_4 + r_6/2)$	1
weighted-plus Szeged entropy (I_{w^+Sz})	$r_1 r_2$	$\deg(x) + \deg(y)$

Proposition 4.1. *If G_{sw} is a connected strength-weighted graph and R a regular function for G_{sw} such that $w_e(f) \geq 0$ and $R(f|G_{sw}) \geq 0$ for any edge $f \in E(G_{sw})$, but for at least one edge $f \in E(G_{sw})$ it holds $w_e(f)R(f|G_{sw}) > 0$, then the Szeged-like entropy of G_{sw} can be computed as*

$$I_h(G_{sw}) = \log(TI_R(G_{sw})) - \frac{TI_{R'}(G_{sw}) + TI_R(G'_{sw})}{TI_R(G_{sw})}.$$

Proof. By Eq. (4) we get

$$\begin{aligned} I_h(G_{sw}) &= -\sum_{f\in E(G)} \frac{w_e(f)R(f|G_{sw})}{TI_R(G_{sw})} \log\left(\frac{w_e(f)R(f|G_{sw})}{TI_R(G_{sw})}\right) \\ &= \frac{-1}{TI_R(G_{sw})} \sum_{f\in E(G)} w_e(f)R(f|G_{sw})(\log(w_e(f)R(f|G_{sw})) - \log(TI_R(G_{sw}))) \\ &= \frac{-1}{TI_R(G_{sw})} \sum_{f\in E(G)} w_e(f)\log(w_e(f))R(f|G_{sw}) + \frac{-1}{TI_R(G_{sw})} \sum_{f\in E(G)} w_e(f)R(f|G_{sw})\log(R(f|G_{sw})) \\ &\quad + \frac{\log(TI_R(G_{sw}))}{TI_R(G_{sw})} \sum_{f\in E(G)} w_e(f)R(f|G_{sw}). \end{aligned}$$

By using Eqs. (6), (5), and (1) we finally arrive to

$$\begin{aligned} I_h(G_{sw}) &= \frac{-1}{TI_R(G_{sw})}\left(\sum_{f\in E(G)} w'_e(f)R(f|G_{sw}) + \sum_{f\in E(G)} w_e(f)R'(f|G_{sw})\right) + \frac{\log(TI_R(G_{sw}))}{TI_R(G_{sw})} TI_R(G_{sw}) \\ &= \log(TI_R(G_{sw})) - \frac{TI_{R'}(G_{sw}) + TI_R(G'_{sw})}{TI_R(G_{sw})}, \end{aligned}$$

which completes the proof. □

We can easily see from Eq. (6) that if $w_e \equiv 1$ (which is used to compute many well known topological indices), then $w'_e \equiv 0$ and also $TI_R(G'_{sw}) = 0$. As a consequence, we obtain the next corollary.

Corollary 4.2. *Let G_{sw} be a connected strength-weighted graph with a regular function R such that $w_e \equiv 1$. If $R(f|G_{sw}) \geq 0$ for any $f \in E(G_{sw})$, but for at least one edge $f \in E(G_{sw})$ it holds $R(f|G_{sw}) > 0$, then the Szeged-like entropy of G_{sw} can be computed as*

$$I_h(G_{sw}) = \log(TI_R(G_{sw})) - \frac{TI_{R'}(G_{sw})}{TI_R(G_{sw})}.$$

4.2. Cut method

Here we present a computational procedure for the Szeged-like entropies of graphs by using quotient graphs. Let all the notation be the same as in the previous subsection. Moreover, suppose that $\{C_1, \ldots, C_k\}$ is a c-partition of $E(G_{sw}) = E(G'_{sw})$. In addition, for arbitrary $j \in \{1, \ldots, k\}$, let the quotient graph G_{sw}/C_j be denoted as G_j and the quotient graph G'_{sw}/C_j as G'_j (see Section 3 for the definition of these quotient graphs). By Proposition 4.1 and Theorem 3.1 we arrive to the main result of the paper, which enables us to calculate Szeged-like entropies from quotient graphs.

Theorem 4.3. *Let G_{sw} be a connected strength-weighted graph and R its regular function such that $w_e(f) \geq 0$ and $R(f|G_{sw}) \geq 0$ for any $f \in E(G_{sw})$, but for at least one edge $f \in E(G_{sw})$ it holds $w_e(f)R(f|G_{sw}) > 0$. For a c-partition $\{C_1, \ldots, C_k\}$ of $E(G_{sw})$, the Szeged-like entropy of G_{sw} can be computed as*

$$I_h(G_{sw}) = \log\left(\sum_{j=1}^{k} TI_R(G_j)\right) - \frac{\sum_{j=1}^{k}\left(TI_{R'}(G_j) + TI_R(G'_j)\right)}{\sum_{j=1}^{k} TI_R(G_j)}.$$

5. Application to a molecular graph

In this section, we compute the Szeged entropy I_{Sz} and the weighted-plus Szeged entropy I_{w^+Sz} of a molecular graph by using the cut method described in the previous section.

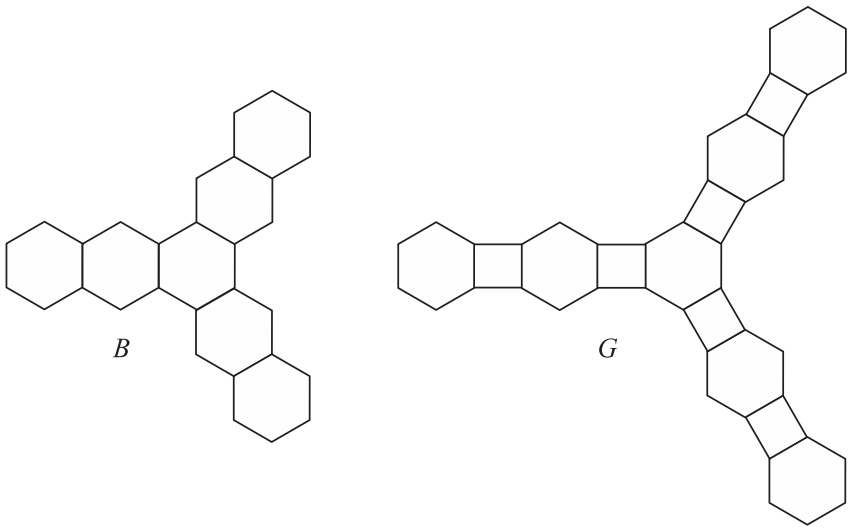


Fig. 1. A benzenoid system B and the corresponding phenylene G.

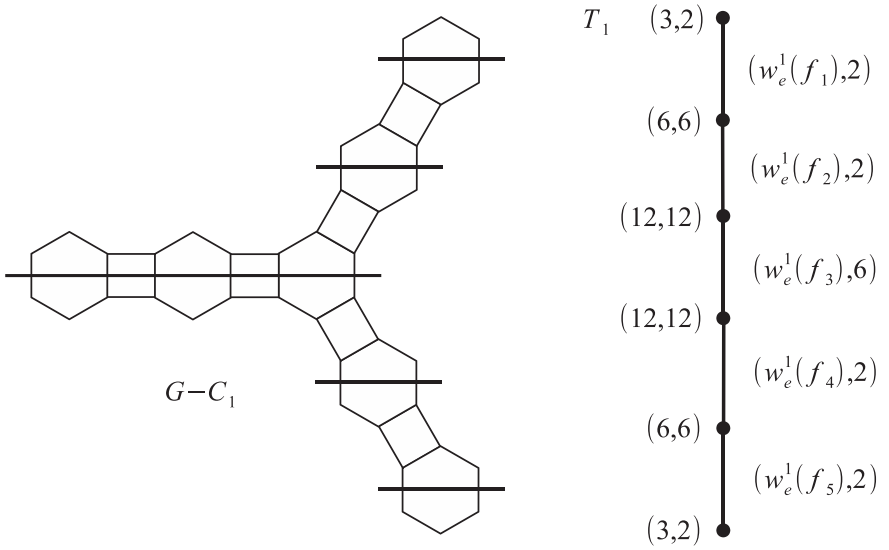


Fig. 2. The edges in the set C_1 (crossed by line segments) and the strength-weighted quotient tree T_1.

A *benzenoid system* is a subgraph of a regular hexagonal lattice determined by a cycle. If every vertex of a benzenoid system lies on this cycle, we call it a *catacondensed benzenoid system*, see graph B in Fig. 1. Every catacondensed benzenoid system can be associated with a molecular graph called *phenylene*, see graph G in Fig. 1. Note that some Szeged-like TI of such graphs were calculated in [3].

Obviously, we can partition the edge set of a (catacondensed) benzenoid system into three subsets regarding their direction. These three subsets correspond to the three subsets of edges C_1, C_2, C_3 of a phenylene. One of these sets, C_1, can be seen in the left part of Fig. 2.

Moreover, we define $C_4 = E(G) \setminus (C_1 \cup C_2 \cup C_3)$, see the left part of Fig. 3. It turns out that $\{C_1, C_2, C_3, C_4\}$ is a c-partition of the set $E(G)$ and that the corresponding quotient graphs are trees, for more details see [3]. For a phenylene G_{sw}, we denote these quotient trees as $T_j = G_{sw}/C_j$, $j \in \{1, 2, 3, 4\}$. Because of the symmetry, it is easy to see that for a graph G in Fig. 1, the quotient trees T_1, T_2, T_3 are isomorphic. Strength-weighted quotient trees T_1 and T_4 are depicted on right parts of Figs. 2 and 3.

Firstly, we compute the Szeged entropy, for which $R(f|G_{sw}) = n_x(f|G_{sw})n_y(f|G_{sw})$ and $w_e(f) = 1$ for any $f = xy \in E(G)$. Therefore, $w_e^1(f_1) = w_e^1(f_2) = w_e^1(f_4) = w_e^1(f_5) = 2$ and $w_e^1(f_3) = 6$ (see Fig. 2). Moreover, $w_e^4(f) = 2$ for any edge f in the quotient tree T_4 (see Fig. 3).

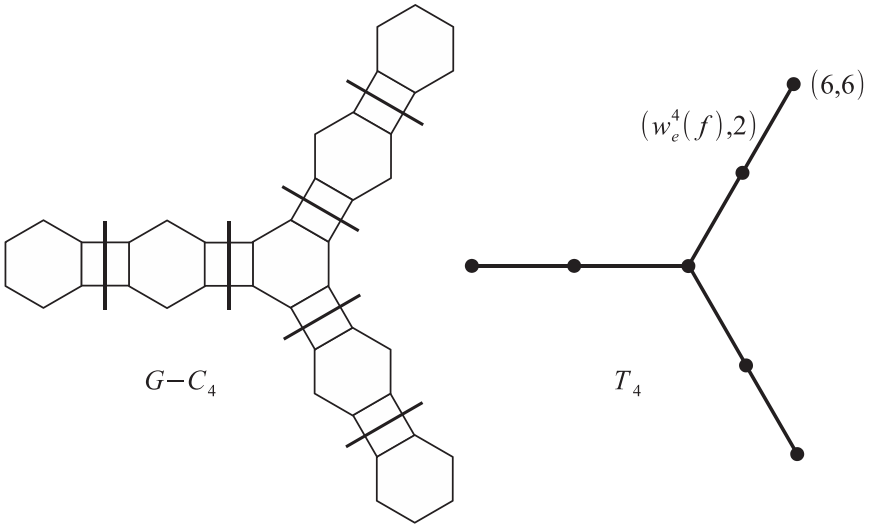


Fig. 3. The edges in the set C_4 (crossed by line segments) and the strength-weighted quotient tree T_4.

Next, we compute the corresponding topological indices of quotient trees:

$$TI_R(T_1) = 2\cdot 2\cdot 3\cdot 39 + 2\cdot 2\cdot 9\cdot 33 + 6\cdot 21\cdot 21 = 4302,$$
$$TI_{R'}(T_1) = 2\cdot 2\cdot 3\cdot 39\cdot\log(3\cdot 39) + 2\cdot 2\cdot 9\cdot 33\cdot\log(9\cdot 33) + 6\cdot 21\cdot 21\cdot\log(21\cdot 21) = 36218.086,$$

and

$$TI_R(T_4) = 3\cdot 2\cdot 6\cdot 36 + 3\cdot 2\cdot 12\cdot 30 = 3456,$$
$$TI_{R'}(T_4) = 3\cdot 2\cdot 6\cdot 36\cdot\log(6\cdot 36) + 3\cdot 2\cdot 12\cdot 30\cdot\log(12\cdot 30) = 28392.737.$$

Therefore, we get

$$TI_R(G) = 3\cdot TI_R(T_1) + TI_R(T_4) = 16362,$$
$$TI_{R'}(G) = 3\cdot TI_{R'}(T_1) + TI_{R'}(T_4) = 137046.994.$$

Finally, by Proposition 4.1 or Theorem 4.3 we obtain

$$I_{Sz}(G) = 5.622.$$

For the weighted-plus Szeged entropy we have $R(f|G_{sw}) = n_x(f|G_{sw})n_y(f|G_{sw})$ and $w_e(f) = \deg(x) + \deg(y)$ for any $f = xy \in E(G)$. Therefore, $w_e^1(f_1) = w_e^1(f_5) = 9$, $w_e^1(f_2) = w_e^1(f_4) = 10$, and $w_e^1(f_3) = 34$ (see Fig. 2). Moreover, $w_e^4(f) = 12$ for any edge f in the quotient tree T_4 (see Fig. 3). Similarly, $(w'_e)^1(f_1) = (w'_e)^1(f_5) = 4\cdot\log(4) + 5\cdot\log(5)$, $(w'_e)^1(f_2) = (w'_e)^1(f_4) = 2\cdot 5\cdot\log(5)$, and $(w'_e)^1(f_3) = 5\cdot 6\cdot\log(6) + 4\cdot\log(4)$. In addition, $(w'_e)^4(f) = 2\cdot 6\cdot\log(6)$ for any edge f in the tree T'_4.

The topological indices of quotient trees are:

$$TI_R(T_1) = 2\cdot 9\cdot 3\cdot 39 + 2\cdot 10\cdot 9\cdot 33 + 34\cdot 21\cdot 21 = 23040,$$
$$TI_{R'}(T_1) = 2\cdot 9\cdot 3\cdot 39\cdot\log(3\cdot 39) + 2\cdot 10\cdot 9\cdot 33\cdot\log(9\cdot 33) + 34\cdot 21\cdot 21\cdot\log(21\cdot 21) = 194978.859,$$
$$TI_R(T'_1) = 2\cdot(4\cdot\log(4) + 5\cdot\log(5))\cdot 3\cdot 39 + 2\cdot 2\cdot 5\cdot\log(5)\cdot 9\cdot 33 + (5\cdot 6\cdot\log(6) + 4\cdot\log(4))\cdot 21\cdot 21 = 56107.963,$$

and

$$TI_R(T_4) = 3\cdot 12\cdot 6\cdot 36 + 3\cdot 12\cdot 12\cdot 30 = 20736,$$
$$TI_{R'}(T_4) = 3\cdot 12\cdot 6\cdot 36\cdot\log(6\cdot 36) + 3\cdot 12\cdot 12\cdot 30\cdot\log(12\cdot 30) = 170356.421,$$
$$TI_R(T'_4) = 3\cdot 2\cdot 6\cdot\log(6)\cdot 6\cdot 36 + 3\cdot 2\cdot 6\cdot\log(6)\cdot 12\cdot 30 = 53601.782.$$

Therefore, we deduce

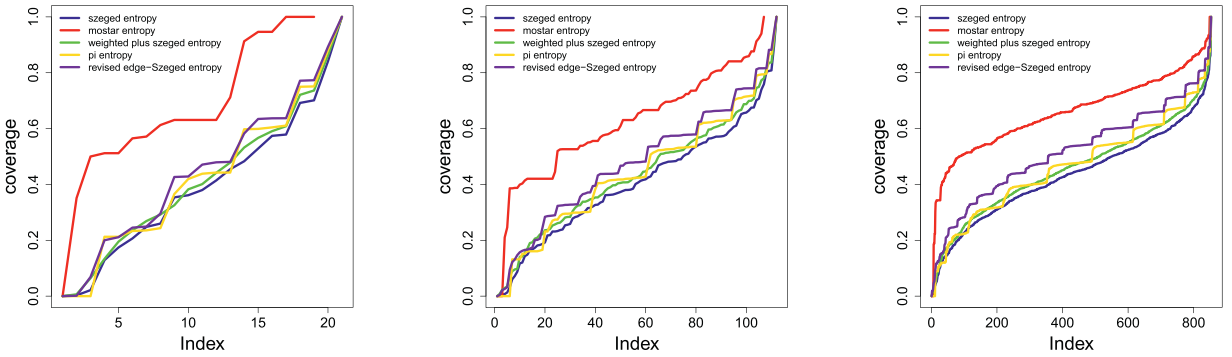
$$TI_R(G) = 3\cdot TI_R(T_1) + TI_R(T_4) = 89856,$$
$$TI_{R'}(G) = 3\cdot TI_{R'}(T_1) + TI_{R'}(T_4) = 755292.997,$$
$$TI_R(G') = 3\cdot TI_R(T'_1) + TI_R(T'_4) = 221925.670.$$

Finally, by Proposition 4.1 or Theorem 4.3 we compute

$$I_{w^+Sz}(G) = 5.580.$$

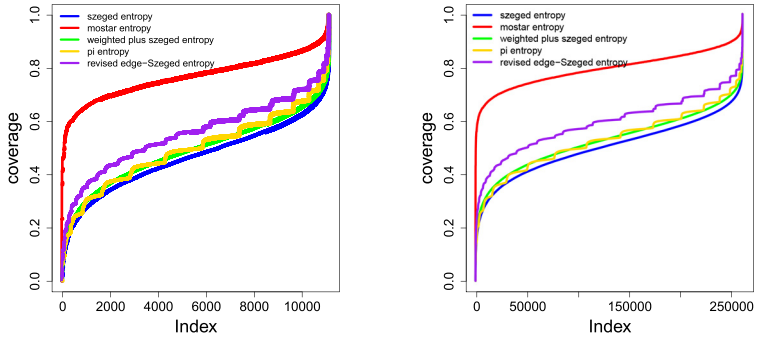
Table 2
ndv and sensitivity for selected Szeged-like entropies of different general graph classes and tree classes.

	Szeged entropy		Mostar entropy		weig.-plus Sz. ent.		PI entropy		rev. edge-Sz. ent.	
Graph Class	ndv	sens.	ndv	sens.	ndv	sens.	ndv	sens.	ndv	sens.
					All connected graphs					
N_5	0	1.0000	11	0.4762	0	1.0000	7	0.6667	0	1.0000
N_6	2	0.9821	60	0.4643	0	1.0000	40	0.6429	4	0.9643
N_7	43	0.9496	546	0.3599	0	1.0000	345	0.5955	40	0.9531
N_8	615	0.9447	9625	0.1342	47	0.9958	4833	0.5653	1055	0.9051
N_9	36123	0.8616	250306	0.0413	6327	0.9758	120223	0.5395	33706	0.8709
					All tree structures					
T_{12}	445	0.1924	445	0.1924	12	0.9782	551	0.0000	445	0.1924
T_{13}	1154	0.1130	1154	0.1130	74	0.9431	1301	0.0000	1154	0.1130
T_{14}	2884	0.0871	2898	0.0826	217	0.9313	3159	0.0000	2884	0.0871
T_{15}	7425	0.0408	7425	0.0408	868	0.8879	7741	0.0000	7425	0.0408
T_{16}	18650	0.0347	18667	0.0338	2469	0.8722	19320	0.0000	18650	0.0347
T_{17}	47824	0.0166	47824	0.0166	8600	0.8232	48629	0.0000	47824	0.0166
T_{18}	122328	0.0124	122419	0.0117	24357	0.8034	123867	0.0000	122328	0.0124



(a) N_5. (b) N_6. (c) N_7.

Fig. 4. Coverage of different distance measures for general graph classes.



(a) N_8. (b) N_9.

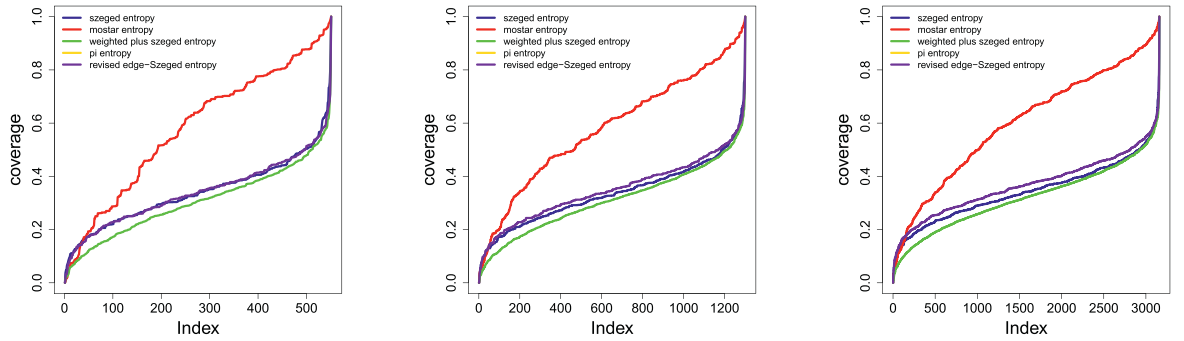
Fig. 5. Coverage of different distance measures for general graph classes.

6. Numerical results

In this section, we present numerical results regarding the discrimination power for different types of Szeged-like entropies. We consider the following measures: the Szeged entropy, the Mostar entropy, the weighted-plus Szeged entropy, the PI entropy, and the revised edge-Szeged entropy.

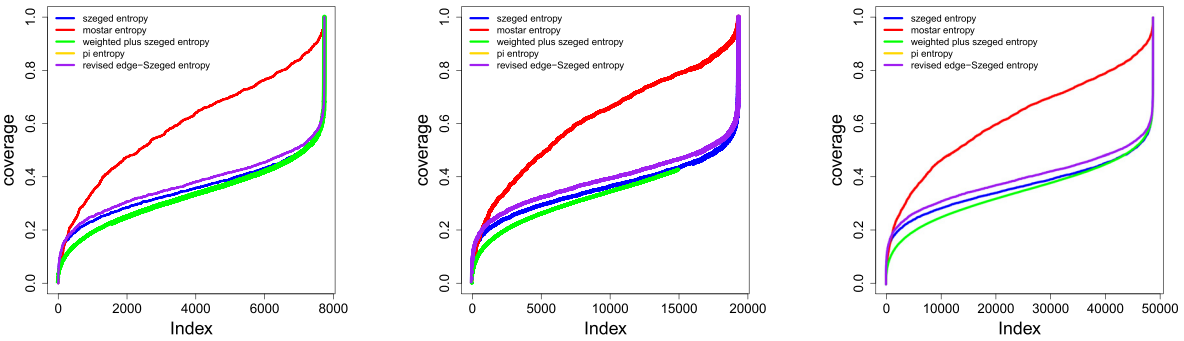
In particular, for the mentioned entropies we investigate the sensitivity defined by Konstantinova in 1996 [14]. For a finite class of graphs $\mathcal{G}$, the *sensitivity* S of a measure I is defined as

$$S = \frac{|\mathcal{G}| - \text{ndv}}{|\mathcal{G}|},$$



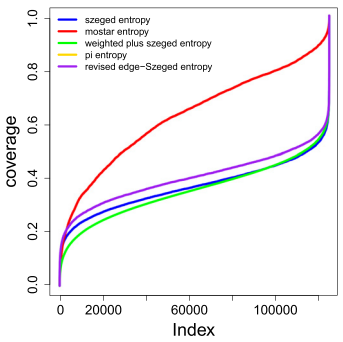
(a) T_{12}. (b) T_{13}. (c) T_{14}.

Fig. 6. Coverage of different distance measures for tree classes.



(a) T_{15}. (b) T_{16}. (c) T_{17}.

Fig. 7. Coverage of different distance measures for tree classes.



(a) T_{18}.

Fig. 8. Coverage of different distance measures for tree class T_{18}.

where ndv is the number of non-distinguishable values, i.e. the number of non-isomorphic graphs in $\mathcal{G}$ which cannot be distinguished by I. More precisely, ndv is the cardinality of the set of all graphs $G \in \mathcal{G}$ for which there exist a graph $H \in \mathcal{G}$ such that $I(G) = I(H)$, but G and H are not isomorphic.

Let N_k be the class of all connected graphs with k vertices and T_k the set of all trees on k vertices. Table 2 shows the number ndv and the sensitivity for graph classes N_i, $i \in \{5, 6, 7, 8, 9\}$, and T_i, $i \in \{12, 13, 14, 15, 16, 17, 18\}$. Note that the graphs for which the Mostar index equals 0 (which means that the Mostar entropy is not defined) are counted as non-distinguishable values.

We notice from the obtained results that the Szeged entropy and the revised edge-Szeged entropy can distinguish more than 86% of graphs in the class N_9, but less than 2% of trees in T_{18}. Moreover, the PI entropy can distinguish almost 54% of graphs in N_9, but can not distinguish trees with the same number of vertices. Also, the Mostar entropy behaves badly in the sense of uniqueness, since it can distinguish just around 4% of graphs in N_9 and around 1% of graphs in T_{18}. But on the

other hand, the weighted-plus Szeged entropy can discriminate more than 97% of graphs in N_9 and around 80% of graphs in T_{18}. Therefore, we can conclude that among the considered distance measures, the most appropriate one to distinguish different graphs in various classes is the weighted-plus Szeged entropy.

To conclude the section, we present the value-domain coverage for the considered Szeged-like graph measures, see Figs. 4, 5, 6, 7, 8. These figures show the normalized values of the corresponding entropies for all graphs from the considered classes.

7. Conclusion

In this paper, we dealt with several distance-based graph measures such as the Mostar entropy and other Szeged-like graph measures. In particular, we defined the general concept of a Szeged-like entropy based on the information functional approach due to Dehmer [4]. We also applied the cut method [3] for Szeged-like TI of strength-weighted graphs.

As a result, we see that the new measures namely the Szeged entropy, the Mostar entropy, the weighted-plus Szeged entropy, the PI entropy, and the revised edge-Szeged entropy follow the same principle described in [4]. As discussed, they possess meaningful properties such as a high discrimination power and a very good value-domain coverage. Given the fact that these measures are distance-based, their discrimination power is high compared to other standard distance-based indices, see [6]. Also, their value-domain coverage is excellent, that means, those measures cover the entire value domain. If not, one could argue that they can be only used for very special cases and from a theoretical point of view, an insufficient value-domain coverage is simply undesired.

As future work, the Szeged entropy, the Mostar entropy, the weighted-plus Szeged entropy, the PI entropy, and the revised edge-Szeged entropy could be investigated more deeply. For instance, interrelations based on inequalities can be proved. Also a study on a large scale should be performed, i.e., the mentioned measures should be compared with many other known measures in terms of correlation ability and exploring other mathematical properties. However, as already mentioned, among all investigated entropies, the weighted-plus Szeged entropy has the best discrimination power in all considered graph classes. Therefore, the emphasis of the future work should be related to this entropy.

Data Availability

All the computational results are included in the paper.

Acknowledgement

N. Tratnik and P. Žigert Pleteršek acknowledge the financial support from the Slovenian Research Agency (programme no. P1-0297). The research was also a part of a bilateral project between Slovenia and Austria co-founded by the Slovenian Research Agency (Grant No. BI-AT/20-21-001). M. Dehmer thanks the Austrian Agency OeAD of the just mentioned bilateral project (Project No. BI-AT/20-21-001) for financial support.

References

[1] M. Arockiaraj, J. Clement, K. Balasubramanian, Topological indices and their applications to circumcised donut benzenoid systems, Kekulenes and drugs, Polycycl. Aromat. Comp. 40 (2020) 280–303.
[2] F. Aziz, E.R. Hancock, R.C. Wilson, Network entropy using edge-based information functionals, J. Complex Netw. 8 (2020). Cnaa015.
[3] S. Brezovnik, N. Tratnik, General cut method for computing Szeged-like topological indices with applications to molecular graphs, Int. J. Quantum Chem. 121 (2021) E26530.
[4] M. Dehmer, Information processing in complex networks: graph entropy and information functionals, Appl. Math. Comput. 201 (2008) 82–94.
[5] M. Dehmer, F. Emmert-Streib, Quantitative Graph Theory: Mathematical Foundations and Applications, CRC Press, Taylor & Francis Group, Boca Raton, 2015.
[6] M. Dehmer, M. Grabner, K. Varmuza, Information indices with high discriminative power for graphs, PLoS ONE 7 (2012) E31214.
[7] M. Dehmer, A. Mowshowitz, A history of graph entropy measures, Inform. Sci. 181 (2011) 57–78.
[8] T. Došlić, I. Martinjak, R. Škrekovski, S.T. Spužević, I. Zubac, Mostar index, J. Math. Chem. 56 (2018) 2995–3013.
[9] A. Ghalavand, M. Eliasi, A.R. Ashrafi, First degree-based entropy of graphs, J. Appl. Math. Comput. 59 (2019) 37–46.
[10] I. Gutman, A formula for the wiener number of trees and its extension to graphs containing cycles, Graph Theory Notes N. Y. 27 (1994) 9–15.
[11] R. Hammack, W. Imrich, S. Klavžar, Handbook of Product Graphs, second ed., CRC Press, Taylor & Francis Group, Boca Raton, 2011.
[12] R. Kazemi, Entropy of weighted graphs with the degree-based topological indices as weights, MATCH Commun. Math. Comput. Chem. 76 (2016) 69–80.
[13] P.V. Khadikar, On a novel structural descriptor PI, Nat. Acad. Sci. Lett. 23 (2000) 113–118.
[14] E.V. Konstantinova, The discrimination ability of some topological and information distance indices for graphs of unbranched hexagonal systems, J. Chem. Inf. Comput. Sci. 36 (1996) 54–57.
[15] S. Manzoor, M.K. Siddiqui, S. Ahmad, On computation of entropy measures for phthalocyanines and porphyrins dendrimers, Int. J. Quantum Chem. (2021) E26854.
[16] Y.M. Omar, P. Plapper, A survey of information entropy metrics for complex networks, Entropy 22 (2020) 1417.
[17] C.E. Shannon, W. Weaver, The Mathematical Theory of Communication, University of Illinois Press, 1949.