World Applied Sciences Journal 27 (12): 1625-1629, 2013
ISSN 1818-4952

DOI: 10.5829/idosi.wasj.2013.27.12.13728

Positional Analysis and Mapping of Scientific Networks

Alexey Tselykh and Gennady Veselov

Southern Federal University, Taganrog, 347922, Russia

Abstract: In this paper, we consider prospective research directions in the field of mapping scientific and technical activities of the university. We approach the problem of role discovery in a research group using methods of positional analysis in the co-authorship network based on equivalence relation. We carry out positional analysis in scientific networks based on the concept of a fuzzy (weak) cutpoint and provide an algorithm for enumerating all the fuzzy cutpoints in a fuzzy hypergraph.

Key words: Science metrics . network analysis

INTRODUCTION

With the development of information technologies and the Internet, the key assessment tools to measure scientific and technical activities of the university are scientometric analysis and science mapping [1]. These are state-of-the-art methods and techniques for the analysis and visualization of structured scientific information arrays, e.g. international citation indices. The best-known approach is based on co-citation methods, or methods of prospective links between publications, that allow for the identification of new emerging scientific areas and advanced research frontiers. These methods were simultaneously and independently developed in USSR by I.V. Marshakova [2] and in USA by G. Small [3] in 1973.

Another rapidly growing research area are socio-semantic knowledge networks [4]. They allow formalizing and solving the problems of data analysis and knowledge management in science and technology based on the collective intelligence of social networks and Semantic Web technologies. Such knowledge networks are based on hybrid organization of metadata that combines different classification approaches, including well-established taxonomies and ontologies as well as weakly structured tag-based folksonomies. Socio-semantic knowledge networks are formally described using graph and hypergraph models [5, 6]. Social network analysis methods are widely applied then. In this regard, we can refer to VIVO international research and education network, Sci2 Tool and other projects supervised by K. Börner [1] which are the main sources for scientific infographics today.

The market of scientometric analysis and science mapping software is currently led by SciVal Spotlight and InCites services provided respectively by Web of Science and Thompson Reuters. However, their possibilities for the assessment of scientific and technical activities in Russian universities are limited.

Firstly, these tools deal exclusively with English language citation indices. They do not adequately reflect the results of Russian scientists whose share in global science and technology (according to Web of Science) is less than 2 percent.

The functionality of Russian Scientific Citation Index (RSCI) in this area is limited to basic indicators and statistical reports on co-authored works, referenced and citing publications. RSCI does not explicitly provide co-citation data. Our bibliographic review has shown that there are no informative studies on the analysis of co-citation and co-authorship networks in RSCI.

Secondly, there is an open problem of visualizing hierarchical cognitive structure of science with the support for drill-down, i.e., "research area"-"research frontier"-"cluster of publications"-"scientific unit". Hereafter, a scientific unit is either a university, institute, research center, laboratory, research team or an individual researcher. The interconnection between basic and applied research based on the network and temporal analysis has not been researched yet.

Thirdly, in terms of small research groups the scientometric analysis and science mapping technologies mainly address the problem of revealing "invisible colleges", i.e., the networks of personal contacts between scientists without clearly defined borders. Such networks emerge by virtue of communication between researchers working on the same or similar problems in different organizational structures. Thus, geographically distributed "invisible

Corresponding Author: Tselykh, Southern Federal University, 44 Nekrasovsky Street, Taganrog, 347922, Russia

colleges" are a broader concept than a scientific school or a research team that constitute the core of the "college".

Overall, we can make a conclusion on the impossibility to apply existing tools for complex assessment of scientific activity at the level of a university, faculty, or a research team.

Science of team science: Social activity of research teams and their leaders is the key focus of research aimed at increasing the overall efficiency of scientific research activities.

Studies on research teams are conducted within the research field of "the science of team science" [7]. The research methodology includes methods of scientometric and bibliometric analysis, methods of social network analysis [8], techniques for mapping and visualization as well as surveys, ethnographic studies, case-based analysis and interviews.

At the beginning of XX century 82% of scientific publications belonged to a single author. By the 60s, 40% of works had two authors and 17% had three or more authors [4]. Today, researchers working on their own are just a rare exception to the rule. Most of the research projects are interdisciplinary while monodisciplinary projects assume separation of intensional and organizational functions within the project. Since the 1990s, the amount of investments in interdisciplinary research carried out by distributed international research teams had significantly increased due to the globalization processes. The world's top universities are on a "bounty hunt" for the leading scientists who have an influence on the process of communication in a research team, promote its intensification and, to a certain degree, productivity.

In the process of carrying out research, there arises an objective need for separation of intensional and organizational responsibilities in a small research group. Each of its members either solves his own scientific problem or performs a scientific role-a specific set of activities this employee fulfils better than the rest of the group members.

Each researcher occupies a position that is characterized by a degree of his (her) involvement in the process of scientific communication. The highest scientific status corresponds to the position of a leader and the group member occupying this position is correspondingly the team leader. This leader is usually an employee acknowledged by the majority of other group members as the author of the research program. However, the supervisor does not have to be a team leader. He might also play the role of a research manager. In this role, his objective is to identify employees who already have a research program or are capable to generate it and to organize and provide implementation of a program. With that, a manager and an author of the research program should cooperate and clearly divide the responsibilities for the implementation of the program.

Positional analysis in scientific networks based on the relation of equivalency: It is possible to conduct in-depth analysis of research team structure using methods of positional analysis in co-author networks based on the equivalency relation.

We can define a position in a scientific network as a set of structurally indistinguishable actors (individual researchers, research teams, etc.) that have similar relations and the same patterns of ties with other actors in the network. Hence, we can define a role [9] as a type of relations between actors and/or positions.

An important property of equivalency is the possibility to divide set into non-intersecting classes of equivalent elements and later on during the set transformations to consider any single element of the class instead of that class.

Hereinafter, we refer to equivalency classes as positions and denote them as B_k. An affiliation of actor i with position B_k can be written in the form $\varphi(i) = B_k$. If actors i and j are equivalent, $i \equiv j$, they belong to the same position B_k.

Actors are structurally equivalent [10] if they have the same relations with all the other actors in the network.

Suppose there R relations $\chi_1, \chi_2, \ldots, \chi_R$. The tie between actors i and j on relation χ_R is denoted by $i \xrightarrow{\chi_r} j$. Then, actors i and j are structurally equivalent, $i \equiv j$, if $(i \xrightarrow{\chi_r} k) \leftrightarrow (j \xrightarrow{\chi_r} k)$ and $(k \xrightarrow{\chi_r} i) \leftrightarrow (k \xrightarrow{\chi_r} j)$ for every actor $k = 1,2,\ldots,g$ $(k \neq i,j)$ and every relation $r = 1,2,\ldots,R$. It is easy to verify that intersection and transitive union closure of two structural equivalencies is a structural equivalency.

A trivial example of structural equivalency is an identity partition φ_Δ where each actor occupies a separate position and a complete partition φ_Π where all the actors belong to the same equivalency class and occupy the same position. Bipartition of a complete bipartite graph is another example of structural equivalency.

It should be mentioned that the distance between two structurally equivalent not isolated vertices is not greater than 2, because if i is adjacent to k, then k is adjacent to j. It follows that structural equivalency allows to determine equivalency only for adjacent vertices. Thus, structurally equivalent vertices classes are either independent sets or cliques.

The search for stable research teams based on bibliometric analysis can be formalized as a problem of enumerating all the cliques (maximal complete subgraphs) in a co-authorship graph. Cliques will correspond to groups of researchers who had (at least one) joint works.

It should be mentioned that a condition of having identical connections with the same actors in the network is a significant limitation. In our example, a researcher occupies a position of a team leader in a group. Leaders of different research groups will not be structurally equivalent.

Thus, structural equivalency does not allow to compare positions and roles across populations and to formalize the definition of a position. We should mention, though, that it is possible to measure structural equivalency, e.g. using Euclidian distance between the pairs of actors in a multidimensional space of dyadic network ties, instead of measuring presence or absence of the structural equivalency.

Suppose $\varphi = (X,F)$ is a functional and total relation and a certain operation ? is defined on the set X. If $\varphi(A_1, A_2) = \varphi(A_1), \varphi(A_2)$ is true for any two subsets A_1, $A_2 \subseteq X$ and the relation is injective and surjective at the same time, i.e., it is bijective, then relation φ is an isomorphism, or a bijective homomorphism of the set X on itself, i.e. an automorphism.

We can define automorphic equivalency in a network using the definition of automorphism.

Actors i and j are automorphically equivalent, $i \equiv j$, if there is an automorphism φ, $\varphi(i) = j$. Hence, $(i \xrightarrow{\chi_r} k) \rightarrow (j \xrightarrow{\chi_r} \varphi(k))$ and $(k \xrightarrow{\chi_r} i) \rightarrow (\varphi(k) \xrightarrow{\chi_r} j)$ for $\forall k, \chi_r$.

Automorphically equivalent actors have the same graph-theoretic characteristics, i.e., indegree and outdegree, centrality coefficient, etc.

It is easy to verify that structurally equivalent vertices are automorphically equivalent. The opposite, generally speaking, is not true.

To analyze equivalency of actors from different groups, we can define isomorphic equivalency in a network using the definition of isomorphism.

Suppose there are relations $\varphi(X,F)$ and $\psi = (Y,P)$ and mapping *f*: $X \rightarrow Y$. If for any two elements $x_1, x_2 \in X$, such that $x_1 \varphi x_2$, $x_1 \varphi x_2 \rightarrow f(x_1) \psi f(x_2)$ is true where

$$f(x_1) = y_1, f(x_2) = y_2, y_1, y_2 \in Y$$

and *f* is a bijection, then mapping *f* is a bijective homomorphism, or an isomorphism between the relations φ and ψ.

Actors $i \in X$ and $j \in Y$ are isomorphically equivalent, $i \equiv j$, if there is an isomorphism *f*, *f*(i) = j.

Positional analysis in scientific networks based on the concept of a fuzzy cutpoint: Using notions of a fuzzy (strong) cutpoint and a fuzzy (strong) bridge given by Mordeson [11] we define a notion of a fuzzy weak cutpoint and a fuzzy weak bridge in a fuzzy graph. The main difference between these two concepts is in the usage of degree of reachability for fuzzy weak cutpoints and weak bridges instead of degree of mutual reachability for (strong) cutpoits and bridges.

Definition 1: Node b is regarded as fuzzy reachable from node a in a fuzzy graph $\tilde{H} = (X,U)$ if there is a fuzzy graph walk $\tilde{M}(a,b)$ from node a to node b.

The degree of reachability of node b from node a can be defined as

$$\gamma(a,b) = \max_{\alpha}(\mu(\tilde{M}_{\alpha}(a,b)), \alpha = 1,2,...,n$$

where n is a number of different directed paths from a to b.

Definition 2: We define node $x \in X$ as a fuzzy weak cutpoint in a fuzzy graph if its removal and removal of incident arcs leads to the reduction of the degree of reachability $\gamma(y,z)$ between some nodes y and z, $x \neq y \neq z$.

Property 1: Node $x \in X$ is a fuzzy weak cutpoint if and only if there are such nodes y and z, other than x, that x belongs to each fuzzy walk $\tilde{M}(y,z)$ with the maximal conjunctive strength.

Property 2: All fuzzy walks connecting nodes from different components of connectivity always pass a fuzzy cutpoint.

Definition 3: Arc $\tilde{u} \in U$ is called a fuzzy weak bridge if its removal reduces the degree of reachability $\gamma(y,z)$ between some nodes y and z that are not incident to that arc.

Property 3: Arc $\tilde{u} \in U$ is a fuzzy weak bridge if and only if there are nodes y ? z that are not incident to that arc, such that $\tilde{u} \in U$ belongs to each fuzzy walk $\tilde{M}(y,z)$ with a maximal conjunctive strength.

Property 4: Terminal node of a weak bridge is a weak cutpoint if there are other arcs in a graph that are incident to that node.

Theorem 1: Suppose there is a fuzzy graph $\tilde{X}(\tilde{H})$ and it is a cycle. Node x is a fuzzy weak cutpoint if and only if it shares node for two fuzzy weak bridges.

Theorem 2: In case a node is incident to a pair of fuzzy weak bridges, it is a fuzzy weak cutpoint.

Theorem 3: If $<\mu_U(x,y)/<x,y>>$ is a fuzzy weak bridge, the degree of reachability of node y from node x is equal to the degree of incidence $\mu_U(x,y)$.
Proofs of theorems are given in [6].

We now provide an algorithm for enumerating all the fuzzy cutpoints in a fuzzy hypergraph.

Property 5: Node y is a cutpoint from the standpoint of node x if its removal breaks each and every elementary cycle that contains x in hypergraph $\tilde{H}_\alpha$. In other words, if node y belongs to each and every cycle x belongs to.

We can derive this property from Definition 2.

Hence, in order to identify all the fuzzy cutpoints we set a single-valued representation of a fuzzy hypergraph $\tilde{H}_\alpha$ as a fuzzy vertex graph $\tilde{X}(\tilde{H})$. By choosing all the different degrees of incidence in a graph, we can split it into crisp graphs of α-level [12] and find all the elementary cycles. Then, we find cutpoints and assign the degree of fuzziness α to them. Such cutpoints make a set of fuzzy cutpoints in a fuzzy hypergraph.

We can write down the algorithm for enumerating all the fuzzy cutpoints in a fuzzy hypergraph as follows.

1. Specify single-valued representation of a fuzzy hypergraph as a fuzzy vertex graph $\tilde{X}(\tilde{H})$.
2. Choose all the different degrees of incidence $\mu(x_i,x_j)$ present in $\tilde{X}(\tilde{H})$, then grade them and put down a sequence $0 < \alpha_n < ... < \alpha_1 = h(\tilde{X}(\tilde{H}))$.
3. Use the rule of zero semidegree, i.e., if any node in a graph has zero indegree or zero outdegree, it does not belong to any of the cycles, thus it is a cutpoint and can be excluded. That allows to significantly reduce in-depth search in a graph and avoid repeated cycles in all graphs of α-level.
4. Find all elementary cycles in graph $X_{\alpha_n}(\tilde{H})$ using algorithm given in [5].
5. Using Property 5 we find cutpoints in graph $\tilde{X}_\alpha(\tilde{H})$ and assign a degree of fuzziness α to them.
6. Suppose n = n-1. If $n \geq 1$, go to Step 3 of the algorithm. Otherwise, go to Step 7.
7. The end of algorithm.

Based on the set of fuzzy cutpoints $\tilde{A}$, we can construct a fuzzy oriented graph of cutpoints that is a fuzzy oriented graph of a second type [13] both with fuzzy nodes and fuzzy edges. The degree of fuzziness coincides with a maximal degree of fuzziness of a cutpoint and the strength of arcs coincides with the degree of mutual reachability of the nodes.

CONCLUSION

The next prospective area of research in the field of mapping scientific and technical activities of the university are methods and tools for cognitive audit of scientific and technical knowledge. They allow for measuring knowledge gap between the volume of key scientific and technical knowledge necessary to achieve the strategic goal, e.g. entering the world university rankings and existing scientific and technical potential of the university.

Our suggested approach to this problem is to build maps of scientific and technical knowledge (knowledge assets) that include:

- Ontology maps that allow for the comparison of similar research projects conducted by several research teams in order to formalize and integrate scientific and technical knowledge;
- maps of technical competencies that facilitate the search for experts;
- visualizations of social knowledge networks that allow for the analysis of knowledge networks and models of communication between different structural units and individual researchers, international partners and other members of scientific and technical knowledge exchange;
- process-oriented knowledge maps that sketch knowledge of the university in the context of its key business processes.

The results of such mapping can be the basis for creating strategic business plans and roadmaps for the university.

REFERENCES

1. Börner, K., M. Conlon, J. Corson-Rikert and Y. Ding, 2012. VIVO: A Semantic Approach to Scholarly Networking and Discovery. Morgan & Claypool Publishers, pp: 250.
2. Marshakova-Shaikevich, I. Meeting and R. Merton, 2010. Discussion of Co-citation Maps of Science and the Evaluation of Scientific Journals. Sociology of Science and Technology, 1 (4): 118-125.
3. Small, H.G., 1973. Co-citation in the Scientific Literature: A New Measure of the Relationship between Two Documents. JASIST, 24 (4): 265-269.

4. Tselykh, A., 2012. Graph-Hypergraph Model of a Semantic Social Network. Izvestiya Yuzhnogo federalnogo universiteta. Tekhnicheskie nauki, 4 (129): 225-229.
5. Christofides, N., 1976. Graph Theory. An Algorithmic Approach. Academic Press, London, pp: 432.
6. Bershtein, L.S. and A.V. Bozhenyuk, 2008. Fuzzy Graphs and Fuzzy Hypergraphs. In: Dopico, J., de la Calle, J., Sierra, A. (Eds.). Encyclopedia of Artificial Intelligence, Information SCI, Hershey, New York, pp: 704-709.
7. Taramasco, C., J.-P. Cointet and C. Roth, 2010. Academic Team Formation as Evolving Hypergraphs. Scientometrics, 85: 721-740.
8. Wasserman, S. and K. Faust, 1994. Social Network Analysis: Methods and Applications. Cambridge University Press, pp: 825.
9. Brandes, U. and J. Lerner, 2007. Role-equivalent Actors in Networks. In the Proceedings of ICFCA Satellite Workshop on Social Network Analysis and Conceptual Structures: Exploring Opportunities. Date Views 10.11. 2013 www.inf.uni-konstanz.de/algo/publications/bl-rean-07.pdf.
10. Lorrain, F. and H.C. White, 1971. Structural Equivalence of Individuals in Social Networks. Journal of Mathematical Sociology, 1: 49-80.
11. Mordeson, J.N. and P.S. Nair, 2000. Fuzzy Graphs and Fuzzy Hypergraphs, Physica-Verlag, Heidelberg, pp: 248.
12. Durfee, A. and A. Tselykh, 2011. Evaluating Operational Risk Exposure Using Fuzzy Number Approach to Scenario Analysis. In Proceedings of the 7th Conference of the European Society for Fuzzy Logic and Technology. EUSFLAT 2011 and French Days on Fuzzy Logic and Applications, LFA 2011, 1 (1): 1045-1051.
13. Bozhenyuk, A., I. Rozenberg and D. Yastrebinskaya, 2013. Finding of Service Centers in GIS Described by Second Kind Fuzzy Graphs. World Applied Sciences Journal (Special Issue on Techniques and Technologies), DOI: 10.5829/idosi.wasj.2013.22.tt.22144, 22: 82-86.