Science-Metrix

Measure of homophily in scientific collaboration networks

Application to the co-participation network of the 7th Framework Programme for Research and Technological Development (FP7) of the European Commission

B. St-Louis, G. Roberge, R. Lavoie, D. Campbell, G. Côté, E. Archambault

Sunbelt 2016 | April 3, 2016

Outline of presentation

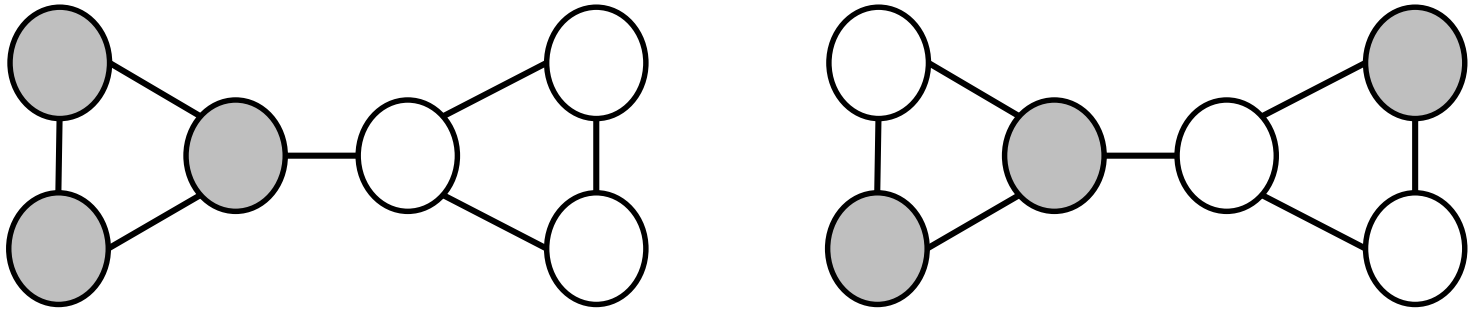
- What is this about?
- Measure of homophily
- Scientific collaboration networks of FP7
- Homophily across sectoral research organizations
- Summary/Conclusion

What is this about ?

- "Birds of a feather flock together." (Lazarsfeld & Merton, 1954)
- In social networks, homophily can be defined as the tendency of actors to exhibit preferential attachments toward actors sharing a common characteristic for some attribute.
- Evidence observed in studies on social networks, including homophily based on ethnicity, age, gender, religion, aspiration, etc.
- This presentation focuses on sectoral homophily.

Metric

- Principle:
  - Observe more ties of a particular type than expected



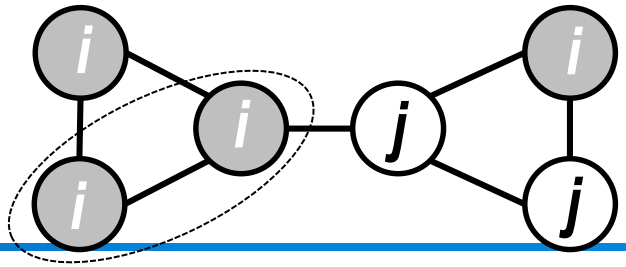
- Measure with:
  - Relative frequency of pairs in the same category compared to pairs from differing categories
  - Deviation from the expected values that a random network would exhibit
  - Compare avg. path length of in category nodes to avg. path length of out category nodes
  - Etc.

Metric

- We measure homophily by comparing homogenous pairs to heterogeneous pairs (with twist to account for the deviation from expected values).
- Homogenous pairs share a common characteristic for some type or attribute, i.
- Considering an actor has equal chance to tie to any other actor in the network, the expected frequency of homophilic pair ii is given by

$$F_{ii} = \left(\frac{N_i}{N}\right)\left(\frac{N_i - 1}{N - 1}\right) * 2$$

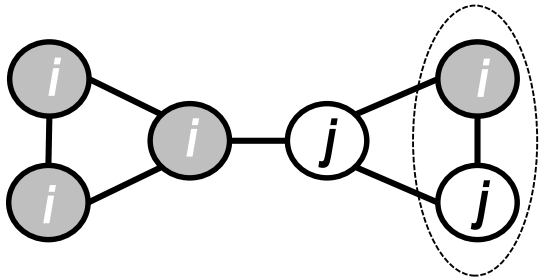
Where N is the total number of actors in the network and N_i is the number of actors of type i



Metric

- Heterogenous pair between node i and j

$$F_{ij} = \left(\frac{N_i}{N}\right)\left(\frac{N_j}{N-1}\right) + \left(\frac{N_j}{N}\right)\left(\frac{N_i}{N-1}\right)$$



- Expected occurrence, O, obtained by multiplying the expected frequency by the weighted sum of all ties, T.
  - Expected homophilic pair occurrence: $O_{ii\ exp} = F_{ii} * T$
  - Expected heterogeneous pairs occurrence: $O_{ij\ exp} = F_{ij} * T$

- Total expected occurrence: $O_{hom\ exp} = \sum O_{ii\ exp} = O_{ii\ exp} + O_{jj\ exp}$

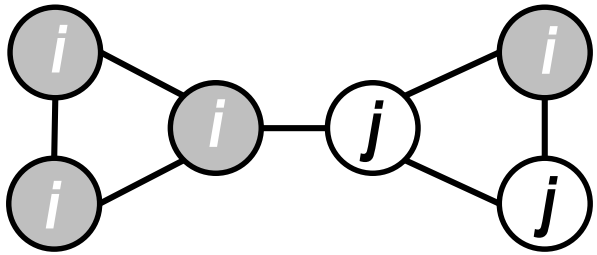
Metric

- The ratio of observed and expected occurrence

$$R_{Hom} = \frac{O_{Hom\ obs}}{O_{Hom\ exp}}$$

- Overall homophily of the network

$$H = \log\left(\frac{R_{Hom}}{R_{Het}}\right)$$



- H indicator
  - H > 0; bias in homophilic pairs stronger than bias in heterogeneous pair
  - H = 0; bias in homophilic pairs equals bias in heterogeneous pair
  - H < 0; bias in heterogeneous pairs stronger than bias in homogeneous pairs

Scientific collaboration networks of FP7

- Science-Metrix was commissioned by the DG Research of the European Commission to carry out an evaluative study of the scientific network of researchers supported by the EC's 7th Framework Programme for Research and Technological Development (FP7).

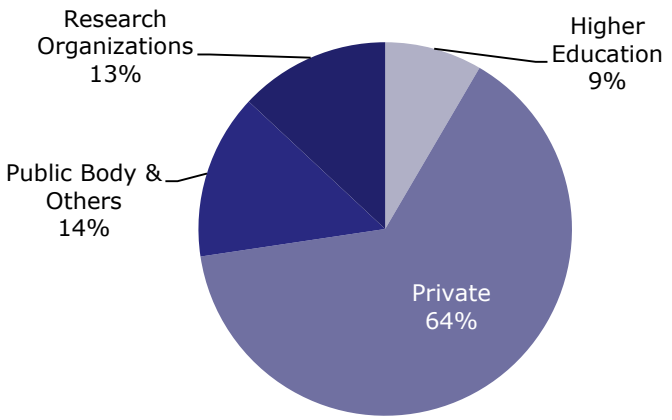
- What is FP7?
  - Europe-wide funding program
  - €50+ billion in grants to support basic research, technological development, and demonstration projects

- Key objective:
  - Promote collaboration at different levels –> Boost innovation
    - International
    - Inter-sectoral
  - Improve competitiveness through innovation

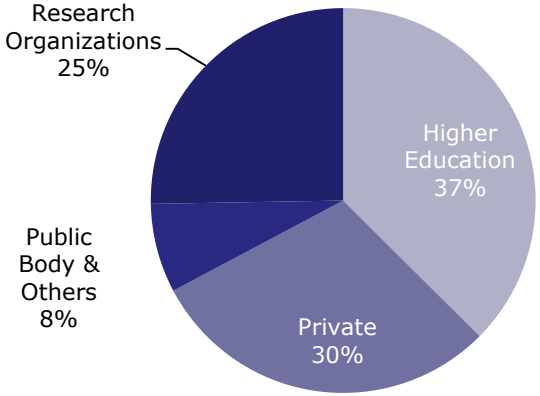
Scientific collaboration networks of FP7

- CORDA database of FP participants
  - 25,000 actors
  - Tied by 22,000 projects
- Actor affiliation
  - Higher education (universities)
  - Research organizations (e.g., Max Planck Society)
  - Public body (e.g., USDA labs)
  - Private for profit (companies)
  - And Others
- Project categorized by thematic areas (next slide)

Distribution of participating organizations per sector for FP7



Distribution of participations per sector for FP7



Scientific collaboration networks of FP7

- Ideas (€8 bn)
  - *Supports risk and high impact research*
- Capacities (€4 bn)
  - Research infrastructures
  - Research for the benefit of small and medium sized enterprises (SMEs)
  - Regions of knowledge
  - Research potential
  - Science in society
  - Support for the coherent development of research policies
  - Activities of international cooperation
- People (€5 bn)
  - *Training and career development of researchers*
- Cooperation (€32 bn)
  - Health
  - Food, Agriculture and Fisheries, Biotechnology
  - Information and Communication Technologies
  - Nanosciences, Nanotechnologies, Materials and New Production Technologies
  - Energy
  - Environment (including Climate Change)
  - Transport (including Aeronautics)
  - Socio-economic Sciences and Humanities
  - Space
  - Security

Homophily across sectoral research organizations

- Did FP7 link together actors from different sectors and bridge the flow of innovative ideas?
- Metric applied to measure difference in sectoral homophily
- More heterogeneous the better -> increased sectoral flow
- Measured across thematic areas and for whole network (FP6 vs. FP7)

Homophily across sectoral research organizations

Programme	Homophily Indicator
FP6 (overall)	**0.14**
FP7 (overall)	**-0.11**
COOPERATION	**-0.11**
Socio-economic sciences and Humanities	0.29
Environment (including Climate Change)	0.23
Health	0.17
Space	0.17
Security	0.10
Food, Agriculture, and Biotechnology	0.09
Information and Communication Technologies	-0.06
Energy	-0.06
Transport (including Aeronautics)	-0.09
Joint Technology Initiatives (Annex IV-SP1)	-0.10
Nanosciences, Nanotechnologies, Materials and new Production Technologies	-0.23
CAPACITIES	**0.06**
Science in Society	0.27
Research Infrastructures	0.26
Coherent development of research policies	0.25
Activities of International Cooperation	0.23
Research Potential	0.15
Regions of Knowledge	0.06
Research for the benefit of SMEs	-0.25
PEOPLE	**0.27**
IDEAS	**0.23**

More homophilic — More heterogeneous

Homophily across sectoral research organizations

Programme	Homophily Indicator
FP6 (overall)	**0.14**
FP7 (overall)	**-0.11**
COOPERATION	**-0.11**
Socio-economic sciences and Humanities	0.29
Environment (including Climate Change)	0.23
Health	0.17
Space	0.17
Security	0.10
Food, Agriculture, and Biotechnology	0.09
Information and Communication Technologies	-0.06
Energy	-0.06
Transport (including Aeronautics)	-0.09
Joint Technology Initiatives (Annex IV-SP1)	-0.10
Nanosciences, Nanotechnologies, Materials and new Production Technologies	-0.23
CAPACITIES	**0.06**
Science in Society	0.27
Research Infrastructures	0.26
Coherent development of research policies	0.25
Activities of International Cooperation	0.23
Research Potential	0.15
Regions of Knowledge	0.06
Research for the benefit of SMEs	-0.25
PEOPLE	**0.27**
IDEAS	**0.23**

More homophilic — More heterogeneous

- Decrease from FP6 to FP7 – better integration of sectors
- Nanosciences, Nanotech., Materials and new Prod. Tech. is highly heterogeneous
- Research for the benefit of SMEs unsurprisingly score low for homophily

Other findings

- Regression between presence of SMEs and innovative outputs (measured by survey)
  - ↑ Knowledge transfer from research to market
  - ↑ Introduction of innovation in the form of new products or processes

- Other interesting findings from the study
  - >70% of projects involved 3+ sectors
  - >50% of survey respondents strengthened collaborations with researchers from other sectors (70% strengthened collaborations with same sector)
  - Share of project participation in the private sector 23% (FP6) -> 30% (FP7)
  - SMEs account for 54% (FP6) -> 60% (FP7)
  - High participation rate of SMEs in Capacities pillar (driven by thematic area Research for benefits of SMEs)
  - Highest participation rate in thematic Nanoscience, nanotech, mat. & new prod. tech. as well as in Joint technology initiatives

Summary

- Homophily indicator was developed
  - Advantages
    - Simple to compute
    - Applicable to large and disconnected networks
  - Improvement
    - Consider direct ties only
    - Model of expected frequency could be improved
- Applied to a set of real world data
  - Decrease in homophily FP6 to FP7
  - Thematic areas designed for cooperation
  - Presence of private companies significantly increases application of innovative ideas to market opportunity

Thank you !

Questions ?

Link to complete study: http://www.science-metrix.com/en/publications/reports#/en/publications/reports/study-on-network-analysis-of-the-7th-framework-programme-participation

Made possible by the contract 30-CE-0603443/00-49

Contact Information

Bastien St-Louis, Research Analyst

bastien.st-louislalonde@science-metrix.com

Telephone : 514-495-6505 ext. 136

Science-Metrix

Montréal – Ottawa – Washington

WEBSITE

www.science-metrix.com

PHONE

1.514.495.6505

1.800.994.4761