Digital Enterprise Research Institute

www.deri.org

Life-Cycles and Mutual Effects of Scientific Communities

Václav Belák, Marcel Karnstedt, and Conor Hayes

first.last@deri.org

Introduction

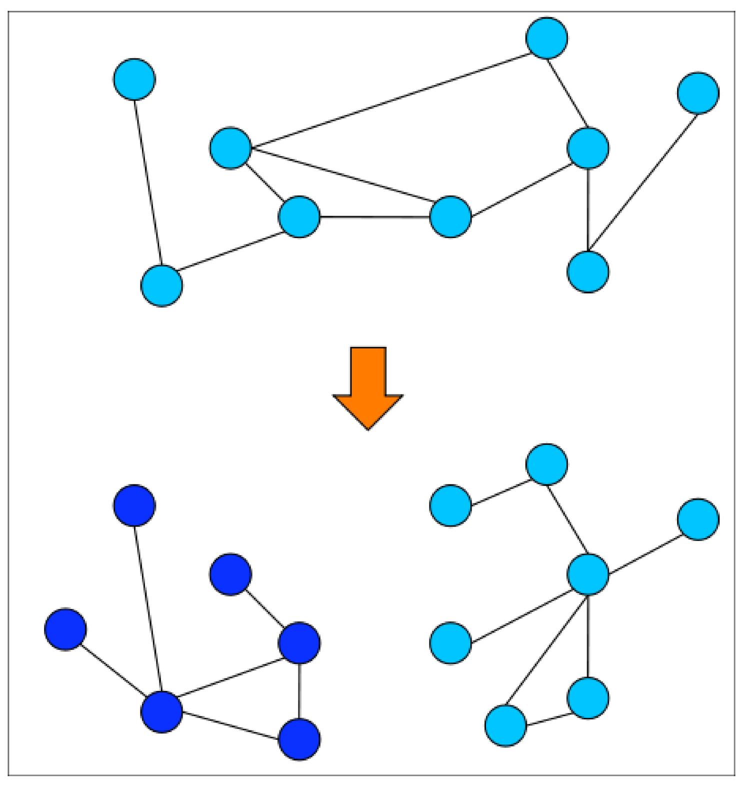
Assessment of scientific progress based on citation measures is static. Analysis and modelling of dynamic mutual effects of scientific communities promises to be a more viable form of evaluation of progress and impact of a research community.

We propose a general framework uniquely combining topological and content analysis with special visualization techniques to analyse and explain such dynamics.

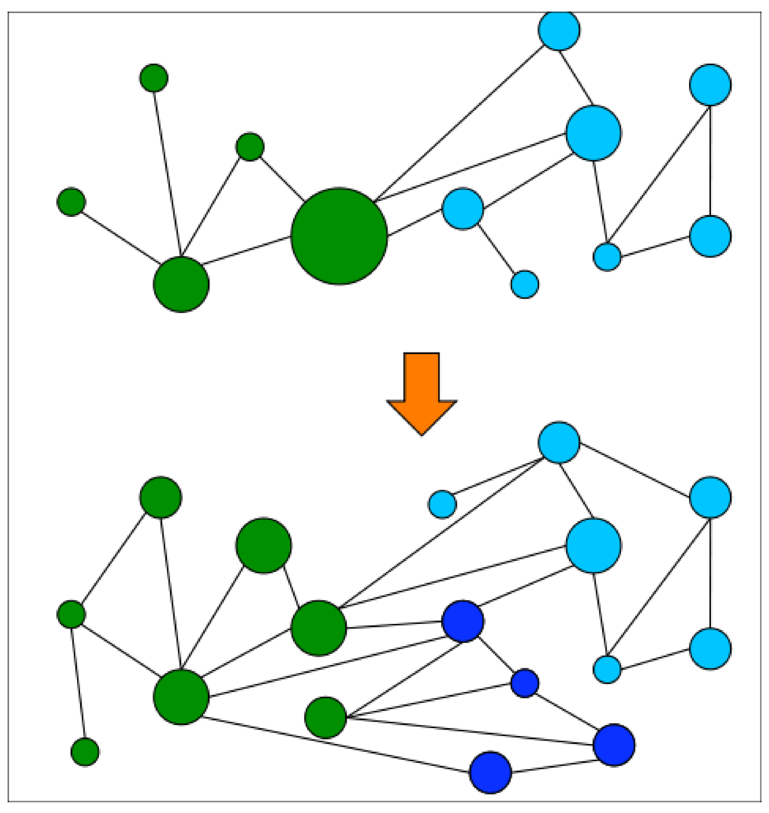
Aim

Inspired by Thomas Kuhn's work, we identified several interesting cross-community phenomena, which themselves we mined in an automated manner:

- Community shift and shift/merge
- Community specialization



Community shift *The upper part shows co-citation relations between different authors at a specific point in time. Over the time, a sub-community detaches from its original community (lower part of the figure). This means, authors from both communities are not cited together any more, with ongoing time the sub-community splits.*



Community shift/merge *Community merge is an opposite to the community shift. Even a combination of shift and merge is thinkable: over the time, from one large community only a sub-community approaches the originally separated one—whereby in parallel detaching from its original field of research.*

Data

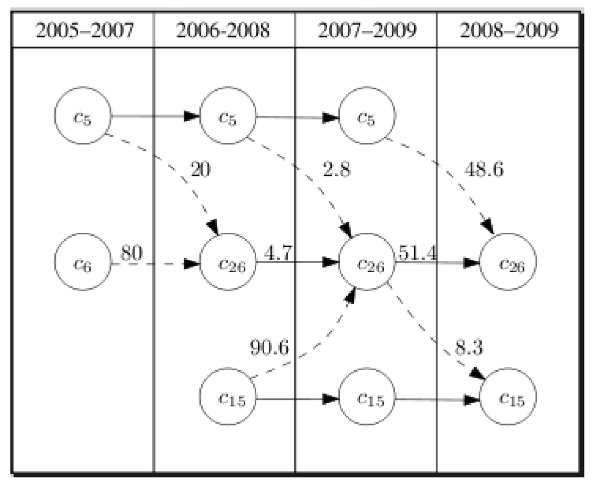
- 5772 authors and 817642 links in total
- 10 slices between 2000-2009
- Average slice had 2027 nodes and 81764.2 links
- Nearly 70% coverage by keywords

Methodology

- We extracted co-citation network from publications of two related disciplines: Semantic Web (SW) and Information Retrieval (IR)
- Overlapping time-slices of the network were generated from the network
- Communities were identified in each slice using Infomap and Louvain methods, but **any other can be used**
- Communities were matched according to the highest Jaccard coefficient
- Then, we detected important ancestors and descendants
- Each author was described by a ranked (TF-IAF) vector of keywords mined from the publications

Cross-community effects were analysed and detected using:

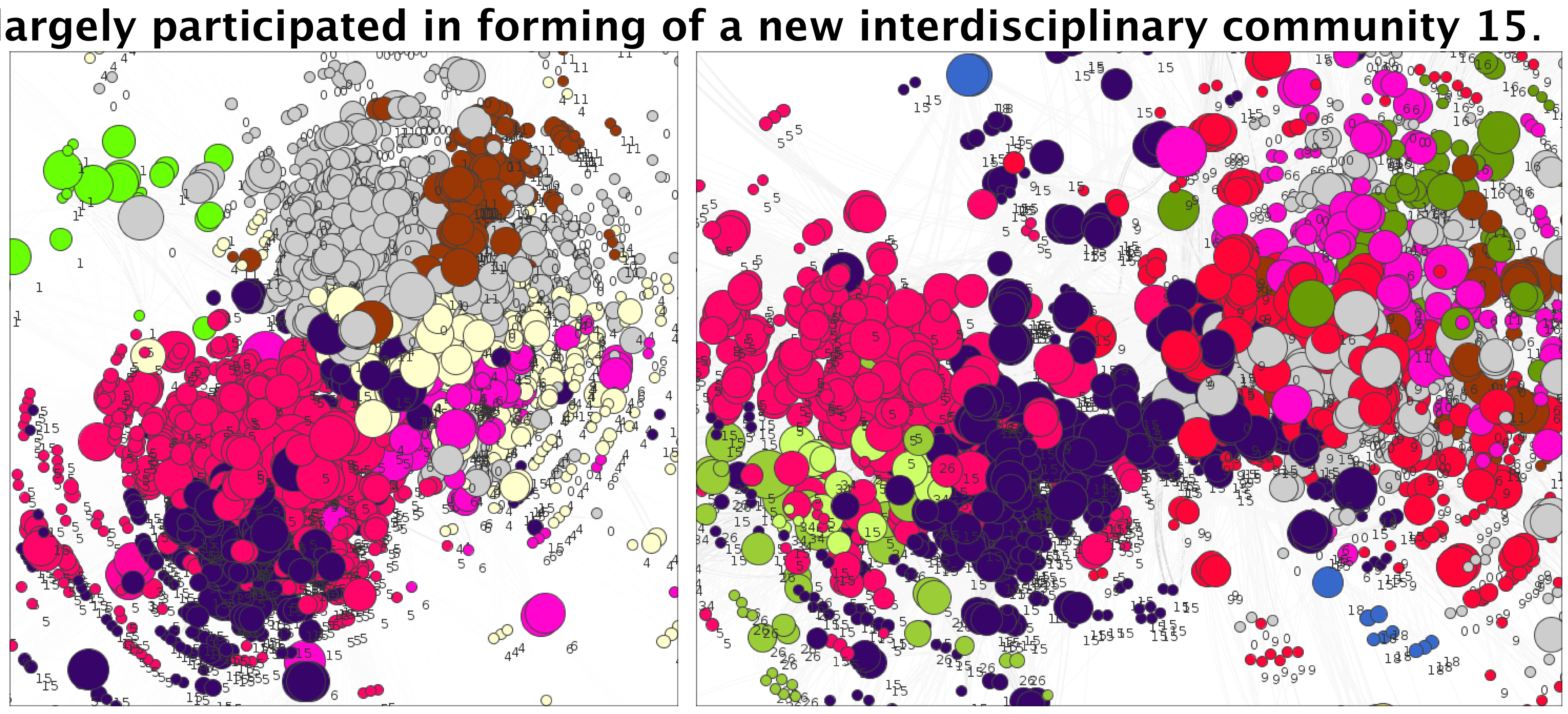
- *Community life-cycle measures*



- *Community topic evolution measures*

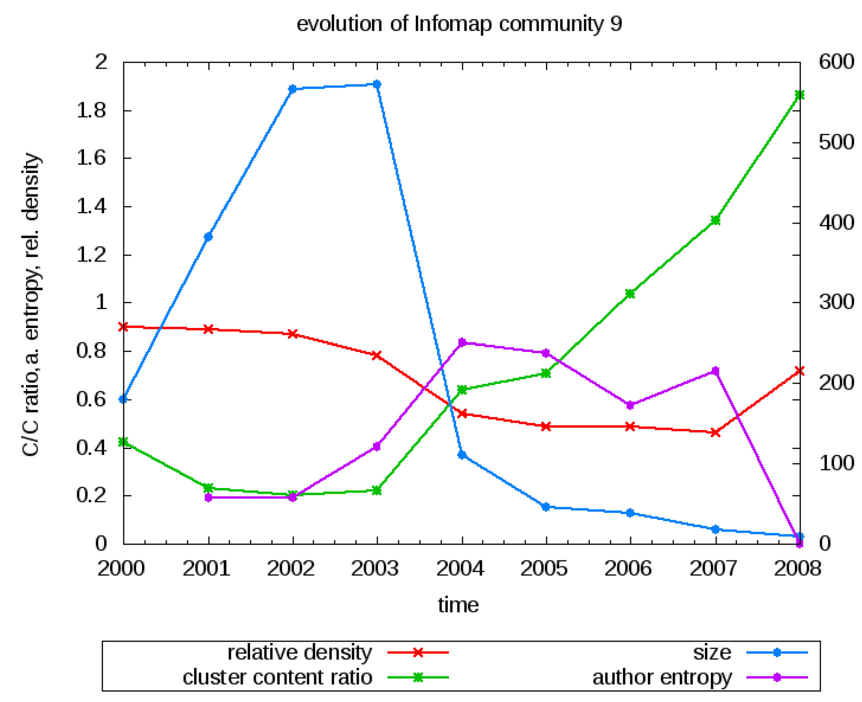
Results

Emergence of intermediary Louvain community 15 was first identified as a community shift from IR community 4. It then changed its topic under an influence of SW community 5, which kept its focus on core SW topics, while **it largely participated in forming of a new interdisciplinary community 15.**



Illustrations of community 5 "SW" (red—left side), "IR" communities 0, 4, 6 and 9 (grey, beige, pink and red—right side, respectively) and their intermediary community 15 (violet) in years 2004 (left) and 2007 (right).

Specialization of Infomap community 9 was characterized by a diminishing size with rise of topical cohesiveness (C/C ratio). Since 2002, when this development started, until 2004, we found 3 community shifts. Therefore, **the original big SW community started to specialize** while it produced couple of more tightly focused communities.



Conclusion

We developed a **general and scalable framework** for analysis of cross-community phenomena **uniquely combining *topological* and *content* analysis**. The developed measures proved to be useful in detection of phenomena like community shift.

Acknowledgements

Supported by the Science Foundation Ireland under Grant No. SFI/08/CE/I1380 and 08/SRC/I1407.