

Derived bibliographic networks

The impact of multi-person units and truncated networks

Vladimir Batagelj

UP FAMNIT Koper and IMFM Ljubljana

1340. sredin seminar

Ljubljana, 22. november 2023

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

- 1 Multiple units
- 2 Fractional approach
- 3 Truncated networks
- 4 Linking through a network
- 5 Conclusions
- 6 References



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

Current version of slides (November 22, 2023 at 03:35): PDF

<https://github.com/bavla/biblio/>

The multipersonality's effect on the results of bibliographic analyses

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

We would like to study the effect of multipersons in derived networks [5]. Let $\mathbf{M} = [m[u, v]]$ is a matrix on $U \times V$ and $\mathbf{C}_U = \{C_1, C_2, \dots, C_k\}$ a partition of the set U , $\emptyset \subset C_i \subseteq U$, $C_i \cap C_j = \emptyset$ for $i \neq j$, and $\bigcup_i C_i = U$. The set U is the (ground truth) set of real units (persons). The partition \mathbf{C}_U corresponds to units (for example authors) identified by the network construction process. A cluster $C \in \mathbf{C}_U$ with $|C| > 1$ represents a multi-unit; and for $|C| = 1$ a correctly identified unit.

We introduce the *shrinking* transformation S of matrix \mathbf{M} by partition \mathbf{C}_U into $S_r(\mathbf{M}, \mathbf{C}_U) = \mathbf{S} = [s[C, v]]$ on $\mathbf{C}_U \times V$ determined by the rule

$$s[C, v] = \sum_{u \in C} m[u, v]$$

The shrinking transformation can be extended to a partition \mathbf{C}_V of the set V by

$$s[u, C] = \sum_{v \in C} m[u, v]$$

Multiunits' effect

or

$$S_c(\mathbf{M}, \mathbf{C}_V) = S_r(\mathbf{M}^T, \mathbf{C}_V)^T$$

and to partitions \mathbf{C}_U and \mathbf{C}_V of both sets by

$$S(\mathbf{M}, (\mathbf{C}_U, \mathbf{C}_V)) = S_c(S_r(\mathbf{M}, \mathbf{C}_U), \mathbf{C}_V)$$

Consider now the case of two compatible matrices $\mathbf{M} = [m[u, t]]$ on $U \times T$ and $\mathbf{N} = [n[t, v]]$ on $T \times V$. For a partition \mathbf{C}_U of the set U it holds

$$S_r(\mathbf{M} \cdot \mathbf{N}, \mathbf{C}_U) = S_r(\mathbf{M}, \mathbf{C}_U) \cdot \mathbf{N}$$

To check this let's denote with \mathbf{L} and \mathbf{R} the left and right sides of this expression. We have

$$l[C, v] = \sum_{u \in C} \mathbf{M} \cdot \mathbf{N}[u, v] = \sum_{u \in C} \sum_{t \in T} m[u, t] \cdot n[t, v]$$

and

$$r[C, v] = \sum_{t \in T} S_r(\mathbf{M}, \mathbf{C}_U)[C, t] \cdot n[t, v] = \sum_{t \in T} \left(\sum_{u \in C} m[u, t] \right) \cdot n[t, v] = l[C, v]$$

Multiunits' effect

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

For the partition \mathbf{C}_V of the set V we get

$$\begin{aligned} S_c(\mathbf{M} \cdot \mathbf{N}, \mathbf{C}_V) &= S_r((\mathbf{M} \cdot \mathbf{N})^T, \mathbf{C}_V)^T = S_r(\mathbf{N}^T \cdot \mathbf{M}^T, \mathbf{C}_V)^T = \\ &= (S_r(\mathbf{N}^T, \mathbf{C}_V) \cdot \mathbf{M}^T)^T = \mathbf{M} \cdot S_r(\mathbf{N}^T, \mathbf{C}_V)^T = \mathbf{M} \cdot S_c(\mathbf{N}, \mathbf{C}_V) \end{aligned}$$

Therefore

$$S_c(\mathbf{M} \cdot \mathbf{N}, \mathbf{C}_V) = \mathbf{M} \cdot S_c(\mathbf{N}, \mathbf{C}_V)$$

For partitions of both sets U and V we have

$$\begin{aligned} S(\mathbf{M} \cdot \mathbf{N}, (\mathbf{C}_U, \mathbf{C}_V)) &= S_c(S_r(\mathbf{M} \cdot \mathbf{N}, \mathbf{C}_U), \mathbf{C}_V) = \\ &= S_c(S_r(\mathbf{M}, \mathbf{C}_U) \cdot \mathbf{N}, \mathbf{C}_V) = S_r(\mathbf{M}, \mathbf{C}_U) \cdot S_c(\mathbf{N}, \mathbf{C}_V) \end{aligned}$$

and finally

$$S(\mathbf{M} \cdot \mathbf{N}, (\mathbf{C}_U, \mathbf{C}_V)) = S_r(\mathbf{M}, \mathbf{C}_U) \cdot S_c(\mathbf{N}, \mathbf{C}_V)$$

Multiunits' effect

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

For $C_u \in \mathbf{C}_U$ and $C_v \in \mathbf{C}_V$ we have

$$S(\mathbf{M} \cdot \mathbf{N}, (\mathbf{C}_U, \mathbf{C}_V))[C_u, C_v] = S_c(S_r(\mathbf{M} \cdot \mathbf{N}, \mathbf{C}_U), \mathbf{C}_V)[C_u, C_v] = \sum_{z \in C_v} S_r(\mathbf{M} \cdot \mathbf{N}, \mathbf{C}_U)[C_u, z] = \sum_{z \in C_v} \sum_{w \in C_u} \mathbf{M} \cdot \mathbf{N}[w, z] = \sum_{w \in C_u} \sum_{z \in C_v} \mathbf{M} \cdot \mathbf{N}[w, z]$$

In a special case of singleton clusters $C_u = \{u\}$ and $C_v = \{v\}$ we get

$$S(\mathbf{M} \cdot \mathbf{N}, (\mathbf{C}_U, \mathbf{C}_V))[\{u\}, \{v\}] = \sum_{w \in \{u\}} \sum_{z \in \{v\}} \mathbf{M} \cdot \mathbf{N}[w, z] = \mathbf{M} \cdot \mathbf{N}[u, v]$$

We see that **the multi-units don't affect the values of relations between singletons in the derived networks.**

Fractional approach

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

Assume that we have the authorship network represented by a matrix $\mathbf{WA} = [wa[w, a]]$ where $wa[w, a] = 1$ iff the author a is (co)author of the work w , and $wa[w, a] = 0$ otherwise.

We will discuss two normalizations. The *standard* normalization $n(\mathbf{WA}) = [nwa[w, a]]$ where

$$nwa[w, a] = \frac{wa[w, a]}{\max(1, \deg(w))}$$

and *strict* or *Newman's* normalization $n'(\mathbf{WA}) = [nwa'[w, a]]$ where

$$nwa'[w, a] = \frac{wa[w, a]}{\max(1, \deg(w) - 1)}$$

We have $\sum_{a \in A} nwa[w, a] = \deg(w)$ and

$$\sum_{a \in A} nwa'[w, a] = \frac{\deg(w)}{\max(1, \deg(w) - 1)}.$$

... Fractional approach

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

The meaning of the weighted degree of node a in the network $n(\mathbf{WA})$, $\text{wdeg}_{n(\mathbf{WA})}(a) = \sum_{w \in W} nwa[w, a]$, is the *fractional contribution of the author* a to all works.

Note that if $\deg(w) = 0$ then both $nwa[w, a] = 0$ and $nwa'[w, a] = 0$, and if $\deg(w) = 1$ then both $nwa[w, a] = 1$ and $nwa'[w, a] = 1$.

The standard co-appearance network matrix $\mathbf{Cn} = [cn[a, b]]$ is obtained as

$$\mathbf{Cn} = n(\mathbf{WA})^T \cdot n(\mathbf{WA})$$

and the *strict* co-appearance network matrix $\mathbf{Ct} = [ct[a, b]]$ is obtained as

$$\mathbf{Ct} = D_0(n(\mathbf{WA})^T \cdot n'(\mathbf{WA}))$$

where $D_0(\mathbf{M})$ sets the diagonal of matrix \mathbf{M} to 0.

Let's look at an entry of \mathbf{Cn}

$$cn[a, b] = \sum_{w \in W} nwa^T[a, w] \cdot nwa[w, b] = \sum_{w \in W} nwa[w, a] \cdot nwa[w, b] = cn[b, a]$$

... Fractional approach

and for **Ct**

$$\begin{aligned} ct[a, b] &= \sum_{w \in W} nwa^T[a, w] \cdot nwa'[w, b] = \sum_{w \in W} \frac{wa[w, a]}{\max(1, \deg(w))} \cdot \frac{wa[w, b]}{\max(1, \deg(w) - 1)} \\ &= \sum_{w \in W} \frac{wa[w, b]}{\max(1, \deg(w))} \cdot \frac{wa[w, a]}{\max(1, \deg(w) - 1)} = ct[b, a] \end{aligned}$$

The fractional co-appearance matrices **Cn** and **Ct** are symmetric.

From

$$\begin{aligned} wdeg_{Cn}(a) &= \sum_{b \in A} cn[a, b] = \sum_{w \in W} nwa[w, a] \cdot \sum_{b \in A} nwa[w, b] = \\ &= \sum_{w \in W} nwa[w, a] \cdot \text{sign}(\deg(w)) = \sum_{w \in W} nwa[w, a] = wdeg_{n(WA)}(a) \end{aligned}$$

we see that the authors have the same weighted degree in networks $n(\mathbf{WA})$ and **Cn**.

... Fractional approach

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

Similarly for the network **Ct**. Because by definition, $ct[a, a] = 0$, we have

$$wdeg_{Ct}(a) = \sum_{b \in A \setminus \{a\}} ct[a, b] = \sum_{w \in W} \frac{nwa[w, a]}{\max(1, \deg(w) - 1)} \sum_{b \in A \setminus \{a\}} wa[w, b] =$$

If $wa[w, a] = 0$ the term in the $\sum_{w \in W}$ has value 0. So we can assume $wa[w, a] = 1$. This means that $a \in N(w)$ and $wa[w, b] = 1$ means that also $b \in N(w)$. Therefore

$$\sum_{b \in A \setminus \{a\}} wa[w, b] = |N(w) \setminus \{a\}| = \deg(w) - 1$$

Now, we can continue

$$= \sum_{w \in W} \frac{nwa[w, a]}{\max(1, \deg(w) - 1)} (\deg(w) - 1) = \sum_{w \in W} nwa[w, a] = wdeg_{n(WA)}(a)$$

The authors have the same weighted degree also in networks $n(\mathbf{WA})$ and **Ct**.

... Fractional approach

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

Package **bibmat**

```
> normalize <- function(M) t(apply(M,1,function(x) x/max(1,sum(x))))
> newman <- function(M) t(apply(M,1,function(x) x/max(1,sum(x)-1)))
> D0 <- function(M) {diag(M) <- 0; return(M)}
> binary <- function(M) {B <- t(apply(M,1,function(x) as.integer(x!=0)))
+   colnames(B) <- colnames(M); return(B)}
> wodeg <- function(M) apply(M,1,sum)
> wideg <- function(M) apply(M,2,sum)
> odeg <- function(M) wodeg(binary(M))
> ideg <- function(M) wideg(binary(M))
> wdeg <- wodeg
```

... Fractional approach

Multiunits and truncated networks

V. Batagelj

Multiple units

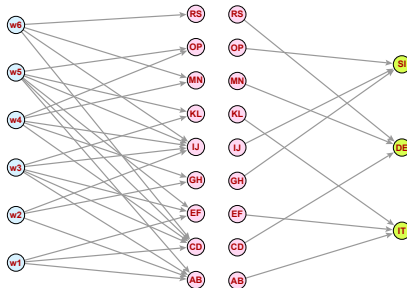
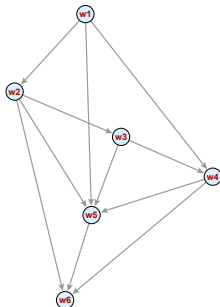
Fractional approach

Truncated networks

Linking through a network

Conclusions

References



```
> Ci
      w1 w2 w3 w4 w5 w6
w1    0  1  0  1  1  0
w2    0  0  1  0  1  1
w3    0  0  0  1  1  0
w4    0  0  0  0  1  1
w5    0  0  0  0  0  1
w6    0  0  0  0  0  0
```

```
> WA
      AB CD EF GH IJ KL MN OP RS
w1    1  1  1  0  0  0  0  0  0
w2    1  1  1  1  1  0  0  0  0
w3    1  1  1  1  1  1  0  0  0
w4    0  1  0  1  1  0  1  1  0
w5    1  1  1  0  1  1  0  1  0
w6    0  1  0  0  1  0  1  0  1
```

```
> AC
      IT DE SI
AB    1  0  0
CD    1  1  0
EF    1  0  0
GH    0  0  1
IJ    0  0  1
KL    1  0  0
MN    1  1  0
OP    0  0  1
RS    0  1  0
```

[Github/bavla/biblio](https://github.com/bavla/biblio)

... Fractional approach

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

```
> wdir <- "C:/Users/vlado/test/biblio"
> setwd(wdir)
> source(
  "https://raw.githubusercontent.com/bavla/biblio/master/code/bibmat.R")
> urlEx <-
  "https://github.com/bavla/biblio/raw/master/Eu/Data/ExNets.RDS"
> download.file(url=urlEx,destfile=paste0(wdir,"/ExNets.RDS",sep=""))
> Ex <- readRDS("ExNets.RDS")
> Ci <- Ex$Ci; WA <- Ex$WA; AC <- Ex$AC

> WAn <- normalize(WA)
> Cn <- t(WAn)%*%WAn
> wideg(WAn)
      AB      CD      EF      GH      IJ      KL      MN      OP      RS
1.0333 1.1500 0.7000 0.5333 1.1500 0.3667 0.4500 0.3667 0.2500
> wideg(Cn)
      AB      CD      EF      GH      IJ      KL      MN      OP      RS
1.0333 1.1500 0.7000 0.5333 1.1500 0.3667 0.4500 0.3667 0.2500
> WAt <- newman(WA)
> Ct <- DO(t(WAn)%*%WAt)
> wideg(Ct)
      AB      CD      EF      GH      IJ      KL      MN      OP      RS
1.0333 1.1500 0.7000 0.5333 1.1500 0.3667 0.4500 0.3667 0.2500
```

For a matrix \mathbf{M} on $U \times V$ we define its *total* $T(\mathbf{M})$ as the sum of all its entries, $T(\mathbf{M}) = \sum_{u \in U} \sum_{v \in V} m[u, v]$. Let's compute

$$T(\mathbf{Cn}) = \sum_{a \in A} \sum_{b \in A} cn[a, b] = \sum_{w \in W} \sum_{a \in A} \sum_{b \in A} nwa[w, a] \cdot nwa[w, b] = \sum_{w \in W} T(w)$$

where

$$\begin{aligned} T(w) &= \sum_{a \in A} \sum_{b \in A} nwa[w, a] \cdot nwa[w, b] = \sum_{a \in A} nwa[w, a] \cdot \sum_{b \in A} nwa[w, b] = \\ &= \text{sign}(\deg(w))^2 = \text{sign}(\deg(w)) \end{aligned}$$

We see that the contribution of each work $w \in W$ with $\deg(w) > 0$ is 1. Therefore

$$T(\mathbf{Cn}) = \sum_{w \in W} \text{sign}(\deg(w)) = |W_1|$$

where $W_1 = \{w \in W : \deg(w) \geq 1\}$.

Totals

For matrices $n(\mathbf{WA})$ and \mathbf{Cn} we have

$$\begin{aligned} T(n(\mathbf{WA})) &= \sum_{a \in A} \sum_{w \in W} nwa[w, a] = \sum_{a \in A} \text{wdeg}_{n(\mathbf{WA})}(a) = \\ &= \sum_{a \in A} \text{wdeg}_{\mathbf{Cn}}(a) = \sum_{a \in A} \sum_{b \in A} cn[a, b] = T(\mathbf{Cn}) \end{aligned}$$

And

$$\begin{aligned} T(\mathbf{Ct}) &= \sum_{a \in A} \sum_{b \in A \setminus \{a\}} ct[a, b] = \\ &= \sum_{w \in W} \sum_{a \in A} \sum_{b \in A \setminus \{a\}} nwa[w, a] \cdot nwa'[w, b] = \sum_{w \in W} T'(w) \end{aligned}$$

where

$$\begin{aligned} T'(w) &= \sum_{a \in A} \sum_{b \in A \setminus \{a\}} nwa[w, a] \cdot nwa'[w, b] = \\ &= \sum_{a \in A} \frac{nwa[w, a]}{\max(1, \deg(w) - 1)} \cdot \sum_{b \in A \setminus \{a\}} wa[w, b] \end{aligned}$$

If $wa[w, a] = 0$ or $wa[w, b] \leq 1$ the term in $\sum_{a \in A}$ has value 0. For $wa[w, a] = 1$ and $wa[w, b] \geq 2$ (or $\deg(w) \geq 2$) we have $\sum_{b \in A \setminus \{a\}} wa[w, b] = \deg(w) - 1$. Therefore

$$T'(w) = \begin{cases} 1 & \deg(w) \geq 2 \\ 0 & \text{otherwise} \end{cases}$$

and finally

$$T(\mathbf{Ct}) = \sum_{w \in W} T'(w) = |W_2|$$

where $W_2 = \{w \in W : \deg(w) \geq 2\}$.

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

```
> sum(Cn)
[1] 6
> sum(Ct)
[1] 6
> empty <- rep(0,length(A))
> WA1 <- rbind(WA,empty,empty)
> rownames(WA1)[7:8] <- c("w7","w8")
> WA1["w8","CD"] <- 1
> WA1
  AB CD EF GH IJ KL MN OP RS
w1 1 1 1 0 0 0 0 0 0
w2 1 0 0 1 1 0 0 0 0
w3 1 1 1 0 1 1 0 0 0
w4 0 1 0 1 1 0 1 1 0
w5 1 1 1 0 1 1 0 1 0
w6 0 1 0 0 1 0 1 0 1
w7 0 0 0 0 0 0 0 0 0
w8 0 1 0 0 0 0 0 0 0
> WAN1 <- normalize(WA1)
> Cn1 <- t(WAN1)%*%WAN1
> sum(Cn1)
[1] 7
> WAt1 <- newman(WA1)
> Ct1 <- DO(t(WAN1)%*%WAt1)
> sum(Ct1)
[1] 6
```

Truncated fractional co-appearance networks

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

A usual approach to the analysis of a two-mode network is its projection to its selected mode. The obtained weighted one-mode network is afterward analyzed using standard methods. Most bibliographic two-mode networks are sparse – they have a small average degree. If the other mode has some nodes of very large degree the projection can "explode" – it is not a sparse network (increased time and space complexity) [3]. In fractional projections, the contribution of the nodes of large degree is very small and mostly doesn't affect the resulting important subnetworks – the important part of the result can be obtained by projection of a two-mode subnetwork on a subset of important nodes – a truncated projection. This idea is elaborated in the following.

Truncated standard fractional network

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

Let's split the set of authors A into two sets A_1 (selected authors) and A_0 (remaining authors), $A_1 \cup A_0 = A$ and $A_1 \cap A_0 = \emptyset$. We call a *truncated (standard) fractional network* the network

$$\mathbf{Cn}_{11} = \mathbf{Cn}[A_1, A_1] = n(\mathbf{WA})[W, A_1]^T \cdot n(\mathbf{WA})[W, A_1]$$

For a selected author $a \in A_1$, we denote with $\alpha(a) = \text{wdeg}_{\mathbf{Cn}_{11}}(a)$ her/his *internal* fractional contribution, and with $\beta(a) = \text{wdeg}_{n(\mathbf{WA})}(a) - \alpha(a)$ her/his *external* fractional contribution.

We reorder the nodes of the network \mathbf{Cn} according to the A_1, A_0 split (see next slide). The network matrix is split into four submatrices $\mathbf{Cn}_{ij}, i, j \in \{0, 1\}$. We denote their totals $T_{ij} = T(\mathbf{Cn}_{ij})$. Because the matrix \mathbf{Cn} is symmetric we have $T_{01} = T_{10}$. T_{11} is the fractional contribution of collaboration among selected authors, $T_{10} + T_{01} = 2T_{10}$ is the fractional contribution of collaboration of selected authors with remaining authors, and T_{00} is the fractional contribution of collaboration among remaining authors.

... Truncated standard fractional network

Multiunits and truncated networks

V. Batagelj

Multiple units

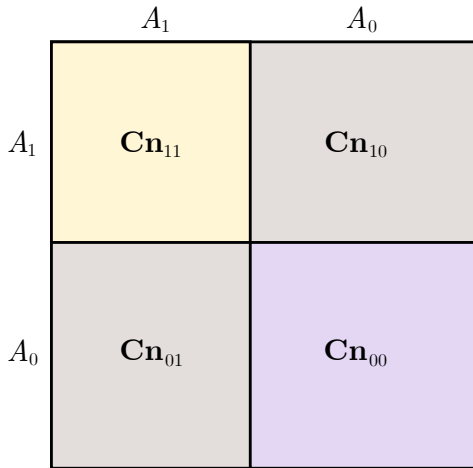
Fractional approach

Truncated networks

Linking through a network

Conclusions

References



... Truncated standard fractional network

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

We can compute $T_{11} = T(\mathbf{Cn}_{11})$ and $T_{10} = \sum_{a \in A_1} \beta(a)$, and finally $T_{00} = T(n(\mathbf{WA})) - T_{11} - 2T_{10}$. Note that we used only information from $n(\mathbf{WA})$ and \mathbf{Cn}_{11} .

A primary application of the truncated standard fractional network scheme is for the set of the most active authors

$$A_1 = \{a \in A : \text{wideg}_{n(\mathbf{WA})}(a) \geq t\}$$

where t is a selected threshold value.

Note that the computation of the vector **wideg** $_{n(\mathbf{WA})}$ is cheap.

Truncated strict fractional network

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

Again we have $\text{wdeg}_{C_t}(a) = \text{wdeg}_{n(WA)}(a)$. Therefore the scheme used for the truncated standard fractional network can be applied also for *truncated strict fractional network*

$$\mathbf{Ct}_{11} = \mathbf{Ct}[A_1, A_1] = D_0(n(\mathbf{WA})[W, A_1]^T \cdot n'(\mathbf{WA})[W, A_1])$$

Sizes of truncated networks

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

iMetrics network [2] and Nataliya's HKUST1 network.

<i>interval</i>	<i>iMetrics</i>			<i>HKUST1</i>		
	<i>n</i>	<i>m</i> = <i>A</i>	<i>avdeg</i>	<i>n</i>	<i>m</i> = <i>A</i>	<i>avdeg</i>
≥ 0	33919	225931	13.32	28108	45365272	3227.92
$\geq 1/10$	32418	191888	11.84	17656	3216796	364.39
$\geq 1/5$	26247	134049	10.21	10213	86529	16.94
$\geq 1/3$	14381	71967	10.01	5171	45845	17.73
$\geq 1/2$	12781	60587	9.48	4032	32806	16.27
≥ 1	6211	32395	10.43	1799	13723	15.26
≥ 2	1832	14900	16.27	689	4195	12.18
≥ 3	964	9306	19.31	369	1743	9.45
≥ 5	446	4646	20.83	172	646	7.51
≥ 10	162	1450	17.90	55	125	4.55

Density of $\log_{10}(\text{wdeg}_{WA_n}(a))$

Multiunits and truncated networks

V. Batagelj

Multiple units

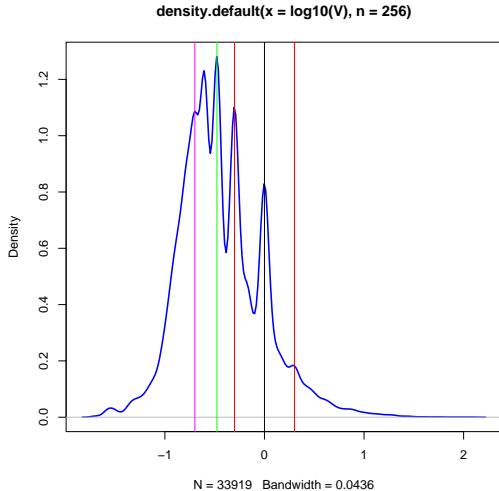
Fractional approach

Truncated networks

Linking through a network

Conclusions

References



Linking through a network

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

Assume that another network \mathbf{S} on $W \times W$ is given. The network

$$\mathbf{Q} = n(\mathbf{WA})^T \cdot \mathbf{S} \cdot n(\mathbf{WA})$$

links authors to authors *through* the network \mathbf{S} .

From [1] we know that

- If \mathbf{S} is symmetric, $\mathbf{S}^T = \mathbf{S}$, then also \mathbf{Q} is symmetric, $\mathbf{Q}^T = \mathbf{Q}$.
- $T(\mathbf{Q}) = T(\mathbf{S})$.

... Linking through a network

Let's look at

$$\text{wdeg}_Q(a) = \sum_{b \in A} q[a, b] = \sum_{b \in A} \sum_{w \in W} \sum_{z \in W} nwa[w, a] \cdot s[w, z] \cdot nwa[z, b] =$$

$$= \sum_{w \in W} \sum_{z \in W} nwa[w, a] \cdot s[w, z] \cdot \text{sign}(\text{deg}(w)) =$$

$$\sum_{w \in W} nwa[w, a] \cdot \sum_{z \in W} s[w, z] = \sum_{w \in W} nwa[w, a] \cdot \text{wdeg}_S(w)$$

or in a vector form

$$\mathbf{wdeg}_Q = n(\mathbf{WA})^T \cdot \mathbf{wdeg}_S$$

The most active authors are

$$A_1 = \{a \in A : \text{wdeg}_Q(a) \geq t\}$$

Again the truncation scheme can be applied.

Authors co-citation

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

The co-citation network is defined as $\mathbf{CoCi} = \mathbf{Ci}^T \cdot \mathbf{Ci}$ and the fractional co-citation network as $\mathbf{CoCin} = \mathbf{Cin}^T \cdot \mathbf{Cin}$. Both \mathbf{CoCi} and \mathbf{CoCin} are symmetric. Authors fractional co-citation network is obtained by linking authors through the fractional co-citation network

$$\begin{aligned}\mathbf{CoCan} &= n(\mathbf{WA})^T \cdot \mathbf{CoCin} \cdot n(\mathbf{WA}) = \\ &= n(\mathbf{WA})^T \cdot \mathbf{Cin}^T \cdot \mathbf{Cin} \cdot n(\mathbf{WA}) = \mathbf{Can}^T \cdot \mathbf{Can}\end{aligned}$$

where $\mathbf{Can} = \mathbf{Cin} \cdot n(\mathbf{WA})$.

As in the case of \mathbf{WA} we have $wdeg_{Cn}(a) = wdeg_{n(WA)}(a)$, for \mathbf{Ci} it holds also $wdeg_{CoCin}(w) = wdeg_{n(Ci)}(w)$. Therefore

$$wdeg_{CoCan} = n(\mathbf{WA})^T \cdot wdeg_{CoCin} = n(\mathbf{WA})^T \cdot wdeg_{n(Ci)}$$

and it is easy to see that also

$$wdeg_{Can} = n(\mathbf{WA})^T \cdot wdeg_{n(Ci)} = wdeg_{CoCan}$$

... Authors co-citation

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

```
> Cin <- normalize(Ci)
> CoCin <- t(Cin)%*%Cin
> CoCan <- t(WAn)%*%CoCin%*%WAn
> wdeg_CoCan <- (t(WAn)%*%wdeg(Cin))[,1]
> wdeg_CoCan
      AB      CD      EF      GH      IJ      KL      MN      OP      RS
0.4556 0.9694 0.3444 0.2778 1.0806 0.3444 0.6250 0.4444 0.4583
> wodeg(CoCan)
      AB      CD      EF      GH      IJ      KL      MN      OP      RS
0.4556 0.9694 0.3444 0.2778 1.0806 0.3444 0.6250 0.4444 0.4583
> Can <- Cin%*%WAn
> wdeg(Can)
      AB      CD      EF      GH      IJ      KL      MN      OP      RS
0.4556 0.9694 0.3444 0.2778 1.0806 0.3444 0.6250 0.4444 0.4583
```

Conclusions

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

- the fractional version(s) of bibliographic coupling $\mathbf{biCo} = \mathbf{Ci} \cdot \mathbf{Ci}^T$ are more complicated [1]. The corresponding “through” constructions are still on the “to do” list.
- efficient implementation as Pajek macros and in Python Nets package.

Acknowledgments

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References

This work is supported in part by the Slovenian Research Agency (research program P1-0294, research program CogniCom (0013103) at the University of Primorska, and research projects J5-2557, J1-2481, and J5-4596), and prepared within the framework of the COST action CA21163 (HiTEc).

Multiunits and truncated networks

V. Batagelj

Multiple units

Fractional approach

Truncated networks

Linking through a network

Conclusions

References



Batagelj, V: On fractional approach to the analysis of linked networks. *Scientometrics* 123 (2020), 621–633. [Springer](#)



Maltseva, D., Batagelj, V. iMetrics: the development of the discipline with many names. *Scientometrics* 125 (2020), pages 313–359.
<https://doi.org/10.1007/s11192-020-03604-4>



Batagelj, V, Cerinšek, M: On bibliographic networks. *Scientometrics* 96 (2013) 3, 845-864.
<https://doi.org/10.1007/s11192-012-0940-1>



Anne-Wil Harzing: harzing.com



Anne-Wil Harzing: Health warning: Might contain multiple personalities. The problem of homonyms in Thomson Reuters Essential Science Indicators. *Scientometrics* 105(3):2259-2270. [paper](#)