Information Sciences 181 (2011) 57–78

Contents lists available at ScienceDirect

Information Sciences

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

A history of graph entropy measures

Matthias Dehmer [a,b,*], Abbe Mowshowitz [c]

[a] Institute for Bioinformatics and Translational Research, UMIT, Eduard Wallnoefer Zentrum 1, A-6060, Hall in Tyrol, Austria
[b] Institute of Discrete Mathematics and Geometry, Vienna University of Technology, Wiedner Hauptstrasse 8-10, A-1040 Vienna, Austria
[c] Department of Computer Science, The City College of New York (CUNY), 138th Street at Convent Avenue, New York, NY 10031, USA

ARTICLE INFO

Article history:
Received 16 February 2010
Received in revised form 5 July 2010
Accepted 23 August 2010

Keywords:
Graphs
Information theory
Information measures
Information inequalities
Entropy
Graph entropy
Graph complexity
Structural complexity

ABSTRACT

This survey seeks to describe methods for measuring the entropy of graphs and to demonstrate the wide applicability of entropy measures. Setting the scene with a review of classical measures for determining the structural information content of graphs, we discuss graph entropy measures which play an important role in a variety of problem areas, including biology, chemistry, and sociology. In addition, we examine relationships between selected entropy measures, illustrating differences quantitatively with concrete examples.

1. Introduction

A variety of problems in, e.g., discrete mathematics, computer science, information theory, statistics, chemistry, biology etc. deal with investigating entropies of relational structures. Thus it is not surprising to find variation in the way researchers define the term 'graph entropy'. For example, graph entropy has been used extensively to characterize the structure of graph-based systems in mathematical chemistry [11]. In these applications the entropy of a graph is interpreted as its structural information content and serves as a complexity measure. Such a measure is associated with an equivalence relation defined on a finite graph. The partition induced by the equivalence relation allows for defining a probability distribution [11,69,77,86]. Applying Shannon's entropy formula [79] with the probability distribution one obtains a numerical value that serves as an index of the structural feature captured by the equivalence relation. In particular, with X representing a graph invariant and α an equivalence relation that partitions X into k subsets of cardinality $|X_i|$, a measure $\bar{I}(G,\alpha)$ may be defined as follows:

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

$$\bar{I}(G,\alpha) = -\sum_{i=1}^{k}P_i\log(P_i) = -\sum_{i=1}^{k}\frac{|X_i|}{|X|}\log\left(\frac{|X_i|}{|X|}\right). \tag{2}$$

* Corresponding author at: Institute for Bioinformatics and Translational Research, UMIT, Eduard Wallnoefer Zentrum 1, A-6060, Hall in Tyrol, Austria. Tel./fax: +43 (0) 50 8648 3851.
E-mail addresses: matthias.dehmer@umit.at, mdehmer@geometrie.tuwien.ac.at (M. Dehmer), abbe@cs.ccny.cuny.edu (A. Mowshowitz).

0020-0255/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.ins.2010.08.041

Rashevsky [77], Trucco [86], Mowshowitz [66–69] were the first researchers to define and investigate the entropy of graphs. After this seminal work, Körner [55] introduced a different definition of graph entropy closely linked to problems in information and coding theory. The underlying problem was to determine the performance of a best possible encoding of messages emitted by an information source where the symbols belong to a finite vertex set V. Another definition of Körner's entropy that first appeared in [24] is based on the so-called stable set problem that is strongly related to minimum entropy colorings of graphs [80,81].

As the foregoing discussion suggests, there are competing notions of graph entropy. In fact, there may be no 'right' one, since what may be useful in one domain may not be serviceable in another. Since graph entropy in one form or another plays an important role in a variety of problem areas, a broad survey of the concept is warranted. Two survey papers [80,81] have already appeared, but these are narrowly focused on aspects and properties of Körner's entropy measures. We aim to provide a broad overview of the most well-known graph entropy measures that have been defined and applied thus far. Apart from reviewing classical definitions [77,86,69,66–68,55], we will discuss graph entropy measures which have been applied in such diverse fields as computer science, sociology, chemistry, and biology.

1.1. Outline of the survey

We start our survey on graph entropy by providing some mathematical preliminaries in Section 2.1. This is followed by a discussion of classical measures for determining the structural information content of graphs in Section 2.2. Entropy measures on graphs designed to characterize chemical structures are described in Section 2.3. In contrast to measures based on metrical properties of graphs (see Section 2.3.3) for describing graphs by their structural information content, local entropies are defined in Section 2.3.4. Section 2.4 discusses entropy measures to analyze social network structures. By assigning probability values to each individual vertex in a graph using certain information functions, families of entropy measures can be obtained. This approach is presented in Section 2.5. Then, Section 2.5.1 shows concrete examples and Section 2.5.2 presents information measures based on graph decompositions. The latter approach leads to entropy measures on hierarchical graphs making use of 'natural' vertex partitions. Section 2.5.3, focuses on so-called information inequalities for graphs. In this section, we demonstrate relations between entropy measures on graphs. In order to show that different entropy measure capture structural information differently, we present some numerical results in Section 3. To identify further areas dealing with graph entropy measures, we shed light on some examples in Section 4. Section 5 summarizes the survey and offers concluding remarks.

2. Entropy measures on graphs

2.1. Preliminaries

We begin with some basic definitions drawn from [23,46,45,82]. Note that all the graphs discussed in this paper are assumed to be connected.

Definition 2.1. $G = (V, E), |V| < \infty,\ E \subseteq \binom{V}{2}$ is called a finite undirected graph. If $G = (V,E)$, $|V| < \infty$, and $E \subseteq V \times V$, then G is called a finite directed graph. $\mathcal{G}_{UC}$ denotes the set of finite undirected graphs.

Definition 2.2. A tree is a connected, acyclic undirected graph. A tree $T = (V,E)$ with a distinguished vertex $r \in V$ is a rooted tree. r is called the root of the tree. The level of a vertex v in a rooted tree T equals the length of the path from r to v. The maximum path length d from the root r to any vertex in the tree is called the height of T. A leaf is a vertex incident to exactly one edge in a tree.

We now state the definition of an undirected generalized tree [40] that extends the concept of an ordinary rooted tree. We remark that directed generalized trees have been introduced in [33,63].

Definition 2.3. Let $T = (V,E_1)$ be an undirected finite rooted tree, and let $|L|$ denote the cardinality of the level set $L := \{l_0, l_1, \ldots, l_d\}$. The maximum length of a path in T is denoted by d. $\mathcal{L} : V \to L$ is a surjective mapping and it is called a multi level function if it assigns to each vertex an element of the level set L. Clearly, $d = |L| - 1$. A graph $H = (V,E_{GT})$ is called a finite, undirected generalized tree if its edge set can be represented by the union $E_{GT} := E_1 \cup E_2 \cup E_3$, where

- E_1 forms the edge set of the underlying undirected rooted tree T.
- E_2 denotes the set of horizontal across-edges, i.e., an edge whose incident vertices are at the same level i.
- E_3 denotes the set of edges whose incident vertices are at different levels.

Fig. 1 shows an undirected rooted tree as well as an undirected generalized tree. Note that special undirected generalized trees (see Definition 2.3) will be used to decompose an undirected graph and to define graph entropy measures in Section 2.5.2.

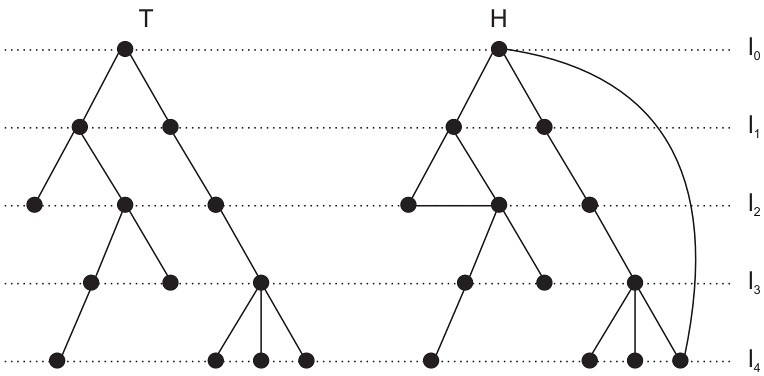

Fig. 1. Left: T represents an ordinary undirected rooted tree. Right: An undirected generalized tree H whose underlying rooted tree equals T.

Definition 2.4. The quantity $\delta(v)$ is called the degree of a vertex $v \in V$ where $\delta(v)$ equals the number of edges $e \in E$ which are incident with v.

Definition 2.5. We call a graph k-regular iff $\delta(v) = k \forall v \in V$. $\mathcal{G}_k$ denotes the set of finite k-regular graphs.

Definition 2.6. $d(u, v)$ denotes the distance between $u \in V$ and $v \in V$ expressed as the minimum length of a path between u, v. Note that $d(u, v)$ is an integer metric. We call the quantity $\sigma(v) = \max_{u \in V} d(u, v)$ the eccentricity of $v \in V$. $\rho(G) = \max_{v \in V} \sigma(v)$ is called the diameter of G.

Definition 2.7. The j-sphere of a vertex v_i in G is defined by the set

$$S_j(v_i, G) := \{v \in V | d(v_i, v) = j, \ j \geqslant 1\}. \tag{3}$$

Definition 2.8. Let X be a discrete random variable by using alphabet A, and $p(x_i) = \Pr(X = x_i)$ the probability mass function of X. The mean entropy of X is then defined by

$$H(X) := -\sum_{x_i \in A} p(x_i) \log(p(x_i)), \quad x_i \in A. \tag{4}$$

Remark 2.1. Throughout this paper, logarithms are always taken to the base two.

2.2. The first entropy measures for graphs

The concept of graph entropy introduced by Rashevsky [77] and Trucco [86] was used to measure structural complexity. Several graph invariants such as the number of vertices, the vertex degree sequence, and extended degree sequences (i.e., second neighbor, third neighbor etc.) have been used in the construction of entropy-based measures. These information measures for graphs have been defined by Rashevsky [77]

$${}^{V}I(G) := |V| \log(|V|) - \sum_{i=1}^{k} N_i \log(N_i), \tag{5}$$

$${}^{V}\bar{I}(G) := -\sum_{i=1}^{k} \frac{|N_i|}{|V|} \log\left(\frac{|N_i|}{|V|}\right). \tag{6}$$

Note that the entropy measure represented by Eq. (6) was originally called the topological information content [77] of a graph G. According to Rashevsky [77], $|N_i|$ denotes the number of topologically equivalent vertices in the ith vertex orbit of G, where k is the number of different orbits. Vertices are considered as topologically equivalent if they belong to the same orbit of a graph G. By applying this principle to the edge automorphism group, Trucco [86] introduced similar entropy measures

$${}^{E}I(G) := |E| \log(|E|) - \sum_{i=1}^{k} N_i^E \log(N_i^E), \tag{7}$$

$${}^{E}\bar{I}(G) := -\sum_{i=1}^{k} \frac{|N_i^E|}{|E|} \log\left(\frac{|N_i^E|}{|E|}\right). \tag{8}$$

Correspondingly, $|N_i^E|$ stands for the number of edges which belong to the *i*-th edge orbit [11,85] of *G*.

Mowshowitz [66–69] explored the properties of structural information measures relative to two different equivalence relations defined on the vertices of a graph. One is the measure $I_a(G)$, based on Rashevsky's topological information content, that computes the entropy of a graph relative to the vertex partition induced by the automorphism group. The other $I_c(G)$ is defined relative to a chromatic decomposition of the vertices. The automorphism-based measure captures the symmetry structure of a graph. At one extreme there is the complete graph which has the full symmetric group and thus zero information content; at the other one finds the graph whose group consists of the identity alone and thus has the maximum log (|*V*|) information content. Note that to compute the information content of an arbitrary graph is non-trivial. In an effort to establish properties of entropy-based measures and to simplify computation, Mowshowitz [69] examined the action of the measures on various graph operations and products. For the automorphism-based measure the following results were obtained.

Theorem 2.2. *Let* G_i $(1 \leqslant i \leqslant n)$ *be isomorphic to G. Then*

$$I_a(G_1 \cup G_2 \cup \cdots \cup G_n) = I_a(G) \tag{9}$$

and

$$I_a(G_1 + G_2 + \cdots + G_n) = I_a(G), \tag{10}$$

where $\cup$ *denotes the sum and + the join operation on graphs.*

These equations show that information content is unaffected by the obvious kinds of repetition. Further, it can be shown that the information measure is semi-additive on the cartesian product and on the composition of two graphs.

Theorem 2.3. *For graphs G and H*

$$I_a(G \times H) \leqslant I_a(G) + I_a(H) \tag{11}$$

and

$$I_a(G \circ H) \leqslant I_a(G) + I_a(H), \tag{12}$$

where $\times$ *and* $\circ$ *represent the cartesian product and composition, respectively.*

In the case of cartesian product, a sufficient condition for equality is that *G* and *H* are relatively prime with respect to the product. A somewhat more complicated sufficient condition for equality holds in the case of composition. These results are generalized for arbitrary 'well-behaved' product operations on graphs. In particular, the information content of such a product graph is shown to be the sum of the information contents of the respective components minus a certain conditional entropy defined relative to the cartesian products of the respective orbits of the component graphs. This result can be used to derive the information content of graphs such as the hypercube that are defined in terms of product operations. The automorphism-based entropy measure applies to directed (digraphs) as well as to undirected graphs [66]. Binary operations extend to digraphs in a natural way, and the information measure has properties that are analogous to those mentioned above for binary operations on graphs. Entropy values derived from partitions of integers can be realized as the information content of some digraph the cardinalities of whose orbits correspond to the elements of the partition. It is also possible to extend the information measure to an infinite graph defined as the limit of a sequence of finite graphs. This extension is quite useful in demonstrating the dependence of the information measure on a particular structural feature of a graph. As a measure of complexity, entropy, computed on a partition of graph elements, is relative to the structural feature that induces the partition.

Computing the automorphism-based information measure for a particular graph entails determining the respective cardinalities of the orbits of the graph's automorphism group. An obvious way of doing this is to construct the automorphism group and then determine the number of elements in each orbit explicitly. The adjacency matrix of a graph can be used to simplify the construction. This follows from the fact that a permutation of the vertices of a graph is an automorphism if and only if the corresponding permutation matrix commutes with the graph's adjacency matrix. The automorphisms can be computed with the aid of a canonical form for the adjacency matrix together with a transforming matrix. This approach does not yield an efficient algorithm in general, but is useful in special cases [67]. As an example, Fig. 2 shows the values of I_a and I_c for a cycle and an identity graph. For G_1, we easily obtain $I_a(G_1) = \log(1) = 0$ and $I_c(G_1) = \log(2) = 1$. Further, we calculate $I_a(G_2) = \log(6)$ and $I_c(G_2) = \frac{1}{6}\log(6) + \frac{1}{3}\log(3) + \frac{1}{2}\log(2)$.

Theorem 2.4. *Let G and H be n-vertex digraphs with A = A (G) and B = A(H). Suppose the elementary divisors of both A and B are co-prime in pairs, and AB = BA. Then* $I_a(G) = I_a(H)$.

This result holds for undirected graphs *G* and *H* whose adjacency matrices *A* and *B* have distinct eigenvalues. Graph colorings (or independent sets) offer a structural feature, quite different from automorphisms, on which to base an entropy measure. The following definition is from [68].

Definition 2.9. Let *G* be a graph with |*V*| vertices, and let

$$\widehat{V} = \{V_i | 1 \leqslant i \leqslant h\}, \quad |V_i| = n_i(\widehat{V}) \tag{13}$$

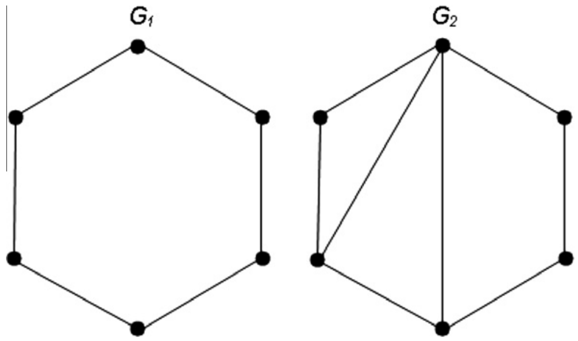

Fig. 2. Left: The cycle graph G_1. Right: An identity graph G_2.

be an arbitrary chromatic decomposition of G where $h = \chi(G)$ is the chromatic number of G. Then the chromatic information content $I_c(G)$ of G is given by

$$I_c(G) = \min_{\widehat{V}} \left\{ -\sum_{i=1}^{h} \frac{n_i(\widehat{V})}{|V|} \log\left(\frac{n_i(\widehat{V})}{|V|}\right) \right\}. \tag{14}$$

Roughly speaking, the chromatic information content of a graph is inversely related to the number of vertices in a maximally independent set. Note that $I_c(G)$ does not necessarily give the minimum over all possible chromatic decompositions of an arbitrary graph G. However, if G does not have a complete k-coloring for $k \geqslant \chi(G)$ then $I_c(G)$ is indeed the minimum. In general, $I_c(G) \leqslant \log(\chi(G))$ and $I_c(G) \leqslant \log(d_0 + 1)$ where d_0 is the maximum degree of any vertex in G. Additional upper bonds are given in [68].

Also, graph operations such as cartesian product and composition are useful for the automorphism-based measure because the orbits of the combined graphs are closely related to the cartesian products of the orbits of the component graphs. However, these products do not seem useful for studying the chromatic-based measure. Other graph operations such as the Kronecker product appear to be more useful [68]. Further, the radical difference between the measures $I_a(G)$ and $I_c(G)$ can be illustrated with respect to trees. Since the chromatic number of a tree is 2, $I_c(G) \leqslant 1$. However, for every integer $n \geqslant 7$ there exists an identity tree on n vertices, which means $I_a(G) = \log(n)$. Thus, the numerical difference between the two measures is unbounded. Clearly, the complexity of a graph as measured by an entropy function is relative to the structural feature that gives rise to the partition used in the entropy computation.

2.3. Entropy measures in biology and chemistry

Shannon's seminal work [79] in the late nineteen-forties marks the starting point of modern information theory. Following early applications in linguistics and electrical engineering, information theory was applied extensively in biology and chemistry, see, e.g., [65,73,77]. Here, the main novelty was the idea of considering a structure as an outcome of an arbitrary communication [13]. With the aid of this insight, Shannon's entropy formulas [79] were used to determine the structural information content of a network [66–69,77,86]. As a result, this method has been used for exploring living systems, e.g., biological and chemical systems by means of graphs. These applications are closely related to the work of Rashevsky [77] and Trucco [86] discussed in Section 2.2. In what follows, we review in chronological order graph entropy measures that have been used for studying biological and chemical networks.

2.3.1. Classical measures for detecting molecular complexity

As noted in Section 1, complexity measures defined on graphs [69,75,76] can be obtained by applying Shannon's entropy formula based on partitions induced by structural characteristics a graph. Hence, each measure outlined in Section 2.2 can be used for measuring molecular complexity. In particular, Bertz [8] developed an extension of Rashevsky's measure designed to analyze molecular structures. A known weak point of a measure like the one in Eq. 5 is that it does not properly reflect the number of the invariants used because one obtains $I(G) = 0$ when all invariants are equal [8]. This holds independently of the size of the graph. To overcome this problem, Bertz chose as graph invariant the number of two-edge subgraphs and added the term $|V|\log(|V|)$ resulting in

$${}^{V}C(G) = 2|V|\log(|V|) - \sum_{i=1}^{k} N_i \log(N_i), \tag{15}$$

or more generally

$${}^{X}C(G) = 2|X|\log(|X|) - \sum_{i=1}^{k} N_i \log(N_i). \tag{16}$$

X represents any graph invariant [8]. Other measures of the molecular complexity of graphs can be found in [13,64,75,76].

2.3.2. Entropy measures based on graph decompositions

In 1971, Hosoya [48] introduced the topological index Z to characterize molecular branching [15], namely,

$$Z = \sum_{i=0}^{\lfloor |V|/2 \rfloor} P(G, i), \tag{17}$$

where $P(G,i)$ denotes the number of selections of i mutually non-adjacent edges in G. In the case of acyclic graphs, Z was defined by the sum of the absolute values of the polynomial coefficients $p(G,k)$ of the characteristic polynomial

$$P(G, x) = \sum_{k=0}^{s} (-1)^k p(G, k) x^{|V|-2k}. \tag{18}$$

Here, s represents the largest number of mutually non-incident edges in the acyclic graph, see also [87]. Usually, the characteristic polynomial can be calculated from its adjacency matrix. Now, one can argue that applying the Hosoya-Index to a graph G induces a decomposition of a graph [11]. By using the definition of Z, Bonchev and Trinajstić [15] defined information contents for polynomial coefficients of the characteristic polynomial of a graph G by

$$I_{pc}(G) = Z \log(Z) - \sum_{k=0}^{[|V|/2]} p(G, k) \log(p(G, k)), \tag{19}$$

$$\bar{I}_{pc}(G) = -\sum_{k=0}^{[|V|/2]} \frac{p(G, k)}{Z} \log\left(\frac{p(G, k)}{Z}\right). \tag{20}$$

In [15], numerical results are given that compare Z, $I_{pc}(G), \bar{I}_{pc}(G)$, and other entropy measures on graphs. Additional entropy measures based on graph decompositions are discussed in Section 2.5.

2.3.3. Entropy measures based on metrical properties of graphs

The measures discussed in the preceding sections are based on classical graph invariants, e.g., number of vertices, edges, connections, etc. As stated in [13], a limitation of the resulting classical measures is that structurally non-equivalent graphs may have the same information content. For example, two non-isomorphic graphs can have the same information content using the measure defined in Eq. (2). In mathematical chemistry, this problem deals with evaluating the degree of the so-called degeneracy [15,85] of a topological index. An index, i.e., a graph complexity measure is called degenerate if the index possesses the same value for more than one structure. In order to overcome this problem, Bonchev and Trinajstić [15] developed a variety of so-called magnitude-based graph entropy measures which are based on weighted probability distributions taking several structural graph features into account, e.g., distances and vertex degrees etc.

In this section, we discuss only the most well-known entropy measures based on graph distances, starting with the work of Bonchev and Trinajstić [15]. The basis of the following information measure is the distance matrix

$$D = (d(v_i, v_j))_{ij}, \quad 1 \leqslant i \leqslant |V|, \quad 1 \leqslant j \leqslant |V|. \tag{21}$$

By defining the probability values $p_0 = \frac{1}{|V|}$ and $p_i = \frac{2k_i}{|V|^2}$, Bonchev and Trinajstić obtained [15]

$$I_D(G) = |V|^2 \log(|V|^2) - |V| \log(|V|) - \sum_{i=1}^{\rho(G)} 2k_i \log(2k_i), \tag{22}$$

$$\bar{I}_D(G) = -\frac{1}{|V|} \log\left(\frac{1}{|V|}\right) - \sum_{i=1}^{\rho(G)} \frac{2k_i}{|V|^2} \log\left(\frac{2k_i}{|V|^2}\right). \tag{23}$$

Here, a value i in the distance matrix D appears $2k_i$ times. As a result, it turns out that these measures are more sensitive than other classical topological indices used in mathematical chemistry [15]. Yet another pair of entropy measures for graphs has been defined [15] as

$$I_t^W(G) = W(G) \log(W(G)) - \sum_{i=1}^{\rho(G)} i k_i \log(i), \tag{24}$$

$$\bar{I}_D^W(G) = -\sum_{i=1}^{\rho(G)} \frac{i k_i}{W(G)} \log\left(\frac{i}{W(G)}\right), \tag{25}$$

where $W(G)$ is called the Wiener-Index [92],

$$W(G) = \sum_{i=1}^{\rho(G)} i k_i. \tag{26}$$

As noted earlier, the strength of these information measures lies in their high discrimination power [15,54] evidenced by empirical tests using appropriate data sets. Another distance-based entropy measure was developed by Balaban and Balaban

[5]. The definitions shown below are designed to compensate for the fact that information measures defined for graphs may be highly degenerate. Balaban and Balaban [5] first defined the mean information on the magnitude of distances for each vertex v_i as

$$u(v_i) = -\sum_{j=1}^{\sigma(v_i)} \frac{jg_j}{d(v_i)} \log\left(\frac{j}{d(v_i)}\right). \tag{27}$$

Moreover,

$$d(v_i) = \sum_{j=1}^{|V|} d(v_i, v_j) = \sum_{j=1}^{\sigma(v_i)} jg_j, \tag{28}$$

where g_j indicates the number of vertices whose distance from v_i is j. Additionally, the local information on the magnitude of distances is defined as

$$w(v_i) = d(v_i)\log(d(v_i)) - u(v_i). \tag{29}$$

Finally, applying Randiĉ's formula [74] one obtains

$$U_1(G) = \frac{|E|}{\mu+1} \sum_{(v_i,v_j)\in E} [u(v_i)u(v_j)]^{-\frac{1}{2}}, \tag{30}$$

$$U_2(G) = \frac{|E|}{\mu+1} \sum_{(v_i,v_j)\in E} [w(v_i)w(v_j)]^{-\frac{1}{2}}, \tag{31}$$

where μ denotes the cyclomatic number defined by $\mu := |E| + 1 - |V|$, see [5].

2.3.4. Local entropy measures based on metrical properties

The entropy measures presented thus far have been designed to characterize a graph G by determining its global information content. However, it is also useful to define information measures on local features or substructures of a graph. For example, one can define an entropy measure for each vertex of a graph. Such a measure can be interpreted as a kind of vertex complexity [78] that here depends on the distances to the remaining vertices in the graph. These kind of measures have been developed by, e.g., Konstantinova and Paleev [53], Raychaudhury et al. [78] and Balaban and Balaban [5], see Section 2.3.3. For example, the following entropy measure [53]

$$I_D(v_i) := -\sum_{j=1}^{|V|} \frac{d(v_i, v_j)}{d(v_i)} \log\left(\frac{d(v_i, v_j)}{d(v_i)}\right) \tag{32}$$

represents the information distance of the vertex $v_i \in V$. Correspondingly, the entropy of G is defined by summing up the information distances for each vertex,

$$I_D^{\star}(G) := \sum_{i=1}^{|V|} I_D(v_i). \tag{33}$$

By applying the same principle to the matrix

$$S = (|S_j(v_i, G)|)_{ij}, \quad i = 1, \ldots, |V|,\ j = 1, \ldots, \rho(G) \tag{34}$$

of j-sphere cardinalities of a graph G, Konstantinova and Paleev [53] also obtained

$$I_S(v_i) = -\sum_{j=0}^{\sigma(v_i)} \frac{|S_j(v_i, G)|}{|V|} \log\left(\frac{|S_j(v_i, G)|}{|V|}\right), \tag{35}$$

$$I_S^{\star}(G) = \sum_{i=1}^{|V|} I_S(v_i). \tag{36}$$

These measures discriminate between graphs by means of their additive entropies. Recent work by Dehmer and Emmert-Streib [32] on local entropy for graphs has led to the construction of parametric information measures,

$$I_{g_\mu}(v_i) := -\sum_{j=1}^{|V|} \frac{g_\mu^j(v_i)}{\sum_{j=1}^{|V|} g_\mu^j(v_i)} \log\left(\frac{g_\mu^j(v_i)}{\sum_{j=1}^{|V|} g_\mu^j(v_i)}\right), \tag{37}$$

where

$$g_1^j(v_i) := d(v_i, v_j), \quad 1 \leqslant i \leqslant |V|, \tag{38}$$

$$g_2^j(v_i) := c_j d(v_i, v_j), \quad 1 \leqslant i \leqslant |V|, \quad c_i > 0. \tag{39}$$

$I_{g_\mu}(v_i)$ is a local vertex entropy. By setting $c_j = 1$ in $g_2^j(v_i)$, we get Eq. (32) as special case. Finally, the entropy of G can also be defined by

$$I_{g_\mu}(G) := \frac{\sum_{i=1}^{|V|} I_{g_\mu}(v_i)}{|V|}. \tag{40}$$

These measures represent families of local entropies. In particular, we define special information measures by choosing concrete coefficients [35]:

$$I_{loc}^1(G) := \frac{\sum_{i=1}^{|V|} I_{g_1}(v_i)}{|V|}, \tag{41}$$

$$I_{loc}^2(G) := \frac{\sum_{i=1}^{|V|} I_{g_2}(v_i)}{|V|}, \tag{42}$$

where the coefficients are linearly decreasing, e.g.,

$$c_1 := \rho(G), \quad c_2 := \rho(G) - 1, \ldots, c_{\rho(G)} := 1. \tag{43}$$

Finally, I_{loc}^3 is also defined by using Eq. (42) where the coefficients are exponentially decreasing, e.g.,

$$c_1 := \rho(G), \quad c_2 := \rho(G)e^{-1}, \ldots, c_{\rho(G)} := \rho(G)e^{-\rho(G)+1}. \tag{44}$$

Other local entropies can be found in [85].

2.4. Entropy measures in sociology and psychology

Structural approaches play an important role in the analysis of social networks and in mathematical psychology [60,91]. One area of continuing interest in the theory of social networks centers on measuring the complexity of social networks [19]. However, most of the contributions are related to non-information-theoretic measures, e.g., see [19,22,44,58]. In contrast, relatively little work has been done on measuring complexity by using entropy measures [19]. In this section, we review the few existing contributions that apply entropy measures to social network analysis. As noted above, an important application area focuses on measuring structural complexity of such networks. For investigating this problem, Everett [43] applied the entropy measure developed by Mowshowitz [69] for detecting so-called role complexity. This notion of role is used to analyze relations between individuals in social networks. Based on numerical results obtained from sample graphs, it turned out that the non-information-theoretic complexity measure

$$R_C(G) := 1 - \frac{|\mathrm{Aut(G)}|}{|V|!}, \tag{45}$$

could reflect role complexity more meaningfully than an entropy measure. One can see that Eq. (45) is related to the number of positions spanned by the vertex orbits [43]. Another definition of an entropy-based measure designed to capture behavioral diversity among robots was developed by Balch [6]. In order to quantify robot team diversity [6], the concept of 'hierarchical social entropy' was introduced based on entropies of simple partitions induced by a hierarchical clustering among robots. The connection with topological graph entropy can be explained as follows. By performing agglomerative clustering [49], a tree-like structure called a 'dendogram' is created. Hence, for each resulting partition $\mathcal{P}_i$ of the dendogram a probability distribution $p_1, \ldots, p_k$ is defined, where

$$p_\mu = \frac{|c_\mu|}{\sum_{j=1}^{k} |c_j|}. \tag{46}$$

$|c_\mu|$ denotes the cardinality of the collection of robots in the μth subset on $\mathcal{P}_i$. Hence, the entropy of $\mathcal{P}_i$ can be expressed by

$$I(\mathcal{P}_j) = -\sum_{j=1}^{k} p_j \log(p_j). \tag{47}$$

This entropy measure characterizes a partition of a tree-like graph structurally representing a dendogram obtained from a clustering process [6]. Fig. 3 provides an example of such a dendogram with its partitions. To conclude this section, we briefly mention an information-theoretic approach developed by Tutzauer [88] for detecting the centrality of vertices by means of transfer and flows along paths. This method was applied to analyze vertex centrality in networks representing gangs. Based on flow probabilities p_{v_i,v_j} starting at v_i and ending at v_i, the path-transfer centrality of v_i is [88]

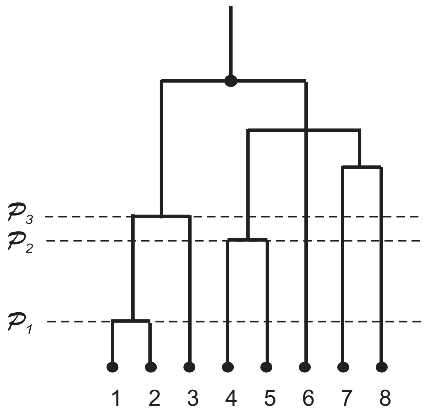

Fig. 3. A tree-like graph resulting from agglomerative clustering.

$$I_C(v_i) = -\sum_{j=1}^{|V|} p_{v_i,v_j} \log(p_{v_i,v_j}). \tag{48}$$

To explore phenomena in special social networks as well as traffic networks, entropy measures based on link distributions have also been introduced [50,59,94].

2.5. More recent entropy measures for graphs

The entropy measures presented in Section 2.3 are based on grouping the elements (partitions) of a given graph invariant X using an equivalence criterion. Examples for concrete invariants are vertices, edges, degrees, and distances in a graph, see [11]. In this section, we outline a different approach for deriving graph entropy measures recently developed in [29,31]. The main idea can be summarized as follows. Instead of inducing partitions and determining their probabilities, we assign a probability value to each individual vertex in a graph. One way to do this is by means of certain information functions which capture structural features of a graph. This procedure avoids the problem of determining partitionings associated with an equivalence relation. Several parametric information functions based on metrical properties of graphs have been defined in [28]. A notable feature of this approach is that the resulting graph entropy measures can be used for solving machine learning problems because the existing parameters can be learned by using appropriate data sets. In what follows, we present salient definitions and results [28,29].

2.5.1. Parametric graph entropy measures

Definition 2.10. Let G be an arbitrary finite graph and let S be a given set, e.g., a set of vertices or paths etc. Functions f of the form $f : S \to \mathbb{R}_+$ play a role in defining information measures on graphs, so we call them abstract information functions of G.

Definition 2.11. Let f be an abstract information function of G. Then

$$p^f(v_i) := \frac{f(v_i)}{\sum_{j=1}^{|V|} f(v_j)}. \tag{49}$$

Obviously,

$$p^f(v_1) + p^f(v_2) + \cdots + p^f(v_{|V|}) = 1. \tag{50}$$

Hence, $(p^f(v_1),\ldots,p^f(v_{|V|})$ forms a probability distribution.

Definition 2.12. Let G be a finite arbitrary graph and let f be an abstract information function.

$$I_f(G) := -\sum_{i=1}^{|V|} \frac{f(v_i)}{\sum_{j=1}^{|V|} f(v_j)} \log\left(\frac{f(v_i)}{\sum_{j=1}^{|V|} f(v_j)}\right), \tag{51}$$

$$I_f^{\lambda}(G) := \lambda\left(\log(|V|) + \sum_{i=1}^{|V|} \frac{f(v_i)}{\sum_{j=1}^{|V|} f(v_j)} \log\left(\frac{f(v_i)}{\sum_{j=1}^{|V|} f(v_j)}\right)\right), \tag{52}$$

are families of information measures representing the structural information content of G. $\lambda > 0$ is a scaling constant. I_f is the entropy of G and I_f^λ its information distance between maximum entropy and I_f.

The meaning of I_f and I_f^λ has been investigated by calculating the information content of real and synthetic chemical structures, see [35]. Also, the information measures were calculated using specific graph classes to study extremal values and, hence, to detect the kind of structural information captured by the measures. Clearly Eqs. (51), (52) represent families of information measures to index the information content of graphs. Special entropy measures for graphs can be obtained by specifying particular information functions. For instance, such functions (first defined in [29]) are based on j-sphere cardinalities and can be expressed as exponential and linear functions.

Definition 2.13. Let $G = (V, E) \in \mathcal{G}_{UC}$. For a vertex $v_i \in V$, we define f^{V_j}, $j = 1,2$ as

$$f^{V_1}(v_i) := \alpha^{c_1|S_1(v_i,G)|+c_2|S_2(v_i,G)|+\cdots+c_{\rho(G)}|S_{\rho(G)}(v_i,G)|}, \quad c_k > 0,\ 1 \leqslant k \leqslant \rho(G),\ \alpha > 0, \tag{53}$$

$$f^{V_2}(v_i) := c_1|S_1(v_i,G)| + c_2|S_2(v_i,G)| + \cdots + c_{\rho(G)}|S_{\rho(G)}(v_i,G)|, \quad c_k > 0,\ 1 \leqslant k \leqslant \rho(G). \tag{54}$$

Applying Definition 2.12, we obtain the corresponding entropies.

Definition 2.14. Let $G = (V, E) \in \mathcal{G}_{UC}$.

$$I_{f^{V_j}}(G) := -\sum_{i=1}^{|V|} p^{f^{V_j}}(v_i) \log\left(p^{f^{V_j}}(v_i)\right), \tag{55}$$

$$I^\lambda_{f^{V_j}}(G) := \lambda\left(\log(|V|) + \sum_{i=1}^{|V|} p^{f^{V_j}}(v_i) \log\left(p^{f^{V_j}}(v_i)\right)\right). \tag{56}$$

The fact that the underlying information functions as well as the resulting entropies are parametric gives us the possibility of weighting structural differences or characteristics of a graph. In particular, the c_k must be chosen such that not all values are equal, e.g.,

$$c_1 > c_2 > \cdots > c_\rho \quad \text{or } c_1 < c_2 < \cdots < c_\rho. \tag{57}$$

Numerical examples of the computation of $I_{f^{V_1}}$ for chemical graphs are presented in [31]. Note that $I_{f^{V_1}}$ is shown to reflect the structural complexity of graphs in a chemically meaningfully way [31]. More complex information functions depend on the 'local information graph' of $G = (V, E) \in \mathcal{G}_{UC}$ for a vertex v_i. This concept [28] is based mainly on the j-sphere and the observation that in many real world networks, information is distributed via shortest paths from a given vertex.

Definition 2.15. Let $G = (V, E) \in \mathcal{G}_{UC}$ and $v \in V$. Suppose further that $S_j(v, G) = \{u_1, u_2, \ldots, u_k\}$. For each $u_t \in S_j(v, G)$, $1 \leqslant t \leqslant k$, let $P_t^j(v)$ denote the path of length j from v to u_t. $P_t^j(v) = (v, w_1, w_2, \ldots, w_j)$ where $w_j = u_t$. The set of edges E_t^j in the *path is given by*

$$E_t^j = \{\{v, w_1\}, \{w_1, w_2\}, \ldots, \{w_{j-1}, w_j\}\} \tag{58}$$

and the set of vertices V_t^j in the *path is given by* $V_t^j = \{v, w_1, w_2, \ldots, w_j\}$. *Let*

$$V^j_{L_G} := V^j_1 \cup V^j_2 \cdots \cup V^j_k \tag{59}$$

and

$$E^j_{L_G} := E^j_1 \cup E^j_2 \cup \cdots \cup E^j_k. \tag{60}$$

The local information graph $L_G(v, j)$ of G regarding v is defined by

$$L_G(v,j) = \left(V^j_{L_G}, E^j_{L_G}\right). \tag{61}$$

An example should help to clarify the construction of the graph entropy. Fig. 4 shows the j-spheres of a graph G as concentric circles of vertices. Fig. 5 illustrates local information graphs $L_G(v_i, 1)$, $L_G(v_i, 2)$, and $L_G(v_i, 3)$. Note that the local information graph for $v_i \in V$ can not always be uniquely defined because there may exist more than one path from v_i to some vertex in the corresponding j-sphere [28]. In such cases, additional constraints are required. Finally, local property measures, e.g., vertex centrality measures [18], allow for using the local information graph to define more complex information functions on graphs and, hence, to obtain novel families of graph entropy measures.

Definition 2.16. Let $G \in \mathcal{G}_{UC}$. For each vertex $v_i \in V$, let $L_G(v_i, j)$ be the local information graph of G. We define $f^{C_j}(v_i)$, $j = 1, 2$ as

$$f^{C_1}(v_i) := \alpha^{c_1\beta^{L_G(v_i,1)}(v_i)+c_2\beta^{L_G(v_i,2)}(v_i)+\cdots+c_{\rho(G)}\beta^{L_G(v_i,\rho(G))}(v_i)}, \quad c_k > 0,\ 1 \leqslant k \leqslant \rho(G),\ \alpha > 0, \tag{62}$$

$$f^{C_2}(v_i) := c_1\beta^{L_G(v_i,1)}(v_i) + c_2\beta^{L_G(v_i,2)}(v_i) + \cdots + c_{\rho(G)}\beta^{L_G(v_i,\rho(G))}(v_i), \quad c_k > 0,\ 1 \leqslant k \leqslant \rho(G). \tag{63}$$

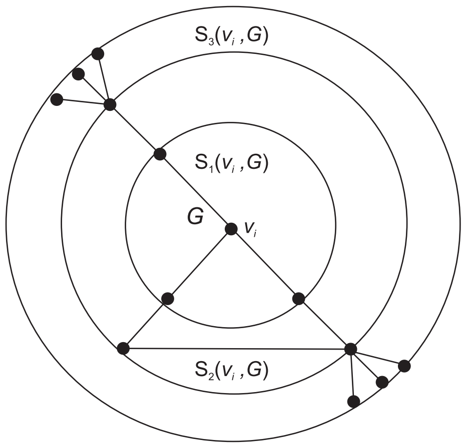

Fig. 4. j-spheres of v_i for j = 1, 2, 3.

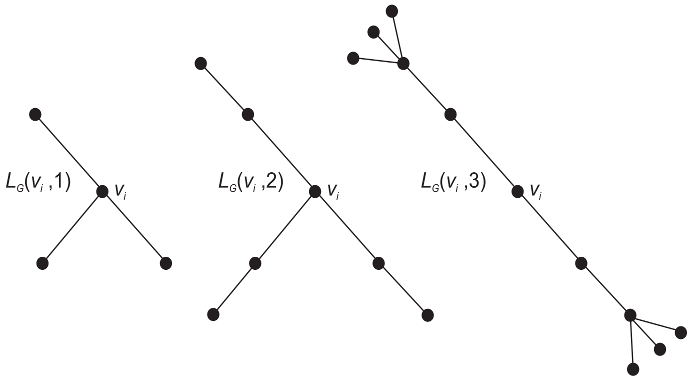

Fig. 5. Local information graphs of G for v_i where j = 1, 2, 3.

β represents an arbitrary vertex centrality measure. c_k are real positive coefficients. The value resulting in application of β to the vertex v_i regarding $L_G(v_i,j)$ is denoted by $\beta^{L_G(v_i,j)}(v_i)$.

Definition 2.17. Let $G = (V,E) \in \mathcal{G}_{UC}$.

$$I_{f^{C_j}}(G) := -\sum_{i=1}^{|V|} p^{f^{C_j}}(v_i)\log(p^{f^{C_j}}(v_i)), \tag{64}$$

$$I^{\lambda}_{f^{C_j}}(G) := \lambda\left(\log(|V|) + \sum_{i=1}^{|V|} p^{f^{C_j}}(v_i)\log(p^{f^{V_j}}(v_i))\right). \tag{65}$$

Remark 2.5. Note that the definition of information functions and the resulting graph entropy measures can be extended to finite directed graphs.

2.5.2. Entropy measures based on graph decompositions

Another type of entropy measure is based on a decomposition of a graph into special subgraphs, see [28]. One such decomposition can be obtained by deriving tree-like structures of given height. Instead of determining the entropy of a graph G, we calculate the entropies of derived hierarchical structures $H_1, \ldots, H_k$ [28]. The key steps of the procedure are as follows:

- Let $G \in \mathcal{G}_{UC}$ be a finite undirected graph.
- Chose a vertex v_i as starting point. Then, derive a generalized tree H_i of height d (see Algorithm 2.1).
- Performing this step for all vertices of G, we obtain the tree set

$$S^H_G := \{H_1, H_2, \ldots, H_{|V|}\}.$$

- Now, the entropy of each H_i can be calculated by using, e.g., Eqs. (66) and (67). At this point, we make use of use of 'natural' vertex partitions (on each level i).
- Finally, the entropy of G can be computed by, e.g., Eqs. (68) and (69).

The idea of determining the topological entropy of hierarchical structures has been elaborated in [39]. In particular, these authors describe a method for computing the entropy of so-called undirected universal graphs. Such graphs are related to generalized trees (see Definition 2.3). An advantage of this measure is that vertex partitions necessary for calculating the entropy in hierarchical graphs are obtained in a natural way. The collection of the respective vertex sets on the different level of a hierarchical graph forms a partition. Since the computation of entropy measures based on vertex partitions may be inefficient, the overall time complexity can be reduced by using the previously mentioned decomposition method. In the following, we present the definitions of entropy measures for undirected generalized trees. These measures, originally given in [28], are similar to those defined in [39].

Definition 2.18. Let H be a generalized tree of height d, and suppose $|V_i|$ denotes the number of vertices on the ith level. A probability distribution associated with H is determined as follows. Let $p_i^V := \frac{|V_i|}{|V|-1}$. Then, the vertex entropy of a generalized tree H is defined by

$$I^V(H) := -\sum_{i=1}^{d} p_i^V \log\left(p_i^V\right). \tag{66}$$

Definition 2.19. Let H be a generalized tree of height, and suppose d. $|E_i|$ denotes the number of edges on the ith level. A probability distribution can be associated with H as follows. Let $p_i^E := \frac{|E_i|}{2|E|-\delta(r)}$. Then, the edge entropy of a generalized tree H is defined by

$$I^E(H) := -\sum_{i=1}^{d} p_i^E \log\left(p_i^E\right). \tag{67}$$

These definitions can be applied directly to calculate the entropies of given generalized trees. However, to apply these measures for determining the entropy of a non-hierarchical undirected graph, an algorithm (such as given below) is needed to decompose such a graph into a set of undirected generalized trees [42].

Algorithm 2.1. A graph $G \in \mathcal{G}_{UC}$ with $|V|$ vertices can be locally decomposed into a set of generalized trees as follows: Assign labels from *1* to $|V|$ to the vertices. We call $L_S = \{1, \ldots, |V|\}$ the label set. Choose a desired height d of the trees, and select an arbitrary label from L_S, e.g. i. The vertex with this label is the root vertex of a tree. Now, perform the following steps:

1. Calculate the shortest distance from vertex i to all other vertices in the graph G, e.g. by the algorithm of Dijkstra [36].
2. The vertices at distance k are the vertices on the kth level of the resulting generalized trees. Select all vertices (together with incident edges) in the graph up to distance d. Paths of length > d are deleted.
3. Delete label i from the label set L_S.
4. Repeat this procedure if L_S is not empty by choosing an arbitrary label from L_S, otherwise terminate.

Fig. 7 shows the outcome of applying Algorithm 2.1 to the graph G of Fig. 6. We see that decomposing G by means of Algorithm 2.1 results in a set of special generalized trees. In general, the cardinality of the generalized tree set obtained equals the number of vertices of the graph being decomposed. From Definition 2.18, we see that this measure attains its maximum if a generalized tree H possesses the same number of vertices on each level i, $1 \leqslant i \leqslant d$. The same follows for the edge entropy defined by Definition 2.19. Note that vertex numbering labels have been omitted (see Figs. 6 and 7) since the entropy measures defined above are invariant under permutations of the vertices on a given level i. Finally, by using the proposed method for decomposing undirected graphs into sets of special generalized trees, we define the structural information content of undirected graphs as follows [28].

Definition 2.20. Let $G \in \mathcal{G}_{UC}$ and $S_G^H := \{H_1, H_2, \ldots, H_{|V|}\}$ be the associated set of generalized trees. We define the structural information content of G by

$$I^V(G) := \sum_{i=1}^{|V|} I^V(H_i) \tag{68}$$

and

$$I^E(G) := \sum_{i=1}^{|V|} I^E(H_i). \tag{69}$$

Parameterized entropy measures based on the generalized tree decomposition discussed here have also been defined in [28].

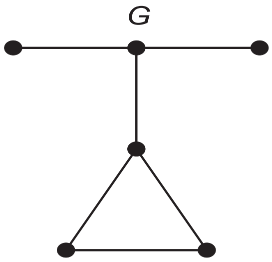

Fig. 6. An undirected graph G.

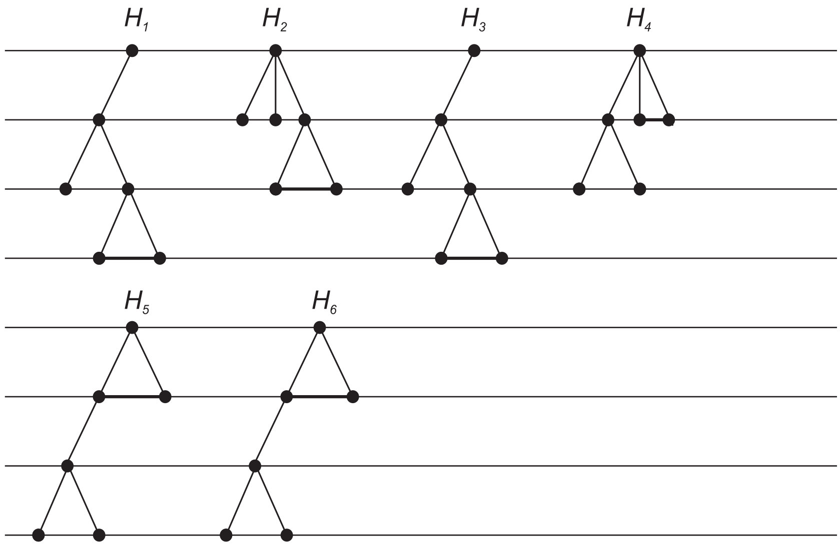

Fig. 7. Decomposing the graph G, see Fig. 6. The decomposed graphs represent special undirected generalized trees.

2.5.3. Information inequalities for graphs

In this section, we sketch some results to derive so-called information inequalities for graphs [30,34]. Generally, information inequalities describe relations between information measures for graphs. A major objective in studying information inequalities is to obtain bounds on the entropies of special classes of graphs. Also, specific information inequalities can be obtained by using different information functions, see [30] In the following, we discuss only so-called implicit information inequalities which can be considered as a special type of an information inequality. We call such an inequality implicit because the entropy of a graph will be estimated by a quantity that contains another graph entropy expression. In the following, we state some implicit information inequalities which have been proven in [34]. We note that the technique shown in [34] aims to develop a general method for proving inequalities between graph entropy measures. For instance, a special application of this approach is to characterize graph classes by using such information inequalities, see, e.g., [30].

Theorem 2.6. *Let f and $f^\star$ be information functions. Let $\mathcal{G}$ be a class of graphs. If*

$$p^f(v) < \psi \cdot p^{f^\star}(v), \quad \forall G \in \mathcal{G}, \tag{70}$$

then

$$I^f(G) + \psi \sum_{i=1}^{|V|} p^{f^\star}(v_i) \cdot \log\left(\psi \cdot p^{f^\star}(v_i) + 1\right) + \sum_{i=1}^{|V|} \log\left(\psi \cdot p^{f^\star}(v_i) + 1\right) > 0, \quad \forall G \in \mathcal{G}. \tag{71}$$

ψ is a constant expression.

Theorem 2.7. *Let f and $f^\star$ be information functions. Let $\mathcal{G}$ be a class of graphs. If*

$$p^f(v) < \psi \cdot p^{f^\star}(v), \quad \forall G \in \mathcal{G} \tag{72}$$

and

$$\psi \cdot p^{f^\star}(v) > \frac{1}{2^c - 1} > 0, \quad c > 0, \tag{73}$$

then

$$I^f(G) > \psi \cdot I^{f^\star}(G) - \psi \cdot \log(\psi) - c \cdot \psi - \sum_{i=1}^{|V|} \left(\psi \cdot p^{f^\star}(v_i) + 1\right). \tag{74}$$

Theorem 2.8. *Let f and $f^\star$ be information functions. Let $\mathcal{G}$ be a class of graphs. If*

$$p^f(v) < \psi \cdot p^{f^\star}(v), \quad \forall G \in \mathcal{G}, \tag{75}$$

then

$$I^f(G) > \psi \cdot I^{f^\star}(G) - \psi \cdot \log(\psi) - \psi \cdot \sum_{i=1}^{|V|} p^{f^\star}(v_i) \log\left(1 + \frac{1}{\psi \cdot p^{f^\star}(v_i)}\right) - \sum_{i=1}^{|V|} \log\left(p^{f^\star}(v_i) + 1\right). \tag{76}$$

We state the following assertions without proof since the underlying procedure is similar to the one just presented. These results also describe implicit information inequalities when using a different relation for the vertex probabilities.

Theorem 2.9. *Let f and $f^\star$ be information functions. Let $\mathcal{G}$ be a class of graphs. If*

$$p^f(v) > \psi \cdot p^{f^\star}(v), \quad \forall G \in \mathcal{G}, \tag{77}$$

then

$$I^{f^\star}(G) + \frac{1}{\psi} \sum_{i=1}^{|V|} p^f(v_i) \cdot \log\left(p^f(v_i) + 1\right) + \frac{1}{\psi} \sum_{i=1}^{|V|} \log\left(p^f(v_i) + 1\right) - \log(\psi) > 0, \quad \forall G \in \mathcal{G}. \tag{78}$$

Theorem 2.10. *Let f and $f^\star$ be information functions. Let $\mathcal{G}$ be a class of graphs. If*

$$p^f(v) > \psi \cdot p^{f^\star}(v), \quad \forall G \in \mathcal{G} \tag{79}$$

and

$$p^f(v) > \frac{1}{2^c - 1} > 0, \quad c > 0, \tag{80}$$

then

$$I^{f^\star}(G) > \frac{I^f(G)}{\psi} - \frac{c}{\psi} - \frac{1}{\psi} \sum_{i=1}^{|V|} \left(p^f(v_i) + 1\right) + \log(\psi). \tag{81}$$

Theorem 2.11. *Let f and $f^\star$ be information functions. Let $\mathcal{G}$ be a class of graphs. If*

$$p^f(v) > \psi \cdot p^{f^\star}(v), \quad \forall G \in \mathcal{G}, \tag{82}$$

then

$$I^{f^\star}(G) > \frac{I^f(G)}{\psi} - \frac{1}{\psi} \sum_{i=1}^{|V|} p^f(v_i) \log\left(1 + \frac{1}{p^f(v_i)}\right) - \frac{1}{\psi} \sum_{i=1}^{|V|} \log(p^f(v_i)) + \log(\psi). \tag{83}$$

The next theorem demonstrates that such information inequalities can be derived by assuming characteristic properties of the functions involved [34]. The following statement is a consequence of the concave property of logarithmic function, see [38].

Theorem 2.12. *Let f and $f^\star$ be information functions. Then*

$$I^f(G) \geqslant -\sum_{i=1}^{|V|} p(v_i) \log(p^{f^\star}(v_i)) - \frac{1}{\ln(2)} \sum_{i=1}^{|V|} \frac{(p^f(v_i))^2 - p^{f^\star}(v_i) \cdot p^f(v_i)}{p^{f^\star}(v_i)}. \tag{84}$$

Theorem 2.13. *Let f and $f^\star$ be information functions. Then*

$$I^{f^\star}(G) \leqslant -\sum_{i=1}^{|V|} p^{f^\star}(v_i)\log(p(v_i)) + \frac{1}{\ln(2)}\sum_{i=1}^{|V|}\frac{p^f(v_i)\cdot p^{f^\star}(v_i) - (p^{f^\star}(v_i))^2}{p^{f^\star}(v_i)}. \tag{85}$$

3. Evaluation and interpretation of selected entropy measures

In this section, we evaluate some selected entropy measures and interpret the results. The most important question we want to tackle here is which kind of structural information the entropy measures do detect. Before starting, we emphasize that graph complexity can not be uniquely defined because there exist a lot of different structural features which contribute to the complexity of a graph. Similarly, to detect molecular complexity by taking the topological complexity of the underlying molecule into account is a challenging undertaking, see, e.g., [9,71]. Following Nikolić and Trinajstić [71], the topological complexity of a molecular graph is characterized by its number of vertices and edges, branching, cyclicity etc. For instance, concrete topological measures for capturing topological complexity can be found in [8,15,92].

Consider the following scatter plots (Figs. 8 and 9) describing correlations between the graph entropy measures. For calculating the graph entropies, we use a set of 2265 unlabeled graphs (MS 2265) selected from a mass spectral database. For the graphs in this set, we obtain

$$4 \leqslant |V| \leqslant 19, \quad 2 \leqslant \rho(G) \leqslant 15. \tag{86}$$

Further details of this set of graphs can be found in [35]. The first two scatter plots (see Figs. 8 and 9) show correlations between I_a and $I^{\lambda}_{f^{V_2}_{lin}}, I^W_D$ (instead of the symbol $\bar{I}^W_D$, we here use the simplified notation I^W_D). First of all, we see that these measures are highly uncorrelated, which means that they capture structural information differently. Also, it is evident that I_a is highly degenerate as indicated by the vertical strips in the scatter plots. This is clear from the definition of I_a: In order to calculate this graph entropy measure, we need to determine the vertex orbits. This corresponds to the problem of partitioning a graph into sets that contain only topologically equivalent vertices. Note that graphs of significantly different structure may have the same vertex partition. We turn now to some properties of I_a and $I^{\lambda}_{f^{V_j}}$. It easily follows that $I_a = 0$ for vertex transitive graphs. An example of such a graph is C_7 (cycle with 7 vertices) as shown in Fig. 10. Note that C_7 has a rich symmetry structure inasmuch as its automorphisms form the dihedral group D_7. It is also the case that for graphs $G \in \mathcal{G}_k$, $I^{\lambda}_{f^{V_j}}$, $j = 1, 2$ is zero, independent of the choice of the parameters c_k. By applying f^{V_j}, the resulting graph entropy $I_{f^{V_j}}$ (see Eq. (55)) is $\log(|V|)$ (the maximum possible value) and, hence by definition (see Eq. (56)), $I^{\lambda}_{f^{V_j}}(G) = 0$, $G \in \mathcal{G}_k$. Interestingly, Fig. 11 represents the

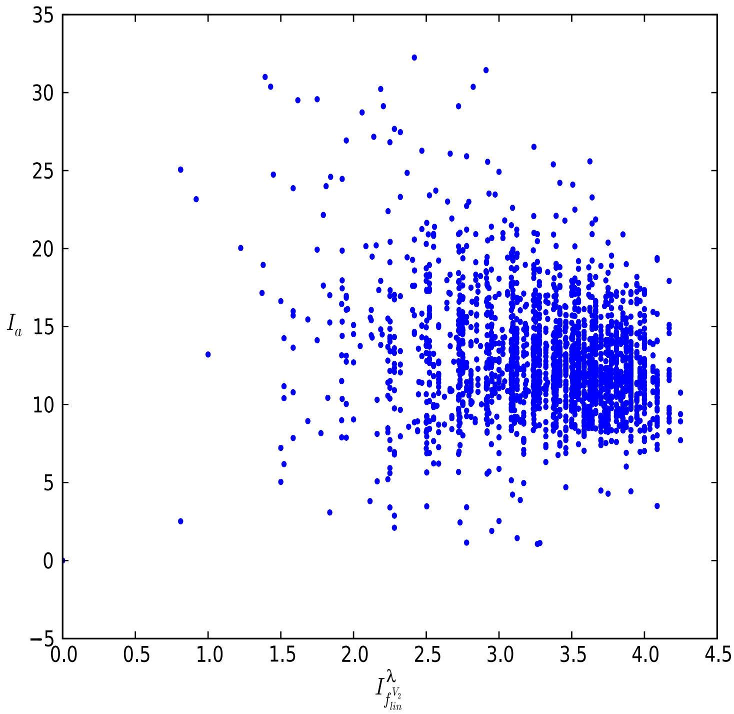

Fig. 8. I_a versus $I^{\lambda}_{f^{V_2}_{lin}}$.

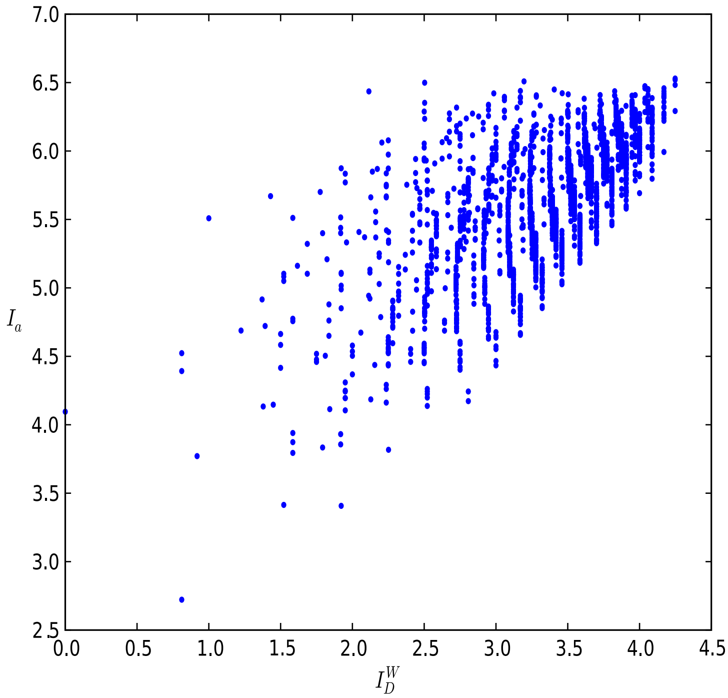

Fig. 9. I_a versus I_D^W.

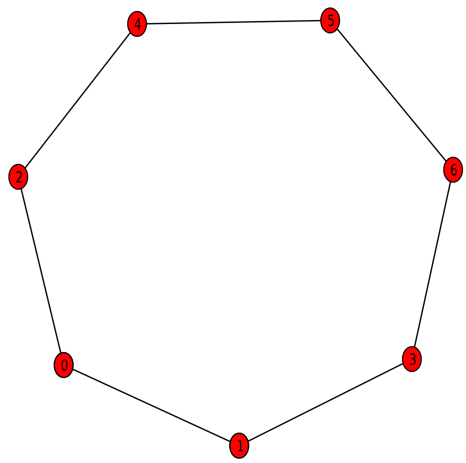

Fig. 10. A vertex transitive and 2-regular graph.

graph which maximizes I_a because all vertex partitions are singleton sets. Based on the characteristics of the vertex partition, this graph is highly asymmetrical.

The information measure $I^{\lambda}_{f^{V_2}_{lin}}$ has the maximum value for the graph shown in Fig. 12. The detailed computation of $I^{\lambda}_{f^{V_2}_{lin}}$ for this graph reveals that its vertices are topologically very different with respect to neighborhood properties defined by the j-sphere cardinalities used to compute f^{V_2} (Eq. (54)) and, finally, $I^{\lambda}_{f^{V_2}}$. Hence, the higher the value of $I^{\lambda}_{f^{V_2}}$, the more topologically different are the vertices in the graph and the fewer the symmetries [35]. The graphs with minimum and maximum entropy with respect to I_D^W are shown in Figs. 13 and 14, respectively. The graph in Fig. 13 is acyclic, whereas the one in Fig. 14 has several cycles. As observed for $I_{f^{V_j}}$ (see Eq. (55)), a high value of I_D^W is associated with a highly symmetrical graph whereas in this case, the molecular complexity can be interpreted relative to cycle structure. This is in accordance with Bonchev's observation [12] that molecular complexity increases with the number of rings, multiple bonds, branches, as well as with molecular size.

Finally, consider the relationships shown in Figs. 15 and 16. When starting with $I^{\lambda}_{f^{V_2}_{lin}}$ versus I_D^W, we see that they are not correlated at all. However, I^1_{loc} versus I_D^W reveals the existence of highly correlated clusters. Thus, the measures capture

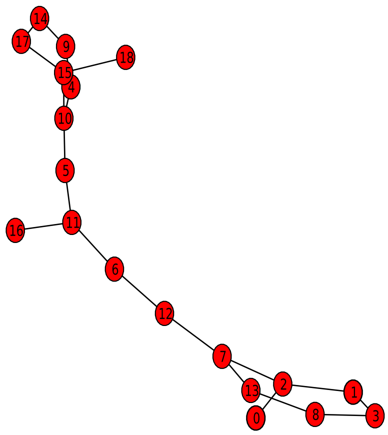

Fig. 11. Graph that has maximum entropy for I_a.

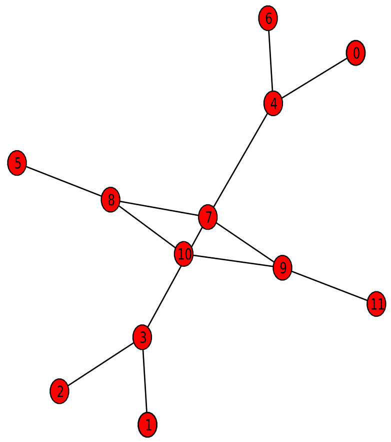

Fig. 12. Graph which maximizes $I^{\lambda}_{f^{v_2}_{lin}}$.

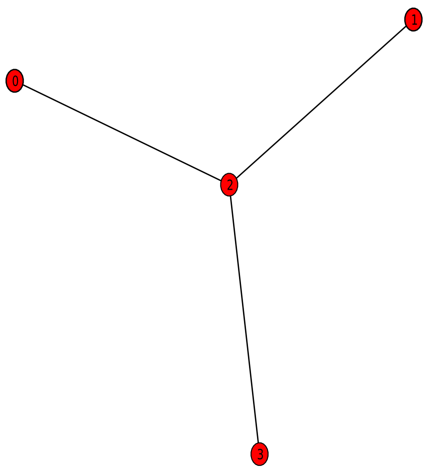

Fig. 13. Graph which minimizes I_D^W.

structural information in a similar way. It is important to emphasize that the procedures used to construct these two measures differ significantly. I_D^W (see Eq. (23)) is a so-called magnitude-based measure constructed by assigning weighted probabilities derived from distances in the graph; by contrast, the graph entropy I_{loc}^1 is defined as the mean of local vertex entropies (see Eqs. (32) and (33)).

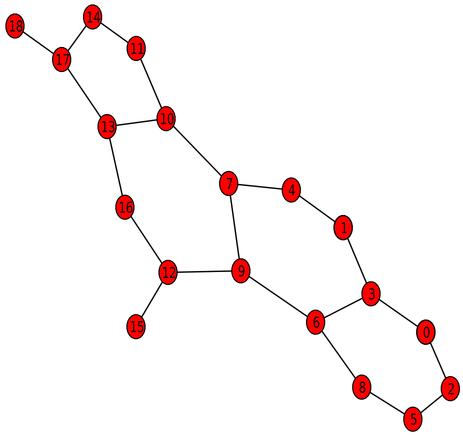

Fig. 14. Graph which maximizes I_D^W.

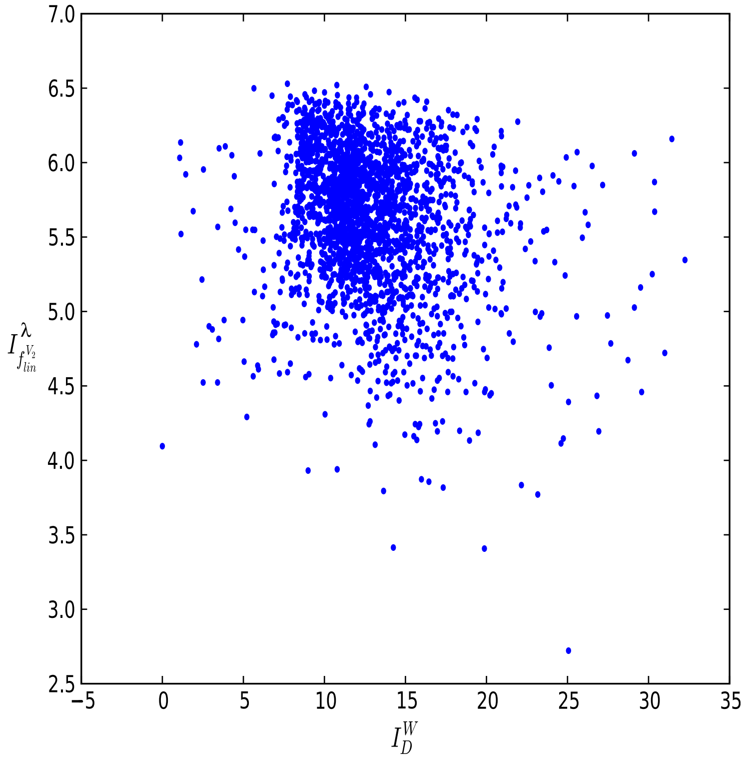

Fig. 15. $I^{\lambda}_{f_{lin}^{V_2}}$ versus I_D^W.

4. Network information measures: further applications

In Section 2, we examined the use of network information measures in biology, chemistry, and social network analysis. Here we look at some important applications of such measures in other domains.

Networks have been used to model characteristics of living organisms, evolutionary processes and general properties of complex systems, see, e.g., [17,37]. In particular, networks representing natural systems have often been investigated from the perspective of statistical physics. These investigations have given rise to insights into scale free networks, and to the development of new methods for analyzing random networks [7,17,37,70]. More generally, models for determining the complexity of networks have turned out to have wide applicability [25,52]. For example, Thurner [84] has studied topological phase transitions and network entropy in the context of statistical mechanics. To develop an information-theoretic analysis of networks, Anand and Bianconi [3] defined Shannon entropies for network ensembles and also discussed physically-based

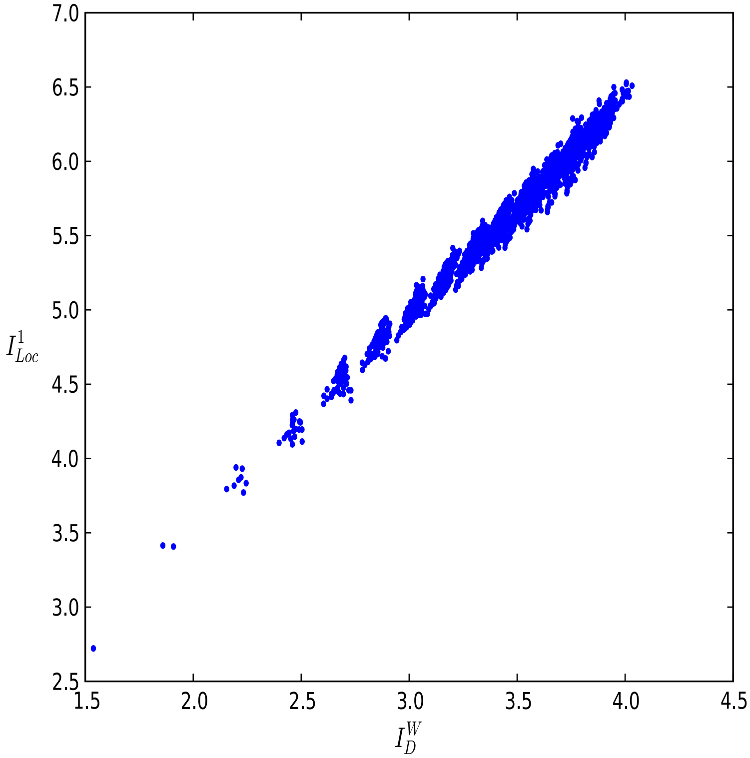

Fig. 16. I^1_{loc} versus I^W_D.

network entropies such as Gibbs entropy or von Neumann entropy [72]. Moreover, information-theoretic complexity measures for directed graphs such as the 'Medium Articulation' have been developed [93]. Claussen [21] has defined an entropy measure called 'Offdiagonal complexity'. For purposes of determining the entropy of a network, Claussen [21] used the so-called offdiagonal elements of the vertex-vertex link correlation matrix. Using the number of spanning trees of a network, Kim and Wilhelm [52] explored an entropy measure based on calculating a quantity for each edge taking account of the number of spanning trees of the graph and the number of spanning trees of the corresponding one-edge-deleted subgraph. The result can be then interpreted as a kind of spanning tree sensitive complexity measure of a network, see [52]. Note that many similar information measures for describing disorder and complexity of networks can be found in [14,25,83]. Moreover, various non-information-theoretic techniques for determining the complexity of networks have been investigated recently. For further details, see, e.g., [4,10,27,26,52].

In computer-related disciplines, graph entropies have also been proven to be useful. For example, information measures have been used for measuring the complexity of software abstractions [1]. In this work, graph patterns are represented as hypergraphs and the resulting entropies are used to measure size, complexity, coupling, and their significance, see [1]. In a more recent contribution, Borgert et al. [16] investigated business processes represented by undirected graphs and determined their structural complexity by using information-theoretic and non-information-theoretic graph measures. In this research, the challenge is to find measures which can encode structural information uniquely, i.e., those whose discrimination power is high [16]. As in applications in mathematical chemistry [11,35,85], graph entropy measures were found to be the best measures to distinguish non-isomorphic process graphs uniquely [16]. Ideally, complexity measures for process models could be used to detect errors, see, e.g., [1,57] and, hence, they could serve as useful tools when designing and analyzing process models such as graphs inferred from real-life business or software processes. In computational linguistics, Mehler [62] employed graph entropy measures as balance or imbalance measures by using so-called 'social ontology graphs' representing complex hierarchical structures. Entropy measures have been found to be useful for detecting significant structural characterizes of social ontologies [62].

In addition to the classical contributions mentioned in Section 2.3, information theory has been applied in modern biologically-related disciplines such as systems biology, computational biology and ecology [95,90]. For example, Ulanowicz [89,90] discussed various measures for the quantitative analysis of flow graphs representing directed ecological networks. Examples include the use of Shannon's entropy to measure uncertainty in the flows and the biodiversity of an ecological system, see [89,90]. Moreover, other information measures like conditional entropy and mutual information have been applied to the analysis of ecological networks [89,90]. For a discussion of other information-theoretic measures in this context, see [47,89,90]. Special biological systems such as gene networks are yet another arena for application of entropy measures [2]. In

[2], Altay and Emmert-Streib performed a statistical analysis using information-theoretic techniques to investigate network inference algorithms for inferring gene networks. Also, Emmert-Streib and Dehmer [41] explored the information spread in a gene network by performing single gene knockouts. Finally, the deviation between perturbed and unperturbed communication in networks was measured using the Kullback–Leibler distance [56]. Further related work can be found in, e.g., [20,51,61].

5. Summary and conclusion

Shannon's entropy measure has been used in diverse contexts to characterize graphs and properties of graphs. This survey has attempted to capture the variety of applications and to highlight underlying similarities and differences between the entropy measures. Beginning with the classical work on structural information content of graphs, we proceeded to examine the extensive research dealing with entropy measures designed to characterize graphs representing chemical structures. This was followed by an examination of entropy measures based on local as opposed to global graph properties. The small body of research on the use of entropy measures to analyze social network structures was included in the survey because of its potential importance. Current research on parametric entropy measures to quantify the information content of graphs extends the range of applications by allowing for the definition of families of entropy measures. Graph decomposition offers yet another method for defining entropy measures, in particular by making use of 'natural' vertex partitions on hierarchical graphs. The last part of the survey examined information inequalities, relations between entropy measures, and correlations between some selected graph entropies with a view to illuminating the problem of determining the nature of the structural information captured by the measures.

The wide applicability of graph-based models offers a virtually limitless field for the use of Shannon's entropy to measure structural differences. Identification and classification of structural configurations in networks pose challenging problems for which entropy measures have proven useful. Further development of the theory of entropy measures and progress in designing efficient algorithms for computing entropy are needed to meet this challenge.

Acknowledgments

Research was sponsored by US Army Research laboratory and the UK Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the US Army Research Laboratory, the U.S. Government, the UK Ministry of Defense, or the UK Government. The US and UK Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon. Further, we are grateful to Danail Bonchev, Armin Graber, Frank Emmert-Streib and Alexander Mehler for fruitful discussions. We also thank Kurt Varmuza for providing us the data base containing the chemical graphs and Nicola Barbarini for helping us with numerical calculations.

References

[1] E.B. Allen, Measuring graph abstractions of software: an information-theory approach, in: Proceedings of the Eighth International Symposium on Software Metrics table of contents, IEEE Computer Society, 2002, p. 182.
[2] G. Altay, F. Emmert-Streib, Revealing differences in gene network inference algorithms on the network-level by ensemble methods, Bioinformatics 26 (2010) 1738–1744.
[3] K. Anand, G. Bianconi, Entropy measures for networks: toward an information theory of complex topologies, Physical Review E 80 (2009) 045102(R).
[4] L. Antiqueira, L. da F. Costa, Characterization of subgraph relationships and distribution in complex networks, New Journal of Physics 11 (2009).
[5] A.T. Balaban, T.S. Balaban, New vertex invariants and topological indices of chemical graphs based on information on distances, Journal of Mathematical Chemistry 8 (1991) 383–397.
[6] T. Balch, Hierarchic social entropy: an information theoretic measure of robot group diversity, Autonomous Robots 8 (2001) 209–237.
[7] A.L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.
[8] S.H. Bertz, The first general index of molecular complexity, Journal of the American Chemical Society 103 (1981) 3241–3243.
[9] S.H. Bertz, On the complexity of graphs and molecules, Bulletin of Mathematical Biology 45 (1983) 849–855.
[10] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Complex networks: structure and dynamics, Physics Reports 424 (4–5) (2006) 175–308.
[11] D. Bonchev, Information Theoretic Indices for Characterization of Chemical Structures, Research Studies Press, Chichester, 1983.
[12] D. Bonchev, Overall connectivities and topological complexities: a new powerful tool for QSPR/QSAR, Journal of Chemical Information and Computer Science 40 (4) (2000) 934–941.
[13] D. Bonchev, Complexity in Chemistry, Introduction and Fundamentals, Taylor and Francis, Boca Raton, FL, USA, 2003.
[14] D. Bonchev, D.H. Rouvray, Complexity in Chemistry, Biology and Ecology, Mathematical and Computational Chemistry, Springer, New York, NY, USA, 2005.
[15] D. Bonchev, N. Trinajstić, Information theory distance matrix and molecular branching, Journal of Chemical Physics 67 (1977) 4517–4533.
[16] S. Borgert, M. Dehmer, E. Aitenbichler, On quantitative network complexity measures for business process models, Acta Universitates Apulensis, 18, in press.
[17] S. Bornholdt, H.G. Schuster, Handbook of Graphs and Networks: From the Genome to the Internet, John Wiley & Sons Inc., New York, NY, USA, 2003.
[18] U. Brandes, A faster algorithm for betweenness centrality, Journal of Mathematical Sociology 25 (2) (2001) 163–177.
[19] C.T. Butts, The complexity of social networks: theoretical and empirical findings, Social Networks 23 (2001) 31–71.
[20] P. Chanda, L. Sucheston, S. Liu, A. Zhang, M. Ramanathan, Information-theoretic gene-gene and gene-environment interaction analysis of quantitative traits, BMC Genomics 10 (1) (2009) 509.
[21] J.C. Claussen, Characterization of networks by the offdiagonal complexity, Physica A 375 (2007) 365–373.
[22] W.J. Cook, W.H. Cunningham, W.R. Pullyblank, A. Schrijver, Combinatorial Optimization, Wiley, 1998.
[23] T.M. Cover, J.A. Thomas, Elements of Information Theory, Wiley Series in Telecommunications and Signal Processing, Wiley & Sons, 2006.

[24] I. Csiszár, J. Körner, L. Lovász, K. Marto, G. Simonyi, Entropy splitting for antiblocking corners and perfect graphs, Combinatorica 10 (1) (1990) 27–40.
[25] L. da F. Costa, F. Rodrigues, G. Travieso, Characterization of complex networks: a survey of measurements, Advances in Physics 56 (2007) 167–242.
[26] L. da F. Costa, F.A. Rodrigues, Seeking for simplicity in complex networks, Europhysics Letters 85 (2009) 48001(1)–48001(6).
[27] L. da F. Costa, F.A. Rodrigues, C.C. Hilgetag, M. Kaiser, Beyond the average: detecting global singular nodes from local features in complex networks, Europhysics Letters 87 (2009) 18008(1)–18008(6).
[28] M. Dehmer, Information-theoretic concepts for the analysis of complex networks, Applied Artificial Intelligence 22 (7 & 8) (2008) 684–706.
[29] M. Dehmer, A novel method for measuring the structural information content of networks, Cybernetics and Systems 39 (2008) 825–843.
[30] M. Dehmer, S. Borgert, F. Emmert-Streib, Entropy bounds for molecular hierarchical networks, PLoS ONE 3 (8) (2008) e3079.
[31] M. Dehmer, F. Emmert-Streib, Structural information content of networks: graph entropy based on local vertex functionals, Computational Biology and Chemistry 32 (2008) 131–138.
[32] M. Dehmer, F. Emmert-Streib, Towards network complexity, in: Complex Sciences, in: J. Zhou (Ed.), Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 4, Springer, Berlin/Heidelberg, Germany, 2009, pp. 707–714.
[33] M. Dehmer, A. Mehler, F. Emmert-Streib, Graph-theoretical characterizations of generalized trees, in: Proceedings of the International Conference on Machine Learning: Models, Technologies & Applications (MLMTA'07), Las Vegas/USA, 2007, pp. 113–117, 2007.
[34] M. Dehmer, A. Mowshowitz, Inequalities for entropy-based measures of network information content, Applied Mathematics and Computation 215 (2010) 4263–4271.
[35] M. Dehmer, K. Varmuza, S. Borgert, F. Emmert-Streib, On entropy-based molecular descriptors: statistical analysis of real and synthetic chemical structures, Journal of Chemical Information and Modelling 49 (2009) 1655–1663.
[36] E.W. Dijkstra, A note on two problems in connection with graphs, Numerische Math. 1 (1959) 269–271.
[37] S.N. Dorogovtsev, J.F.F. Mendes, Evolution of Networks, From Biological Networks to the Internet and WWW, Oxford University Press, 2003.
[38] S.S. Dragomir, C.J. Goh, Some bounds on entropy measures in information theory, Applied Mathematics Letters 10 (1997) 23–28.
[39] F. Emmert-Streib, M. Dehmer, Information theoretic measures of UHG graphs with low computational complexity, Applied Mathematics and Computation 190 (2007) 1783–1794.
[40] F. Emmert-Streib, M. Dehmer, Topological mappings between graphs trees and generalized trees, Applied Mathematics and Computation 186 (2) (2007) 1326–1333.
[41] F. Emmert-Streib, M. Dehmer, Fault tolerance of information processing in gene networks, Physica A 338 (4) (2009) 541–548.
[42] F. Emmert-Streib, M. Dehmer, J. Kilian, Classification of large graphs by a local tree decomposition, in: H.R. Arabnia, A. Scime (Eds.), Proceedings of DMIN'05, International Conference on Data Mining, Las Vegas, USA, 2005, pp. 200–207.
[43] M.G. Everett, Role similarity and complexity in social networks, Social Networks 7 (1985) 353–359.
[44] L.C. Freeman, Spheres cubes and boxes: graph dimensionality and network structure, Social Networks 5 (1983) 139–156.
[45] R. Halin, Graphentheorie, Akademie Verlag, Berlin, Germany, 1989.
[46] F. Harary, Graph Theory, Addison Wesley Publishing Company, Reading, MA, USA, 1969.
[47] H. Hirata, R.E. Ulanowicz, Information theoretical analysis of ecological networks, International Journal of Systems Science 15 (1984) 261–270.
[48] H. Hosoya, Topological index. A newly proposed quantity characterizing the topological nature of structural isomers of saturated hydrocarbons, Bulletin of the Chemistry Society of Japan 44 (1971) 2332–2339.
[49] A.K. Jain, R.C. Dubes, Algorithms for Clustering Data, Prentice-Hall Inc., Upper Saddle River, NJ, USA, 1988.
[50] B. Jiang, C. Claramunt, A structural approach to the model generalization of an urban street network, Geoinformatica 8 (2) (2004) 157–171.
[51] D.C. Kim, X. Wang, C.R. Yang, J. Gao, Learning biological network using mutual information and conditional independence, BMC Bioinformatics 11 (3) (2010) S9.
[52] J. Kim, T. Wilhelm, What is a complex graph?, Physica A 387 (2008) 2637–2652
[53] E.V. Konstantinova, A.A. Paleev, Sensitivity of topological indices of polycyclic graphs, Vychisl. Sistemy. 136 (1990) 38–48 (in Russian).
[54] E.V. Konstantinova, V.A. Skorobogatov, M.V. Vidyuk, Applications of information theory in chemical graph theory, Indian Journal of Chemistry 42 (2002) 1227–1240.
[55] J. Körner, Coding of an information source having ambiguous alphabet and the entropy of graphs, in: Transactions of the Sixth Prague Conference on Information Theory, 1973, pp. 411–425,.
[56] S. Kullback, R.A. Leibler, On information and sufficiency, Annals of Mathematical Statistics 22 (1) (1951) 79–86.
[57] A.M. Latva-Koivisto. Finding a complexity measure for business process models, Technical report, Helsinki University of Technology, 2001.
[58] M. Li, P.M.B. Vitányi, Combinatorics and kolmogorov complexity, in: Proceedings of the Sixth Annual IEEE Conference on Structure in Complexity Theory, Springer, Berlin, New York, 1991.
[59] J.-L. Lu, F. Valois, M. Dohler, D. Barthel, Quantifying organization by means of entropy, Communications Letters IEEE 12 (3) (2008) 185–187.
[60] R.D. Luce, R.R. Bush, E. Galanter, Handbook of Mathematical Psychology, John Wiley and Sons, 1967.
[61] M. Maruyama, P. Fenwick, A. Ioannides, Interocular yoking in human saccades examined by mutual information analysis, Nonlinear Biomedical Physics 4 (1) (2010) S10.
[62] A. Mehler. A quantitative graph model of social ontologies by example of Wikipedia. in M. Dehmer, F. Emmert-Streib, A. Mehler (Eds.), Towards an Information Theory of Complex Networks: Statistical Methods and Applications, Birkhäuser, Boston/Basel, 2009.
[63] A. Mehler, M. Dehmer, R. Gleim, Towards logical hypertext structure — A graph-theoretic perspective, in: Proceedings of the Fourth International Workshop on Innovative Internet Computing Systems (I2CS '04), in: T. Böhme, G. Heyer (Eds.), Lecture Notes in Computer Science, vol. 3473, Springer, Berlin/New York, 2004, pp. 136–150.
[64] D. Minoli, Combinatorial graph complexity, Atti. Accad. Naz. Lincei, VIII. Ser., Rend., Cl. Sci. Fis. Mat. Nat. 59 (1975) 651–661.
[65] H. Morowitz, Some order-disorder considerations in living systems, Bulletin of Mathematical Biophysics 17 (1953) 81–86.
[66] A. Mowshowitz, Entropy and the complexity of graphs II: the information content of digraphs and infinite graphs, Bulletin of Mathematical Biophysics 30 (1968) 225–240.
[67] A. Mowshowitz, Entropy and the complexity of graphs III: graphs with prescribed information content, Bulletin of Mathematical Biophysics 30 (1968) 387–414.
[68] A. Mowshowitz, Entropy and the complexity of graphs IV: entropy measures and graphical structure, Bulletin of Mathematical Biophysics 30 (1968) 533–546.
[69] A. Mowshowitz, Entropy and the complexity of the graphs I: an index of the relative complexity of a graph, Bulletin of Mathematical Biophysics 30 (1968) 175–204.
[70] M.E.J. Newman, A.L. Barabási, D.J. Watts, The Structure and Dynamics of Networks, Princeton Studies in Complexity, Princeton University Press, 2006.
[71] S. Nikolić, N. Trinajstić, Complexity of molecules, J. Chem. Inf. Comput. Sci. 40 (2000) 920–926.
[72] F. Passerini, S. Severini, The von neumann entropy of networks, International Journal of Agent Technologies and Systems 1 (2009) 58–67.
[73] H. Quastler, Information theory in biology, Bulletin of Mathematical Biology (1954) 183–185.
[74] M. Randić, On characterization of molecular branching, Journal of the American Chemical Society 97 (1975) 6609–6615.
[75] M. Randić, D. Plavšić, Characterization of molecular complexity, International Journal of Quantum Chemistry 91 (2002) 20–31.
[76] M. Randić, D. Plavšić, On the concept of molecular complexity, Croatica Chemica Acta 75 (2002) 107–116.
[77] N. Rashevsky, Life information theory and topology, Bulletin of Mathematical Biophysics 17 (1955) 229–235.
[78] C. Raychaudhury, S.K. Ray, J.J. Ghosh, A.B. Roy, S.C. Basak, Discrimination of isomeric structures using information theoretic topological indices, Journal of Computational Chemistry 5 (1984) 581–588.
[79] C.E. Shannon, W. Weaver, The Mathematical Theory of Communication, University of Illinois Press, 1949.

[80] G. Simonyi, Graph entropy: a survey, in: W. Cook, L. Lovász, P. Seymour (Eds.), Combinatorial Optimization, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, vol. 20, 1995, pp. 399–441.
[81] G. Simonyi, Perfect graphs and graph entropy. An updated survey, in: J. Ramirez-Alfonsin, B. Reed (Eds.), Perfect Graphs, John Wiley & Sons, 2001, pp. 293–328.
[82] V.A. Skorobogatov, A.A. Dobrynin, Metrical analysis of graphs, Communications in Mathematical and in Computer Chemistry 23 (1988) 105–155.
[83] R.V. Solé, S. Valverde, Information theory of complex networks: on evolution and architectural constraints, in: Lecture Notes in Physics, vol. 650, 2004, pp. 189–207.
[84] S. Thurner, Statistical mechanics of complex networks, in: M. Dehmer, F. Emmert-Streib (Eds.), Analysis of Complex Networks: From Biology to Linguistics, Wiley-VCH, 2009, pp. 23–45.
[85] R. Todeschini, V. Consonni, R. Mannhold, Handbook of Molecular Descriptors, Wiley-VCH, Germany,Weinheim, 2002.
[86] E. Trucco, A note on the information content of graphs, Bulletin of Mathematical Biology 18 (2) (1956) 129–135.
[87] L. Türker, Contemplation on the hosoya indices, Journal of Molecular Structure (Theochem) 623 (2003) 75–77.
[88] F. Tutzauer, Entropy as a measure of centrality in networks characterized by path-transfer flow, Social Networks 29 (2007) 249–265.
[89] R.E. Ulanowicz, Information theory in ecology, Computers and Chemistry 25 (2001) 393–399.
[90] R.E. Ulanowicz, Quantitative methods for ecological network analysis, Computational Biology and Chemistry 28 (2004) 321–339.
[91] S. Wasserman, K. Faust, Social Network Analysis: Methods and Applications, Structural Analysis in the Social Sciences, Cambridge University Press, 1994.
[92] H. Wiener, Structural determination of paraffin boiling points, Journal of the American Chemical Society 69 (17) (1947) 17–20.
[93] T. Wilhelm, R. Brueggemann, Information theoretic measures for the maturity of ecosystems, in: M. Matthies, H. Malchow, J. Kriz (Eds.), Integrative Systems Approaches to Natural and Social Sciences Systems Science2000, Springer, Berlin,Germany, 2001, pp. 263–273.
[94] F. Xie, D. Levinson, Measuring the structure of road networks, Geographical Analysis 39 (3) (2007) 336–356.
[95] H.P. Yockey, Information Theory and Molecular Biology, Cambridge University Press, Cambridge,UK, 1992.