MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **82** (2019) 561-580

ISSN 0340 - 6253

Exhaustive and Metaheuristic Exploration of Two New Structural Irregularity Measures

Paulo Boaventura–Netto[a], Leonardo de Lima[b], Gilles Caporossi[c]

[a] *Department of Production Engineering, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil*
boaventu@po.coppe.ufrj.br

[b] *Department of Production Engineering, Federal Center of Technological Education Celso Suckow da Fonseca, Rio de Janeiro, Brazil*
leonardo.lima@cefet-rj.br

[c] *GERAD and HEC Montréal, Montréal, Canada*
Gilles.Caporossi@gerad.ca

(Received November 13, 2018)

Abstract

Let $G(V,E)$ be a graph with vertex set V, $|V| = n$, and edge set E. In this paper, we introduce two new polynomial irregularity measures:

$$IRR_m(G) = \frac{(\xi - 1)}{n} + \sum_{(i,j)\in E} |(\ d_i\mu(d_i) - d_j\mu(d_j)\ )|$$

and

$$IRR_d(G) = \frac{(\xi - 1)}{n} + \sum_{(i,j)\in E} \left|\frac{d_i}{\mu(d_i)} - \frac{d_j}{\mu(d_j)}\right|,$$

where d_i is the degree of the vertex $v_i \in V$, $\mu(d_i)$ is the degree multiplicity of v_i in the degree sequence and ξ is the number of (different) degree values of G. The results of two explorations: one, exhaustive, of the graph sets from 4 to 10 vertices, and other, using AGX-III program on graphs from 11 to 30 vertices, both looking for extremal graphs of two new polynomial irregularity measures are presented. Some discussion on the obtained values and structures is presented. The use of AGX-III allowed us to identify typical structures for the extremal graphs associated with these measures. Some improvements were obtained through the variation of a parameter, with the aid of manual graph building by using an optimal strategy. These structures we built were of the type indicated by the heuristic. For the second measure, AGX-III showed extremal graphs based on unigraphic sequences which generate threshold graphs.

1 Introduction

1.1 Objectives and content

The irregularity in graphs is a property of relatively recent study. It was motivated by applications in different contexts as one can see in [2] and the references therein, and particularly in chemistry, where it is a consequence of the different valences of the chemical elements present in the graph structures associated with chemical formulas, [14]. Despite this structural motivation, the literature has shown, with a single exception - the Albertson measure, [2] - expressions aimed at measuring the irregularity based only on degree sequences, which leads to obtaining the same values for different graphs. Recently, the σ irregularity index, also based on degree sequences, was introduced in [15] and some basic properties were shown. More results on σ graph with maximal irregularity, especially for extremal graphs can be found in [1]. The proposal discussed here, as in our previous work, [9], [11], is based on Albertson measure. This section contains a quick discussion of the nomenclature and the notation relevant for the work. Section 2 presents the proposed irregularity measures. Section 3 is dedicated to the exhaustive exploration conducted on orders 4 to 10. Section 4 presents the results obtained through AGX-III exploration, on orders 11 to 30. Section 5 aligns some conclusions and suggestions for future research.

1.2 Nomenclature and notation

We consider, in this text, simple graphs $G = (V, E)$ (non-oriented, without multiple edges and without loops) where $V = \{v_i,\ i = 1, ..., n\}$ is the vertex set, $E = \{(v_i, v_j),\ i, j = 1, ..., n,\ i \neq j\}$ is the edge set, $n = |V|$ is the order and $m = |E|$ is the *size* of G. The theory includes other equivalent definitions. We call $G(n)$ the set of all graphs G of order n and $G(n, m)$ the set of all graphs with n vertices and m edges. The *degree* d_i of a vertex v_i is the number of edges from which it participates and we define the *degree sequence* $d = \{(d_i; i = 1, ..., n)\}$ as a non increasing sequence such that $d_1 \geq d_2 \geq \cdots \geq d_n$ producing an *ordered degree sequence* (ODS). For convenience, we also consider the graphical sequence as a non increasing sequence. A graph G is k-regular if every vertex in G has the same degree k. If there is no $k \in N$, such that G is k-regular, then G is *irregular*. A graph G on n vertices is *antiregular* if its degree sequence has $n - 1$ different values. For every $n \geq 2$ there is a connected antiregular graph on n vertices, [20], which we denote by AR_n. The *complement* of a graph G is a graph $\overline{G}$ on the same vertices such that two distinct vertices of $\overline{G}$ are adjacent if and only if they are not adjacent in G. For every n the complement $\overline{AR_n}$ of AR_n is also antiregular and no other graph in $G(n)$ is antiregular. There is only one graph for each order, the *complete graph* (K_n), which contains all possible edges. Two graphs are *isomorphic* if there exists a bijection preserving their edge sets. Two graphs are *isomeric* if they have the same ODS. A walk in a graph is a sequence $v_0, e_1, v_1, \ldots, e_l, v_l$ of vertices v_i and edges e_i such that for $1 \leq i \leq l$, the edge e_i has endpoints v_{i-1} and v_i. A *path* is a walk with no repeated vertex. A graph with n vertices and $n-1$ edges, which is a path, is denoted P_n. A *cycle* is a closed path in a graph which does not repeat any of its elements. A graph with n vertices and n edges which is a cycle is denoted C_n. A graph is *connected* if for every pair v_i, v_j of vertices there is a path joining v_i to v_j and is *not connected*, or *disconnected*, if this is not true. An *independent set* $S \subseteq V$ is a vertex set where no vertex pair defines an edge. Several matrices associated with a given graph can be defined. The most immediate is the *adjacency matrix* $A = [a_{ij}]$, where $a_{ij} = 1$ if $\exists (v_i, v_j) \in E$ and $a_{ij} = 0$ on the contrary. The *diversity* $\xi(G)$ of a graph is the number of (different) degree values of its sequence, $\xi(G) = 1$ if G is k-regular, $\xi(G) = |\{d_i|, d_i \neq d_j, i = 1, ..., n-1, j = i+1, ..., n\}|$, if G is irregular. The *multiplicity* $\mu(x)$ of a given value x in a sequence associated with a graph is the number of times x appears in the sequence. Here, we apply this concept to the degree sequence of a graph. A *split graph* is a graph where the vertex set can be partitioned into a complete graph and an independent set. More details can be found in [5], [6], [10] and [19].

2 The proposed irregularity measures

An *irregularity measure*(IM) of a graph G is a real function $F : I(G) \rightarrow R$ of a G invariant set I, such that $F(G) = 0$ if and only if G is regular. The work on this subject involves not only the definition of new measures which better express the irregularity, but also concerns the search for extremal graphs associated with the existing measures, i.e., graphs that present maximum value for a given IM. These extremal graphs would then be the most irregular for the corresponding measure. This last topic proved to be quite difficult, without complete success until today. Details concerning the known extremal graph families for a number of IMs are in [23] and [24]. The authors in [2] defined the imbalance measure,

$$irr(G) = \sum_{(i,j)\in E} |d_i - d_j|. \tag{1}$$

The module of the difference between i and j degrees is the *unbalancing* of the edge (i, j). This measure presents some zero values for disconnected graphs with regular components of different degrees. This violates the necessity condition to define an IM but, even with this drawback, it has been included in the IM literature, which can be understood by the fact that it was, among the existing polynomial IMs, the one in which the definition involves the edges and therefore properties of structure. We call a measure having this property a *structural* one.

We propose two new measures (*division* and *multiplication* measures), which are also structural and take into account the degree multiplicities. The second terms of their expressions are based on that of Albertson measure and they should present the same drawback previously discussed concerning that measure. The first terms were introduced in order to avoid this problem, that is, the case when the graph is disconnected with regular connected components of different sizes. The division IM (IRR_d) is given by

$$IRR_d(G) = \frac{(\xi - 1)}{n} + \sum_{(i,j)\in E} \left| \frac{d_i}{\mu(d_i)} - \frac{d_j}{\mu(d_j)} \right|, \tag{2}$$

where ξ is the number of different degrees and $\mu(d_k)$ is the degree multiplicity of k in the degree sequence. It is an IM, since for a r-regular graph G with n vertices, we have $\mu(r) = n$ and all difference modules are $|r/n - r/n| = 0$. In this case the first term, being itself a (nonstructural) IM, will be also null, then $IRR_d(G) = 0$ for a regular graph. This term marks the presence of irregular graphs even when the second term is null. The second term of $IRR_d(G)$ presents zeroes for P_6 and for graphs within a family given by the ODS $(4, 4, 4, 4, 4, 4, 2, 2, 2)$ for $n = 9$, $(4, 4, 4, 4, 4, 4, 4, 4, 2, 2, 2, 2)$ for $n = 12$ and so on, provided the obtained sequences are graphic (like these examples). The multiplication IM (IRR_m) is given by

$$IRR_m(G) = \frac{(\xi - 1)}{n} + \sum_{(i,j)\in E} |(\ d_i\mu(d_i) - d_j\mu(d_j)\ )| \tag{3}$$

It is easy to see that $IRR_m(G) = 0$ for regular graphs. Here, once again, the first term acts as a correction for the case of disconnected graphs having regular components with different degrees and other possible zeroes. The sum of all degree multiplicities is equal to n. For vertices of equal degree, the degree multiplicities being equal, the results will be null. On the other hand, the greater the diversity of the graph, the smaller the number of zero differences, but the final value depends both on degrees and multiplicities values.

3 The exhaustive exploration

3.1 The graphs dataset

The calculation involved all graphs (connected and non-connected) with orders from 4 to 10. Table 1, [26], below, presents the number of different graphs for each order. As it can be seen, the number of connected graphs grows very quickly with the order. More detailed investigations for small orders as, for instance, determining not connected extremal graphs, can be easily done. The obtained extremals are chiefly connected, which in some sense confirm that connexity constraints are not so important to the general computations here.

Order	4	5	6	7	8	9	10
All graphs	11	34	156	1,044	12,346	274,668	12,005,168
Connected graphs	6	21	112	853	11,117	261,080	11,716,571
% of connected graphs	54.5	61,8	71,8	81,7	90,0	95,0	97,6
Maximum of $(\xi - 1)/n$	0.500	0.600	0.667	0.714	0.750	0.778	0.800

Table 1. Amount of graphs and connected graphs with orders from 4 to 10

We can observe that $0 \leq (\xi - 1)/n \leq (n - 2)/n$, [11]. Table 1 shows the maximum values of this term as a function of graph order. Given the low values obtained, its calculation was limited to the second term. It is easily seen that second-term IRR_m values are always integer. Unless we need a more detailed investigation of the final values, the sum of the first term is not needed for an extremality study. For the extremal second-term values, vastly outnumbered, the final verification can be done separately, looking for different diversities.

Regarding to IRR_d, some doubts about the extremality may appear for smaller orders; however, it turns out that their second-term values are integer and, for these graphs, it is worth the same argument of the previous case. Additional discussion is presented together with AGX-III results (Section 4).

3.2 Some details on the associated programming

We used the well-known $GENG$ procedure from *nauty* routines, [22], in order to generate all non-isomorphic graphs of orders from 4 to 10. For each order n, $GENG$ generates a text file containing the adjacency matrices of all graphs. We also implemented a function irregularity $(G, IrrFunc)$ that, given a graph G and the IM, $IrrFunc$ computes the sequence degree of G and its multiplicities and return the value of the chosen IM. The pseudocode of the implemented algorithm is presented below.

Algorithm 1: ComputeIrregularity(IrrFunc, n)

```
ObjFunc* ← 0 File ← GENG(n)
initialization
while not EndOfFile(File) do
    G ← ReadAdjMatrix(file)
    ObjFunc ← irregularity(G,IrrFunc)
    if ObjFunc* < ObjFunc then
        G* ← G
        ObjFunc* ← ObjFunc
    end
end
```

The experiments were performed in Matlab R2014b on 2.5 GHz Intel Core i5 processor (Mac OS X 10.9.4) and 8 GB of RAM. The extremal graphs for the two IMs, within each order, were shown graphically. The results are as follows.

3.3 Results

3.3.1 Extremal graphs and their IRR_d values

Figure 1 displays the extremal graphs and Table 2 shows the IM values and a note about the intersection with the split graph family.

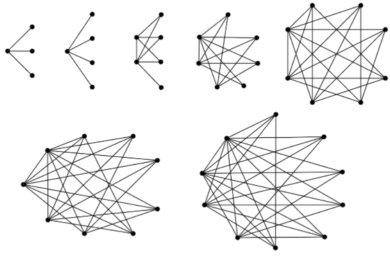

Figure 1. Extremal graphs for IRR_d ($4 \leq n \leq 10$)

Order	4	5	6	7	8	9	10
Second term	8	15	28	46	76.80	113.6	161.333
$(\xi - 1)/n$	0.250	0.200	0.500	0.571	0.375	0.444	0.400
IRR_d value	8.250	15.200	28.500	46.571	77.175	114.044	161.733
Split $(\|K\|, \|I\|)$	1,3	1,4	2,4	2,5	no	no	no

Table 2. IRR_d values and possible split structure for extremal graphs, $4 \leq n \leq 10$

We found extremal graphs presenting both low and high diversity values. The most interesting is the presence of a $G(8)$ extremal with $\xi = 2$, the same value shown by the $G(4)$ and $G(5)$ extremals. On the other hand, $G(6)$ and $G(7)$ extremals have $\xi = 4$ while $G(9)$ and $G(10)$ extremals have $\xi = 5$. Other interesting observation, for the extremals from 4 to 7 vertices, is that they are all split graphs, while for $G(8)$ through $G(10)$ extremals it is not true.

3.3.2 Extremal graphs and their IRR_m values

We also calculated separately the second term values. Figure 2 displays the extremal graphs.

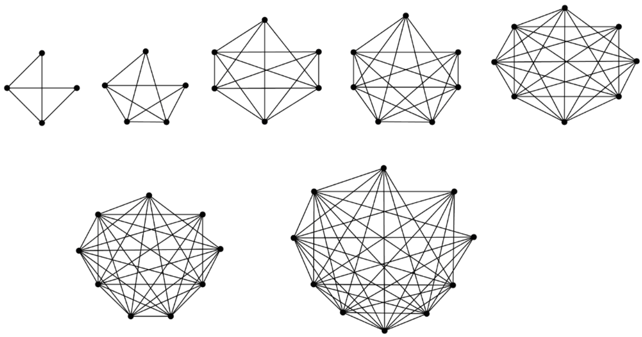

Figure 2. Extremal graphs for IRR_m $(4 \leq n \leq 10)$

Table 3 gives the extremal IRR_m values for these datasets.

Order	4	5	6	7	8	9	10
Second term	8	36	96	200	380	660	1056
$(\xi - 1)/n$	0.250	0.200	0.167	0.142	0.250	0.222	0.333
IRR_d value	8.250	36.200	96.167	200.142	380.250	660.222	1056.333

Table 3. IRR_m values for extremal graphs, $4 \leq n \leq 10$

Unlike that observed with IRR_d, here all the extremals have low diversity, 2 for $G(4)$ to $G(7)$, 3 for $G(8)$ and $G(9)$ and 4 for $G(10)$. Their general structure is near to that of K_n, with one or two less edges.

4 The use of AGX-III

4.1 Using the multiobjective AGX-III capacity

A systematic exploring of IRR_d and IRR_m landscapes, from orders 11 to 30, was done with the aid of AGX-III, [12], Version 3.1.6. This software applies the Variable Neighborhood Search (VNS) metaheuristic, [21], [13], [16] to a universe of graphs automatically added in order to look, in our case, for maximum IRR_d and IRR_m values. Depending on the problem, the optimization may fail when no better solution is found in the neighborhood of the current one. Diversification of the search is then needed. Since the perturbation phase of VNS aims at handling this difficulty, we used the new multi-objective capability of AGX-III to help the search. Indeed, a secondary criterion based on a very discriminating invariant to be minimized and maximized was added to the problem. A solution then becomes a tentative Pareto front, which yields a large number of graphs with rather good quality. The neighborhoods of this set of graphs being larger, the chances to find a better solution are increased. The Balaban index [3], which is known to be very discriminating, was used, which clearly improved the efficiency of the search.

For IRR_d, and chiefly for IRR_m, the results were strongly consistent with the structures obtained by the initial exhaustive exploration. The IRR_d landscape showed to be more complex, while the IRR_m one showed the most interesting properties. Through that exploration we can see that the vertex set V of a IRR_d extremal graph can be, since $n = 4$, partitioned according to their multiplicity values, into two subsets, which we call A and B (Table 4). The cases $n = 6$ and 7 are exceptions, since the set A presented one degree value lesser than that of set B. All results from $n = 8$ on, both with exhaustive research and metaheuristic application, show the set B formed by equal integers lesser than A minimum value.

Order	A	B	Divers.	Order	A	B	Divers.
4	3	1^3	2	8	7,6,5	4^5	4
5	4	1^4	2	9	8,7,6,5	4^5	5
6	5,4,1	2^3	4	10	9,8,7,6	4^6	5
7	6,5,4,2	3^3	4				

Table 4. Partitioning of vertex sets according to degrees ($n = 4$ to 10)

Table 5 below gives the best IRR_d values found and the degree sequences corresponding to the associated graphs. A continuously descending degree subsequence is indicated by $a \to b$, where a and b are its maximum and minimum values. Multiplicity is shown as an exponent in B column. Orders 8 to 10 were included in benefit of general comparison. Let d_A and d_B be the minimum and maximum degrees in the sets A and B, respectively. Some properties such as size, diversity and the difference $d_A - d_B$, are also presented.

n	IRR_d value	Degree seq. subsets		Size	Diversity	$\|A\|$	$d_A - d_B$
		A	B	m			
8	77.175	$7 \to 5$	4^5	19	4	3	1
9	114.044	$8 \to 5$	4^5	23	5	4	1
10	161.733	$9 \to 6$	4^6	27	5	4	2
11	219121	$10 \to 6$	5^6	35	6	5	1
12	291845	$11 \to 7$	5^7	40	6	5	2
13	374635	$12 \to 8$	5^8	45	6	5	3
14	467468	$13 \to 9$	5^9	50	6	5	4
15	579333	$14 \to 9$	7^9	66	7	6	2
16	713104	$15 \to 9$	8^9	78	8	7	1
17	849804	$16 \to 9$	8^9	86	9	8	1
18	1033.64	$17 \to 10$	8^{10}	94	9	8	2
19	1195.88	$18 \to 11$	10^{11}	113	9	8	1
20	1379.43	$19 \to 13$	8^{13}	108	8	7	5
21	1640.07	$20 \to 13$	8^{13}	113	9	8	5
22	1875.18	$21 \to 13$	11^{13}	148	10	9	2
23	2178.68	$22 \to 14$	10^{14}	151	10	9	4
24	2477.98	$23 \to 15$	9^{15}	153	10	9	6
25	2846.02	$24 \to 13$	12^{13}	189	13	12	1
26	3192.96	$25 \to 15$	14^{15}	215	12	11	1
27	3537.90	$26 \to 17$	11^{17}	201	11	10	6
28	3956.16	$27 \to 17$	12^{17}	223	12	11	5
29	4477.00	$28 \to 17$	14^{17}	254	13	12	3
30	4809.73	$29 \to 18$	15^{18}	276	13	12	3

Table 5. Maximum IRR_d values found by AGX-III and their degree sequences

4.2 Some discussion on the results

Table 5 allows us to observe some interesting points in what concerns the structure of (presumably) IRR_d-extremal graphs. In what follows, we call $|A| = k$, then $d_k = d_{min(A)}$, $\mu(d_B)$=$n - k$. First of all, every order between 8 and 30 presents a set A formed by descending consecutive degree values. All results but those of $n = 25$ were obtained by AGX-III. We managed the exception through building a graph with the same characteristics found in other orders from Table 5, using the same reasoning presented later in this item when discussing an upper bound calculation. We can observe that $d_k = \mu(d_B)$, which is a consequence of set complementarity, given A and B definitions. A number of cases present $d_k - d_B > 1$, which allows us to explore similar structures based on different d_B values. For instance, the difference $d_k - d_B$ is 5 for $n = 20$ and 21 and 6 for $n = 24$ and 27. These points were used as a basis to state Conjecture 1.

Conjecture 1. *For any order, an IRR_d-extremal graph $G = (V, E)$ has a partition $V = (A, B)$, such that the vertex degrees in A can be consecutively decreasing ordered from $n - 1$ to d_k and the vertex degrees in B are all equal and lesser than d_k. Besides that, we have $|B| = \mu(d_B) = d_k$.*

Figure 3 shows extremal graphs with orders 10 and 17, where the (A, B) partition can be observed. The difference $d_k - d_B$ is 2 for the graph on the left and 1 for that on the right.

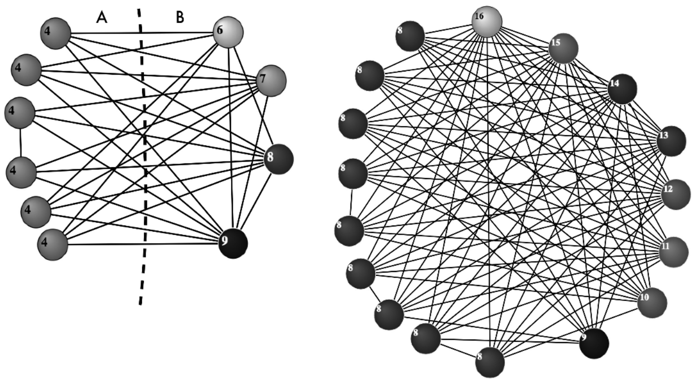

Figure 3. IRR_d-extremal graphs of orders 10 and 17

4.3 Discussion on $d_k - d_B > 1$ cases

The edge set E can be partitioned into three subsets:

$(A, A) = \{(u, v)|u \in A, v \in A\}; (A, B) = \{(u, v)|u \in A, v \in B\}; (B, B) = \{(u, v)|u \in B, v \in B\}$.

A number of graph orders investigated by AGX-III showed $d_k - d_B > 1$, which points to the possibility of better IRR_d values be found by examining the graph structure from Conjecture 3, with every feasible $d_k - d_B$ difference. (The feasibility of these values depends on the parity of the sequences). This should be done by graph construction based on graphic sequences, by using an optimality criterion.

From IRR_d definition, we conclude that the edges giving the greater contribution for its value should be the (A, B) ones, because the multiplicities 1 of A degrees and $|B|$ of the single degree in B shall correspond to greater edge values. The lesser contribution of (A, B) edges is $d_k - d_B/(n - k)$, while the (A, A) edges contribute with values between 1 and $d_k - d_B = k - 1$. We have to find conditions for the cheapest (A, B) edge to have a greater or equal value than the most high-valued (A, A) edge. This is given by the following lemma:

Lemma 2. *Let $G = (V, E)$ be a graph whose structure follows Conjecture 3. Then there is a minimum $|A|$ value, such that the subgraph containing all possible (A, B) edges, to which are added the admissible (A, A) edges, is able to show an upper bound for IRR_d value on G, when the edge values are calculated by using the original degree values.*

Proof. This condition will be valid everywhere the corresponding expression, based on IRR_d definition, is valid every time we have to eliminate (A, B) edges to create residual degrees to put (A, A) edges:

$$d_{min(A)} - \frac{d_B}{|B|} \geq (n - 1) - d_{min(A)} \tag{4}$$

With $|A| = k$, from the definitions of A and B, we have $|B| = d_{min(A)} = n - k$

Then, from (4),

$$n - k - \frac{d_B}{n - k} \geq k - 1 \tag{5}$$

Solving for k, we have

$$k \geq \frac{1}{4}(3n + 1) \pm \sqrt{(3n + 1)^2 - 4(n^2 + n - d_B)} \tag{6}$$

The value under the radical sign is $O(5n^2)$: then, the sum is $O(\frac{(3n+n\sqrt{5}+1)}{4} > n$, which is senseless since $k = |A|$, $A \subset V$. Then the correct solution is given by the subtraction,

$$k \geq \frac{1}{4}(3n+1) - \sqrt{(3n+1)^2 - 4(n^2+n-d_B)} \tag{7}$$

■

The value given by (7) approaches $(3n - n\sqrt{5}+1)/4 = O(0.2n)$ when n grows. Then the cardinality of B would be $O(0.8n)$ and this puts a limit on d_B value for the obtained graphs to be extremal with k near the limit: when d_B grows, the number of (B, B) edges will also grow and they count null for IRR_d value. The conclusion is that this limit is very weak for practical uses. Table 6 gives the minimum $|A|$ values, based on (7), for $d_B = \lceil n/2 \rceil$ (a common value, since d_B has little influence on k limit). The integer values are the lesser feasible ones.

Order	LS	Int	Order	LS	Int
8	1.553	2	20	3.846	4
9	1.756	2	21	4.042	5
10	1.935	2	22	4.228	6
11	2.136	3	23	4.423	5
12	2.317	3	24	4.610	5
13	2.517	3	25	4.805	5
14	2.699	3	26	4.992	5
15	2.898	3	27	5.187	6
16	3.081	4	28	5.374	6
17	3.279	4	29	5.568	6
18	3.463	4	30	5.756	6
19	3.660	4			

Table 6. Minimum A cardinality for (7) to be valid

4.3.1 The use of Lemma 2

When building an (A, B) set, for each new added edge (a, b), $a \in A$, $b \in B$, we will discount a unity from d_a and another from d_b. In order to have the highest edge contributions, we have to work with the current highest degrees in A. After the process is finished, every vertex will have a residual degree value. Let's consider the set A and the residual degree sequence S_A. Looking for an upper bound for IRR_d, we can add (A, A) edges and sum their contribution to the current (A, B)-obtained value. This can be done, even if S_A is not graphic – when there will be at least one degree value not completely fulfilled. Let Ψ_P be the provisional IRR_d value thus obtained. If S_A is graphic, (A, A) and (A, B) edge contributions will be added to give the final IRR_d value (the eventually present (B, B) edges give no contribution). Otherwise, we will have to eliminate (A, B) edges in order to modify S_A. This can result in the addition of new (B, B) edges to compensate the degree changes. Let $S_{A'}$ be the graphic sequence obtained as described. We had already a provisional IRR_d value, which is not correct as discussed before. To obtain S_A we had to replace some (the most expensive) (A, B) edges, for (the cheaper) (A, A) ones – and, if we follow Lemma 2 condition for $|A|$, it guarantees this trade result as a value loss. Then Ψ_P is an upper bound for IRR_d. We defined a percent slack, $Slack\% = 100(\Psi_P - AGX - IIIextremal)/AGX - IIIextremal$. If the slack is negative, the corresponding sequence cannot be extremal. If it is positive, it can be extremal or not, depending on the effect of the edge changes when trying to complete the graph construction. In Table 7, the columns under the head “AGX-III extremals data” show AGX-III better results, their upper bounds Ψ_P and their corresponding percent slacks. Under the head “Optimal strategy results” we show the number of results stronger than AGX-III, their corresponding optimal sequences and the overall best result value. The presence of isomers is indicated by an asterisk: since they are built using the optimal strategy, their values are better than those obtained by the metaheuristic, but they give not the best result, excepting for $n = 17$ and 18.

n	Possible seqs	Seqs UB	AGX-III extremal	$Slack\%$	Stronger results	Optimal seqs	Best result
10	$9 \to 6, 4^6$	167.400	161.733	3.504	———	Same	Same
11	$10 \to 6, 5^6$	222.455	219.121	1.521	———	Same	Same
12	$11 \to 7, 5^7$	297.417	291.845	1.909	———	Same	Same
13	$12 \to 8, 5^8$	382.385	374.635	2.069	———	Same	Same
14	$13 \to 9, 5^9$	477.357	467.468	2.116	———	Same	Same
15	$14 \to 9, 7^9$	592.400	579.333	2.256	———	Same	Same
16	$15 \to 9, 8^9$	722.437	713.104	1,309	———	Same	Same
17	$16 \to 9, 8^9$	870.470	847.804	2.673	1*	Same*	860.026*
18	$17 \to 10, 8^{10}$	1050.44	1033.64	1.626	1*	Same*	1035.64*
19	$18 \to 11, 10^{11}$	1230.42	1195.88	2.888	3*	$18 \to 11, 8^{11}$	1227.33
20	$19 \to 13, 8^{13}$	1422.35	1379.43	3.112	7*	$19 \to 11, 9^{11}$	1439.72
21	$20 \to 13, 8^{13}$	1686.38	1640.07	2.824	5*	$20 \to 12, 9^{12}$	1678.43
22	$21 \to 13, 11^{13}$	1940.41	1875.18	3.478	6*	$21 \to 13, 9^{13}$	1935.18
23	$22 \to 14, 10^{14}$	2228.39	2178.68	2.282	5*	$22 \to 14, 9^{14}$	2209.96
24	$23 \to 15, 9^{15}$	2534.37	2477.98	2.276	3*	$23 \to 13, 12^{13}$	2517.07
25	$24 \to 13, 12^{13}$	2867.48	2846.02	0.754	1*	$24 \to 15, 11^{15}$	2849.07
26	$25 \to 15, 14^{15}$	3241.42	3192.98	1.518	4*	$25 \to 15, 12^{15}$	3232.42
27	$26 \to 17, 11^{17}$	3615.38	3537.90	2.190	5*	$26 \to 15, 12^{15}$	3637.24
28	$27 \to 17, 12^{17}$	4077.40	3956.16	3.065	8*	$27 \to 16, 12^{16}$	4068.93
29	$28 \to 17, 14^{17}$	4547.42	4477.00	1.573	6*	$28 \to 16, 13^{16}$	4534.32
30	$29 \to 18, 15^{18}$	5021.40	4809.73	4.401	12*	$29 \to 17, 13^{17}$	5004.40

Table 7. Comparison between optimal strategy and AGX-III results

Remark 1. *The sequences could be defined by the single term* $(d_B)^{n-k}$*. The extended form used before aims for better clarity.*

We can see that for all orders up to $n = 19$, AGX-III found structures corresponding to extremal IRR_d values. For $n = 25$ and $d_B = 12$, the upper bound was so close to the corresponding IRR_d value that AGX-III was unable to find an extremal structure.

4.4 Exploring IRR_m landscape

We can divide the degree sequences in a way similar to that used with IRR_d. Here, we call A the set of greater degrees with their multiplicity values and B the set of the remaining, and value-non-increasing, degrees. Preliminary results for small graphs are given in Table 8, where we can notice some facts: from $n = 4$ to 10 the diversities form a non-decreasing sequence; A has only one repeated element equal to $n - 1$; the minimum degree of the set B is equal to $\mu(d_A)$.

Order	A	B	Divers.	Order	A	B	Divers.
4	3^2	2^2	2	8	7^5	$6^2, 5$	3
5	4^3	3^2	2	9	8^6	$7^2, 6$	3
6	5^4	4^2	2	10	9^6	$8, 7^2, 6$	4
7	6^5	5^2	2				

Table 8. Partitioning of vertex sets according to degrees ($n = 4$ to 10)

Table 9 below gives the best IRR_m values found and the degree sequences corresponding to the associated graphs. A continuously descending degree subsequence is indicated by $a \to b$, where a and b are their maximum and minimum values.

n	IRR_m value	Degree seq. subsets		Size	Diversity
		A	B	m	
10	1056.30	9^6	$8, 7^2, 6$	41	4
11	1638.27	10^7	$9, 8^2, 7$	51	4
12	2286.33	11^8	$10, 9^2, 8$	62	4
13	3366.23	12^9	$11, 10^2, 9$	74	4
14	4612.29	13^9	$12, 11^2 \to 9$	85	5
15	6204.27	14^10	$13, 12^2 \to 10$	99	5
16	8122.25	15^10	$14, 13^2, 12, 11$	114	5
17	10522.30	16^11	$15 \to 13^2, 12, 11$	127	6
18	13398.30	17^12	$16 \to 14^2 \to 12$	144	6
19	16750.30	18^13	$17 \to 15^2 \to 13$	162	6
20	20712.30	19^13	$18 \to 16^2 \to 13$	178	7
21	25412.30	20^14	$19 \to 17^2 \to 14$	198	7
22	30776.30	21^15	$20 \to 18^2 \to 15$	219	7
23	36816.30	22^16	$21 \to 19^2 \to 16$	241	7
24	44182.30	23^16	$22 \to 19^2 \to 16$	260	8
25	52212.30	24^17	$23 \to 20^2 \to 17$	284	8
26	61150.30	25^18	$24 \to 21^2 \to 18$	309	8
27	71680.30	26^18	$25 \to 22^2 \to 18$	331	9
28	83146.30	27^19	$26 \to 23^2 \to 19$	358	9
29	95760.30	28^20	$27 \to 24^2 \to 20$	386	9
30	110452.0	29^20	$28 \to 24^2 \to 20$	410	10

Table 9. Maximum IRR_m values found by AGX-III and their degree sequences

4.4.1 Some discussion on the results

From the results obtained in Table 7, we observe some interesting points on the structure of (presumably) IRR_m-extremal graphs: all of them are supergraphs of a complete split graph where the set A induces a clique, say of order $|A| = n - t$, and the set B induces the complement of an antiregular graph of order $|B| = t$, that is, G is isomorphic to $K_{n-t} \bowtie \overline{AR_t}$; the degree sequence of B decreases from $n-2$ to $n-t$ and the vertex of degree $n - \lfloor t/2 \rfloor - 1$ has multiplicity 2 and the others multiplicity one. These points were used as a basis for the following conjecture:

Conjecture 3. *For any order, an IRR_m-extremal graph $G = (V, E)$ is a join of a clique of order $n-t$ and a complement of an antiregular graph with order t. Besides, the degree sequence of G is given by $d_G = (n-1)^{n-t}, n-2, n-3, \cdots, (n - \lfloor t/2 \rfloor - 1)^2, \cdots, n-t$.*

Figure 4 shows extremal graphs with orders 12 and 18, where the (A, B) partition, the antiregular structure and the degree sequence can be observed.

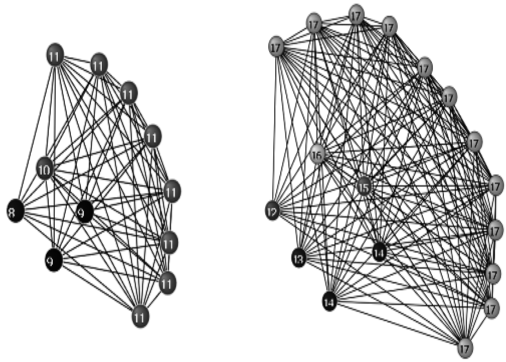

Figure 4. IRR_m-extremal graphs of orders 12 and 18

Remark 2. *From Merris [20], every antiregular graph is threshold. Besides that, if we iteratively add universal vertices to a threshold graph, the graph thus obtained is always threshold. As a consequence, the extremal graphs of Conjecture 3 are threshold.*

Discussion: The edge set E can be partitioned into three subsets:

$$(A,A) = \{(u,v)|u \in A, v \in A\}; (A,B) = \{(u,v)|u \in A, v \in B\};$$
$$(B,B) = \{(u,v)|u \in B, v \in B\}.$$

In what follows, let $d_{min(B)}$ and $d_{max(B)}$ respectively be the minimum and maximum degrees of B. From the results, the following equalities hold: $\mu(d_A) = d_{min(B)} = n-t$, which implies that the diversity is given by t, and $d_{max(B)} = n-2$. Also, $d_A = n-1$. From IRR_m definition, the edges (A,A) do not contribute to the summation. The edge in (A,B) that contributes the most to the irregularity is the one connecting the vertices of degrees d_A and $d_{min(B)}$ and its contribution is equal to $(n-t)(n-2)$. The lesser contribution is $(n-1)(n-t-1)+1$, given by connecting the vertices of degrees d_A and $d_{max(B)}$. The edges of the set (B,B) contribute with values between 0 and $t-2$. It is interesting to note that, unless the IRR_d-extremals, the degree sequences of IRR_m-extremal graphs are unigraphic. This fact occurs for the following reasons: the $n-t$ vertices of degree $n-1$ should be connected to every vertex which generates a complete split graph; the residual degrees are $0,1,2,\cdots,t-1$, and the remaining edges should only connect vertices in B. Since both (complementary) antiregular sequences are unigraphic, and the residual degrees correspond to the sequence of a disconnected antiregular graph, the whole degree sequence is unigraphic. Then it should be fruitless to search for graphs with better IRR_m values as we have done for IRR_d.

5 Conclusions and suggestions for future research

5.1 An interesting feature of IRR_d-extremal graphs is the existence of a two-subset partition, $V = (A, B)$, defined according to their degree values. The descending-degree definition for set A, starting with $d_1 = n - 1$, obliges $|B|$ cardinality to be equal to the lesser degree in A. Moreover, we always have $d_{min(A)} - d_B \geq 1$ and, exploring the strict inequality values we can find, for some orders, better IRR_d values. The extremal structures, nevertheless, were always similar to those proposed by AGX-III and we conjecture there is no other structure with higher IRR_d values. Using an extremal-value strategy for graph building of AGX-III typical structure, we were able to derive upper bounds for IRR_d. This was initially used when we looked for an extremal 25-vertex graph, an order where the very low slack given by the upper bound shows the problem as particularly difficult. For this order, AGX-III found a local optimum of value 2839.58 with structure $A = 24 \to 15, 13$ and $B = 12^{14}$, only 0.21% far from a better value we found to correspond to $A = 24 \to 13$ and $B = 12^{13}$ by following the lead of the typical structure. Further exploration gave an even higher value, as shown by Table 7.

5.2 In what concerns IRR_m, it seems that the sequence degree of the extremal graphs follows some sort of a pattern where the vertex set can be partitioned into two sets A and B, $V = (A, B)$, where A is a clique of order $n - t$ and B is isomorphic to $\overline{AR_t}$. The complementary graph $\overline{G}$ of the extremal one has an even clearer structure since it is a union of an independent set of order $n - t$ (then, a trivial graph) to AR_t. We conjecture that the same structural properties hold for any IRR_m-extremal graph and, consequently, that the degree sequences of the extremal ones are unigraphic, since this is valid both for antiregular and trivial graphs.

5.3 Both from the graph-theoretical and the numerical comparison point of view, we think there are multiple interesting points to explore in what concerns these two measures. They can certainly allow for the research of still better measures, since the extremal graphs for our measure are not antiregular graphs, which, from a numerical point of view, are the most irregular among all graphs.

References

[1] H. Abdo, D. Dimitrov, I. Gutman, Graphs with maximal σ irregularity, *Discr. Appl. Math.* **250** (2018) 57–64.

[2] M. O. Albertson, The irregularity of a graph, *Ars Comb.* **46** (1997) 219–225.

[3] A. Balaban, Distance connectivity index, *Chem. Phys. Lett.* **89** (1982) 399–404.

[4] F. K. Bell, On the maximal index of connected graphs, *Lin. Algebra Appl.* **144** (1991) 135–151.

[5] C. Berge, *Graphes et Hypergraphes*, Dunod, Paris, 1971.

[6] J. A. Bondy, U. S. R. Murty, *Graph Theory and Applications*, Macmillan, London, 1978.

[7] P. O. Boaventura–Netto, P-antiregular graphs, *Adv. Appl. Discr. Math.* **1** (2008) 133–148.

[8] P. O. Boaventura–Netto, QAPV: a polynomial invariant for graph isomorphism testing, *Pesquisa Operacional.* **33** (2013) 163–184.

[9] P. O. Boaventura–Netto, Structural versus nonstructural irregularity measures, *Proceedings of the ALIO-EURO, Uruguay,* 2014.

[10] P. O. Boaventura–Netto, *Grafos: Teoria, modelos, algoritmos.*, Blucher, São Paulo, 2012.

[11] P. O. Boaventura–Netto, Graph irregularity: discussion, graph extensions and new proposals, *Rev. Mat. Teoría Apl.* **22** (2015) 293–310.

[12] G. Caporossi, Variable neighborhood search for extremal vertices: the AutoGraphiX-III system, *Comput. Oper. Res.* **78** (2017) 431–438.

[13] G. Caporossi, P. Hansen, Variable neighborhood search for extremal graphs: 1 The AutoGraphiX system, *Discr. Math.* **212** (2000) 29–44.

[14] I. Gutman, P. Hansen, H. Mélot, Variable neighborhood search for extremal graphs. 10. Comparison of irregularity indices for chemical trees, *J. Chem. Inf. Model.* **45** (2005) 222–230.

[15] I. Gutman, M. Togan, A. Yurttas, A. Cevik, I. Cangul, Inverse problem for sigma index, *MATCH Commun. Math. Comput. Chem.* **79** (2018) 491–508.

[16] P. Hansen, N. Mladenović, Variable neighborhood search: principles and applications, *Eur. J. Oper. Res.* **130** (2001) 449–467.

[17] F. Harary, *Graph Theory*, Addison–Wesley, Reading, 1987.

[18] P. Hansen, H. Mélot, Variable neighborhood search for extremal graphs. 9. Bounding the irregularity of a graph, in: S. Fajtlowicz (Ed.), *Graphs and Discovery*, Am. Math. Soc., Providence, 2005, pp. 253–264.

[19] R. Merris, Split graphs, *Eur. J. Comb.* **24** (2003) 413–430.

[20] R. Merris, Antiregular graphs are universal for trees, *Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat.* **14** (2003) 1–3.

[21] N. Mladenović, P. Hansen, Variable neighborhood search, *Comput. Oper. Res.* **24** (1997) 1097–1100.

[22] B. D. McKay, A. Piperno, Practical graph isomorphism II, *J. Symb. Comput.* **60** (2013) 94–112.

[23] J. A. Oliveira, *Medidas de irregularidade em grafos*, D.Sc. Thesis, COPPE/UFRJ, Rio de Janeiro, 2012.

[24] J. A. Oliveira, C. S. Oliveira, C. Justel, N. M. M. de Abreu, Measures of irregularity of graphs, *Pesq. Oper.* **33** (2013) 383–398.

[25] D. C. Porumbel, Isomorphism testing via polynomial-time graph extensions, *J. Math. Model. Alg.* **10** (2011) 119–143.

[26] R. C. Read, R. J. Wilson, *An Atlas of Graphs*, Clarendon, Oxford, 1998.

[27] G. Sierksma, H. Hoogeven, Seven criteria for integer sequences being graphic, *J. Graph Theory* **15** (1991) 223–231.