



Structural Differentiation of Graphs Using Hosoya-Based Indices

Matthias Dehmer^{1,2*}, Abbe Mowshowitz³, Yongtang Shi^{4,5}

1 Department of Computer Science, Universität der Bundeswehr München, Neubiberg, Germany, **2** Institute for Bioinformatics and Translational Research, UMIT, Hall in Tirol, Austria, **3** Department of Computer Science, The City College of New York (CUNY), New York, New York, United States of America, **4** Center for Combinatorics and LPMC-TJKLC, Nankai University, Tianjin, China, **5** College of Computer and Control Engineering, Nankai University, Tianjin, China

Abstract

In this paper, we introduce the Hosoya-Spectral indices and the Hosoya information content of a graph. The first measure combines structural information captured by partial Hosoya polynomials and graph spectra. The latter is a graph entropy measure which is based on blocks consisting of vertices with the same partial Hosoya polynomial. We evaluate the discrimination power of these quantities by interpreting numerical results.

Citation: Dehmer M, Mowshowitz A, Shi Y (2014) Structural Differentiation of Graphs Using Hosoya-Based Indices. PLoS ONE 9(7): e102459. doi:10.1371/journal.pone.0102459

Editor: Matjaz Perc, University of Maribor, Slovenia

Received: May 9, 2014; **Accepted:** June 19, 2014; **Published:** July 14, 2014

Copyright: © 2014 Dehmer et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that, for approved reasons, some access restrictions apply to the data underlying the findings. The data sets (graphs) can be generated by using the papackage Nauty (subpackage geng) by McKay.

Funding: Matthias Dehmer thanks the Austrian Science Funds for supporting this work (project P26142). Matthias Dehmer gratefully acknowledges financial support from the German Federal Ministry of Education and Research (BMBF) (project RiKoV, Grant No. 13N12304). Yongtang Shi are supported by NSFC, PCSIRT, China Postdoctoral Science Foundation and China Scholarship Council. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Matthias Dehmer is a PLOS ONE Editorial Board member. This does not alter the authors' adherence to PLOS ONE Editorial policies and criteria.

* Email: matthias.dehmer@umit.at

Introduction

Structural differentiation entails the classification of graphs according to structural features captured by quantitative measures, see, e.g., [1–5]. One way to demonstrate a classification procedure is to apply a measure (or index) to a special class of graphs and show that the measure discriminates between non-isomorphic graphs with high probability. A prominent example is the Balaban J index [6–8] which is highly discriminating on chemical graphs. However, this index has limitations as shown by Dehmer et al. [2] by means of a statistical analysis of the performance of the J and other indices on an exhaustively generated set of graphs without structural constraints, see [2]. This analysis shows that the discrimination power (also called uniqueness [2]) of graph measures depends on the underlying class of graphs [2].

This paper is an investigation of the discriminating power of structural indices based on the zeros of partial Hosoya polynomials and graph spectra. Also, we introduce and evaluate the *Hosoya information content* of a graph. To position this investigation we begin with a survey of literature dealing with eigenvalues and entropy-based measures of graphs. Classical results in the theory of graph spectra are due to Cvetković et al. [9]. The main concern of this theory is to explore structural properties of graphs and complex networks captured by graph spectra [10]. More recent results have been presented and surveyed by Chung [11] and Cioabă [12]. Interdisciplinary applications of graph spectra, e.g., the analysis of biological networks and web graphs can be found in [10,13]. Various graph measures incorporating eigenvalues have been discussed by Randić et al. [14] and Dehmer et al. [15]. One example of a measure is defined as the sum of the moduli of non-

zero eigenvalues of the adjacency matrix of a graph; another is given by graph entropies based on the eigenvalues of matrices associated with a graph [14–16]. Yet another well-known measure is the Estrada index [17–20] which has been explored in bioinformatics, mathematical chemistry and applied mathematics. A more recent review of this quantity is due to Gutman et al. [21]. Variants of this measures using other matrices have been discussed by Li et al. [22]. A related measure is the so-called energy of a graph is an important quantity defined in relation to the eigenvalues of matrices associated with a graph, see [23–25]. Extremal properties of graph energy have been studied by [23–25]. A recent book on graph energy summarizing classical and new results is [26]. Inequalities for eigenvalue-based graph measures have been discussed in [12]. Elphick and Wocjan [27] analyzed a novel spectral measure for determining network irregularity [27].

Graph entropy measures have been explored extensively in various disciplines. Rashevsky and Mowshowitz did seminal work when developing the first graph entropy measures based on vertex orbits [28,29]. Körner introduced a graph entropy measure that has been used in information theory [30]. Bonchev et al. developed the magnitude-based information indices and various others based on graph invariants such as vertex degrees and distances in graphs [31–34]. Also, Bonchev et al. [1] proposed an information index for graphs which is based on the Hosoya graph decomposition. However, this information index (using Hosoya index Z [35] to define the probabilities of the induced partition) is quite different from the one we introduce here in section ‘Hosoya-based Indices’. Many other graph entropy measures can be found in [36–38]. To study results towards the Hosoya polynomial, we refer to [39,40].

Table 1. Exhaustively generated sets of non-isomorphic trees: $|T_{10}|=106, |T_{11}|=235, |T_{12}|=551, |T_{13}|=1301$.

Measure	T_{10}		T_{11}		T_{12}		T_{13}	
	ndv	S	ndv	S	ndv	S	ndv	S
HS_1	0	1,000000	0	1,000000	0	1,000000	0	1,000000
HS_2	0	1,000000	0	1,000000	0	1,000000	0	1,000000
HS_3	0	1,000000	0	1,000000	0	1,000000	0	1,000000
I_H	76	0,283018	201	0,144680	499	0,094373	1237	0,049192

doi:10.1371/journal.pone.0102459.t001

In an earlier paper [3], we explored the discrimination power of measures (see section ‘Hosoya-based Indices’) that are based on the moduli of the zeros of the partial Hosoya polynomial. The main contribution of this paper is to define the *Hosoya-Spectral indices* combining structural information captured by partial Hosoya polynomials with graph spectra. Also, we examine the discrimination power of these indices and of the Hosoya information content of a graph. We discuss and compare the numerical results with the earlier ones produced in [3]. Further we elaborate on the usage of these measures as highly discriminating graph invariants.

Methods and Results

Hosoya-Based Indices

In this section, we reproduce the graph indices based on partial Hosoya polynomials, see [3]. As outlined in [3], the partial Hosoya polynomial of a vertex v_i in the graph $G=(V,E)$ is given by [41,42]

$$H^{v_i}(G,z) := \sum_{\substack{v \in V \\ v \neq v_i}} z^{d(v_i,v)}, \tag{1}$$

where $d(v_i,v)$ is the distance (i.e., length of a shortest path) between the vertices v_i and v . Solving the equation

$$H^{v_i}(G,z)=0, \tag{2}$$

yields the complex zeros $z_1^{v_i}, z_2^{v_i}, \dots, z_{k_i}^{v_i}$ which are not equal to zero. We infer $k_i < |V|$ by applying the well-known fundamental theorem of Algebra [43,44] stating that a complex polynomial

$$f(z) := a_k z^k + a_{k-1} z^{k-1} + a_1 z + a_0, a_k \neq 0, a_k \in \mathbb{C}, \tag{3}$$

with degree $\deg(f)=k$ has k complex zeros.

Also in [3], Dehmer et al. introduced the following indices:

$$M_1(G) := (|z_1^{v_1}| + |z_2^{v_1}| + \dots + |z_{k_1}^{v_1}|) + (|z_1^{v_2}| + |z_2^{v_2}| + \dots + |z_{k_2}^{v_2}|) + \dots + (|z_1^{v_{|V|}}| + |z_2^{v_{|V|}}| + \dots + |z_{k_{|V|}}^{v_{|V|}}|), \tag{4}$$

$$M_2(G) := \sqrt{|z_1^{v_1}| + |z_2^{v_1}| + \dots + |z_{k_1}^{v_1}|} + \sqrt{|z_1^{v_2}| + |z_2^{v_2}| + \dots + |z_{k_2}^{v_2}|} + \dots + \sqrt{|z_1^{v_{|V|}}| + |z_2^{v_{|V|}}| + \dots + |z_{k_{|V|}}^{v_{|V|}}|}, \tag{5}$$

and

$$M_3(G) := (|z_1^{v_1}| + |z_2^{v_1}| + \dots + |z_{k_1}^{v_1}|) \log(|z_1^{v_1}| + |z_2^{v_1}| + \dots + |z_{k_1}^{v_1}|) + (|z_1^{v_2}| + |z_2^{v_2}| + \dots + |z_{k_2}^{v_2}|) \log(|z_1^{v_2}| + |z_2^{v_2}| + \dots + |z_{k_2}^{v_2}|) + \dots + (|z_1^{v_{|V|}}| + |z_2^{v_{|V|}}| + \dots + |z_{k_{|V|}}^{v_{|V|}}|) \log(|z_1^{v_{|V|}}| + |z_2^{v_{|V|}}| + \dots + |z_{k_{|V|}}^{v_{|V|}}|). \tag{6}$$

Here, M_1 is the sum of the sums of the moduli of all partial Hosoya polynomials $H^{v_1}, H^{v_2}, \dots, H^{v_{|V|}}$. M_2 is the sum of the square roots of the sums of the moduli of all partial Hosoya polynomials. M_3 represents an entropy-like measure taking the sums of the moduli of all partial Hosoya polynomials into account.

Spectra of graphs have been investigated extensively [9,12]. As already mentioned, well-known spectral based indices are the Estrada index [18–20] and various forms of graph energy due to Gutman, see [23–25]. Since Hosoya polynomials and graph spectra capture different aspects of graph structure, we propose to combine the two in one index. So, let $\lambda_1 \dots \lambda_{|V|}$ be the eigenvalues of $\chi(G) := \det(A - \lambda E)$; A the adjacency matrix of G . The *Hosoya-Spectral indices* $HS_i(G)$ are defined as follows:

$$HS_1(G) := \lambda_1 (|z_1^{v_1}| + |z_2^{v_1}| + \dots + |z_{k_1}^{v_1}|) + \lambda_1 (|z_1^{v_2}| + |z_2^{v_2}| + \dots + |z_{k_2}^{v_2}|) + \dots + \lambda_{|V|} (|z_1^{v_{|V|}}| + |z_2^{v_{|V|}}| + \dots + |z_{k_{|V|}}^{v_{|V|}}|), \tag{7}$$

$$HS_2(G) := \lambda_1 \sqrt{|z_1^{v_1}| + |z_2^{v_1}| + \dots + |z_{k_1}^{v_1}|} + \lambda_2 \sqrt{|z_1^{v_2}| + |z_2^{v_2}| + \dots + |z_{k_2}^{v_2}|} + \dots + \lambda_{|V|} \sqrt{|z_1^{v_{|V|}}| + |z_2^{v_{|V|}}| + \dots + |z_{k_{|V|}}^{v_{|V|}}|}, \tag{8}$$

$$HS_3(G) := \lambda_1 (|z_1^{v_1}| + |z_2^{v_1}| + \dots + |z_{k_1}^{v_1}|) \log(|z_1^{v_1}| + |z_2^{v_1}| + \dots + |z_{k_1}^{v_1}|) + \lambda_2 (|z_1^{v_2}| + |z_2^{v_2}| + \dots + |z_{k_2}^{v_2}|) \log(|z_1^{v_2}| + |z_2^{v_2}| + \dots + |z_{k_2}^{v_2}|) + \dots + \lambda_{|V|} (|z_1^{v_{|V|}}| + |z_2^{v_{|V|}}| + \dots + |z_{k_{|V|}}^{v_{|V|}}|) \log(|z_1^{v_{|V|}}| + |z_2^{v_{|V|}}| + \dots + |z_{k_{|V|}}^{v_{|V|}}|). \tag{9}$$

Inspired by studying information-theoretic complexity measures [1,28,31–33,45], we define the *Hosoya information content* of a graph G . Let V_i for $1 \leq i \leq k$ be the set of all vertices in G with the i -th of k partial Hosoya polynomials of the vertices of $G=(V,E)$; V_i is then the i -th block in a partition of V . The *Hosoya information content* of G is defined by

Table 2. Exhaustively generated sets of non-isomorphic trees: $|T_{14}|=3159$, $|T_{15}|=7741$, $|T_{16}|=19320$, $|T_{17}|=48629$.

Measure	T_{14}		T_{15}		T_{16}		T_{17}	
	ndv	S	ndv	S	ndv	S	ndv	S
HS_1	0	1,000000	0	1,000000	0	1,000000	0	1,000000
HS_2	0	1,000000	0	1,000000	0	1,000000	0	1,000000
HS_3	0	1,000000	0	1,000000	0	1,000000	0	1,000000
I_H	3067	0,029123	7637	0,013434	19178	0,007349	48629	0,003783

doi:10.1371/journal.pone.0102459.t002

$$I_H(G) := - \sum_{i=1}^k \frac{|V_i|}{|V|} \log\left(\frac{|V_i|}{|V|}\right). \quad (10)$$

In the next section, we evaluate the discrimination power of this new measure and the Hosoya-Spectral indices on exhaustively generated graphs.

Numerical Results

As in [3], we present the numerical values resulting from the evaluation of the discrimination power of the Hosoya-Spectral indices and the Hosoya information content. In order to do so, we use the same graph classes in order to make direct comparisons [3]; $T_i, 10 \leq i \leq 18$ are the sets of all non-isomorphic trees with i vertices. N_9 is the set of all non-isomorphic graphs with 9 vertices, see [2].

To evaluate the discrimination power quantitatively, we use the same measures as in [3]: ndv stands for the number of non-distinguishable graphs according to the values of the indices. From this, we also compute $S(I) := \frac{|G| - \text{ndv}}{|G|}$ where $I : G \rightarrow \mathbb{R}$, see [46].

As in [3], the measures HS_i are fully unique on T_{10}, \dots, T_{13} , see Table 1. We obtain the same result by applying the Hosoya-Spectral indices to T_{14}, \dots, T_{17} , see Table 2. See also [3]. Moreover, the Hosoya-Spectral indices can discriminate the tree class T_{18} uniquely (see Table 3). Note that the earlier defined measures M_i (see Equation 4–6) [3] produced the first degeneracies on the set T_{18} .

The exhaustively generated graphs with 9 vertices (N_9) warrants special attention. In contrast to the previously introduced measures M_i , Table 3 shows that the uniqueness of HS_i is high. More precisely, HS_1 and HS_2 can discriminate 99% of the graph uniquely. The discrimination power of HS_3 is approximately 95%. This marks a considerable improvement compared with the measures M_i evaluated in [3] on the same classes of graphs. The

improvement of the discrimination power of the new measures can be explained by the fact that partial Hosoya polynomials and graph spectra capture quite different aspects of graph structure. In particular, the partial Hosoya polynomial captures local graph properties related to distances in a graph, and the indices M_i take account of the moduli of the zeros of these polynomials. By contrast, the spectrum of a graph captures connectivity properties linked to its adjacency matrix. The combination of these graph properties in the measures HS_i plausibly accounts for their superior performance over the single property measures (M_i).

Evidently, the discrimination power of I_H declines as the graph classes grow in size, i.e., the greater the cardinality of the graph class, the lower is index's discrimination power (measured by ndv and S), see Table 1–3. Even for small classes, the degeneracy is high. For N_9 , the Hosoya information content I_H cannot discriminate at all and, hence, $S=0$. These results are not surprising in view of the definition of Hosoya information content. The blocks of the partitions consist of vertices with the same partial Hosoya polynomial. Thus, the more cycles in a graph, the greater the likelihood of obtaining large blocks of vertices with the same partial Hosoya polynomial. The occurrence of such large blocks results in high values for the quantity ndv (and low values for S).

Summary and Conclusions

In this paper, we defined the Hosoya-Spectral indices as well as the Hosoya information content of a graph. The former measures combine structural information captured by partial Hosoya polynomials and graph spectra. It is evident that those two graph features capture structural information differently and, hence, the resulting measures may be more unique than the ones (M_i) used in earlier work, see [3]. The numerical study reported here has confirmed this conjecture for both trees and graphs. Finally, as expected, the discrimination power of Hosoya information content was found to be very low.

In future research, we plan to explore extremal properties of both measures. In particular, Hosoya information content is

Table 3. Exhaustively generated sets of non-isomorphic trees and graphs: $|T_{18}|=123757$, $|N_9|=261080$.

Measure	T_{18}		N_9	
	ndv	S	ndv	S
HS_1	0	1,000000	24	0,999908
HS_2	0	1,000000	18	0,999931
HS_3	0	1,000000	12393	0,952723
I_H	123512	0,001979	261080	0,000000

doi:10.1371/journal.pone.0102459.t003

related to the orbit structure of a graph, and this calls for studying the automorphism groups of certain classes of graphs.

Acknowledgments

We thank Lukas Huber for help regarding the data analysis and fruitful discussions

References

- Bonchev D, Mekenyan O, Trinajstić N (1981) Isomer discrimination by topological information approach. *J Comp Chem* 2: 127–148.
- Dehmer M, Grabner M, Varmuza K (2012) Information indices with high discriminative power for graphs. *PLoS ONE* 7: e31214.
- Dehmer M, Mowshowitz A, Shi Y (2014) Discrimination power of graph measures based on complex zeros of the partial Hosoya polynomial. submitted for publication.
- Konstantinova EV, Paleev AA (1990) Sensitivity of topological indices of polycyclic graphs. *Vychisl Sistemy* 136: 38–48.
- Petrov E, Stoyanov B, Kochev N, Bangov I (2014) Use of cti index for perception of duplicated chemical structures in large chemical databases. *MATCH Commun Math Comput Chem* 71: 645–656.
- Balaban AT (1982) Highly discriminating distance-based topological index. *Chem Phys Lett* 89: 399–404.
- Randić M (2013) d_{MAX} matrix of dominant distances in a graph. *MATCH Commun Math Comput Chem* 70: 221–238.
- Randić M, Orel R, Balaban AT (2013) d_{MAX} matrix graph invariants as graph descriptors. graphs having the same balaban j index. *MATCH Commun Math Comput Chem* 70: 239–258.
- Cvetković DM, Doob M, Sachs H (1980) Spectra of Graphs. Theory and Application. Deutscher Verlag der Wissenschaften. Berlin, Germany.
- Estrada E (2011) The Structure of Complex Networks. Theory and Applications. Oxford University Press.
- Chung F (1997) Spectral Graph Theory, volume 12 of *Cbms Regional Conference Series in Mathematics*. American Mathematical Society.
- Cioabă SM (2010) Some applications of eigenvalues of graphs. In: Dehmer M, editor, *Structural Analysis of Complex Networks*, Boston/Basel: Birkhäuser. pp. 357–379.
- Wu J, Deng HZ, Tan YJ (2010) Spectral measure of robustness for internet topology. In: 2010 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT). volume 6, pp. 50–54.
- Randić M, Vracko M, Novic M (2001) Eigenvalues as molecular descriptors. In: Diudea MV, editor, *QSPR/QSAR Studies by Molecular Descriptors*, Nova Publishing. pp. 93–120. Huntington, NY, USA.
- Dehmer M, Sivakumar L, Varmuza K (2012) Uniquely discriminating molecular structures using novel eigenvalue-based descriptors. *MATCH Communications in Mathematical and in Computer Chemistry* 67: 147–172.
- Janežič D, Miležević A, Nikolić S, Trinajstić N (2009) Topological complexity of molecules. In: Meyers R, editor, *Encyclopedia of Complexity and System Science*, Springer, volume 5. pp. 9210–9224.
- Das KC, Lee SG (2009) On the Estrada index conjecture. *Linear Algebra and its Applications* 431: 1351–1359.
- Estrada E (2000) Characterization of 3d molecular structure. *Chemical Physics Letters* 319: 713–718.
- Estrada E (2002) Characterization of the folding degree of proteins. *Bioinformatics* 18: 697–704.
- Gutman I, Estrada E, Rodríguez-Velázquez JA (2007) On a graph-spectrum-based structure descriptor. *Croatica Chemica Acta* 80: 151–154.
- Gutman I, Deng H, Randenković S (2011) The Estrada index: An updated survey. In: Cvetković D, Gutman I, editor, *Selected Topics on Applications of Graph Spectra*, Mathematical Institute of the Serbian Academy of Sciences and Arts, Belgrade. pp. 155–174.
- Li J, Shiu WC, Chan WH (2011) Note on the Laplacian Estrada Index of a Graph. *MATCH Commun Math Comput Chem* 6: 777–784.
- Gutman I (2001) The energy of a graph: Old and new results. In: Betten A, Kohnert A, Laue R, Wassermann A, editor, *Algebraic Combinatorics and Applications*, Springer Verlag. pp. 196–211. Berlin.
- Gutman I, Zhou B (2006) Laplacian energy of a graph. *Linear Algebra and its Applications* 414: 29–37.
- Gutman I, Li X, Zhang J (2009) Graph energy. In: Dehmer M, Emmert-Streib F, editor, *Analysis of Complex Networks: From Biology to Linguistics*, Wiley-VCH. pp. 145–174.
- Li X, Shi Y, Gutman I (2012) *Graph Energy*. Springer.
- Elphick C, Wocjan P (2014) New measures of graph irregularity. *Electronic Journal of Graph Theory and Applications* 2: 52–65.
- Mowshowitz A (1968) Entropy and the complexity of the graphs I: An index of the relative complexity of a graph. *Bull Math Biophys* 30: 175–204.
- Rashevsky N (1955) Life, information theory, and topology. *Bull Math Biophys* 17: 229–235.
- Körner J (1973) Coding of an information source having ambiguous alphabet and the entropy of graphs. *Transactions of the 6-th Prague Conference on Information Theory*: 411–425.
- Bonchev D, Trinajstić N (1977) Information theory, distance matrix and molecular branching. *J Chem Phys* 67: 4517–4533.
- Bonchev D (1979) Information indices for atoms and molecules. *MATCH Commun Math Comp Chem* 7: 65–113.
- Bonchev D (1983) *Information Theoretic Indices for Characterization of Chemical Structures*. Research Studies Press, Chichester.
- Bonchev D (2000) Overall connectivities and topological complexities: A new powerful tool for QSPR/QSAR. *J Chem Inf Comput Sci* 40: 934–941.
- Hosoya H (1971) Topological index. A newly proposed quantity characterizing the topological nature of structural isomers of saturated hydrocarbons. *Bull Chem Soc Jpn* 44: 2332–2339.
- Dehmer M, Mowshowitz A (2011) A history of graph entropy measures. *Information Sciences* 1: 57–78.
- Devillers J, Balaban AT (1999) *Topological Indices and Related Descriptors in QSAR and QSPR*. Gordon and Breach Science Publishers. Amsterdam, The Netherlands.
- Mehler A, Weiß P, Lücking A (2010) A network model of interpersonal alignment. *Entropy* 12: 1440–1483.
- Gutman I, Zhang Y, Dehmer M, Ilić A (2012) Altenburg, Wiener, and Hosoya polynomials. In: Gutman I, Furtula B, editor, *Distance in Molecular Graphs - Theory*, Univ. Kragujevac. pp. 49–70. Kragujevac.
- Deutsch E, Klavžar S (2013) Computing the hosoya polynomial of graphs from primary subgraphs. *MATCH Commun Math Comput Chem* 70: 627–644.
- Došlić T (2008) Vertex-weighted Wiener polynomials for composite graphs. *Ars Math Contemp* 1: 66–80.
- Hosoya H (1988) On some counting polynomials. *Discrete Applied Mathematics* 19: 239–257.
- Dehmer M, Mowshowitz A (2011) Bounds on the moduli of polynomial zeros. *Applied Mathematics and Computation* 218: 4128–4137.
- Fricke R (1924) *Lehrbuch der Algebra*. Friedrich Vieweg & Sohn Verlag. Braunschweig, Germany.
- Dehmer M, Borgert S, Emmert-Streib F (2008) Entropy bounds for molecular hierarchical networks. *PLoS ONE* 3: e3079.
- Konstantinova EV, Skorobogatov VA, Vidyuk MV (2002) Applications of information theory in chemical graph theory. *Indian Journal of Chemistry* 42: 1227–1240.

Author Contributions

Contributed to the writing of the manuscript: MD YS AM.