Modelling of Scientific Collaboration based on Graphical Analysis

Veslava Osinska[1], Grzegorz Osinski[2] and Wojciech Tomaszewski[2]

[1] *wieo@umk.pl*
Nicolaus Copernicus University, Institute of Information Science and Book Studies
ul. Bojarskiego 1, 87-100 Torun (Poland)

[2] *grzegorz.osinski@wsksim.edu.pl; wojciech.tomaszewski@wsksim.edu.pl*
Institute of Computer Science, College of Social and Media Culture
sw. Jozefa 23/35, 87-100 Torun (Poland)

Introduction

An analysis of the interrelationships between elements within dynamic structure typically involves perturbation methods based on the minimum energy. In result, the researchers use minimum distance-based algorithms and therefore the shortest path between the various components of the system. However, the history of science development shows that collaboration between the researchers in different disciplines becomes effective and fruitful when scientific explorations do not follow the "shortest possible" roads.

In current work authors present a novel approach, how to analyse and evaluate the possible collaborations ways in a small team of researchers (number of nodes is less than 100) participating in the project network KnowEscape COST Action.[1]

Data, metrics and assumption

Analysed dataset consists of 83 records characterized each member of COST network. Input data organized in 83x83 matrix, describe two years collaboration within such activities as: mobility, events organization, publishing (also for former years) and project management. The dataset was gathered using KnowEscape website (knowescape.org), ResearchGate and Mendeley services.

To describe the mutual relationships between members the graph based on Mycielski concept was constructed (Larsen, Propp & Ullman, 1995). The authors identified graphically four attractors of maximum energy. The clique represents each researcher's pair, and arbitrarily large chromatic number means any combination of disciplines. Presented visualisation (Fig. 1) was generated by using the Poincare section (PS) of the 3D space which is defined by all ties between team's members (Tamassia, 2000).

The main problem concerns identification subgroups categories with regard to scientific activity. The matrix was generated using selected nodes and links through Poincare projection (Clifford, Azuaje, & McSharry, 2006).

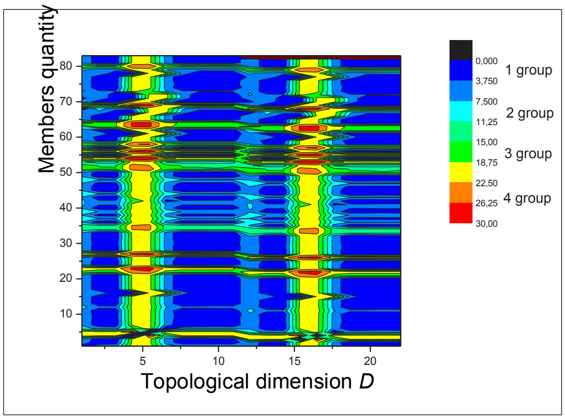

Figure 1. An iterated visualization of discrete distance routes.

Obtained iterated visualization of discrete distance routes is shown on Figure 1. As a final result we observe four clear clusters. All participants were divided on four groups by describing appropriate roles in social network: leaders, connectors, performers and outliers.

This approach was tested using algorithms adopted from medical data analysis for time series (Swierkocka-Miastkowska & Osinski, 2007, Mazur, Osinski, Swierkocka, 2009).

The authors evaluate also the dynamics of total activity by using fractal dimension (FD) of each PS image. FD is the measure of nonclassical geometry shapes and can be used as a pattern's complexity parameter (Osinska 2012).

Fractal dimension was obtained by Higuchi algorithm, so the resulting maps help to discover possible opportunities for further development of cooperation between the scientists.

Visual results

All members' activities represented by matrixes are summarized and full collaboration is weighted by appropriate real numbers. Popular application *Gephi* allows finding collaboration groups and revealing the scientists with basic roles: leader,

[1] This research is sponsored by National Science Center (NCN) under grant 2013/11/B/HS2/03048/ Information Visualization methods in digital knowledge structure and dynamics study.

subgroup leader, connector, outsider and so on. By using force directed layout (*force atlas 2*) the authors have obtained clarify configuration presented on Figure 2. As expected, the central point is occupied by the real team's leader. The closer node to central one represents the scientist who is more active in collaboration with the team's leader.

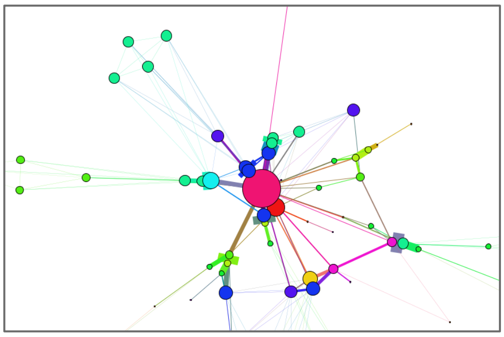

Figure 2. The graph of full activity of team's members.

Network visualisation exposes also some subgroups where intrinsic collaboration (mainly in publishing) is significant. The scientists within these groups share a common feature: geographic localisation. They work in the same country.

Simple quantitative proportional correlations between identified groups on a graph are compatible with the ones visualised on Figure 1.

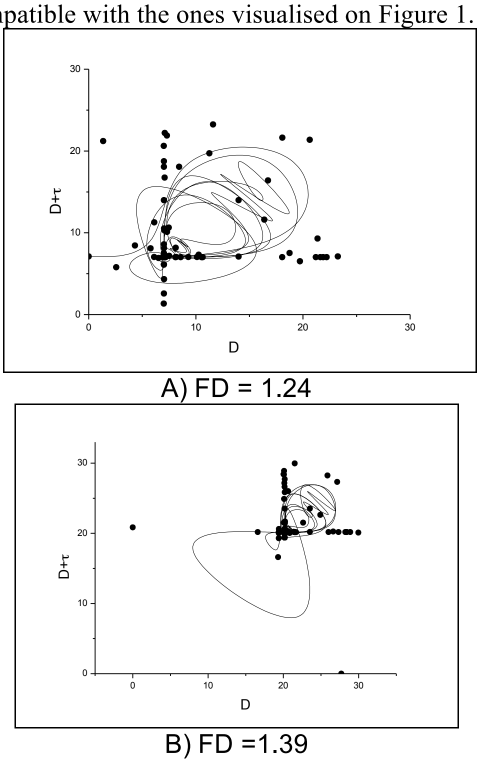

Figure 3. Two variations of collaboration between scientists with different social roles: A) Leader-performer; B) performer-performer.

Next step, calculation of fractal dimension, was accomplished for combinations of representatives of different groups, for example: leader-performer, subleader-leader, connector-performer and so on.

Two variations of collaboration with appropriate FD are shown on Figure 3. Fractal dimension is always lower for every pairs composed from the leader or subleader compared to the performers and connectors.

Conclusions

The authors propose new parameters for the prediction of a stable way of scientific collaboration. First is the shape of Poincare section (Return Map Poincare). For inhomogeneous academic groups where there is no self-consistency (like in this work), the level of nonlinearity can also reflect collaboration potential. It is proportional to the quantity of curves on Figure 3. The second indicator – FD shows the possibility to cooperate as well as its dynamics.

Higher fractal dimension in the case of performers can be explained by larger dynamics of predictive collaboration. This indicates the pattern is more complex. It means the pair covers significant collaboration potential.

Visualisation can help discover possible opportunities for further development of scientific cooperation. Therefore, we can observe common career landscapes of the various members and groups.

References

Clifford, G.D., Azuaje, F. & McSharry, P. (2006). *Advanced Methods and Tools for ECG Data Analysis.* Artech House Publishers.

Larsen, M., Propp, J., & Ullman, D. (1995). The Fractional Chromatic Number of Mycielski's Graphs. *Journal of Graph Theory*, *19*, 411-416.

Mazur, R., Osinski, G., Swierkocka, M. & Mikolaiczik, G. (2009). Evaluation of the dynamics of energetic changes in the brain stem respiratory centre in the course of increasing disorders of consciousness. *Activitas Nervosa Superior. 51*(5112), 69-72.

Osinska, V. (2012). Fractal Analysis of Knowledge Organization in Digital Library. In A. Katsirikou & Ch. Skiadas (Eds.) *New Trends in Qualitative and Quantitative Methods in Libraries* (pp. 17-23). World Scientific Publishing Company.

Swierkocka-Miastkowska, M. & Osinski, G. (2007). Nonlinear analysis of dynamic changes in brain spirography. Results in patients with ischemic stroke. *Clinical Neurophysiology*, *118*(12), 2822.

Tamassia, R. (2000*)*. Graph Drawing. Ch. 21 In J.-R. Sack & J. Urrutia (Eds.) *Handbook of Computational Geometry* (pp. 937-971). Amsterdam, Netherlands: North-Holland.