NEMO
Network Models, Governance and R&D Collaboration Networks

Newsletter #3
April 2008

Searching for Communities in EU Framework Programme Networks

Structural differences or similarities of large networks like those generated in the European Framework Programmes are not obvious at a glance. The cluster identification algorithm BRIM developed in NEMO reveals thematic communities within the bipartite networks of organisations and projects.

Bipartite networks of organisations and projects are a better representation of collaboration networks than (their projection) networks of organisations only. For these bipartite networks, a new algorithm for identifying sub-communities has been developed. The algorithm is applied to networks derived from the EU Framework Programmes. Community groups identified are compared using information-theoretic methods. In Fig.1, the communities are shown as vertices in a network, with the vertex positions determined using spectral methods. Vertex size corresponds to intra-community linkage, edge width with inter-community linkage. The communities can be meaningfully differentiated with respect to their thematic profile resulting from the projects involved (colour).

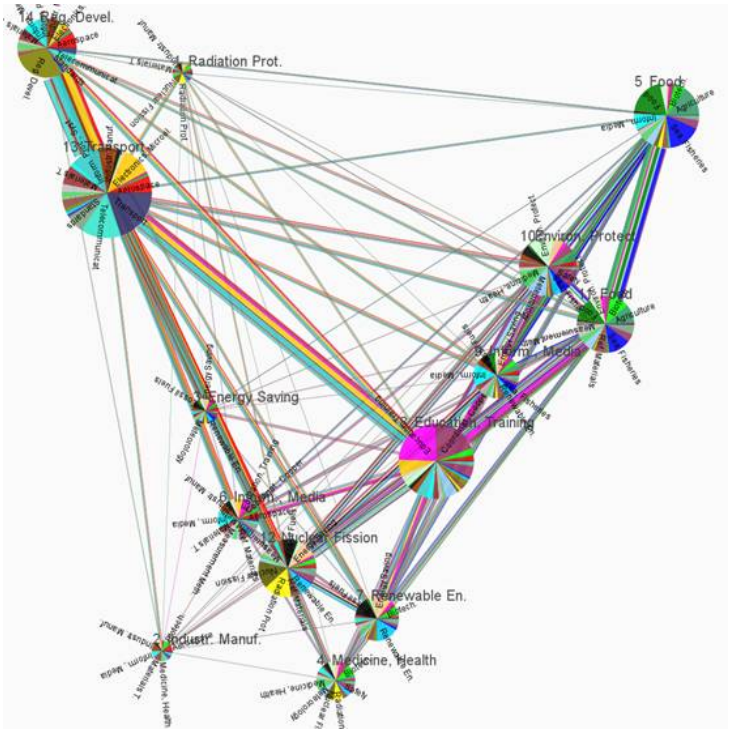

Fig.1: Communities in the bipartite network in FP3. [Michael J. Barber, Margarida Faria, Ludwig Streit and Oleg Strogan (2008), Searching for Communities in Bipartite Networks, NEMO WP#9]

The NEMO Newsletter

The objective of the periodic NEMO Newsletter is to provide a platform of interdisciplinary information exchange and discourse for all sciences concerned with complex interorganisational R&D collaboration networks, and to promote the NEMO project worldwide. The Newsletter will offer regular insights into the NEMO project and document its results including previews of NEMO publications.

Other continuous features include the publication of short articles, comments on interesting links, and information about events or publications which are located in the area of our research. External contributions are welcome and should be addressed to the editor. The Newsletter is published quarterly on the NEMO website
http://www.nemo-net.eu

Contents of this Issue

Project Progress 2
WP 2 - Structure and dynamics of complex random graphs and associated processes 2
WP 4 - Empirical network analysis 2
Recent NEMO Publications (Selected) 3
Recent Meetings 4
Workshop Funchal, Madeira (7/8 Mar 2008) 4
Comments and Feedback 5
Imprint 6
Disclaimer 6

PROJECT PROGRESS

Newsletter #1/2 has started to draw a picture of the main results NEMO has achieved so far, focusing on three workpackages (WP 1, WP 3, and WP 5). The current newsletter follows up and adds an account of the remaining workpackages (WP 2 and WP 4).

WP 2 - STRUCTURE AND DYNAMICS OF COMPLEX RANDOM GRAPHS AND ASSOCIATED PROCESSES

This workpackage aims at developing mathematical tools for precise analysis of the structure of and dynamic processes on complex random networks, having generic features of real-world social and economic networks. Models of knowledge production and diffusion on such networks are implemented and matched with the empirical data.

(a) Different classes of combinatorial structures, such as generalised plane oriented recursive trees, generalised bucket recursive trees, or random bipartite graphs and their projections were developed and analysed to serve as models for real-world networks. Moreover, the mathematical framework of information theory and statistics was applied to complex networks for obtaining methods to characterise such networks structurally. The ultimate goal was to develop methods which are easily computable (i.e., in polynomial time complexity) and lead to a better understanding of information processing in networks.

(b) In order to understand the differing performance of different networks related to information propagation, a model of generalised epidemic processes was developed. Here, a new network indicator, the communication index, was introduced and its statistical properties regarding the diffusion of information were studied both analytically and numerically.

(c) With the final aim to generate formal models for knowledge exchange and network formation, a first model of knowledge production and exchange was formulated and analysed with computer simulations. Hereby, traders and givers of knowledge were distinguished, yielding interesting results on the effectiveness of different network architectures: For example, networks with structural holes perform well, when knowledge is scarce; when knowledge is abundant, however, the networks with high social capital (high clustering) perform well.

WP 4 - EMPIRICAL NETWORK ANALYSIS

The aim of this workpackage is to improve available data on politically induced R&D collaboration networks, to construct different types of empirical networks by, e.g., governance regime, thematic classification, or geographical characteristics, and to analyse their structural properties as well as their temporal evolution. To this end, appropriate software tools for network investigation and representation are developed.

The sysres EUPRO database presently comprises information on 50,590 research projects of the EU FPs and 49,644 participating organisations. It contains systematic information on project objectives and achievements, project costs, project funding, temporal location and contract type and provides detailed insight on participating organisations including the full name, the geographical location, the participating department, the contact person and the type of the organisation. Thus, the sysres EUPRO database represents an invaluable source for any kind of empirical network analysis within NEMO. Four different lines of analysis have been performed up to now:

(a) Spectral analysis of the networks reveals information on the structural characteristics and the capacity for flow (of information) within the network, and thus allows quality considerations for network simulations.

(b) Similarity measures for graphs and random graph families were developed in order to find sub-communities and new measures of complexity, characterizing the structure of large networks.

(c) Cluster identification and characterization algorithms were developed and implemented in order to divide the network into distinct community groups of high inner similarity, which simplifies analysis of the network as a whole, and lends insight into mesoscopic structures of a network. A new algorithm, BRIM, was developed for identifying and visualising communities directly in bipartite networks, yielding higher-quality clusters within the collaboration networks. The algorithm was used to identify communities with respect to organisation type and thematic area.

(d) A spatial analysis of the collaboration networks – via a Poisson spatial interaction model – revealed geographical and technological effects on cross-regional collaborative activities: Geographical proximity and co-localisation of organisations in neighbouring regions are important determinants of cross-region collaborations, but less important than technological proximity. To a smaller extent, R&D collaborations are also impeded by language barriers, while foreign partnership in

general is a small but observable barrier to collaboration.

RECENT NEMO PUBLICATIONS (SELECTED)

The sysres EUPRO database manual

by Michael Barber, Barbara Heller-Schuh, Thomas Roediger-Schluga, and Thomas Scherngell (ARC sys)

Workpackage 4, Deliverable D4.2, submission date: 2008-01-29

In this deliverable we describe in some detail the sysres EUPRO database, a novel data source on funded research projects of the first six EU Framework Programmes (FPs). The EU FPs are set up to foster economic competitiveness and to stimulate knowledge diffusion across European countries. We shed some light on data acquisition and construction as well as on some structural properties of the database. The sysres EUPRO database is one of the key information sources in NEMO, in particular for the empirical analysis of network structures and network processes within the European research area.

The core data source for the construction of the sysres EUPRO database is the CORDIS projects database (CORDIS search 2006). It contains detailed information on funded projects and project participants of EU FPs. For the funded projects the CORDIS database lists information on project objectives and achievements, project costs, project funding, temporal location and contract type. It also provides valuable information on the project participants including the full name, the geographical location, the participating department, the contact person and the type of the organisation. The sysres EUPRO database presently comprises information on 50,590 research projects of the EU FPs and 49,644 participating organisations. Considering also subentities the number of different organisations amounts to 55,454.

Modularity and community detection in bipartite networks

by Michael J. Barber (ARC sys)

Workpackage 4

Phys. Rev. E 76, 066102 (2007), preprint available at arXiv:0707.1616v3 [physics.data-an]

The modularity of a network quantifies the extent, relative to a null model network, to which vertices cluster into community groups. We define a null model appropriate for bipartite networks, and use it to define a bipartite modularity. The bipartite modularity is presented in terms of a modularity matrix **B**; some key properties of the eigenspectrum of **B** are identified and used to describe an algorithm for identifying modules in bipartite networks. The algorithm is based on the idea that the modules in the two parts of the network are dependent, with each part mutually being used to induce the vertices for the other part into the modules. We apply the algorithm to real-world network data, showing that the algorithm successfully identifies the modular structure of bipartite networks.

Determinants of collaboration in European R&D networks: Empirical evidence from a binary choice model perspective

By Manfred Paier and Thomas Scherngell (ARC sys)

Workpackage 1, April 2008

Available at SSRN: http://ssrn.com/abstract=1120081

We adopt an econometric perspective to identify determinants of link formation, including various actor characteristics, relational and network effects as well as geographical effects. We employ a binary choice model estimated by means of logistic regressions, with a dependent variable that represents the establishment of a formal cooperation between two organizations in FP5. We use data on EU FP projects from the sysres EUPRO database and from a representative survey of FP 5 participants.

The study produces statistically significant evidence that R&D collaboration choices of organizations participating in European FPs are affected by geography, FP experience and relational factors including network characteristics. Thematic proximity matters more than geographic proximity, while most influential for collaboration appears to be prior acquaintance of the actors. Also, network effects sig-

nificantly determine the collaboration choice, but to a slightly smaller extent than geographical effects.

Integrated Projects in FP6: Intra-project linkages and project performance

By Petra Wagner-Luptacik, Barbara Heller-Schuh, Manfred Paier, and Jeannette Müller (ARC sys).

Workpackage 1, Milestone 1.3b, NEMO Working Paper #8, September 2007

In a case study approach, qualitative and quantitative research revealed stylised facts on intra-project linkages and project performance, thus representing a valuable complementary data source to the large-scale collaboration data. According to this, real life projects are often not fully connected networks but may – dependent on project size – consist of several weakly interlinked cliques, especially regarding the joint production or exploitation of knowledge.

The Degree Distribution of Thickened Trees

By Matthias Dehmer, Michael Drmota, Bernhard Gittenberger, and Alois Panholzer (Vienna University of Technology)

Workpackage 2, NEMO Working Paper #9, October 2007

We develop a combinatorial structure to serve as model of random real world networks. Starting with recursive trees or their variants known in the combinatorial literature we substitute the nodes by more complex graphs. In such a way we obtain graphs having a global tree-like structure while locally looking clustered. This fits with observations obtained from real-world networks.

Random Bipartite Networks and their Projections

By Michael Drmota, Bernhard Gittenberger (Vienna University of Technology), Tyll Krüger, and Rainer Siegmund-Schultze (University of Bielefeld)

Workpackage 2, Deliverable D2.1, April 2008

Starting from the specific bipartite cooperation network of relations between research organisations and research projects, which is the topic of the NEMO project, this paper aims to investigate aspects of the general random graph model in the special situation of a bipartite random graph. Bipartite networks appear in many real world situations involving two natural partitions (evident examples: co-authors-papers, actors-movies, gender relations). We investigate the bipartite version of the model with respect to the special question of the degree distribution (i.e. the frequency of sites with a given number of links) for the corresponding projection graph. This projection is obtained by connecting two sites of the first partition, if they are linked to the same site in the second partition.

Recent Meetings

Workshop Funchal, Madeira (7/8 Mar 2008)

The NEMO project meeting took place at the Universidade da Madeira, during the latest Madeira Math Encounters (XXXIV), held at the Centro de Ciências Matemáticas (CCM) in the first half of March 2008. The meeting was organised by Prof. Dr. Ludwig Streit and Dr. Oleg Strogan from the Universidade da Madeira. ARC sys, the co-ordinator of the NEMO project, was represented by Dr. Matthias Weber, Dr. Michael Barber, Dr. Thomas Scherngell, and Manfred Paier. The University of Bielefeld was equally strongly represented by Prof. Dr. Phillippe Blanchard, Dr. Andreas Krüger, Dr. Sascha Delitzscher, and Dr. habil Rainer Siegmund-Schultze. Other participants were Prof. Dr. Robert Cowan (University of Maastricht), Dr. Nicholas Jonard (Luxembourg University), Dr. Terhi Nokkala (University of Surrey), and Ramon Scholz (University of Bremen).

The Madeira Math Encounters itself was focused on analysing complex networks. Particular attention was paid towards:

- Dynamics of Networks - how they evolve, identification of suitable models and simulations to describe real world networks
- Dynamics on Networks - how information can flow on them
- Structure of Networks - cliques, communities, visual representations.
- Spatial analyses of European R&D collaboration networks

The Madeira meeting was mainly used to coordinate the combination of different working parts into one another as well as to inform extensively the other members of the consortium of the sometimes, due to the interdisciplinary approach, very distinct work of various work packages. The second important issue

was related to improving NEMO's dissemination efforts.

NEMO Workshop in Madeira 7/8 March 2008: Robin Cowan and Nicolas Jonard (foreground from left to right)

At the beginning of the meeting, the current project progress was laid out (for further details on specific workshops see above), and the work of the focus groups established during the 1st project meeting in Vienna was briefly discussed. These focus groups, aiming at the integration of the different parts of the NEMO project, are:

- Structural analysis (Ludwig Streit (CCM) and Michael Barber (ARC sys))
- Motivation and cooperation (Andreas Pyka (UniHB) and Nigel Gilbert (UniS))
- Output performance and dynamics on networks (Matthias Weber (ARC sys) and Andreas Krüger (BiBoS))
- Governance rules and clustering (Petra Ahrweiler (UCD) and Lena Kruckenberg (UCD))
- Random graphs, thickend trees and empirical networks (Bernhard Gittenberger (TU Vienna) and Manfred Paier (ARC sys))

The next important issue on the agenda was the dissemination of project results. It was agreed that the second newsletter should be published shortly after the first one and that it will focus on empirical studies within the NEMO project. Moreover, it was argued that the website needs to be reworked in order to make publications of the NEMO consortium more easily available to the public.

Additionally, the next meeting of NEMO's User Group was scheduled to take place parallel to the next general NEMO meeting in Lisbon on 7-8 October 2008. The group comprises policy makers, programme managers, evaluation practitioners, and experts from research performing organisations (industry and research organisations running their own research programmes), i.e., people likely to have a keen interest in the principles of development and design of collaborative research networks. It ensures the practical relevance of the proposed project's findings and helps gear the results and tools developed in the NEMO project such that they promise to meet further application needs. The NEMO consortium involves the User Group on a regular basis.

The Madeira meeting was also used to identify upcoming conferences where parts of the NEMO consortium might take place and further present the project (forthcoming events were listed in Newsletter #1/2).

Like at the WP3 meeting in Hamburg, the Madeira NEMO meeting also set up some more workshops to advance NEMO's research work. Hence, Workpackage 4 will meet in June jointly with the UCD and UniHB. Likewise, Workpackage 1 and 2 are also planning to conduct, yet undefined, workshops in the near future.

The final, but very important, topic briefly addressed at the Madeira NEMO workshop was future prospects of the NEMO project after the year 2009; currently, it seems that the consortia is eager to carry on further research in the same direction. However, more substantial debate on this issue is still to come during the next, especially general, NEMO meetings.

COMMENTS AND FEEDBACK

We welcome feedback on our newsletter and work. Do not hesitate to contact us if you have ideas or comments about what you would like to see covered by the newsletter, or if you would like to write a contribution yourself.

For further information about the project visit www.nemo-net.eu or contact the Project Co-ordinator:

Dr. Matthias Weber
Austrian Research Centers GmbH – ARC
Department of Technology Policy
Tech Gate Vienna
Donau-City-Straße 1
A-1220 Vienna, Austria
Phone: +43 (0)50 550-4561
Fax: +43 (0)50 550-4599
Email: matthias.weber[ät]arcs.ac.at

IMPRINT

Publisher:
NEMO Consortium

Represented by:
Austrian Research Centers GmbH – ARC
Division systems research
Department Technology Policy
Tech Gate Vienna
Donau-City-Straße 1
A-1220 Vienna, Austria
Tel: +43(0)50 550-4500
Fax: +43(0)50 550-4599

Company registration number (Firmenbuchnummer): FN 115980i
Court of registry (Sitz- und Registergericht): Wien
VAT identification number (Umsatzsteuer-ID): ATU14703506

Managing Directors:
Dr. Hans Rinnhofer and Prof. Dr. Erich Gornik

Owned by:
BMVIT Austrian Federal Ministry for Transport, Innovation and Technology (50,5 %)

Content:
Manfred Paier
Department Technology Policy
Austrian Research Centers GmbH – ARC
Donau-City-Straße 1
A-1220 Vienna, Austria

Concept and Design:
Karsten Pötschke
National Institute of Technology Management
University College Dublin
Carysfort Avenue, Blackrock
Co. Dublin, Ireland

DISCLAIMER

CONTENT

The publisher and the authors reserve the right not to be responsible for the topicality, correctness, completeness or quality of the information provided. Liability claims regarding damage caused by the use of any information provided, including any kind of information which is incomplete or incorrect, will therefore be rejected. All offers are not-binding and without obligation.

REFERRALS AND LINKS

The publisher and the authors are not responsible for any contents linked or referred to in this newsletter – unless they have full knowledge of illegal contents and would be able to prevent the readers of this newsletter from accessing those information. If any damage occurs by the use of information presented there, only the author of the respective information might be liable, not the one who has linked or referred to these information.

COPYRIGHT

The publisher and the authors intended not to use any copyrighted material for the newsletter or, if not possible, to indicate the copyright of the respective object.

PRIVACY POLICY

The publisher may collect customer data from newsletter subscribers at the time of subscription to this service, such as name and e-mail address. The data collected will exclusively be used to provide the voluntarily requested newsletter service. As a general principle, no information gathered on subscribers is to be used other than for the provision of the voluntarily requested newsletter service and under no circumstances are third parties to be granted access to the personal details of subscribers; not withstanding legal obligations of the publisher. The use of published postal addresses like telephone or fax numbers and email addresses for marketing purposes is prohibited, offenders sending unwanted spam messages will be punished.

LEGAL VALIDITY OF THIS DISCLAIMER

This disclaimer is to be regarded as part of newsletter. If sections or individual terms of this statement are not legal or correct, the content or validity of the other parts remain uninfluenced by this fact.