Scientific collaboration dynamics

Anuška Ferligoj
Franc Mali, Luka Kronegger
Toni Pustavrh, Marjan Cugmas, Hajdeja Iglič

Faculty of Social Sciences, University of Ljubljana

Torino - Ljubljana Meeting
Ljubljana, February 4-5, 2016

Introduction

Slovenian researchers have studied the last eight years scientific collaboration using

- bibliometric analysis (Ferligoj and Kronegger 2009; Mali et al. 2010, Kronegger et al. 2011, 2012, 2014, Ferligoj et al. 2015, Cugmas et al. 2016)
- survey analysis (Iglič et al. 2014)
- qualitative approach (Groboljšek et al. 2014)

of co-authorship networks using longitudinal data on the Slovenian science system in order to explore and explain their dynamics, first, across four scientific disciplines:

- **Mathematics** - an old discipline where research takes place primarily in offices
- **Physics** - an old discipline where the research occurs mostly organized into research groups within laboratories
- **Sociology** - an old discipline where research also occurs mostly in offices
- **Biotechnology** - a new laboratory discipline

Seven scientific fields in Slovenia

Later the main goal was to identify the key factors driving collaboration and the main differences in collaboration behavior across all scientific fields and disciplines.

ID	Scientific field	No. of disciplines
1	Natural sciences and mathematics	9
2	Engineering sciences and technologies	19
3	Medical sciences	9
4	Biotechnical sciences	6
5	Social sciences	11
6	Humanities	12
7	Interdisciplinary studies	2

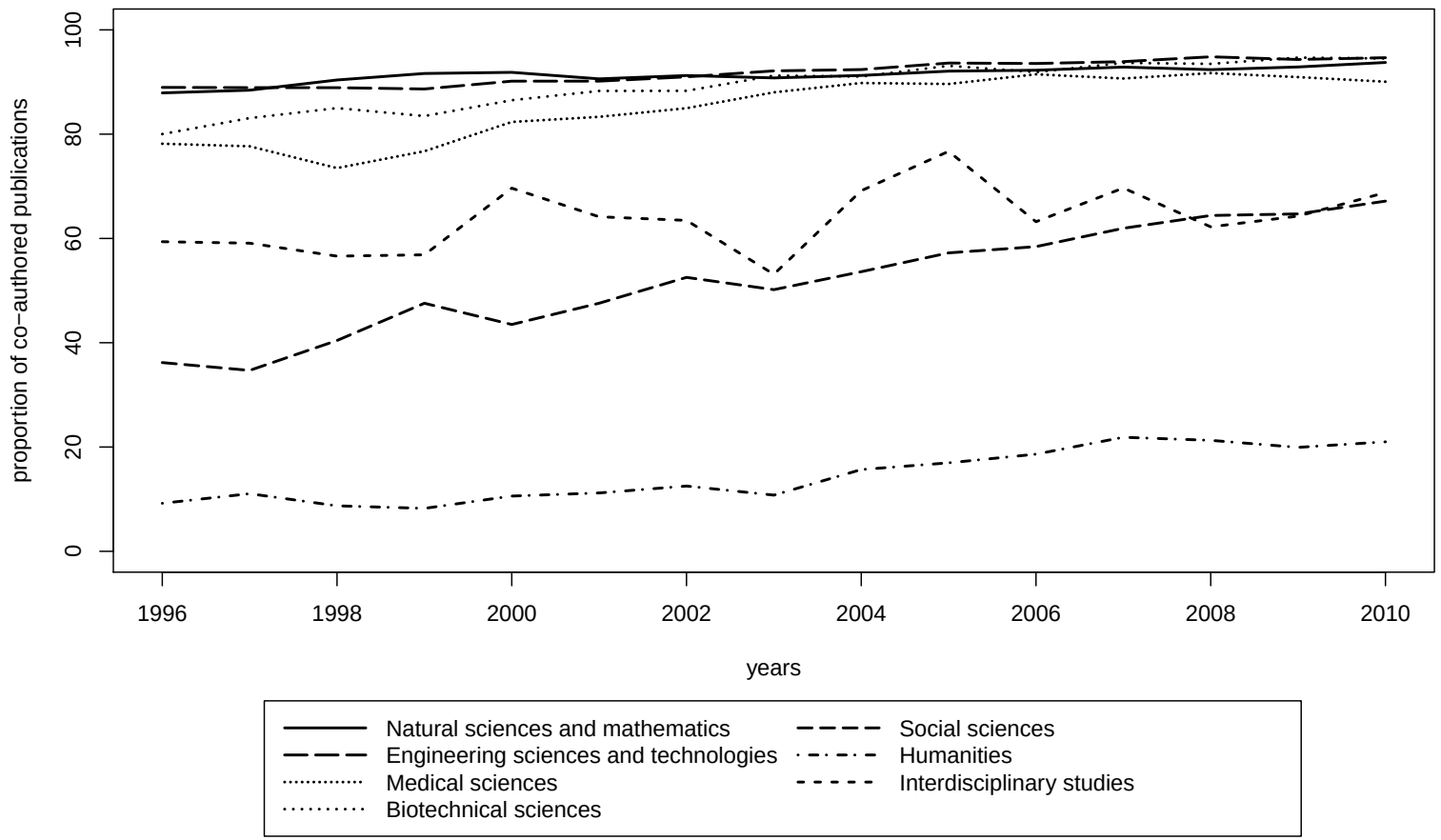
Bibliometric analysis

- **Current Research Information System (SICRIS)** which includes information on all current and former researchers registered with the Slovenian Research Agency and
- **co-operative On-Line Bibliographic System & Services (COBISS)** which is an officially maintained database of all publications available in Slovenian libraries. From this system, we collected complete scientific bibliographies of all Slovenian researchers who had ever been given a research identification number (ARRS ID) by the Slovenian Research Agency.

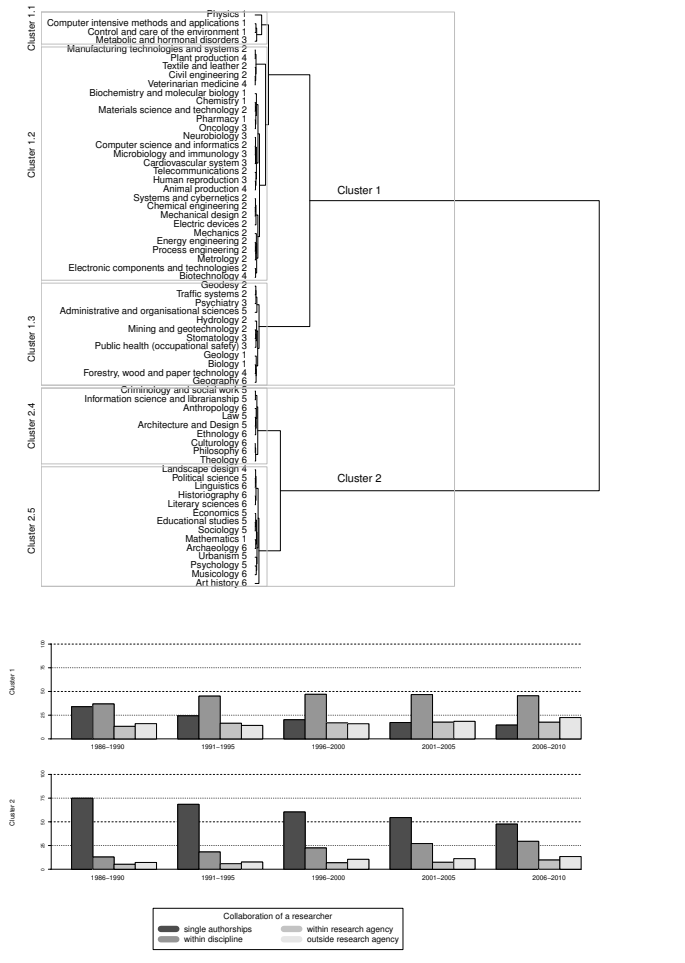
A tie was defined if two researchers appeared together as authors in at least one publication.

- The total number of researchers with an ARRS ID who published in the time period 1996–2010 was 15,424.
- These researchers collaborated with another 48,191 authors not registered with ARRS.
- Together, they published 170,118 publications that are, according to the evaluation criteria of ARRS, treated as scientific outputs.
- The data about discipline memberships were provided by the researchers themselves when they applied for an identification number.

Percentages of co-authored publications in seven scientific fields in Slovenia from 1996 to 2010



Symbolic clustering of disciplines according to collaboration structures in 5 time periods



Modelling co-authorship network dynamics

Kronegger et al. (2012) and Ferligoj et al. (2015) combined two approaches for modelling co-authorship network dynamics. They used the small-world model (Watts and Strogatz 1998) and the mechanism of preferential attachment, also known as the process of cumulative advantage (Price 1963, 1965; Garfield and Merton 1979). One dimension of the small world was measured by its clustering level, and preferential attachment was operationalized through the coauthorship of researchers within and across disciplines.

Actor-oriented model

Between many used network approaches we used actor-oriented model (Snijders 2001, 2005; Snijders et al. 2007, 2010) for longitudinal co-authrorship network data. The model is defined as a continuous-time Markov process.
Since our data are **non-directed networks**, a modification to the models of Snijders (2001, 2005) was required. To obtain a non-directed network, the assumption was made that at random moments, a randomly chosen actor ('ego') chooses another actor ('alter') to propose a new tie or to drop an existing tie; if a new tie is proposed, alter can decide to accept or reject the proposal. The choice by ego of alter is a multinomial choice, and the acceptance decision by alter is a binary choice.

Program

The probability models for these choices are based on a linear predictor similar to generalized linear models.

Stochastic-actor-based model (SAOM) is implemented in the SIENA program.

We used RSiena.

Three 5-year intervals

- **Period 1, 1996–2000**: a period of harmonization with the European Union (EU) and the OECD standards;
- **Period 2, 2001–2005**: in 2004, Slovenia became a member of the EU. The Slovenian Research Agency was established in the same year followed by many positive effects on R&D evaluation procedures due to its policies;
- **Period 3, 2006–2010**: a more stable period.

Results

Scientists collaborating at one point in time can choose their co-authorship tie at a later time. We considered the possibility that ties can be **created or maintained** since this is a feature that characterizes co-authorship networks.

Estimated parameters for the six scientific fields

parameters	1 Nat		2 Eng		3 Med		4 Bio		5 Soc		6 Hum	
rate 1	22.061	(0.840)	18.492	(0.577)	45.007	(1.580)	29.298	(1.125)	32.655	(2.407)	14.838	(6.879)
rate 2	27.210	(1.009)	26.031	(0.630)	54.946	(1.287)	32.976	(1.332)	35.265	(1.270)	16.579	(5.738)
degree (density)	-2.360	(0.020)	-2.550	(0.018)	-2.108	(0.017)	-1.657	(0.029)	-2.400	(0.029)	-3.448	(1.248)
transitive triads	0.458	(0.010)	0.710	(0.010)	0.352	(0.007)	0.371	(0.015)	0.450	(0.017)	1.734	(0.354)
same research group	1.540	(0.038)	2.017	(0.035)	1.263	(0.028)	0.924	(0.052)	1.494	(0.048)	2.290	(1.155)
degree of alter	-0.025	(0.002)	-0.064	(0.003)	-0.018	(0.002)	-0.025	(0.003)	-0.047	(0.004)	-0.104	(0.033)
degree out	0.171	(0.012)	0.212	(0.012)	0.170	(0.010)	0.128	(0.015)	-0.029	(0.013)	-0.145	(0.040)
excellence	-0.117	(0.033)	-0.009	(0.028)	-0.009	(0.025)	-0.034	(0.035)	0.534	(0.034)	0.537	(0.152)
first publication year	0.012	(0.001)	0.010	(0.001)	0.022	(0.001)	0.018	(0.002)	0.013	(0.002)	0.008	(0.007)
first pub. similarity	0.111	(0.054)	-0.070	(0.073)	0.023	(0.052)	0.241	(0.095)	0.055	(0.067)	-0.327	(0.798)
PhD (yes)	0.988	(0.043)	1.015	(0.031)	0.711	(0.023)	0.482	(0.041)	0.921	(0.049)	0.501	(0.131)
gender (male)	0.102	(0.027)	0.166	(0.039)	0.169	(0.020)	0.174	(0.029)	-0.244	(0.030)	0.187	(0.109)

Shaded estimates are not statistically significant;

there are standard errors in parentheses.

Survey analysis

A web survey of researchers from the four scientific disciplines was designed to understand which kinds of incentives, perceptions, and personal strategies help account for collaboration from the perspectives of individual scientists (Iglič et al. 2015).
The results of the analysis showed that the differences between disciplines in the proportion of researchers active work time spent collaborating with others were much smaller when assessed through interviews compared to results from co-authorship data. Researchers in the social sciences do not necessarily collaborate less as the attributions of authorship are different in the social and natural sciences.

Disciplines vary according to the nature of collaboration partners. Physicists and mathematicians from the basic sciences have much wider and far-reaching collaboration networks than sociologists or biotechnologists from applied sciences. The extensive collaboration networks of biotechnologists seem to be more limited and focused on local partners.

Qualitative approach

For an even more detailed analysis of collaboration practices, a qualitative investigation among key representatives of the four scientific disciplines and research policymakers was conducted to gain an insight into their views on the importance of scientific collaboration (Groboljšek et al. 2014).
While the policy mechanisms aiming to encourage scientific collaboration are important, the interviewed scientists and policymakers believed that longterm and successful collaborations derive from researcher's efforts and their individual engagements—but only where suitable conditions have been created by policy mechanisms which encourage international mobility, along with interdisciplinary and interinstitutional/ intersectoral collaboration.

New developments

- The effects of funding and coauthorskih on research performance in a small scientific community (Mali, Pustavrh, Platinovšek, Kronegger, Ferligoj)
- Stabilty of co-authorship structures (Cugmas, Ferligoj, Kronegger)
- Identification of hierarchies in scientific community (Kronegger)