

Evolution of the Social Structure of Nanotechnology

Staša Milojević
University of California, Los Angeles
Department of Information Studies
Los Angeles, CA 90095-1520

stasa@ucla.edu

ABSTRACT

New disciplines require the presence of a social infrastructure for the research to become institutionalized. I study the change of structural arrangements to determine the social development of nanotechnology as an emerging discipline. In order to understand structural properties of social infrastructure, I have conducted the empirical study of the time evolution of co-authorship network of scientists in nanotechnology. In particular I have examined static topological properties of networks and underlying dynamical processes leading to particular topologies.

The problem of empirical delineation of a discipline in this study is to a great extent addressed by using a database of field-specific publications. Namely, I built networks using data on half a million nanotechnology articles covering 35 years (1970-2004) from NanoBank database at UCLA. The articles have been selected from Science Citation Index Expanded, Social Science Citation Index, and Arts and Humanities Citation Index using two methods: search for 379 predefined terms and probabilistic procedure for the automatic identification of terms [3]. My dataset contains over 460,000 different authors, or more precisely distinct names in a database. After evaluating the uncertainty due to author name disambiguation, I decided to use author's last name and all the initials as a unit of analysis.

My analysis is based on non-directed networks where the nodes are authors who have collaborated (coauthors), edges are coauthoring events, and the edge weight is the number of articles on which the two authors have collaborated. For the time period studied the number of coauthors varies from 255 in 1970-71 to 184,448 in 2003-2004. Since the fraction of non-connected authors, i.e. authors who have not coauthored at all, has been declining from 11-12% in the 1970s and early 1980s to less than 2% in 2004, the number of researchers that cannot be studied with coauthorship networks is relatively small. Therefore, it is safe to say that the results obtained by the analysis of coauthorship networks can be used to make inferences about the social

structure of the field as a whole.

The distribution of the number of papers people publish and the number of authors per paper are considered two primary constraints on the shape of the collaboration network [1]. The average number of collaborators per author in nanotechnology has been increasing, especially since 1990s, reaching 13.7 and 15.8 collaborators per author (for 1995-99 and 2000-04, respectively). The number of collaborators per author is strongly related with the number of authors per paper ($R^2=0.86$) as well as with author productivity ($R^2=0.83$). However, the productivity is not independent from the number of authors per article, but is very tightly correlated ($R^2=0.93$), making the correlation between the coauthorship and productivity more fundamental than the correlation of each with the level of collaboration.

For each author in coauthorship network we can get the number of authors connected to it (i.e. the degree of a node). The degree distribution is the probability distribution of these degrees over the whole network. The analysis of two most recent time periods (1995-99 and 2000-04) shows that most authors have three collaborators, which gives scale to the distribution. The number of authors with an increasing number of collaborators falls monotonically in a power-law fashion. The part of the distribution below 20 collaborators follows very closely the lognormal distribution. The lognormal distribution stops being a good model when there are 20 or more collaborators per author. From that point the power law is an excellent description of the distribution.

A cluster is defined as a subset of nodes interconnected by links. The largest cluster is also known as a giant component. In nanotechnology the largest component started very small and similar in size to the 2nd largest component, but increased dramatically in 1980-85. In 1995-99, the largest component grew to >80%, reaching the size similar to giant components in other established scientific disciplines [2]. The existence of a giant component indicates wide-ranging collaboration not constrained by research specialty. The formation of the giant component in this study coincides with generally held view that nanotechnology formed as a field in the early 1990s. The robustness of the giant component was tested by random removal of constituents of the network. The results show that the giant component falls below 50% only after 93% of the original number of actors are removed.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference '04, Month 1–2, 2004, City, State, Country.
Copyright 2004 ACM 1-58113-000-0/00/0004...\$5.00.

This shows that the giant component is not sensitive to a significant removal of its structural elements.

The shortest path is the path (a collection of edges that connect two nodes) that passes through the minimal number of other nodes. The network diameter is the longest distance between the researchers connected by collaboration ties. Except for the first time interval, before the recognized emergence of nanotechnology, this separation has been steadily decreasing, despite the enormous growth of the field (researchers in collaboration network) and is currently (2000-04) around 6. In other words, for authors in the giant component (which represents 90% of all authors) we need on average 6 and a maximum of 23 steps to connect one to another. Small-world structure enables quick and efficient transmission of information.

I confirm the emergence in the 1990s of the robust giant component and of small-world structure in the coauthorship networks in the field of nanotechnology. These results indicate that nanotechnology has formed as a field at that time.

Keywords

Co-authorship networks, network topology, small-world networks, nanotechnology.

1. ACKNOWLEDGMENTS

Certain data included herein are derived from Nanobank (Lynne G. Zucker and Michael R. Darby, Nanobank Data Description, release 1.0 (beta-test), Los Angeles, CA: UCLA Center for International Science, Technology, and Cultural Policy and Nanobank, Jan 17, 2007.) and from the Science Citation Index Expanded, Social Sciences Citation Index, and Arts & Humanities Citation Index of the Institute for Scientific Information®, Inc. (ISI®), Philadelphia, Pennsylvania, USA: © Copyright Institute for Scientific Information®, Inc. 2006. All rights reserved.

2. REFERENCES

- [1] Moody, J. 2004. The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999. *American Sociological Review* 69(2): 213-238.
- [2] Newman, M. E. J. 2004. Who is the best connected scientist? A study of scientific coauthorship networks. E. Ben-Naim, H. Frauenfelder and Z. Toroczkai. Berlin, Springer: 337-370.
- [3] Zucker, L. G., M. R. Darby, et al. 2006. Minerva unbound: Knowledge stocks, knowledge flows and new knowledge production. NBER Working Paper No. 12669.