Networks of Scientific Collaborations and Citations

Petter Holme

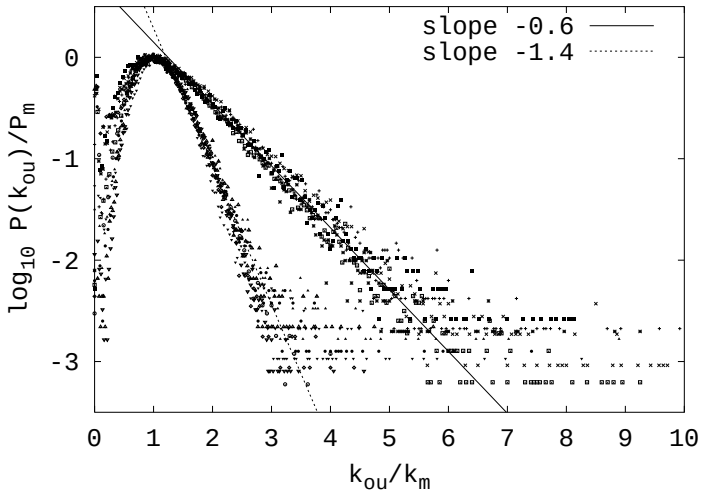
11th May 2001

References:

[1.] M. E. J. Newman, *Who is the best connected scientist? A study of scientific coauthorship networks*, cond-mat/0011144.
[2.] A. L. Barabási *et al.*, *Evolution of the network of scientific collaborations*, cond-mat/0104162.
[3.] A. Vazquez, *Statistics of citation networks*, cond-mat/0105031.
[4.] A. Vazquez, *Knowing a network by walking on it: Emergence of Scaling*, cond-mat/0006132.

Citation Networks

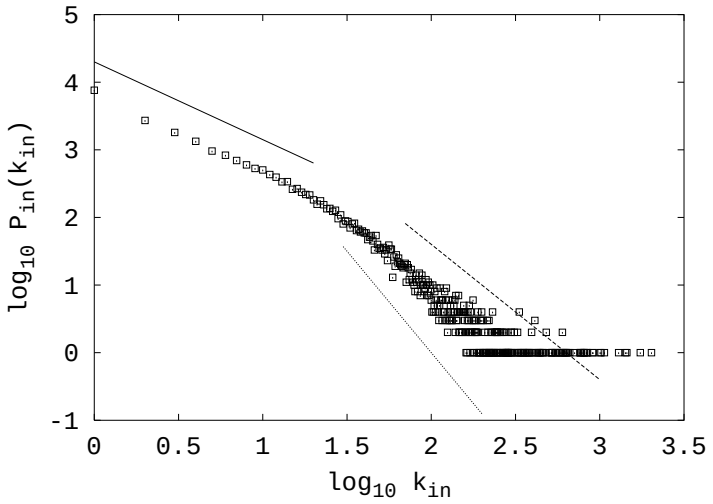
Out Degree



The connectivity distribution of one paper's citations to other papers.

Large k power-law with non-integer exponent.

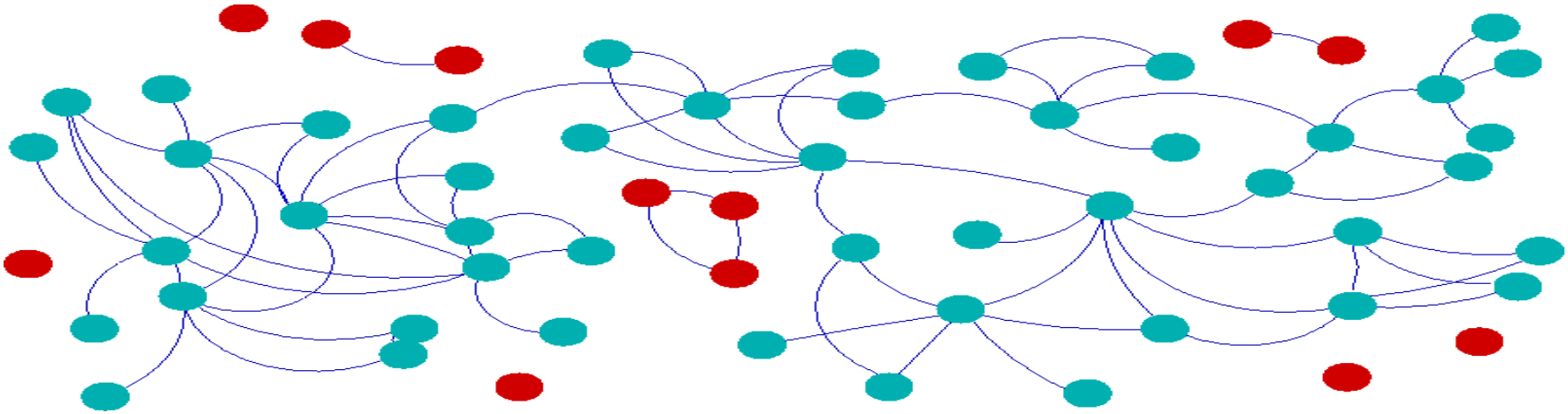
In Degree



The connectivity distribution of citations from other papers to a specific paper.

A slope 2 $\rightarrow$ slope 3 power-law.

Fact and Fiction about the Giant Component



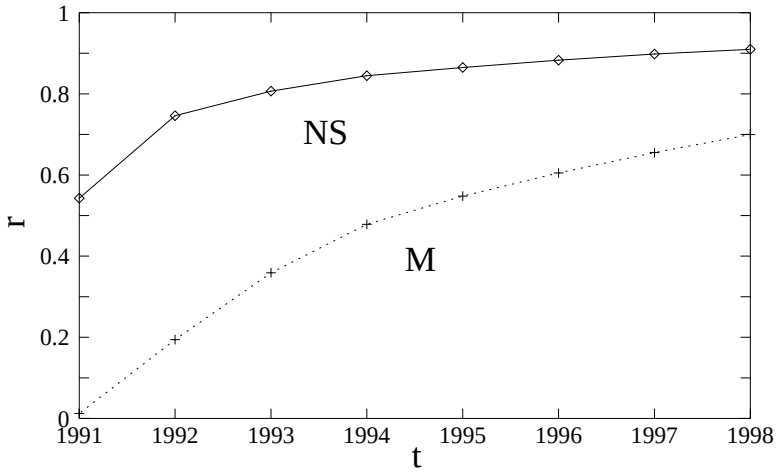
Random graph theory (ER graphs and percolation on connected graphs):

- Random graph theory: There is 2°-phase transition (for increasing number of edges) to a state with an extensively growing largest cluster, a giant component.

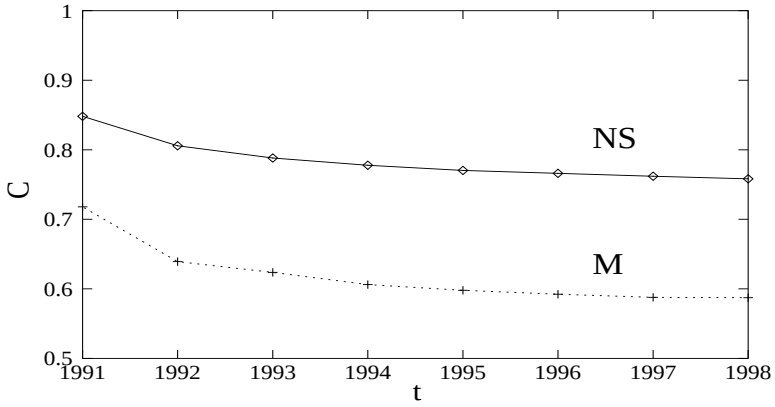
Real networks:

- Are disconnected.
- A vast majority belongs to one connected giant component.
- Characteristic lengths and clustering coefficients are calculated within the giant component.

Giant Component



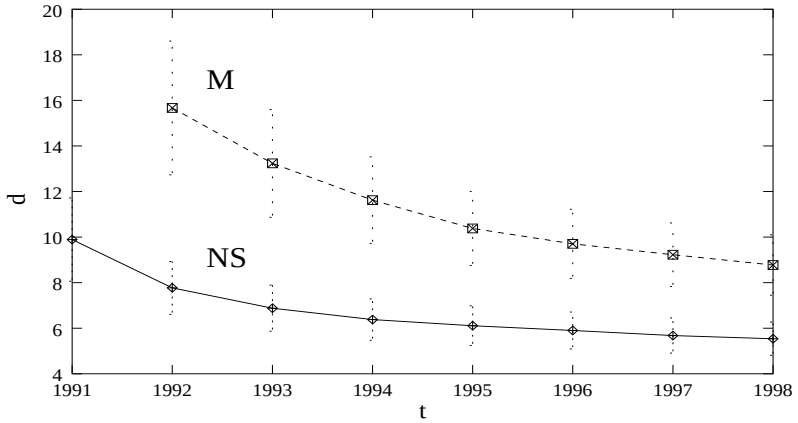
Clustering Coefficient



Cumulative data up to indicated year.

- Large k power-law tail with slope ~ 2.
- Increasing average connectivity.

Average Shortest Distance

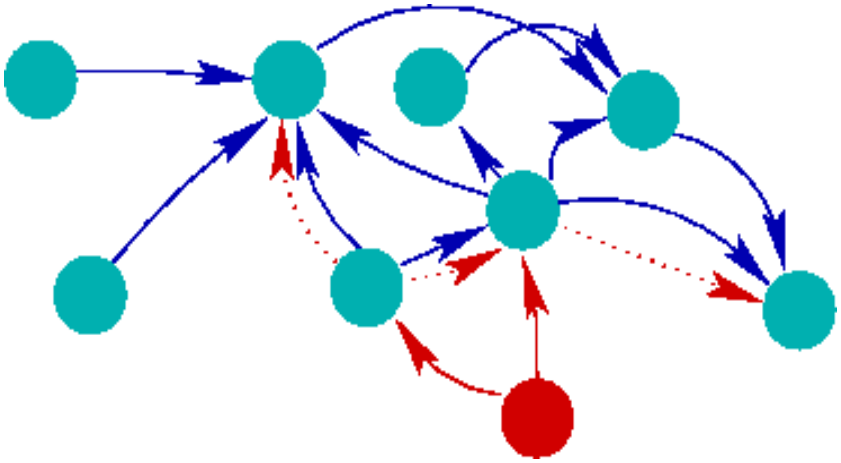


New Scale Free Network Models

Recursive Search Model

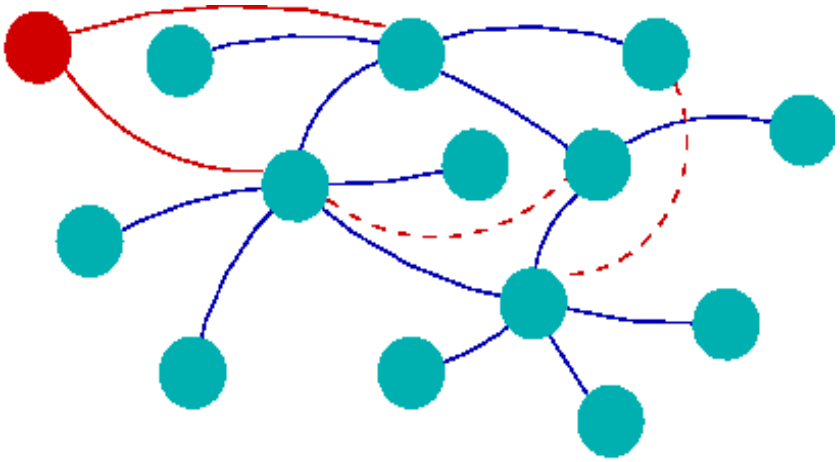
Directed graph.

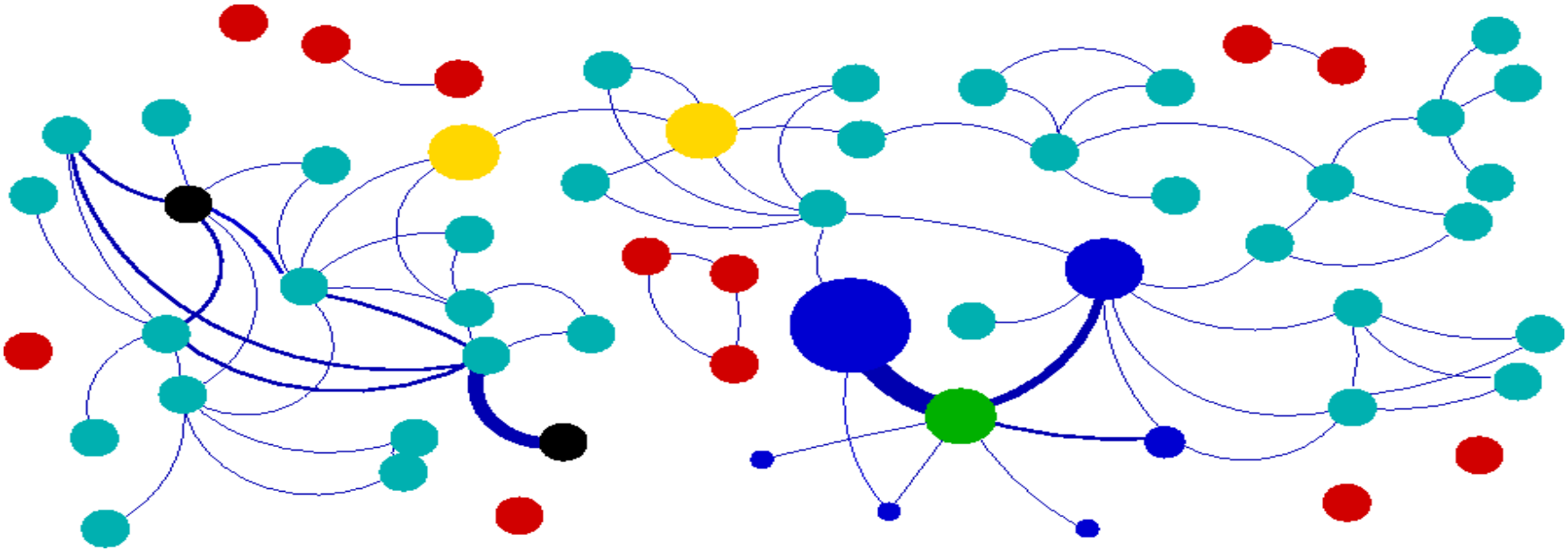
- *Initial condition.* One node, no links.
- *Adding rule.* Add one node, and a link to another (randomly selected).
- *Walking rule.* The new node "walks" through the nodes pointed to by the selected node and creates links to them with probability p.



Generalized Scale Free Model

- *Initial condition.* m nodes, no links.
- *Preferential attachment.* A node is added and linked to another with a probability proportional to the connectivity of the other node.
- *Intra-linking.* More nodes are added between nodes in the graph with a probability proportional to the product of the nodes connectivities.





Betweenness

The total number of shortest paths between all nodes through i defines the betweenness of i. If many shortest paths exists the weight is split between the nodes.

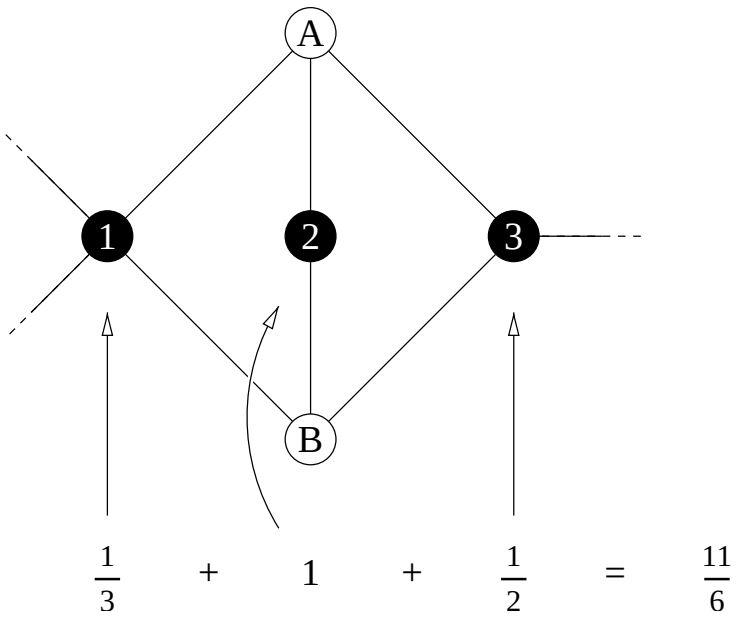
Points of high betweenness makes the average shortest distance much longer if removed.

Funneling

Shortest paths from a particular node passes through only a few neighbors.

Weighted Collaboration Networks

- More authors → weaker ties. (Assume tie strength $\propto 1/(n-1)$, where n is the number of authors.)
- Many papers together → stronger ties. (Assume tie strength $\propto N$, the number of coauthored papers.)



The weight, w_{ij}, of a collaboration tie:

$$w_{ij} = \sum_k \frac{\delta_i^k \delta_j^k}{n_k - 1}$$

Weighted Betweenness

Sheer number of collaborators is no longer necessary for being well connected.

Strong connections to other highly connected collaborators also scores highly.

Hi-Score Lists

number of papers		number of co-workers		betweenness ($\times 10^6$)		collaboration weight	
112	Fabian, A.C.	360	Frontera, F.	2.33	Kouveliotou, C.	16.5	Moskalenko, I.V./Strong, A.W.
101	van Paradijs, J.	353	Kouveliotou, C.	2.15	van Paradijs, J.	15.0	Hernquist, L./Heyl, J.S.
81	Frontera, F.	329	van Paradijs, J.	1.80	Filippenko, A.V.	14.0	Mathews, W.G./Brighenti, F.
80	Hernquist, L.	299	Piro, L.	1.57	Beaulieu, J.P.	13.4	Labini, F.S./Pietronero, L.
79	Gould, A.	296	Costa, E.	1.52	Nomoto, K.	12.2	Piran, T./Sari, R.
78	Silk, J.	291	Feroci, M.	1.52	Pian, E.	11.8	Zaldarriaga, M./Seljak, U.
78	Klis, M.V.D.	284	Pian, E.	1.49	Frontera, F.	11.4	Hernquist, L./Katz, N.
73	Kouveliotou, C.	284	Hurley, K.	1.35	Silk, J.	11.1	Avila-Reese, V./Firmani, C.
70	Ghisellini, G.	244	Palazzi, E.	1.33	Kamionkowski, M.	10.9	Dai, Z.G./Lu, T.
66	Piro, L.	244	Heise, J.	1.28	McMahon, R.G.	10.8	Ostriker, J.P./Cen, R.
116	Parisi, G.	107	Uchida, S.	4.11	MacDonald, A.H.	22.3	Belitz, D./Kirkpatrick, T.R.
79	Scheffler, M.	103	Ueda, Y.	3.96	Bishop, A.R.	17.0	Shrock, R./Tsai, S.
75	Das Sarma, S.	96	Revcolevschi, A.	3.36	Das Sarma, S.	15.0	Yukalov, V.I./Yukalova, E.P.
74	Stanley, H.E.	94	Eisaki, H.	2.96	Tosatti, E.	14.7	Martín-Delgado, M.A./Sierra, G.
70	MacDonald, A.H.	84	Cheong, S.	2.52	Wang, X.	14.3	Krapivsky, P.L./Ben-Naim, E.
68	Sornette, D.	83	Isobe, M.	2.38	Revcolevschi, A.	14.1	Beenakker, C.W.J./Brouwer, P.W.
60	Volovik, G.E.	78	Stanley, H.E.	2.30	Uchida, S.	13.8	Weng, Z.Y./Sheng, D.N.
56	Beenakker, C.W.J.	76	Shirane, G.	2.21	Sigrist, M.	13.7	Sornette, D./Johansen, A.
53	Dagotto, E.	76	Scheffler, M.	2.19	Cheong, S.	13.6	Rikvold, P.A./Novotny, M.A.
50	Helbing, D.	76	Menovsky, A.A.	2.18	Stanley, H.E.	13.0	Scalapino, D.J./White, S.R.
78	Odintsov, S.D.	50	Ambjorn, J.	0.98	Odintsov, S.D.	34.0	Lu, H./Pope, C.N.
73	Lu, H.	44	Ferrara, S.	0.88	Ambjorn, J.	29.0	Odintsov, S.D./Nojiri, S.
72	Pope, C.N.	43	Vafa, C.	0.88	Kogan, I.I.	18.7	Lee, H.W./Myung, Y.S.
69	Cvetic, M.	39	Odintsov, S.D.	0.84	Henneaux, M.	18.3	Schweigert, C./Fuchs, J.
68	Ferrara, S.	39	Kogan, I.I.	0.73	Douglas, M.R.	14.7	Ovrut, B.A./Waldram, D.
65	Vafa, C.	36	Proeyen, A.V.	0.67	Ferrara, S.	14.7	Kleihaus, B./Kunz, J.
65	Tseytlin, A.A.	35	Fre, P.	0.63	Vafa, C.	12.9	Mavromatos, N.E./Ellis, J.
65	Mavromatos, N.E.	35	Ellis, J.	0.60	Khare, A.	12.4	Kachru, S./Silverstein, E.
63	Witten, E.	35	Douglas, M.R.	0.58	Tseytlin, A.A.	11.7	Kakushadze, Z./Tye, S.H.H.
54	Townsend, P.K.	34	Lu, H.	0.58	Townsend, P.K.	11.6	Arefeva, I.Y./Volovich, I.V.

rank	name	co-workers	papers
1	Rees, M. J.	31	36
2	Miralda-Escude, J.	36	34
3	Fabian, A. C.	156	112
4	Waxman, E.	15	30
5	Celotti, A.	119	45
6	Narayan, R.	65	58
7	Loeb, A.	33	64
8	Reynolds, C. S.	45	38
9	Hernquist, L.	62	80
10	Gould, A.	76	79
1	Fisher, M. P. A.	21	35
2	Balents, L.	24	29
3	MacDonald, A. H.	64	70
4	Senthil, T.	9	13
5	Das Sarma, S.	51	75
6	Millis, A. J.	43	37
7	Ioffe, L. B.	16	27
8	Sachdev, S.	28	44
9	Lee, P. A.	24	34
10	Jungwirth, T.	27	17
1	Cvetic, M.	33	69
2	Behrndt, K.	22	41
3	Tseytlin, A. A.	22	65
4	Bergshoeff, E.	21	39
5	Youm, D.	3	30
6	Lu, H.	34	73
7	Klebanov, I. R.	29	47
8	Townsend, P. K.	31	54
9	Pope, C. N.	33	72
10	Larsen, F.	11	27