Iranian Journal of Mathematical Chemistry, Vol. 3, No. 1, February 2012, pp. 35 – 43 **IJMC**

Centric connectivity index by shell matrices

MIRCEA V. DIUDEA•

Faculty of Chemistry and Chemical Engineering Babes-Bolyai University, 3400 Cluj, Romania

(*Received January 10, 2012*)

ABSTRACT

Relative centricity *RC* values of vertices/atoms are calculated within the Distance Detour and Cluj-Distance criteria on their corresponding Shell transforms. The vertex *RC* distribution in a molecular graph gives atom equivalence classes, useful in interpretation of NMR spectra. Timed by vertex valences, *RC* provides a new index, called Centric Connectivity *CC*, which can be useful in the topological characterization of graphs and in QSAR/QSPR studies.

Keywords: Graph theory, Cluj matrix, Relative centricity, Centric connectivity index.

1. INTRODUCTION

Let $G = (V, E)$ be a connected graph, with no multiple bonds and loops. V is the set of vertices and E is the set of edges in G, $v = |V(G)|$ and $e = |E(G)|$ being their cardinalities. A *walk w* is an alternating string of vertices and edges: $w_{1,n} = (v_1, e_1, v_2, e_2, ..., v_{n-1}, e_m, v_n)$, with the property that any subsequent pair of vertices represent an edge: $(v_{i-1}, v_i) \in E(G)$.

A *path p* is a walk having all its vertices and edges distinct: $v_i \neq v_j$, $(v_{i-1}, v_i) \neq (v_{j-1}, v_j)$ for any $1 \leq i < j \leq n$. The *length* of a path is $l(p_{1,n}) = |E(p_{1,n})| = |V(p_{1,n})| - 1$, with $V(p_{1,n})$ being the vertex set of the path $p_{1,n}$. A closed path is a *cycle* (i.e., *circuit*).

The *distance* d_{ij} is the length of a *shortest* path joining vertices v_i and v_j: $d_{ij} = min\ l(p_{ij})$; otherwise $d_{ij} = \infty$. The *detour* δ_{ij} is the length of a *longest* path between vertices v_i and v_j: $\delta_{ij} = max\ l(p_{ij})$; otherwise $\delta_{ij} = \infty$ [1,2].

The square arrays that collect the distances and detours, in G are called the *Distance* **DI** and *Detour* **DE** matrix**,** respectively [3-5].

$$\left[\mathbf{DI}(G)\right]_{i,j} = \begin{cases} \min l(p_{i,j}), \text{ if } i \neq j \\ 0 \text{ if } i = j \end{cases} \quad (1)$$

•Corresponding author (Email: diudea@gmail.com).

$$|\mathbf{DE}(G)|_{i,j} = \begin{cases} \max l(p_{i,j}), \text{if } i \neq j \\ 0 \quad \text{if } i = j \end{cases} \tag{2}$$

In words, these matrices collect the number of edges separating the vertices *i* and *j* on the shortest and longest path $p_{i,,j}$, respectively. The half sum of entries in the Distance and Detour matrices provide the well-known Wiener index *W* [6] and its analogue, the detour number *w* [7,8].

The *Cluj fragments* are sets of vertices obeying the relation [3-5,9-12]:

$$CJ_{i,j,p} = \left\{ v \,\middle|\, v \in V(G); D_{(G-p)}(i,v) < D_{(G-p)}(j,v) \right\} \tag{3}$$

The entries in the Cluj matrix **UCJ** are taken, by definition, as the maximum cardinality among all such fragments

$$\mathbf{[UCJ]}_{i,j} = \max_{p} \left| CJ_{i,j,p} \right| \tag{4}$$

It is because, in graphs containing rings, more than one path can join the pair (*i, j*), thus resulting more than one fragment related to *i* (with respect to *j* and path *p*).

The Cluj matrix is defined by using either *distances* or *detour*s [13]: when the path *p* belongs to the set of distances DI(*G*), the suffix DI is added to the name of matrix, as in UCJDI. When the path *p* belongs to the set of detours DE(*G*), the suffix is DE.

Two graphs are called *isomorphic*, G ≈ G', if there exists a mapping $f : V \rightarrow V'$ that preserves adjacency (*i.e.*, if (*i,j*)∈ E(G), then (*f* (*i*), *f* (*j*))∈ E'(G')). The function *f* provides a one-to-one correspondence between the vertices of the two sets. The isomorphism of *G* with itself is called an automorphism. It is demonstrated that all the automorphisms of *G* form a group, *Aut*(G) [3,4].

The symmetry of a graph is often called a topological symmetry. It is defined in terms of *connectivity*, as a constitutive principle of molecules and expresses equivalence relationships among elements of the graph: vertices, bonds, faces or larger subgraphs. The topological symmetry does not fully determine the molecular geometry and it does not need to be the same as (*i.e.*, isomorphic to) the molecular point group symmetry. However, it represents the maximal symmetry which the geometrical realization of a given topological structure may posses [14-16].

Given a graph *G*=(*V,E*) and a group Aut(*G*), two vertices, *i, j*∈V are called *equivalent* if there is a group element, aut(v_i)∈Aut(*G*), such that *j* aut(v_i) *i*. The set of all

vertices *j* (obeying the *equivalence relation*) is called the *i*'s class of equivalence. Two vertices *i* and *j*, showing the same vertex invariant $In_i=In_j$ belong to the same *invariant class IC*. The process of vertex partitioning in *IC*-s leads to *m* classes, with $v_1, v_2,...v_m$ vertices in each class. Note that invariant-based partitioning may differ from the orbits of automorphism since no vertex invariant is known so far to discriminate two non-equivalent vertices in any graph [3,4].

In the chemical field, the isomorphism search could answer to the question if two molecular graphs represent or not one and the same chemical compound. Two isomorphic graphs will show the same value of a topological index, so that they cannot be distinguished by topological descriptors.

2. CENTRIC CONNECTIVITY CC INDEX

In studies on the centrality/centricity of graphs, Bonchev *et al*. [17,18] have proposed the distance-based criteria *1D-3D* as follows:

1D*:* minimum vertex eccentricity: min ecc_i

2D*:* minimum vertex distance sum*:* min DIS_i

3D*:* minimum number of occurrence of the largest distance*:* min $\mathbf{[LM, ShM]}_{i,j\ \max}$

When applied hierarchically, the above criteria lead to the center(s) of a graph. In the above, **LM**, **ShM** denote the layer matrix and the shell matrix (of a given square info-matrix M), defined as follows [19-21].

The entries in the layer matrix (of vertex property) **LM** are defined as

$$\left[\mathbf{LM}\right]_{i,k} = \sum_{v|d_{i,v}=k} p_v \tag{5}$$

Layer matrix is a collection of the above defined entries

$$\mathbf{LM}(G) = \left\{ \left[\mathbf{LM}\right]_{i,k}; i \in V(G); k \in [0,1,..,d(G)] \right\} \tag{6}$$

with $d(G)$ being the diameter of the graph (*i.e*., the largest distance in *G*). Any atomic/vertex property can be considered as p_i. More over, any square matrix **M** can be taken as *info matrix*, *i.e.*, the matrix supplying local/vertex properties as row sum *RS* or column sum *CS*. The zero column is just the column of vertex properties, $\left[\mathbf{LM}\right]_{i,0} = p_i$. When the vertex property is 1 (*i.e*., the counting property), the **LM** matrix is called **LC** (the Layer matrix of Counting).

Let define the entries in the shell matrix **ShM** (of pair vertex property) as [21]

$$\left[\mathbf{ShM}\right]_{i,k} = \sum_{v \mid d_{i,v}=k} [\mathbf{M}]_{i,v} \tag{7}$$

The shell matrix is a collection of the above defined entries

$$\mathbf{ShM}(G) = \left\{ [\mathbf{ShM}]_{i,k}; i \in \mathrm{V}(G); k \in [0,1,..,d(G)] \right\} \tag{8}$$

A shell matrix **ShM**(*G*) partitions the entries of the square matrix **M** according to the vertex (distance) partitions in the graph. It represents a true decomposition of the property collected by the info square matrix according to the contributions brought by pair vertices pertaining to shells located at the distance *k* around each vertex. The zero column entries $\left[\mathbf{ShM}\right]_{i,0}$ are just the diagonal entries in the info matrix.

The distance-based functions, expressing the topology related to the center of the graph, are as follows:

$$P(i)_k = [\mathbf{LM}, \mathbf{ShM}]_{i,k} \tag{9}$$

$$CP(i) = \sum_k P(i)_k \cdot k^{-n}; k = 1,2,...d(G); n = 1,2,.. \tag{10}$$

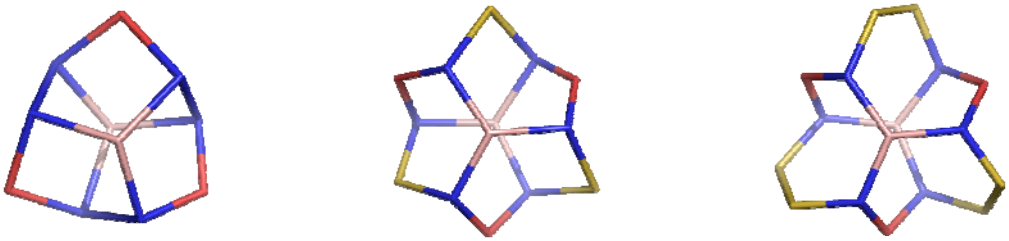
$$RC(i) = CP(i) / CP(i)_{\max} \tag{11}$$

$$RC(G) = \sum_i RC(i) \tag{12}$$

$$CC(i) = RC(i) \cdot d(i) \tag{13}$$

$$CC(G) = \sum_i CC(i) \tag{14}$$

The property *P*(*i*) (collected in **LM** or **ShM**, (9)) is transformed in a centric property *CP*(i) (10) by multiplying with the rank *k* of a given layer, raised at a negative power, summation running up to the diameter of *G*. *CP*(i) is thus calculated in the sense of the Bonchev's 1D-3D criteria, by virtue of the involved **LM, ShM** matrices. There is a clear difference between *CP*(i) and the eccentricity $\varepsilon(i)$ (counting the largest topological distance from *i* to any other vertex in *G*), used in the construction of "Eccentric Connectivity index" [22]. The relative vertex centricity *RC*(i) (11) accounts for the deviation to the maximum centricity, equaling 1 in case of vertices being centers of the graph. The global value RC(G) (12) is useful in characterizing the distribution of the centricity function (10), particularly when is normalized by the number of vertices *v* of *G*, *RC*(*G*)/*v* or by the product between the number of vertices and the number of invariant classes, *RC*(*G*)/ (*v*×*IC*).



G_1 ; v=11; IC=3 G_2 ; v=14; IC=4 G_3 ; v=17; IC=4

Figure. Graphs with Only Pentagons G_1, Hexagons G_2 and Heptagons G_3.

Finally, the centric connectivity CC index (13,14) is hoped to be useful in QSAR/QSPR studies, their values being of the same order of magnitude as the number of vertices/atoms in the molecular graph. Relation (13) can be generalized by changing d(i) by the "remote" degrees [4,23] or by degrees of "extended connectivity" [24-28].

Tables 1 to 3 exemplify the above formulas for the molecular graphs illustrated in Figure. The invariant classes of equivalence *IC*-s are given at the bottom of tables, by their population (i.e., the number of vertices in each class). *IC*-s are important in NMR spectra interpretation.

Table 1. Shell matrix of Cluj matrix **CJDI**, vertex Centric property $CP(i)$, Relative centricity $RC(i)$ and Centric connectivity $CC(i)$ values, the vertex degree $d(i)$, number of invariant classes *IC*-s (in decreasing order of Centricity), the global relative centricity *RC* for the graph G_1 and its normalized values *RC/v* and *RC/vIC*

	ShCJDI				ShCJDI(i) × k^{-2}						
#G_1	1	2	3	4	1	2	3	$d(i)$	CP_i	RC_i	CC_i
1	13	21	9	13	13	5.25	1	3	19.2500	0.8390	2.5169
2	13	21	9	13	13	5.25	1	3	19.2500	0.8390	2.5169
3	13	21	9	13	13	5.25	1	3	19.2500	0.8390	2.5169
4	15	30	4	15	15	7.5	0.4444	3	22.9444	1	3
5	4	12	14	4	4	3	1.5556	2	8.5556	0.3729	0.7458
6	13	21	9	13	13	5.25	1	3	19.2500	0.8390	2.5169
7	4	12	14	4	4	3	1.5556	2	8.5556	0.3729	0.7458
8	13	21	9	13	13	5.25	1	3	19.2500	0.8390	2.5169
9	13	21	9	13	13	5.25	1	3	19.2500	0.8390	2.5169
10	15	30	4	15	15	7.5	0.4444	3	22.9444	1	3
11	4	12	14	4	4	3	1.5556	2	8.5556	0.3729	0.7458
								RC	*CC*	8.1525	*23.3390*
								RC/v	*RC/vIC*	0.7411	0.2470

Table 2. Shell matrix of Detour matrix **DE**, vertex Centric property *CP*(*i*), Relative centricity *RC*(*i*) and Centric connectivity *CC*(*i*) values, the vertex degree *d*(*i*), number of invariant classes *IC*-s (in decreasing order of Centricity), the global relative centricity *RC* for the graph G_2 and its normalized values *RC*/*v* and *RC*/*vIC*

ShDE					ShDE(*i*)× k^{-2}					*IC*=4:((2,6,6*); 6*= (3,3))		
#G_2	1	2	3	4	1	2	3	4	*d*(*i*)	CP_i	RC_i	CC_i
1	33	40	55	10	33	10	6.1111	0.6250	3	49.7361	0.8975	2.6925
2	33	40	55	10	33	10	6.1111	0.6250	3	49.7361	0.8975	2.6925
3	33	40	55	10	33	10	6.1111	0.6250	3	49.7361	0.8975	2.6925
4	33	72	33	12	33	18	3.6667	0.7500	3	55.4167	1	3
5	22	48	44	36	22	12	4.8889	2.2500	2	41.1389	0.7424	1.4847
6	22	48	44	36	22	12	4.8889	2.2500	2	41.1389	0.7424	1.4847
7	33	40	55	10	33	10	6.1111	0.6250	3	49.7361	0.8975	2.6925
8	22	48	44	36	22	12	4.8889	2.2500	2	41.1389	0.7424	1.4847
9	33	40	55	10	33	10	6.1111	0.6250	3	49.7361	0.8975	2.6925
10	33	40	55	10	33	10	6.1111	0.6250	3	49.7361	0.8975	2.6925
11	33	72	33	12	33	18	3.6667	0.7500	3	55.4167	1	3
12	22	48	44	36	22	12	4.8889	2.2500	2	41.1389	0.7424	1.4847
13	22	48	44	36	22	12	4.8889	2.2500	2	41.1389	0.7424	1.4847
14	22	48	44	36	22	12	4.8889	2.2500	2	41.1389	0.7424	1.4847
									RC	*CC*	11.8391	31.0632
									RC/*v*	*RC*/*vIC*	0.8456	0.2114

Table 3. Shell matrix of Distance matrix **DI**, vertex Centric property *CP*(*i*), Relative centricity *RC*(*i*) and Centric connectivity *CC*(*i*) values, no. of "invariant classes" *IC*-s (in decreasing order of Centricity), global relative centricity *RC* for the graph G_3 and its normalized values *RC*/*v* and *RC*/*vIC*

ShDI							ShDI(*i*) = ShDI× k^{-2}							*IC*=4: (2,6,3,6)		
#G_3	1	2	3	4	5		1	2	3	4	5		*d*(*i*)	CP_i	RC_i	CC_i
1	3	8	21	8	0	41	3	2	2.333	0.5	0	3	3	7.8333	0.9495	2.8485
2	3	8	21	8	0	41	3	2	2.333	0.5	0	3	3	7.8333	0.9495	2.8485
3	3	8	21	8	0	41	3	2	2.333	0.5	0	3	3	7.8333	0.9495	2.8485
4	3	12	18	4	0	38	3	3	2	0.25	0	3	3	8.25	1	3
5	2	6	15	20	5	49	2	1.5	1.667	1.25	0.2	2	2	6.6167	0.8020	1.6040
6	2	8	18	16	0	45	2	2	2	1	0	2	2	7	0.8485	1.6970
7	3	8	21	8	0	41	3	2	2.333	0.5	0	3	3	7.8333	0.9495	2.8485
8	2	8	18	16	0	45	2	2	2	1	0	2	2	7	0.8485	1.6970
9	3	8	21	8	0	41	3	2	2.333	0.5	0	3	3	7.8333	0.9495	2.8485
10	2	6	15	20	5	49	2	1.5	1.667	1.25	0.2	2	2	6.6167	0.8020	1.6040
11	3	8	21	8	0	41	3	2	2.333	0.5	0	3	3	7.8333	0.9495	2.8485
12	3	12	18	4	0	38	3	3	2	0.25	0	3	3	8.25	1	3
13	2	6	15	20	5	49	2	1.5	1.667	1.25	0.2	2	2	6.6167	0.8020	1.6040
14	2	8	18	16	0	45	2	2	2	1	0	2	2	7	0.8485	1.6970
15	2	6	15	20	5	49	2	1.5	1.667	1.25	0.2	2	2	6.6167	0.8020	1.6040
16	2	6	15	20	5	49	2	1.5	1.667	1.25	0.2	2	2	6.6167	0.8020	1.6040
17	2	6	15	20	5	49	2	1.5	1.667	1.25	0.2	2	2	6.6167	0.8020	1.6040
													RC	*CC*	15.0545	37.8061
													RC/*v*	*RC*/*vIC*	0.8856	0.2214

According to the normalized relative centricity *RC/vIC*, the ordering of the above graphs (cf. Tables 1 to 3) is: G_1 (0.2470) > G_3 (0.2214) > G_2 (0.2114). Remark that the detour **DE** matrix provides degenerate *IC*-classes, so that in the above calculation $IC(G_2)=4$ was taken.

3. CONCLUSIONS

The relative centricity *RC*(*i*) of vertices in a molecular graph were calculated within the Distance, Detour and Cluj-Distance criteria. The *RC*(*i*) distribution in a molecular graph gives information on the equivalence classes (as vertex invariant classes) of atoms, useful in the interpretation of NMR spectra. Timed by vertex valences, *RC* provides an index, called Centric Connectivity *CC*. By definition, there is a clear difference between the Centric Connectivity CC(G) index and the older Eccentric Connectivity index or its newer versions [29-31]. This index represents a new descriptor (an improvement of the above version, ref. [32]) which can be useful in the topological characterization of graphs and in QSAR/QSPR studies.

ACKNOWLEDGEMENT. The work was supported by the Romanian CNCSIS-UEFISCSU project PN-II-ID-PCE-2011-3-0346.

REFERENCES

1. N. Trinajstić, *Chemical Graph Theory,* CRC Press: Boca Raton, FL, 1983.
2. F. Harary, *Graph Theory*, Addison-Wesley, Reading, M.A., 1969.
3. M. V. Diudea, I. Gutman, and L. Janschi, *Molecular Topology*, Nova Science, Huntington, N. Y., 2001.
4. M. V. Diudea, Nanomolecules and Nanostructures, Polynomials and Indices, MCM, No. **10,** Univ. Kragujevac and Fac. Sci. Kragujevac, Serbia, 2010.
5. M. V. Diudea, M. S. Florescu, and P. V. Khadikar, *Molecular Topology and Its Applications*, **EFICON**, Bucharest, 2006.
6. H. Wiener, Structural Determination of Paraffin Boiling points, *J.Am.Chem.Soc.* **1947,** *69*, 17-20.
7. I. Lukovits, The Detour Index, *Croat. Chem. Acta*, **1996**, *69*, 873-882.
8. I. Lukovits and M. Razinger, On Calculation of the Detour Index, *J.Chem.Inf.Comput.Sci*. , **1997**, *37*, 283-286.
9. M. V. Diudea, Cluj Matrix Invariants. *J. Chem. Inf. Comput. Sci*. **1997**, *37*, 300-305.
10. M. V. Diudea, Cluj Matrix $\mathbf{CJ_u}$: source of various graph descriptors, *Commun. Math. Comput. Chem.* (*MATCH*), **1997**, *35*, 169-183.

11. M. V. Diudea and I. Gutman, Wiener-Type Topological Indices. *Croat. Chem. Acta,* **1998**, *71*, 21-51.
12. M. V. Diudea, Valencies of Property. *Croat. Chem. Acta,* **1999**, *72*, 835-851.
13. M. V. Diudea, B. Parv, and I. Gutman, Detour-Cluj Matrix and Derived Invariants *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 1101-1108.
14. C. Y. Hu and L. Xu, Algorithm for Computer Perception of Topological Symmetry. *Anal. Chim. Acta,* **1994**, *295*, 127-134.
15. G. S. Ezra, *Symmetry Properties of Molecules*, Lecture Notes in Chemistry 28, Springer, 1982.
16. M. Razinger, K. Balasubramanian, and M. E. Munk, Graph Automorphism Perception Algorithms in Computer-Enhanced Structure Elucidation. *J. Chem. Inf. Comput. Sci.*, **1993**, *33*, 197-201.
17. Bonchev, D.; Balaban, A.T.; Randić, M. The Graph Center Concept for Polycyclic Graphs, *Int. J. Quantum Chem.* **1981**, *19*, 61-82.
18. Bonchev, D.; Mekenyan, O.; Balaban, A.T. Iterative Procedure for the Generalized Graph Center in Polycyclic Graphs, *J. Chem. Inf. Comput. Sci*. **1989**, *29*, 91-97.
19. M. V. Diudea, Layer Matrices in Molecular Graphs, *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1064-1071.
20. M. V. Diudea, M. Topan, and A. Graovac, Layer Matrices of Walk Degrees, *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1071 -1078.
21. M. V. Diudea and O. Ursu, Layer matrices and distance property descriptors. *Indian J. Chem*., 42*A*, **2003**, 1283-1294.
22. V. Sharma, R. Goswami, A. K. Madan, Eccentric connectivity index: A novel highly discriminating topological descriptor for structure property and structure activity studies, J. Chem. Inf. Comput. Sci. 37 (1997) 273-282.
23. P. E. John and M. V. Diudea, The second distance matrix of the graph and its characteristic polynomial, *Carpath. J. Math.*, **2004**, *20* (2), 235-239.
24. T. Balaban, O. Mekenyan, and D. Bonchev, Unique Description of Chemical Structures Based on Hierarchically Ordered Extended Connectivities (HOC Procedures). I. Algorithms for Finding graph Orbits and Cannonical Numbering of Atoms, *J. Comput. Chem.* **1985**, *6*, 538-551.
25. A. T. Balaban, O. Mekenyan, and D. Bonchev, Unique Description of Chemical Structures Based on Hierarchically Ordered Extended Connectivities (HOC Procedures). II. Mathematical Proofs for the HOC Algorithm, *J. Comput. Chem.* **1985**, *6*, 552-561.
26. O. Mekenyan, A. T. Balaban, and D. Bonchev, Unique Description of Chemical Structures Based on Hierarchically Ordered Extended Connectivities (HOC Procedures). VI. Condensed Benzenoid Hydrocarbons and Their 1H-NMR Chemical Shifts. *J. Magn. Reson.* **1985**, *63*, 1-13.

27. A.T. Balaban, A. T.; Moțoc, I.; Bonchev, D.; Mekenyan, O. Topological Indices for Structure - Activity Correlations, *Top. Curr. Chem*. **1993**, *114*, 21-55.
28. H Morgan, The generation of a unique machine description for chemical structures. A technique developed at Chemical Abstracts Service, *J. Chem. Doc.* **1965**, *5*, 107-113.
29. A. Ilić, I. Gutman, Eccentric Connectivity Index of Chemical Trees, *MATCH Commun. Math. Comput. Chem*. 65 (2011), 731-744.
30. A. Ilić, Eccentric connectivity index, in: I. Gutman, B. Furtula, Novel Molecular Structure Descriptors -Theory and Applications II, MCM 9, University of Kragujevac, 2010.
31. G. Yu, L. Feng, A. Ilić, On the eccentric distance sum of trees and unicyclic graphs, J. Math. Anal. Appl. 375 (2011), 934-944.
32. M. V. Diudea, Centric connectivity index, *Studia Univ. Babes-Bolyai, Chemia*, **2010**, *55* (*4*), 319-324.