Article

Information Theory of Networks

Matthias Dehmer

UMIT, Institute for Bioinformatics and Translational Research, Eduard Wallnöfer Zentrum 1, 6060, Hall in Tyrol, Austria; E-Mail: Matthias.Dehmer@umit.at; Tel./Fax: +43-050-8648-3851

Abstract: The paper puts the emphasis on surveying information-theoretic network measures for analyzing the structure of networks. In order to apply the quantities interdisciplinarily, we also discuss some of their properties such as their structural interpretation and uniqueness.

Keywords: information theory; networks; quantitative graph analysis

1. Introduction

Information theory has been proven useful to solve interdisciplinary problems. For example, problems in biology, chemistry, computer science, ecology, electrical engineering, and neuroscience have been tackled by using information-theoretic methods such as entropy and mutual information, see [1–5]. In particular, advanced information-measures such as the Jensen–Shannon divergence have also been used for performing biological sequence analysis [6].

In terms of investigating networks, information-theoretic techniques have been applied in an interdisciplinary manner [7–11]. In this paper, we put the emphasis on reviewing information-theoretic measures to explore the network structure and shed light on some of their strong and weak points. But note that the problem of exploring the dynamics of networks by using information theory has also been tackled, see [12].

Interestingly, the problem of exploring graphs quantitatively emerged in the fifties when investigating structural aspects of biological and chemical systems [13,14]. In this context, an important problem is to quantify the structural information content of graphs by using Shannon's information measure [14–19]. This groundbreaking work led to numerous measurements of network complexity by using Shannon's information measure [15,20–22]. Particularly this task firstly appeared when

studying the complexity of chemical and biological systems [13,23–25]. Besides studying chemical and biological questions, the structural complexity of networks have been also explored in computer science [1,26], ecology [10,27–29], information theory [30], linguistics [31,32], sociology [33,34], and mathematical psychology [33,35]. Also, information-theoretic approaches for investigating networks have been employed in network physics, see [7,36,37].

As the measures have been explored interdisciplinarily, it is particularly important to understand their strong and weak points. Otherwise, the results of applications involving the measures can not be understood properly. Besides surveying the most important measures, the main contribution is to highlight some of their strong and weak points. In this paper, this relates to better understand their *structural interpretation* and to gain insights about their *uniqueness*. The uniqueness, often called the discrimination power or degeneracy of an information-theoretic graph measure (and of course of any graph measure) relates to the property how well it can discriminate non-isomorphic graphs by its values, see [38–40]. An important problem is to evaluate the degree of degeneracy of a measure by several quantities such as the sensitivity measure due to Konstantinova [40]. Note that the discrimination power of a measure clearly depends on the graph class in question, see [38,41].

2. Graph Entropies

2.1. Measures Based on Equivalence Criteria and Graph Invariants

To find such measures, seminal work was done by Bonchev [8,42], Mowshowitz [15–18], Rashevsky [19] and Trucco [14]. Chronologically, Rashevsky [19], MacArthur [27] and Trucco [14] were the first who applied Shannon's information measure to derive an entropy of a graph characterizing its topology. Then, Mowshowitz [15–18] called it *structural information content* of a graph and developed a theory to study the properties of such graph entropies under certain graph operations such as product, join *etc*. So far, numerous related quantities have been defined by applying the general approach of deriving partitions based on a graph invariant which is due to Mowshowitz [15]. As a result of these developments, Bertz [43], Basak *et al*. [44,45] and Bonchev [8,42,46] contributed various related measures which are all based on the idea of deriving partitions by using a graph invariant, e.g., vertices, edges, degrees, and distances.

Let $G = (V, E)$ be a graph, X be a graph invariant, and τ be an equivalence criterion. Then, G can be partitioned with respect to the elements of the graph invariant under consideration. From this procedure, one also obtains probability values for each partition [8,15] given by $p_i := \frac{|X_i|}{|X|}$. By applying Shannon's information measure [5], we yield the graph entropies as follows [8]:

$$I_t(G, \tau) := |X| \log(|X|) - \sum_{i=1}^{k} |X_i| \log(|X_i|) \tag{1}$$

$$I_m(G, \tau) := - \sum_{i=1}^{k} \frac{|X_i|}{|X|} \log\left(\frac{|X_i|}{|X|}\right) \tag{2}$$

where k equals the number of different partitions. I_t is called total information content and I_m is called the mean information content of G, respectively.

In the following, we survey graph entropy measures exemplarily by applying this principle. Besides well-known quantities, we also mention more recently developed indices.

1. Topological information content due to Rashevsky [19]:

$$I_a(G) := - \sum_{i=1}^{k} \frac{|N_i|}{|V|} \log \left(\frac{|N_i|}{|V|} \right) \tag{3}$$

   $|N_i|$ denotes the number of topologically equivalent vertices in the i-th vertex orbit of G. k is the number of different orbits. This measure is based on symmetry in a graph as it relies on its automorphism group and vertex orbits. It can be easily shown that I_a vanishes for vertex transitive graphs. Also, it attains maximum entropy for asymmetric graphs. However, it has been shown [41] that these symmetry-based measures possess little discrimination power. The reason for this is that many non-isomorphic graphs have the same orbit structure and, hence, they can not be distinguished by this index. Historically seen, the term *topological information content* was proposed by Rashevski [19]. Then, Trucco [14] redefined the measure in terms of graph orbits. Finally, Mowshowitz [15] studied extensively mathematical properties of this information measure for graphs (e.g., the behavior of I_a under graph operations) and generalized it by considering infinite graphs [18].

2. Symmetry index for graphs due to Mowshowitz *et al.* [47]:

$$S(G) := (\log(|V|) - I_a(G)) + \log(|\mathbf{Aut}(G)|) \tag{4}$$

   In [47], extremal values of this index and formulas for special graph classes such as wheels, stars and path graphs have been studied. As conjectured, the discrimination power of S turned out to be higher than by using I_a as a discriminating term $\log(|\mathbf{Aut}(G)|)$ has been added, see Equation (4). In particular, we obtained this result by calculating S on a set of 2265 chemical graphs whose order range from four to nineteen. A detailed explanation of the dataset can be found in [48].

3. Chromatic information content due to Mowshowitz [15,16]:

$$I_c(G) := \min_{\hat{V}} \left\{ - \sum_{i=1}^{h} \frac{n_i(\hat{V})}{|V|} \log \left(\frac{n_i(\hat{V})}{|V|} \right) \right\} \tag{5}$$

   where $\hat{V} := \{V_i | 1 \leq i \leq h\}$, $|V_i| := n_i(\hat{V})$ denotes an arbitrary chromatic decomposition of a graph G. $h = \chi(G)$ is the chromatic number of G. Graph-theoretic properties of I_c and its behavior on several graph classes have been explored by Mowshowitz [15,16]. To our knowledge, the structural interpretation of this measure as well as the uniqueness has not yet been explored extensively.

4. Magnitude-based information indices due to Bonchev *et al.* [49]:

$$I_D(G) := -\frac{1}{|V|} \log \left(\frac{1}{|V|} \right) - \sum_{i=1}^{\rho(G)} \frac{2k_i}{|V|^2} \log \left(\frac{2k_i}{|V|^2} \right) \tag{6}$$

$$I_D^W(G) := W(G) \log(W(G)) - \sum_{i=1}^{\rho(G)} i k_i \log(i) \tag{7}$$

where k_i is the occurrence of a distance possessing value i in the distance matrix of G. The motivation to introduce these measures was to find quantities which detect branching well, see [49]. In this context, branching of a graph correlates with the number of terminal vertices. By using this model, Bonchev *et al.* [49] showed numerically and by means of inequalities that these indices detect branching meaningfully. Also, it turned out that magnitude-based information indices possess high discrimination power for trees. But recent studies [50] have shown that the uniqueness of the magnitude-based information indices deteriorate tremendously when being applied to large sets of graphs containing cycles. More precisely, Dehmer *et al.* [50] evaluated the uniqueness of several graph entropy measures and other topological indices by using almost 12 million non-isomorphic, connected and unweighted graphs possessing ten vertices.

5. Vertex degree equality-based information index found by Bonchev [8]:

$$I_{\deg}(G) := \sum_{i=1}^{\bar{k}} \frac{|N_i^{k_v}|}{|V|} \log \left(\frac{|N_i^{k_v}|}{|V|} \right) \tag{8}$$

where $|N_i^{k_v}|$ is the number of vertices with degree equal to i and $\bar{k} := \max_{v \in V} k_v$. Note that this quantity is easy to determine as the time complexity of the calculation of the degrees is clearly polynomial. But it is intuitive that a simple comparison of the degree distribution of graphs is not meaningful to discriminate their structure. In [50], it has been shown that this measure possesses little discrimination power when applying the quantity to several sets of graphs.

6. Overall information indices found by Bonchev [46,51]:

$$OX(G) := \sum_{k=0}^{|E|} {}^k X; \quad \{X\} := \{{}^0 X, {}^1 X, \ldots, {}^{|E|} X\} \tag{9}$$

$$I(G, OX) := OX \log(OX) - \sum_{k=0}^{|E|} {}^k X \log \left({}^k X \right) \tag{10}$$

The index calculates the overall value OX of a certain graph invariant X by summing up its values in all subgraphs, and partitioning them into terms of increasing orders (increasing number of subgraph edges k). In the simplest case, we have $OX = SC$, *i.e.*, it is equal to the subgraph count [51]. Several more overall indices and their informational functionals have been calculated, such as overall connectivity (the sum of total adjacency of all subgraphs), overall Wiener index (the sum of total distances of all subgraphs), the overall Zagreb indices, and the overall Hosoya index [51]. They all share (with some inessential variations) the property to increase in value with the increase in graph complexity. The properties of most of these information functionals will not be studied here in detail.

Clearly, we only surveyed a subset of existing graph entropy measures. Further measures which are based on the same criterion can be found in [51–53]. Also, we would like to mention that information measures for graphs based on other entropy measures have been studied [54]. For instance, Passerini and Severini [54] explored the von Neumann entropy of networks in the context of network physics. Altogether, the variety of existing network measures bears great potential for analyzing complex networks quantitatively. But in the future, the usefulness and ability of these measures must be investigated more extensively to gain further theoretical insights in terms of their properties.

2.2. *Körner Entropy*

The definition of the Körner entropy is rooted in information theory and has been introduced to solve a particular coding problem, see [30,55]. Simony [55] discussed several definitions of this quantity which have been proven to be equivalent. One definition thereof is

$$H(G, P) := \lim_{t \longrightarrow \infty} \min_{U \subseteq V^t, P^t(U) > 1 - \epsilon} \frac{1}{t} \log(\chi(G^t(U))) \tag{11}$$

For $V' \subseteq V(G)$, the induced subgraph on V' is denoted by $G(V')$ and $\chi(G)$ is the chromatic number [56] of G, G^t the t-th co-normal power [30] of G and

$$P^t(U) := \sum_{x \in U} P^t(x) \tag{12}$$

Note that $P^t(x)$ is the probability of the string x, see [55]. Examples and the interpretation of this graph entropy measure can be found in [30,55]. Due to the fact that its calculation relies on the stable set problem, its computational complexity may be insufficient. To our knowledge, the Körner entropy has not been used as a graph complexity measure in the sense of the quantities described in the previous section. That means, it does not express the structural information content of a graph (as the previously mentioned graph entropies) as it has been used in a different context, see [30,55]. Also, its computational complexity makes it impossible to apply this quantity on a large scale and to investigate properties such as correlation and uniqueness.

2.3. *Entropy Measures Using Information Functionals*

Information-theoretic complexity measures for graphs can also be inferred by assigning a probability value to each vertex of a graph in question [9,21]. Such probability values have been defined by using information functionals [9,21,48]. In order to define these information functionals, some key questions must be answered:

- What kind of structural features (e.g., vertices, edges, degrees, distances *etc.*) should be used to derive meaningful information functionals?
- In this context, what does "meaningful" mean?
- In case the functional is parametric, how can the parameters be optimized?
- What kind of structural information does the functional as well as the resulting entropy detect?

To discuss the first item, see [9,21,48] and note that metrical properties have been used to derive such information functionals. In order to prove whether a functional as well as the resulting entropy measures captures structural information meaningfully, an optimality criterion is needed. For example, suppose there exists a data set where the class labels of its entities (graphs) are known. By employing supervised machine learning techniques, the classification error can be optimized. Note that the last item relates to investigate the structural interpretation of the graph entropy measure. Indeed, this question could be raised for any topological index.

In order to reproduce some of these measures, we start with a graph $G = (V, E)$ and let f be an information functional representing a positive function that maps vertices to the positive reals. Note that f captures structural information of G. If we define the vertex probabilities as [9,21]

$$p(v_i) := \frac{f(v_i)}{\sum_{j=1}^{|V|} f(v_j)} \tag{13}$$

we yield the families of information-theoretic graph complexity measures [9,48]:

$$I_f(G) := -\sum_{i=1}^{|V|} \frac{f(v_i)}{\sum_{j=1}^{|V|} f(v_j)} \log\left(\frac{f(v_i)}{\sum_{j=1}^{|V|} f(v_j)}\right) \tag{14}$$

$$I_f^\lambda(G) := \lambda\left(\log(|V|) + \sum_{i=1}^{|V|} \frac{f(v_i)}{\sum_{j=1}^{|V|} f(v_j)} \log\left(\frac{f(v_i)}{\sum_{j=1}^{|V|} f(v_j)}\right)\right) \tag{15}$$

$\lambda > 0$ is a scaling constant. Typical information functionals are [9,21,48]

$$f_1(v_i) := \alpha^{c_1 |S_1(v_i, G)| + c_2 |S_2(v_i, G)| + \cdots + c_{\rho(G)} |S_{\rho(G)}(v_i, G)|}, \ c_k > 0, 1 \le k \le \rho(G), \alpha > 0 \tag{16}$$

and

$$f_2(v_i) := c_1 |S_1(v_i, G)| + c_2 |S_2(v_i, G)| + \cdots + c_{\rho(G)} |S_{\rho(G)}(v_i, G)|, \ c_k > 0, 1 \le k \le \rho(G) \tag{17}$$

The parameters $c_k > 0$ to weight structural characteristics or differences of G in each sphere have to be chosen such that at least $c_i \ne c_j$ holds. Otherwise the probabilities become $\frac{1}{|V|}$ leading to maximum entropy $\log(|V|)$. For instance, the setting $c_1 > c_2 > \cdots > c_{\rho(G)}$ have often been used, see [9,21,48]. Also, other schemes for the coefficients can be chosen but need to be interpreted in terms of the structural interpretation of the resulting entropy measure. As the measures are parametric (when using a parametric information functional), they can be interpreted as generalizations of the aforementioned partition-based measures.

By applying Equation (15), concrete information measures to characterize the structural complexity chemical structures have been derived in [48]. For example, if we choose the coefficients linearly decreasing, e.g.,

$$c_1 := \rho(G), \ c_2 := \rho(G) - 1, \ldots, \ c_{\rho(G)} := 1 \tag{18}$$

or exponentially decreasing, e.g.,

$$c_1 := \rho(G), c_2 := \rho(G)e^{-1}, \ldots, c_{\rho(G)} := \rho(G)e^{-\rho(G)+1} \tag{19}$$

the resulting measures are called $I_{f_{lin}^V}^\lambda$ and $I_{f_{exp}^V}^\lambda$, respectively. Importantly, it turned out that $I_{f_{lin}^V}^\lambda$ and $I_{f_{exp}^V}^\lambda$ possess high discrimination power when applying them to real and synthetic chemical graphs, see [48].

To obtain more advanced information functionals, the concept outlined above has been extended in [57]. The main idea for deriving these information functionals is based on the assumption that starting from an arbitrary vertex $v_i \in V$, information spreads out via shortest paths in the graph which can be determined by using Dijkstra's algorithm [58]. Then, more sophisticated information functionals as well as complexity measures have been defined [57] by using local property measures, e.g., vertex centrality

measures [59]. In particular, some of them turned out to be highly unique when applying the measures to almost 12 million non-isomorphic, connected and unweighted graphs possessing ten vertices [50]. Interestingly, the just mentioned information-theoretic complexity measures showed a constantly high uniqueness that does not depend much on the cardinality of the underlying graph set. This property is desirable as we found that the uniqueness of the most existing measures deteriorates dramatically if the cardinality of the underlying graph set increases.

2.4. Information-Theoretic Measures for Trees

In this section, we sketch a few entropic measures which have been developed to characterize trees structurally. For example, Emmert-Streib *et al.* [60] developed an approach to determine the structural information content of rooted trees by using the natural partitioning of the vertices in such a tree. That means the number of vertices can be counted on each tree level which leads to a probability distribution and, thus, to an entropy characterizing the topology of a rooted tree. Dehmer [57] used this idea to calculate the entropy of arbitrary undirected graphs by applying a decomposition approach. Mehler [31] also employed entropic measures as balance and imbalance measures of tree-like graphs in the context of social network analysis. Other aspects of tree entropy have been tackled by Lions [61].

2.5. Other Information-Theoretic Network Measures

Apart from information-theoretic measures mostly used in mathematical and structural chemistry, several other entropic networks measures for measuring disorder relations in complex networks have been explored in the context of network physics, see [62]. If $P(k_v)$ denotes the probability of a vertex v possessing degree k, the distribution of the so-called remaining degree was defined by [62]

$$q(k_v) := \frac{(k+1)P_{k_v+1}}{<k>} \tag{20}$$

$<k> := \sum_k kP(k_v)$. By applying Shannon's information measure, the following graph entropy measure has been obtained [62]:

$$I(G) := \sum_{i=1}^{|V|} q(i) \log(q(i)) \tag{21}$$

It can be interpreted as a measure for determining the heterogeneity of a complex network [62]. In order to develop information indices for weighted directed networks, Wilhelm *et al.* [63] defined the measure called Medium Articulation that obtains its maximum for networks with a medium number of edges. It has been defined by [63]

$$\mathrm{MA}(G) := R(G) \cdot I(G) \tag{22}$$

where

$$R(G) := -\sum_{i,j} T_{v_i v_j} \log \left(\frac{T_{v_i v_j}^2}{\sum_k T_{v_k v_j} \sum_l T_{v_i v_l}} \right) \tag{23}$$

represents the redundancy and [63]

$$I(G) := \sum_{i,j} T_{v_i v_j} \log \left(\frac{T_{v_i v_j}}{\sum_k T_{v_k v_j} \sum_l T_{v_i v_l}} \right) \tag{24}$$

the mutual information.

Finally, the normalized flux from v_i to v_j is

$$T_{v_i v_j} := \frac{t_{v_i v_j}}{\sum_{k,l} t_{v_k v_l}} \qquad (25)$$

$t_{v_i v_j}$ is the flux (edge weight) between v_i and v_j. It can be easily shown that R vanishes for a directed ring but attains its maximum for the complete graph [63]. The behavior of I is just converse. This implies that MA vanishes for extremal graphs and attains its maximum in between [63]. We remark that a critical discussion of MA and modified measures have been recently contributed by Ulanowicz *et al.* [64].

For finalizing this section, we also reproduce the so-called offdiagonal complexity (OdC) [65] that is based on determining the entropy of the offdiagonal elements of the vertex-vertex link correlation matrix [65,66]. Let $G = (V, E)$ be a graph and let $(c_{ij})_{ij}$ be the vertex-vertex link correlation matrix, see [65]. Here c_{ij} denotes the number of all neighbors possessing degree $j > i$ of all vertices with degree i [66]. $\bar{k} := \max_{v \in V} k_v$ stands for the maximum degree of G. If one defines [66]

$$a_{|V|} := \sum_{i=1}^{\bar{k}-|V|} c_{i,i+|V|} \qquad (26)$$

and

$$b_{|V|} := \frac{a_{|V|}}{\sum_{|V|=0}^{\bar{k}-1} a_{|V|}} \qquad (27)$$

OdC can be defined by [66]

$$OdC := \frac{-\left(\sum_{|V|=0}^{\bar{k}-1} b_{|V|} \log(b_{|V|}) \right)}{\log(|V|-1)} \in [0, 1] \qquad (28)$$

As the measure depends on correlations between degrees of pairs of vertices [65], it is not surprising that its discrimination power is low, see [41].

3. Structural Interpretation of Graph Measures

We already mentioned the problem of exploring the structural interpretation of topological graph measures exemplarily in the preceding sections. In general, this relates to explore what kind of structural complexity a particular measure does detect. The following listing shows a few such *types of structural complexity* of measures which have already been explored:

- Branching in trees [49,67,68]. Examples for branching measures are the Wiener index [69], the magnitude-based measures also known as Bonchev–Trinajstić indices [49] and others outlined by Janežić *et al.* [68].
- Linear tree complexity depending on their size and symmetry [68]. Examples for such measures are the MI and MB indices, TC and TC1 Indices *etc.*, see [68].
- Balance and imbalance of tree-like graphs [31]. For examples, see [31].
- Cyclicity in graphs [23,38,68,70,71]. Note that in the context of mathematical chemistry, this graph property has been introduced and studied by Bonchev *et al.* [38]. Examples for branching measures are the BT and BI Indices, and the F index, see [70].

- Inner symmetry and symmetry in graphs [15,47,48,72]. Examples for such measures are I_a, S (see Section (2.1)) and I_{f_2} (see Section (2.3)).

In view of the vast amount of topological measures developed so far, determining their structural interpretation is a daunting problem. Evidently, it is important to contribute to this problem as measures could be then classified by this property. This might be useful when designing new measures or finding topological indices for solving a particular problem.

4. Summary and Conclusion

In this paper, we surveyed information-theoretic measures for analyzing networks quantitatively. Also, we discussed some of their properties, namely the structural interpretation and uniqueness. Because a vast number of measures have been developed, the former problem has been somewhat overlooked when analyzing topological network measures. Also, the uniqueness of information-theoretic and non-information-theoretic measures is a crucial property. Applications thereof might be interesting for applications such as problems in combinatorial chemistry [73]. In fact, many papers exist to tackle this problem [40,74–76] but not on a large scale. Interestingly, a statistical analysis has been recently shown [50] that the uniqueness of many topological indices strongly depends on the cardinality of a graph set in question. Also, it is clear that the uniqueness property depends on a particular graph class. This implies that results may not be generalized when the measure gives feasible results for a special class only, e.g., trees, isomers *etc*.

References

1. Allen, E.B. Measuring Graph Abstractions of Software: An Information-Theory Approach. In *Proceedings of the 8-th International Symposium on Software Metrics*, IEEE Computer Society: Ottawa, ON, Canada, 4–7 June 2002; p. 182.
2. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; Wiley & Sons: Hoboken, NJ, USA, 2006.
3. McDonnell, M.D.; Ikeda, S.; Manton, J.H. An introductory review of information theory in the context of computational neuroscience. *Biol. Cybern.* **2011**, *105*, 55–70.
4. Mathar, R.; Schmeink, A. A Bio-Inspired Approach to Condensing Information. In *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, IEEE Xplore: Saint-Petersburg, Russia, 31 July–5 August 2011; pp. 2524–2528.
5. Shannon, C.E.; Weaver, W. *The Mathematical Theory of Communication*; University of Illinois Press: Champaign, IL, USA, 1949.
6. Grosse, I.; Galván, P.B.; Carpena, P.; Roldán, R.R.; Oliver, J.; Stanley, H.E. Analysis of symbolic sequences using the Jensen-Shannon divergence. *Phys. Rev. E* **2002**, *65*, 041905:1–041905:16.

7. Anand, K.; Bianconi, G. Entropy measures for networks: Toward an information theory of complex topologies. *Phys. Rev. E* **2009**, *80*, 045102(R):1–045102(R):4.

8. Bonchev, D. *Information Theoretic Indices for Characterization of Chemical Structures*; Research Studies Press: Chichester, UK, 1983.

9. Dehmer, M. Information processing in complex networks: Graph entropy and information functionals. *Appl. Math. Comput.* **2008**, *201*, 82–94.

10. Hirata, H.; Ulanowicz, R.E. Information theoretical analysis of ecological networks. *Int. J. Syst. Sci.* **1984**, *15*, 261–270.

11. Kim, D.C.; Wang, X.; Yang, C.R.; Gao, J. Learning biological network using mutual information and conditional independence. *BMC Bioinf.* **2010**, *11*, S9:1–S9:8.

12. Barnett, L.; Buckley, L.; Bullock, C.L. Neural complexity and structural connectivity. *Phys. Rev. E* **2009**, *79*, 051914:1–051914:12.

13. Morowitz, H. Some order-disorder considerations in living systems. *Bull. Math. Biophys.* **1953**, *17*, 81–86.

14. Trucco, E. A note on the information content of graphs. *Bull. Math. Biol.* **1956**, *18*, 129–135.

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

16. Mowshowitz, A. Entropy and the complexity of graphs II: The information content of digraphs and infinite graphs. *Bull. Math. Biophys.* **1968**, *30*, 225–240.

17. Mowshowitz, A. Entropy and the complexity of graphs III: Graphs with prescribed information content. *Bull. Math. Biophys.* **1968**, *30*, 387–414.

18. Mowshowitz, A. Entropy and the complexity of graphs IV: Entropy measures and graphical structure. *Bull. Math. Biophys.* **1968**, *30*, 533–546.

19. Rashevsky, N. Life, information theory, and topology. *Bull. Math. Biophys.* **1955**, *17*, 229–235.

20. Bonchev, D. Information Theoretic Measures of Complexity. In *Encyclopedia of Complexity and System Science*; Meyers, R., Ed.; Springer: Berlin, Heidelberg, Germany, 2009; Volume 5, pp. 4820–4838.

21. Dehmer, M. A novel method for measuring the structural information content of networks. *Cybern. Syst.* **2008**, *39*, 825–842.

22. Mowshowitz, A.; Mitsou, V. Entropy, Orbits and Spectra of Graphs. In *Analysis of Complex Networks: From Biology to Linguistics*; Dehmer, M., Emmert-Streib, F., Eds.; Wiley-VCH: Hoboken, NJ, USA, 2009; pp. 1–22.

23. Bonchev, D.; Rouvray, D.H. *Complexity in Chemistry, Biology, and Ecology*; Springer: New York, NY, USA, 2005.

24. Dancoff, S.M.; Quastler, H. Information Content and Error Rate of Living Things. In *Essays on the Use of Information Theory in Biology*; Quastler, H., Ed.; University of Illinois Press: Champaign, IL, USA, 1953; pp. 263–274.

25. Linshitz, H. The Information Content of a Battery Cell. In *Essays on the Use of Information Theory in Biology*; Quastler, H., Ed.; University of Illinois Press: Urbana, IL, USA, 1953.

26. Latva-Koivisto, A.M. *Finding a Complexity Measure for Business Process Models*. Helsinki University of Technology: Espoo, Finland, 2001; doi:10.1.1.25.2991.

27. MacArthur, R.H. Fluctuations of animal populations and a measure of community stability. *Ecology* **1955**, *36*, 533–536.

28. Solé, R.V.; Montoya, J.M. Complexity and Fragility in Ecological Networks. *Proc. R. Soc. Lond. B* **2001**, *268*, 2039–2045.

29. Ulanowicz, R.E. Information theory in ecology. *Comput. Chem.* **2001**, *25*, 393–399.

30. Körner, J. Coding of an Information Source Having Ambiguous Alphabet and the Entropy of Graphs. In *Proceedings of the Transactions of the 6th Prague Conference on Information Theory*; Academia: Prague, Czechoslovakia, 19–25 September 1973; pp. 411–425.

31. Mehler, A. Social Ontologies as Generalized Nearly Acyclic Directed Graphs: A Quantitative Graph Model of Social Tagging. In *Towards an Information Theory of Complex Networks: Statistical Methods and Applications*; Dehmer, M., Emmert-Streib, F., Mehler, A., Eds.; Birkhäuser Boston: MA, USA, 2011; pp. 259–319.

32. Mehler, A.; Weiß, P.; Lücking, A. A network model of interpersonal alignment. *Entropy* **2010**, *12*, 1440–1483.

33. Balch, T. Hierarchic social entropy: An information theoretic measure of robot group diversity. *Auton. Robot.* **2001**, *8*, 209–237.

34. Butts, C.T. The complexity of social networks: Theoretical and empirical findings. *Soc. Netw.* **2001**, *23*, 31–71.

35. Sommerfeld, E.; Sobik, F. Operations on Cognitive Structures—Their Modeling on the Basis of Graph Theory. In *Knowledge Structures*; Albert, D., Ed.; Springer: Berlin, Heidelberg, Germany, 1994; pp. 146–190.

36. Krawitz, P.; Shmulevich, I. Entropy of complex relevant components of Boolean networks. *Phys. Rev. E* **2007**, *76*, 036115:1–036115:7.

37. Sanchirico, A.; Fiorentino, M. Scale-free networks as entropy competition. *Phys. Rev. E* **2008**, *78*, 046114:1–046114:10.

38. Bonchev, D.; Mekenyan, O.; Trinajstić, N. Topological characterization of cyclic structures. *Int. J. Quantum Chem.* **1980**, *17*, 845–893.

39. Dehmer, M.; Sivakumar, L.; Varmuza, K. Uniquely discriminating molecular structures using novel eigenvalue-based descriptors. *MATCH Commun. Math. Comput. Chem.* **2012**, *67*, 147–172.

40. Konstantinova, E.V. The discrimination ability of some topological and information distance indices for graphs of unbranched hexagonal systems. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 54–57.

41. Dehmer, M.; Barbarini, N.; Varmuza, K.; Graber, A. A large scale analysis of information-theoretic network complexity measures using chemical structures. *PLoS One* **2009**, *4*, doi:10.1371/journal.pone.0008057.

42. Bonchev, D. Information indices for atoms and molecules. *MATCH Commun. Math. Comp. Chem.* **1979**, *7*, 65–113.

43. Bertz, S.H. The first general index of molecular complexity. *J. Am. Chem. Soc.* **1981**, *103*, 3241–3243.

44. Basak, S.C.; Magnuson, V.R. Molecular topology and narcosis. *Drug Res.* **1983**, *33*, 501–503.

45. Basak, S.C. Information-Theoretic Indices of Neighborhood Complexity and their Applications. In *Topological Indices and Related Descriptors in QSAR and QSPAR*; Devillers, J., Balaban, A.T., Eds.; Gordon and Breach Science Publishers: Amsterdam, The Netherlands, 1999; pp. 563–595.

46. Bonchev, D. Overall connectivities and topological complexities: A new powerful tool for QSPR/QSAR. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 934–941.

47. Mowshowitz, A.; Dehmer, M. A symmetry index for graphs. *Symmetry Cult. Sci.* **2010**, *21*, 321–327.

48. Dehmer, M.; Varmuza, K.; Borgert, S.; Emmert-Streib, F. On entropy-based molecular descriptors: Statistical analysis of real and synthetic chemical structures. *J. Chem. Inf. Model.* **2009**, *49*, 1655–1663.

49. Bonchev, D.; Trinajstić, N. Information theory, distance matrix and molecular branching. *J. Chem. Phys.* **1977**, *67*, 4517–4533.

50. Dehmer, M.; Grabner, M.; Varmuza, K. Information indices with high discrimination power for arbitrary graphs. *PLoS One* submitted for publication, **2011**.

51. Bonchev, D. The Overall Topological Complexity Indices. In *Advances in Computational Methods in Science and Engineering*; Simos, T., Maroulis, G., Eds.; VSP Publications: Boulder, USA, 2005; Volume 4B, pp. 1554–1557.

52. Bonchev, D. My life-long journey in mathematical chemistry. *Internet Electron. J . Mol. Des.* **2005**, *4*, 434–490.

53. Todeschini, R.; Consonni, V.; Mannhold, R. *Handbook of Molecular Descriptors*; Wiley-VCH: Weinheim, Germany, 2002.

54. Passerini, F.; Severini, S. The von Neumann entropy of networks. *Int. J. Agent Technol. Syst.* **2009**, *1*, 58–67.

55. Simonyi, G. Graph Entropy: A Survey. In *Combinatorial Optimization*; Cook, W., Lovász, L., Seymour, P., Eds.; ACM: New York, NY, USA, 1995; Volume 20, pp. 399–441.

56. Bang-Jensen, J.; Gutin, G. *Digraphs. Theory, Algorithms and Applications*; Springer: Berlin, Heidelberg, Germany, 2002.

57. Dehmer, M. Information-theoretic concepts for the analysis of complex networks. *Appl. Artif. Intell.* **2008**, *22*, 684–706.

58. Dijkstra, E.W. A note on two problems in connection with graphs. *Numer. Math.* **1959**, *1*, 269–271.

59. Brandes, U.; Erlebach, T. *Network Analysis*; Springer: Berlin, Heidelberg, Germany, 2005.

60. Emmert-Streib, F.; Dehmer, M. Information theoretic measures of UHG graphs with low computational complexity. *Appl. Math. Comput.* **2007**, *190*, 1783–1794.

61. Lyons, R. Identities and inequalities for tree entropy. *Comb. Probab. Comput.* **2010**, *19*, 303–313.

62. Solé, R.V.; Valverde, S. *Information Theory of Complex Networks: On Evolution and Architectural Constraints*; Springer: Berlin, Heidelberg, Germany, 2004; Volume 650, pp. 189–207.

63. Wilhelm, T.; Hollunder, J. Information theoretic description of networks. *Physica A* **2007**, *388*, 385–396.

64. Ulanowicz, R.E.; Goerner, S.J.; Lietaer, B.; Gomez, R. Quantifying sustainability: Resilience, efficiency and the return of information theory. *Ecol. Complex.* **2009**, *6*, 27–36.

65. Claussen, J.C. Characterization of networks by the offdiagonal complexity. *Physica A* **2007**, *365-373*, 321–354.

66. Kim, J.; Wilhelm, T. What is a complex graph? *Physica A* **2008**, *387*, 2637–2652.

67. Bonchev, D. Topological order in molecules 1. Molecular branching revisited. *J. Mol. Strut. THEOCHEM* **1995**, *336*, 137 – 156.

68. Janežić, D.; Miležević, A.; Nikolić, S.; Trinajstić, N. Topological Complexity of Molecules. In *Encyclopedia of Complexity and System Science*; Meyers, R., Ed.; Springer: Berlin, Heidelberg, Germany, 2009; Volume 5, pp. 9210–9224.

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

70. Balaban, A.T.; Mills, D.; Kodali, V.; Basak, S.C. Complexity of chemical graphs in terms of size, branching and cyclicity. *SAR QSAR Environ. Res.* **2006**, *17*, 429–450.

71. Finn, J.T. Measures of ecosystem structure and function derived from analysis of flows. *J. Theor. Biol.* **1976**, *56*, 363–380.

72. Garrido, A. Symmetry of complex networks. *Adv. Model. Optim.* **2009**, *11*, 615–624.

73. Li, X.; Li, Z.; Wang, L. The inverse problems for some topological indices in combinatorial chemistry. *J. Comput. Biol.* **2003**, *10*, 47–55.

74. Bonchev, D.; Mekenyan, O.; Trinajstić, N. Isomer discrimination by topological information approach. *J. Comput. Chem.* **1981**, *2*, 127–148.

75. Diudea, M.V.; Ilić, A.; Varmuza, K.; Dehmer, M. Network analysis using a novel highly discriminating topological index. *Complexity* **2011**, *16*, 32–39.

76. Konstantinova, E.V.; Paleev, A.A. Sensitivity of topological indices of polycyclic graphs. *Vychisl. Sist.* **1990**, *136*, 38–48.