

Learning Layers

Scaling up Technologies for Informal Learning in SME Clusters

Social Activity versus Academic Activity : A Case Study of Computer Scientist on Twitter

Subhash Chandra Pujari¹, Asmelash Teka Hadgu², Elisabeth Lex¹, Robert Jäschke²

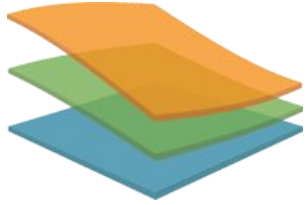
FWF



Austrian Science Fund: P 24866-N15

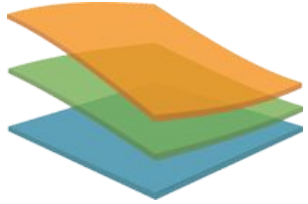
<http://Learning-Layers.eu> – Scaling up Technologies for Informal Learning in SME Clusters – layers@learning-layers.eu



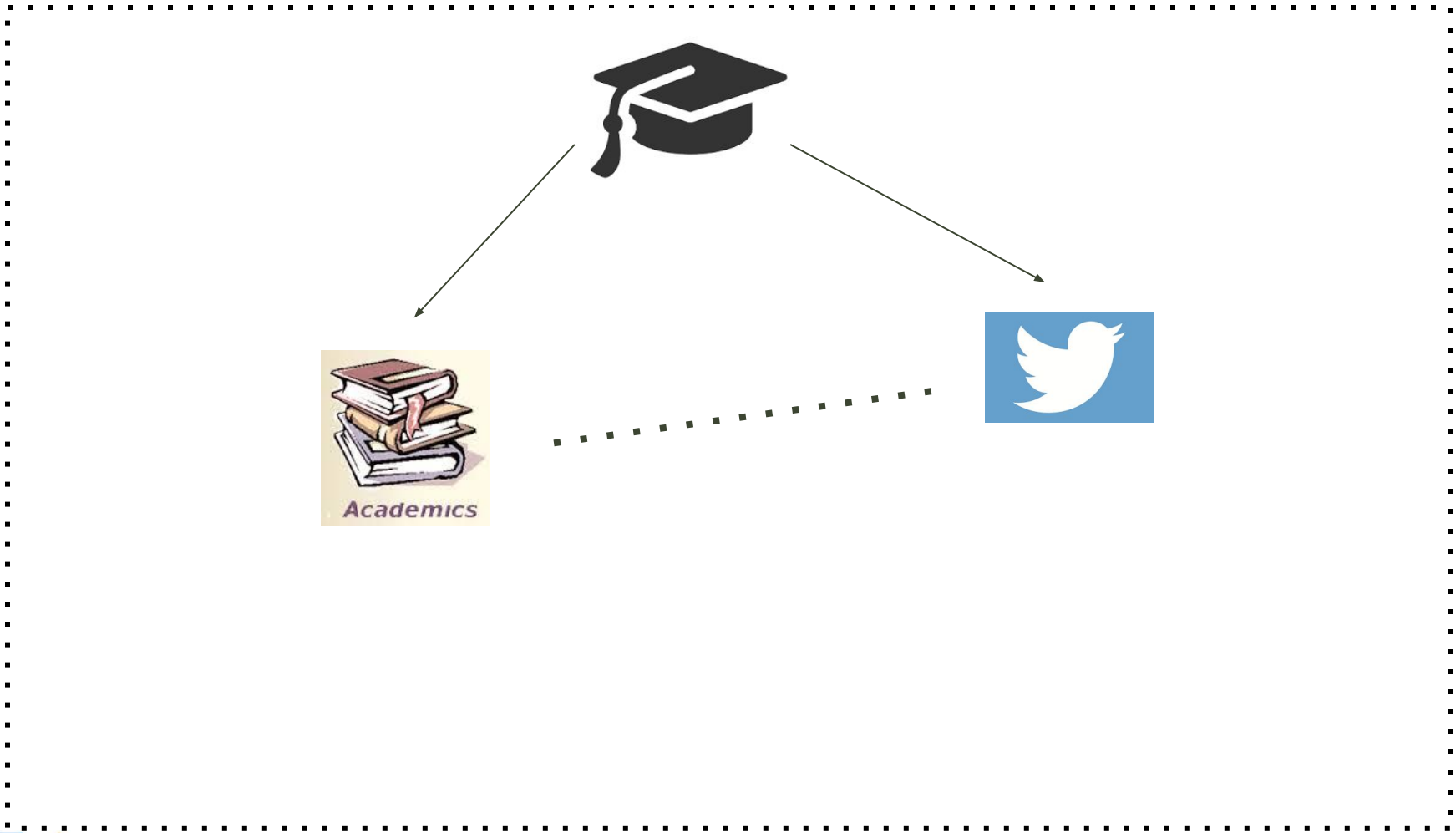


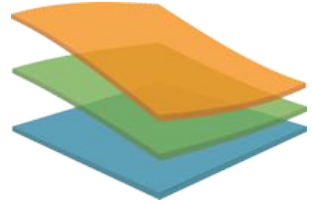
Agenda

- Introduction
- Research Questions
- Experimental Setup
- Dataset
- Results
- Conclusion



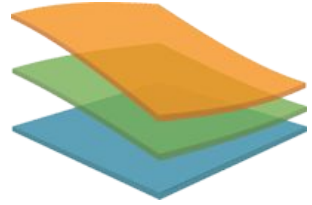
Researchers in Social Arena





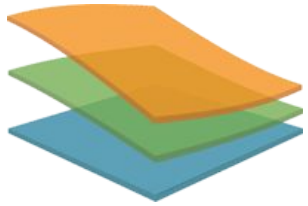
Formal and informal networks

- Formal networks (Academic)
 - Citation
 - Coauthor
- Informal networks (Twitter)
 - Follow (subscribing for information)
 - Retweet (endorsing the content)
 - Mention (content somehow relevant to person mentioned)

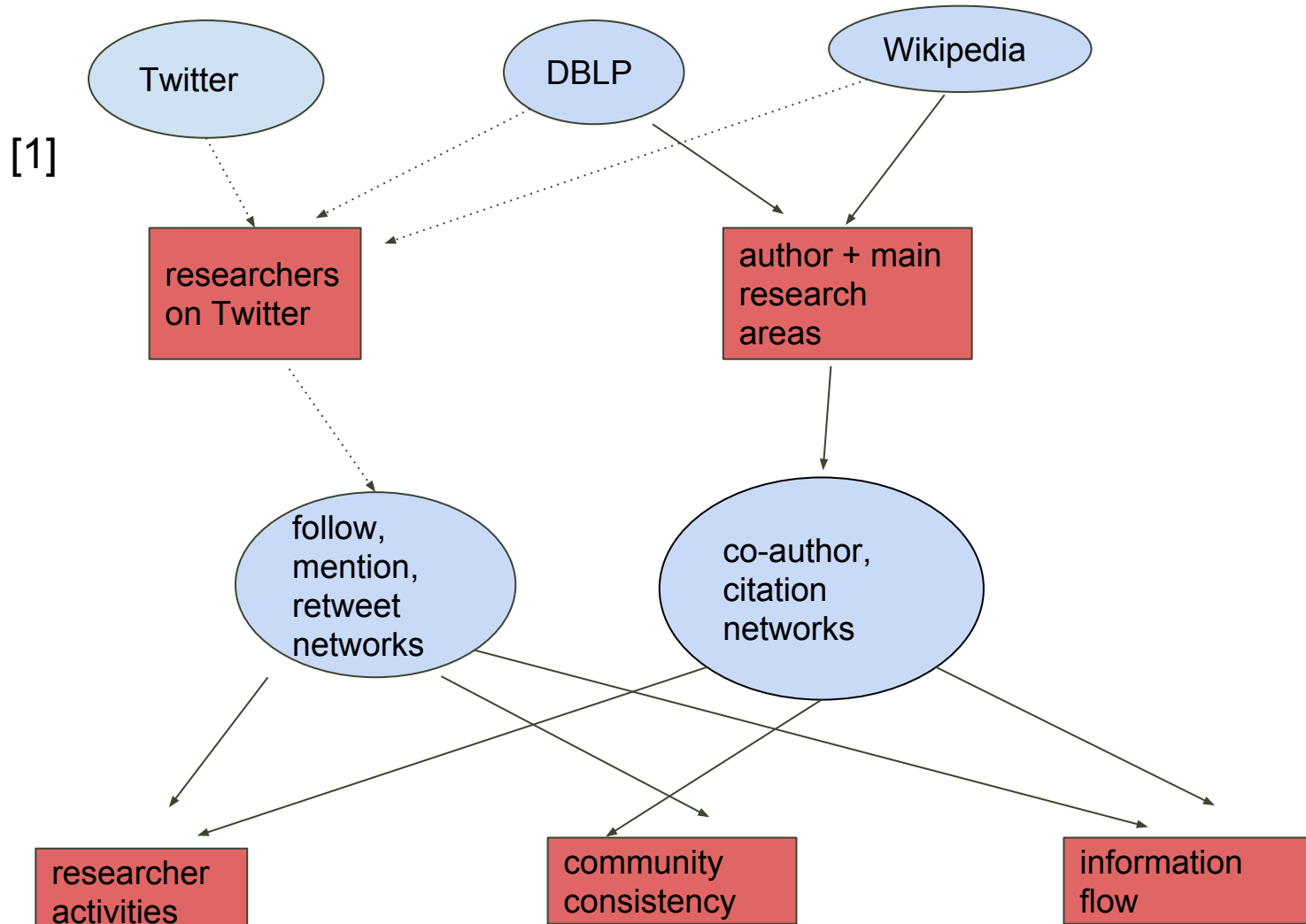


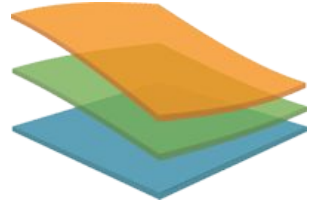
Research Questions

- RQ1: How related are activities and interaction in research and on Twitter?
- RQ2: How do researchers form communities and how consistent are those communities in different networks on Twitter and traditional academic networks?
- RQ3: Which patterns of information flow can be found between different areas of Computer Science?



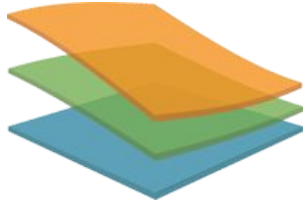
Experimental Setup





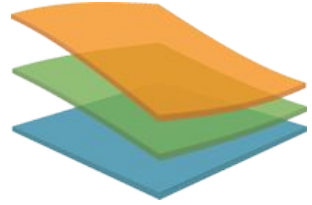
Dataset Collection

- **Wikipedia Conference List:** http://en.wikipedia.org/wiki/List_of_computer_science_conferences
- Search of conference handles
 - google search
 - twitter lookup for potential conference handles
 - 98 conferences, 170 Twitter handles
- DBLP to Twitter mapping:
 - 9191 users
 - 73% accuracy on manual evaluation of 150 users from 9191 users



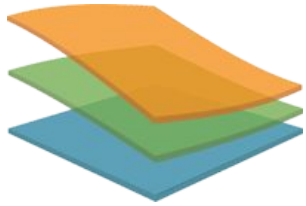
Dataset

Network	 V 	 E
follow	7,969	135,282
mention	6,030	73,357
retweet	5,050	48,592
citation	5,163	105,004
co-author	1,313,098	5,124,388



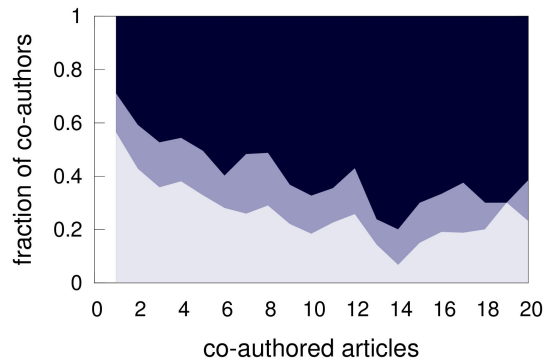
Research Question 1

How related are activities and interaction in research and on Twitter?

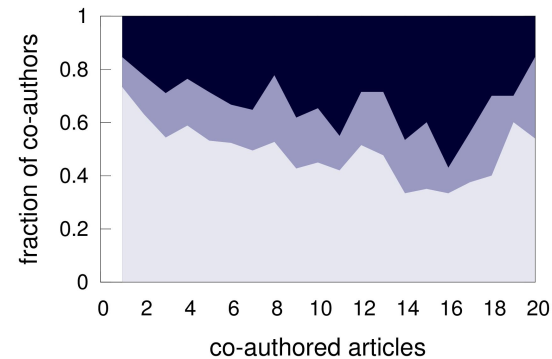


Comparison between academic and Twitter relations

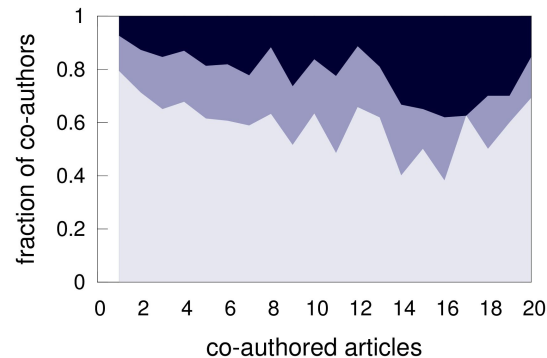
following



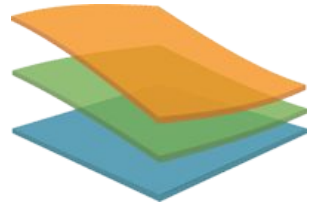
mentioning



retweeting

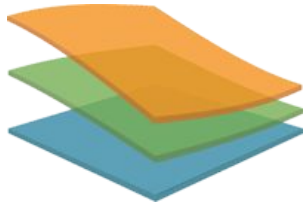


reciprocal ■
one-way ■
none ■



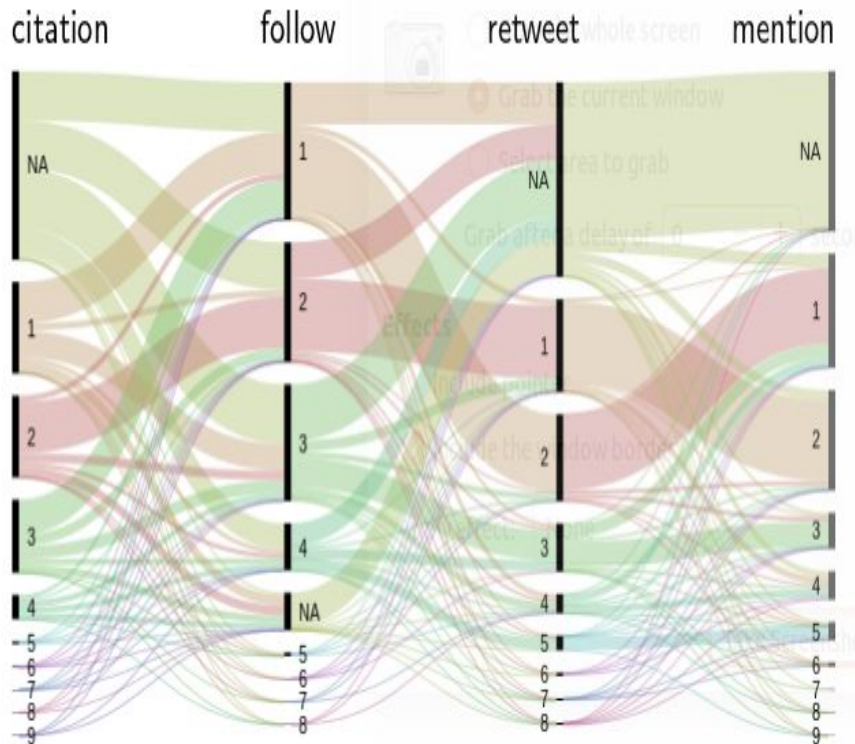
Research Question 2

How do researchers form communities and how consistent are those communities in different networks on Twitter and traditional academic networks?

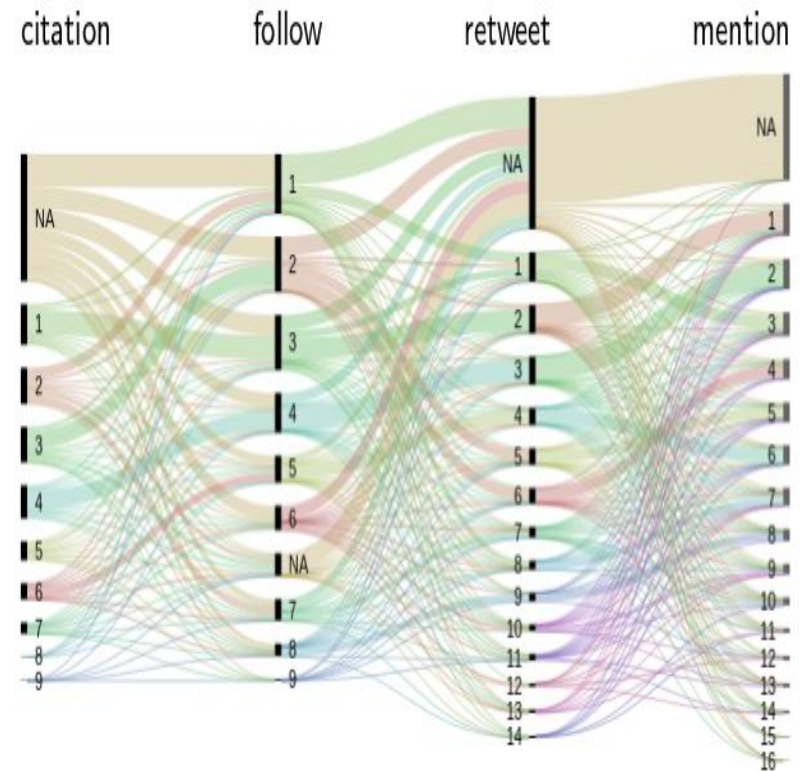


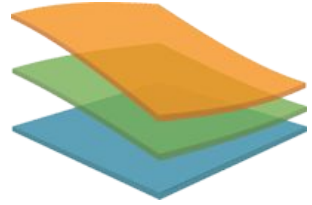
Overlap between communities in different Networks

CNM [3]



Louvian [2]

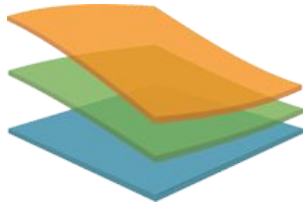




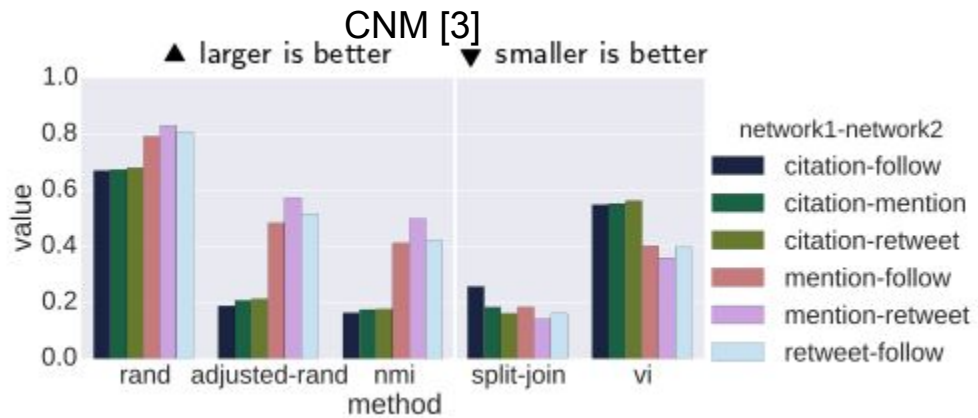
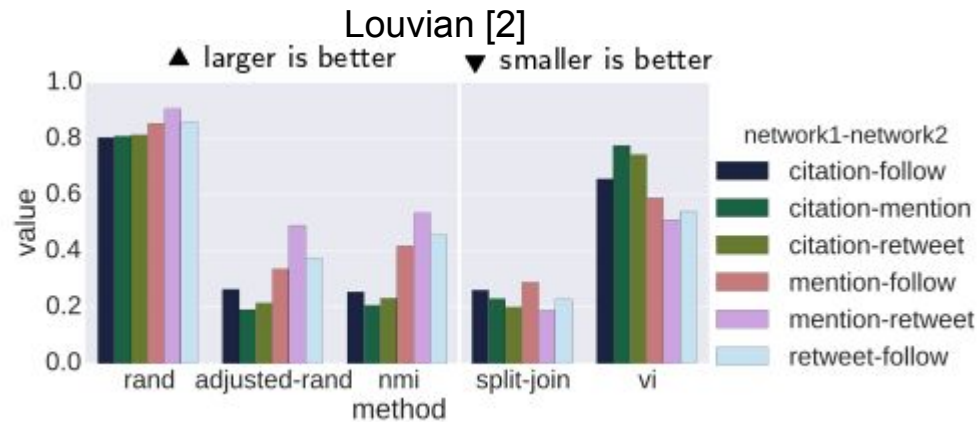
Community consistency scores

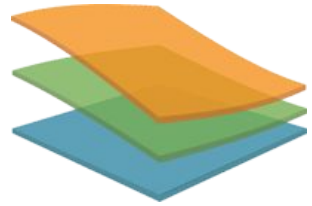
1. Larger values means communities are more similar
 - a. the Rand index (rand) [4]
 - b. the adjusted Rand index (adjusted-rand) [5]
 - c. normalized mutual information (nmi) [6]

2. Smaller value means communities are more similar
 - a. the split-join score (split-join) [7]
 - b. variation of information (vi) [8]



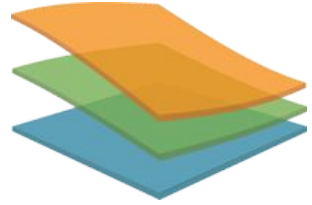
Comparing community consistency score





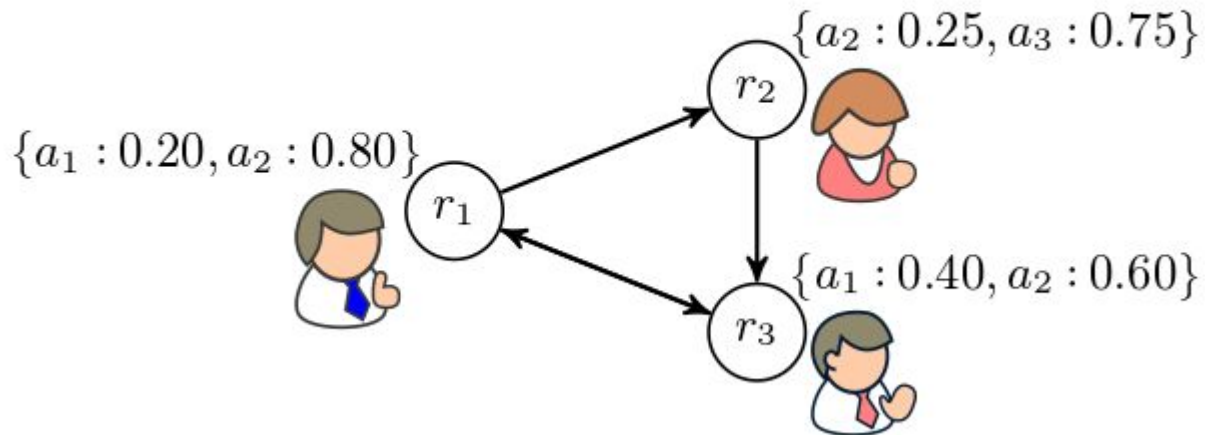
Research Question 3

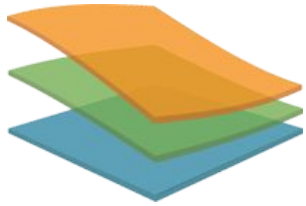
Which patterns of information flow can be found between different areas of Computer Science?



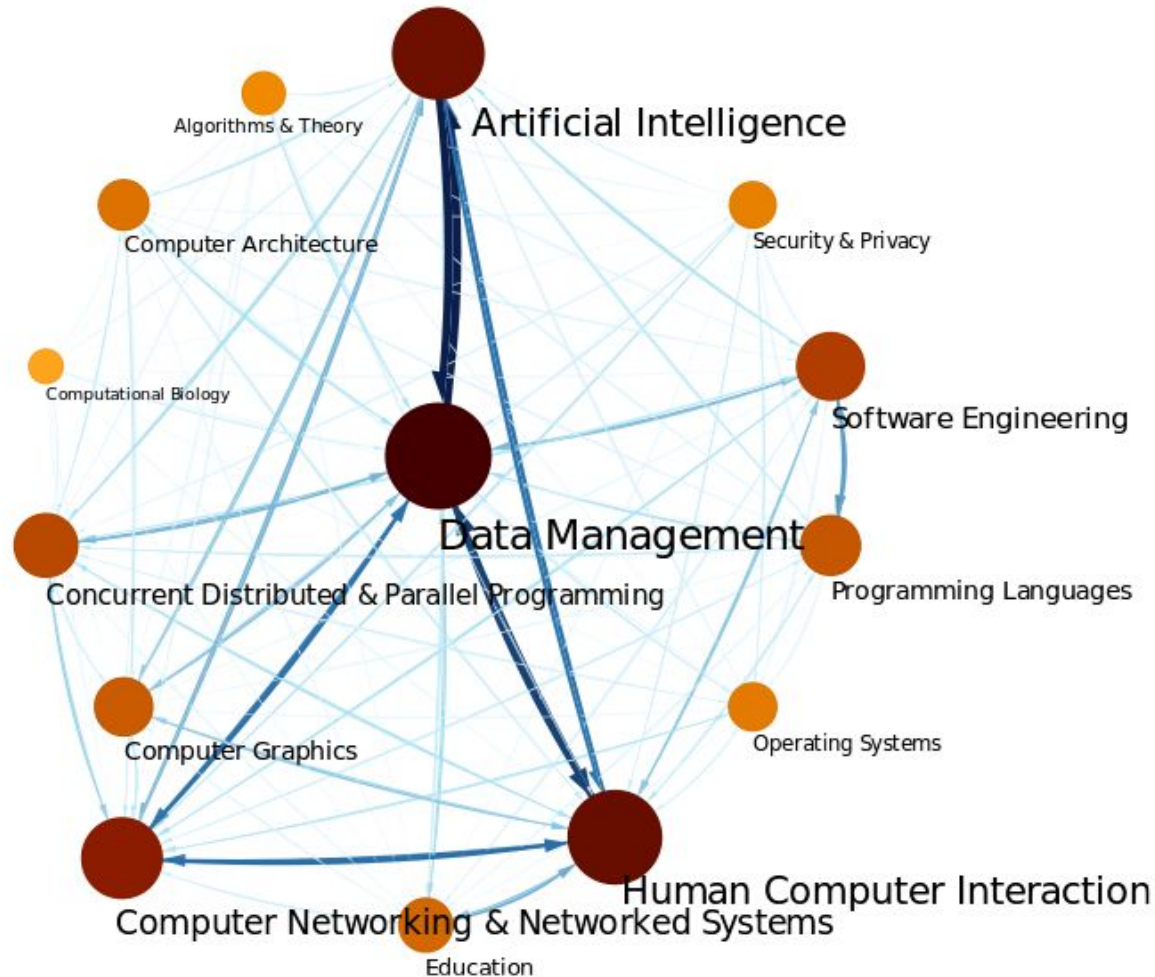
Computing Area Score

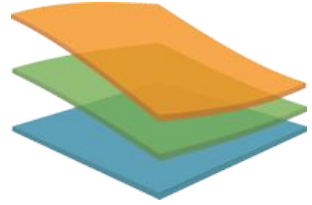
- Users are mapped to area of research with soft assignment score based on fraction of publication in an area.





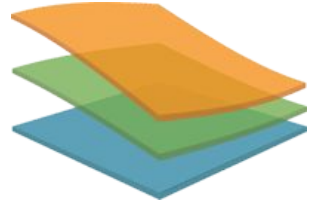
Flow of information between areas of research





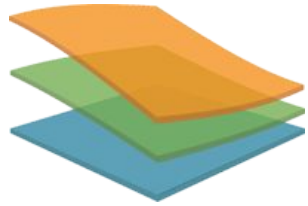
Take Away Messages

- Researchers activity on Twitter is weakly correlated to their activity in academics.
- The structure of network on Twitter is more similar to each other than academic networks. Also, dynamic networks like mention and retweet network are more similar to each other than follow network.
- Researcher move out from their core area of research to connect with other researchers.



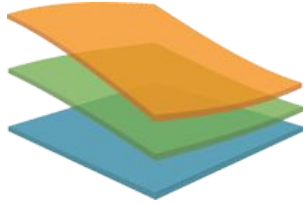
Future Work

- Studying change in structure of network and behaviour of researchers on Twitter with time.
- Understanding flow of information on Twitter between different areas of science.
- Studying patterns and temporal changes in content dissemination over network of computer scientists on Twitter.



References

1. A. T. Hadgu and R. Jäschke. Identifying and analyzing researchers on Twitter. In Proc. Web Science, pages 23–32. ACM, 2014.
2. V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, 2008.
3. A. Clauset, M. E. J. Newman, and C. Moore. Finding community structure in very large networks. *Physical Review E*, 70:066111, 2004.
4. W. Rand. Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*, 66(336):846–850, 1971.
5. L. Hubert and P. Arabie. Comparing partitions. *Journal of Classification*, 2(1):193–218, 1985.
6. L. Danon, A. Díaz-Guilera, J. Duch, and A. Arenas. Comparing community structure identification. *J. Stat. Mech.*, 9:8, 2005.
7. S. Dongen. Performance criteria for graph clustering and markov cluster experiments. Technical report, CWI (Centre for Mathematics and Computer Science), Amsterdam, The Netherlands, 2000.
8. M. Meilă. Comparing clusterings by the variation of information. In *Learning Theory and Kernel Machines*, pages 173–187. Springer, 2003.



Thank you for your attention!

Questions?

Subhash Chandra Pujari

spujari@know-center.at

Knowledge Technologies Institute, KTI

Graz University of Technology

Social Computing Team, Know-Center (Austria)