


On Marginal Entropy of Graphs

 Bünyamin Şahin

Department of Mathematics, Faculty of Science, Selçuk University, Konya, Turkey
 ✉ shnbnymn25@gmail.com

RECEIVED: July 6, 2021 ★ REVISED: November 12, 2021 ★ ACCEPTED: November 20, 2021

Abstract: Marginal entropy is one of the distances based on the graph entropy. Then, this entropy is computed by the Wiener index of graphs. In this paper, we obtain the marginal entropy of paths, stars, double stars, cycles and vertex-transitive graphs.

Keywords: graph entropy, distance, marginal entropy.

INTRODUCTION

THE entropy concept was introduced by Shannon in 1948.^[1] The Shannon entropy can be applied to different networks via the possibility of constructing a finite probability scheme for each network. The graph entropy concept was defined by Rashevsky^[2] in 1955. His entropy measure is based on partitioning the vertices with respect to equivalent classes of vertex degrees.

Many molecular properties of materials are obtained by molecular topologies.^[3] These measures are called topological indices or molecular descriptors in chemical graph theory. Chemical, physical and biological properties of molecules have good correlations with these topological indices. Therefore, many researchers from a wide range of sciences study this topic. The first topological index was introduced by Wiener in 1947.^[4] The Wiener index equals half of the total distances between every pair of vertices in a graph.

Too many topological indices have been introduced in the last 50 years. It is understood that they have usually correlated more or less with the relative molecular properties of molecules, but the same index can not have a high discrimination ability for different molecules.^[3] Bonchev and Trinajstić introduced an entropy measure that is based on distances to interpret the molecular branching of molecular graphs.^[5] Later they applied the information theory in characterization of chemical structures.^[6,7] These molecular descriptors were called information indices, and it was shown that the information indices have greater

discriminating power for molecules than the respective topological indices.^[8]

Hosoya entropy is based on distance-related partitions of the vertices.^[9] Computation of Hosoya entropy requires partitions of vertices concerning the number of vertices at the same distance to each vertex.^[10] More details about the graph entropies can be found in the paper.^[11] Moreover, the meeting of the information theory with the chemical graph theory was expressed by Bonchev in the paper.^[12] Another application of information theory in chemical graph theory was obtained by Konstantinova and Diudea^[13] in comparing the Wiener polynomial derivatives, information indices and topological indices.

In this paper, we study the marginal entropy which was defined by Konstantinova.^[3] We obtain the marginal entropy for paths, stars, double stars, cycles and vertex-transitive graphs.

PRELIMINARIES

Let G be a simple graph with the vertex set $V(G)$ and the edge set $E(G)$. For a vertex $u \in V(G)$, the notation $N_G(u) = \{v \mid uv \in E(G)\}$ denotes the vertices which are adjacent to u and $N_G[u] = \{u\} \cup N_G(u)$. The degree of a vertex u is the cardinality of $N_G(u)$ and it is denoted by $\deg_G(u)$ or simply $\deg(u)$. A vertex which has degree one is called a leaf. Moreover, the distance between the vertices u and v is denoted by $d(u,v)$.^[10]

The number of vertices of a graph G is called order, and it is denoted by n . The paths, cycles and stars of order

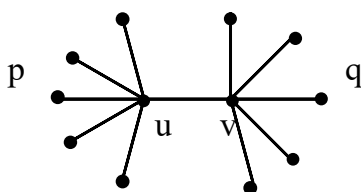


Figure 1. The double star graph $S_{p,q}$.

n are denoted by P_n , C_n and S_n , respectively. The double star graphs $S_{p,q}$ of order n are consisted of the stars $S_{1,p}$ and $S_{1,q}$ such that $n = p + q + 2$.

Definition 2.1. For a vertex $u \in V(G)$, the total distance of u is introduced^[4]

$$D(u) = \sum_{v \in V(G)} d(u, v).$$

Definition 2.2. The Wiener index of a graph G is introduced^[4]

$$W(G) = \frac{1}{2} \sum_{u \in V(G)} D(u).$$

Definition 2.3. The entropy of a graph G can be defined by Dehmer's information functional approach.^[11] Let G be a graph and $f: S \rightarrow R_+$ be an information functional defined on $S = \{s_1, s_2, \dots, s_k\}$ such that S is a set of elements of G . Then the entropy is defined as

$$\begin{aligned} I_f(G) &= - \sum_{i=1}^k \frac{f(s_i)}{\sum_{j=1}^k f(s_j)} \log \left(\frac{f(s_i)}{\sum_{j=1}^k f(s_j)} \right) \\ &= \log \left(\sum_{i=1}^k f(s_i) \right) - \frac{\sum_{i=1}^k f(s_i) \log(f(s_i))}{\sum_{j=1}^k f(s_j)}, \end{aligned}$$

as logarithmic phrases have base 2.

Definition 2.4. The marginal entropy of a graph G is denoted by $I_D(G)$ and it is computed by the following equation:^[3]

$$\begin{aligned} I_D(G) &= - \sum_{u \in V(G)} \frac{D(u)}{2W(G)} \log \left(\frac{D(u)}{2W(G)} \right) \\ &= 1 + 2 \log W(G) - \frac{1}{2W(G)} \sum_{u \in V(G)} D(u) \log D(u). \end{aligned}$$

Since $W(G) = \frac{1}{2} \sum_{u \in V(G)} D(u)$, the marginal entropy is given by the equality

$$I_D(G) = - \sum_{u \in V(G)} \frac{D(u)}{\sum_{u \in V(G)} D(u)} \log \left(\frac{D(u)}{\sum_{u \in V(G)} D(u)} \right).$$

MAIN RESULTS

Theorem 3.1. The marginal entropy of the path graph P_n of order n is given by the following formula for a vertex $v_i (1 \leq i \leq n)$

$$I_D(P_n) = - \sum_{i=1}^n \frac{3D(v_i)}{n^3 - n} \log \left(\frac{3D(v_i)}{n^3 - n} \right)$$

such that

$$D(v_i) = \frac{i(i-1) + (n-i)(n-i+1)}{2}.$$

Proof. For a vertex $v_i (1 \leq i \leq n)$ on a path, the distance from v_i to others as follows.

$$d(v_i, v_j) = \begin{cases} i-j, & \text{for } j < i \\ j-i, & \text{for } i < j. \end{cases}$$

Therefore, we obtain that

$$\begin{aligned} D(v_i) &= i-1 + i-2 + \dots + 1 + 1 + 2 + \dots + n-i \\ &= \frac{i(i-1) + (n-i)(n-i+1)}{2}. \end{aligned}$$

Since the Wiener index of a path^[14] of order n equals to $\frac{n^3-n}{6}$, the result is obtained.

Theorem 3.2. The marginal entropy of the star graph S_n of order n is given by the following formula

$$\begin{aligned} I_D(S_n) &= - \frac{n-1}{2n^2 - 4n + 2} \log \left(\frac{n-1}{2n^2 - 4n + 2} \right) \\ &\quad - (n-1) \frac{2n-3}{2n^2 - 4n + 2} \log \left(\frac{2n-3}{2n^2 - 4n + 2} \right). \end{aligned}$$

Proof. A star graph is consisted of a central vertex of degree $n-1$ and $n-1$ leaves. Then, the distance from the central vertex to leaves is one, and the total distance of the central vertex is $n-1$ to other vertices. Moreover, the distance from a leaf to another leaf is two in a star graph. Then, the total distance of a leaf to other vertices is $2n-3$. Therefore, we obtain the Wiener index of the stars as in the following equation.

$$W(S_n) = \frac{(n-1)(2n-3) + n-1}{2} = (n-1)^2.$$

By Definition 2.4, we obtain that

$$\begin{aligned} I_D(S_n) &= - \frac{n-1}{2n^2 - 4n + 2} \log \left(\frac{n-1}{2n^2 - 4n + 2} \right) \\ &\quad - (n-1) \frac{2n-3}{2n^2 - 4n + 2} \log \left(\frac{2n-3}{2n^2 - 4n + 2} \right). \end{aligned}$$

The proof is completed with this equation.

Theorem 3.3. The marginal entropy of the double star graph $S_{p,q}$ of order n is given by the following formula

$$I_D(S_{p,q}) = -p \frac{2p+3q+1}{2W(S_{p,q})} \log\left(\frac{2p+3q+1}{2W(S_{p,q})}\right) - q \frac{3p+2q+1}{2W(S_{p,q})} \log\left(\frac{3p+2q+1}{2W(S_{p,q})}\right) - \frac{p+2q+1}{2W(S_{p,q})} \log\left(\frac{p+2q+1}{2W(S_{p,q})}\right) - \frac{2p+q+1}{2W(S_{p,q})} \log\left(\frac{2p+q+1}{2W(S_{p,q})}\right)$$

Proof. A double star graph $S_{p,q}$ is illustrated in Figure 1 with the central vertices u, v and the leaves p, q . If a leaf a is taken from p leaves, then the total distance from a to other vertices is computed by

$$D(a) = 1 + 2(p-1) + 2 + 3q = 2p + 3q + 1.$$

Similarly, if a leaf b is taken from q leaves, then the total distance from b to other vertices is computed by

$$D(b) = 1 + 2(q-1) + 2 + 3p = 3p + 2q + 1.$$

We obtain the total distances of vertices u, v to other vertices as follows.

$$D(u) = p + 2q + 1,$$

$$D(v) = 2p + q + 1.$$

By these equations, the Wiener index of $S_{p,q}$ is computed such that

$$W(S_{p,q}) = \frac{1}{2}[pD(a) + qD(b) + D(u) + D(v)] = \frac{1}{2}[p(2p+3q+1) + q(2q+3p+1) + p+2q+1 + 2p+q+1] = p^2 + q^2 + 3pq + 2p + 2q + 1.$$

Finally, we obtain the marginal entropy of $S_{p,q}$ as

$$I_D(S_{p,q}) = -p \frac{2p+3q+1}{2W(S_{p,q})} \log\left(\frac{2p+3q+1}{2W(S_{p,q})}\right) - q \frac{3p+2q+1}{2W(S_{p,q})} \log\left(\frac{3p+2q+1}{2W(S_{p,q})}\right) - \frac{p+2q+1}{2W(S_{p,q})} \log\left(\frac{p+2q+1}{2W(S_{p,q})}\right) - \frac{2p+q+1}{2W(S_{p,q})} \log\left(\frac{2p+q+1}{2W(S_{p,q})}\right).$$

The proof is completed with this equation.

Theorem 3.4. The marginal entropy of the cycle graph C_n of order n is given by the following formula

$$I_D(C_n) = \log n.$$

Proof. We know that Wiener index of cycles^[14] equals to

$$W(C_n) = \begin{cases} \frac{n^3}{8}, & n \text{ even} \\ \frac{n^3-n}{8}, & n \text{ odd} \end{cases}$$

If the order of a cycle is even, the maximum distance between two vertices is $\frac{n}{2}$. Then, for a vertex u , the total distance from u to others equals to

$$D(u) = 1 + 2 + \dots + \frac{n}{2} + \dots + 2 + 1 = 2[1 + 2 + \dots + \frac{n}{2} - 1] + \frac{n}{2} = \frac{n^2}{4}.$$

If the order is even, then marginal entropy of cycles is obtained in the following formula.

$$I_D(C_n) = - \sum_{u \in V(G)} \frac{D(u)}{2W(C_n)} \log\left(\frac{D(u)}{2W(C_n)}\right) = -n \frac{\frac{n^2}{4}}{2 \times \frac{n^3}{8}} \log\left(\frac{\frac{n^2}{4}}{2 \times \frac{n^3}{8}}\right) = -\log \frac{1}{n} = \log n.$$

If the order of a cycle is odd, the maximum distance between two vertices is $\frac{n-1}{2}$. Thus, for a vertex u , the total distance from u to others equals to

$$D(u) = 1 + 2 + \dots + \frac{n-1}{2} + \frac{n-1}{2} + \dots + 2 + 1 = 2[1 + 2 + \dots + \frac{n-1}{2}] = \frac{n^2-1}{4}.$$

If the order is odd, then the marginal entropy of cycles is computed by the following equation.

$$I_D(C_n) = - \sum_{u \in V(G)} \frac{D(u)}{2W(C_n)} \log\left(\frac{D(u)}{2W(C_n)}\right) = -n \frac{\frac{n^2-1}{4}}{2 \times \frac{n^3-n}{8}} \log\left(\frac{\frac{n^2-1}{4}}{2 \times \frac{n^3-n}{8}}\right) = -\log \frac{1}{n} = \log n.$$

The proof is completed by this equation.

We know that the cycles are vertex-transitive graphs, and we obtain that the marginal entropy of cycles equals to $\log n$ in the previous theorem. It was shown in the paper,^[10] if G is a vertex-transitive graph, for any two vertices u, v in $V(G)$, we have $D(u) = D(v)$. From this property and Definition 2.2, we immediately have the following formula

$$W(G) = \frac{n}{2} D(u)$$

for an arbitrary vertex $u \in V(G)$. We use this equation in the following theorem.

Theorem 3.5. If G is a vertex-transitive graph of order n , then the marginal entropy of G is presented by the following formula

$$I_b(G) = \log n.$$

Proof. If G is a vertex-transitive graph of order n , we know that $W(G) = \frac{n}{2} D(u)$. If we write this equation in the following formula, the marginal entropy of vertex-transitive graphs is presented as in the following phrase.

$$\begin{aligned} I_b(G) &= - \sum_{u \in V(G)} \frac{D(u)}{2W(G)} \log \left(\frac{D(u)}{2W(G)} \right) \\ &= -n \frac{D(u)}{nD(u)} \log \left(\frac{D(u)}{nD(u)} \right) \\ &= -\log \frac{1}{n} \\ &= \log n. \end{aligned}$$

The proof is completed with this result.

CONCLUSION

In this paper, the marginal entropies of paths, stars, double stars, cycles and vertex-transitive graphs are obtained. This paper is a contribution to the computation of information indices in graph theory. Furthermore, marginal entropy is a possible tool for measuring the uncertainty of different networks. Therefore, the results of the this paper may be used in the computation of marginal entropy of different networks.

Acknowledgment. I would like to thank the anonymous reviewers for their helpful comments that helped to improve the quality of this article.

REFERENCES

- [1] C. Shannon and W. Weaver, *Mathematical Theory of Communications*, University of Illinois, Urbana, **1949**.
- [2] N. Rashevsky, *Bull. Math. Biophys.* **1955**, *17*, 229–235. <https://doi.org/10.1007/BF02477860>
- [3] E. V. Konstantinova in *General Theory of Information Transfer and Combinatorics, Lecture Notes in Computer Science, vol. 4123* (Eds: R. Ahlswede, L. Baumer, N. Cai, H. Aydinian, V. Blinovsky, C. Deppe, H. Mashurian) Berlin, Springer **2006**, pp 831–852. https://doi.org/10.1007/11889342_53
- [4] A. H. Wiener, *J. Am. Chem. Soc.* **1947**, *69*, 17–20. <https://doi.org/10.1021/ja01193a005>
- [5] D. Bonchev, N. Trinajstić, *J. Chem. Phys.* **1977**, *38*, 4517–4533. <https://doi.org/10.1063/1.434593>
- [6] D. Bonchev, N. Trinajstić, *Int. J. Quantum Chem.* **1978**, *S12*, 293–303. <https://doi.org/10.1002/qua.560140823>
- [7] D. Bonchev, N. Trinajstić, *Intern. J. Quantum Chem.: Quantum Chem Symp.* **1982**, *16*, 463–480. <https://doi.org/10.1002/qua.560220845>
- [8] E. V. Konstantinova, *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 54–57. <https://doi.org/10.1021/ci9502461>
- [9] A. Mowshowitz, M. Dehmer, *Entropy* **2015**, *17*, 1054–1062. <https://doi.org/10.3390/e17031054>
- [10] M. Ghorbani, M. Dehmer, A. Mowshowitz, J. Tao, F. Emmert-Streib, *Symmetry* **2019**, *11*(8), 1013. <https://doi.org/10.3390/sym11081013>
- [11] M. Dehmer, A. Mowshowitz, *Information Sciences* **2011**, *181*, 57–78. <https://doi.org/10.1016/j.ins.2010.08.041>
- [12] D. Bonchev, *Croat. Chem. Acta* **2016**, *89*, 391–392. <https://doi.org/10.5562/cca3035>
- [13] E. V. Konstantinova, M. V. Diudea, *Croat. Chem. Acta* **2000**, *73*, 383–403.
- [14] Y. N. Neh, I. Gutman, *Discrete Mathematics* **1994**, *135*, 359–365. [https://doi.org/10.1016/0012-365X\(93\)E0092-I](https://doi.org/10.1016/0012-365X(93)E0092-I)