

Analiza omrežij

V. Batagelj

Osnovni pojm

Pajek

)pisi omrežij

Prikazi omrež

Vrste omrežii

Pajek, Excel in

Analiza omrežij

2. Vrste in opis omrežij

Vladimir Batagelj

Magistrski program Uporabna statistika Ljubljana, april 2024



Kazalo

Analiza omrežij

V. Batagelj

Osnovni pojm

Pajek

Opisi omrežij

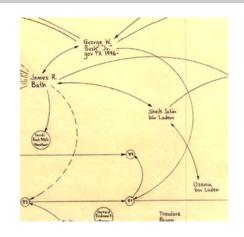
Prikazi omrežij

Vrste omrežij

vibre onnern

1 Osnovni pojmi

- 2 Pajek
- Opisi omrežij
- 4 Prikazi omrežij
- 5 Vrste omrežij
- 6 Pajek, Excel in R



prof. Vladimir Batagelj: vladimir.batagelj@fmf.uni-lj.si
prosojnice (PDF)

22. april 2024 ob 01:11/ marec 2013



Omrežja osnovni pojmi

Analiza omrežij V. Batageli

Osnovni pojmi

- --5 ---

opisi omrezij

Prikazi omrez

Vrste omrežij

Viste Onnezij

Osnovni sestavini *omrežja* sta množica *vozlišč*, ki predstavljajo izbrane *enote*, in množica *povezav*, ki predstavljajo *odnose* (relacije) med enotami. Vozlišča in povezave določajo *graf*.

Pri povezavi je lahko smer pomembna – *usmerjena* povezava, ali pa ni *neusmerjena* povezava.

O vozliščih in povezavah lahko poznamo dodatne podatke – njihove *lastnosti*. Npr. oznaka, vrsta, vrednost, . . .

Omrežje = Graf + Podatki

Ti podatki so lahko izmerjeni ali izračunani.



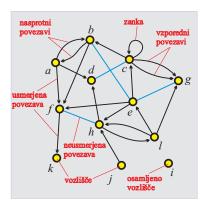
Graf

Analiza omrežij

V. Batagelj

Osnovni poimi

Prikazi omrežij



enota – *vozlišče*, točka relacija – *povezava* usmerjena povezava (a, d) a je njen začetek d pa njen konec. neusmerjena povezava (c:d) c in d sta njeni krajišči.



Omrežja / Formalno

Analiza omrežij V. Batageli

Osnovni poimi

Omrežje $\mathcal{N} = (\mathcal{V}, \mathcal{L}, \mathcal{P}, \mathcal{W})$ je določeno z:

- grafom $\mathcal{G} = (\mathcal{V}, \mathcal{L})$, kjer je \mathcal{V} množica vozlišč, \mathcal{A} je množica usmerjenih povezav, in \mathcal{E} je množica neusmerjenih povezav. Z $\mathcal{L} = \mathcal{E} \cup \mathcal{A}$ označimo množico vseh povezav. $n = |\mathcal{V}|, m = |\mathcal{L}|$
- \mathcal{P} lastnosti / vozliščne vrednosti: $p: \mathcal{V} \to A$
- W uteži / povezavne vrednosti: $w: \mathcal{L} \to B$



Velikost omrežia

Analiza omrežij

V. Batagelj

Osnovni pojmi

Velikost omrežja/grafa je določena z dvema številoma: številom vozlišč $n = |\mathcal{V}|$ in številom povezav $m = |\mathcal{L}|$.

V enostavnem neusmerjenem grafu (ni vzporednih povezav in zank) je $m \leq \frac{1}{2}n(n-1)$; v enostavnem usmerjenem grafu (ni vzporednih povezav) pa $m < n^2$.

Razmerje $\gamma = \frac{m}{m}$ je gostota grafa.

Omrežja na nekaj deset vozliščih so majhna – lahko jih narišemo in za njihovo analizo uporabimo veliko algoritmov (UCINET, NetMiner). Tudi srednje velika omrežja (nekaj sto vozlišč) še lahko narišemo (preglednost !?), nekateri postopki že odpovejo.

Do začetka 90. let je bila večina omrežij majhnih – raziskovalci so jih zbrali z anketami, opazovanjem, iz arhivskih zapisov, ... Razvoj IT je omogočil, da je bilo mogoče ustvariti omrežja iz podatkov že zbranih na računalnikih. Velika omrežja so postala dejstvo. Velika omrežja ne moremo naenkrat podrobno prikazati v celoti. Za njihovo analizo in prikaz potrebujem posebne, nove pristope. Za to je bil razvit program Pajek.



Velika omrežja

Analiza omrežij

V. Batagelj

Osnovni poimi

Veliko omrežje – več tisoč ali celo milijonov vozlišč. V celoti ga lahko shranimo v pomnilnik računalnika – sicer ogromno omrežje. 64-bitni računalniki!

Jure Leskovec: SNAP - Stanford Large Network Dataset Collection

Social networks

Name	Туре	Nodes	Edges	Description
ego-Facebook	Undirected	4,039	88,234	Social circles from Facebook (anonymized)
ego-Gplus	Directed	107,614	13,673,453	Social circles from Google+
ego-Twitter	Directed	81,306	1,768,149	Social circles from Twitter
soc-Epinions1	Directed	75,879	508,837	Who-trusts-whom network of Epinions.com
soc-LiveJournal1	Directed	4,847,571	68,993,773	LiveJournal online social network
soc-Pokec	Directed	1,632,803	30,622,564	Pokec online social network
soc-Slashdot0811	Directed	77,360	905,468	Slashdot social network from November 2008
soc-Slashdot0922	Directed	82,168	948,464	Slashdot social network from February 2009
wiki-Vote	Directed	7,115	103,689	Wikipedia who-votes-on-whom network

→ Networks with ground-truth communities

Name	Туре	Nodes	Edges	Communities	Description
com-LiveJournal	Undirected, Communities	3,997,962	34,681,189	287,512	LiveJournal online social network
com-Friendster	Undirected, Communities	65,608,366	1,806,067,135	957,154	Friendster online social network
com-Orkut	Undirected, Communities	3,072,441	117,185,083	6,288,363	Orkut online social network



Dunbarovo število

Analiza omrežij

V. Batagelj

Osnovni pojmi

_ . .

O-:-: ----

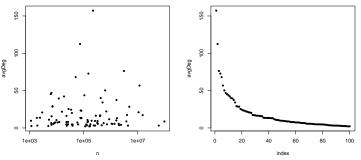
Drikazi omraž

v.,

Vrste omrežij

Detal Food

Povprečne stopnje omrežij SNAP in Konect:



Stopnja $\deg(v)$ je enaka številu povezav, ki imajo vozlišče v za krajišče. Povprečna stopnja $\bar{d} = \frac{1}{n} \sum_{v \in V} \deg(v) = \frac{2m}{n}$. Večina dejanskih omrežij je redkih – število vozlišč n in število povezav m sta istega velikostnega reda.

To opažanje je znano kot Dunbarovo število. Izhaja iz naslednjega razmisleka: recimo, da vsako vozlišče porabi za vzdrževanje posamezne povezave določeno "energijo" in da ima omejeno celotno energijo. Tedaj je tudi število povezav omejeno. V človeških združbah je Dunbarovo število običajno med 100 in 150.



Analiza omrežij

V. Batagelj

Osnovni pojmi

Paiek

Prikazi omre

Vrste omreži

VISCO OIIIICZI

Pajek, Excel i

Zahtevnost algoritmov

Poglejmo časovne zahtevnosti nekaj značilnih algoritmov:

	T(n)	1.000	10.000	100.000	1.000.000	10.000.000
LinAlg	O(n)	0.00 s	0.015 s	0.17 s	2.22 s	22.2 s
LogAlg	$O(n \log n)$	0.00 s	0.06 s	0.98 s	14.4 s	2.8 m
SqrtAlg	$O(n\sqrt{n})$	0.01 s	0.32 s	10.0 s	5.27 m	2.78 h
SqrAlg	$O(n^2)$	0.07 s	7.50 s	12.5 m	20.8 h	86.8 d
CubAlg	$O(n^3)$	0.10 s	1.67 m	1.16 d	3.17 L	3.17 kL

Za interaktivno uporabo na velikih podatkih so že kvadratični algoritmi, $O(n^2)$, prezahtevni.

Pri algoritmih eksponentne zahtevnosti, npr. $O(2^n)$, časovne zahteve naraščajo izjemno hitro. Recimo, da za n=20 potrebujemo 1 s, tedaj za n=30 potrebujemo 2^{10} s = 1024 s = 17 m; za n=40 dobimo 1048576 s = 12.14 d; za n=50 že 1073741824 s = 34.05 L; . . . za n=80 pa z 36.56 10^9 L presežemo domnevno starost Zemlje.



Pajek in velika omrežja

Analiza omrežij

V. Batagelj

Osnovni poimi

Pajek

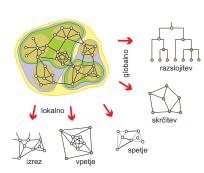
0....

Prikazi omre

vrste omrezij

.

Glavni cilji v zasnovi programa Pajek so:



- podpora abstrakciji z (rekurzivnim) razčlenjevanjem velikega omrežja v več manjših, ki jih lahko v nadaljevanju posamično obdelamo z zahtevnejšimi postopki;
- ponuditi uporabniku izbor orodij za prikaz omrežij;
- izbrati/razviti nabor učinkovitih podkvadratičnih algoritmov za analizo velikih omrežij.

S Pajekom lahko: *določimo* zanimive skupine (komponente, soseščine 'pomembnih' vozlišč, sredice, ...) v omrežju, *izrežemo* del omrežja, ki pripada izbranim skupinam in ga ločeno *prikažemo*; lahko s 'povzetkom vpetosti' v celotno omrežje (lokalni pogled), *skrčimo* vozlišča, ki pripadajo isti skupini, in prikažemo odnose med skupinami (globalni pogled).



Vrste podatkov v Pajek u

Analiza omrežij

V. Batagelj

Osnovni poimi

Pajek

Opisi omreži

Prikazi omrež

Vrste omrežii

Viste Officezij

Pajek, Excel i

Pri svojem delu se Pajek naslanja na 6 vrst podatkov:



- omrežje/network (graf),
- razbitje/partition (imenske ali urejenostne lastnosti vozlišč),
- vektor/vector (številske lastnosti vozlišč),
- skupina/cluster (podmnožica vozlišč),
- urejenost/permutation (urejenost vozlišč, urejenostna lastnost), in
- hierarhija/hierarchy (drevo nad vozlišči).

Novembra 2004 je bila vgrajena tudi podpora *večrelacijskim omrežjem* in *razbitjem povezav*.



... Vrste podatkov v Pajek u

Analiza omrežij

V. Batageli

Pajek

Prikazi omrežij

Moč programa Pajek temelji na velikem številu operacij, ki omogočajo različne pretvorbe med temi vrstami podatkov. Na njih je osnovana tudi zgradba uporabniškega vmesnika programa Pajek. Pajekovo glavno okno je nekakšno 'računalo' s seznamskimi 'registri' za vsako vrsto podatkov. Operacije se opravijo nad tekočimi (izbranimi) podatki v teh seznamih in vanje tudi vrnejo rezultate.

Operacije najdemo v izbirah glavnega okna. Razvrščene so glede na vrste podatkov, ki jih zahtevajo.

Pogosto uporabljana zaporedja operacij lahko združimo v makroje. To omogoča prilagoditve programa Pajek različnim skupinam uporabnikov (družboslovje, analiza besedil, rodoslovje, kemija, biologija, transport, računalništvo, matematika, ...) in posebnim obdelavam. Pajek podpira tudi ponavljanje operacij na zaporedjih omrežij.



Opis grafa – množice / NET

Analiza omrežij

V. Batageli

Opisi omrežij

Prikazi omrežij

$$\mathcal{V} = \{a, b, c, d, e, f, g, h, i, j, k, l\}
\mathcal{A} = \{(a, b), (a, d), (a, f), (b, a), (b, f), (c, b), (c, c), (c, g), (c, g), (e, c), (e, f), (e, h), (f, k), (h, d), (h, l), (j, h), (l, e), (l, g), (l, h)\}
\mathcal{E} = \{(b: e), (c: d), (e: g), (f: h)\}
\mathcal{G} = (\mathcal{V}, \mathcal{A}, \mathcal{E})$$

 $\mathcal{A} = \emptyset$ - neusmerien graf; $\mathcal{E} = \emptyset$ - usmerien graf.

Pajek: GraphSet; TinaSet;

WWW: GraphSet / net; TinaSet / net, slika picture.

 $f = A \cup E$



Graf – Množice / NET

Analiza omrežij

V. Batagelj

Osnovni poimi

Paie

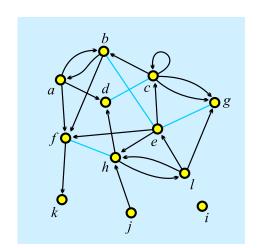
Opisi omrežij

Prikazi omrežij

Vrste omrežij

viste omrezij

Pajek, Excel in



```
*Vertices
   "a"
           0.1020 0.3226
           0.2860
                   0.0876
           0.5543
                   0.4770
           0.4789 0.9055
           0.1175 0.9032
           0.7095 0.6475
```



Graf – Sosedi / NET

Analiza omrežij

V. Batageli

Opisi omrežij

$$\begin{split} N_{\mathcal{A}}(a) &= \{b,d,f\} \\ N_{\mathcal{A}}(b) &= \{a,f\} \\ N_{\mathcal{A}}(c) &= \{b,c,g,g\} \\ N_{\mathcal{A}}(e) &= \{c,f,h\} \\ N_{\mathcal{A}}(f) &= \{k\} \\ N_{\mathcal{A}}(h) &= \{d,l\} \\ N_{\mathcal{A}}(j) &= \{h\} \\ N_{\mathcal{A}}(l) &= \{e,g,h\} \end{split}$$

$$N_{\mathcal{E}}(e) = \{b, g\}$$

 $N_{\mathcal{E}}(c) = \{d\}$

$$N_{\mathcal{E}}(c) = \{a\}$$

 $N_{\mathcal{E}}(f) = \{h\}$

Pajek: GraphList; TinaList;

WWW: GraphList / net; TinaList / net.



Graph - Sosedi / NET

Analiza omrežij

V. Batagelj

Osnovni pojmi

Paie

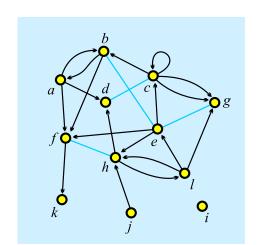
Opisi omrežij

Prikazi omrežij

Vrste omrežii

Vrste omrezij

Pajek, Excel in



```
*Vertices
           12
          0.1020
                   0.3226
          0.2860
                   0.0876
          0.5322
                   0.2304
    "d"
          0.3259
                   0.3917
          0.5543
                   0.4770
    "f"
          0.1552
                   0.6406
    "a"
          0.8293
                   0.3249
          0.4479
                   0.6866
          0.8204
                   0.8203
          0.4789
                   0.9055
    "k"
          0.1175
                   0.9032
12 "1"
          0.7095
                   0.6475
*Arcslist
        4
     4 12
10
12
*Edgeslist
```



Graf – Matrika / NET

Analiza omrežij

V. Batagelj

Opisi omrežij Prikazi omrežij

	а	Ь	с	d	е	f	g	h	i	j	k	1
а	0	1	0	1	0	1	0	0	0	0	0	0
Ь	1	0	0	0	1	1	0	0	0	0	0	0
с	0	1	1	1	0	0	2	0	0	0	0	0
d	0	0	1	0	0	0	0	0	0	0	0	0
e	0	1	1	0	0	1	1	1	0	0	0	0
f	0	0	0	0	0	0	0	1	0	0	1	0
g	0	0	0	0	1	0	0	0	0	0	0	0
h	0	0	0	1	0	1	0	0	0	0	0	1
i	0	0	0	0	0	0	0	0	0	0	0	0
j	0	0	0	0	0	0	0	1	0	0	0	0
k	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	1	0	1	1	0	0	0	0

Pajek: GraphMat; TinaMat, slika picture;

WWW: GraphMat / net; TinaMat / net, paj.

Graf G je enostaven ntk. vse vrednosti v matriki so 0 ali 1.



Graph - Matrika / NET

Analiza omrežij

V. Batagelj

Osnovni pojmi

Paie

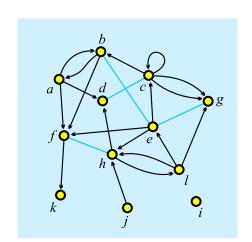
Opisi omrežij

Prikazi omrežij

Vrste omrežij

Viste Officezij

Pajek, Excel in



*Ve	ert	iic	ces	3	12	2					
1 "a"						.10)2()	0 .	. 32	226
2		'b'	•		0.	. 28	360)	0 .	. 08	376
3	,	'c'	•		0.	. 53	322	2	0 .	. 23	304
4		'd'	•		0.	. 32	259	9	0.	. 39	917
5	,	'e'	•		0.	. 55	543	3	0 .	. 47	770
6	,	'f'	•		0.	. 15	552	2	0 .	. 64	106
7	•	'g'	•		0.	. 82	293	3	0 .	. 32	249
8		''n'			0.	. 44	179	9	0 .	. 68	366
9	,	'i'	•		0.	. 82	204	1	0 .	. 82	203
10	, ,	'j'	•		0.4789			0.9055			
11		١ķ٠			0.	.1:	175	5	0.	. 90	32
12		'1'	•		0.	.70	95	5	0.	. 64	175
*Ma	tı	rix	ζ.								
0	1	0	1	0	1	0	0	0	0	0	0
1	0	0	0	1	1	0	0	0	0	0	0
0	1	1	1	0	0	2	0	0	0	0	0
0	0	1	0	0	0	0	0	0	0	0	0
0	1	1	0	0	1	1	1	0	0	0	0
0	0	0	0	0	0	0	1	0	0	1	0
0	0	0	0	1	0	0	0	0	0	0	0
0	0	0	1	0	1	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	1	1	0	0	0	0



Lastnosti vozlišč / CLU, VEC, PER

Analiza omrežij

V. Batagelj

Osnovni pojm

- -- -- ---

Opisi omrežij

Prikazi omrez

Vrste omrežij

vibto omitozij

Vse tri vrste datotek imajo enako zgradbo:

*vertices n

n število vozlišč

 v_1 vozlišče 1 ima vrednost v_1

.. V_n

v_n

CLUstering – razbitje vozlišč – *imenska* ali *urejenostna* lastnost vozlišč

 $v_i \in \mathbb{N}$: vozlišče *i* pripada skupini v_i ;

VECtor – <u>številska</u> lastnost vozlišč

 $v_i \in \mathbb{R}$: lastnost ima vrednost v_i na vozlišču i;

PERmutation – *urejenost* vozlišč

 $v_i \in \mathbb{N}$: vozlišče i je na v_i -tem mestu.

Ko zbiramo podatke o omrežjih, ne pozabimo zajeti tudi čimveč lastnosti in uteži.



Primer: Wolfe Monkey Data

Analiza omrežij

V. Batagelj

Osnovni pojm

Pajel

Opisi omrežij

Prikazi omrežij

Vrste omrežij

.

Pajek, Excel in R

inter.net	inter.net	sex.clu	age.vec	rank.per
*Vertices 20 1 'm01" 2 'm02" 3 'm03" 4 'm04" 5 'm05" 6 'f06" 7 'f07" 8 'f08" 9 'f09" 10 'f10" 11 'f11" 12 'f12" 13 'f13" 14 'f14" 15 'f15" 16 'f16" 17 'f17" 18 'f18" 19 'f19" 20 'f20" *Edges 1 3 10 1 4 4	1 6 5 1 7 9 1 8 7 1 9 4 1 10 3 1 11 3 1 12 7 1 13 3 1 14 5 1 15 5 1 16 4 1 17 4 1 18 1 2 3 5 2 4 1 2 5 1 2 6 1 2 7 1 2 8 2 2 9 2 2 10 2 2 11 5 2 12 4 2 13 3 2 14 2	*vertices 20 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2	*vertices 20 15 10 10 8 7 15 5 11 8 9 16 10 14 5 7 11 7 5 15 4	*vertices 20 1 2 3 4 5 10 11 6 12 9 7 8 18 19 20 13 14 15 16 17

Pozor: 0 ni dovoljena številka vozlišča.





Pajekova projektna datoteka / PAJ

Analiza omrežij

V. Batageli

Opisi omrežij

Vse vrste podatkov o omrežju lahko združimo v eno samo datoteko - Pajekovo *projektno* datoteko *file*.paj. Najenostavneje to naredimo tako:

- preberite v program Pajek vse podatkovne datoteke,
- izračunajte še morebitne dodatne podatke,
- odstranite (dispose) morebitne odvečne podatke,
- shranite vse skupaj na projektno datoteko z File/Project file/Save.

Naslednjič lahko obnovite stanje z eno samo zahtevo File/Project file/Read.

Wolfe-vo omrežje kot Pajekova projektna datoteka (PDF/paj).



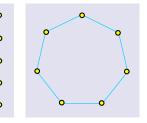
Posebni grafi – pot, cikel, zvezda, polni

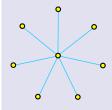
Analiza omrežij

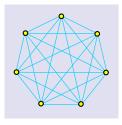
V. Batagelj

Opisi omrežij

Prikazi omrežij







Na slikah so zaporedoma prikazani grafi: pot P_5 , cikel C_7 , zvezda S_8 in polni graf K_7 .



Prikazi omrežij

Analiza omrežij

V. Batageli

Paiek

Prikazi omrežii

Ne prevelika omrežja

Options/Read-Write [Large Network (Vertices)]

lahko narišemo v *prikaznem oknu* z ukazi iz izbire Draw, npr. Draw/Network.

Če omrežje nima določenih koordinat vozlišč, ga Pajek nariše *krožno*. Sliko lahko spreminjamo s premikanjem vozlišč z miško.

Za večja omrežja uporabimo postopke za prikaz omrežij, ki jih najdemo v izbiri [Draw] Layout.





Lepe risbe

Analiza omrežij

V. Batagelj

Osnovni poimi

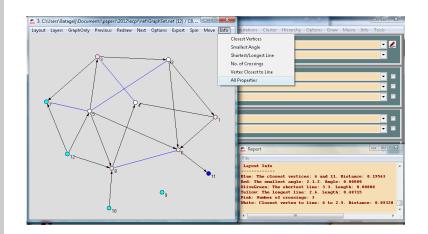
Opisi omreži

Prikazi omrežij

Vrste omrežii

viste omrezi

Pajek, Excel in R





Prikazi lastnosti

Analiza omrežij

V. Batagelj

Prikazi omrežii

Za omrežje $\mathcal{N} = (\mathcal{V}, \mathcal{L}, \mathcal{P}, \mathcal{W})$ so lastnosti vozlišč \mathcal{P} in povezav \mathcal{W} lahko merjene v različnih lestvicah. Lahko so vnešene z ostalimi podatki o omrežju, ali pa *izračunane* pri analizi omrežja.

V programu Pajek podamo posamezno številsko lastnost vozlišč kot vektor, imensko lastnost pa ali kot razbitje ali kot oznako vozlišč. Na sliki lahko številsko lastnost prikažemo kot *velikost*(i) vozlišča ali njegovo koordinato ali velikost imena/oznake; imensko lastnost pa kot barvo ali obliko lika, ali kot oznako vozlišča, ali kot barvo oznake.

Vrednosti povezav so številske. Na sliki jih prikažemo z izpisom vrednosti, debelino črte ali sivino. Imenske vrednosti lahko dodamo v opisu omrežja kot oznake, barvo ali vzorec (glejte priročnik programa Pajek, razdelek 4.3).



Nekaj ukazov v programu Pajek

Analiza omrežij

V. Batagelj

Osnovni pojm

Paje

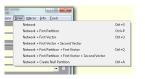
oisi omrežij

Prikazi omrežij

Vrste omrežii

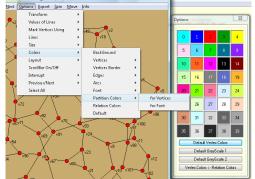
VISCO OTTICZIJ

Pajek, Excel in R



Operations/Network+Vector/Transform/Put Coordi Network/Create Vector/Get Coordinate [Draw] Options

[Draw] Layout/Energy/Kamada-Kawai/Free [Draw] Export/2D/EPS-PS





Prikaz lastnosti – šolarji (Moody)

Analiza omrežij

V. Batagelj

Osnovni pojm

Pajel

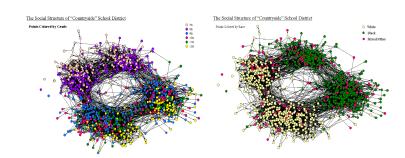
)pisi omrež

Prikazi omrežij

Vrste omrežii

viste omrezi

Pajek, Excel in





Analiza omrežij

V. Batageli

Prikazi omrežij

Primer: Snyder in Kick-ovo omrežie svetovne trgovine

Omrežje je dostopno kot Pajek-ova projektna datoteka SaKtrade.paj

Omrežje opisuje trgovanje med državami (118 vozlišč, 515 usmerjenih in 2116 neusmerjenih povezav). Omrežje je podrobneje opisano v članku: Snyder, David and Edward Kick (1979). The World System and World Trade: An Empirical Exploration of Conceptual Conflicts, Sociological Quaterly, 20.1, 23-36.

Za omrežje je podano tudi razbitje na (pod)kontinente: 1 -Evropa, 2 - Severna Amerika, 3 - Srednja Amerika, 4 - Južna Amerika, 5 - Azija, 6 - Afrika, 7 - Oceanija.

Novejše podatke je mogoče dobiti na NBER / Feenstra; in primer analize na Science / Hidalgo et. al.



Draw / Partition

Analiza omrežij

V. Batagelj

Osnovni pojm

Paiek

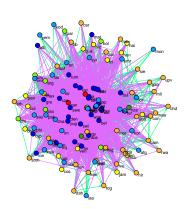
pisi omrež

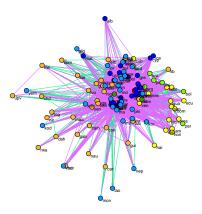
Prikazi omrežij

Vrste omrežij

VISLE OITHEZI

Pajek, Excel i





Draw/Network + First Partition Layout/Energy/Kamada-Kawai/Free Layout/Energy/Fruchterman Reingold/2D



Povečave

Analiza omrežij

V. Batagelj

Prikazi omrežii

1. SaK trade (118) / C1. (Sub)continents (118)

Z desnim gumbom na miški izberite območje povečave. Običajni prikaz obnovite z Redraw.

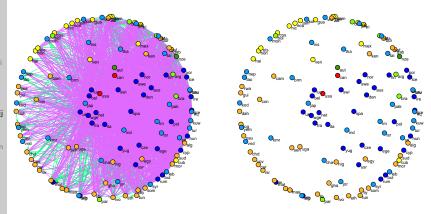


Fruchterman Reingold / factor = 9

Analiza omrežij

V. Batagelj

Prikazi omrežij



Layout/Energy/Fruchterman Reingold/3D

3D picture / King





Matrični prikaz

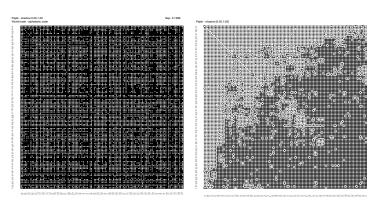
Snyder & Kick's World trade network / n = 118, m = 514

Analiza omrežij

V. Batagelj

Prikazi omrežij

Vrste omrežij



Abecedni vrstni red (levo) in preureditev (desno)



Časovna omrežja

Analiza omrežij

V. Batagelj

Osnovni pojmi

Pajek

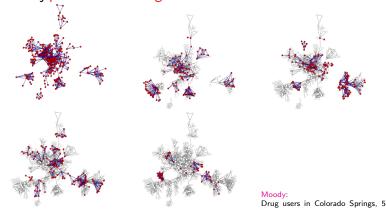
Opisi omrež

Prikazi omrež

Vrste omrežij

viste omiezi

V časovnem omrežju se prisotnost vozlišča/povezave v omrežju spreminja skozi čas. Pajek ponuja dva načina opisa časovnih omrežij *prisotnostni* in *dogodkovni*.





Časovna omrežja

Analiza omrežij

V. Batagelj

Osnovni pojm

Pajek

Opisi omreži

Prikazi omrež

Vrste omrežij

viste omrezi

Pajek, Excel

Časovno omrežje

$$\mathcal{N}_T = (\mathcal{V}, \mathcal{L}, \mathcal{P}, \mathcal{W}, T)$$

dobimo, če navadnemu omrežju dodamo *čas* T . T je množica *časovnih točk* ali *trenutkov* $t \in T$.

V časovnem omrežju vozlišča $v \in \mathcal{V}$ in povezave $I \in \mathcal{L}$ niso nujno vseskozi dejavna/prisotna. Če je povezava I(u, v) dejavna v trenutku t, morata biti dejavni v t tudi njeni krajišči u in v.

Omrežje sestavljeno iz vozlišč in povezav dejavnih v trenutku $t \in T$ bomo označevali z $\mathcal{N}(t)$ in mu rekli *časovna rezina* v trenutku t. V programu Pajek dobimo zaporedje časovnih rezin z zahtevo

Network/Temporal Network/Generate in time



Časovna omrežja – prisotnostni opis

Analiza omrežij

V. Batageli

Vrste omrežij

Vertices 3 [5-10, 12-14]"b" [1-3.7] 3 "e" [4-] *Edges 1 2 1 [7] 1 3 1 [6-8]

Podatke o obdobjih prisotnosti/ dejavnosti posameznega vozlišča/ povezave navedemo na koncu pripadajoče vrstice v oglatih oklepajih [in]. Obdobja so oštevilčena od 1 naprej. Pri naštevanju jih ločimo z vejico ,. Zaporedna obdobja od začetka z do konca k lahko krajše zapišemo z-k. Znak * pomeni neskončno. Vozlišče a je prisotno v obdobjih 5 do 10 in 12 do 14. Povezava (1 : 3) je prisotna v obdobjih 6 do 8.

Povezava je prisotna, če sta prisotni obe njeni krajišči.

Time.net

Opis časovnih omrežij uporabljen v programu Pajek je omejen spreminjajo se lahko tudi vrednosti lastnosti in uteži. V razvoju je poseben program Nets / TQ, ki bo omogočal opis in analizo splošnejših časovnih omrežij. 1, 2



Časovna omrežja – dogodkovni opis

Analiza omrežij

V. Batagelj

Osnovni pojm

Pajel

Opisi omreži

Prikazi omreż

Vrste omrežij

VISCO OTTICZIJ

Pajek, Excel in R

Event	Explanation
TI t	initial events – following events happen when
	time point t starts
TE t	end events – following events happen when
	time point t is finished
AV vns	add vertex v with label n and properties s
HV v	hide vertex v
SV v	show vertex v
DV v	delete vertex v
AA uvs	add arc (u,v) with properties s
HA uv	hide arc (u,v)
SA uv	show arc (u,v)
DA uv	delete arc (u,v)
AE uvs	add edge (u:v) with properties s
HE uv	hide edge (u:v)
SE uv	show edge (u:v)
DE uv	delete edge (u:v)
CV vs	change property of vertex v to s
CA uvs	change property of arc (u,v) to s
CE uvs	change property of edge (u:v) to s
CT uv	change (un)directedness of line (u,v)
CD uv	change direction of arc (u,v)
PE IIVS	replace pair of arcs (u,v) and (v,u) by single edge $(u:v)$
	with properties s
AP uvs	add pair of arcs (u,v) and (v,u)
	with properties s
DP IIV	delete pair of arcs (u,v) and (v,u)
EP uvs	replace edge (u:v) by pair of arcs (u,v) and (v,u)
	with properties s

s je lahko prazen.

Pri vzporednih povezavah :k označuje k-to povezavo — HE:3 14 37 skrije tretjo neusmerjeno povezavo med vozliščema 14 in 37.

Time.tim Friends.tim

*Ve	ert	ic	es	3
*Ev		ts	3	
ΤI	1			
AV	2	"ł)"	
ΤE	3			
HV	2			
TE HV TI AV	4			
ΑV	3	"€	∍"	
TI AV	5			
TI	Ţ	" a	1	
Y E	1	3	1	
AE TI	7	3	1	
GM	2			
SV AE TE	1	2	1	
TE	7	-	-	
DE DV TE	1	2		
DV	2			
ΤE	8			
DE	1	3		
ΤE)		
ΗV				
ΤI		!		
SV	1			
TE				
DV	1			



Časovna omrežja / 11. september

Analiza omrežij

V. Batagelj

Osnovni pojmi

Pajel

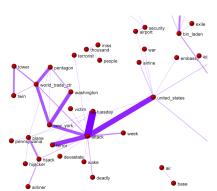
pisi omreži

Prikazi omrež

Vrste omrežij

viste omrezij

Pajek, Excel i R



Slike v SVG: 66 dni

V časovnih omrežjih posamezna vozlišča in povezave niso nujno ves čas dejavne/ prisotne. Steve Corman s sodelavci z Arizona State University je s Centering Resonance Analysis (CRA) predelal dnevne Reutersove novice (66 dni) o 11. septembru v časovno omrežje sopojavljanja besed.



Večkratna omrežja

Analiza omrežij

V. Batagelj

Osnovni pojm

Рајек

Opisi omrezi

Prikazi omreži

Vrste omrežij

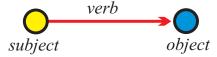
viste omrezij

V večkratnem ali večrelacijskem omrežju

$$\mathcal{N} = (\mathcal{V}, (\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_k), \mathcal{P}, \mathcal{W})$$

je množica povezav \mathcal{L} razbita na podmnožice (*relacije*) \mathcal{L}_i .

Večrelicijska omrežja omogočajo prevedbo besedil na osnovi enostavnih stavkov S-V-O (Subject-Verb-Object) in njih izboljšav.



Subjects ∪ Objects določajo vozlišča, Verbs pa relacije.

Primeri: Roberto Franzosi; *KEDS*, *Tabari*, KEDS / Gulf. Glej še trojice RDF na pomenskem spletu in SPARQL.



Analiza omrežij

V. Batagelj

Osnovni pojm

1 dj Ch

Opisi omrezi

Prikazi omreż

Vrste omrežij

vrste omrezij

Pajek, Excel i

... Večkratna omrežja

V Pajeku jih lahko zapišemo na dva načina:

 geslom, ki napovedujejo opis povezav, (*arcs, *edges, *arcslist, *edgeslist, *matrix) dodamo številko relacije in lahko tudi njeno ime. Npr.

*arcslist :3 "posojanje gradiv"

Vse geslu podrejene povezave pripadajo navedeni relaciji. (Sampson, SampsonL)

 Med povezavami, podrejenimi gesloma *arcs ali *edges, lahko posamezno povezavo pripišemo izbrani relaciji, tako da njen opis začnemo s številko relacije

3: 47 14 5

Povezava s krajiščema 47 in 14 ter utežjo 5 pripada relaciji 3.



Večrelacijsko časovno omrežje – KEDS/WEIS

Analiza omrežij

V. Batagelj

Osnovni pojmi

Pajek

Opisi omreži

Prikazi omrež

Vrste omrežij

viole oimezij

Pajek, Excel in

```
% Recoded by WEISmonths, Sun Nov 28 21:57:00 2004
% from http://www.ku.edu/~keds/data.dir/balk.html
*vertices 325
1 "AFG" [1-*]
 "AFR" [1-*
 "ALB" [1-*]
 "ALBMED" [1-*]
318 "YUGGOV"
320 "YUGMED"
               Γ1-*
    "YUGSER"
322 "10655...
323 "ZAI" [1-*]
*arcs :0 "*** ABANDONED"
*arcs :10 "YTELD"
*arcs :11 "SURRENDER"
*arcs :12 "RETREAT"
*arcs : 223 "MIL ENGAGEMENT"
*arcs : 224 "RIOT"
*arcs : 225 "ASSASSINATE TORTURE"
*arcs
224: 314 153 1 [4]
212: 314 83 1 [4]
                                        890402
                                                 YUG
                                                           KSV
                                                                    224
                                                                           (RIOT) RIOT-TORN
                                        890404
                                                                    212
                                                                           (ARREST PERSON) ALB ETHNIC JAILEI
                                                           ETHALB
224: 3 83 1 [4]
123: 83 153 1 [4]
                                        890407
                                                 ALB
                                                           ETHALB
                                                                    224
                                                                           (RIOT) RIOTS
                                        890408
                                                 ETHALB
                                                           KSV
                                                                    123
                                                                           (INVESTIGATE)
                                                                                             PROBING
42: 105 63 1 [175]
                                                 GER
                                                                    042
                                                                           (ENDORSE)
                                                                                             GAVE SUPPORT
                                        030731
                                                           CYP
                                        030731
                                                                           (ARREST PERSON)
212: 295 35 1 [175]
                                                 UNWCT
                                                           BOSSER
                                                                    212
                                                                                             SENTENCED TO PRIS
43: 306 87 1 [175]
13: 295 35 1 [175]
121: 295 22 1 [175]
                                        030731
                                                 VAT
                                                           EUR
                                                                    043
                                                                           (RALLY) RALLIED
                                        030731
                                                           BOSSER
                                                 UNWCT
                                                                    013
                                                                           (RETRACT)
                                                                                             CLEARED
                                        030731
                                                           BAL
                                                                    121
                                                                           CRITICIZE)
                                                  UNWCT
                                                                                             CHARGES
122: 246 295 1 [175]
                                        030731
                                                 SER
                                                           UNWCT
                                                                    122
                                                                           (DENTGRATE)
                                                                                             TESTIFIED
121: 35 295 1 [175]
                                        030731
                                                 ROSSER
                                                           UNWCT
                                                                    121
                                                                           (CRITICIZE)
                                                                                             ACCUSED
```

Kansas Event Data System KEDS

<□ > < □ > < □ > < ≥ > < ≥ >



Dvovrstna omrežja

Analiza omrežij

V. Batagelj

Osnovni pojm

Pajek

Opisi omiczij

Prikazi omrez

Vrste omrežij

viste omiezij

Pajek, Excel in R V *dvovrstnem* omrežju $\mathcal{N}=((\mathcal{V}_1,\mathcal{V}_2),\mathcal{L},\mathcal{P},\mathcal{W})$ je množica vozlišč sestavljena iz dveh ločenih množic vozlišč \mathcal{V}_1 in \mathcal{V}_2 , $\mathcal{V}=\mathcal{V}_1\cup\mathcal{V}_2$, povezave iz množice \mathcal{L} pa imajo eno krajišče v \mathcal{V}_1 drugo pa v \mathcal{V}_2 . Običajno je znana tudi *utež* $w:\mathcal{L}\to\mathbb{R}\in\mathcal{W}$; če ni, privzamemo w(u,v)=1 za vse povezave $(u,v)\in\mathcal{L}$.

Dvovrstno omrežje lahko opišemo tudi s pravokotno matriko

$$\mathbf{A}=[a_{uv}]_{\mathcal{V}_1\times\mathcal{V}_2}.$$

$$a_{uv} = egin{cases} w_{uv} & (u,v) \in \mathcal{L} \ 0 & ext{otherwise} \end{cases}$$

Primeri: (članki, avtorji), (ljudje, društva, leta članstva), (kupci, dobrine, količina), (poslanci, vprašanje, pozitivni glas), (ljudje, revije, branost).

Dvovrstno omrežje napovemo z *vertices n $n_{\mathcal{U}}$.

Avtorji in dela.



Analiza omrežij

V. Batagelj

Osnovni poimi

Onisi omreži

Prikazi omrežij

Vrste omrežij

Pajek, Excel in

Južnakinje



Najbolj znan primer dvovrstnega omrežja so Davisove južnakinje.

Davis.paj. Freemanov pregled.

Names of Participants of Group I	CODE NUMBERS AND DATES OF SOCIAL EVENTS REPORTED IN Old City Heroid													
	(1) 6/27	3/2	(3) 4/12	(4) 9/26	(5) 2/25	(6) 5/19	3/25	(8) 9/16	(9) 4/8	(10) 6/10	溫	(12) 4/7	(13) 11/21	(14) 8/3
1. Mrs. Evelyn Jefferson	×	×	×	×	××	×		×	×					
2. Miss Laura Mandeville	l x	X	X		X	X	X	X						
3. Miss Theresa Anderson		X	×	×	X	×××	×	×	×					
4. Miss Brenda Rogers	X		×	×	×	X	×	×						
5. Miss Charlotte McDowd				X	X		×							
6. Miss Frances Anderson			X		×	X		×						l
7. Miss Eleanor Nye						×	×	×						
8. Miss Pearl Oglethorpe						X		X	X					
9. Miss Ruth DeSand								X	×					
O. Miss Verne Sanderson								X	X			×		
11. Miss Myra Liddell									X	X		×		
2. Miss Katherine Rogers								X	X	1x		×	×	×
13. Mrs. Sylvia Avondale								X	X	X		×	X	×
14. Mrs. Nora Fayette						X	×		X	×	×	XXX	X	X
15. Mrs. Helen Lloyd							X	X		×	X	×		
16. Mrs. Dorothy Murchison									X	l				
17. Mrs. Olivia Carleton									X		X			
18. Mrs. Flora Price	l					L	l		Ix		l x		l	l



Večvrstna omrežja in povezana omrežja

Analiza omrežij

V. Batagelj

Osnovni pojm

Vrste omrežij

Pajek, Excel in

V večvrstnem omrežju $\mathcal{N} = ((\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_k), \mathcal{L}, \mathcal{P}, \mathcal{W})$ je množica vozlišč sestavljena iz več ločenih množic vrst vozlišč \mathcal{V}_i , $i \in 1 : k$.

Tako omrežje lahko razstavimo v zbirko povezanih enovrstnih

$$\mathcal{N}_j = (\mathcal{V}_j, \mathcal{L}_j, \mathcal{P}_j, \mathcal{W}_j)$$

in dvovrstnih omrežij

$$\mathcal{N}_{i,j} = ((\mathcal{V}_i, \mathcal{V}_i), \mathcal{L}_{i,j}, \mathcal{P}_{i,j}, \mathcal{W}_{i,j}), \qquad i < j$$



Omrežja in razpredelnice

Analiza omrežij

V. Batagelj

Osnovni pojm

1 4 9 0 11

Opisi omrezi

Prikazi omrez

Vrste omrežij

Pajek, Excel in

Kako opisati omrežje \mathcal{N} ? Odgovor je, kot vemo, preprost – podati moramo množice \mathcal{V} , \mathcal{L} , \mathcal{P} , and \mathcal{W} .

Pogosto lahko to storimo tako, da omrežje $\mathcal N$ podamo s tabelama $(\mathcal V,\mathcal P)$ in $(\mathcal L,\mathcal W)$.

Kot primer opišimo omrežje, ki ga določajo naslednja dela: Generalized blockmodeling, Clustering with relational constraint, Partitioning signed social networks, The Strength of Weak Ties

Imamo več vrst vozlišč: osebe, članke, knjige, zbirke, revije, založnike; in različne odnose (relacije) med njimi: author_of, editor_of, contained_in, cites, published_by.

Tabeli lahko pripravimo in vzdržujemo v razpredelniškem programu (npr. Excel). Shranimo jih lahko v znakovni obliki CSV (Comma Separated Values).



bibNodes.csv

Analiza omrežij

V. Batagelj

Osnovni pojm

Pajek

pisi omrežij

Prikazi omrež

Vrste omrežij

Pajek, Excel in

```
name; mode; country; sex; year; vol; num; fPage; lPage; x; y
"Batagelj, Vladimir"; person; SI; m;;;;; 809.1; 653.7
"Doreian, Patrick"; person; US; m;;;;; 358.5; 679.1
"Ferligoj, Anuška"; person; SI; f;;;;;619.5;680.7
"Granovetter, Mark"; person; US; m; ;; ;; ;145.6; 660.5
"Moustaki, Irini";person;UK;f;;;;;783.0;228.0
"Mrvar, Andrej"; person; SI; m;;;;;;478.0;630.1
"Clustering with relational constraint";paper;;;1982;47;;413;426;684.1;3
"The Strength of Weak Ties";paper;;;1973;78;6;1360;1380;111.3;329.4
"Partitioning signed social networks";paper;;;2009;31;1;1;11;408.0;337.8
"Generalized Blockmodeling"; book;;;2005;24;;1;385;533.0;445.9
"Psychometrika"; journal;;;;;;;741.8;086.1
"Social Networks"; journal;;;;;;;321.4;236.5
"The American Journal of Sociology"; journal;;;;;;;111.3;168.9
"Structural Analysis in the Social Sciences"; series;;;;;;;310.4;082.8
"Cambridge University Press"; publisher; UK;;;;;;534.3;238.2
"Springer"; publisher; US;;;;;884.6;174.0
```

bibNodes.csv

V velikih omrežjih, v izogib praznim celicam, omrežje razcepimo na podomrežja – zbirka omrežij.



bibLinks.csv

Analiza omrežij

V. Batagelj

Osnovni pojm

· ·

Vrste omrežij

Pajek, Excel in

from;relation;to
"Batagelj, Vladimir";authorOf;"Generalized Blockmodeling"
"Doreian, Patrick";authorOf;"Generalized Blockmodeling"

"Doreian, Patrick";authorOf;"Generalized Blockmodeling"
"Ferligoj, Anuška";authorOf;"Generalized Blockmodeling"

"Batagelj, Vladimir"; authorOf; "Clustering with relational constraint"
"Ferligoj, Anuška"; authorOf; "Clustering with relational constraint"

"Granovetter, Mark"; authorOf; "The Strength of Weak Ties"
"Granovetter, Mark"; editorOf; "Structural Analysis in the Social Sciences

"Doreian, Patrick"; authorOf; "Partitioning signed social networks"

"Mrvar, Andrej"; authorOf; "Partitioning signed social networks" "Moustaki, Irini"; editorOf; "Psychometrika"

"Doreian, Patrick"; editorOf; "Psychometrika"
"Doreian, Patrick"; editorOf; "Social Networks"

"Generalized Blockmodeling"; containedIn; "Structural Analysis in the Soci
"Clustering with relational constraint"; containedIn; "Psychometrika"

"The Strength of Weak Ties"; containedIn; "The American Journal of Sociolo "Partitioning signed social networks"; containedIn; "Social Networks"

"Partitioning signed social networks"; cites; "Generalized Blockmodeling"

"Generalized Blockmodeling"; cites; "Clustering with relational constraint

"Structural Analysis in the Social Sciences";publishedBy; "Cambridge Univ "Psychometrika";publishedBy; "Springer"

hibl inks csv



Faktorizacija in opis velikih omrežij

Analiza omrežij

V. Batagelj

Osnovni pojm

rajek

Opisi omrezi

Prikazi omrež

Vrste omrežij

Pajek, Excel in

V analizi podatkov za varčevanje s pomnilnikom in večjo računsko učinkovitost pogosto zamenjamo vrednosti imenskih (kategoričnih) spremenljivk s (naravno) številskimi kodami. V R-ju temu rečemo *faktorizacija*.

Vse možne vrednosti imenske spremenljivke oštevilčimo (kodna tabela) in nato zamenjamo vrednosti s pripadajočimo indeksi v tabeli.

Ta pristop se uporablja tudi v večini programov za delo z velikimi omrežji. Žal je pogosto kodna tabela pozabljena ali ohranjena kot meta podatki.

Branje Pajkovega omrežja v R (wiki).



CSV2Pajek.R

transforming CSV file to Pajek files

Analiza omrežij

V. Batagelj

Osnovni pojm

Paiek

Prikazi omrež

Vrste omrežij

Pajek, Excel in

```
# by Vladimir Batageli, June 2016
# setwd("C:/Users/batagelj/work/Python/graph/SVG/EUSN")
# colC <- c(rep("character",4).rep("numeric",7)): nas=c("","NA","NaN")
colC <- c(rep("character",4),rep("numeric",5)); nas=c("","NA","NAN")
nodes <- read.csv2("bibNodes.csv",encoding='UTF-8',colClasses=colC,na.strings=nas)
n <- nrow(nodes): M <- factor(nodes$mode): S <- factor(nodes$sex)
mod <- levels(M): sx <- levels(S): S <- as.numeric(S): S[is.na(S)] <- 0
links <- read.csv2("bibLinks.csv".encoding='UTF-8'.colClasses="character")
F <- factor(links$from.levels=nodes$name.ordered=TRUE)
T <- factor(links$to.levels=nodes$name.ordered=TRUE)
R <- factor(links$relation): rel <- levels(R)
net <- file("bib.net", "w"); cat('*vertices ',n,'\n',file=net)</pre>
clu <- file("bibMode.clu","w"); sex <- file("bibSex.clu","w")</pre>
cat('%',file=clu); cat('%',file=sex)
for(i in 1:length(mod)) cat(' ',i,mod[i],file=clu)
cat('\n*vertices '.n.'\n'.file=clu)
for(i in 1:length(sx)) cat(' '.i.sx[i].file=sex)
cat('\n*vertices '.n.'\n'.file=sex)
for(v in 1:n) {
 cat(v, '"', nodes$name[v], '"\n', sep='', file=net);
 cat(M[v],'\n',file=clu): cat(S[v],'\n',file=sex)
for(r in 1:length(rel)) cat('*arcs :',r,' "',rel[r],'"\n',sep='',file=net)
cat('*arcs\n'.file=net)
for(a in 1:nrow(links))
  cat(R[a],': ',F[a],' ',T[a],' 1 l "',rel[R[a]],'"\n',sep='',file=net)
close(net): close(clu): close(sex)
```

CSV2Pajek.R



bib.net

*arcs :5 "publishedBy"

Analiza omrežij V. Batagelj

Osnovni pojm

Pajek

Opisi omreži

Prikazi omrež

Vrste omrežii

Viste Officezij

Pajek, Excel in

```
*vertices 16
1 "Batagelj, Vladimir"
2 "Doreian, Patrick"
                                                            2 10 1 1 "authorOf"
3 "Ferligoj, Anuška"
 "Granovetter, Mark"
 "Moustaki, Irini"
6 "Mrvar, Andrei"
7 "Clustering with relational constraint"
                                                                 1 1 "editorOf"
8 "The Strength of Weak Ties"
                                                              9 1 1 "authorOf"
9 "Partitioning signed social networks"
                                                                  1 "authorOf"
10 "Generalized Blockmodeling"
                                                                 1 1 "editorOf"
11 "Psychometrika"
                                                            2 12 1 1 "editorOf"
12 "Social Networks"
                                                            10 14 1 1 "containedIn"
13 "The American Journal of Sociology"
                                                         3: 7 11 1 1 "containedIn"
14 "Structural Analysis in the Social Sciences"
                                                         3: 8 13 1 1 "containedIn"
15 "Cambridge University Press"
                                                         3: 9 12 1 1 "containedIn"
16 "Springer"
                                                         2: 9 10 1 1 "cites"
*arcs :1 "authorOf"
                                                         2: 10 7 1 1 "cites"
                                                         5: 14 15 1 1 "publishedBy"
*arcs :2 "cites"
*arcs :3 "containedIn"
                                                         5: 11 16 1 1 "publishedBy"
*arcs :4 "editorOf"
```

bib.net, bibMode.clu, bibSex.clu; bib.paj, bib.ini.



Bibliografsko omrežje – slika / Pajek

Analiza omrežij

V. Batagelj

Osnovni pojm

D . 1

Onici omrož

Prikazi omrež

Vrste omreži

viste omrezi

Pajek, Excel in

