Information Sciences 278 (2014) 22–33

Contents lists available at ScienceDirect

Information Sciences

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

Extremality of degree-based graph entropies

Shujuan Cao [a], Matthias Dehmer [b,*], Yongtang Shi [c]

[a] School of Mathematical Sciences, Nankai University, Tianjin 300071, PR China
[b] Institute for Bioinformatics and Translational Research, UMIT, Hall in Tyrol, Austria
[c] Center for Combinatorics and LPMC-TJKLC, Nankai University, Tianjin 300071, PR China

ARTICLE INFO

Article history:
Received 20 September 2013
Received in revised form 18 January 2014
Accepted 29 March 2014
Available online 25 April 2014

Keywords:
Information theory
Entropy
Shannon's entropy
Graph entropy
Degree sequence
Degree power

ABSTRACT

Many graph invariants have been used for the construction of entropy-based measures to characterize the structure of complex networks. Based on Shannon's entropy, we study graph entropies which are based on vertex degrees by using so-called information functionals. When considering Shannon entropy-based graph measures, there has been very little work to find their extremal values. The main contribution of this paper is to prove some extremal values for the underlying graph entropy of certain families of graphs and to find the connection between the graph entropy and the sum of degree powers. Further, conjectures to determine extremal values of graph entropies are given.

1. Introduction

Studies of the information content of graphs and networks have been initiated in the late fifties based on the seminal work due to Shannon [70]. The concept of graph entropy [20,24] introduced by Rashevsky [67] and Trucco [76] has been used to measure the structural complexity of graphs [11,21,22]. The entropy of a graph is an information-theoretic quantity that has been introduced by Mowshowitz [58]. Here the complexity of a graph [25] is based on the well-known Shannon's entropy [18,20,70,58]. Importantly, Mowshowitz interpreted his graph entropy measure as the structural information content of a graph and demonstrated that this quantity satisfies important properties when using product graphs, etc., see, e.g., [58–61]. Note the Körner's graph entropy [52] has been introduced from an information theory point of view and has not been used to characterize graphs quantitatively. An extensive overview on graph entropy measures can be found in [24]. A statistical analysis of topological graph measures has been performed by Emmert-Streib and Dehmer [29].

Several graph invariants, such as the number of vertices, the vertex degree sequences, extended degree sequences (i.e., the second neighbor, third neighbor, etc.), edges, and connections, have been used for developing entropy-based measures [20,24]. In this paper, we introduce a novel graph entropy, which is based on a new information functional by using degree powers. Degree powers is one of the most important graph invariants, which has been proven useful in information theory, social networks, network reliability and mathematical chemistry, see [9,10]. In view of the vast of amount of existing graph entropy measures [11,20], there has been very little work to find their extremal values [23]. A reason for this might be the fact that Shannon's entropy represents a multivariate function and all probability values are not equal to zero when considering graph entropies. Inspired by Dehmer and Kraus [23], it turned out that determining minimal values of graph entropies

* Corresponding author.
E-mail addresses: sj.cao@163.com (S. Cao), matthias.dehmer@umit.at (M. Dehmer), shi@nankai.edu.cn (Y. Shi).

http://dx.doi.org/10.1016/j.ins.2014.03.133

is intricate because there is a lack of analytical methods to tackle this particular problem. Other related work is due to Shi [71], who proved a lower bound of quantum decision tree complexity by using Shannon's entropy. Dragomir and Goh [28] obtained several general upper bounds for Shannon's entropy by using Jensen's inequality [43]. Finally, Dehmer and Kraus [23] proved some extremal results for graph entropies which are based on information functionals.

The main contribution of the paper is to study novel properties of graph entropies which are based on an information functional by using degree powers of graphs. In particular, we determine the extremal values for the underlying graph entropy of certain families of graphs and find the connection between graph entropy and the sum of degree powers, which is well-studied in graph theory and some related disciplines. Further, conjectures to determine extremal values of graph entropies are proposed.

The paper is organized as follows. In Section 2, some concepts and notation in graph theory are introduced. In Section 3, we introduce some results on the sum of degree powers. In Section 4, we state the definitions of graph entropies based on the given information functional by using degree powers. In Sections 5 and 6, extremal properties of graph entropies have been studied. Further, we express some conjectures to find extremal values of trees. We discuss some potential applications of degree-based entropies in Section 7. The paper finishes with a summary and conclusion in Section 8.

2. Preliminaries

A *graph* G is an ordered pair of sets $V(G)$ and $E(G)$ such that the elements $uv \in E(G)$ are a sub-collection of the unordered pairs of elements of $V(G)$. For convenience, we denote a graph by $G = (V, E)$ sometimes. The elements of $V(G)$ are called *vertices* and the elements of $E(G)$ are called *edges*. If $e = uv$ is an edge, then we say vertices u and v are *adjacent*, and $u,\ v$ are two endpoints (or ends) of e. A *loop* is an edge whose two endpoints are the some one. Two edges are called *parallel*, if both edges have the same endpoints. A *simple graph* is a graph containing no loops and parallel edges. If G is a graph with n vertices and m edges, then we say the *order* of G is n and the *size* of G is m. A graph of order n is addressed as an *n-vertex graph*, and a graph of order n and size m is addressed as an (n, m)-*graph*. A graph F is called a *subgraph* of a graph G, if $V(F) \subseteq V(G)$ and $E(F) \subseteq E(G)$, denoted by $F \subseteq G$. In this paper, we only consider simple graphs.

Let $G = (V(G), E(G))$ and $H = (V(H), E(H))$ be two simple graphs. A *graph isomorphism* from G to H is a bijection $f:\ V(G) \to V(H)$ such that $uv \in E(G)$ if and only if $f(u)f(v) \in E(H)$. If there is a graph isomorphism from G to H, then G is said to be *isomorphic* to H, denoted by $G \cong H$.

A graph is *connected* if, for every partition of its vertex set into two nonempty sets X and Y, there is an edge with one end in X and one end in Y. Otherwise, the graph is *disconnected*. In other words, a graph is *disconnected* if its vertex set can be partitioned into two nonempty subsets X and Y so that no edge has one end in X and one end in Y.

A *path graph* is a simple graph whose vertices can be arranged in a linear sequence in such a way that two vertices are adjacent if they are consecutive in the sequence, and are nonadjacent otherwise. Likewise, a *cycle graph* on three or more vertices is a simple graph whose vertices can be arranged in a cyclic sequence in such a way that two vertices are adjacent if they are consecutive in the sequence, and are nonadjacent otherwise. Denote by P_n and C_n the path graph and the cycle graph with n vertices, respectively.

A connected graph without any cycle is a *tree*. Actually, the path P_n is a tree of order n with exactly two pendent vertices. The *star* of order n, denoted by S_n, is the tree with $n-1$ pendent vertices. A simple connected graph is called *unicyclic* if it has exactly one cycle. We use S_n^+ to denote the unicyclic graph obtained from the star S_n by adding to it an edge between two pendent vertices of S_n. Observe that a tree and a unicyclic graph of order n have exactly $n-1$ and n edges, respectively. A *bicyclic graph* is a graph of order n with $n+1$ edges.

The *length* of a path is the number of its edges. For two vertices u and v, the *distance* between u and v in a graph G, denoted by $d_G(u, v)$, is the length of the shortest path connecting u and v. The *diameter* of a graph G is the greatest distance between two vertices of G.

All vertices adjacent to vertex u are called *neighbors* of u. The *neighborhood* of u is the set of the neighbors of u. The number of edges adjacent to vertex u is the *degree* of u, denoted by $d(u)$. Vertices of degrees 0 and 1 are said to be *isolated* and *pendent vertices*, respectively. A pendent vertex is also referred to as a *leaf* of the underlying graph. A vertex of degree i is also addressed as an *i-degree vertex*. The minimum and maximum degree of G is denoted by $\delta(G)$ and $\Delta(G)$, respectively. If G has a_i vertices of degree d_i $(i = 1, 2, \ldots, t)$, where $\Delta(G) = d_1 > d_2 > \cdots > d_t = \delta(G)$ and $\sum_{i=1}^{t} a_i = n$, we define the *degree sequence* of G as $D(G) = [d_1^{a_1}, d_2^{a_2}, \ldots, d_t^{a_t}]$. If $a_i = 1$, we use d_i instead of $d_i^{a_i}$ for convenience.

In chemical graph theory, a *chemical graph* or *molecular graph* is a representation of the structural formula of a chemical compound in terms of graph theory. Here, a graph corresponds to a chemical structural formula, in which a vertex and an edge correspond to an atom and a chemical bond, respectively. Since carbon atoms are 4-valent, we obtain graphs in which no vertex has degree greater than four. Analogously, a *chemical tree* is a tree T with maximum degree at most four. For a more thorough introduction on chemical graphs, we refer to [12,77].

Let G be a graph of order n. The *adjacency matrix* of a graph G is the $n \times n$ matrix $A(G) := (a_{uv})$, where a_{uv} is the number of edges joining vertices u and v, each loop counting as two edges. If G is simple, then $A(G)$ is a $(0, 1)$-matrix. The *eigenvalues* of G is the eigenvalues of its adjacency matrix $A(G)$. All eigenvalues of G forms the spectrum of G, which is widely studied in algebraic graph theory. For this topic, we refer to [34].