Social Network Analysis:

Bibliographic Network Analysis of the Field and its Evolution

Daria Maltseva, Vladimir Batagelj

IMFM Ljubljana, IAM UP Koper and NRU HSE Moscow

Sredin Seminar

November 21, 2018

Outline

Inroduction

Social Network Analysis (SNA) has moved from a fragmented direction represented by the works of individual scientific groups unrelated to each other, to a discipline whose representatives by 1990 have formed an "invisible college" and achieved the status of what Kuhn had labeled a "normal science" [Freeman, 2004; Hummon and Carley, 1993].

Starting from that time, the field has grown significantly, which can be seen by the number of scientific publications [Otte and Rousseau, 2002] in different scientific fields, including Natural Sciences, which lead to the so called "physicists' invasion" into SNA [Batagelj et al., 2014] and resulted with the development of Network Science discipline.

This calls into a question whether the field remains unified and which scientific groups (by disciplines, thematic agenda, etc.) it is currently formed of. Thus, the aim of the current study is to trace the evolution of the field of Social Network Analysis using bibliographic approach.

Previous studies

1. Hummon N.P., Doreian P., Freeman L.C. (1990). Analyzing the Structure of the Centrality-Productivity Literature Created Between 1948 and 1979 / Science Communication. 11, 4, 459 – 480.
2. Freeman, L. (2004). The development of social network analysis. A Study in the Sociology of Science, 1.
3. Hummon, N. P., Carley, K. (1993). Social networks as normal science. Social networks, 15(1), 71-106.
4. Otte, E., Rousseau, R. (2002). Social network analysis: a powerful strategy, also for the information sciences. Journal of information Science, 28(6), 441-453.
5. Borgatti, S. P., Foster, P. C. (2003). The network paradigm in organizational research: A review and typology. Journal of management, 29(6), 991-1013.
6. Leydesdorff L., Schank T., Scharnhorst A., De Nooy W. (2008). Animating the development of Social Networks over time using a dynamic extension of multidimensional scaling / El Profesional de Informacion, 17(6).
7. Lazer, D., Mergel, I., and Friedman, A. (2009). Co-citation of prominent social network articles in sociology journals: The evolving canon. Connections, 29(1):43:64.
8. Brandes, U., Pich, C. (2011). Explorative visualization of citation patterns in social network research. Journal of social structure, 12(8), 1-19.
9. Varga, A. V., Nemeslaki, A. (2012) Do organizational network studies constitute a cohesive communicative field? Mapping the citation context of organizational network research. Journal of Sociology and Social Anthropology, 5(64), XV, 349-364.
10. Batagelj, V., Doreian P., V., Ferligoj, A., Kejžar N. (2014). Understanding Large Temporal Networks and Spatial Networks: Exploration, Pattern Searching, Visualization and Network Evolution.
11. Groenewegen, P., Hellsten, I., Leydesdorff, L. (2015) Social Networks as a looking glass on the social networks community. International Sunbelt XXXV Conference. Hilton Metropole, Brighton, UK, June 23 – 28, 2015. Abstracts, 118.

WoS

To the Web of Science (WoS), Clarivate Analytics's multidisciplinary databases of bibliographic information, we put the query

```
"social network*"
```

Additionally, all the articles from the following journals were collected:

```
Social Networks, Network Science,
Computational Social Networks, Applied Network Science,
Social Network Analysis and Mining,
Online Social Networks and Media, Journal of Complex
Networks, Journal of Social Structure, Connections
```

We limited the search to the Web of Science Core Collection because for other data bases from WoS the CR-fields (containing citation information) can not be exported.The first data set was collected in 2007, second - in June, 2018.

WoS record

```
PT J
AU GRANOVET.MS
TI STRENGTH OF WEAK TIES
SO AMERICAN JOURNAL OF SOCIOLOGY
LA English
DT Article
C1 JOHNS HOPKINS UNIV, BALTIMORE, MD 21218 USA.
CR BARNES JA, 1969, SOCIAL NETWORKS URBA
   BECKER MH, 1970, AM SOCIOL REV, V35, P267
   BERSCHEID E, 1969, INTERPERSONAL ATTRAC
   BOISSEVAIN J, 1968, MAN, V3, P542
   BOTT E, 1957, FAMILY SOCIAL NETWOR
NR 61
TC 2156
PU UNIV CHICAGO PRESS
PI CHICAGO
PA 5720 S WOODLAWN AVE, CHICAGO, IL 60637
SN 0002-9602
J9 AMER J SOCIOL
JI Am. J. Sociol.
PY 1973
VL 78
IS 6
BP 1360
EP 1380
PG 21
SC Sociology
GA P7726
UT ISI:A1973P772600003
ER
SK IP
```

WoS

We call a *terminal* node a node without a description in the collected data set – it appears only in the WoS CR field as a reference.

We additionally collected on WoS and Google data for terminal nodes with large indegree in the citation network – highly cited works without description in the collected data set. If a description of a node was not available in WoS we manually constructed a corresponding description without CR data (using RIS biblographic format and converting it to WoS).

As the data set of 2007 was already completed, we made this additional search only for works 2008-* in July 2018.

Types of networks and partitions

We applied the WoS2Pajek 1.5 to the collected data.

The following networks were constructed:

1. the authorship network *WA* on works × authors (from the field AU),
2. the journalship network *WJ* on works × journals (from the field CR or J9),
3. the keywordship network *WK* on works × keywords (from the field ID or DE or TI),
4. the citation network *Cite* on works (from the field CR).

We obtained also the following partitions:

1. partition *year* of works by publication year,
2. the *DC* partition distinguishing between works with complete description (DC=1) and the cited only works (DC=0),
3. the vector of number of pages *NP*.

ISI names

The usual *ISI name* of a work (field CR)

`GRANOVETTER M, 1985, AM J SOCIOL, V91, P481`

has the following structure

AU + ', ' + PY + ', ' + SO[:20] + ', V' + VL + ', P' + BP

All its elements are in upper case.
In WoS the same work can have different ISI names. To improve the precission the program `WoS2Pajek` supports also *short names* (similar to the names used in HISTCITE output). They have the format:

LastNm[:8] + '_' + FirstNm[0] + '(' + PY + ')' + VL + ':' + BP

For example: `GRANOVET_M(1985)91:481`
From the last names with prefixes `VAN`, `DE`, ... the space is deleted.
Unusual names start with character * or $.

Equivalent works, authors and journals


```
BOYD_D(2007)13 | BOYD_D(2008)13:210
GRANOVET_M(1973)78:1360 | GRANOVET_M(1973)78:6
COLEMAN_J(1988)94:95 | COLEMAN_J(1988)94:S95

GRANOVET_M
GRANOVET_

63656 1312696 10849 SONEANMI | SOCIAL NETWORK ANAL
63657 1330776 3 SONEANMI | SOCIAL NETWORKS ANAL
63658 1311789 645 SONEANMI | SOC NETW ANAL MIN
63659 1313366 7 SONEANMI | SOCIAL NETW ANAL MIN
63660 1315722 7 SONEANMI | SOC NETW ANAL MINING
...
55366 1351847 54714 1 PSPOSC  | PS POLITICAL SCIENCE
55768 1320199 23066 5 POSC | POLITICAL SCI
55769 1320573 23440 3 POSC | POLIT SCI
56082 1297982 849 224 PSSCPO | PS-POLIT SCI POLIT
56083 1298064 931 110 PSSCPO | PS-POLITICAL SCI POL
```

Equivalent works, authors and journals

There are two possibilities how to correct the data:

- to make corrections in the local copy of original data (WoS file);
- to make the equivalence partition of nodes and shrink the set of works accordingly in all obtained networks.

For the **works** with largest counts we prepared lists of possible equivalents and manually determined equivalence classes. With a program in R we produced a Pajek's partition *EQ.clu* file used for shrinking the set of works. Using the partition $p = worksEQ$, we shrink using Pajek the Citation network *cite*, *WA*, *WJ*, and *WK*. We had to shrink also partitions *year*, *DC* and the vector *NP*.

We manually inspected all **journals** with at least one of their names cited 200 and more times. To reduce the number of titles to inspect we considered only titles that appeared in at least 3 citations. We got the list *journalK*100.*csv* with 3714 titles for inspection.

We also produced a list of frequent journal names of length at most 5, have chosen all the cases that could be regarded as abbreviations (*CACM, JACM, JASA*), and manually searched for them.

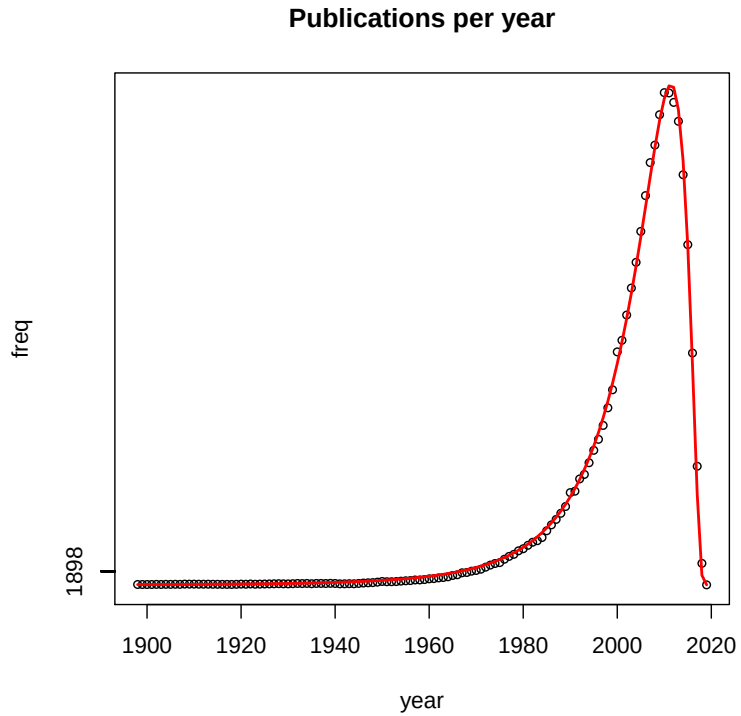
Sizes of Original cleaned and Reduced networks

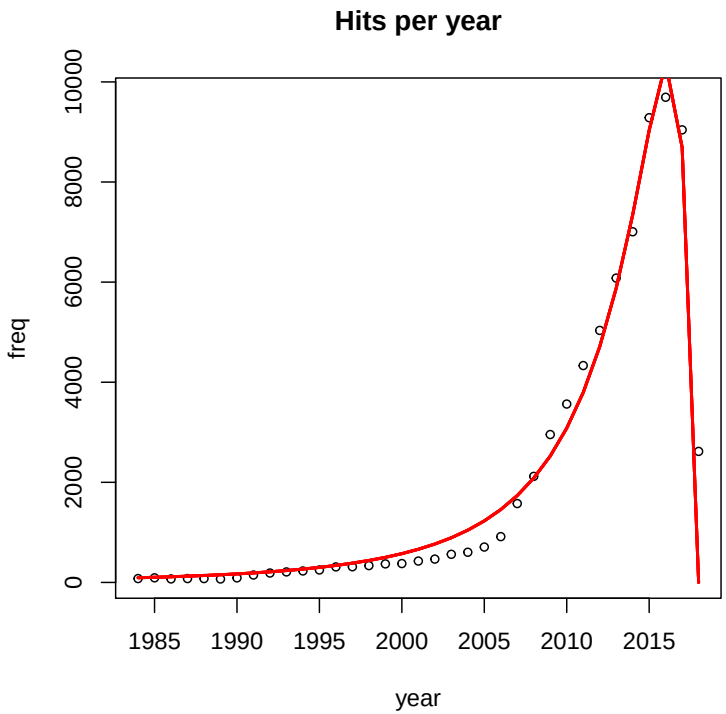
	# nodes (sum)	# nodes 1	# nodes 2	# arcs
WKn	1,329,542	1,297,133	32,409	1,167,670
WKr	103,201	**70,792**	**32,409**	1,167,666
WJn	1,366,279	1,297,133	69,146	720,044
WJr	79,735	**70,792**	**8,943**	61,741
WAn	1,693,104	1,297,133	395,971	1,442,240
WAr	163,803	**70,792**	**93,011**	215,901
CiteN	1,297,133			2,753,633
CiteR	**70,792**			398,199

An important property of all these networks is that they share as the first node set the same set of works (papers, reports, books, etc.) - they are *linked*.

Cite network

Distribution of works by years





The distribution fits the *log normal* distribution

$c \cdot \mathrm{dlnorm}(2019 - year, a, b)$

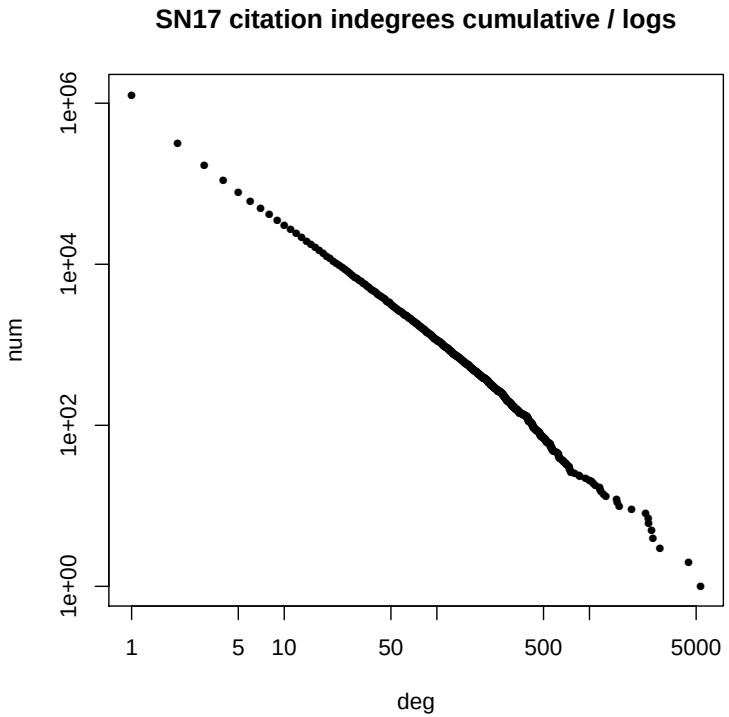
$a = 2.543$

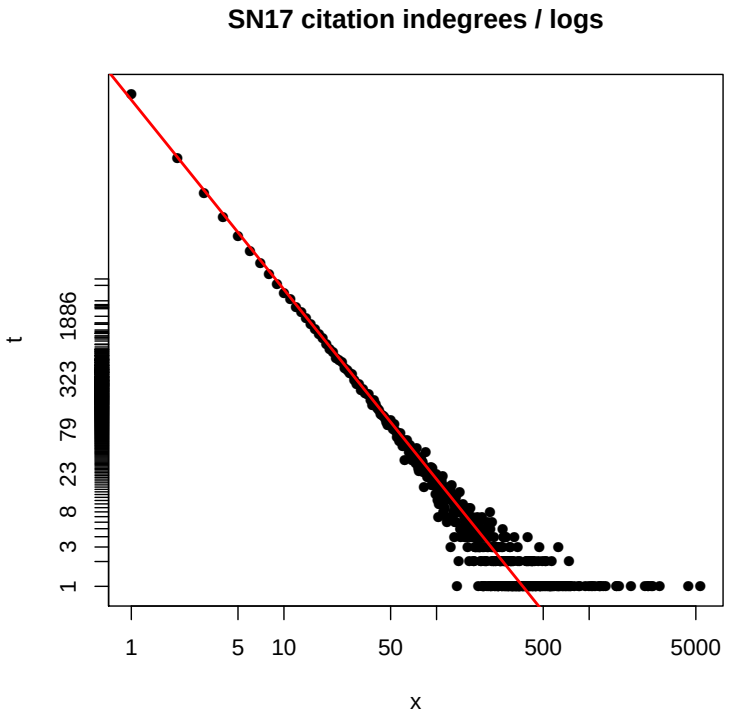
$b = 0.7206$

$c = 1.27810^6$

Cite network

Indegree distribution





The indegree distribution in citation network follows the *power law* $f = c \cdot n^{-\alpha}$.

Fitted $\alpha = 2.3007$, $c = 749338$.

Cite net

The most cited works - indegree

i	freq	id	i	freq	id
1	5348	WASSERMA_S(1994):	31	734	NEWMAN_M(2001)98:404
2	4471	GRANOVET_M(1973)78:1360	32	719	NEWMAN_M(2010):
3	2906	WATTS_D(1998)393:440	33	701	PORTES_A(1998)24:1
4	2614	BARABASI_A(1999)286:509	34	687	BLEI_D(2003)3:993
5	2561	FREEMAN_L(1979)1:215	35	670	BURT_R(2004)110:349
6	2447	BOYD_D(2007)13:210	36	654	HANSEN_M(1999)44:82
7	2429	MCPHERSO_M(2001)27:415	37	639	PALLA_G(2005)435:814
8	2330	BURT_R(1992):	38	634	CLAUSET_A(2004)70:066111
9	1886	COLEMAN_J(1988)94:95	39	629	BONACICH_P(1987)92:1170
10	1572	NEWMAN_M(2003)45:167	40	628	ERDOS_P(1959)6:290
11	1520	GIRVAN_M(2002)99:7821	41	628	UZZI_B(1997)42:35
12	1510	PUTNAM_R(2000):	42	628	ROGERS_E(2003):
13	1285	ALBERT_R(2002)74:47	43	613	PUTNAM_R(1993):
14	1240	GRANOVET_M(1985)91:481	44	593	BERKMAN_L(1979)109:186
15	1192	SCOTT_J(2000):	45	583	ZACHARY_W(1977)33:452
16	1171	EVERETT_M(2002):	46	572	BORGATTI_S(2009)323:892
17	1166	NEWMAN_M(2004)69:026113	47	569	NEWMAN_M(2001)64:025102
18	1093	COLEMAN_J(1990):	48	565	BURT_R(2005):
19	1058	STEINFIE_C(2007)12:1143	49	561	ADLER_P(2002)27:17
20	1034	FORTUNAT_S(2010)486:75	50	559	CHRISTAK_N(2008)358:2249
21	999	BORGATTI_S(2002):	51	555	ROGERS_E(1995):
22	945	CHRISTAK_N(2007)357:370	52	554	MILGRAM_S(1967)1:61
23	867	FREEMAN_L(1977)40:35	53	553	BARON_R(1986)51:1173
24	854	HANNEMAN_R(2005):	54	550	GRANOVET_M(1978)83:1420
25	800	LIN_N(2001):	55	539	FISCHER_C(1982):
26	757	KAPLAN_A(2010)53:59	56	537	BRIN_S(1998)30:107
27	756	BLONDEL_V(2008):P10008	57	524	MARSDEN_P(1990)16:435
28	742	NAHAPIET_J(1998)23:242	58	523	KEMP_D(2003):137
29	740	FORNELL_C(1981)18:39	59	523	KLEINBER_J(1999)46:604
30	740	NEWMAN_M(2006)103:8577	60	517	BOCCALET_S(2006)424:175

Cite net

The most citing work - outdegree

i	freq	id	i	freq	id
1	1572	CHAPMAN_C(2016):1	11	731	TSATSOU_P(2014):1
2	1406	HRUSCHKA_D(2010)5:1	12	654	GOODALE_E(2017):IX
3	1293	COWARD_F(2015):1	13	649	PEPPER_G(2017)40:S0140525X1700190X
4	1254	FITZGERA_P(2008):1	14	632	STROM_R(2012):1
5	1207	DAVIES_N(2015):V	15	613	SCHACHNE_G(2015)23:49
6	1055	MARSH_C(2009):1	16	597	**COSTA_L(2011)60:329**
7	942	YUS_F(2011)213:1	17	593	**BRANDES_U(2005)3418:1**
8	929	**BOCCALET_S(2006)424:175**	18	586	ROBERTS_J(2014):1
9	799	REEVES_M(2017):1	19	557	GUNTER_B(2016):1
10	768	GROSS_J(2007):1	20	547	CASTELLA_C(2009)81:591

- MUIJS, D., Reynolds, D., CHAPMAN, C. (2015). Educational effectiveness and improvement research and practice: The emergence of the discipline. In The Routledge International Handbook of Educational Effectiveness and Improvement (pp. 33-56). Routledge.
- Hruschka, D. J. (2010). Friendship: Development, ecology, and evolution of a relationship (Vol. 5). Univ of California Press.
- Coward, F., Hosfield, R., Pope, M., Wenban-Smith, F. (Eds.). (2015). Settlement, society and cognition in human evolution. Cambridge University Press.
- Fitzgerald, P., Lambkin, B. (2008). Migration in Irish history 1607-2007. Springer.
- Davies, N.B. Animal Social Networks Foreword. In: Krause, J., James, R., Franks, D. W., Croft, D. P. (Eds.). (2015). Animal social networks. Oxford University Press, USA.
- Marsh, C. J. (2009). Key concepts for understanding curriculum. Routledge.

WA net

Authors with the largest number of papers - indegree

Rank	Value	Id	Rank	Value	Id
1	1169	WANG_Y	21	552	KIM_H
2	883	ZHANG_Y	22	550	CHEN_J
3	868	CHEN_Y	23	536	LIU_X
4	847	LI_Y	24	533	WANG_L
5	838	WANG_X	25	509	LI_H
6	819	ZHANG_J	26	490	KIM_Y
7	788	WANG_J	27	485	ZHANG_Z
8	786	LIU_Y	28	474	WANG_Z
9	766	LEE_J	29	471	WANG_S
10	765	LEE_S	30	471	CHEN_X
11	749	LI_J	31	471	**NEWMAN_M**
12	708	LI_X	32	462	CHEN_L
13	696	CHEN_C	33	461	ZHANG_L
14	690	KIM_J	34	450	YANG_Y
15	620	WANG_H	35	450	ZHANG_H
16	611	ZHANG_X	36	432	WU_J
17	611	LIU_J	37	431	LEE_H
18	570	CHEN_H	38	420	LI_Z
19	557	KIM_S	39	420	WANG_W
20	554	WANG_C	40	417	LI_L

The large number of Chinese authors in the list is a "three Zhang, four Li" effect. It is out of our resources to drill into this. We can only make a warning.

WA net

Number of authors in works - outdegree

outdeg	Freq	Freq%	outdeg	Freq	Freq%
1	1239496	95.5566	21	4	0.0003
2	18637	1.4368	22	3	0.0002
3	16661	1.2844	23	4	0.0003
4	10617	0.8185	24	2	0.0002
5	5759	0.4440	25	1	0.0001
6	2802	0.2160	26	2	0.0002
7	1322	0.1019	27	5	0.0004
8	686	0.0529	28	2	0.0002
9	384	0.0296	29	1	0.0001
10	247	0.0190	31	3	0.0002
11	155	0.0119	36	1	0.0001
12	90	0.0069	41	1	0.0001
13	70	0.0054	42	1	0.0001
14	54	0.0042	43	1	0.0001
15	32	0.0025	48	1	0.0001
16	12	0.0009	53	1	0.0001
17	14	0.0011	126	1	0.0001
18	9	0.0007			
19	6	0.0005			
20	2	0.0002			
SUM				1297133	100

Works with the largest number of authors:

Rank	Freq	Id
1	126	WANG_M(2016)34:828
2	53	VASHISHT_R(2012)7:0039808
3	48	SNIJDERS_T(2007)170:322
4	43	GUSTAVSS_A(2011)21:718
5	42	DOLL_L(1992)29:1
6	41	MAGLIANO_L(2006)15:219

WA net

Works with the largest number of authors - outdegree

Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking / Nature Biotechnology volume 34, pages 828–837 (2016)

Mingxun Wang, Jeremy J Carver, Vanessa V Phelan, Laura M Sanchez, Neha Garg, Yao Peng, Don Duy Nguyen, Jeramie Watrous, Clifford A Kapono, Tal Luzzatto-Knaan, Carla Porto, Amina Bouslimani, Alexey V Melnik, Michael J Meehan, Wei-Ting Liu, Max Crüsemann, Paul D Boudreau, Eduardo Esquenazi, Mario Sandoval-Calderón, Roland D Kersten, Laura A Pace, Robert A Quinn, Katherine R Duncan, Cheng-Chih Hsu, Dimitrios J Floros, Ronnie G Gavilan, Karin Kleigrewe, Trent Northen, Rachel J Dutton, Delphine Parrot, Erin E Carlson, Bertrand Aigle, Charlotte F Michelsen, Lars Jelsbak, Christian Sohlenkamp, Pavel Pevzner, Anna Edlund, Jeffrey McLean, Jörn Piel, Brian T Murphy, Lena Gerwick, Chih-Chuang Liaw, Yu-Liang Yang, Hans-Ulrich Humpf, Maria Maansson, Robert A Keyzers, Amy C Sims, Andrew R Johnson, Ashley M Sidebottom, Brian E Sedio, Andreas Klitgaard, Charles B Larson, Cristopher A Boya P, Daniel Torres-Mendoza, David J Gonzalez, Denise B Silva, Lucas M Marques, Daniel P Demarque, Egle Pociute, Ellis C O'Neill, Enora Briand, Eric J N Helfrich, Eve A Granatosky, Evgenia Glukhov, Florian Ryffel, Hailey Houson, Hosein Mohimani, Jenan J Kharbush, Yi Zeng, Julia A Vorholt, Kenji L Kurita, Pep Charusanti, Kerry L McPhail, Kristian Fog Nielsen, Lisa Vuong, Maryam Elfeki, Matthew F Traxler, Niclas Engene, Nobuhiro Koyama, Oliver B Vining, Ralph Baric, Ricardo R Silva, Samantha J Mascuch, Sophie Tomasi, Stefan Jenkins, Venkat Macherla, Thomas Hoffman, Vinayak Agarwal, Philip G Williams, Jingqui Dai, Ram Neupane, Joshua Gurr, Andrés M C Rodríguez, Anne Lamsa, Chen Zhang, Kathleen Dorrestein, Brendan M Duggan, Jehad Almaliti, Pierre-Marie Allard, Prasad Phapale, Louis-Felix Nothias, Theodore Alexandrov, Marc Litaudon, Jean-Luc Wolfender, Jennifer E Kyle, Thomas O Metz, Tyler Peryea, Dac-Trung Nguyen, Danielle VanLeer, Paul Shinn, Ajit Jadhav, Rolf Müller, Katrina M Waters, Wenyuan Shi, Xueting Liu, Lixin Zhang, Rob Knight, Paul R Jensen, Bernhard Ø Palsson, Kit Pogliano, Roger G Linington, Marcelino Gutiérrez, Norberto P Lopes, William H Gerwick, Bradley S Moore, Pieter C Dorrestein, Nuno Bandeira.

WJ net

The most used journals - indegree

Rank	Value	Id	Rank	Value	Id
1	7757	LECT NOTES COMPUT SC	31	1278	**BRIT J PSYCHIAT**
2	3866	SOC SCI MED	32	1267	**AM J PSYCHIAT**
3	3414	**J PERS SOC PSYCHOL**	33	1244	**STRATEGIC MANAGE J**
4	2741	P NATL ACAD SCI USA	34	1225	**MANAGE SCI**
5	2734	COMPUT HUM BEHAV	35	1221	**J BUS RES**
6	2631	SCIENCE	36	1189	**ACAD MANAGE REV**
7	2609	AM J PUBLIC HEALTH	37	1188	**J CONSULT CLIN PSYCH**
8	2208	NATURE	38	1154	**ORGAN SCI**
9	2111	**AM SOCIOL REV**	39	1150	ADDICTION
10	1945	PHYSICA A	40	1123	CYBERPSYCHOL BEHAV
11	1825	ANIM BEHAV	41	1092	COMPUT EDUC
12	1812	**AM J SOCIOL**	42	1087	**J GERONTOL B-PSYCHOL**
13	1780	JAMA-J AM MED ASSOC	43	1075	PEDIATRICS
14	1763	LANCET	44	1067	AM J EPIDEMIOL
15	1759	SCIENTOMETRICS	45	1024	**DEV PSYCHOL**
16	1703	**ACAD MANAGE J**	46	1022	**PSYCHOL BULL**
17	1668	LECT NOTES ARTIF INT	47	1020	INFORM SCI
18	1642	***SOC NETWORKS***	48	1016	J ADOLESCENT HEALTH
19	1573	**J APPL PSYCHOL**	49	1009	ARCH GEN PSYCHIAT
20	1517	AM ECON REV	50	997	**J MARKETING**
21	1450	**J MARRIAGE FAM**	51	994	AIDS BEHAV
22	1441	EXPERT SYST APPL	52	972	PERS INDIV DIFFER
23	1403	BRIT MED J	53	949	**PERS SOC PSYCHOL B**
24	1399	CHILD DEV	54	947	J BUS ETHICS
25	1379	**RES POLICY**	55	939	**J MARKETING RES**
26	1372	COMMUN ACM	56	925	**HARVARD BUS REV**
27	1365	NEW ENGL J MED	57	915	IEEE T KNOWL DATA EN
28	1311	PHYS REV E	58	914	DRUG ALCOHOL DEPEN
29	1287	**SOC FORCES**	59	908	J ADV NURS
30	1279	GERONTOLOGIST	60	906	MIS QUART

WK net

The most used keywords - indegree

Rank	Value	Id	Rank	Value	Id
1	51333	**social**	31	3485	**structure**
2	46191	**network**	32	3479	life
3	11751	**analysis**	33	3444	risk
4	10219	**model**	34	3358	research
5	8104	**community**	35	3143	learn
6	8090	use	36	3116	influence
7	7596	base	37	3054	student
8	7439	information	38	3054	impact
9	7061	health	39	3049	perspective
10	7023	behavior	40	3042	complex
11	6745	online	41	3024	theory
12	6087	networking	42	2859	organization
13	5833	media	43	2828	**relationship**
14	5404	support	44	2802	algorithm
15	5101	communication	45	2776	education
16	5013	study	46	2714	group
17	4759	datum	47	2704	mobile
18	4376	management	48	2698	**tie**
19	4372	internet	49	2695	adult
20	4164	knowledge	50	2633	approach
21	4126	user	51	2608	care
22	4023	facebook	52	2551	adolescent
23	3984	technology	53	2479	role
24	3907	site	54	2472	state
25	3888	web	55	2467	innovation
26	3855	self	56	2434	pattern
27	3784	**graph**	57	2385	effect
28	3676	performance	58	2339	people
29	3534	service	59	2333	trust
30	3512	dynamics	60	2332	family

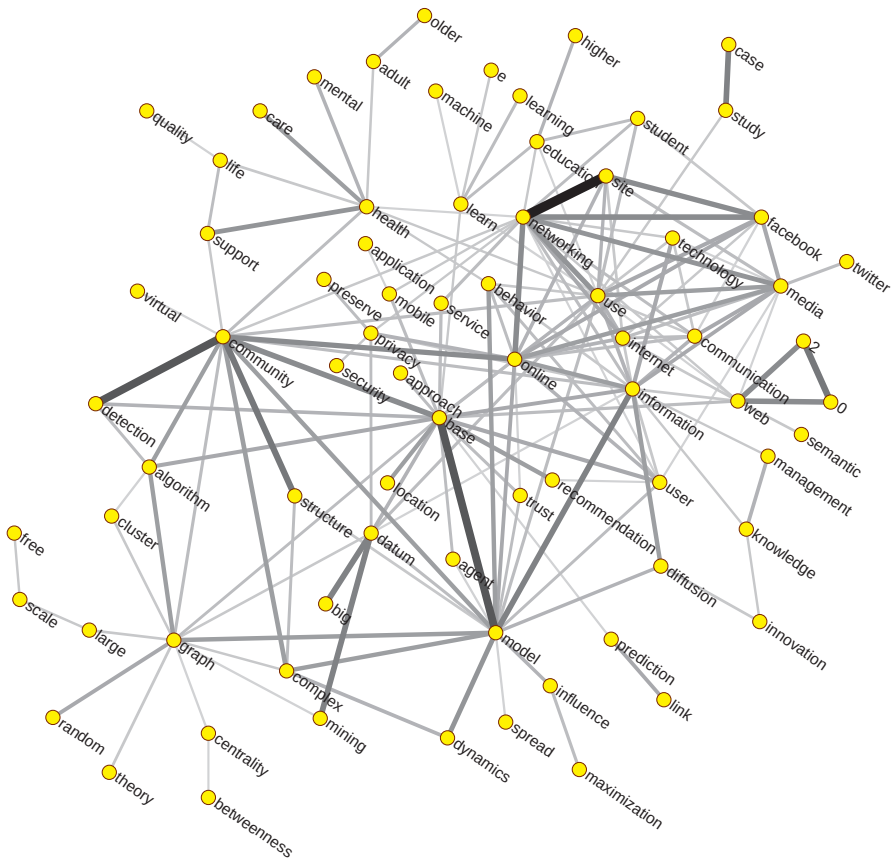
Topical structure

KK net

```
KKn = t(n(WK))*n(WK), where n(W,K)[w,k] = WK[w,k]/
(outdeg(w)-1)
```

Topical structure

KK net Main island



Cite net

Boundary problem

The network Cite has 1,297,133 nodes and 2,753,767 arcs.

indeg	Freq	Freq%	CumFreq	CumFreq%
0	41954	3.2344	41954	3.2344
1	933315	71.9521	975269	75.1865
2	154895	11.9413	1130164	87.1278
3	58141	4.4823	1188305	91.6101
4	29885	2.3039	1218190	93.9140
5	17651	1.3608	1235841	95.2748

Most of nodes are terminal ($DC = 0$) or nodes cited only once (indegree=1). We decided (*boundary problem*) to include in our networks nodes with $DC > 0$ or indeg > 2 (partition boundary). They determine a subnetwork **CiteB** with 222,086 nodes and 1,521,434 arcs.

Cite net

Components

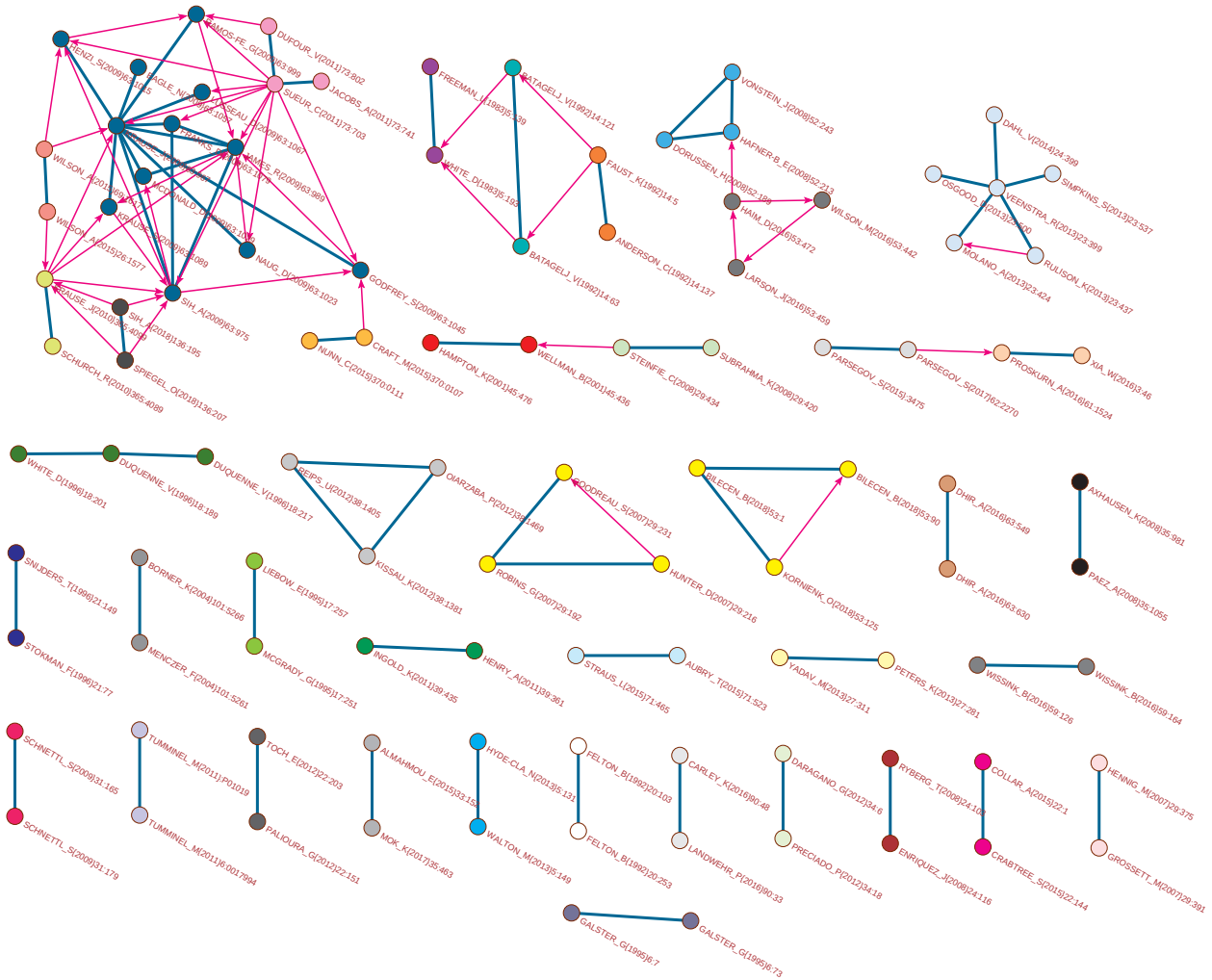
The citation network CiteB has 41 nontrivial strong components (see figure).
To get an acyclic network we applied the *preprint transformation* to CiteB. The resulting network **CiteT** has 222,189 nodes and 1,521,658 arcs.
We computed the SPC weights on network arcs, and determined

- CPM path / Main path = 59 nodes
- Key-routes = 127 nodes
- SPC link islands [Line weights] of sizes [20, 200] = 5 islands of 138, 65, 13, 12, and 11 nodes
- SPC node islands [Vertex weights] of sizs [20, 200] = 1 island of 200 nodes

We computed the Probabilistic flow on weighted network, and determined Node islands [Vertex weights] of sizes [10, 200] = 1 island of 200 nodes

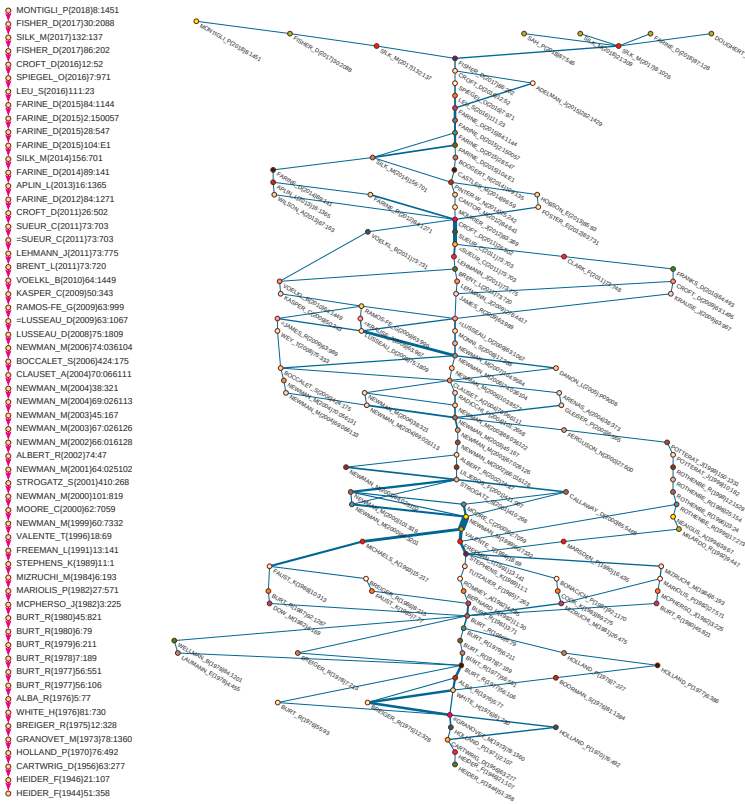
Strong components

from SPC network



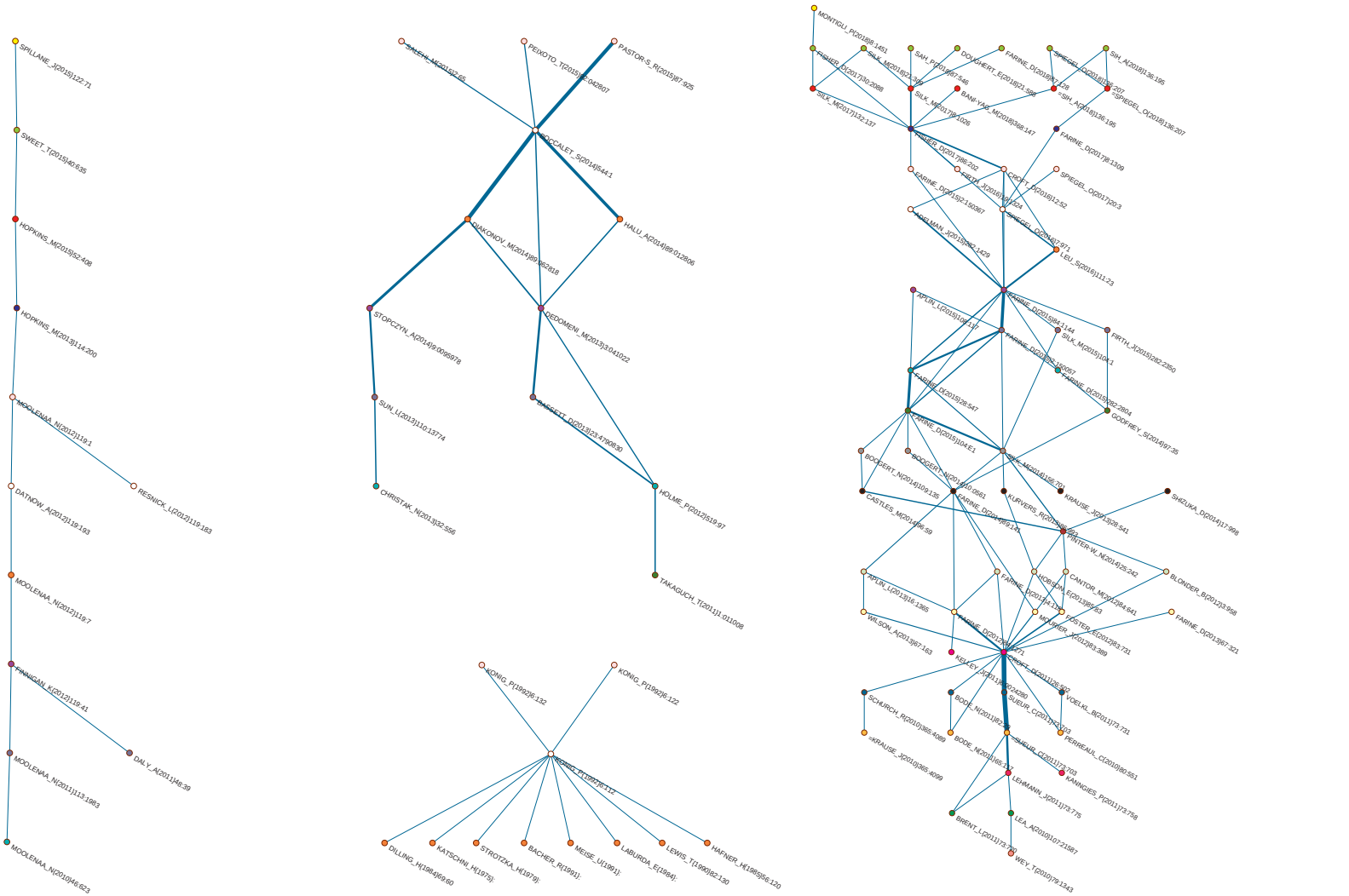
Main path, Key Routes, and Island 4

from SPC network



Islands 1-3, 5

from SPC network



Most important works

from Probabilistic Flow network

Rank	Value	Id	Rank	Value	Id
1	4691	WASSERMA_S(1994):	31	545	BLONDEL_V(2008):P10008
2	2941	WATTS_D(1998)393:440	32	527	KATZ_L(1953)18:39
3	2676	GRANOVET_M(1973)78:1360	33	526	NEWMAN_M(2010):
4	2445	BOYD_D(2007)13:210	34	520	STROGATZ_S(2001)410:268
5	2241	BARABASI_A(1999)286:509	35	517	PALLA_G(2005)435:814
6	1926	FREEMAN_L(1979)1:215	36	499	CLAUSET_A(2004)70:066111
7	1396	GIRVAN_M(2002)99:7821	37	497	ERDOS_P(1960)5:17
8	1299	NEWMAN_M(2003)45:167	38	488	ROGERS_E(2003):
9	1227	MCPHERSO_M(2001)27:415	39	485	NEWMAN_M(2006)103:8577
10	1158	ALBERT_R(2002)74:47	40	481	COLEMAN_J(1990):
11	1105	SCOTT_J(2000):	41	478	BRIN_S(1998)30:107
12	1098	BURT_R(1992):	42	477	AMARAL_L(2000)97:11149
13	1045	MILGRAM_S(1967)1:61	43	475	ERDOS_P(1959)6:290
14	1013	NEWMAN_M(2004)69:026113	44	465	WATTS_D(1999):
15	928	KAPLAN_A(2010)53:59	45	462	LAVE_J(1991):
16	878	FREEMAN_L(1977)40:35	46	460	KLEINBER_J(1999)46:604
17	852	PUTNAM_R(2000):	47	449	SCOTT_J(1991):
18	847	COLEMAN_J(1988)94:95	48	446	BOLLOBAS_B(1985):
19	835	BLEI_D(2003)3:993	49	442	PAGE_L(1999):
20	742	GRANOVET_M(1985)91:481	50	440	NEWMAN_M(2001)64:025102
21	731	CHRISTAK_N(2007)357:370	51	436	NEWMAN_M(2004)69:066133
22	727	EVERETT_M(2002):	52	431	REDNER_S(1998)4:131
23	726	NEWMAN_M(2001)98:404	53	429	CHRISTAK_N(2008)358:2249
24	719	ALBERT_R(1999)401:130	54	424	ADOMAVIC_G(2005)17:734
25	701	O'REILLY_T(2005):	55	424	KEMP_D(2003):137
26	669	BORGATTI_S(2002):	56	423	DOMINGOS_P(2001):57
27	667	FORTUNAT_S(2010)486:75	57	423	MITCHELL_J(1969):
28	633	HANNEMAN_R(2005):	58	415	ALBERT_R(2000)406:378
29	569	STEINFIE_C(2007)12:1143	59	415	GLASER_B(1967):
30	549	ZACHARY_W(1977)33:452	60	410	ROGERS_E(1995):

Cite net

Overlapping of components

i	name	title	jour	comp
1	Granovet M	Strength of weak ties	amer j sociol	1, 2, 4, 5, 6
2	Newman M	The structure and function of complex networks	siam rev	1, 2, 4, 5, 6
3	Albert R	Statistical mechanics of complex networks	rev mod phys	1, 2, 4, 5, 6
4	Boccaletti S	Complex networks: structure and dynamics	phys rept	1, 2, 4, 5, 6
5	White H	Soc. str. from mult. nets. Blockmodels	amer j sociol	1, 2, 4, 5, 6
6	Newman M	Clustering and pref.l attach. in growing nets	phys rev e	1, 2, 4, 5, 6
7	Newman M	Finding and evaluating comm. struct. in nets	phys rev e	1, 2, 4, 5, 6
8	Newman M	Mixing patterns in networks	phys rev e	1, 2, 4, 5, 6
9	Strogatz S	Exploring complex networks	nature	1, 2, 4, 5, 6
10	Newman M	Detecting community structure in nets	eur phys j b	1, 2, 4, 5, 6
11	Newman M	Spread of epidemic disease on nets	phys rev e	1, 2, 4, 5, 6
12	Newman M	Finding community str. in nets using eigenvectors	phys rev e	1, 2, 4, 5, 6
13	Cartwright D	Structural balance - a generaliz. of heider theory	psychol rev	1, 2, 4, 5, 6
14	Clauset A	Finding community struct. in very large nets	phys rev e	1, 2, 4, 5, 6
15	Newman M	Models of the small world	j statist phys	1, 2, 4, 5
16	Newman M	Scaling and percolation in small-world net model	phys rev e	1, 2, 4, 5
17	Valente T	Social net thresholds in the diff. of innov.	soc networks	1, 2, 4, 5
18	Burt R	Cohesion versus structural equivalences as a basis for net subgroups	soc meth res	1, 2, 4, 5
19	Stephenson K	Rethinking centrality - methods and examples	soc networks	1, 2, 4, 5
20	Breiger R	Algorithm for clustering relational data	j math psychl	1, 2, 4, 5
21	Freeman L	Centrality in valued graphs - a measure of betweenness based on net flow	soc networks	1, 2, 4, 5
22	Burt R	Models of network structure	annu rev soc	1, 2, 4, 5
23	Holland P	Method for detecting structure in sociom. data	amer j sociol	1, 2, 4, 5
24	Alba R	Intersection of social circles	socl meth res	1, 2, 4, 5
25	Moore C	Exact solution of site and bond percolation on small-world net	phys rev e	1, 2, 4, 5
26	Mcpherson J	Hypernetwork sampling - duality and differentiation among voluntary organizations	soc networks	1, 2, 4, 5
27	Mariolis P	Centrality in corporate interlock networks	adm sci quart	1, 2, 4, 5
28	Burt R	Positions in multiple network systems 1. General conception of stratification and prestige	soc forces	1, 2, 4, 5
29	Burt R	Positions in multiple network systems 2. Stratification and prestige among elite	soc forces	1, 2, 4, 5
30	Mizruchi M	Interlock groups, cliques, or interest-groups	soc networks	1, 2, 4, 5

1- Key Routes, 2- Main Path (CPM), 3- Island5, 4 - Island 4, Node Island, 5 - Prob Flow Island

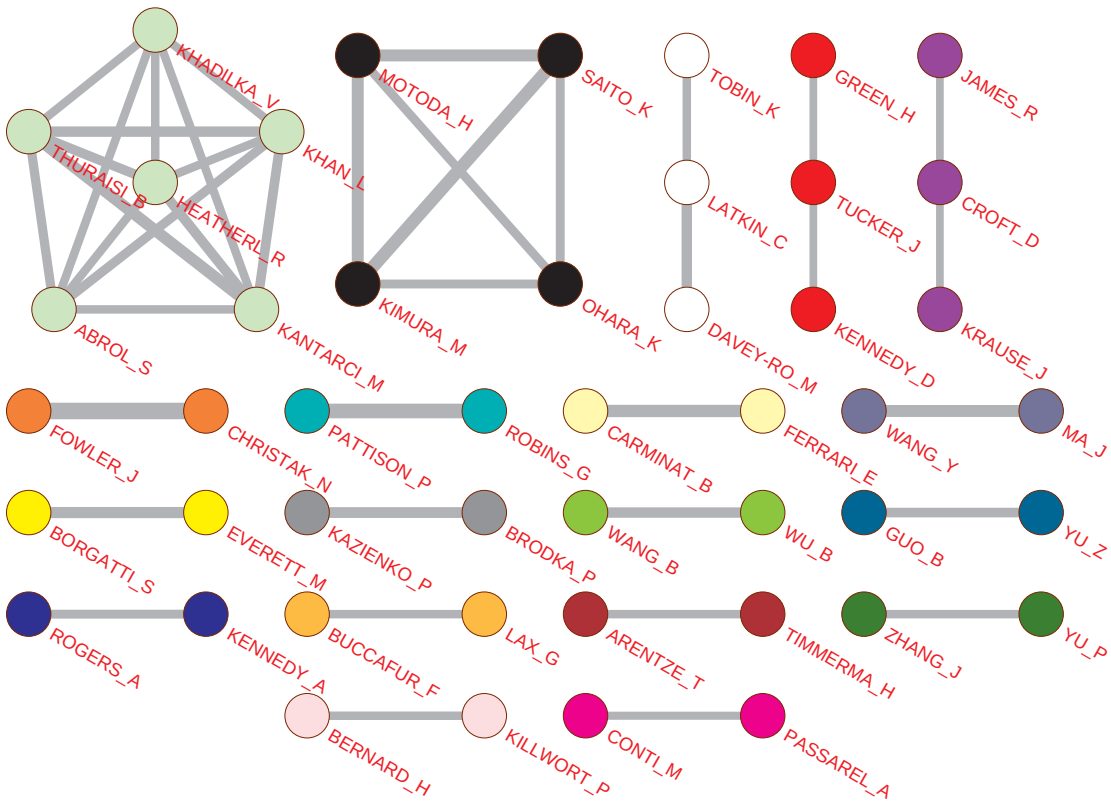
Authors Collaboration

from WA net

```
Co = t(WAsr) * WAsr = AWsr * WAsr = AA
Cn = t(WAsr) * n(WAsr), where n(WAsr)[w,a] = WAsr[w,a]
/ outdeg(w)
Ct‘ = t(n(WAsr)) * n’(WAsr), where n’(WAsr)[w,a] =
WAsr[w,a] / (outdeg(w)-1)
```

Authors Collaboration

from Co net (20 and more works written together)



Authors Collaboration

Collaborativeness index from Cn net

#	Author	Tot Contr	Tot #works	Collab	#	Author	Tot Contr	Tot #works	Collab
1	BURT_R	55,73	71	0,22	31	**PATTISON_P**	18,94	58	0,67
2	NEWMAN_M	50,02	81	0,38	32	THELWALL_M	18,41	37	0,5
3	DOREIAN_P	46,19	72	0,36	33	KRACKHAR_D	18,24	38	0,52
4	**PARK_H**	41,94	113	0,63	34	FALOUTSO_C	17,86	60	0,7
5	DUNBAR_R	40,02	91	0,56	35	JACKSON_M	17,78	38	0,53
6	WELLMAN_B	36,43	63	0,42	36	**GONZALEZ_M**	17,76	52	0,66
7	**VALENTE_T**	34,96	97	0,64	37	MOODY_J	17,7	40	0,56
8	**PARK_S**	34,59	109	0,68	38	SCOTT_J	17,54	28	0,37
9	BONACICH_P	34	46	0,26	39	MORRIS_M	17,22	43	0,6
10	LEYDESDO_L	33,28	51	0,35	40	**RODRIGUE_J**	15,9	52	0,69
11	**LATKIN_C**	32,99	130	0,75	41	WASSERMA_S	15,64	35	0,55
12	LITWIN_H	32,42	50	0,35	42	KLEINBER_J	15,05	34	0,56
13	MARSDEN_P	30,17	39	0,23	43	BATAGELJ_V	14,64	33	0,56
14	BORGATTI_S	29,72	71	0,58	44	WILLIAMS_A	14,5	31	0,53
15	SNIJDERS_T	29,63	67	0,56	45	**SINGH_A**	14,5	36	0,60
16	FRIEDKIN_N	28,17	36	0,22	46	BRANDES_U	14,39	35	0,59
17	CARLEY_K	28,11	72	0,61	47	**BERKMAN_L**	14,3	39	0,63
18	BARABASI_A	27,61	67	0,59	48	MASUDA_N	14,26	28	0,49
19	WHITE_H	27,28	42	0,35	49	**SMITH_A**	14,2	40	0,65
20	**CHRISTAK_N**	22,89	74	0,69	50	LAZEGA_E	14,17	26	0,46
21	EVERETT_M	22,58	44	0,49	51	**CONTRACT_N**	14,15	43	0,67
22	**KAZIENKO_P**	21,97	64	0,66	52	**GONZALEZ_A**	14,13	35	0,60
23	MARTINEZ_M	21,9	53	0,59	53	**PENTLAND_A**	14,12	41	0,66
24	**JOHNSON_J**	21,19	54	0,61	54	FARINE_D	14,04	34	0,59
25	**FOWLER_J**	20,14	65	0,69	55	SCHNEIDE_J	13,89	52	0,73
26	SKVORETZ_J	20,07	42	0,52	56	WATTS_D	13,67	27	0,49
27	FREEMAN_L	20,03	27	0,26	57	FAUST_K	13,5	25	0,46
28	BREIGER_R	19,73	31	0,36	58	**SMITH_M**	13,29	39	0,66
29	**ROBINS_G**	19,67	64	0,69	59	**RODRIGUE_M**	13,21	46	0,71
30	**RAHMAN_M**	19,18	59	0,67	60	**RICE_E**	13,09	48	0,73

Authors Collaboration

from Ct‘ net

Key words in coauthorship islands

from AK net (nAWr x nWKr)

	BORGATTI_S		BARABASI_A		CHRISTAKIS_K	
Rank	Value	Id	Value	Id	Value	Id
1	4.9303	network	7.0709	network	3.1788	network
2	2.5918	social	2.0782	social	2.9358	social
3	2.0858	graph	1.7068	dynamics	1.0204	spread
4	1.4210	centrality	1.6670	complex	1.0192	behavior
5	1.4202	analysis	1.6362	scale	0.7261	health
6	1.3399	role	1.5946	web	0.5512	large
7	1.2780	regular	1.5516	community	0.5169	model
8	1.2424	equivalence	1.4709	world	0.4778	smoking
9	1.0530	semigroup	1.3622	internet	0.4522	human
10	1.0000	correction	1.1906	model	0.4479	cooperation
11	0.9891	structure	1.1858	free	0.4313	obesity
12	0.7755	clique	1.0210	evolve	0.4125	influence
13	0.7576	homomorphism	1.0087	science	0.3973	life
14	0.7241	relation	0.9808	random	0.3728	dynamics
15	0.6346	power	0.9476	wide	0.3715	evolution
16	0.6301	betweenness	0.8178	human	0.3463	analysis
17	0.6287	exchange	0.8076	theory	0.3286	cosponsorship
18	0.6232	algorithm	0.7561	small	0.3044	norm
19	0.6167	similarity	0.7536	graph	0.3036	trial
20	0.5595	ebloc	0.6603	phenomenon	0.2985	study
Sum:	63.0810		76.6373		46.8865	

Key words in coauthorship islands

from AK net (nAWr x nWKr)

	PATTISON_P		SNIJDERS_T		VALENTE_T	
Rank	Value	Id	Value	Id	Value	Id
1	2.2196	network	2.6375	network	2.5536	network
2	2.0729	social	2.0902	social	1.9553	social
3	1.7567	model	1.6702	model	1.0000	untitled
4	1.3084	graph	1.0692	graph	0.9419	health
5	0.8939	random	0.8857	dynamics	0.8737	diffusion
6	0.8583	markov	0.7390	markov	0.7802	behavior
7	0.8531	logit	0.6903	random	0.7402	innovation
8	0.8220	logistic	0.6734	friendship	0.6974	model
9	0.8220	regression	0.6228	datum	0.6521	use
10	0.8012	exponential	0.5932	statistical	0.6349	peer
11	0.7055	analysis	0.5780	behavior	0.6216	adolescent
12	0.6752	p	0.5547	analysis	0.5717	influence
13	0.5530	statistical	0.5423	peer	0.5610	smoking
14	0.5038	structure	0.5383	inference	0.5371	analysis
15	0.3561	semigroup	0.5346	influence	0.5247	prevention
16	0.3522	asterisk	0.4623	stochastic	0.4987	cigarette
17	0.3368	process	0.4612	actor	0.4979	opinion
18	0.3333	multirelational	0.4480	selection	0.4860	leader
19	0.3249	family	0.4372	longitudinal	0.4545	risk
20	0.3031	dynamics	0.3785	orient	0.4491	intervention
Sum:	38.6110		46.6732		44.8812	

Citation among authors

from WA and Cite nets

```
CiteA = t(WAsr) * CiteR * WAsr

CiteAn = t(WAsr) * nCiteR * WAsr
```

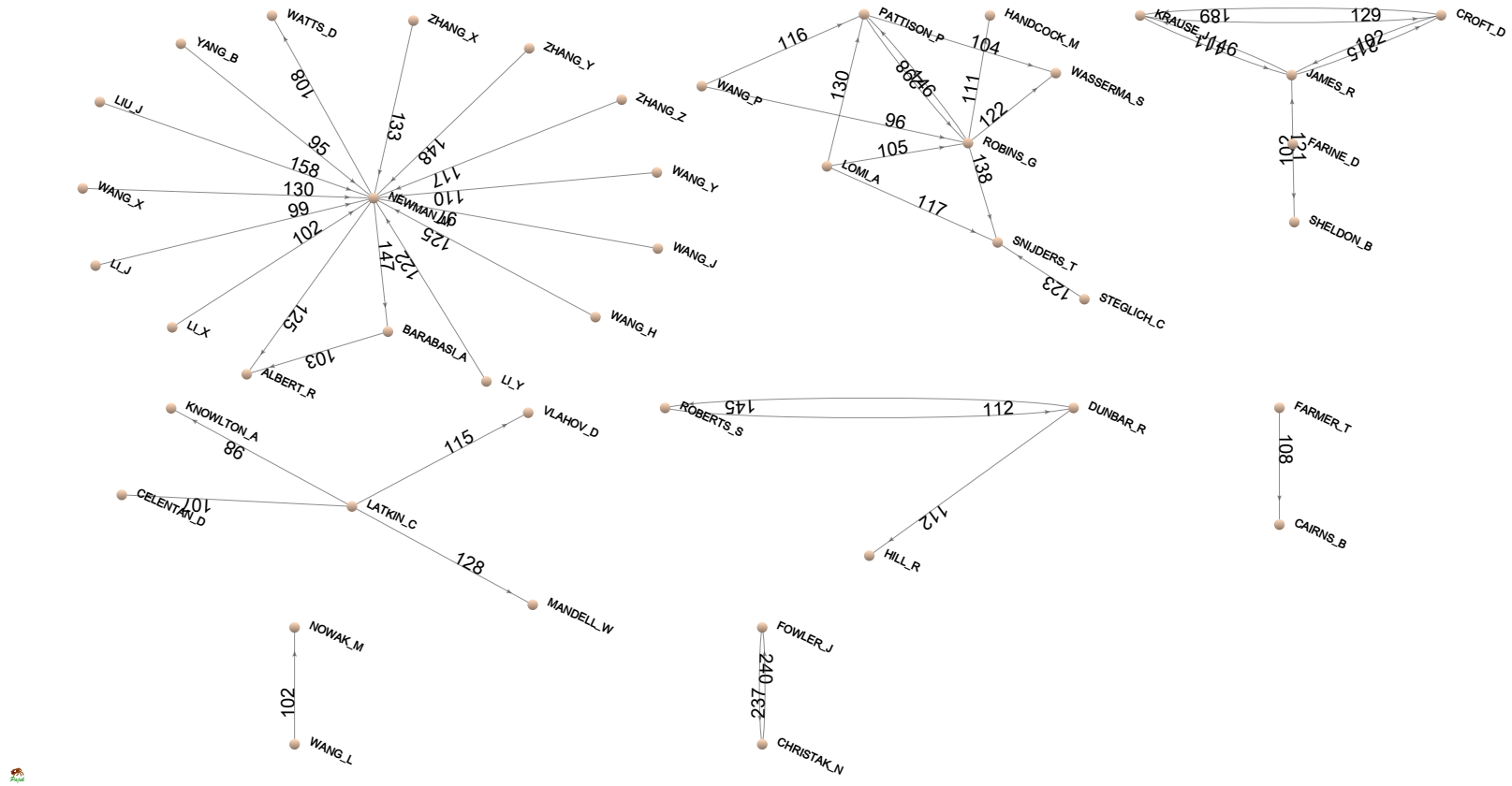
Citation among authors

self-citation from CiteA net

Rank	Value	Id	Rank	Value	Id
1	589	DUNBAR_R	11	201	BARABASI_A
2	387	LATKIN_C	12	191	FARINE_D
3	292	CHRISTAK_N	13	188	SNIJDERS_T
4	280	VALENTE_T	14	153	WELLMAN_B
5	268	BURT_R	15	148	DOREIAN_P
6	248	NEWMAN_M	16	146	BORGATTI_S
7	232	ROBINS_G	17	146	ZENOU_Y
8	224	PATTISON_P	18	143	RICE_E
9	221	FOWLER_J	19	142	JAMES_R
10	204	CROFT_D	20	141	KRAUSE_J

Citation among authors

authors with largest line weights from CiteA net



Citation among authors

self-citation from CiteAn net

#	Author	Value	%	#	Author	Value	%
1	TANG_J	9.786	0.025	26	FALOUTSO_C	8.169	0.014
2	FARINE_D	9.778	0.074	27	HANSON_B	8.091	0.041
3	PENTLAND_A	9.746	0.016	28	SUEUR_C	8.025	0.092
4	MARATHE_M	9.528	0.034	29	FRANK_K	7.995	0.062
5	ZENOU_Y	9.466	0.077	30	LI_X	7.966	0.020
6	EVERETT_M	9.342	0.012	31	MORENO_M	7.724	0.034
7	KRAUSE_J	9.035	0.022	32	THELWALL_M	7.698	0.033
8	CHEN_H	8.975	0.032	33	SHEN_X	7.679	0.039
9	BERKMAN_L	8.949	0.007	34	KENNEDY_A	7.619	0.070
10	POTTERAT_J	8.899	0.027	35	GARLAND_S	7.606	0.087
11	MORRIS_M	8.861	0.025	36	ZHANG_D	7.586	0.041
12	KAZIENKO_P	8.802	0.067	37	NOWAK_M	7.554	0.021
13	SHEN_H	8.799	0.049	38	MAGLIANO_L	7.542	0.051
14	LIU_J	8.763	0.034	39	BONACICH_P	7.540	0.020
15	XU_Q	8.667	0.085	40	LU_R	7.458	0.045
16	TUCKER_J	8.496	0.061	41	WANG_J	7.414	0.023
17	SKVORETZ_J	8.481	0.056	42	WANG_L	7.345	0.038
18	THAI_M	8.453	0.068	43	SAITO_K	7.335	0.055
19	BATAGELJ_V	8.421	0.032	44	CHEN_W	7.245	0.012
20	MUTH_S	8.382	0.026	45	FERRARI_E	7.209	0.038
21	MARTINEZ_M	8.313	**0.141**	46	COHEN_S	7.202	0.010
22	LITWIN_H	8.297	0.052	47	RYAN_L	7.193	0.080
23	STANTON_N	8.250	**0.216**	48	MEYBODI_M	7.175	**0.314**
24	TUREL_O	8.227	**0.161**	49	KIMURA_M	7.139	0.053
25	ABDELZAH_T	8.176	0.085	50	KIM_H	7.121	0.037

Islands

Citation among journals

from WJ and Cite nets

```
CiteJ = t(WJsr) * CiteR * WJsr

CiteJn = t(WJsr) * n(CiteR) * WJsr
```

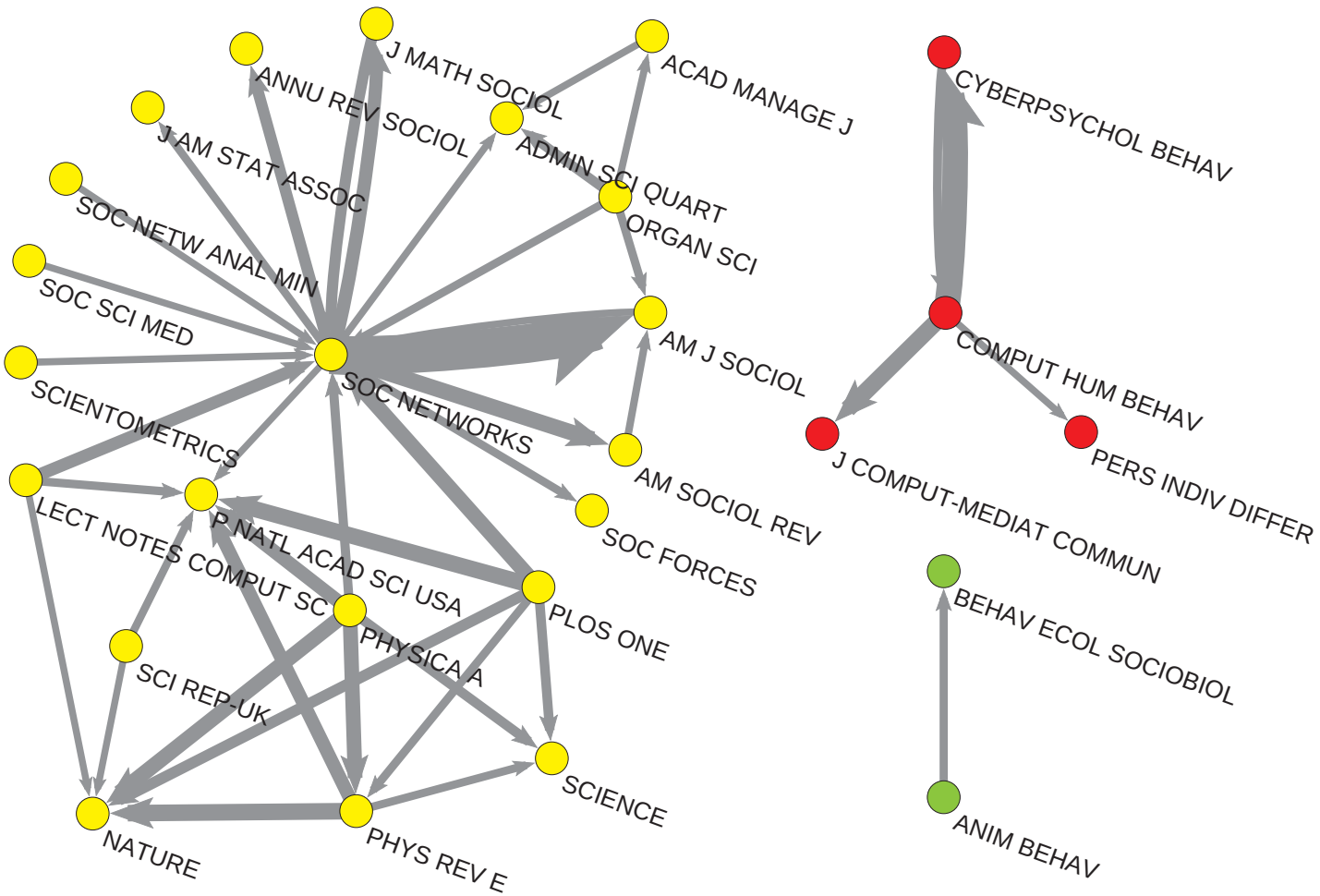
Citation among journals

self-citation from CiteJ net

Rank	Value	Id
1	4443	SOC NETWORKS
2	2058	COMPUT HUM BEHAV
3	569	PHYSICA A
4	429	PHYS REV E
5	382	LECT NOTES COMPUT SC
6	339	CYBERPSYCHOL BEHAV
7	328	SOC SCI MED
8	315	AM J SOCIOL
9	303	PLOS ONE
10	258	ANIM BEHAV
11	246	SCIENTOMETRICS
12	232	J MED INTERNET RES
13	226	P NATL ACAD SCI USA
14	209	ORGAN SCI
15	194	BEHAV ECOL SOCIOBIOL

Citation among journals

authors with largest line weights from CiteJ net



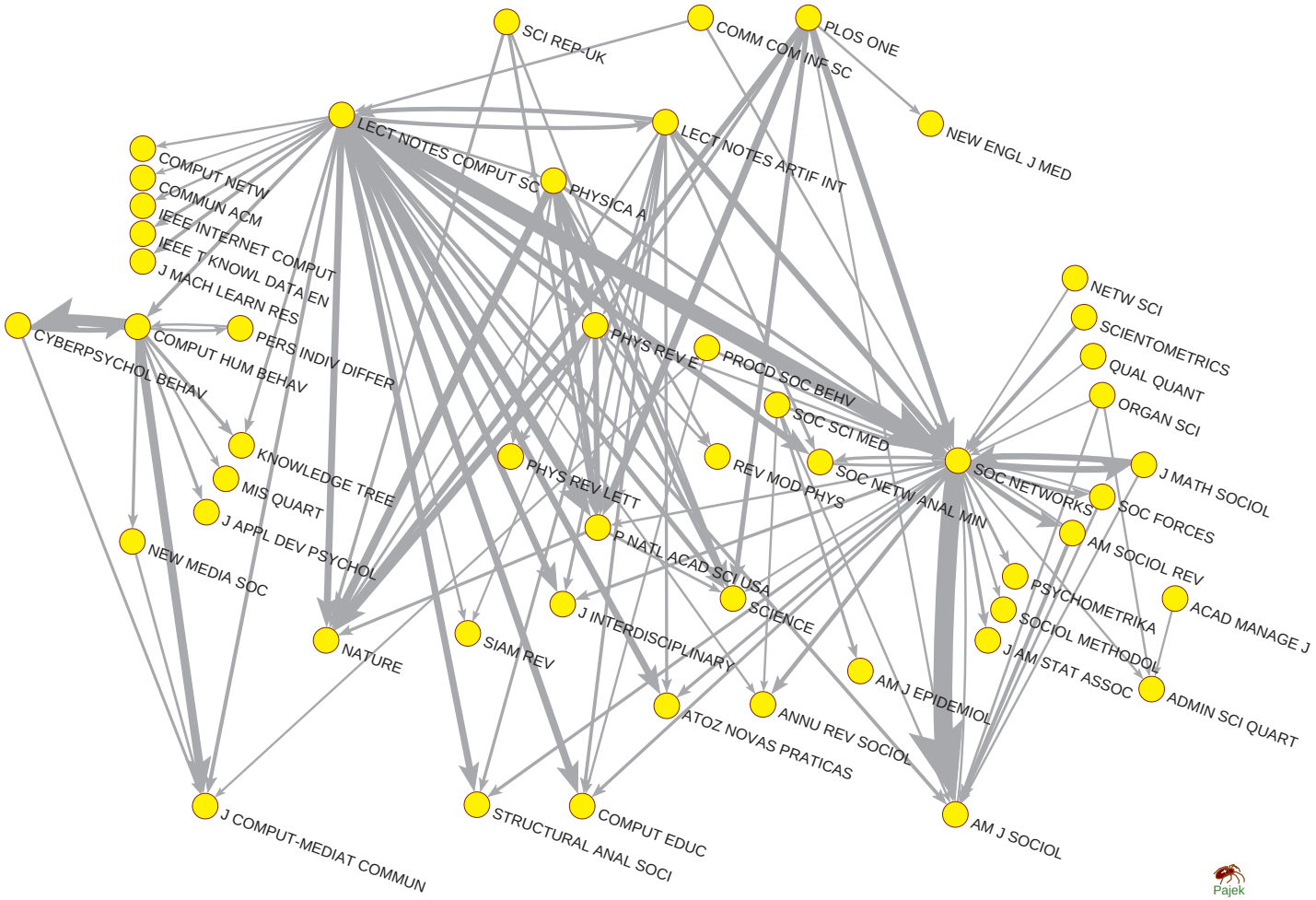
Citation among journals

self-citation from CiteJn net

#	Value	%	Journal	#	Value	%	Journal
1	**355.65**	**0.34**	SOC NETWORKS	16	18.35	0.17	ANIM BEHAV
2	**168.39**	0.22	COMPUT HUM BEHAV	17	17.03	0.12	AIDS BEHAV
3	**122.57**	0.09	LECT NOTES COMPUT SC	18	16.03	0.19	AM J COMMUN PSYCHO
4	57.75	0.13	PHYSICA A	19	14.87	0.10	INFORM SCI
5	43.00	0.14	SOC SCI MED	20	14.14	0.14	KNOWL-BASED SYST
6	42.18	**0.24**	J MED INTERNET RES	21	12.64	0.19	PROF INFORM
7	41.49	0.21	CYBERPSYCHOL BEHAV	22	12.35	**0.23**	COMUNICAR
8	33.16	0.05	PLOS ONE	23	12.00	0.18	BEHAV ECOL SOCIOBI
9	32.93	0.11	PHYS REV E	24	11.87	**0.25**	AM J EPIDEMIOL
10	30.22	0.13	SCIENTOMETRICS	25	11.01	0.11	DECIS SUPPORT SYST
11	24.16	0.14	P NATL ACAD SCI USA	26	10.58	0.14	J ETHN MIGR STUD
12	23.15	**0.26**	AM J SOCIOL	27	10.43	0.13	COMPUT EDUC
13	20.04	0.05	LECT NOTES ARTIF INT	28	10.31	0.18	SEX TRANSM DIS
14	19.31	0.12	EXPERT SYST APPL	29	10.19	**0.28**	NATURE
15	18.77	0.14	NEW MEDIA SOC	30	9.85	0.09	ORGAN SCI

Citation among journals

Main island from CiteJn net



Citation among journals

Pairs of journals with largest line weights from CiteJn net

#	value	from journal	to journal
1	8,1	IEEE GLOB COMM CONF	IEEE INFOCOM SER
2	6,26	HIST COMUN SOC	COMUNICAR
3	4,63	J YOUTH ADOLESCENCE	J RES ADOLESCENCE
4	4,44	INT J GERIATR PSYCH	J PSYCHIAT RES
5	4,38	INT MIGR	INT MIGR REV
6	4,31	J BUS ETHICS	ACAD MANAGE REV
7	3,99	DEMOGR RES	DEMOGRAPHY
8	3	J INTELL FUZZY SYST	J APPL MATHE COMPUT
9	3	J INT DEV	TROP MED INT HEALTH
10	3	PERVASIVE MOB COMPUT	INT CONF PERVAS COMP
11	2,78	J CONSTR ENG M	J CONSTR ENG M ASCE
12	2,68	PHYS EDUC RES CONF	PHYS REV SPEC TOP-PH
13	2,59	ENERGY RES SOC SCI	ENERG POLICY
14	2,5	INT P ECON DEV RES	TECHNOVATION
15	2,37	COMPUT ASSIST LANG L	LANG LEARN TECHNOL
16	2,33	INFORM SOC-ESTUD	PERSPECT CIENC INF
17	2,33	WORLD DEV	ECON J
18	2,31	J PEACE RES	J CONFLICT RESOLUT
19	2,22	HEALTH RES POLICY SY	HEALTH POLICY PLANN
20	2,1	SEX HEALTH	INT J STD AIDS
21	2	REV LAT COMUN SOC	PALABRA CLAVE
22	2	J RETAIL CONSUM SERV	AUSTRALAS MARK J
23	2	ETHN DIS	HEART LUNG
24	2	IEEE INT SYMP INFO	IEEE T INFORM THEORY
25	2	REV BRAS ENFERM	REV LAT-AM ENFERM

Bibliographic coupling

from Cite and WJ / WA nets

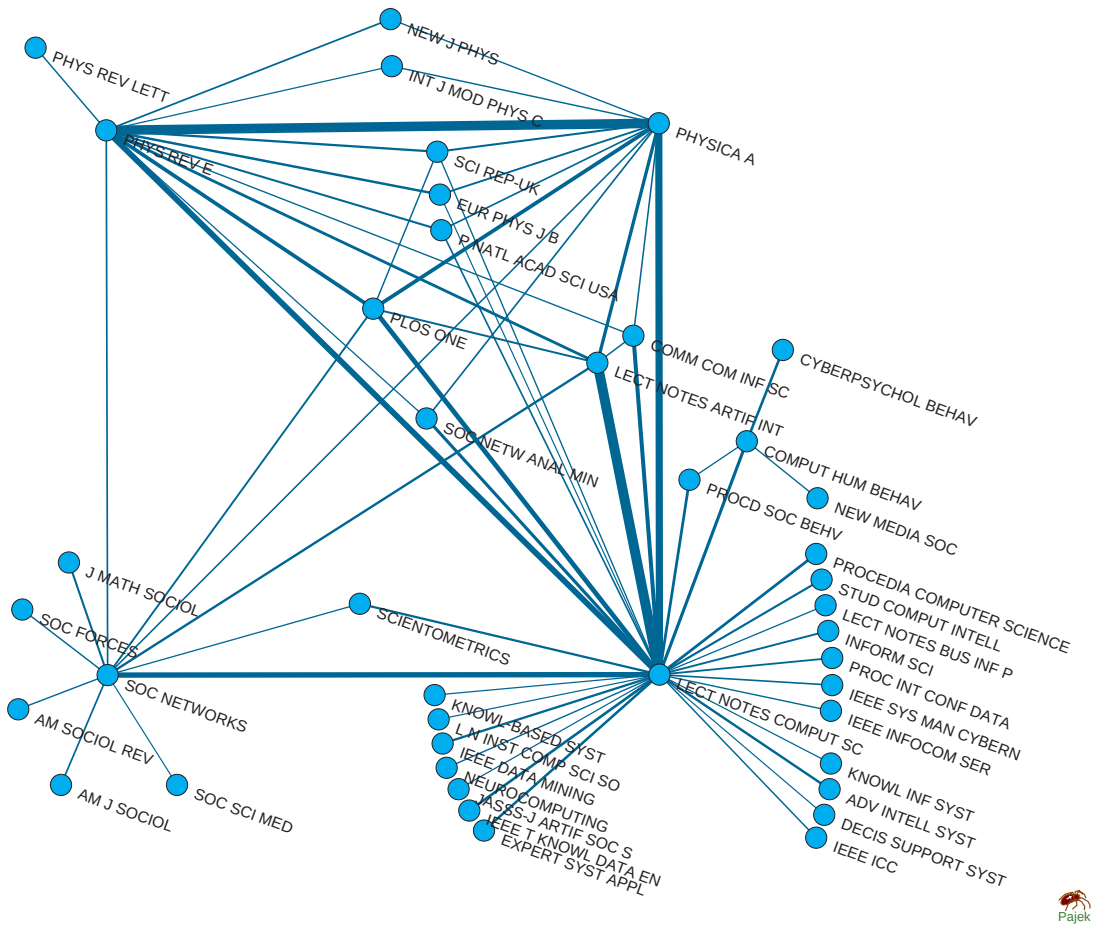
```
biCo = Cite * (Cite)T

JCoj = n(WJsr)T * biCoj * n(WJsr)

ACoj = n(WAsr)T * biCoj * n(WAsr)
```

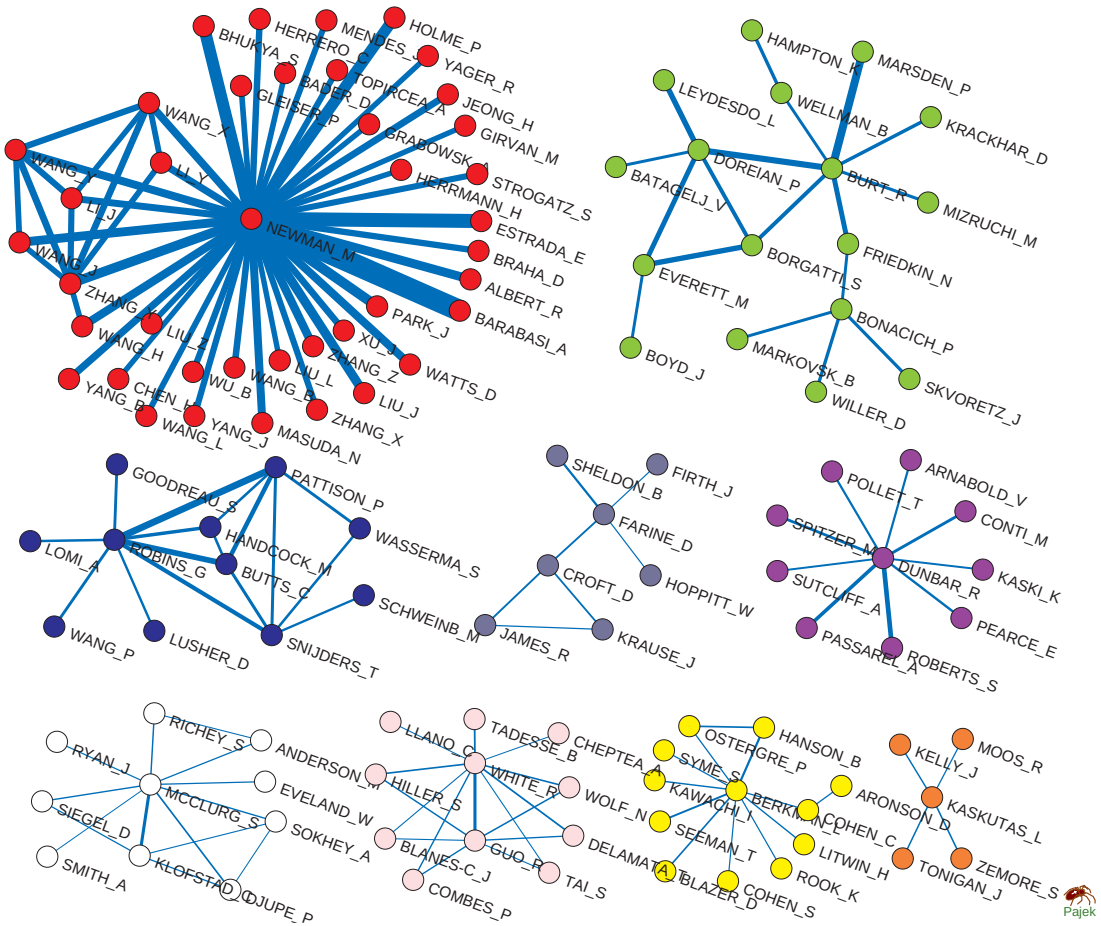
Bibliographic coupling

Co-citation among journals



Bibliographic coupling

Co-citation among authors



Bibliography

1. Batagelj, V. (2007) WoS2Pajek. Networks from Web of Science. Version 0.3. Manual. URL: http://vlado.fmf.uni-lj.si/pub/networks/pajek/WoS2Pajek/WoS2Pajek.pdf
2. Batagelj V., Cerinšek M.(2013). On bibliographic networks. Scientometrics. 96 (3), 845-864
3. Batagelj, V., Doreian P., V., Ferligoj, A., Kejzar N. Understanding Large Temporal Networks and Spatial Networks: Exploration, Pattern Searching, Visualization and Network Evolution, 2014.
4. Batagelj, V., Ferligoj, A. Squazzoni, F. The emergence of a field: a network analysis of research on peer review. Scientometrics (2017) 113: 503. https://doi.org/10.1007/s11192-017-2522-8