Fractional approach to derived networks

Vladimir Batagelj

IMFM Ljubljana and IAM UP Koper

1274. Sredin seminar
Ljubljana, 29. March 2017

Outline

Names of Participants of Group I	Code Numbers and Dates of Social Events Reported in *Old City Herald* (1) 6/27	(2) 3/2	(3) 4/12	(4) 9/26	(5) 2/25	(6) 5/19	(7) 3/15	(8) 9/16	(9) 4/8	(10) 6/10	(11) 2/23	(12) 4/7	(13) 11/21	(14) 8/3
1. Mrs. Evelyn Jefferson	×	×	×	×	×	×		×	×					
2. Miss Laura Mandeville	×	×	×		×	×	×	×						
3. Miss Theresa Anderson		×	×	×	×	×	×	×	×					
4. Miss Brenda Rogers	×		×	×	×	×	×	×						
5. Miss Charlotte McDowd			×	×	×		×							
6. Miss Frances Anderson			×		×	×		×						
7. Miss Eleanor Nye					×	×	×	×						
8. Miss Pearl Oglethorpe						×		×	×					
9. Miss Ruth DeSand					×		×	×	×					
10. Miss Verne Sanderson							×	×	×			×		
11. Miss Myra Liddell								×	×	×		×		
12. Miss Katherine Rogers								×	×	×		×	×	×
13. Mrs. Sylvia Avondale							×	×	×	×		×	×	×
14. Mrs. Nora Fayette						×	×		×	×	×	×	×	×
15. Mrs. Helen Lloyd							×	×		×	×	×		
16. Mrs. Dorothy Murchison								×	×					
17. Mrs. Olivia Carleton									×		×			
18. Mrs. Flora Price									×		×			

Vladimir Batagelj: vladimir.batagelj@fmf.uni-lj.si

Version (April 21, 2017, 00 : 10): sreda1274.pdf

Recent papers on fractional approach

- Perianes-Rodriguez, A., Waltman, L., Van Eck, N.J. (2016). Constructing bibliometric networks: A comparison between full and fractional counting. Journal of Informetrics, 10(4), 1178-1195. paper
- Loet Leydesdorff, Han Woo Park (2016). Full and Fractional Counting in Bibliometric Networks. Journal of Informetrics Volume 11, Issue 1, February 2017, Pages 117–120. arXiv, paper
- Gangan Prathap, Somenath Mukherjee (2016). A conservation rule for constructing bibliometric network matrices. arXiv
- Batagelj, V, Cerinšek, M: On bibliographic networks. Scientometrics 96 (2013) 3, 845-864. paper
- Cerinšek, M., Batagelj, V.: Network analysis of Zentralblatt MATH data. Scientometrics, 102(2015)1, 977-1001. paper

Bibliographic networks

From special bibliographies (BibTEX) and bibliographic services (Web of Science, Scopus, SICRIS, CiteSeer, Zentralblatt MATH, Google Scholar, DBLP Bibliography, US patent office, IMDb, and others) we can derive some two-mode networks on selected topics:

works $\times$ authors (**WA**),
works $\times$ keywords (**WK**);
works $\times$ journals/publishers (**WJ**);

and from some data also the network

works $\times$ classification (**WC**)

and the one-mode citation network

works $\times$ works (**Ci**);

where works include papers, reports, books, patents etc.

Besides this we get also at least the partition of works by the journal or publisher, the partition of works by the publication year, and the vector of number of pages.

Linked / multi-modal networks

Linked or multi-modal networks are collections of networks over at least two sets of nodes (modes) and consist of some one-mode networks and some two-mode networks linking different modes. For example: modes are Persons and Organizations. Two one-mode networks describe collaboration among Persons and among Organizations. The linking two-mode network describes membership of Persons to different Organizations.
An important approach in analysis of linked networks is the use of derived networks obtained by network multiplication.

- Krackhardt, D., Carley, K.M. 1998. A PCANS Model of Structure in Organization. In Proceedings of the 1998 International Symposium on Command and Control Research and Technology Evidence Based Research: 113-119, Vienna, VA. MetaMatrix, paper
- Kathleen M. Carley (2003). Dynamic Network Analysis. in the Summary of the NRC workshop on Social Network Modeling and Analysis, Ron Breiger and Kathleen M. Carley (Eds.), National Research Council. preprint

MetaMatrix

Carley and Diesner

Meta-Matrix Entities	Agent	Knowledge	Resources	Tasks/ Event	Organizations	Location
Agent	Social network	Knowledge network	Capabilities network	Assignment network	Membership network	Agent location network
Knowledge		Information network	Training network	Knowledge requirement network	Organizational knowledge network	Knowledge location network
Resources			Resource network	Resource requirement Network	Organizational Capability network	Resource location network
Tasks/ Events				Precedence network	Organizational assignment network	Task/Event location network
Organizations					Inter-organizational network	Organizatio nal location network
Location						Proximity network

Multiplication of networks

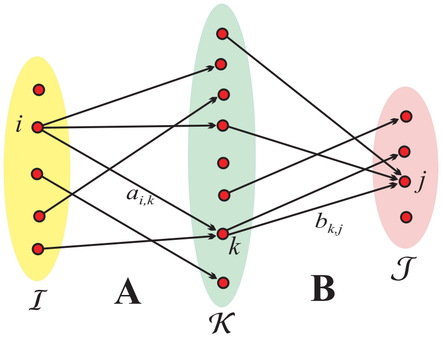
To a simple (no parallel arcs) two-mode *network* $\mathcal{N} = (\mathcal{I}, \mathcal{J}, \mathcal{A}, w)$; where $\mathcal{I}$ and $\mathcal{J}$ are sets of *nodes*, $\mathcal{A}$ is a set of *arcs* linking $\mathcal{I}$ and $\mathcal{J}$, and $w : \mathcal{A} \to \mathbb{R}$ (or some other semiring) is a *weight*; we can assign a *network matrix* $\mathbf{W} = [w_{i,j}]$ with elements: $w_{i,j} = w(i,j)$ for $(i,j) \in \mathcal{A}$ and $w_{i,j} = 0$ otherwise.

Given a pair of compatible networks $\mathcal{N}_A = (\mathcal{I}, \mathcal{K}, \mathcal{A}_A, w_A)$ and $\mathcal{N}_B = (\mathcal{K}, \mathcal{J}, \mathcal{A}_B, w_B)$ with corresponding matrices $\mathbf{A}_{\mathcal{I}\times\mathcal{K}}$ and $\mathbf{B}_{\mathcal{K}\times\mathcal{J}}$ we call a *product of networks* $\mathcal{N}_A$ and $\mathcal{N}_B$ a network $\mathcal{N}_C = (\mathcal{I}, \mathcal{J}, \mathcal{A}_C, w_C)$, where $\mathcal{A}_C = \{(i,j) : i \in \mathcal{I}, j \in \mathcal{J}, c_{i,j} \neq 0\}$ and $w_C(i,j) = c_{i,j}$ for $(i,j) \in \mathcal{A}_C$. The product matrix $\mathbf{C} = [c_{i,j}]_{\mathcal{I}\times\mathcal{J}} = \mathbf{A} * \mathbf{B}$ is defined in the standard way

$$c_{i,j} = \sum_{k\in\mathcal{K}} a_{i,k} \cdot b_{k,j}$$

In the case when $\mathcal{I} = \mathcal{K} = \mathcal{J}$ we are dealing with ordinary one-mode networks (with square matrices).

Multiplication of networks



$$c_{i,j} = \sum_{k \in N_A(i) \cap N_B^-(j)} a_{i,k} \cdot b_{k,j}$$

If all weights in networks $\mathcal{N}_A$ and $\mathcal{N}_B$ are equal to 1 the value of $c_{i,j}$ counts the number of ways we can go from $i \in \mathcal{I}$ to $j \in \mathcal{J}$ passing through $\mathcal{K}$: $c_{i,j} = |N_A(i) \cap N_B^-(j)|$.

Multiplication of networks

The standard matrix multiplication has the complexity $O(|\mathcal{I}| \cdot |\mathcal{K}| \cdot |\mathcal{J}|)$ – it is too slow to be used for large networks. For sparse large networks we can multiply much faster considering only nonzero elements.

For sparse large networks we can multiply much faster considering only nonzero elements.

for k **in** $\mathcal{K}$ **do**
  for (i, j) **in** $N_A^-(k) \times N_B(k)$ **do**
    if $\exists c_{i,j}$ **then** $c_{i,j} := c_{i,j} + a_{i,k} \cdot b_{k,j}$
    else new $c_{i,j} := a_{i,k} \cdot b_{k,j}$

`Networks/Multiply Networks`

In general the multiplication of large sparse networks is a 'dangerous' operation since the result can 'explode' – it is not sparse.

Multiplication of networks – details

If at least one of the sparse networks $\mathcal{N}_A$ and $\mathcal{N}_B$ has small maximal degree on $\mathcal{K}$ then also the resulting product network $\mathcal{N}_C$ is sparse.
If for the sparse networks $\mathcal{N}_A$ and $\mathcal{N}_B$ there are in $\mathcal{K}$ only few nodes with large degree and no one among them with large degree in both networks then also the resulting product network $\mathcal{N}_C$ is sparse.
From the network multiplication algorithm we see that each intermediate node $k \in \mathcal{K}$ adds to a product network a complete two-mode subgraph $K_{N_A^-(k),N_B(k)}$ (or, in the case $\mathcal{I} = \mathcal{J}$, a complete subgraph $K_{N(k)}$). If both degrees $\deg_A(k) = |N_A^-(k)|$ and $\deg_B(k) = |N_B(k)|$ are large then already the computation of this complete subgraph has a quadratic (time and space) complexity – the result 'explodes'.

For more details see the paper.

Two-mode network analysis

by conversion to one-mode network – projections

Often we transform a two-mode network $\mathcal{N} = (\mathcal{U}, \mathcal{V}, \mathcal{E}, w)$ into an ordinary (one-mode) network $\mathcal{N}_1 = (\mathcal{U}, \mathcal{E}_1, w_1)$ or/and $\mathcal{N}_2 = (\mathcal{V}, \mathcal{E}_2, w_2)$, where $\mathcal{E}_1$ and w_1 are determined by the matrix $\mathbf{W}^{(1)} = \mathbf{W}\mathbf{W}^T$, $w^{(1)}_{uv} = \sum_{z\in\mathcal{V}} w_{uz} \cdot w^T_{zv}$. Evidently $w^{(1)}_{uv} = w^{(1)}_{vu}$. There is an edge $(u : v) \in \mathcal{E}_1$ in $\mathcal{N}_1$ iff $N(u) \cap N(v) \neq \emptyset$. Its weight is $w_1(u, v) = w^{(1)}_{uv}$.

The network $\mathcal{N}_2$ is determined in a similar way by the matrix $\mathbf{W}^{(2)} = \mathbf{W}^T\mathbf{W}$.
The networks $\mathcal{N}_1$ and $\mathcal{N}_2$ are analyzed using standard methods.

```
Network/2-Mode Network/2-Mode to 1-Mode/Rows
```

Authorship networks

Let **WA** be the works × authors two mode authorship network; $wa_{pi} \in \{0, 1\}$ is describing the authorship of author i of work p.

$$\forall p \in W : \sum_{i \in A} wa_{pi} = \text{outdeg}_{WA}(p) = \# \text{ authors of work } p$$

Let **N** be its normalized version

$$\forall p \in W : \sum_{i \in A} n_{pi} \in \{0, 1\}$$

obtained from **WA** by $n_{pi} = wa_{pi} / \max(1, \text{outdeg}_{WA}(p))$, or by some other rule determining the author's contribution – the *fractional* approach.

Some transformations of networks

Binarization $b(\mathcal{N})$ is a network obtained from the $\mathcal{N}$ in which all weights are set to 1.

Transposition $\mathcal{N}^T$ or $t(\mathcal{N})$ is a network obtained from $\mathcal{N}$ in which to all arcs their direction is reversed. $\mathbf{AW} = \mathbf{WA}^T$, $\mathbf{KW} = \mathbf{WK}^T$, ...

(Out) normalization $n(\mathcal{N})$ is a network obtained from $\mathcal{N}$ in which the weight of each arc a is divided by the sum of weights of all arcs having the same initial node as the arc a. For binary networks

$$n(\mathbf{A}) = \text{diag}(\frac{1}{\max(1, \text{outdeg}_{WA}(i))})_{i\in\mathcal{I}} * \mathbf{A}$$

$\mathbf{N} = n(\mathbf{WA})$, $\mathbf{WA} = b(\mathbf{N})$

First co-authorship network

$\mathbf{Co} = \mathbf{AW} * \mathbf{WA}$

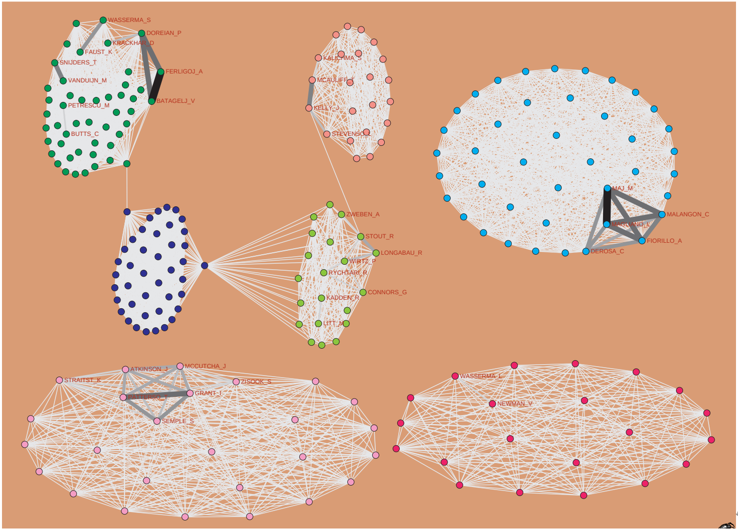
$$co_{ij} = \sum_{p \in W} wa_{pi} wa_{pj} = \sum_{p \in N^-(i) \cap N^-(j)} 1$$

co_{ij} = the number of works that authors i and j wrote together
co_{ii} = the total number of works that author i wrote
It holds: $co_{ij} = co_{ji}$.
Using the weights co_{ij} we can determine the Salton's cosine similarity or Ochiai coefficient between authors i and j as

$$cos(i,j) = \frac{co_{ij}}{\sqrt{co_{ii}co_{jj}}}, \qquad \text{for } co_{ij} > 0$$

Cores of orders 20–47 in **Co**(SN5)

Network SN5 (2008): for "social network*" + most frequent references + around 100 social networkers; $|W| = 193376$, $|C| = 7950$, $|A| = 75930$, $|J| = 14651$, $|K| = 29267$



Papers by number of authors

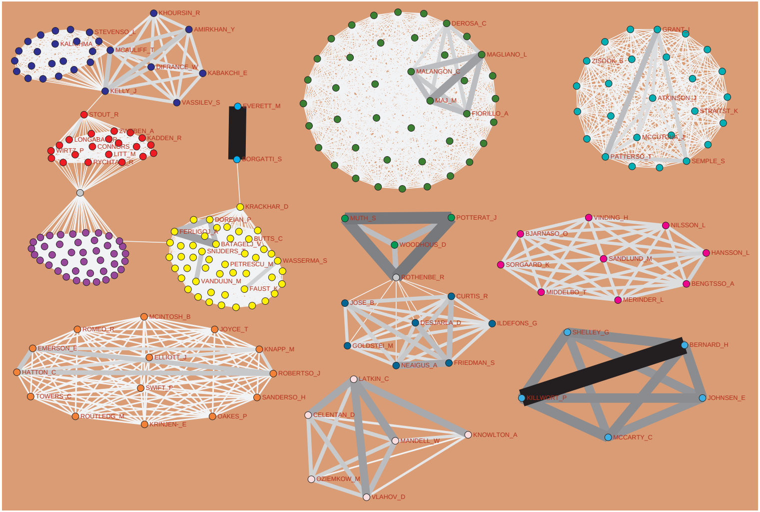
Problem: The **Co** network is composed of complete graphs on the set of work's authors. Works with many authors produce large complete subgraphs and are over-represented, thus bluring the collaboration structure.

outdeg	frequency	outdeg	frequency	paper
1	2637	12	8	
2	2143	13	4	
3	1333	14	3	
4	713	15	2	
5	396	21	1	Pierce et al. (2007)
6	206	22	1	Allen et al. (1998)
7	114	23	1	Kelly et al. (1997)
8	65	26	1	Semple et al. (1993)
9	43	41	1	Magliano et al. (2006)
10	24	42	1	Doll et al. (1992)
11	10	48	1	Snijders et al. (2007)

Snijders et al. (2007)

Snijders et al.(2007): Snijders, T.A.B., Robinson, T., Atkinson, A.C., Riani, M., Gormley, I.C., Murphy, T.B., Sweeting, T., Leslie, D.S., Longford, N.T., Kent, J.T., Lawrance, T., Airoldi, E.M., Besag, J., Blei, D., Fienberg, S.E., Breiger, R., Butts, C.T., Doreian, P., Batagelj, V., Ferligoj, A., Draper, D., van Duijn, M.A.J., Faust, K., Petrescu-Prahova, M., Forster, J.J., Gelman, A., Goodreau, S. M., Greenwood, P.E., Gruenberg, K., Francis, B., Hennig, C., Hoff, P.D., Hunter, D.R., Husmeier, D., Glasbey, C., Krackhardt, D., Kuha, J., Skrondal, A., Lawson, A., Liao, T. F., Mendes, B., Reinert, G., Richardson, S., Lewin, A., Titterington, D.M., Wasserman, S., Werhli, A.V. and Ghazal, P.. *Discussion on the paper by Handcock, Raftery and Tantrum*. Journal of the Royal Statistical Society: Series A - Statistics in Society, 170 (2007), pp. 322-354.

p_S-core at level 20 of **Co**(SN5)



Second co-authorship network

$\mathbf{Cn} = \mathbf{AW} * \mathbf{N}$

$$cn_{ij} = \sum_{p \in W} wa_{pi} n_{pj} = \sum_{p \in N^-(i) \cap N^-(j)} n_{pj}$$

cn_{ij} = contribution of author j to works, that (s)he wrote together with the author i.

It holds $\sum_{j \in A} \sum_{j \in A} wa_{pi} n_{pj} = \text{outdeg}_{WA}(p)$ and $\sum_{j \in A} cn_{ij} = \text{indeg}_{WA}(i)$

$cn_{ii} = \sum_{p \in N(i)} n_{pi}$ is the contribution of author i to his/her works.

Self-sufficiency: $S_i = \frac{cn_{ii}}{\text{indeg}_{WA}(i)}$

Collaborativness: $K_i = 1 - S_i$

$$\sum_{i \in A} \sum_{j \in A} cn_{ij} = \sum_{i \in A} \text{indeg}_{WA}(i) = m_{WA}$$

To compute the table we prepared a macro in `Pajek`.

The "best" authors in Social Networks

i	author	cn_{ii}	total	K_i	i	author	cn_{ii}	total	K_i
1	Burt,R	43.83	53	0.173	26	Latkin,C	10.14	37	0.726
2	Newman,M	36.77	60	0.387	27	Morris,M	9.98	20	0.501
3	Doreian,P	34.44	47	0.267	28	Rothenberg,R	9.82	28	0.649
4	Bonacich,P	30.17	41	0.264	29	Kadushin,C	9.75	11	0.114
5	Marsden,P	29.42	37	0.205	30	Faust,K	9.72	18	0.460
6	Wellman,B	26.87	41	0.345	31	Batagelj,V	9.69	20	0.516
7	Leydesdorf,L	24.37	35	0.304	32	Mizruchi,M	9.67	15	0.356
8	White,H	23.50	33	0.288	33	[Anon]	9.00	9	0.000
9	Friedkin,N	20.00	23	0.130	34	Johnson,J	8.89	21	0.577
10	Borgatti,S	19.20	41	0.532	35	Fararo,T	8.83	16	0.448
11	Everett,M	16.92	31	0.454	36	Lazega,E	8.50	12	0.292
12	Litwin,H	16.00	21	0.238	37	Knoke,D	8.33	11	0.242
13	Freeman,L	15.53	20	0.223	38	Ferligoj,A	8.19	19	0.569
14	Barabasi,A	14.99	35	0.572	39	Brewer,D	8.03	11	0.270
15	Snijders,T	14.99	30	0.500	40	Klovdahl,A	7.96	17	0.532
16	Valente,T	14.80	34	0.565	41	Hammer,M	7.92	10	0.208
17	Breiger,R	14.44	20	0.278	42	White,D	7.83	15	0.478
18	Skvoretz,J	14.43	27	0.466	43	Holme,P	7.42	14	0.470
19	Krackhardt,D	13.65	25	0.454	44	Boyd,J	7.37	13	0.433
20	Carley,K	12.93	28	0.538	45	Kilduff,M	7.25	16	0.547
21	Pattison,P	12.10	27	0.552	46	Small,H	7.00	7	0.000
22	Wasserman,S	11.72	26	0.549	47	Iacobucci,D	7.00	12	0.417
23	Berkman,L	11.21	30	0.626	48	Pappi,F	6.83	10	0.317
24	Moody,J	10.83	15	0.278	49	Chen,C	6.78	12	0.435
25	Scott,J	10.47	15	0.302	50	Seidman,S	6.75	9	0.250

Third co-authorship network

$\mathbf{Ct} = \mathbf{N}^T * \mathbf{N}$

ct_{ij} = the total contribution of 'collaboration' of authors i and j to works.

It holds $ct_{ij} = ct_{ji}$ and

$$\sum_{i \in A} \sum_{j \in A} n_{pi} n_{pj} = 1$$

The total contribution of a complete subgraph corresponding to the authors of a work p is 1.

$\sum_{j \in A} ct_{ij} = \sum_{p \in W} n_{pi}$ = the total contribution of author i to works from W.

$$\sum_{i \in A} \sum_{j \in A} ct_{ij} = |W|$$

Outer product decomposition

$x = [x_1, x_2, \ldots, x_n], \; y = [y_1, y_2, \ldots, y_m]$

$$x \circ y = [x_i \cdot y_j]_{n \times m}$$

$S_x = \sum_i x_i, \qquad S_y = \sum_j y_j$
$S = \sum_{i,j} (x \circ y)_{ij} = \sum_i \sum_j x_i \cdot y_j = \sum_i x_i \cdot \sum_j y_j = S_x \cdot S_y$

$$S_x = S_y = 1 \quad \Rightarrow \quad S = 1$$

$$\mathbf{AK} = \mathbf{AW} * \mathbf{WK} = \sum_w \mathbf{WA}[w, \cdot] \circ \mathbf{WK}[w, \cdot], \qquad \mathbf{AW} = \mathbf{WA}^T$$

$$\mathbf{WA}[w, \cdot] \circ \mathbf{WK}[w, \cdot] = K_{N_{WA}(w), N_{WK}(w)}$$

$x' = x / S_x \quad \Rightarrow \quad S_{x'} = 1$

Example

R code

```
> WA <- rbind(c(1,0,1,0),c(1,1,0,0),c(1,0,1,1),c(0,1,0,1),c(1,0,1,1))
> W <- paste('w',1:5,sep=''); A <- paste('a',1:4,sep='')
> rownames(WA) <- W; colnames(WA) <- A
> WK <- rbind(c(1,1,0,0),c(1,0,1,0),c(0,1,1,1),c(0,0,1,0),c(0,1,0,1))
> K <- paste('k',1:4,sep='')
> rownames(WK) <- W; colnames(WK) <- K

WA  a1  a2  a3  a4     WK  k1  k2  k3  k4
w1   1   0   1   0     w1   1   1   0   0
w2   1   1   0   0     w2   1   0   1   0
w3   1   0   1   1     w3   0   1   1   1
w4   0   1   0   1     w4   0   0   1   0
w5   1   0   1   1     w5   0   1   0   1

> H <- t(WA) %*% WK
> H1 = WA[1,] %o% WK[1,]
> H2 = WA[2,] %o% WK[2,]
> H3 = WA[3,] %o% WK[3,]
> H4 = WA[4,] %o% WK[4,]
> H5 = WA[5,] %o% WK[5,]
> sH <- H1+H2+H3+H4+H5
```

Example

results

WA	a1	a2	a3	a4
w1	1	0	1	0
w2	1	1	0	0
w3	1	0	1	1
w4	0	1	0	1
w5	1	0	1	1

WK	k1	k2	k3	k4
w1	1	1	0	0
w2	1	0	1	0
w3	0	1	1	1
w4	0	0	1	0
w5	0	1	0	1

H	k1	k2	k3	k4
a1	2	3	2	2
a2	1	0	2	0
a3	1	3	1	2
a4	0	2	2	2

sH	k1	k2	k3	k4
a1	2	3	2	2
a2	1	0	2	0
a3	1	3	1	2
a4	0	2	2	2

H	k1	k2	k3	k4
a1	2	3	2	2
a2	1	0	2	0
a3	1	3	1	2
a4	0	2	2	2

=

H1	k1	k2	k3	k4
a1	1	1	0	0
a2	0	0	0	0
a3	1	1	0	0
a4	0	0	0	0

+

H2	k1	k2	k3	k4
a1	1	0	1	0
a2	1	0	1	0
a3	0	0	0	0
a4	0	0	0	0

+

H3	k1	k2	k3	k4
a1	0	1	1	1
a2	0	0	0	0
a3	0	1	1	1
a4	0	1	1	1

+

H4	k1	k2	k3	k4
a1	0	0	0	0
a2	0	0	1	0
a3	0	0	0	0
a4	0	0	1	0

+

H5	k1	k2	k3	k4
a1	0	1	0	1
a2	0	0	0	0
a3	0	1	0	1
a4	0	1	0	1

Example

fractional results


```
> Frac <- function(k){(WA[k,]/sum(WA[k,])) %o% (WK[k,]/sum(WK[k,]))}
> F <- Frac(1)+Frac(2)+Frac(3)+Frac(4)+Frac(5)
> sum(F)
[1] 5
> F1 <- Frac(1); F5 <- Frac(5)

F1   k1  k2  k3  k4       F5  k1  k2  k3  k4
a1  1/4 1/4   0   0       a1   0 1/6   0 1/6
a2    0   0   0   0       a2   0   0   0   0
a3  1/4 1/4   0   0       a3   0 1/6   0 1/6
a4    0   0   0   0       a4   0 1/6   0 1/6


> wkr <- apply(WK, 1, sum); war <- apply(WA, 1, sum)
> wkr
w1 w2 w3 w4 w5
 2  2  3  1  2
```

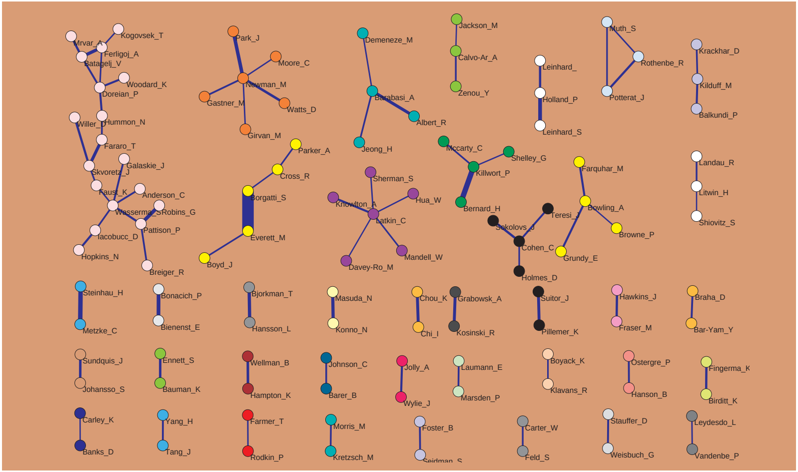
Example

fractional results

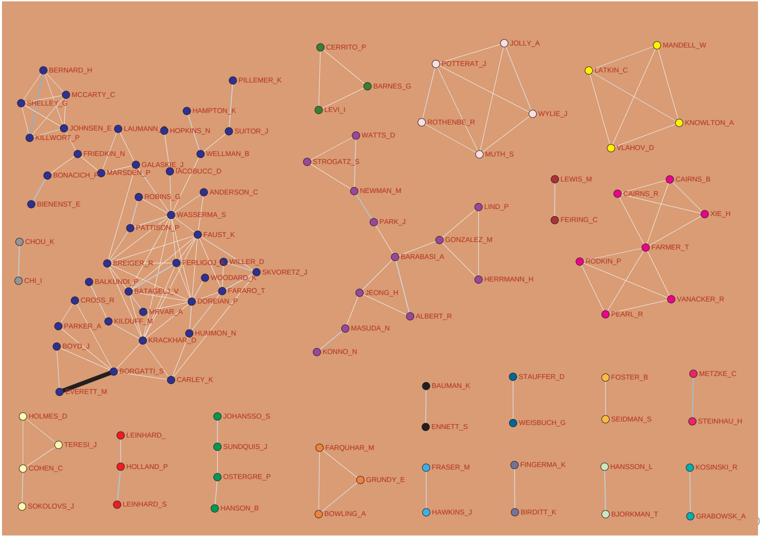
```
> diag(1/wkr)
     [,1] [,2] [,3] [,4] [,5]
[1,]  1/2    0    0    0    0
[2,]    0  1/2    0    0    0
[3,]    0    0  1/3    0    0
[4,]    0    0    0    1    0
[5,]    0    0    0    0  1/2
> WKn <- diag(1/wkr) %*% WK; WAn <- diag(1/war) %*% WA

WAn    a1  a2  a3  a4      WKn   k1  k2  k3  k4
[1,] 1/2   0 1/2   0       [1,] 1/2 1/2   0   0
[2,] 1/2 1/2   0   0       [2,] 1/2   0 1/2   0
[3,] 1/3   0 1/3 1/3       [3,]   0 1/3 1/3 1/3
[4,]   0 1/2   0 1/2       [4,]   0   0   1   0
[5,] 1/33  0 1/3 1/3       [5,]   0 1/2   0 1/2

> AKt <- t(WAn) %*% WKn

AKt   k1        k2        k3        k4        F    k1        k2        k3        k4
a1  0.50 0.5277778 0.3611111 0.2777778        a1 0.50 0.5277778 0.3611111 0.2777778
a2  0.25 0.0000000 0.7500000 0.0000000        a2 0.25 0.0000000 0.7500000 0.0000000
a3  0.25 0.5277778 0.1111111 0.2777778        a3 0.25 0.5277778 0.1111111 0.2777778
a4  0.00 0.2777778 0.6111111 0.2777778        a4 0.00 0.2777778 0.6111111 0.2777778

> apply(F, 1, sum)
      a1       a2       a3       a4
1.666667 1.000000 1.166667 1.166667
> apply(F, 2, sum)
       k1        k2        k3        k4
1.0000000 1.3333333 1.8333333 0.8333333
```

Components in **Ct**(SN5) cut at level 0.5

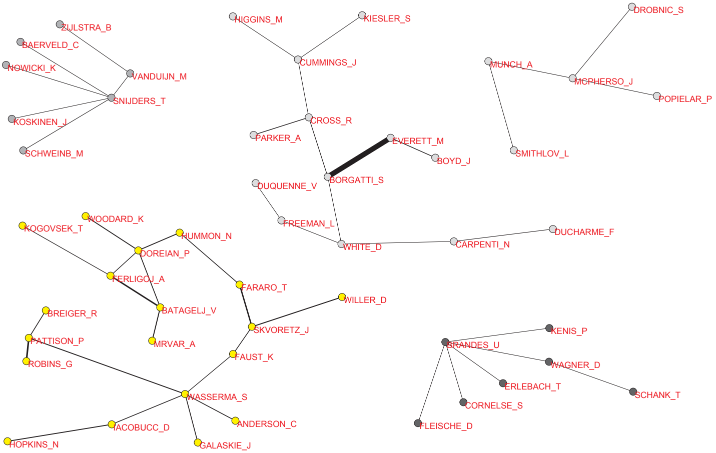
Network SN5 (2008): for "social network*" + most frequent references + around 100 social networkers;
$|W| = 193376$, $|C| = 7950$, $|A| = 75930$, $|J| = 14651$, $|K| = 29267$



p_S-core at level 0.75 in **Ct**(SN5)



Some link islands [5,20] in **Ct**(SN5)



Fourth co-authorship network

$\mathbf{Ct}' = \mathbf{N}^T * \mathbf{N}'$, where $n'_{pi} = wa_{pi} / \max(1, \text{outdeg}_{WA}(p) - 1)$

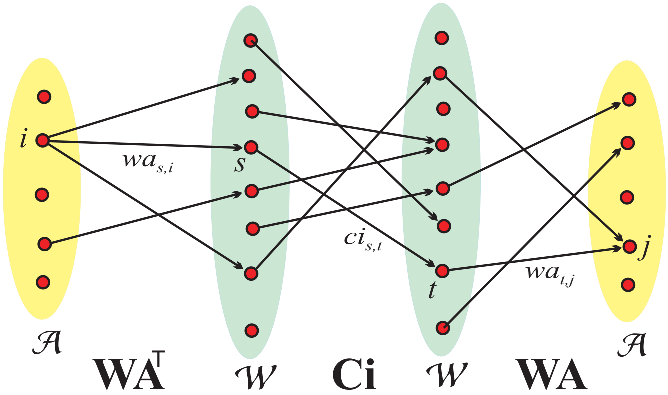
ct'_{ij} = the total contribution of 'strict collaboration' of authors i and j to works.

In Pajek we can use macros to save sequences of commands to produce different co-authorship networks.

The final result is returned as an undirected simple network with weights (for $i \neq j$)

$$ct'_{ij} = \sum_p \frac{2 \cdot wa_{pi} \cdot wa_{pj}}{\max(1, \text{outdeg}_{WA}(p)) \cdot \max(1, \text{outdeg}_{WA}(p) - 1)}$$

Authors' citations network



$\mathbf{Ca} = \mathbf{AW} * \mathbf{Ci} * \mathbf{WA}$ is a network of citations between authors. The weight $w(i,j)$ counts the number of times a work authored by i is citing a work authored by j.

Islands in SN5 authors citation network

Network SN5 (2008): for "social network*" + most frequent references + around 100 social networkers; $|W| = 193376$, $|C| = 7950$, $|A| = 75930$, $|J| = 14651$, $|K| = 29267$