IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 13, 2019, accepted January 22, 2020, date of publication January 28, 2020, date of current version March 2, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2970059

The Orbit-Polynomial: A Novel Measure of Symmetry in Networks

MATTHIAS DEHMER[1,2,3], ZENGQIANG CHEN[2], (Member, IEEE), FRANK EMMERT-STREIB[4,5], ABBE MOWSHOWITZ[6], KURT VARMUZA[7], LIHUA FENG[8], HERBERT JODLBAUER[1], YONGTANG SHI[9], AND JIN TAO[10,11], (Member, IEEE)

[1]University of Applied Sciences Upper Austria, Steyr School of Management, 4400 Steyr, Austria
[2]College of Artificial Intelligence, Nankai University, Tianjin 300350, China
[3]Department of Biomedical Computer Science and Mechatronics, University for Health Sciences, Medical Informatics and Technology (UMIT), 6060 Hall in Tyrol, Austria
[4]Predictive Medicine and Data Analytics Lab, Department of Signal Processing, Tampere University of Technology, 33720 Tampere, Finland
[5]Institute of Biosciences and Medical Technology, 33520 Tampere, Finland
[6]Department of Computer Science, The City College of New York (CUNY), New York, NY 10031, USA
[7]Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, 1040 Vienna, Austria
[8]School of Mathematics and Statistics, Central South University, Changsha 410083, China
[9]Center for Combinatorics and LPMC, Nankai University, Tianjin 300071, China
[10]Department of Electrical Engineering and Automation, Aalto University, Espoo 02150, Finland
[11]College of Engineering, Peking University, Beijing 100871, China

Corresponding author: Matthias Dehmer (matthias.dehmer@umit.at)

The work of Matthias Dehmer was supported by the Austrian Science Funds under Project P30031. The work of Zengqiang Chen was supported by the National Natural Science Foundation of China under Grant 61573199. The work of Yongtang Shi was supported in part by the National Natural Science Foundation of China under Grant 11771221 and Grant 11811540390, in part by the Natural Science Foundation of Tianjin under Grant 17JCQNJC00300, in part by the China-Slovenia Bilateral Project under Grant 12-6, and in part by the Open Project Foundation of Intelligent Information Processing Key Laboratory of Shanxi Province under Grant CICIP2018005. The work of Jin Tao was supported by the Academy of Finland under Grant 315660.

ABSTRACT Research on the structural complexity of networks has produced many useful results in graph theory and applied disciplines such as engineering and data analysis. This paper is intended as a further contribution to this area of research. Here we focus on measures designed to compare graphs with respect to symmetry. We do this by means of a novel characteristic of a graph G, namely an "orbit polynomial." A typical term of this univariate polynomial is of the form cz^n, where c is the number of orbits of size n of the automorphism group of G. Subtracting the orbit polynomial from 1 results in another polynomial that has a unique positive root, which can serve as a relative measure of the symmetry of a graph. The magnitude of this root is indicative of symmetry and can thus be used to compare graphs with respect to that property. In what follows, we will prove several inequalities on the unique positive roots of orbit polynomials corresponding to different graphs, thus showing differences in symmetry. In addition, we present numerical results relating to several classes of graphs for the purpose of comparing the new symmetry measure with existing ones. Finally, it is applied to a set of isomers of the chemical compound adamantane $C_{10}H_{16}$. We believe that the measure can be quite useful for tackling applications in chemistry, bioinformatics, and structure-oriented drug design.

INDEX TERMS Quantitative graph theory, networks, symmetry, graphs, graph measures, data science.

I. INTRODUCTION

The structural analysis of graphs has been an active research topic for half a century, stimulated in large part by the work of Rashevsky [1] and Mowshowitz [2]. In recent decades, the immense influence of the Internet has focused attention on complex networks [3], [4]. The structure of such networks has been investigated by means of a variety of quantitative measures [5]. Features such as branching, symmetry, cyclicity, and connectedness have been used to analyze the structural complexity of graphs, see, e.g., [6]–[10]. Recent symmetry measures, such as Kolmogorov-Sinai entropy [11], based on the automorphism group of a graph, as well as other local and global symmetry measures [12] have been investigated.

The associate editor coordinating the review of this manuscript and approving it for publication was Walter Didimo.

In addition, various measures designed to compare or match graphs according to quantitative similarity, or the distance between them, have been studied extensively, see [13].

Measures that quantify indicators of graph complexity have been defined and applied in several problem areas, including structural chemistry, drug design, and computer programming. In particular, characteristics of chemical structures [14] and hypertext graphs [15] have been explored. Structural information about graphs is not easy to obtain. To capture symmetry or cyclicity properties, for example, it is necessary to craft measures that serve as proxies for these properties. One approach is to compare the values of a given measure with those of others whose structural interpretation is well defined. Mathematical techniques for making such comparisons might take account of correlation or information-theoretic distance like the Kullback-Leibler measures [16].

In this paper, we define a novel graph concept called the *orbit polynomial*, see Section (II-A). A typical term of this univariate polynomial is of the form cz^n, where c is the number of orbits (of the automorphism group of G) of size n. The coefficients are all positive, so subtracting the orbit polynomial from 1 results in a related polynomial that has a unique positive root (denoted by δ), which can serve as a relative measure of the symmetry of a graph. The term ''relative'' signifies the use of the measure to compare two graphs G_1 and G_2 with respect to symmetry. If $\delta(G_1) > \delta(G_2)$, it is reasonable to surmise that G_1 is more symmetric than G_2 based on the numbers and sizes of their respective automorphism group orbits. Starting with a set of graphs $\mathcal{G} = \{G_1, G_2, \ldots, G_n\}$, this symmetry measure can be used to obtain a rank order

$$\delta(G_{[1]}) > \delta(G_{[2]}) > \cdots > \delta(G_{[n]}), \tag{1}$$

of the graphs according to their symmetry. This system of inequalities (1) defines a partial order on the set of graphs. The principal aim of this paper is to show the comparative symmetry of graphs in several different classes by means of their respective δ values. In particular, prove interrelations between well-known graph classes based on δ values. Moreover, we establish several properties of the roots δ, and compute correlations between our novel measure and other symmetry measures. The analysis shows that δ values clearly reflect the symmetry and the system of inequalities (1) induces a symmetry-based rank order on graphs. Computation of the orbit polynomial depends on knowledge of the vertex orbits of a graph, which may be found by first computing the automorphism group. In general, the computational complexity of determining the automorphism group is not known [17], but for simple graphs such as trees, there exist polynomial algorithms to calculate Aut(G).

Since the orbit polynomial offers a new way of measuring symmetry. It is important to explain the motivation for introducing what is genuinely novel about it. Clearly, the entropy-based symmetry measure I_a due to Mowshowitz [2] is also based on the sizes of vertex orbits. More precisely, I_a is the entropy of a finite probability scheme with values p_i [2] based on the relative size of the i-th orbit. That is to say, $I_a := -\sum p_i \log(p_i)$, which provides a quantitative measure of the symmetry of a graph, see [2]. Our approach is different in that we define a polynomial as described above. Therefore, the Equations (2), (20) are graph polynomials and we claim that the zero δ is a novel and useful symmetry measure. Finally, our measure is an algebraic quantity inasmuch as it represents a zero of a graph polynomial. As I_a and the Equations (2), (20) are both based on orbit sizes, it is clear that they are somehow related. However, an examination of the relationship between the two measures for certain classes of graphs is beyond the scope of this paper.

In this paper, we have used the `Nauty`-package due to McKay [18], [19], which calculates the automorphism group as well as the orbits of our graphs efficiently. Note that we calculated the orbit polynomial on exhaustively generated trees; Colbourn and Booth [20] developed algorithms with linear time complexity to determine the automorphism group of trees and related graph classes. So, Aut(G), as well as the vertex orbits, are needed to infer the orbit polynomial, which can be calculated here efficiently. The orbit polynomial can be computed in linear time once the orbits and their sizes are known. In order to compute or approximate δ, a standard algorithm with average time complexity $\log_2(n)$ can be used, see [21], [22].

II. METHODS AND RESULTS

In this section, we define the *orbit polynomial* and derive some of its properties. Note that this polynomial, whose coefficients derive from the automorphism group of the graph, can be said to represent the structure of the graph [23]. We will derive closed forms for special orbit polynomials and prove some properties related to their roots. In addition, we generate and interpret numerical results on the correlation between the novel symmetry measure and other such measures, see Section (II-D).

A. THE ORBIT-POLYNOMIAL AND SPECIAL EXPRESSIONS

The focus of this paper is to use the orbits [24] of the automorphism group defined on the vertices of a graph to define a novel graph polynomial. An automorphism is an edge-preserving bijection of the vertices of a graph [24]. The set of automorphisms under the composition of mappings forms the automorphism group of the graph and is usually denoted by Aut(G), consisting of $|\mathrm{Aut}(G)|$ elements. The equivalence classes of the vertices of a graph under the action of the automorphisms are called vertex orbits [24].

Now, let $G = (V, E)$ be a graph with $|V| < \infty$ and let $V_1, V_2, \ldots, V_\rho$ be its vertex orbits, where ρ is the total number of vertex orbits of G. Let k be the number of different sizes among the orbits, and suppose the number of orbits of size i_j is a_{i_j} for $1 \leq j \leq k$, so that $\sum_{j=1}^{k} i_j a(i_j) = |V|$. $\sum_{j=1}^{k} a(i_j) = \rho$. Now, we state the definition of the orbit polynomial.

Definition 1: The orbit polynomial of G is defined as follows:

$$O_G(z) := \sum_{j=1}^{k} a_{i_j} z^{i_j}. \tag{2}$$

Definition (1) is also valid for directed graphs and graphs with weights associated with vertices or edges. The automorphism group is defined for such graphs as an obvious extension of the case for simple undirected graphs, and orbits can be determined as equivalence classes of vertices under the actions of the automorphisms [25], [26]. For example, the orbit polynomial can be calculated for a molecular graph (see Section (II-E)) with so-called hetero atoms and multiple bonds [26].

By definition, $O_G(z)$ is a real polynomial, but its zeros can be complex-valued as well. By applying the Descartes' rule of signs [27], we see that Equation (2) does not possess any positive zeros as there are no sign changes in its sequence of coefficients. With the substitution $z := -z$ in Equation (2), Descartes' rule of signs can also be used to determine the number of negative zeros. In this paper, we are only concerned with the positive zeros of our novel graph polynomial.

For example, consider the path graph P_5 with five vertices numbered consecutively. The vertex orbits are $\{1, 5\}$, $\{2, 4\}$ and $\{3\}$. This gives $O_{P_5}(z) = 2z^2 + 1z$.

In the following, we establish explicit expressions for special orbit polynomials for simple graph classes such as paths, stars, and other branched trees. The reason for considering these graph classes stems from the fact that these structures have been proven useful in many disciplines like chemistry and bioinformatics, see [26], [28], [29]. They often act as building blocks to understand network-based complex systems. Also, linear and branched trees such as linear and branched alkanes play a major role in chemistry, drug design, and related disciplines [30], [31]. Hence it's important to characterize them using quantitative measures [14], [30], [31] such as δ and to prove interrelations between several classes of graphs.

Theorem 2: Let P_n be the path graph with n vertices. Then

$$O_{P_n}(z) := \frac{n}{2} z^2 \quad (n \text{ is even}), \tag{3}$$

$$O_{P_n}(z) := \left(\left\lceil \frac{n}{2} \right\rceil - 1\right) z^2 + z \quad (n \text{ is odd}). \tag{4}$$

Proof: For simplicity, let n be even. Numbering the vertices of P_n consecutively (from left to right), the set of vertex orbits is

$$S_{P_n} = \left\{\{1, n\}, \{2, n-1\}, \ldots, \left\{n - \frac{n}{2}, n - \frac{n}{2} + 1\right\}\right\}. \tag{5}$$

Hence $|S_{P_n}| = \frac{n}{2}$, if n is even, which establishes Equation (3). If n is odd, we obtain

$$S_{P_n} = \{\{1, n\}, \{2, n-1\}, \ldots, \{n - \frac{n-1}{2} - 1, n - \frac{n-1}{2} + 1\}, \{n - \frac{n-1}{2}\}\}, \tag{6}$$

and $|S_{P_n}| = \left\lceil \frac{n}{2} \right\rceil$. Also, we see that there exist $\left\lceil \frac{n}{2} \right\rceil - 1$ vertex orbits of size two and only one with a single element. Hence, we obtain Equation (4). □

Theorem 3: Let $P_n^{b_1}$ be the first branched path with n vertices. We obtain

$$O_{P_n^{b_1}}(z) := z^2 + (n-2)z. \tag{7}$$

Proof: Consider the branched path $P_n^{b_1}$ labeled consecutively, see Figure (1). For it's orbit set, we obtain

$$S_{P_n^{b_1}} = \{\{1, 3\}, \{2\}, \ldots, \{n-1\}, \{n\}\}. \tag{8}$$

This implies $|S_{P_n^{b_1}}| = n - 1$. So, there exist $n - 2$ orbits of size one and only one of size two. From this, we obtain Equation (7). □

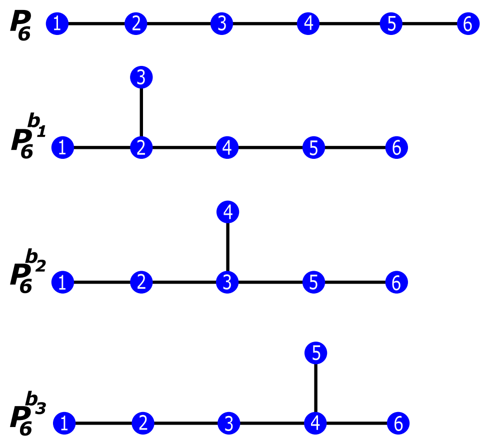


FIGURE 1. Path graph P_6 and branched paths $P_n^{b_1}$, $P_n^{b_2}$, and $P_n^{b_3}$ (see also [32]). Note that, $P_n^{b_i}$ is the tree formed by attaching an end vertex to an 'inner' vertex on the original path. Here, we have $n = 6$ and add an edge to the first, second, and third inner vertex. The remaining vertices are end vertices.

Theorem 4: Let S_n be the star graph with n vertices. We obtain

$$O_{S_n}(z) = z^{n-1} + z \tag{9}$$

Proof: Consider the star graph with n vertices and denote the hub vertex by 1 and the remaining ones by $2, 3, \ldots, n$, respectively. The set of vertex orbits of S_n is

$$S_{S_n} = \{\{1\}, \{2, 3, \ldots, n\}\}. \tag{10}$$

There are two orbits, one consisting of the hub alone, and the other containing the remaining vertices. □

B. ANALYTICAL RESULTS

In this section, we define another graph polynomial based on the orbit polynomial represented by Equation (2).

Definition 5: We define the graph polynomial

$$O_G^{\star}(z) := 1 - O_G(z). \tag{11}$$

Now we investigate the positive roots of this polynomial for the purpose of obtaining information about the symmetry of a graph. It is important to note that the aim of this paper is to establish the existence of roots and provide general bounds and inequalities for them. It is not our purpose here to give precise bounds, and therefore we do not calculate roots numerically using iterative methods, see [33].

Before proving more general results, we state two lemmas, which give the values of δ for vertex-transitive and asymmetric graphs. First, the following general lemma is needed for subsequent development.

Lemma 6: Let $H(z)$ be an arbitrary real polynomial defined by

$$H(z) = a_k z^k + a_{k-1} z^{k-1} + \cdots + a_0, \quad a_i > 0,\ 1 \leq i \leq k. \tag{12}$$

The polynomial $H^\star(z) := 1 - H(z)$ has a unique positive zero $\delta < 1$ if $a_0 < 1$ and

$$a_k + a_{k-1} + \cdots + a_0 > 1. \tag{13}$$

Proof: By definition

$$H^\star(z) \coloneqq 1 - H(z) = 1 - (a_k z^k + a_{k-1} z^{k-1} + \cdots + a_0). \tag{14}$$

Since $H^\star(z)$ has only one sign change in its sequence of coefficients, we conclude from Descartes' rule of signs that $H^\star(z)$ has a unique positive zero. Moreover $H^\star(0) = 1 - a_0$. The case $1 - a_0 > 0$ is equivalent to $a_0 < 1$. Let δ be the unique positive zero of $H^\star(z)$. We obtain $\delta < 1$ if

$$H^\star(1) := 1 - H(z) = 1 - (a_k + a_{k-1} + \cdots + a_0) < 0. \tag{15}$$

This inequality is equivalent to inequality (13). The case of $H^\star(0) = 1 - a_0 < 0$ would lead to a contradiction namely $H^\star(0) < 0$. This follows from the fact that $\lim_{z\to\infty} H^\star(z) = -\infty$ implies that $H^\star(z)$ has more than one or no positive zero, which contradicts the finding above that $H^\star(z)$ has a unique positive zero δ. □

Lemma 7: $\delta(G) = 1$ if and only if G is vertex-transitive.

Proof: Let $G = (V, E)$ be a graph and $|V| = n$. According to the Definitions (5), (1), we obtain

$$O_G^\star(z) := 1 - O_G(z) = 1 - \left(a_{i_1} z^{i_1} + a_{i_2} z^{i_2} + \cdots + a_{i_k} z^{i_k}\right). \tag{16}$$

We assume that $\delta(G) > 0$ is a zero of $O_G^\star(z)$. Now,

$$O_G^\star(\delta) = 1 - \left(a_{i_1} \delta^{i_1} + a_{i_2} \delta^{i_2} + \cdots + a_{i_k} \delta^{i_k}\right). \tag{17}$$

Thus, $\lim_{\delta\to\infty} O_G^\star(\delta) = -\infty$ and $O_G^\star(0) = 1$. For a general graph G with given vertex orbits, Equation (18) allows us to conclude that

$$O_G^\star(1) = 1 - \left(a_{i_1} + a_{i_2} + \cdots + a_{i_k}\right) \leq 0. \tag{18}$$

This means that $\delta(G) < 1$. To demonstrate that $\delta(G) = 1$, we need only observe the a_{i_j} must be natural numbers and set $a_{i_1} = 1$ and $a_{i_2} = \cdots a_{i_k} = 0$. Hence,

$$O_G^\star(\delta) = 1 - \delta^n. \tag{19}$$

So, $O_G^\star(\delta) = 0$ gives $\delta = 1$. Finally, note that the specification of the coefficients given above requires having all n vertices in the same orbit, which implies the graph is vertex transitive.

Now suppose that G is vertex-transitive. This means $\{v_1, \ldots, v_n\}$ is the only vertex orbit of G. From Definitions (5), (1), it follows that $O_G^\star(z) = 1 - z^n$. Hence, $O_G^\star(z) = 0$ gives $z = \delta = 1$. □

Lemma 8: Let $G = (V, E)$ be a graph of order n. If G has the trivial automorphism group, then $\delta(G) = \frac{1}{n}$.

Proof: If G has the trivial automorphism group, all vertex orbits are singleton sets. So, we obtain $\{v_1\}, \{v_2\}, \ldots, \{v_n\}$ and, finally, $O_G^\star(z) = 1 - nz$. $O_G^\star(z) = 0$ gives $z = \delta = \frac{1}{n}$. □

Lemma (7) shows vertex-transitive graphs attain the maximum value of δ. The more asymmetric the graph is, the smaller is the value of δ which tends to $\delta = \frac{1}{n}$. Now we are able to prove the following statement.

Theorem 9: The graph polynomial

$$O_G^\star(z) := 1 - O_G(z), \tag{20}$$

has a unique positive zero $\delta < 1$.

Proof:

$$O_G^\star(z) := 1 - O_G(z) = 1 - (a_k z^k + a_{k-1} z^{k-1} + \cdots + a_1 z). \tag{21}$$

Definition (1) implies $a_0 = 0$. We see that $O_G^\star(0) = 1$ and claim

$$O_G^\star(1) = 1 - (a_k + a_{k-1} + \cdots + a_1) < 0. \tag{22}$$

But if $k > 1$, $a_k, \ldots, a_1 \neq 0$, inequality (22) is satisfied. From Lemma (6), we conclude that Equation (21) has a unique, positive root which we denote by δ. □

1) PROPERTIES AND RELATIONS INVOLVING δ

In this section, we prove properties and relations between the unique, positive roots of polynomials $O_G^\star(z) := 1 - O_G(z)$ corresponding to various graphs. To this end, we state a general result.

Theorem 10: Let $G = (V, E)$ be a graph and $|V| = n$ which is not vertex-transitive. Then, $\delta(G) \geq \frac{1}{n}$.

Proof: If G is not vertex-transitive, it is clear that there exist vertex orbits $V_1, V_2, \ldots, V_\rho$, where $\rho > 1$ is the total number of vertex orbits of G. Again, k denotes the number of different sizes among the orbits; and the number of orbits of size i_j is denoted by a_{i_j} for $1 \leq j \leq k$, so that $\sum_{j=1}^{k} i_j a(i_j) = |V| = n$. Taking account of equation (16), consider

$$\begin{aligned} O_G^\star\left(\frac{1}{n}\right) &:= 1 - O_G(z) \\ &= 1 - \left[a_{i_1}\left(\frac{1}{n}\right)^{i_1} + a_{i_2}\left(\frac{1}{n}\right)^{i_2} + \cdots + a_{i_k}\left(\frac{1}{n}\right)^{i_k}\right]. \end{aligned} \tag{23}$$

Assuminge $\delta(G) > \frac{1}{n}$, we need to investigate the inequality

$$O_G^\star\left(\frac{1}{n}\right) = 1 - \left[a_{i_1}\left(\frac{1}{n}\right)^{i_1} + a_{i_2}\left(\frac{1}{n}\right)^{i_2} + \cdots + a_{i_k}\left(\frac{1}{n}\right)^{i_k}\right] < 0, \tag{24}$$

or

$$-\frac{a_{i_1}}{n^{i_1}} - \frac{a_{i_2}}{n^{i_2}} - \cdots - \frac{a_{i_k}}{n^{i_k}} + 1 < 0. \tag{25}$$

From Inequality (25), we see that

$$\frac{a_{i_1}}{n^{i_1}} + \frac{a_{i_2}}{n^{i_2}} + \cdots + \frac{a_{i_k}}{n^{i_k}} - 1 > 0. \tag{26}$$

Now, reordering the first k-terms of inequality (26), and assuming $i_{[1]} \geq i_{[2]} > \cdots i_{[k]}$, we obtain,

$$\frac{a_{i_1}}{n^{i_{[1]}}} + \frac{a_{i_2}}{n^{i_{[2]}}} + \cdots + \frac{a_{i_k}}{n^{i_{[k]}}} - 1 > 0. \tag{27}$$

Multiplying inequality (27) by $n^{i_{[1]}}$ gives

$$a_{i_1} + a_{i_2} n^{i_{[1]}-i_{[2]}} + \cdots + a_{i_k} n^{i_{[1]}-i_{[k]}} - n^{i_{[1]}} > 0. \tag{28}$$

Clearly, Inequality (28) leads to a contradiction for large values of n. Hence, the statement of the theorem holds. According to Lemma (8), $\delta = \frac{1}{n}$ only for the graph with the trivial automorphism group. □

Now, we continue our analysis by proving interrelations between δ values for special graph classes.

Theorem 11: Let $O_{P_n}(z)$ and $O_{P_n^{b_1}}(z)$ be the orbit polynomials of P_n and $P_n^{b_1}$, respectively; and denote the unique, positive roots of $O^{\star}_{P_n}(z)$ and $O^{\star}_{P_n^{b_1}}(z)$ by $\delta(P_n)$ and $\delta(P_n^{b_1})$, respectively. The inequality

$$\delta(P_n) > \delta(P_n^{b_1}), \tag{29}$$

is satisfied.

Proof: Suppose n is even. We need to determine the unique positive root $\delta(P_n)$ of the equation

$$O^{\star}_{P_n}(z) = 1 - \frac{n}{2}z^2 = 0, \tag{30}$$

and find $\delta(P_n) = \sqrt{\frac{2}{n}}$. Note that the symmetry of the path graph, measured by δ decreases as the length n of the path increases. This results from the fact that as n increases, $\sqrt{\frac{2}{n}}$ decreases. This relationship also holds for the branched trees defined above.

Similarly, determining the unique positive root $\delta(P_n^{b_1})$ of the equation

$$O_{P_n^{b_1}}(z) = 1 - [z^2 + (n-2)z] = 0, \tag{31}$$

produces the two candidates

$$z_{1,2} = -\frac{n-2}{2} \pm \sqrt{\left(\frac{n-2}{2}\right)^2 + 1}. \tag{32}$$

So, we set

$$\delta(P_n^{b_1}) := -\frac{n-2}{2} + \sqrt{\left(\frac{n-2}{2}\right)^2 + 1}, \tag{33}$$

and assume $\delta(P_n^{b_1}) > 0$. According to Inequality (29), we have to show that

$$\sqrt{\frac{2}{n}} > -\frac{n-2}{2} + \sqrt{\left(\frac{n-2}{2}\right)^2 + 1}. \tag{34}$$

However, direct calculation shows that the last inequality is satisfied if $n > 2$. If n is odd we first solve (see Equation (4)),

$$O^{\star}_{P_n}(z) = 1 - \left[\left(\left\lceil\frac{n}{2}\right\rceil - 1\right) z^2 + z\right] = 0. \tag{35}$$

Clearly,

$$\delta(P_n) = -\frac{1}{2\left(\left\lceil\frac{n}{2}\right\rceil - 1\right)} + \sqrt{\left(\frac{1}{2\left(\left\lceil\frac{n}{2}\right\rceil - 1\right)}\right)^2 + \frac{1}{\left(\left\lceil\frac{n}{2}\right\rceil - 1\right)}}. \tag{36}$$

For n is odd, Inequality (29) implies

$$-\frac{1}{2\left(\left\lceil\frac{n}{2}\right\rceil - 1\right)} + \sqrt{\left(\frac{1}{2\left(\left\lceil\frac{n}{2}\right\rceil - 1\right)}\right)^2 + \frac{1}{\left(\left\lceil\frac{n}{2}\right\rceil - 1\right)}} > -\frac{n-2}{2} + \sqrt{\left(\frac{n-2}{2}\right)^2 + 1}. \tag{37}$$

Direct calculation shows that Inequality (37) is satisfied if $n \geq 4$. Note that $P_4^{b_1}$ is the smallest branched path, so $n = 4$. □

In the following, we prove properties of the unique, positive root $\delta(S_n)$ of the equation $O^{\star}_{S_n}(z) = 0$ where S_n is the star graph with n vertices. We compute $\delta(S_4)$ using Cardano's formula [34], and obtain

$$\delta(S_4) = \sqrt[3]{\frac{1}{2} + \sqrt{\frac{31}{108}}} - \sqrt[3]{\sqrt{\frac{31}{108}} - \frac{1}{2}} \doteq 0.6823. \tag{38}$$

Now we prove our main result for the unique, positive root of the star graph S_n.

Theorem 12: Let S_n be the star graph with n vertices, $n > 4$, and $\delta(S_n)$ the unique, positive root of

$$O^{\star}_{S_n}(z) = 1 - O_{S_n}(z) = 1 - (z^{n-1} + z) = 0. \tag{39}$$

Then

$$\delta(S_n) \in (\delta(S_4), 1), \quad n > 4. \tag{40}$$

The value of $\delta(S_4)$ is given by Equation (38).

Proof: We know that $O^{\star}_{S_4}(\delta(S_4)) = 0$. Also, Theorem (9) implies $\delta(S_n) < 1$. To conclude the proof of the theorem, we first show that $O^{\star}_{S_n}(z)$ is strictly decreasing. Assuming

$$(O^{\star}_{S_n}(z))' = -(n-1)z^{n-2} - 1 < 0, \tag{41}$$

we now consider

$$z^{n-2} > -\frac{1}{n-1}. \tag{42}$$

Inequality (42) is satisfied if $z > 0$. In our case, this is satisfied as we only consider positive zeros. Hence, $O^{\star}_{S_n}(z)$ is strictly decreasing. Now, taking $\delta(S_4)$ given by Equation (38) as a lower bound, we show

$$O^{\star}_{S_n}(\delta(S_4)) > 0, \tag{43}$$

if $n > 4$. So,

$$O^{\star}_{S_n}(\delta(S_4)) = 1 - (\delta(S_4))^{n-1} - \delta(S_4) > 0, \tag{44}$$

leads immediately to

$$\ln(1 - \delta(S_4)) > (n-1)\ln(\delta(S_4)), \tag{45}$$

or

$$n > \frac{\ln(1 - \delta(S_4))}{\ln(\delta(S_4))} + 1. \tag{46}$$

If we now plug in the value of $\delta(S_4)$ given by Equation (38), Inequality (46) gives $n > 4$. Thus we have proven that $1 > \delta(S_n) > \delta(S_4)$, if $n > 4$ and Relation (40). □

Lemma (7) states that $\delta(G) = 1$ if and only if G is vertex-transitive. Now, let's consider the class of star graphs with n vertices. We are able to show that $\delta(S_n)$ goes to 1 in the limiting case. According to Theorem (12), we obtain $\delta(S_n) \in (0.6823, 1)$ for $n > 4$. As n increases, $\delta(S_n)$ tends to one. Starting from $O^{\star}_{S_n}(z) = 1-(z^{n-1}+z)$ and by fixing $\delta \in (0, 1)$, we obtain $\lim_{n\to\infty} O^{\star}_{S_n}(\delta) = \lim_{n\to\infty}[1-(\delta^{n-1}+\delta)] = 1-\delta$. So, $1 - \delta = 0$ implies $\delta = 1$. In summary, $\delta(S_n)$ tends to one if n gets sufficiently large. In summary, we have shown

Corollary 13: Let S_n be the star graph with n vertices. $\delta(S_n) \in (0.6823, 1)$ is the unique, positive zero of $O^{\star}_{S_n}(z) = 1 - (z^{n-1} + z)$. $\delta(S_n)$ tends to one if n gets sufficiently large.

Next we turn to edge-transitive graphs. In particular we consider complete bipartite graphs $K_{s,t}$ for $s \neq t$. These graphs contain two vertex orbits. The corresponding polynomial is given by

$$O^{\star}_{K_{s,t}}(z) = 1 - z^s - z^t. \tag{47}$$

The above statement allows for studying the behavior of $\delta(K_{s,t})$, if t is taken to vary while s is fixed but arbitrary.

Theorem 14: Let $K_{s,t}$, $s \neq t$ be a complete bipartite graph. $\delta_1(K_{s,t})$ be the zero of the polynomial given by Equation (47) with fixed s and t. $\delta_2(K_{s,t_2})$ be the zero of polynomial

$$O^{\star}_{K_{s,t_2}}(z) = 1 - z^s - z^{t_2}. \tag{48}$$

Then

$$\delta_2(K_{s,t_2}) > \delta_1(K_{s,t}), \tag{49}$$

and

$$t_2 > \frac{\ln(1 - (\delta_1(K_{s,t})^s)}{\ln(\delta_1(K_{s,t})}. \tag{50}$$

Proof: Fix s and t with $s \neq t$. By assumption $\delta_1(K_{s,t})$ is the zero of $O^{\star}_{K_{s,t}}(z)$, so we get

$$1 - \delta_1(K_{s,t})^s - \delta_1(K_{s,t})^t = 0. \tag{51}$$

First of all, observe that $O^{\star}_{K_{s,t}}(z)$ is strictly decreasing for $z > 0$ as

$$O^{\star}_{K_{s,t}}(z)' = -sz^{s-1} - tz^{t-1} < 0. \tag{52}$$

Finally, we infer $\delta_2(K_{s,t_2}) > \delta_1(K_{s,t})$ if

$$O^{\star}_{K_{s,t_2}}(\delta_1(K_{s,t})) = 1 - (\delta_1(K_{s,t}))^s - (\delta_1(K_{s,t}))^{t_2} > 0. \tag{53}$$

But this yields inequality (50). □

To illustrate Theorem (14), we consider the edge-transitive graph $K_{2,3}$ and note that

$$O^{\star}_{K_{2,3}}(z) = 1 - z^2 - z^3 = 0. \tag{54}$$

Solving this equation numerically gives $\delta_1(K_{2,3}) \doteq 0.754878$. Now, from inequality

$$t_2 > \frac{\ln(1 - (0.754878)^2)}{\ln(0.754878)}, \tag{55}$$

we see that $t_2 > 3$. So, we choose $t_2 = 4$ and now consider $K_{2,4}$. Finally, calculating the root of

$$O^{\star}_{K_{2,4}}(z) = 1 - z^2 - z^4 = 0, \tag{56}$$

gives $\delta_2(K_{2,4}) \doteq 0.786151$. So, $|\delta_2(K_{2,3}) - \delta_2(K_{2,4})| = 0.031273$ and we obtain $\delta_2(K_{2,4}) > \delta_1(K_{2,3})$. Furthermore, the graph $K_{2,20}$ has root $\delta_2(K_{2,20}) \doteq 0.913827$. Applying the same argument as in Corollary (13), we conclude that the zero $\delta_2(K_{s,t})$ tends to one if s, t gets sufficiently large.

Finally we examine the roots associated with graphs of the form $K_{s,t}$.

Corollary 15: $z = -\frac{1}{2} + \frac{\sqrt{5}}{2}$ is the smallest zero of $O^{\star}_{K_{s,t}}(z)$ if $s = 1$ and $t = 2$. It holds $\delta(K_{s,t}) > -\frac{1}{2} + \frac{\sqrt{5}}{2} \doteq 0.61803$, $s > 1, t > 2$.

Proof: We obtain $z = -\frac{1}{2}+\frac{\sqrt{5}}{2}$ as the root of the equation

$$O^{\star}_{K_{1,2}}(z) = 1 - z - z^2 = 0, \tag{57}$$

by direct calculation. To prove that $\delta(K_{s,t}) > -\frac{1}{2} + \frac{\sqrt{5}}{2}$, we need to show

$$O^{\star}_{K_{s,t}}\left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right)=1-\left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right)^s-\left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right)^t>0, \tag{58}$$

which, according Equation (52), says that $O^{\star}_{K_{s,t}}(z)$ is strictly decreasing for $z > 0$. However, we consider instead

$$\left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right)^s+\left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right)^t < 1. \tag{59}$$

and give a proof by induction over t. Let s be fixed but arbitrary. We start the induction by setting $t = 3$. This yields

$$\left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right)^s+\left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right)^3 < 1. \tag{60}$$

If $s = 2$, we obtain $\left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right)^2 + \left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right)^3 \doteq 0.61803 < 1$. This relation is still valid for $s > 2$. By the induction hypothesis, equation (59) is satisfied, but it is easier to examine

$$\left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right)^t < 1-\left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right)^s. \tag{61}$$

We conclude by using the induction hypothesis expressed in inequality (61)

$$\begin{aligned}\left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right)^{t+1} &= \left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right)^{t}\left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right)\\ &< \left(1-\left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right)^{s}\right)\left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right)\\ &< 1-\left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right)^{s}. \qquad (62)\end{aligned}$$

□

C. GENERAL INTERPRETATION OF δ

In this section, we elaborate briefly on interpretations of the measure δ. As already mentioned, $\delta = 1$ if and only if the corresponding graph is vertex-transitive. So, all graphs in this are identical from the standpoint of this measure. The result is independent of the size of a graph. For example, all complete graphs are transitive and thus have $\delta = 1$. If a graph is asymmetric, then $\delta = \frac{1}{n}$. Consequently, if a graph is not vertex-transitive, it is clear that $\delta \geq \frac{1}{n}$. Note also that all graphs with the same number of vertex orbits of the same sizes form an equivalence class. Graphs G_2, G_4 and G_5, G_7 offer two examples. Another equivalence class can be obtained by considering all graphs with the same positive root δ of $O^{\star}_G(z) = 1 - O_G(z)$. Any set of graphs with the same δ would constitute an equivalence class. Differences in the value of the measure become arbitrarily small as the size of the graph increases. This is a consequence of the monomials az^n of which the polynomial $1 - O_G(z)$ is composed.

D. NUMERICAL RESULTS

1) CALCULATING δ FOR SPECIAL GRAPHS

In [32], a set of test graphs was used to investigate cyclicity and branching of graphs [32] (see also [35]). These test graphs are used here to evaluate the symmetry measure δ, see the Figures (2) - (8) introduced in this paper. The positive zeros of the polynomials $O^{\star}_G(z) = 1 - O_G(z)$ have been calculated using the programming language R [36]. In particular, we have used the R-packages igraph [37] and Rmpfr [36], see also [32]. In order to calculate the vertex orbits of our test graphs, we again assume that they are labeled consecutively from left to right.

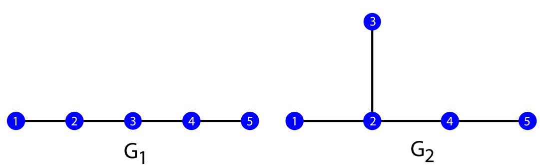


FIGURE 2. Left: $O^{\star}_{G_1}(z) = 1 - z - 2z^2$ and vertex orbits: {1, 5}, {2, 4}, {3}. $\delta(G_1) = \frac{1}{2}$. Right: $O^{\star}_{G_2}(z) = 1 - 3z - z^2$ and vertex orbits: {1, 3}, {2}, {4}, {5}. $\delta(G_2) \doteq 0.3028$.

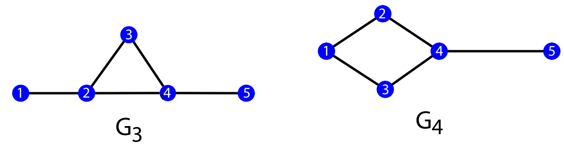


FIGURE 3. Left: $O^{\star}_{G_3}(z) = 1 - z - 2z^2$ and vertex orbits: {1, 5}, {2, 4}, {3}. $\delta(G_3) = \frac{1}{2}$. Right: $O^{\star}_{G_4}(z) = 1 - 3z - z^2$ and vertex orbits: {1}, {2, 3}, {4}, {5}. $\delta(G_4) \doteq 0.3028$.

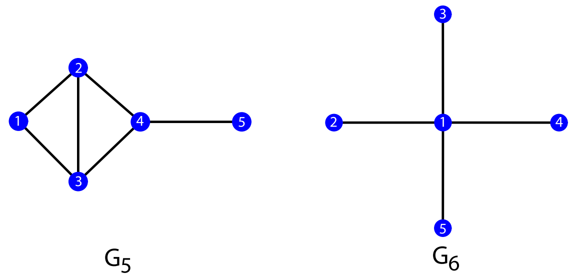


FIGURE 4. Left: $O^{\star}_{G_5}(z) = 1 - 3z - z^2$ and vertex orbits: {1}, {2, 3}, {4}, {5}. $\delta(G_5) \doteq 0.3028$. Right: $O^{\star}_{G_6}(z) = 1 - z - z^4$ and vertex orbits: {1}, {2, 3, 4, 5}. $\delta(G_6) \doteq 0.7245$.

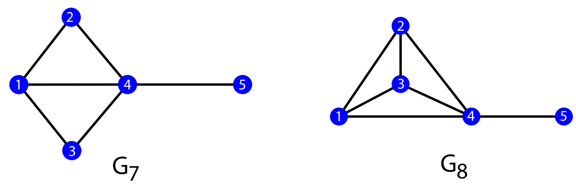


FIGURE 5. Left: $O^{\star}_{G_7}(z) = 1 - 3z - z^2$ and vertex orbits: {1}, {2, 3}, {4}, {5}. $\delta(G_7) \doteq 0.3028$. Right: $O^{\star}_{G_8}(z) = 1 - 2z - z^3$ and vertex orbits: {1, 2, 3}, {4}, {5}. $\delta(G_8) \doteq 0.4534$.

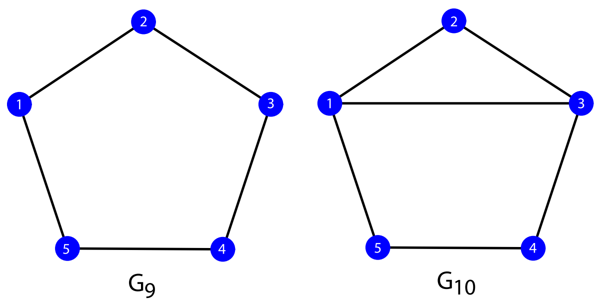


FIGURE 6. Left: $O^{\star}_{G_9}(z) = 1 - z^5$ and vertex orbits: {1, 2, 3, 4, 5}. $\delta(G_9) = 1$. Right: $O^{\star}_{G_{10}}(z) = 1 - z - 2z^2$ and vertex orbits: {1, 3}, {2}, {4, 5}. $\delta(G_{10}) = \frac{1}{2}$.

Now we interpret the results shown in Figures (2) - (8). For the graphs in these figures, the distance matrix D suffices to determine the orbits. The i, j-th entry of D is the number of vertices at distance j from vertex i. The maximum value of j is the diameter of the graph. Vertices in the graphs shown here with the same row values are similar. In general, having the same row values in the distance matrix is a necessary but not a sufficient condition for vertices to be similar (see [41])..

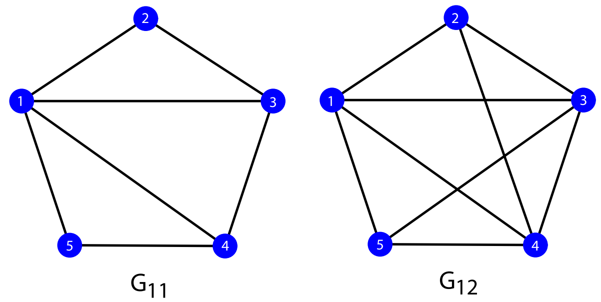


FIGURE 7. Left: $O^{\star}_{G_{11}}(z) = 1 - z - 2z^2$ and vertex orbits: $\{1\}, \{2, 5\}, \{3, 4\}$. $\delta(G_{11}) = \frac{1}{2}$. Right: $O^{\star}_{G_{12}}(z) = 1 - z - z^3$ and vertex orbits: $\{1, 3, 4\}, \{2, 5\}$. $\delta(G_{12}) \doteq 0.7245$.

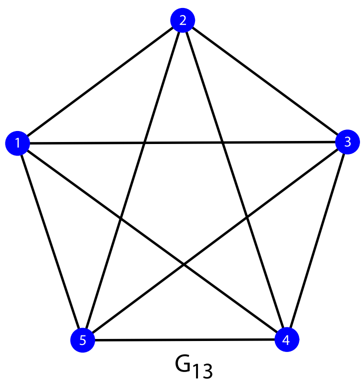


FIGURE 8. $O^{\star}_{G_{13}}(z) = 1 - z^5$ and vertex orbits: $\{1, 2, 3, 4, 5\}$. $\delta(G_{13}) = 1$.

All the graphs under consideration possess the same number of vertices. To begin we compare the results for G_1 and G_2. We see that $\delta(G_1) > \delta(G_2)$ which is expected from Theorem (11). So, it is clear that the set of vertex orbits of G_2 contains more singleton sets and, hence, G_2 is less symmetric than G_1. The extreme case of an asymmetric graph is one with the identity group, see [2], [38]. We proved in Lemma (8) that when G is asymmetric, $\delta = \frac{1}{n}$. Further, we observe that the measure δ fails to distinguish two different graphs with the same sets of vertex orbits. An example of this is $\delta(G_2) = \delta(G_4)$ and $\delta(G_9) = \delta(G_{13})$. Another case is G_7 and G_8. The vertex orbits of these two graphs (see Figure (5)), indicate that G_8 is more symmetric than G_7. This relation of symmetry is reflected by the inequality $\delta(G_8) > \delta(G_7)$. Note that G_9 and G_{13} have transitive groups, meaning there is only one vertex orbit with n elements. Hence, in this case $O^{\star}_G(z) = 1 - z^n$. According to Lemma (7), we obtain $\delta = 1$. Following Theorem (10), wee see that all other values of δ are less than one and $\delta(G) > \frac{1}{5}$. Finally, we showed by using our test graphs that the greater the value of δ, the more symmetric is the graph based on the vertex orbits.

By way of conclusion, we call attention to an important property of the orbit polynomial as well as of δ. The Definitions (1), (5) indicate clearly that the orbit polynomials $O_G(z)$, $O^{\star}_G(z)$ only depend on the vertex orbits of G and their size. Once the graph G and the vertex orbits are given, we may compute δ. So, we see that the Definitions (1), (5) do not rely on the size of the automorphism group. For instance, the graph G_9 depicted in Figure (6) has the same value of δ as a complete graph, which has a much larger automorphism group. We see that this characteristic is not captured by δ. We emphasize that there exists another example for this situation. The well-known entropy measure I_a due to Mowshowitz [2] given by Equation (63) is based on vertex orbits of a given graph to determine the partitions needed to compute the probability values and, finally, the entropy of a graph, see [2]. In fact, I_a does not depend on $|\mathrm{Aut}(G)|$. That's the reason why Mowshowitz and Dehmer [39] defined the so-called symmetry index S to overcome this problem, see Equation (64). So, S possesses another term to incorporate $|\mathrm{Aut}(G)|$. The reason for doing this was to lower the degeneracy of I_a. A measure is said to be degenerate if it fails to distinguish between non-isomorphic graphs. We aimed to improve the discriminating power of the measure [39]. Finally, similar measures could also be developed to relate δ to $|\mathrm{Aut}(G)|$. Lemma (7) states that not all vertex-transitive graphs can be distinguished by δ. Also, vertex-transitive graphs cannot be discriminated by I_a (see Equation (63)) as the measure vanishes. So, we can characterize vertex-transitive graphs (and others) in terms of equivalence classes as explained in Section (II-C); all graphs in a given equivalence class have the same value of δ. Similiarly, graphs can be characterized by I_a in terms of other equivalence classes, where all graphs in a given class have the same value of I_a. This is not a shortcoming of the since we did not set out to produce a measure with zero degeneracy. Had we aimed to do this, we would have had to find a complete set of graph invariants, see [40]. The degeneracy problem can be overcome, for all practical purposes, by constructing a so-called super index [40], [41] that combines several measures based on the same or different graph invariants, thus producing a measure with high discrimination power.

2) CORRELATION ANALYSIS

In this section, we investigate the correlation between δ and other known symmetry measures for graphs, see [42]. To perform our analysis, we take two other known symmetry measures for graphs into consideration namely [2], [38], [39]

$$I_a(G) = \sum_{i=1}^{k} \frac{|V_i|}{|V|} \log\left(\frac{|V_i|}{|V|}\right), \tag{63}$$

and

$$S(G) = [\log(|V|) - I_a(G)] + \log(\mathrm{Aut}(G)). \tag{64}$$

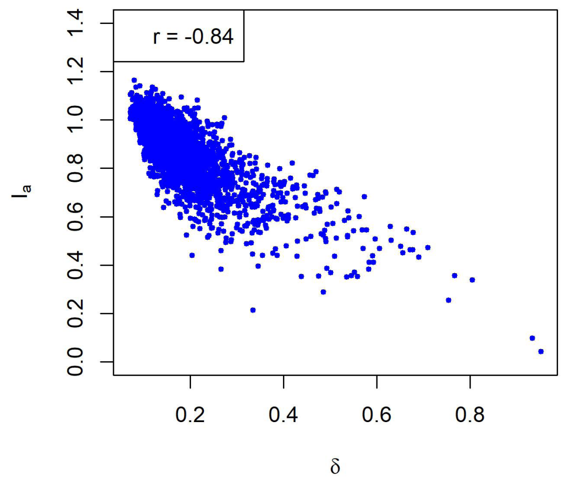
I_a is the well-known topological information content [2] of graphs representing a graph entropy measure. Let $G = (V, E)$ be a graph, where $|V_i|$ the size of the i-th vertex orbit of $\mathrm{Aut}(G)$, and k is the number of different orbits. We note that the symmetry index S [39] has been defined similarly but takes the size of the automorphism group $\mathrm{Aut}(G)$ into account. Evidently, I_a and S are graph entropy measures based on Shannon's entropy [43]. The probability values to define the entropy are inferred from vertex partitions defined by the vertex orbits, see [2], [38], [39]. Thus, the difference between I_a, S and δ is immediate; δ is an algebraic graph measure as it is a zero of a polynomial where the polynomial coefficients rely on the vertex orbits. So, it would be fruitful to show interrelations between the two measures.

TABLE 1. Pearson correlation coefficient between δ and I_a as well as δ and S on exhaustively generated trees.

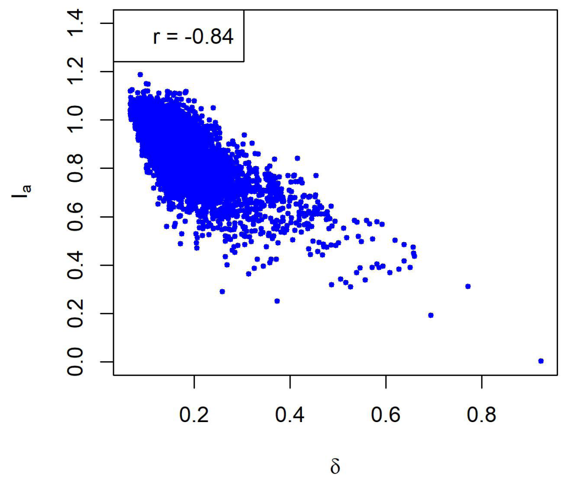
T_i	$r(\delta, S)$	$r(\delta, I_a)$
T_{14}	0.596	-0.840
T_{15}	0.628	-0.840
T_{16}	0.590	-0.821
T_{17}	0.622	-0.829
T_{18}	0.605	-0.818

Table (1) shows the results from calculating the Pearson correlation coefficient r [44]. Before interpreting the results, note that all graph measures have been applied on exhaustively generated trees. We generated these trees using `Nauty` [18], [19]. The sizes of these tree classes are $|T_{14}| = 3159$, $|T_{15}| = 7741$, $|T_{16}| = 19320$, $|T_{17}| = 48629$, $|T_{18}| = 123757$, $|T_{19}| = 317955$, respectively. The reason for using trees in the analysis here is that they are useful for many concrete applications in disciplines such as chemistry, bioinformatics and computer science. That means we have used graphs with a distinct topology rather than random graphs whose applicability is quite limited. To make sure that our results are meaningful, we have generated the graphs exhaustively to generalize the findings as much as possible.

Table (1), shows that the symmetry index representing Equation (64) is very little correlated with δ. T_i is the class of all pairwise non-isomorphic trees with i vertices. Note that $\delta \leq 1$ is the unique, positive root of $O^{\star}_T(z) = 1 - O_T(z)$, $T \in T_i$, $14 \leq i \leq 18$, see Section (II-B). Recall that δ is an algebraic measure for which the symmetry property of an underlying tree is given by the orbit polynomial. By contrast, S and I_a are information-theoretic graph complexity measures based on Shannon's entropy and they rely on partitions representing vertex orbits [2], [38]. Given the fact that all the tree measures are based on vertex orbits, we see clear differences in Table (1). First, we observe in general that δ and S, I_a belong to different categories of topological graph measures (information-theoretic vs. algebraic). When comparing δ and S, we also see that S depends on $|\mathrm{Aut}(G)|$. Especially the latter property might be the reason for the weak correlation between δ and S shown in Table (1). Nevertheless δ and I_a belong to different categories, we see by Table (1) that the correlation between them is stronger than in the previous case; here we just neglect the sign of r. This seems plausible as the two measures have the same input namely just the vertex orbits, and I_a does not rely on $|\mathrm{Aut}(G)|$. The scatter plots represented by the Figures (9) - (10) showing the relation between δ and S indicate a rather weak correlation.



(a) Correlation between δ and I_a based on T_{14}.



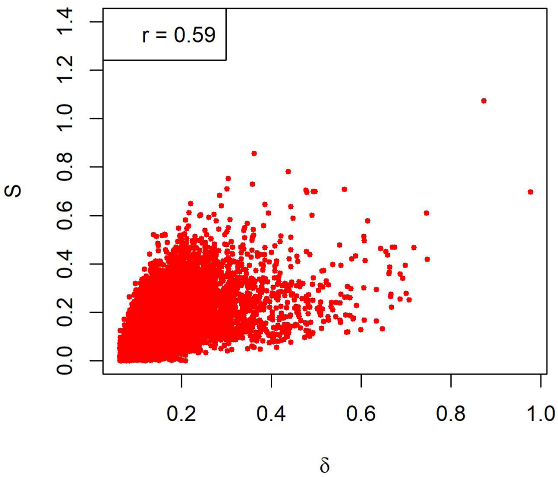
(b) Correlation between δ and I_a based on T_{15}.

FIGURE 9. (a) Correlation between δ and I_a based on T_{14}. (b) Correlation between δ and I_a based on T_{15}.

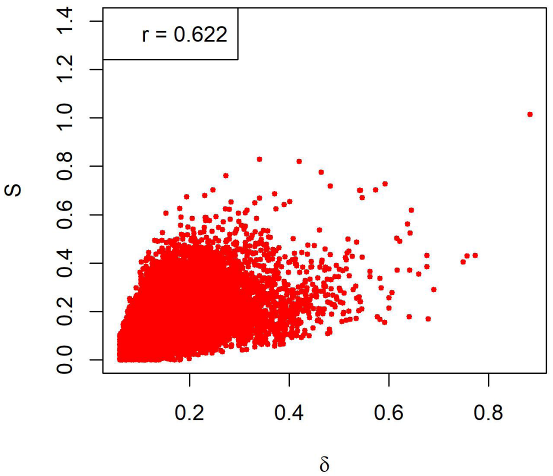
E. APPLICATION TO CHEMICAL STRUCTURES

Chemical structures of carbon-containing compounds (organic compounds) can be represented by colored graphs with the atoms for the vertices and the chemical bonds for the edges [45]. Although this representation is only approximate, several graph invariants have been successfully used for similarity searches of chemical structures [46], [47] or as variables in multiple regression models for the prediction of substance properties from chemical structures [48]. Molecular symmetry and the concept of topologically equivalent atoms or bonds (based on the automorphism group) are for instance important in the interpretation of NMR spectra, in automated synthesis planning, and for the generation of isomers [49]–[52].

Based on the definition of symmetry measures here for uncolored graphs we discuss examples for selected chemical structures (graphs) as follows: (a) skeletons with carbon (C) atoms (H-depleted); (b) only C-C single bonds present (unsaturation is covered by rings); (c) the maximum number of bonds per atom (vertex degree) is four.



(a) Correlation between δ and S based on T_{16}.



(b) Correlation between δ and S based on T_{17}.

FIGURE 10. (a) Correlation between δ and S based on T_{16}. (b) Correlation between δ and S based on T_{17}.

Complete sets of isomers have been generated by software Molgen [50]. Note that not all generated isomers reflect stable molecules.

First we discuss a real chemical compound belonging to the defined restrictions, namely adamantane $C_{10}H_{16}$. See Figure (11). The corresponding graph contains one orbit with 6 vertices $\{1, 3, 5, 7, 9, 10\}$ and one orbit with 4 vertices $\{2, 4, 6, 8\}$. In general, the symmetry measure δ has been defined by the unique, positive root of $O^{\star}_G(z) := 1 - O_G(z)$, see Equation (11). Here, we obtain the equation

$$1 - z^4 - z^6 = 0, \tag{65}$$

and $\delta \doteq 0.869$. The high value (maximum of δ equals 1) corresponds to the high topological symmetry of the adamantane structure.

Next we discuss the complete set of 4875 isomers with the molecular brutto formula of adamantane ($C_{10}H_{16}$). A summary of the symmetry measures δ, S, and I_a, applied to these graphs is given in Table (2). The range of δ is 0.1 to 0.869 with only 23 different values (rounded to 6 decimals). The maximum value appears for only one graph, namely for adamantane. 95 % of the graphs have a low symmetry with $\delta < 0.309$. The smoothed probability density distribution of δ is shown in Figure (12), exhibiting maxima for the frequently occurring values of δ. The second highest value of δ is 0.851, again for only one graph (Figure (13)); however, a corresponding molecule will probably not exist because of the two quaternary C-atoms. The absolute Pearson correlation

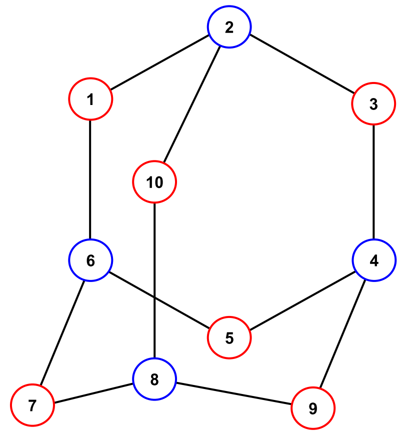


FIGURE 11. Skeleton of the chemical structure of adamantane $C_{10}H_{16}$. The graph contains two vertex orbits, one with 6 atoms (in red), the other with 4 atoms (in blue).

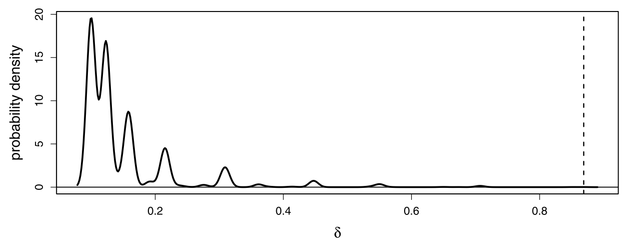


FIGURE 12. Probability density distribution of δ for the 4875 isomers of $C_{10}H_{16}$. One of them is adamantane with the maximum value of 0.869 marked by the vertical line.

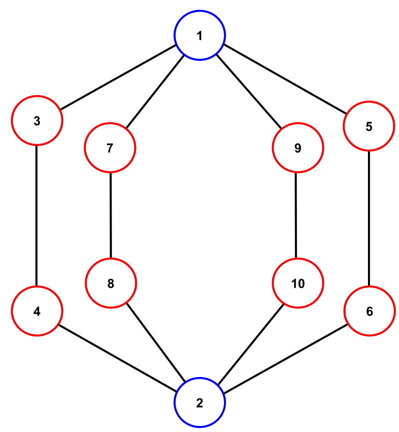


FIGURE 13. Isomer with second highest δ-value (0.851) of the set $C_{10}H_{16}$. The graph contains two vertex orbits, one with 8 atoms (in red), the other with 2 atoms (in blue). The value of δ is obtained as the root of equation $1 - z^2 - z^8 = 0$. This graph does not correspond to a stable molecule.

TABLE 2. Summary of the symmetry measures δ, S, and I_a, applied to the 4875 isomers of $C_{10}H_{16}$.

	Min	Mean	Max	Pearson corr. coeff. r			No. diff. values
				δ	S	I_a	
δ	0.100	0.150	0.869	1.000	0.746	-0.940	23
S	0.000	1.203	8.158	0.746	1.000	-0.892	54
I_a	0.722	3.055	3.322	-0.940	- 0.892	1.000	22

coefficients between δ and the two other symmetry measures are higher than for the exhaustively generated trees (Table 1), i.e., 0.746 between δ and S, and -0.940 between δ and I_a. See Table (2). A reason for this could be the cyclic structure of the molecular graphs which increase the symmetry, but makes δ more degenerate.

III. SUMMARY AND CONCLUSION

This paper has introduced a novel quantitative measure of symmetry based on the orbit polynomial defined in terms of the orbits of the automorphism group of a graph. This measure (i.e., δ, the unique positive root of one minus the orbit polynomial) offers a method for comparing graphs according to their symmetry structure and can be used to establish a partial order on a class of graphs. Properties of the measure have been demonstrated, and the measure has been applied to several classes of trees to analyze correlations between the values of this measure and those of two other quantitative measures of graph symmetry. The relation between δ and the measure I_a warrants further study since both are defined relative to the number and sizes of the orbits of the automorphism group of a graph. In addition, further research is needed to determine the range of structural properties of a graph that can be captured by the orbit polynomial.

In addition, we applied δ as a structural descriptor to chemical structures. The new molecular descriptor δ for topological symmetry is promising if used together with other topological descriptors, for instance, for cluster analysis of molecular skeletons. Further work is necessary for extending this concept to colored graphs that are needed to model general molecular structures.

REFERENCES

[1] N. Rashevsky, "Life, information theory, and topology," *Bull. Math. Biophys.*, vol. 17, no. 3, pp. 229–235, Sep. 1955.
[2] A. Mowshowitz, "Entropy and the complexity of graphs: I. An index of the relative complexity of a graph," *Bull. Math. Biophys.*, vol. 30, no. 1, pp. 175–204, Mar. 1968.
[3] R. Albert, H. Jeong, and A.-L. Barabási, "Diameter of the world-wide Web," *Nature*, vol. 401, pp. 130–131, Sep. 1999.
[4] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world'networks," *Nature*, vol. 393, pp. 440–442, Jun. 1998.
[5] M. Dehmer and F. Emmert-Streib, *Quantitative Graph Theory: Theory and Applications*. Boca Raton, FL, USA: CRC Press, 2014.
[6] D. Bonchev and N. Trinajstić, "Information theory, distance matrix, and molecular branching," *J. Chem. Phys.*, vol. 67, no. 10, pp. 4517–4533, Nov. 1977.
[7] D. Bonchev, "Topological order in molecules 1. Molecular branching revisited," *J. Mol. Struct., THEOCHEM*, vol. 336, nos. 2–3, pp. 137–156, 1995.
[8] M. Dehmer, F. Emmert-Streib, and Y. Shi, "Interrelations of graph distance measures based on topological indices," *PLoS ONE*, vol. 9, no. 4, Apr. 2014, Art. no. e94985.
[9] D. J. Watts, *Small Worlds: The Dynamics of Networks Between Order and Randomness*. Princeton, NJ, USA: Princeton Univ. Press, 1999.
[10] Y. Yang, "Resistance distances and the global cyclicity index of fullerene graphs," *Dig. J. Nanomater. Biostruct.*, vol. 7, pp. 593–598, Apr. 2012.
[11] F. Ball and A. Geyer-Schulz, "Weak invariants of actions of the automorphism group of a graph," *Arch. Data Sci., A*, vol. 2, no. 1, pp. 2363–9881, 2017.
[12] F. Ball and A. Geyer-Schulz, "The impact of graph symmetry on clustering," *Arch. Data Sci., A*, vol. 5, no. 1, pp. 1–19, 2018.
[13] F. Emmert-Streib, M. Dehmer, and Y. Shi, "Fifty years of graph matching, network alignment and network comparison," *Inf. Sci.*, vols. 346–347, pp. 180–197, Jun. 2016.
[14] J. Devillers and A. T. Balaban, *Topological Indices and Related Descriptors in QSAR and QSPR*. Amsterdam, The Netherlands: Gordon and Breach Science Publishers, 1999.
[15] R. A. Botafogo, E. Rivlin, and B. Shneiderman, "Structural analysis of hypertexts: Identifying hierarchies and useful metrics," *ACM Trans. Inf. Syst.*, vol. 10, no. 2, pp. 142–180, Apr. 1992.
[16] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Statist.*, vol. 22, no. 1, pp. 79–86, 1951.
[17] J. Ostrowski, J. Linderoth, F. Rossi, and S. Smriglio, "Orbital branching," *Math. Program.*, vol. 126, no. 1, pp. 147–178, Jan. 2011.
[18] B. D. McKay. (2010). *Nauty*. [Online]. Available: http://cs.anu.edu.au/~bdm/nauty/.
[19] B. D. McKay and A. Piperno. (2013). *Practical Graph Isomorphism, II*. [Online]. Available: http://cs.anu.edu.au/~bdm/nauty/
[20] C. J. Colbourn and K. S. Booth, "Linear time automorphism algorithms for trees, interval graphs, and planar graphs," *SIAM J. Comput.*, vol. 10, no. 1, pp. 203–225, Feb. 1981.
[21] A. S. Householder, *The Numerical Treatment of a Single Nonlinear Equation*. New York, NY, USA: McGrawHill, 1970.
[22] M. A. Jenkins and J. F. Traub, "Algorithm 419: Zeros of a complex polynomial," *Commun. ACM*, vol. 15, no. 2, pp. 97–99, Feb. 1972. [Online]. Available: http://doi.acm.org/10.1145/361254.361262
[23] Y. Shi, M. Dehmer, X. Li, and I. Gutman, *Graph Polynomials*. Weinheim, Germany: Wiley, 2017.
[24] C. Godsil and G. Royle, *Algebraic Graph Theory* (Graduate Texts in Mathematics). New York, NY, USA: Academic, 2001.
[25] A. Mowshowitz, "Entropy and the complexity of graphs II: The information content of digraphs and infinite graphs," *Bull. Math. Biophys.*, vol. 30, pp. 225–240, Jun. 1968.
[26] N. Trinajstić, *Chemical Graph Theory*. Boca Raton, FL, USA: CRC Press, 1992.
[27] M. Marden, "Geometry of polynomials," in *Mathematical Surveys of the American Mathematical Society*, vol. 3. Providence, RI, USA: AMS, 1966.
[28] D. Bonchev, *Information Theoretic Indices for Characterization of Chemical Structures*. Chichester, U.K.: Research Studies Press, 1983.
[29] F. Emmert-Streib and M. Dehmer, Eds. *Analysis of Microarray Data: A Network-Based Approach*. Weinheim, Germany: Wiley, 2010.
[30] S. C. Basak, B. D. Gute, and A. T. Balaban, "Interrelationship of major topological indices evidenced by clustering," *Croatica Chemica Acta*, vol. 77, nos. 1–2, pp. 331–344, 2004.
[31] S. C. Basak, A. T. Balaban, G. D. Grunwald, and B. D. Gute, "Topological indices: Their nature and mutual relatedness," *J. Chem. Inf. Comput. Sci.*, vol. 40, no. 4, pp. 891–898, Jul. 2000.
[32] M. Dehmer, Z. Chen, F. Emmert-Streib, A. Mowshowitz, Y. Shi, S. Tripathi, and Y. Zhang, "Towards detecting structural branching and cyclicity in graphs: A polynomial-based approach," *Inf. Sci.*, vol. 471, pp. 19–28, Jan. 2019.
[33] P. Deuflhard, *Newton Methods for Nonlinear Problems* (Springer Series in Computational Mathematics). Berlin, Germany: Springer, 2006.
[34] B. L. Van der Waerden, *Algebra*, vols. 1–2. Heidelberg, Germany: Springer, 1993.
[35] D. Bonchev, O. Mekenyan, and N. Trinajstić, "Topological characterization of cyclic structures," *Int. J. Quantum Chem.*, vol. 17, no. 5, pp. 845–893, May 1980.
[36] R Core Team, Vienna, Austria. (2013). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*. [Online]. Available: http://www.R-project.org/
[37] G. Csardi and T. Nepusz. (2009). *Package Igraph Network Analysis and Visualization*. [Online]. Available: http://igraph.sourceforge.net

[38] A. Mowshowitz and V. Mitsou, "Entropy, orbits and spectra of graphs," in *Analysis of Complex Networks: From Biology to Linguistics*, M. Dehmer and F. Emmert-Streib, Eds. Weinheim, Germany: Wiley, 2009, pp. 1–22.

[39] A. Mowshowitz and M. Dehmer, "A symmetry index for graphs," *Symmetry, Culture Sci.*, vol. 21, no. 4, pp. 321–327, 2010.

[40] M. Dehmer, F. Emmert-Streib, and M. Grabner, "A computational approach to construct a multivariate complete graph invariant," *Inf. Sci.*, vol. 260, pp. 200–208, Mar. 2014.

[41] D. Bonchev, O. V. Mekenyan, and N. Trinajstic, "Isomer discrimination by topological information approach," *J. Comput. Chem.*, vol. 2, no. 2, pp. 127–148, 1981.

[42] A. Garrido, "Symmetry in complex networks," *Symmetry*, vol. 3, no. 1, pp. 1–15, Jan. 2011.

[43] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*. Champaign, IL, USA: Univ. Illinois Press, 1949.

[44] N. A. Rahman, *A Course in Theoretical Statistics*. London, U.K.: Charles Griffin and Company, 1968.

[45] T. Engel and J. Gasteiger, *Chemoinformatics: Basic Concepts and Methods*. Weinheim, Germany: Wiley, 2018.

[46] K. Varmuza, W. Demuth, M. Karlovits, and H. Scsibrany, "Binary substructure descriptors for organic compounds," *Croatica Chemica Acta*, vol. 78, no. 2, pp. 141–149, 2005.

[47] P. Willett, *Similarity and Clustering in Chemical Information Systems*. Letchworth, U.K.: Research Studies Press, 1987.

[48] K. Varmuza, P. Filzmoser, and M. Dehmer, "Multivariate linear QSPR/QSAR models: Rigorous evaluation of variable selection for PLS," *Comput. Struct. Biotechnol. J.*, vol. 5, pp. 1–10, Feb. 2013.

[49] T. Engel, *Computer Processing of Chemical Structure Information*, T. Engel and J. Gasteiger, Eds. Weinheim, Germany: Wiley, 2018, pp. 43–120.

[50] C. Benecke, T. Grüner, and R. Grund. (2009). *Software Molgen 3.5*. Univ. Göttingen, Göttingen, Germany. [Online]. Available: http://www.molgen.de

[51] M. Razinger, K. Balasubramanian, and M. E. Munk, "Graph automorphism perception algorithms in computer-enhanced structure elucidation," *J. Chem. Inf. Model.*, vol. 33, no. 2, pp. 197–201, Mar. 1993.

[52] M. Randić, M. Nović, and D. Plavšic, *Solved and Unsolved Problems of Structural Chemistry*. Boca Raton, FL, USA: CRC Press, 2016.

MATTHIAS DEHMER received the B.Sc. degree in mathematics from the University of Siegen, Germany, in 1997, and the Ph.D. degree in computer science from the Darmstadt University of Technology, Germany, in 2005.

He is currently a Professor with the University of Applied Sciences Upper Austria, Steyr Campus, and with UMIT—The Health and Life Sciences University, Hall in Tirol, Austria. He has authored and coauthored more than 245 publications. His H-index equals 33 and his i10-index equals 101. His research interests include data science, cybersecurity, disaster management, complex networks, risk analysis, information systems, machine learning, information theory, bioinformatics, visual analytics, and computational statistics.

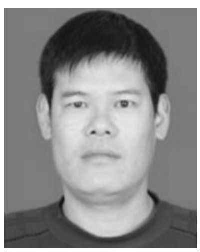
ZENGQIANG CHEN (Member, IEEE) received the B.S. degree in mathematics and the M.S. and Ph.D. degrees in automatic control from Nankai University, Tianjin, China, in 1987, 1990, and 1997, respectively. He is currently a Full Professor with the Department of Automation, College of Computer and Control Engineering, Nankai University. He has authored and coauthored over 200 journal articles. His current research interests include intelligent optimizing control, intelligent computing, complex networks, and multiagent systems.

FRANK EMMERT-STREIB studied physics and mathematics. He received the Diploma degree (Hons.) and the bachelor's degree in theoretical physics from the University of Siegen, Germany, and the Ph.D. degree in theoretical physics from the University of Bremen, Germany.

He is currently a tenured Associate Professor of data science with Tampere University. In addition, he is affiliated with the Institute of Biosciences and Medical Technology (BioMediTech). His research interests include computational biology, network science, digital business, and computational social science.

ABBE MOWSHOWITZ is currently a Professor of computer science with The City College of New York, and a member of the doctoral faculty with the Graduate Center of The City University of New York. In 1984, he joined CUNY, and he has also held academic appointments at the University of Amsterdam, Erasmus University Rotterdam, and the Rotterdam School of Management. Earlier, he held academic appointments at the Rensselaer Polytechnic Institute, The University of British Columbia, the University of Toronto, and the University of Michigan. He has been a major contributor to the analysis of complex networks starting with seminal work in the late 1960s in which he developed and applied an information theoretic measure of graph complexity. From 2006 to 2016, he participated in the International Technology Alliance Research Project dedicated to the further development of complex network technology. He is currently participating in an NSF supported project dealing with secure resilient edge cloud designed networks in collaboration with The City College of New York and the Kyushu Institute of Technology, Japan.

KURT VARMUZA received the degree in chemistry from the Vienna University of Technology, Austria. As one of the pioneers in chemometrics, his research activities in chemoinformatics and chemometrics involve development and applications of methods for spectra-structure relationships (MS and IR), structureproperty relationships (QSPR), and classification of materials in archaeometry and cosmo chemistry. Since 1992, he has been a Professor with the Vienna University of Technology, Austria.



LIHUA FENG received the B.S. degree from Ludong University, Yantai, China, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China. He is currently a Professor with the School of Mathematics and Statistics, Central South University, Hunan. His research interests include graph theory and algebraic graph theory.

HERBERT JODLBAUER received the Dipl.Ing. degree in technical mathematics from the University of Linz, Austria, and the Ph.D. degree in mechanical engineering from the Vienna University of Technology, Austria. He is currently a Full Professor with the University of Applied Sciences Upper Austria, Steyr, Austria, and the Director of the Studies of Production and Operations Management, and he leads the Centre of Excellence for Smart Production.

YONGTANG SHI received the Ph.D. degree in applied mathematics from Nankai University, Tianjin, China. He is currently a Full Professor with Nankai University. His research interests include graph theory and its applications, combinatorial optimization, and discrete mathematics and its applications in computer science, chemistry, and biology.

JIN TAO (Member, IEEE) received the B.Sc. degree in automation from the Qingdao University of Science and Technology, Qingdao, China, in 2008, the M.Sc. degree in control theory and control engineering from the Guangxi University of Science and Technology, Liuzhou, China, in 2011, and the Ph.D. degree in control science and engineering from Nankai University, Tianjin, China, in 2017.

From 2010 to 2011, he spent two years as a Joint Training Student with the Intelligent Control Laboratory, College of Engineering, Peking University, Beijing, China. From 2011 to 2014, he was a Chief Engineer with the Institute of Rotary Drilling Machine, Beijing Sany Heavy Machinery Company Ltd., Beijing. He is currently an Academy of Finland Postdoctoral Researcher with Aalto University, Espoo, Finland. He is also with the College of Engineering, Peking University. He has published more than 40 peer-reviewed articles in international journals and conferences. His research interests include advanced control, evolutionary optimization, dynamic modeling, and their applications.