ACADEMIA ROMÂNĂ
Revue Roumaine de Chimie
http://web.icf.ro/rrch/

Rev. Roum. Chim.,
2016, *61*(4-5), 231-238

Dedicated to Professor Alexandru T. Balaban
on the occasion of his 85th anniversary

ON THE MOLECULAR STRUCTURE OF THE CO-AUTHOR NETWORK OF ALEXANDRU T. BALABAN

Tibor BRAUN,[a,*] András SCHUBERT[b] and Gábor SCHUBERT[c]

[a] ELTE, Budapest, Hungary
[b] MTAKIK, Budapest, Hungary
[c] Stockholm University, Library, Stockholm, Sweden

Received December 21, 2015

Scientometric networks based on the bibliography of Professor Alexandru T. Balaban are analyzed. His position in the topological index research community and the structure of his own co-author network are considered. An analogy is suggested between co-author networks and molecular structures.

INTRODUCTION

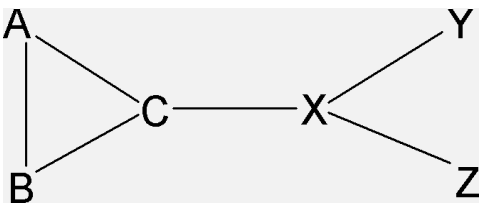
Bibliographic information is usually presented in the form of lists. Lists of publications, lists of authors, lists of citations are compiled and published vastly both in printed and in electronic format. The lists are usually ordered – chronologically, alphabetically, ranked by frequency, etc. –, nevertheless, they are irremediably one-dimensional, linear.

The reality is, however, more complex. Whatever bibliographic items are considered, they form structures. These structures are usually modelled by graphs or even hypergraphs. "Network scientometrics" is nowadays an established area of scientometric research encompassing sophisticated theoretical considerations,[1] mapping techniques[2,3] and even evaluative aspects.[4]

Since all three authors of the present paper are chemists, it is not surprising that the most evident analogy coming into mind is that of atoms and molecules.[5]

One of the most intensively studied examples is co-authorship networks. The usual representation of co-authors of a paper is a set of names, either in alphabetical order or not, occasionally equipped with weights of contribution. Regarding each author as an "atom", this corresponds to the "empirical molecular formula", such as C_2H_6O for methanol. Molecules, however, have their structural formula as well.

Co-author sets have a definite structure, just like molecules have their own. A typical co-author structure may look something like this:



* Corresponding author: braun@mail.iif.hu

"Bonds" may represent direct communication during the research, or simply shared affiliation.

These structural aspects may have crucial significance, *e.g.*, if co-authorship patterns are used for studying the sociology of scientific collaboration.

THE POSITION OF PROFESSOR BALABAN IN HIS RESEARCH AREA

Professor Balaban is highly esteemed for his contributions to mathematical chemistry, more particularly to the area of chemical graph theory and topological indices.

His prominent position in this latter subarea is well reflected in a recent overview of Bodlaj & Batagelj.[6] In a WoS-based sample of 2036 papers on topological indices in the period 1975-2011, Alexandru T. Balaban was the third most prolific author, and three of his articles were included in the 10 most cited papers of the period.

He had an unquestionable central position in the citation network of the topic, as seen in Figure 1 (Fig. 5 [6]).

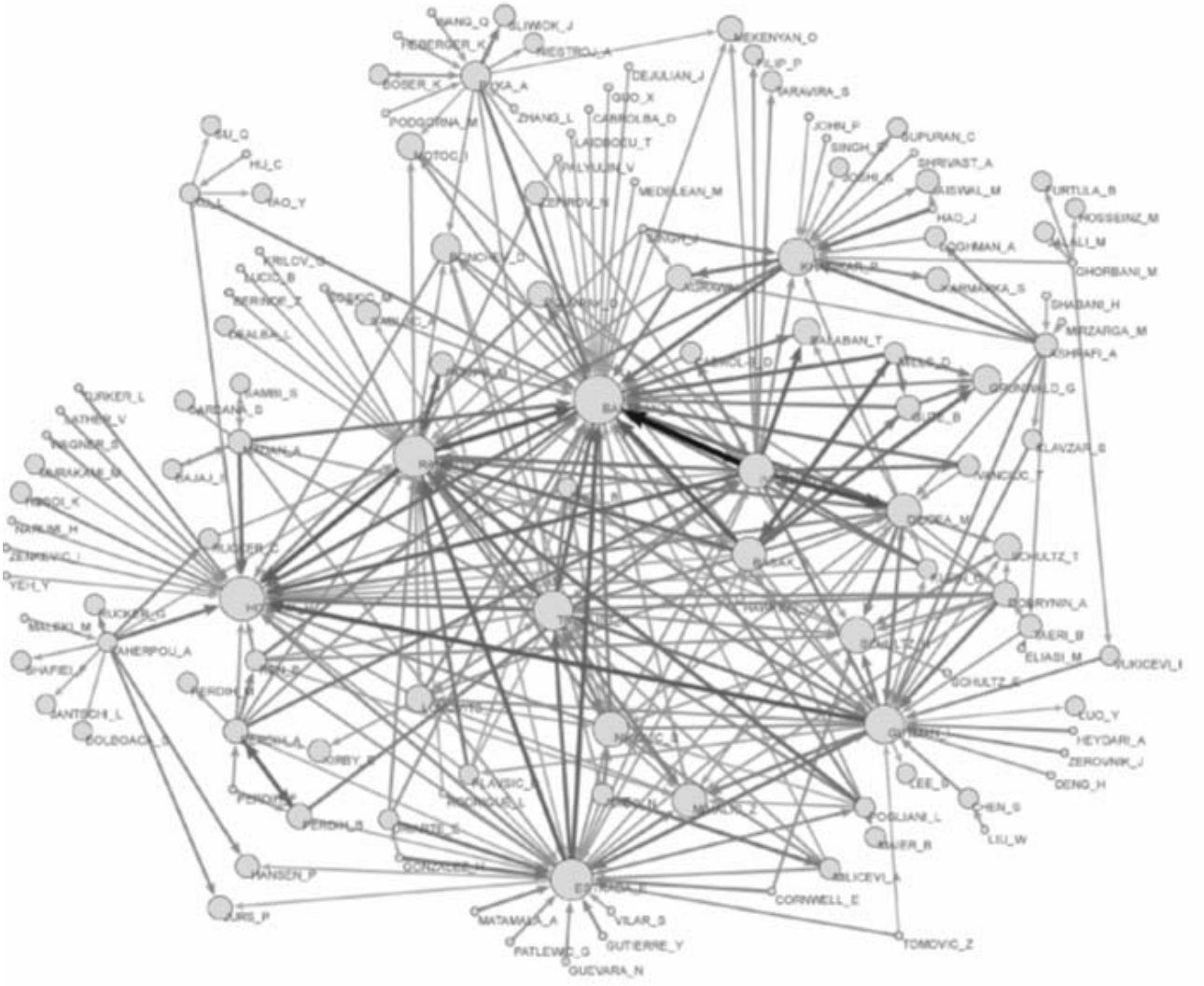


Fig. 1 – The citation network of the most prominent authors in topological index research. The size of each node is proportional to the number of incoming citations. Alexandru T. Balaban is the largest node in the center of the graph.[6]

In the same study, his position in the co-authorship network was found to be far less central, as seen in Figure 2 (Fig. 9 [6]). This may partly be due to the relatively large share (about 25%) of single-authored papers among his contributions to this topic.

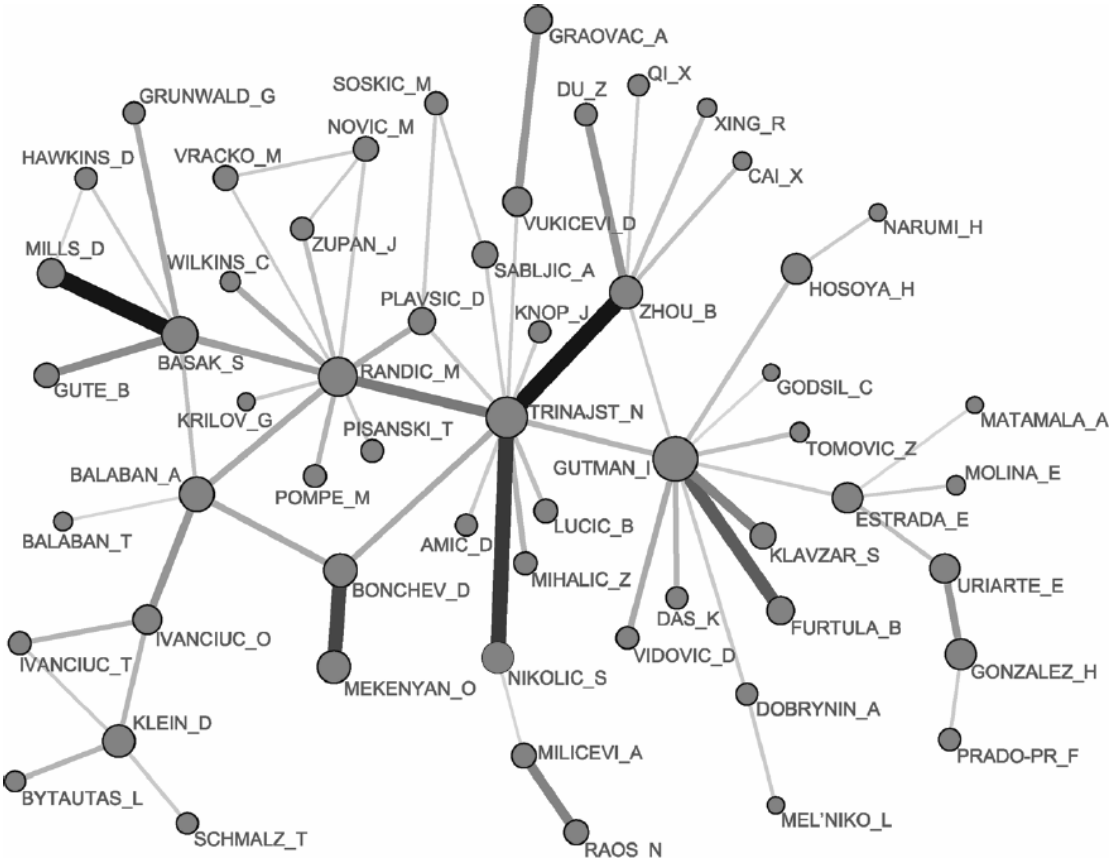


Fig. 2 – The main component of the co-authorship network of topological index research. Weight and shade of links is proportional to the extent of collaboration between authors in question. Size of nodes is proportional to the authors' overall number of articles.[6]

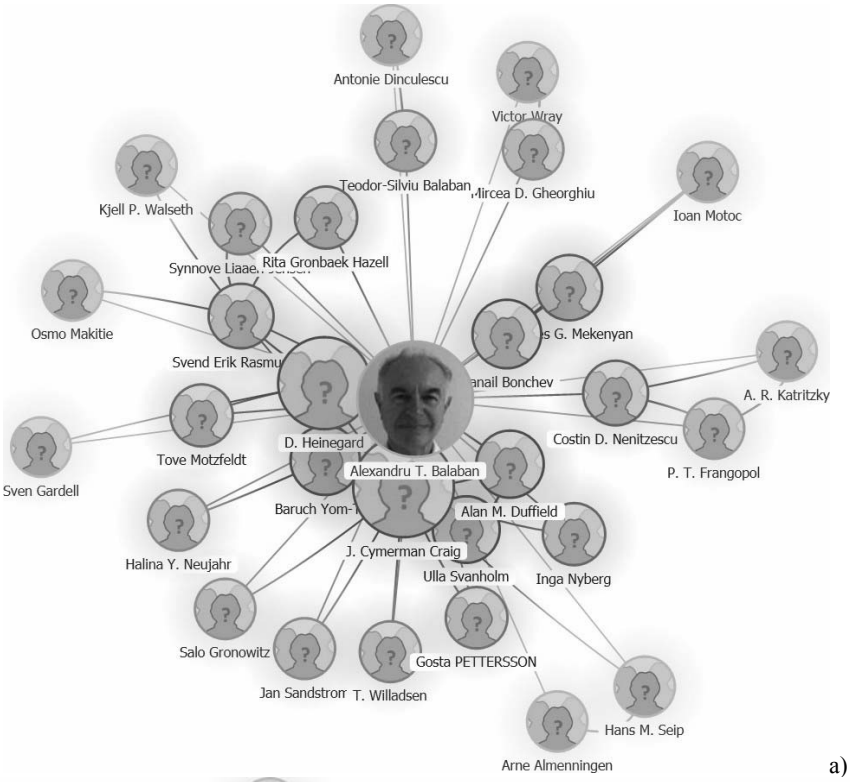
THE CO-AUTHORSHIP NETWORK OF PROFESSOR BALABAN

Google scholar keeps count a total of 897 publications of Alexandru T. Balaban. The Web of Science records 567 papers for the period 1975-2016. 89 of the 567 papers were single authored. A total of 443 co-authors from 34 countries contributed to Professor Balaban's bibliography. The average number of authors per paper was 3.6. Apparently, his co-author network is rather extensive, and his collaboration propensity is high. His Partnership Ability Index (PAB)[7] is 17, *i.e.*, there are 17 co-authors with at least 17 joint publications with him. This is a fairly high value: among the Hevesy-medalists (awarded annually honouring outstanding achievements in radioanalytical and nuclear chemistry by the *Journal of Radioanalytical and Nuclear Chemistry*) studied in,[7] there was only one scientist with higher PAB,[23] and the maximum value among a million computer scientists[8] was 18. As explained in,[6] "low values indicate either a scanty or an inconstant set of co-authors ('poor embedding'), high values suggest a wide and persistent co-authors network ('strong embedding')". Professor Balaban seems thus very strongly embedded in his peer researcher community.

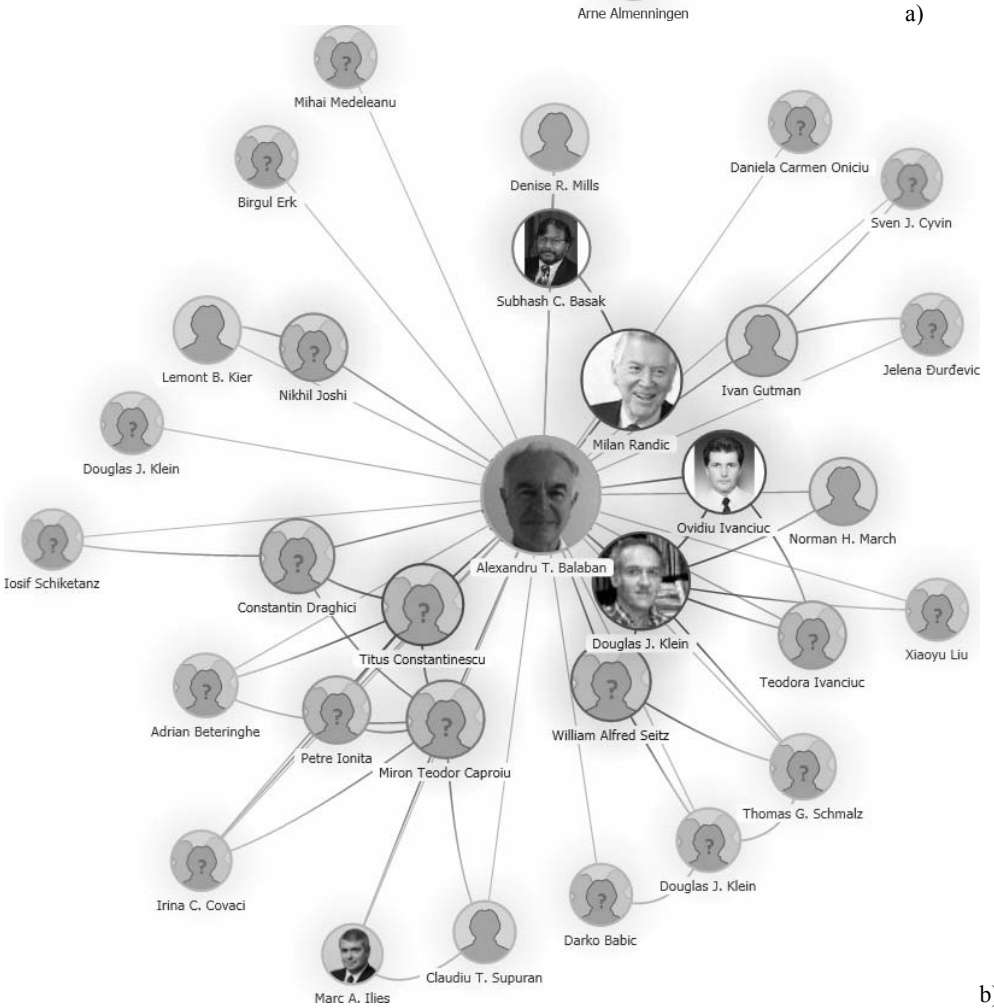
At the Microsoft Academic Search – Visual Explorer website[9] co-authorship networks of researchers can be generated and visualized. Unfortunately, the website has apparently not been updated since 2013, but all data and services are still available. What is even worse, the reliability of the available data is questionable. Alexandru T. Balaban, for instance, has a multiple identity: his Bucharest and Galveston contributions are stored and processed separately. Both networks are presented in Figure 3.

To see a more complete picture, a co-authorship network analysis was made on a well-defined and controlled data set.

Bibliographic data of all papers of Alexandru T. Balaban covered by the Web of Science database in the period 1996-2016 were retrieved. The 271 papers contained 1067 author names (including 271 occurences of the various variants of Professor Balaban's name). The names of all co-authors were manually cleaned; 230 co-authors were identified and a complete co-author graph was constructed. Its structure is presented in Figure 4.



a)



b)

Fig. 3 – The co-author network of Alexandru T. Balaban, Bucharest (a) and Alexandru T. Balaban, Galveston (b) according to the Microsoft Academic Search – Visual Explorer.[9]

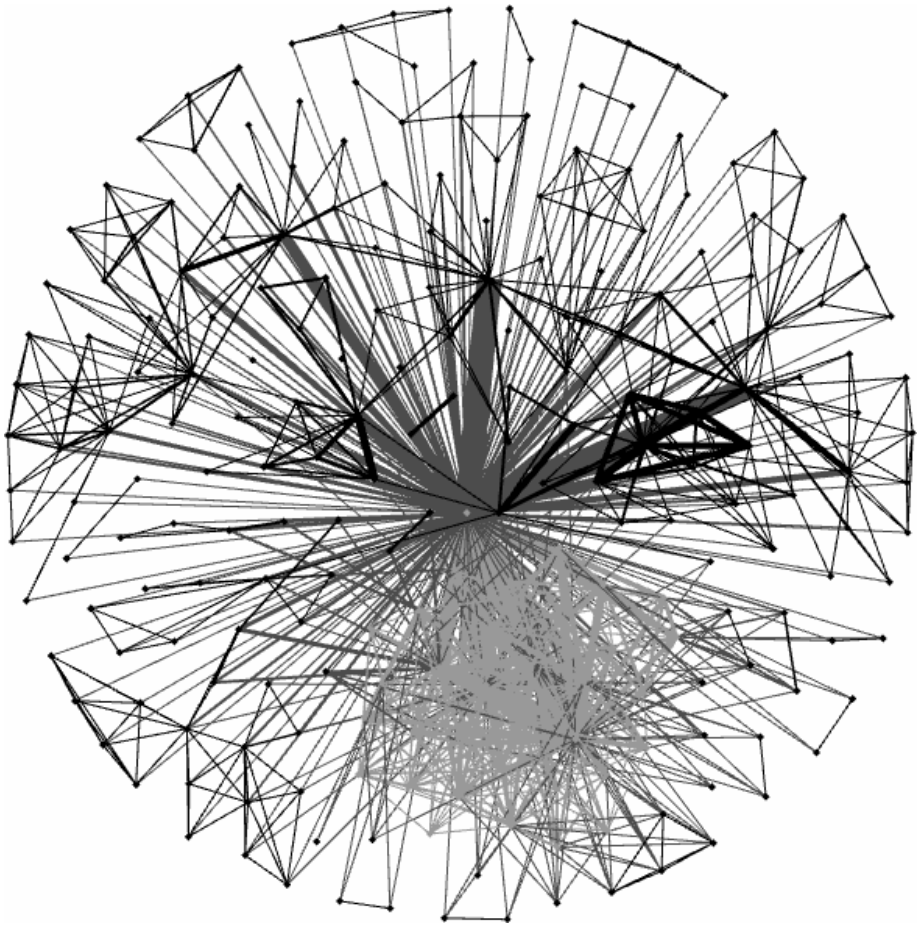


Fig. 4 – The structure of the complete co-author network of Alexandru T. Balaban
(Data: Web of Science, 1996-2016; Visualization: Gephi Ver. 0.9.0, Fruchterman-Reingold layout[10]).

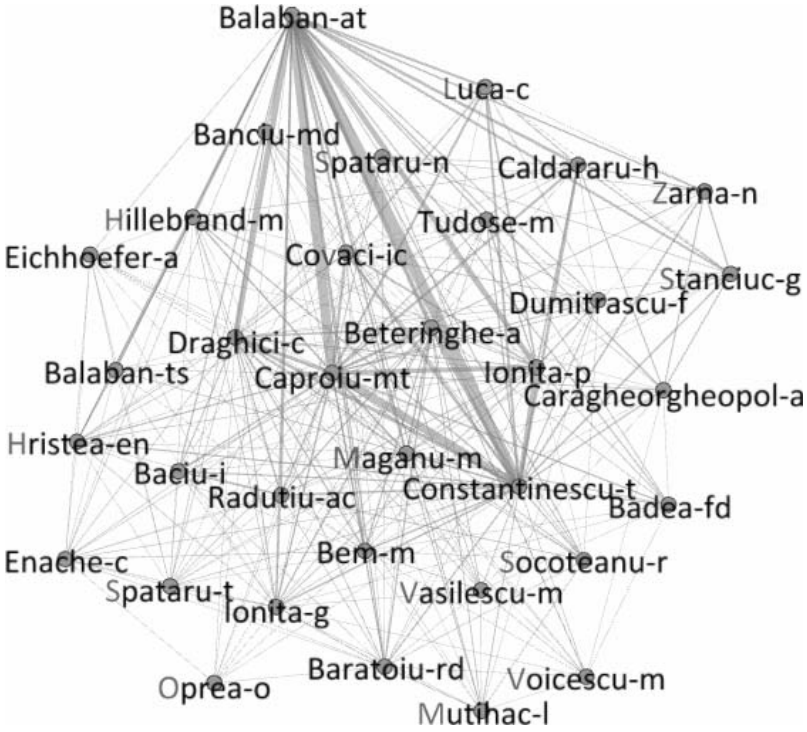


Fig. 5 – The structure of the core co-author subnetwork of Alexandru T. Balaban
(Data and visualization as in Figure 4).

A core subgraph was separated from the complete graph containing only researchers co-authoring at least 10 papers with Professor Balaban. In Figure 4, the nodes of this subgraph are colored in grey. The enlarged picture of this core subgraph completed with the researchers' name as node labels is given in Figure 5.

This subgraph has, indeed, some resemblance to molecular structures; its multiple centers may suggest comparison to multicenter coordination compounds (see, *e.g.*, Figure 6[11]).

Obviously, this comparison is not a strict analogy, just a – if you like, poetic – metaphor. Nevertheless, even such loose analogies may serve well creative thinking. Although the Wiener-index found its way into the glossary of a widely circulated Social Network Analysis textbook,[12] its actual use outside mathematical chemistry has apparently not yet been attempted. In a recent paper,[13] topological features of the co-author network was found to be the most effective in predicting the formation of future co-authorship links. It thus does not seem unreasonable to think that topological indices, Wiener-index, Hosyama-index or – why not? – Balaban's J-index may someday compete in usefulness with Hirsch's h-index in scientometrics.

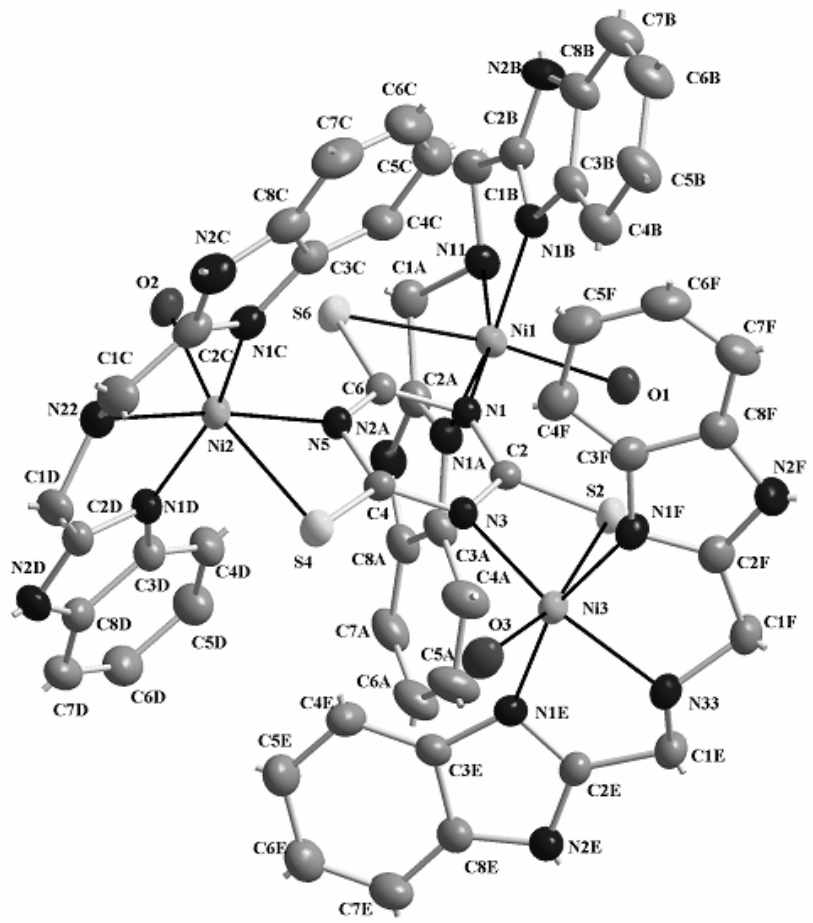


Fig. 6 – The molecular structure of [Ni3(N,N,N',N'',N''-pentamethyldiethylenetriamine)3(μ-trithiocyanurate)]$(ClO_4)_3$.[11]

LONG RANGE INTERACTIONS

Co-author networks considered so far in this paper had rather restricted range, since all researchers included in them are either co-authors or connected through Professor Balaban. Co-author networks of much larger range can, of course, also constructed. For instance, the infamous Erdős Number Project[14] keeps record of mathematicians in the co-author network of Paul Erdős up to a range of 13.

The Microsoft Academic Search – Visual Explorer website[9] makes possible (within the limited reliability mentioned above) to explore co-author paths between researchers at a longer co-authorship distance. Not surprisingly, the default option is the calculation of the Erdős-number.

Alexandru T. Balaban (whether with his Bucharest or Galveston identity) has an Erdős-number of 4. A possible path is shown in Figure 7.

Fig. 7 – A possible co-author path between Alexandru T. Balaban and Paul Erdős
(Source: Microsoft Academic Search – Visual Explorer[9]).

Whether a "Balaban-number" (shortest co-author path to Professor Balaban) may serve as a measure of eminence in chemistry as well as the Erdős-number in mathematics is still to be tested. In a first exploratory test, we could establish that among the nine Nobel Prize winners of the last three years (2013-2015), four (Thomas A. Lindahl, Martin Karplus, Michael Levitt and Arieh Warshel) have a Balaban-number of 3, five (Paul Modrich, Aziz Sancar, Eric Betzig, Stefan W. Hell and William E. Moerner) have a value of 4. One of the authors of the present paper (T. B.) may be, therefore, justly proud of his Balaban-number of 4, helping the other two co-authors to reach a value of 5 (see Figure 8).

Fig. 8 – Co-author path between the authors of the present paper and Alexandru T. Balaban
(Source: Microsoft Academic Search – Visual Explorer[9]).

CLOSING REMARKS

Scientometrics is a "self-reflective" science: it is able to study itself with its own tools.[15] In this sense it is similar to sociology,[16,17] psychology[18] or economics,[19] and different from, say, physics or radiochemistry. This paper suggests that molecular topology may turn to be self-reflective: topological indices may give relevant reflections on the community of topological index research. If this proves to be true, no doubt, this mirror (mirror on the wall) will also show Alexandru T. Balaban as one of the fairest of them all.

REFERENCES

1. A.L. Barabási, H. Jeong, Z. Néda, A. Schubert and T. Vicsek, *Physica A*, **2002**, *311*, 590–614.
2. K. W. Boyack, R. Klavans and K. Börner, *Scientometrics*, **2005**, *64,* 351–374.
3. N. J. Van Eck and L. Waltman, *Scientometrics*, **2010**, *84*, 523-538.
4. Y. Erjia and D. Ying, *Information Processing & Management*, **2011**, *47*, 125–134.
5. A. Schubert, "Atoms and molecules of scientometrics", Authors, publications, references, citations and the bonds among them. Toulouse, November 7, 2013.
6. J. Bodlaj and V. Batagelj, *Molecular Informatics*, **2014**, *33*, 514–535.
7. A. Schubert, *Scientometrics*, **2012**, *91*, 303–308.
8. G. Cabanac, *Scientometrics*, **2013**, *96*, 1–9.
9. http://academic.research.microsoft.com/VisualExplorer
10. https://gephi.org

11. P. Kopel, D. Wawrzak, V. Langer, K. Cihalova, D. Chudobova, R. Vesely, V. Adam and R. Kizek, Molecules, **2015**, *20*, 10360-10376.
12. D. L. Passmore, "Social Network Analysis – Theory and Applications", Wiki Book, 2011.
13. Y. Sun, R. Barber, M. Gupta, C. C. Aggarwal and J. Han, "Co-author Relationship Prediction in Heterogeneous Bibliographic Networks", *Proceedings of the 2011 International Conference on Advances in Social Networks Analysis and Mining*, p. 121–128.
14. http://wwwp.oakland.edu/enp/
15. A. Schubert, "Scientometrics as a quantitative self-reflection of the social sciences", EARLI, 12th Biennial Conference Budapest, 2007, http://earli2007.hu/nq/home/invited_sessions/earli_invited_expert_panel_discussion/scientometrics_as_the_quantitative_selfreflection_of_the_social_sciences/
16. R. W. Friedrichs, "A Sociology of Sociology", NY: Free Press, 1970.
17. E. Morin, *Cahiers Internationaux de Sociologie*, **1985**, *78*, 39–58.
18. G. Richards, *Theory & Psychology*, **2002**, *12*, 7–36.
19. T. Eggertsson, *Kyklos*, **1995**, *48*, 201–210.
20. T. Coupe, *Kyklos*, **2004**, *57*, 197–215.