Topic Affinity Analysis for an Astronomy and Astrophysics Data Set

[1]Theresa Velden, Shiyan Yan, and Carl Lagoze

tvelden@umich.edu, shiyansi@umich.edu, clagoze@umich.edu
[1]School of Information, University of Michigan, 105 S. State Street, Ann Arbor, MI 48109 (USA)

Abstract
In this paper we map the affinity between topics extracted from a body of literature published in Astronomy and Astrophysics journals between 2003-2010. The topics are extracted using the popular information theoretic Infomap clustering algorithm (Rosvall & Bergstrom, 2008) iteratively on the giant component of the direct citation network constructed from the data. The affinity network shows what topics are disproportionally well connected (by citations) to other topics. The topology of the network highlights a large division into astrophysics versus astronomically oriented publications. Bridging between those two domains is a population of smaller topics. Going forward, we plan to create and analyze topic affinity network maps for alternative solutions to the topic extraction challenge on that same data set that are produced by our colleagues and that will be discussed and compared at the proposed special session on 'Same data? Different results? The performative nature of algorithms for topic detection in science' at ISSI 2015. We expect that topic affinity mappings will help to examine the nature of differences between different topic extraction solutions.

Conference Topic
Methods and techniques (special session on algorithms for topic detection)

Introduction
The mapping of research topics and collaborative ties in scientific research fields (Morris 2008) is flourishing for a number of reasons. Increasingly, scholarly publications and their metadata are available from a variety of sources (digital libraries, institutional and disciplinary repositories, along with bibliographic abstracting services such as the long established Web of Knowledge and more recently, Scopus). Complementing this is the emergence of sophisticated algorithms for the analysis of complex networks (Newman 2003b) and the wide availability of advanced user-friendly network analysis and visualization tools like pajek, gephi, or VOS Viewer.

However, many different algorithms for community extraction and topic detection exist and offer different suggestions what the most prominent groupings of publications or authors may be. The special session at ISSI 2015 sets out to systematically compare and evaluate the origin, extent, and implication of differences between topic extraction methods. In this paper we describe the results of our approach to topic detection and topic affinity analysis to the shared 'astronomy and astrophysics' data set. This approach has emerged from research program on studying behavioral patterns in scientific communities and comparing them across fields, and may help to shed light on the nature of differences between topic extraction solutions.

Background
As described in (Velden 2009), we take a mixed method approach to studying field-specific practices and cultures of scientific communities, integrating ethnographic field studies with network analytic methods. The network analytic method we apply here to the 'astronomy and astrophysics' data set is part of an ongoing effort to combine network analytic with ethnographic methods (Velden, Haque & Lagoze, 2010; Velden, 2013). This evolves a tradition of close-up analysis of scientific networks and communication practices started by Crane's work (1972) on invisible colleges and taken up more recently by Zuccala (2006).

Scientific research specialties are a complex social and cognitive phenomenon. Sociologically, they can be characterized as collective production communities that emerge from the indirectly coordinated activity of autonomous actors (research groups) who aim to contribute to a shared knowledge base (Gläser, 2006; Velden, 2013). Therefore, the combined analysis of social and cognitive structures is of particular interest (Ding, 2011). In our work we achieve this in two steps: first by algorithmically extracting major research topics in a research specialty from the direct citation network and generating an affinity network that shows what topics are disproportionally well connected through citations to other topics. In a second step, we overlay the topic information on the group collaboration network (Velden, Haque & Lagoze, 2010) extracted from the co-author network of the research specialty. The resulting maps show how collaborative ties connect groups active in a particular topic area. This paper reports work in progress. At this point, we have produced and analyzed the topic affinity network. Producing the overlay with the group collaboration network will be one of the next steps.

Method

Our approach to topic extraction and topic affinity analysis is discussed in detail in Velden (2013). Below we briefly review the relevant details for the analysis reported in this paper.

Data

The data set used in this study includes papers published 2003-2010 in 59 astrophysical journals indexed by Web of Science. By accepting only documents of type 'Article', 'Letter', and 'Proceedings Paper', the data set comprised the bibliographic data of 111,616 publications.

Network construction

Various citation-based approaches have been used in the past to detect topics in research fields. These include bibliographic coupling, co-citation and direct citation, including or excluding citation environments. The advantages and disadvantages of these approaches have been discussed in Boyack (2010). We base our topic extraction on the direct citation network.

Clustering

We use the Infomap clustering algorithm (Rosvall & Bergstrom, 2008) twice to iteratively extract clusters of clusters of documents. The repeated clustering is necessary to obtain sufficiently large entities (topics) for further visual inspection and analysis. In the resulting topic network, nodes represent clusters of publications based on the direct citation links between them.

Topic affinity network

We evaluate the strength of citation links between topic areas relative to a null model that assumes a random distribution of citation links proportional to topic area sizes. Hence, the existence of a link between topics in the affinity indicates a surplus of connectivity between the two topic areas in question, whereas the absence of a link may either mean 'normal' (random) background connectivity or a negative affinity value ('antagonism').

The affinity between a source topic area and a target topic area is calculated as shown in Figure 1 below.

Assume:
A_{11-i}: Top 11 Areas expect area i
$N_{p(j)}$: Number of papers in topic area j
C_{ij}: Number of Citation from topic area i to topic area j

We define the citation based affinity A between two topic areas i and j as the residual:

$$A_{ij} = \frac{\text{Actual Count}_{ij} - \text{Expected Count}_{ij}}{\sqrt{\text{Expected Count}_{ij}}}$$

where:

$$\text{Actual Count}_{ij} = C_{ij}$$

$$\text{Expected Count}_{ij} = \frac{N_{p(j)}}{\sum_{k \in A_{11-i}} N_{p(k)}} \times \left(\sum_{k \in A_{11-i}} C_{ik}\right)$$

Figure 1. Affinity between a source topic area and a target topic

Topic affinity as defined here is a relative property. It expresses the relative preference for documents in one topic area to cite documents in another area given the choice of topic areas included in the data set and in the affinity calculation. Theoretically, the relative affinity to document clusters outside the set of topic areas selected for this analysis or even outside of the data set (external citations) could be greater than to the ones in the set.

Topic Labeling

To support the interpretation of the resulting topic affinity network, we use a semi-automatic approach to labeling topic areas. To this end, we analyze the frequency of journals that the documents in each topic area are published in. Using a measure based on the concept of *term frequency - inverse document frequency (tf-idf)* to combine popularity with distinctiveness of a journal title within the data set, we produce a ranked list of the 15 most popular journals in each topic area. From those journal titles we then derive labels that typically reflect sub disciplinary orientation of topic areas. A more detailed and specific identification of topic area content either algorithmically or through expert evaluation or would be desirable.

Results

The topic extraction from the giant component of the direct citation network results in 22 document clusters ('topics'). For pragmatic reasons, to support interpretation of the visualized network, we include only the largest eleven topic areas in the affinity network. Given the uneven size distribution of clusters (Fig. 1), these largest clusters account for the large majority of publications in the giant component of the direct citation network, namely 84% (see Table 1 for details on the sizes of various network components).

Table 1. Properties of direct citation network.

	# of nodes (documents)	**% of network**	**% of giant component**
entire network	111,616	100	N.A.
giant component	101,831	91.2	100
11 largest topic areas	85,562	84.0	76.7

The topic affinity network for the largest 11 document clusters is shown in Figure 2. The most striking topological feature regards the relationship between the three largest topics. Notably,

topic 3 (Astronomy/Solar System) is not directly connected with the other two topics, topic 1 (Astronomy/Astrophysics) and topic 2 (Gravitational Physics, Cosmology). Topic 2 has a strong directed link to topic 1, indicating that it borrows disproportionally from the literature in topic 2. Topics 1 and 3 are indirectly linked, via small, astronomically oriented 'proxy topics', essentially topics 7 and 9, and to l lesser degree topics 10 and 11. However, there exists only a very faint indirect affinity link between topic 2 and topic 3, via topic 11.

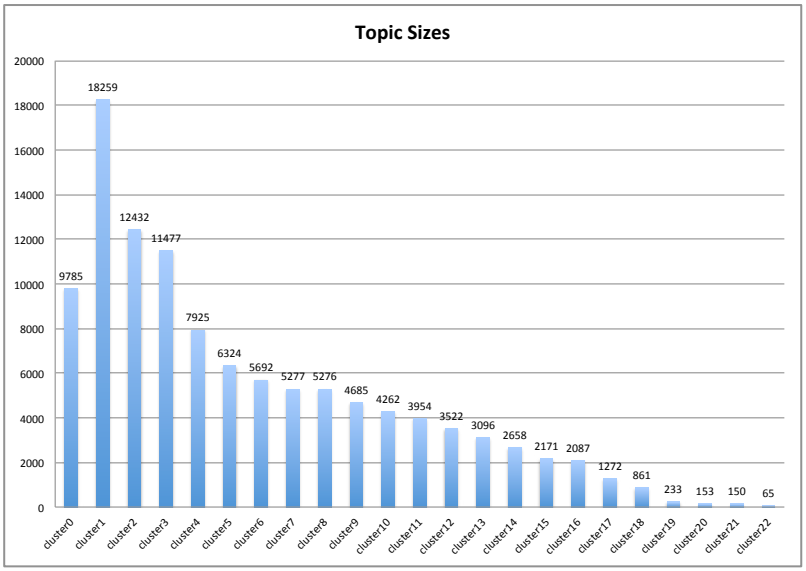

Figure 1. Sizes of the 22 document clusters ('topics') that constitute the giant component of the direct citation network. Cluster '0' shows the number of documents not included in the giant.

Discussion

Based on our own, if limited, expertise in this larger domain of research, we would offer the following speculations about the interpretation of the tripartite structure of the current 2003-2010 literature in the astronomy and astrophysics data set that is suggested by the topology of the affinity network in figure 2. The literature is subdivided into three large domains, with distinct research focus, namely astrophysics - the quest for developing a theoretical understanding of physical and chemical properties of celestial bodies (topic 1), gravitational physics - the quest for understanding the workings of gravitational forces in the universe (topic 2), and planetary science - the quest for understanding the composition, dynamics and history of planets and solar systems (topic 3). As reflected by the affinity network, in the 2003-2010 period, the three domains rely to varying degrees on astronomical observation; this is least the case for gravitational physics. An interesting open question is to what degree the observational astronomy literature has been integrated through citations into these larger topics rather than being identifiable as separate topics. The topic affinity network further underlines that whereas there are strong connections between astrophysics and gravitational physics (such as the role of gravitational forces in the formation of black holes and the puzzle of the nature of black matter), the cognitive links between gravitational physics and planetary science are weak.

Table 2. Ranking of the 15 most popular journals in each topic. This list of journal titles is used to help identify the subject matter of a topic in terms of its subdisciplinary orientation.

Journal titles	# of publications	tf*idf score
Area1		
ASTRONOMICAL JOURNAL	1098	0.104672985
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY	4415	0.091614001
ASTROPHYSICAL JOURNAL SUPPLEMENT SERIES	401	0.06435346
ASTRONOMISCHE NACHRICHTEN	314	0.062939775
PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF AUSTRALIA	116	0.056289489
NEW ASTRONOMY REVIEWS	347	0.043675217
PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC	152	0.037652069
ASTRONOMY REPORTS	164	0.032873003
CHINESE JOURNAL OF ASTRONOMY AND ASTROPHYSICS	171	0.027442498
PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF JAPAN	284	0.027073887
ASTROPHYSICAL JOURNAL LETTERS	510	0.022086996
PHYSICAL REVIEW D	164	0.020641889
ASTROPHYSICS AND SPACE SCIENCE	290	0.006017681
ASTROPHYSICAL JOURNAL	5565	0
ASTRONOMY & ASTROPHYSICS	3148	0
Area2		
PHYSICAL REVIEW D	5616	0.700439718
JOURNAL OF COSMOLOGY AND ASTROPARTICLE PHYSICS	1416	0.533389555
CLASSICAL AND QUANTUM GRAVITY	1533	0.376292436
GENERAL RELATIVITY AND GRAVITATION	543	0.204541334
INTERNATIONAL JOURNAL OF MODERN PHYSICS D	655	0.081693023
GRAVITATION & COSMOLOGY	75	0.036063565
ASTROPARTICLE PHYSICS	78	0.023617218
NEW ASTRONOMY	46	0.017327627
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY	783	0.016100189
NEW ASTRONOMY REVIEWS	122	0.015216105
ASTROPHYSICAL JOURNAL SUPPLEMENT SERIES	49	0.007792228
ASTROPHYSICS AND SPACE SCIENCE	286	0.005880784
ASTROPHYSICAL JOURNAL LETTERS	40	0.001716582
ASTROPHYSICAL JOURNAL	506	0
ASTRONOMY & ASTROPHYSICS	325	0
Area3		
PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC	364	0.160723328
ICARUS	150	0.129745662
ASTRONOMISCHE NACHRICHTEN	361	0.128983753
ASTRONOMICAL JOURNAL	732	0.124387179
NEW ASTRONOMY	107	0.072503543
ASTROPHYSICS	89	0.060306686
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY	1461	0.054039787
ASTRONOMY REPORTS	108	0.038587937
ASTROPHYSICAL JOURNAL SUPPLEMENT SERIES	111	0.031752855
ASTROPHYSICAL JOURNAL LETTERS	318	0.024548551
PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF JAPAN	127	0.021580836
NEW ASTRONOMY REVIEWS	85	0.019070268
ASTROPHYSICS AND SPACE SCIENCE	385	0.014240464
ASTROPHYSICAL JOURNAL	2773	0
ASTRONOMY & ASTROPHYSICS	3122	0
Area4		
SOLAR PHYSICS	1248	2.133094119
ANNALES GEOPHYSICAE	228	0.222784668
ADVANCES IN SPACE RESEARCH	372	0.153453831
GEOPHYSICAL AND ASTROPHYSICAL FLUID DYNAMICS	77	0.131609172
ASTRONOMISCHE NACHRICHTEN	187	0.096348379
SPACE SCIENCE REVIEWS	96	0.093804071
ASTRONOMY REPORTS	119	0.061312605
ASTROPHYSICAL JOURNAL LETTERS	333	0.037069606
CHINESE JOURNAL OF ASTRONOMY AND ASTROPHYSICS	77	0.031763293
ASTRONOMY LETTERS-A JOURNAL OF ASTRONOMY AND SPACE ASTROPHYSICS	95	0.03073523
PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF JAPAN	102	0.024994227
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY	189	0.010080921
ASTROPHYSICS AND SPACE SCIENCE	75	0.004000365
ASTROPHYSICAL JOURNAL	2165	0
ASTRONOMY & ASTROPHYSICS	1609	0
Area5		
ICARUS	2102	2.700439718
PLANETARY AND SPACE SCIENCE	850	1.091995129
ASTROBIOLOGY	258	0.454192886
EARTH MOON AND PLANETS	257	0.330167939
CELESTIAL MECHANICS & DYNAMICAL ASTRONOMY	170	0.299274383
SOLAR SYSTEM RESEARCH	167	0.29399307
SPACE SCIENCE REVIEWS	115	0.115737336
ADVANCES IN SPACE RESEARCH	263	0.111741818
ANNALES GEOPHYSICAE	104	0.104666808
ASTRONOMICAL JOURNAL	231	0.058301094
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY	219	0.012031166
PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF JAPAN	38	0.009590656
ASTROPHYSICAL JOURNAL LETTERS	72	0.008255273
ASTROPHYSICAL JOURNAL	286	0
ASTRONOMY & ASTROPHYSICS	598	0
Area6		
PHYSICAL REVIEW D	4101	0.700439718
JOURNAL OF COSMOLOGY AND ASTROPARTICLE PHYSICS	353	0.182093016
ASTROPARTICLE PHYSICS	430	0.178295313
CLASSICAL AND QUANTUM GRAVITY	33	0.011092632
ADVANCES IN SPACE RESEARCH	45	0.00979976
ASTROPHYSICS	16	0.008253508
NEW ASTRONOMY REVIEWS	45	0.007685878
INTERNATIONAL JOURNAL OF MODERN PHYSICS D	45	0.007685878
COMPTES RENDUS PHYSIQUE	16	0.006634244

Journal titles	# of publications	tf*idf score
Area 6 (contd)		
ASTRONOMY LETTERS-A JOURNAL OF ASTRONOMY AND SPACE ASTROPHYSICS	15	0.002561959
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY	69	0.001942923
ASTROPHYSICAL JOURNAL LETTERS	29	0.001704276
ASTROPHYSICS AND SPACE SCIENCE	25	0.000703958
ASTROPHYSICAL JOURNAL	241	0
ASTRONOMY & ASTROPHYSICS	107	0
Area7		
BALTIC ASTRONOMY	64	0.093118611
REVISTA MEXICANA DE ASTRONOMIA Y ASTROFISICA	39	0.077757085
ASTROPHYSICAL JOURNAL SUPPLEMENT SERIES	131	0.063035619
ASTRONOMICAL JOURNAL	218	0.062312612
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY	686	0.042681773
ASTRONOMY REPORTS	65	0.039065743
PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC	45	0.033422965
SPACE SCIENCE REVIEWS	26	0.029634918
PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF JAPAN	90	0.02572539
ASTROPHYSICAL JOURNAL LETTERS	160	0.02077656
ASTRONOMY LETTERS-A JOURNAL OF ASTRONOMY AND SPACE ASTROPHYSICS	36	0.013586115
CHINESE JOURNAL OF ASTRONOMY AND ASTROPHYSICS	23	0.011067322
ASTROPHYSICS AND SPACE SCIENCE	176	0.010950426
ASTROPHYSICAL JOURNAL	1856	0
ASTRONOMY & ASTROPHYSICS	1359	0
Area8		
PHYSICAL REVIEW D	5208	0.700439718
INTERNATIONAL JOURNAL OF MODERN PHYSICS D	31	0.004169284
CLASSICAL AND QUANTUM GRAVITY	8	0.002117529
JOURNAL OF COSMOLOGY AND ASTROPARTICLE PHYSICS	5	0.002030988
GENERAL RELATIVITY AND GRAVITATION	3	0.001218593
ASTROPHYSICS	3	0.001218593
NUOVO CIMENTO DELLA SOCIETA ITALIANA DI FISICA C-GEOPHYSICS AND SPACE PHYSICS	3	0.001218593
COMPTES RENDUS PHYSIQUE	3	0.000979516
ASTROPARTICLE PHYSICS	2	0.000653011
GRAVITATION & COSMOLOGY	1	0.000518518
CHINESE JOURNAL OF ASTRONOMY AND ASTROPHYSICS	3	0.00051445
NEW ASTRONOMY REVIEWS	1	0.000134493
ASTROPHYSICS AND SPACE SCIENCE	2	4.43E-05
ASTRONOMY & ASTROPHYSICS	2	0
ASTROPHYSICAL JOURNAL	1	0
Area9		
ASTRONOMICAL JOURNAL	571	0.282314914
PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF AUSTRALIA	86	0.216437853
ACTA ASTRONOMICA	50	0.172434283
PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC	104	0.133611565
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY	909	0.097827391
NEW ASTRONOMY	48	0.094634582
ASTRONOMISCHE NACHRICHTEN	79	0.0821274
ASTRONOMY REPORTS	43	0.044702256
ASTRONOMY LETTERS-A JOURNAL OF ASTRONOMY AND SPACE ASTROPHYSICS	58	0.037861606
ASTROPHYSICAL JOURNAL SUPPLEMENT SERIES	45	0.037454628
ASTROPHYSICAL JOURNAL LETTERS	159	0.035713223
PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF JAPAN	42	0.020765721
ASTROPHYSICS AND SPACE SCIENCE	81	0.008717292
ASTROPHYSICAL JOURNAL	1073	0
ASTRONOMY & ASTROPHYSICS	1051	0
Area10		
PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF JAPAN	217	0.086427698
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY	783	0.067881878
CHINESE JOURNAL OF ASTRONOMY AND ASTROPHYSICS	82	0.054979695
ASTRONOMISCHE NACHRICHTEN	65	0.054433948
ADVANCES IN SPACE RESEARCH	72	0.048274854
ASTRONOMY LETTERS-A JOURNAL OF ASTRONOMY AND SPACE ASTROPHYSICS	64	0.033654761
ASTROPHYSICAL JOURNAL SUPPLEMENT SERIES	49	0.03285372
NEW ASTRONOMY REVIEWS	58	0.030499627
PHYSICAL REVIEW D	49	0.025766927
INTERNATIONAL JOURNAL OF MODERN PHYSICS D	49	0.025766927
ASTROPHYSICAL JOURNAL LETTERS	135	0.024426497
ASTRONOMICAL JOURNAL	50	0.019914216
ASTROPHYSICS AND SPACE SCIENCE	106	0.009189628
ASTROPHYSICAL JOURNAL	1332	0
ASTRONOMY & ASTROPHYSICS	897	0
Area11		
NUOVO CIMENTO DELLA SOCIETA ITALIANA DI FISICA C-GEOPHYSICS AND SPACE PHYSICS	105	0.152244762
PHYSICAL REVIEW D	117	0.0561696
PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC	59	0.055745158
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY	596	0.047172325
CHINESE JOURNAL OF ASTRONOMY AND ASTROPHYSICS	70	0.042848482
ASTRONOMICAL JOURNAL	88	0.031998146
ASTRONOMY LETTERS-A JOURNAL OF ASTRONOMY AND SPACE ASTROPHYSICS	58	0.027844759
INTERNATIONAL JOURNAL OF MODERN PHYSICS D	56	0.026884595
ASTROPHYSICAL JOURNAL LETTERS	162	0.026760324
ASTROPHYSICAL JOURNAL SUPPLEMENT SERIES	43	0.026321211
NEW ASTRONOMY REVIEWS	51	0.024484185
PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF JAPAN	49	0.017817149
ASTROPHYSICS AND SPACE SCIENCE	115	0.009102042
ASTROPHYSICAL JOURNAL	1459	0
ASTRONOMY & ASTROPHYSICS	589	0

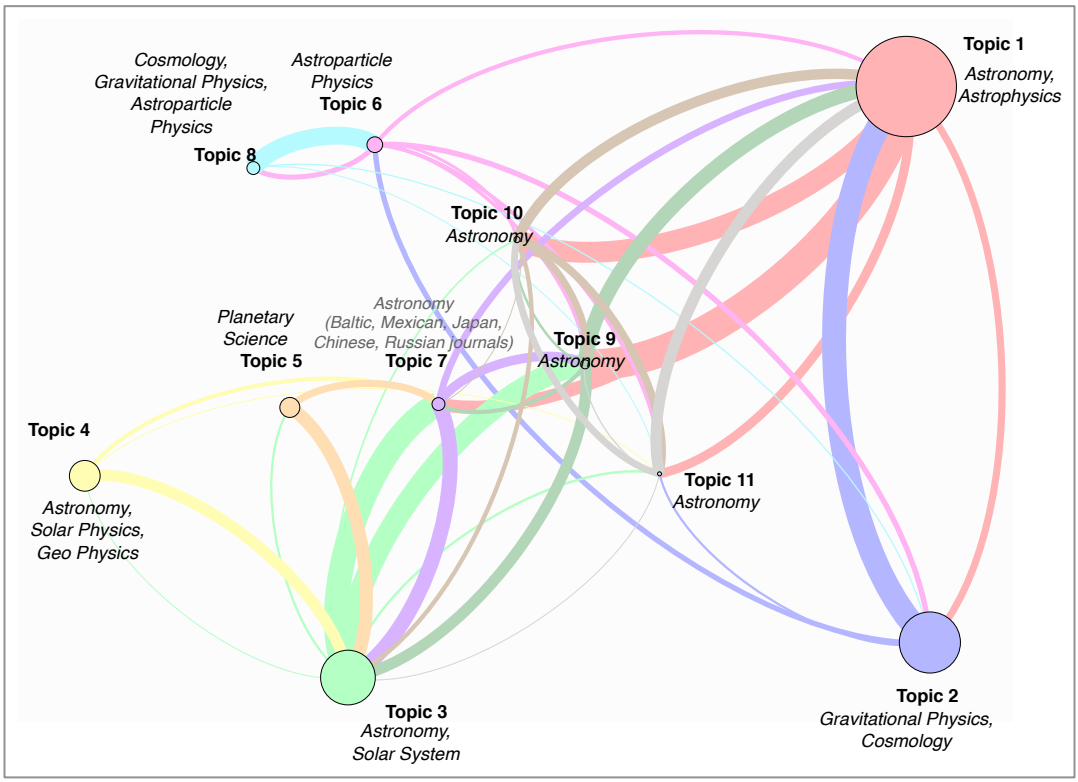

Figure 2: Topic affinity network. Node size indicates number of documents. Link strength indicates relative preference given by publications in one topic to cite publications in another. Links are directed: they are colored by their source node and curve clockwise away from it.

To further validate these hypotheses, a review of the topic contents and interpretation of the topic affinity links by experts could be insightful. Further, an extension of the data set backward in time to show the temporal evolution of affinity links could be informative. This would allow matching the evolution of affinity links over time to reports by experts about major research developments in this domain that may affect the interlinking between topics. One challenge in such an undertaking is that not just the linkages between topics evolve over time, but so does the identity of topics itself.

Conclusions

The topology of the affinity network highlights cognitive links between the topics extracted by our method from the astronomy and astrophysics data set. The interesting question in the context of the special session on the comparison of topic extraction algorithms will be what other cognitive features of this literature will be highlighted, if the affinity network is constructed for alternative groupings of documents into topics produced by other topic extraction algorithms. We suggest that this method of investigating the nature of differences between alternative topic extraction results is useful, in particular for cases where the topic size distribution is such that the large majority of documents, 80-90% is concentrated in 10-30 topics. For more granular topic extraction results the affinity network visualization is likely to become too unwieldy to interpret.

Acknowledgements

We gratefully acknowledge funding from SMA 1258891 EAGER: Collaborative Research: Scientific Collaboration in Time, as well as a travel grant by the intergovernmental framework for European Cooperation in Science and Technology (COST, Action: TD1210).

References

Boyack, K. W. & Klavans, R. (2010). Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? *Journal of the American Society for Information Science and Technology*, *61*(12), 2389–2404.

Crane, D. (1972). *Invisible Colleges - Diffusion of Knowledge in Scientific Communities*. The University of Chicago Press.

Ding, Y. (2011). Community detection: Topological vs. topical. *Journal of Informetrics*, *5*(4), 498–514.

Gläser, J. (2006). *Wissenschaftliche Produktionsgemeinschaften - die soziale Ordnung der Forschung*, Volume 906 of Campus Forschung. Frankfurt: Campus Verlag.

Morris, S. & Van der Veer Martens, B. (2008). Mapping research specialties. *Annual Review of Information Science and Technology*, *42*(1), 213–295.

Newman, M. (2003). The structure and function of complex networks. *SIAM Review*, *45*, 167–256.

Rosvall, M. & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, *105*(4), 1118–1123.

Velden, T. (2013). Explaining field differences in openness and sharing in scientific communities. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, 445–458. ACM.

Velden, T. & Lagoze, C. (2013). The extraction of community structures from publication networks to support ethnographic observations of field differences in scientific communication. *Journal of the American Society for Information Science and Technology*, *64*(12), 2405–2427.

Velden, T. & Lagoze, C. (2009). Patterns of collaboration in co-authorship networks in chemistry - mesoscopic analysis and interpretation. In Larsen, B. & Leta, J. (eds.), *Proceedings of ISSI 2009 - the 12th International Conference of the International Society for Scientometrics and Informetrics, Rio de Janeiro, Brazil, July 14-17, 2009* (2 volumes).

Velden, T., Haque, A. & Lagoze, C. (2010). A new approach to analyzing patterns of collaboration in co-authorship networks: mesoscopic analysis and interpretation. *Scientometrics*, *85*(1), 219–242.

Zuccala, A. (2006) Modeling the invisible college. *Journal of the American Society for Information Science and Technology*, *57*(2), 152 – 168.