Defence Life Science Journal, Vol. 01, No. 2, September 2016, pp. 188-191, DOI : 10.14429/dlsj.1.10087

Co-Authorship Networks among DRDO Life Science Scientists

Suresh Babu[!] and Gitanjali Yadav[*]

[!]*School of Human Ecology, Ambedkar University Delhi, Delhi - 110 006, India*
[*]*Genomics and Systems Biology Laboratory, National Institute of Plant Genome Research, New Delhi - 110 067, India*
[*]*E-mail: gy@nipgr.ac.in*

ABSTRACT

The modern scientific ecosphere increasingly involves collaborative research, wherein diverse talent pools fuse to produce a research output, and 'co-authorship' can be treated a quantifiable measure of scientific collaboration. The present initiative of a Life Science Journal catering to the needs of life-science/bio-medical science researchers offers a suitable occasion to investigate the existing social structure of science in the defence life science research establishment in India. This short communication describes a meta-analysis of co-authorship networks of this community and we find very interesting inter-disciplinary connections that highlight the significance of this new journal for research impact in the long term.

Keywords: Co-authorship networks, DRDO, science collaboration, defence life science research

1. INTRODUCTION

Science can be understood as a social phenomenon involving large numbers of scientists regularly performing specific actions that are consciously coordinated into large schemes of knowledge creation and enrichment[1]. Network Analysis can greatly advance our understanding of the manner in which knowledge is created as well as how knowledge flows through society. A 'co-authorship' network thus can be understood as a social network of scientists consisting of relationships between scientific ideas, in such a way that the cognitive structure of communication is manifested as relationships between scientists[2].

Although co-authorship networks can provide a window on patterns of collaboration within science, they have received far less attention than have citation networks in bibliometrics, even though co-authorship networks contain much important information about cooperation patterns among authors as well as the status and locations of authors in the broader scientific community structures[3,4]. Here, we confine our attention to co-authorship networks among scientists in the national and regional defence research laboratories in order to identify and understand collaborative social patterns structure of the life science community within and outside the establishment. We ask simple questions like, who tends to publish together, or what institutions are most collaborative? We use local as well as global network properties to explain the topology of scientific networks, the most well documented local property being node degree centrality, defined as the number of connections linking a given node to other units in the network, and its corresponding global description being the degree distribution, often known to have a long tail in real world networks[5].

We combine quantitative and qualitative techniques, in terms of graph-theoretic and ethnographic patterns respectively, since these two mutually complement inform each other. We identify and classify scientific collaborations based on structural patterns in observed co-authorship networks in the specialised subfield of life-sciences, by using the available scientific literature (PUBMED April 2016), followed by use of information-theoretic clustering algorithms to extract the modular structure of the network, and overlapping partitions enabling a comprehensive investigation of network organization.

In summary, this is the first meta-analysis on the expanding scientific literature across defence institutions and we find very interesting inter-disciplinary connections that successfully highlight the significance of this new journal for the establishment in the long term.

2. DESIGN AND METHODOLOGY

The existing literature on co-authorship networks in life science community of the Indian defence establishment was identified, evaluated and manually curetted before network construction and interpretation. Data was extracted from PUBMED (in April 2016), a global search engine accessing primarily the MEDLINE database of references and abstracts on life sciences and biomedical topics (https://www.ncbi.nlm.nih.gov/pubmed/). The input dataset for this study was obtained through the keyword DRDO and including all publications that resulted as hits. Data was downloaded in XML format and incorporated into Cytoscape[6] with default parameters for specifying maximum number of authors per publication (Setting the limit to a high value increases the

Received : 03 January 2016, Revised : 29 March 2016
Accepted : 17 April 2016, Online published : 07 October 2016

computational demands of network analysis and is therefore not recommended). For time interval, we suppressed the default limit of five years in order to include all articles in the field published till date. Cytoscape creates a visual summary of how individual authors collaborate by building a co-publication network from the input data where nodes represent authors, edges represent co-authorship and how frequently co-authors collaborate is indicated by the thickness of an edge. Analysis of network topology was done using NEXCADE[7]. Clustering and module identification was performed using Moduland[8]. A module is a set of nodes that have a large influence on each other. In the first step, the influence zone of each edge in the co-authorship network is measured, followed by summing up of all influence zones to determine a community centrality value for the given edge. In the final step, overlapping modules are detected and nodes of the original network are assigned to each module, and a meta-network of modules is created at the next hierarchical level. One can iterate this process and as we go higher and higher in hierarchical levels, the number of meta-nodes becomes smaller and smaller until the whole network will finally coalesce into a single meta-node. The original co-authorship network and modular meta-network along with distinct modules are available from authors upon request.

3. RESULTS

3.1 The Cognitive Structure of Scientific Collaborations

Figure 1 depicts the complete co-authorship network of national and international scientific research collaborations among scientists of the DRDO laboratories in the life science category. It may be visualized as a social interaction structure whose vertices represent authors, and two authors are connected by a tie if they co-authored one or more publications. These ties are necessarily symmetric, or bidirectional. In all the total available data comprise 3004 collaborations between 741 individuals spanning over the past 16 years. Table 1 provides a summary of the basic statistical properties of the network. As can be seen from Fig. 1, the DRDO life-science network is highly interconnected, with few disconnected clusters of small sized networks at the periphery. These smaller networks signify very recent papers published by new scientists, who have not yet initiated collaborations with the larger main component of the network, and are thus either (a) single articles with one main author and his/her students, or (b) articles by new defence scientists co-authored with their national/international collaborators outside the DRDO. Thus, we confine our analysis to the single largest connected component of this network with statistics shown in the last column of Table 1.

This network is also represented by a number of sub-communities as indicated by the node groupings, and these sub communities correspond roughly to topics of research or specific institutions. The highest representation in the network is by INMAS (Delhi), as shown in Fig. 1 inset, followed by DRDE (Gwalior), DIPAS and DRL (Tezpur). One immediate pattern emergent from the network is that collaborations are more frequent within the various defence laboratories, rather than with external organisations. This may arise from the mandate of most defence research programs to be institution centric and fully implementable within the community. However, external collaborations do exist with, emphasis on hospitals and biomedical departments of leading Indian academic agencies, such as Bharathiar University Coimbatore, University of Mysore, AIIMS and Delhi University. A deeper examination of keywords and time scale reveals that majority of these collaborations are quite recent (over the past five years), suggesting that life scientists in defence laboratories are now exploring external ties and newer methodologies, particularly in biomedical application based efforts.

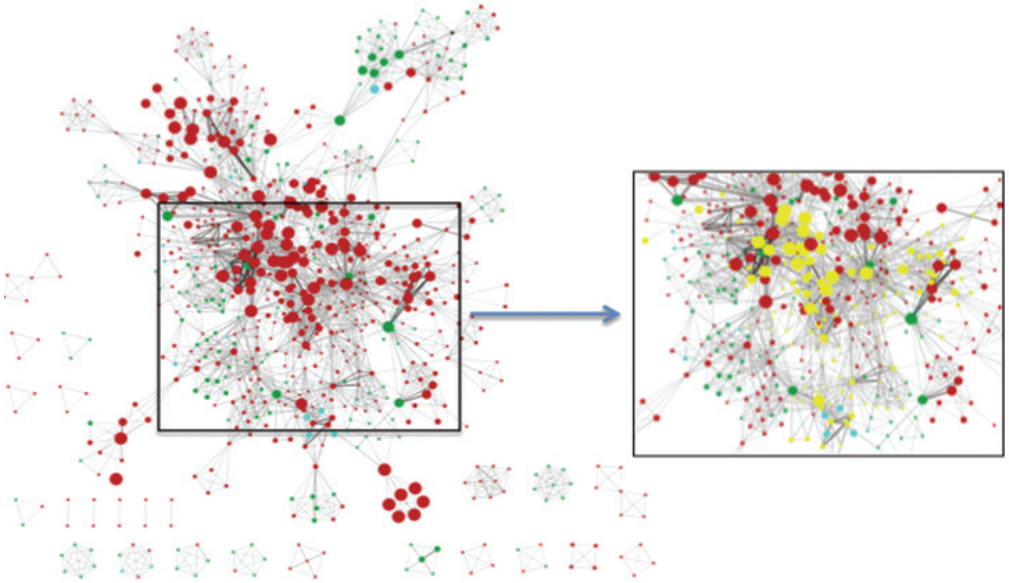


Figure 1. The co-authorship network of national and international collaborations among scientists at Indian DRDO laboratories. Nodes on the network represent individual scientists and a line connecting two of them indicates that they co-authored a paper during the period of study. Node colours represent the parent institutions of authors, namely Red (DRDO), Green (Other Indian organisations) and Blue (International Organisations). Inset on right shows a zoom-in region to signify network coverage by INMAS, the most highly recurring organisation in the dataset.

Table 1. Basic Statistical properties of the Life science specific co-authorship network in Indian DRDO Labs

Property	Full network	Largest connected component
Network nodes	741	638
Edges in network	3004	2808
Isolated nodes	NIL	NIL
Average degree	8.108	8.803
Network density	0.011	0.014
Network diameter/radius	8/1	8/4
Year of publications	2001 to 2016	2001 to 2016

3.2 Overlapping Modules in Co-authorship Data reveal Network dynamics

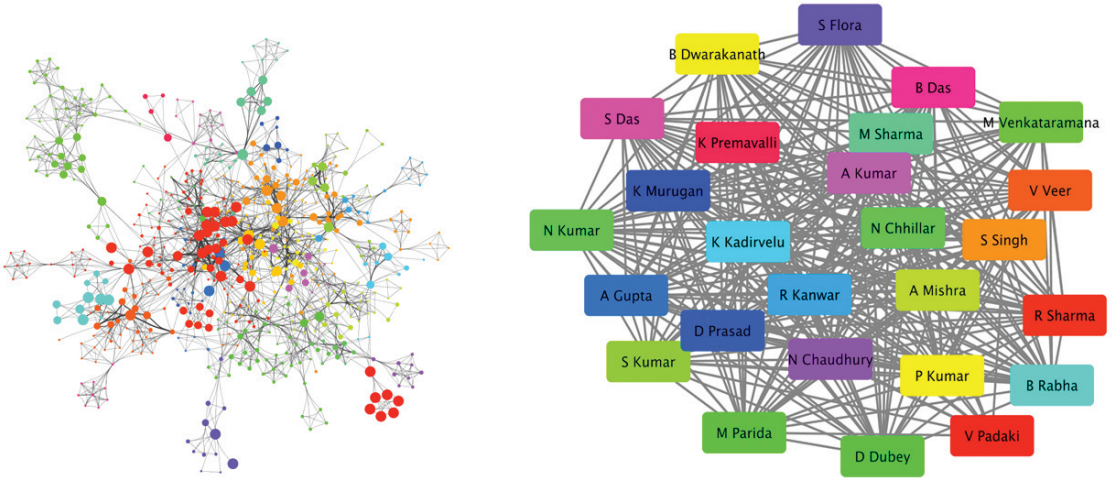
We performed clustering of nodes in the co-authorship network in order to understand diversity of research topics and overlaps in the interaction structure. 26 distinct clusters were identified in the network, resulting in a dynamic 'meta-network' of 26 nodes connected by 299 edges, as shown in Fig. 2. Relationships between nodes of the co-authorship network allowed us to infer network dynamics such as 'bridging' of two existing groups. For example, the name of Dr Ravi B Srivastava (DIHAR, DRDO) emerges over the past decade, as a recurrent strong link between DIHAR and DRL, Tezpur, apart from connecting DIHAR with several other research groups within and outside the DRDO and India, such as DIPAS, INMAS, AIIMS, IVRI Izatnagar and the University of Toronto, Canada. Similarly, the case of Dr Shashi Bala Singh, Director, DIPAS, provides an example of a congruent nexus-node for scientific collaborations between DIPAS and other leading academic institutes like DIHAR (Ladakh), IBER (Haldwani), Jamia Millia Islamia, INMAS, AIIMS and Banasthali University among others. Several other similar cases can be identified directly from the co-authorship network.

In contrast, the meta-network of overlapping modules enables a hierarchical zoom-in analysis of large networks and in this particular case, elucidates significant roles played by individuals in maintaining fruitful productivity as well as structural integrity of the national network, as in case of Dr S.J.S Flora (Senior scientist at DRDE, Gwalior). As can be seen in Fig. 2(b), the first cluster, represented by Dr Flora, overlaps with at least nineteen other functional modules of the meta-network, an association that cannot be derived by topological analysis of the original co-authorship network, but only from the modular meta-network, mainly because it isn't just about one individual, but multiple individuals of a given theme/institution, who may be connecting with multiple nodes of other clusters. Most importantly, such deeper relationships uncover the cognitive structure of the full network at the level of research themes, or institute-specific mandates, as shown in distinct colours Fig. 2(a). Specifically, the edges that connect a given node with a node of a different colour signify integration of expertise and subjects among scientists. A large number of connections are between nodes of the same color, although some nodes of each color radiate out into the larger network and link to other clusters. The Fig. 2(a) network clearly shows links between all possible pairs of modules, suggesting a healthy trend of integrative and collaborative science, a pattern that needs to be further facilitated and encouraged.

In summary, the community-level module concept opens a wide range of possibilities to develop new approaches and applications including network classification, comparison and prediction.

4. DISCUSSION

We have focused on life-science specific co-authorship networks across defence research laboratories in India, in order to investigate the scientific communication cultures.



(a) (b)

Figure 2. The meta-network representing 26 distinct yet overlapping modules of the original co-authorship network. This particular network has several overlapping sub-communities based on research themes and ethnographic collaboration patterns. Each of the nodes here represent modules of the original data and the edges are weighed according to the research topic overlaps between different modules. Panel A shows the original co-authorship network coloured according to identified modules, while panel B shows the color-coded meta-network of individual modules.

Authors represent the individual organisational unit of our co-publication network, and we identify clusters corresponding to specific institutions or prolifically publishing authors, across multi-centre research collaborations. Structurally we find two broad classes of patterns between clusters: 'exclusive' connections where a single node (author) connects two clusters, and 'overlapping' connections where substantial nodes of two clusters overlap. Within these broad classes, further subclasses can be identified that correspond to node-centric collaboration and personnel movement within the defence establishment.

It is important to note that there are many cases of scientific collaborations that do not result in co-authored publications, and that co-authorship in science is not the only form of scientific collaboration; several other indicators of collaboration between scientists have been suggested, including shared supervision of PhD scholars, writing grant proposals together, participation in formal research programs and shared organization of scientific conferences[9]. Accordingly, a large number of scientific collaborations are invisible in formal communication channels either because they do not result in co-authored publications or in formal acknowledgments in scientific texts[10,11] and this aspect must be kept in mind before drawing conclusions from the present co-authorship network analysis. It has also been argued that the organizational units of modern science are groups and not individuals[1], but it remains undisputed that co-authorship in science presents a more substantial indicator than just scientific communication in one way or another. Thus, in this work, we have interpreted the term 'collaboration' primarily to designate research that results in co-authored publications. In summary, this work represents one of the first comprehensive reviews of network analysis of co-authorship in defence research establishment, a fast evolving community with huge potential for basic and applied life science, suggesting a well-timed commencement of a journal dedicated to the subject.

REFERENCES

1. Ziman J. Are debatable scientific questions debatable? *Social Epistemology*, 2000, **14**(2-3), 187-199. doi: 10.1080/02691720050199225
2. Li, E.Y.; Liao, C.H. & Yen, H.R., Co-authorship networks and research impact: A social capital perspective. *Research Policy*, 2013, **42**(9), 1515-1530. doi: 10.1016/j.respol.2013.06.012
3. Newman, M.E. Coauthorship networks and patterns of scientific collaboration. *Proc. Natl. Acad. Sci.*, 2004, 101(suppl 1), 5200-5205. doi: 10.1073/pnas.0307545100
4. Mali. F.; Kronegger, L.; Doreian, P. & Ferligoj, A. Dynamic scientific co-authorship networks. *In* Models of science dynamics. 2012, pp. 195-232. doi: 10.1007/978-3-642-23068-4_6
5. Lambiotte, R. & Panzarasa, P. Communities, knowledge creation, and information diffusion. *J. Inform.*, 2009, **3**(3), 180-190. doi: 10.1016/j.joi.2009.03.007
6. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B. & Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, 2003, **13**(11), 2498-2504. doi: 10.1101/gr.1239303
7. Yadav, G. & Babu, S. NEXCADE: Perturbation analysis for complex networks. *PloS one*. 2012, **7**(8), e41827. doi: 10.1371/journal.pone.0041827
8. Szalay-Bekő, M.; Palotai, R.; Szappanos, B.; Kovács, I.A.; Papp, B. & Csermely, P. ModuLand plug-in for Cytoscape: determination of hierarchical layers of overlapping network modules and community centrality. *Bioinformatics*, 2012, **28**(16), 2202-2204. doi: 10.1093/bioinformatics/bts352
9. De Haan, J. Authorship patterns in Dutch sociology. *Scientometrics*, 1997, **39**(2), 197-208. doi: 10.1007/BF02457448
10. Melin, G. & Persson, O. Studying research collaboration using co-authorships. *Scientometrics*, 1996, **36**(3), 363-377. doi: 10.1007/BF02129600
11. Laudel, G. What do we measure by co-authorships? *Research Evaluation*, 2002, **11**(1), 3-15. doi: 10.3152/147154402781776961

ACKNOWLEDGEMENT

Computational resources provided under the Bio-technology Information System Network (BTISNET) grant of the NIPGR Sub-Distributed Information center (Sub-DIC) of the DBT, Govt. of India, are gratefully acknowledged.

CONTRIBUTORS

Dr Gitanjali Yadav is a specialist in computational and structural biology with specific applications in plant genomics and phytochemistry. She is a group leader in genomics and systems biology at the NIPGR, New Delhi. She has developed of one of the largest repositories of phytochemicals, opening up new avenues of research in plant communication and specialised metabolism. Her interests lie in the area of molecular evolution and social network analysis, wherein she has developed methods to automate induction and measurement of perturbations in complex real world networks.

Dr Suresh Babu is an Associate Professor at the School of Human Ecology (SHE), Ambedkar University Delhi. He is a restoration ecologist and has extensively worked on rainforest ecosystems for the past two decades. His interests are in Systems Ecology and in the analysis of complex systems that can be conceptualised by networks, and represented as interconnected matrices of interactions.