

## Dvovrstna omrežja

Dvovrstno omrežje sestavljata dve množici enot (npr. osebe in dogodki), relacija pa ti dve množici povezuje, npr. vključenost oseb v družabne dogodke.

Primerov takih omrežij je še ogromno:

- *Članstvo v ustanovah* - (ljudje, ustanove, je član); primer je članstvo direktorjev v nadzornih svetih (KdoRola.net).
- *Glasovanje za predloge s strani poslancev* - (poslanci, predlogi, je glasoval za).
- *Omrežje nakupov*, v katerem so prva množica kupci, druga množica artikli, povezava pa pove, kateri artikel je kupec kupil.
- *Bralci in revije, ki jih osebe berejo*.
- *Omrežje citiranj*, katerega prva množica točk so avtorji, druga množica točk članki, povezave med avtorji in članki pa povedo, kateri avtor je citiral kateri članek.
- *Omrežje soavtorstev* - (avtorji, članki, je (so)avtor),

Pripadajoči graf imenujemo *dvodelen graf* – povezave povezujejo samo točke ene s točkami druge množice – znotraj množic ni povezav.



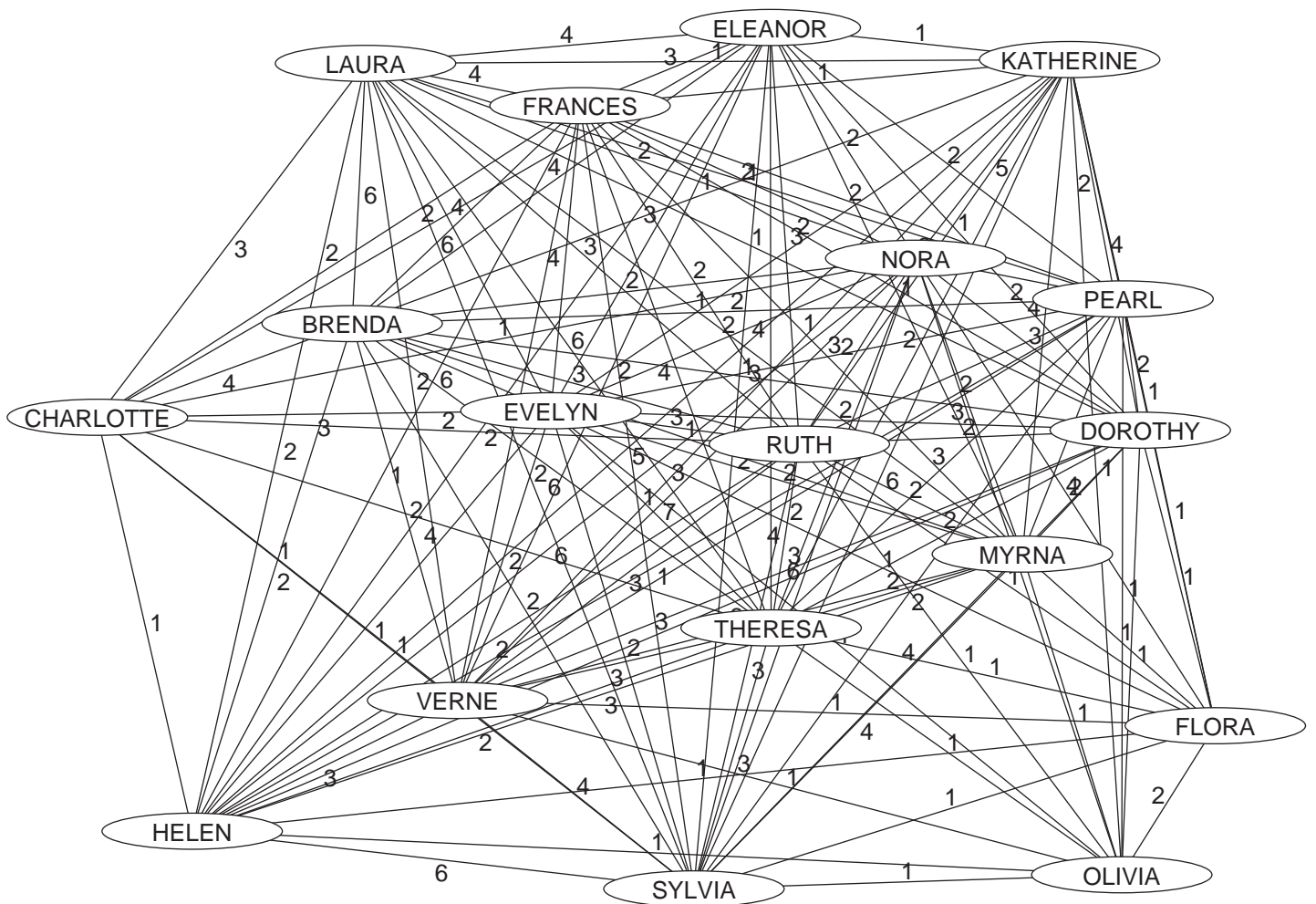
## **Pretvorba dvovrstnih omrežij v običajna omrežja z vrednostmi na povezavah**

Dvovrstno omrežje lahko pretvorimo v običajno omrežje, kjer so enote novega omrežja samo enote iz prve ali samo enote iz druge podmnožice.

Če vzamemo prejšnje dvovrstno omrežje in ga pretvorimo v običajno omrežje, kjer so enote osebe, bosta osebi v relaciji (v ustreznem grafu bo med osebama neusmerjena povezava), če sta se udeležili vsaj enega skupnega dogodka. Vrednost na povezavi predstavlja število dogodkov, ki sta se jih udeležili obe osebi hkrati. Vrednosti zank predstavljajo skupno število dogodkov, ki se jih je posamezna oseba udeležila.

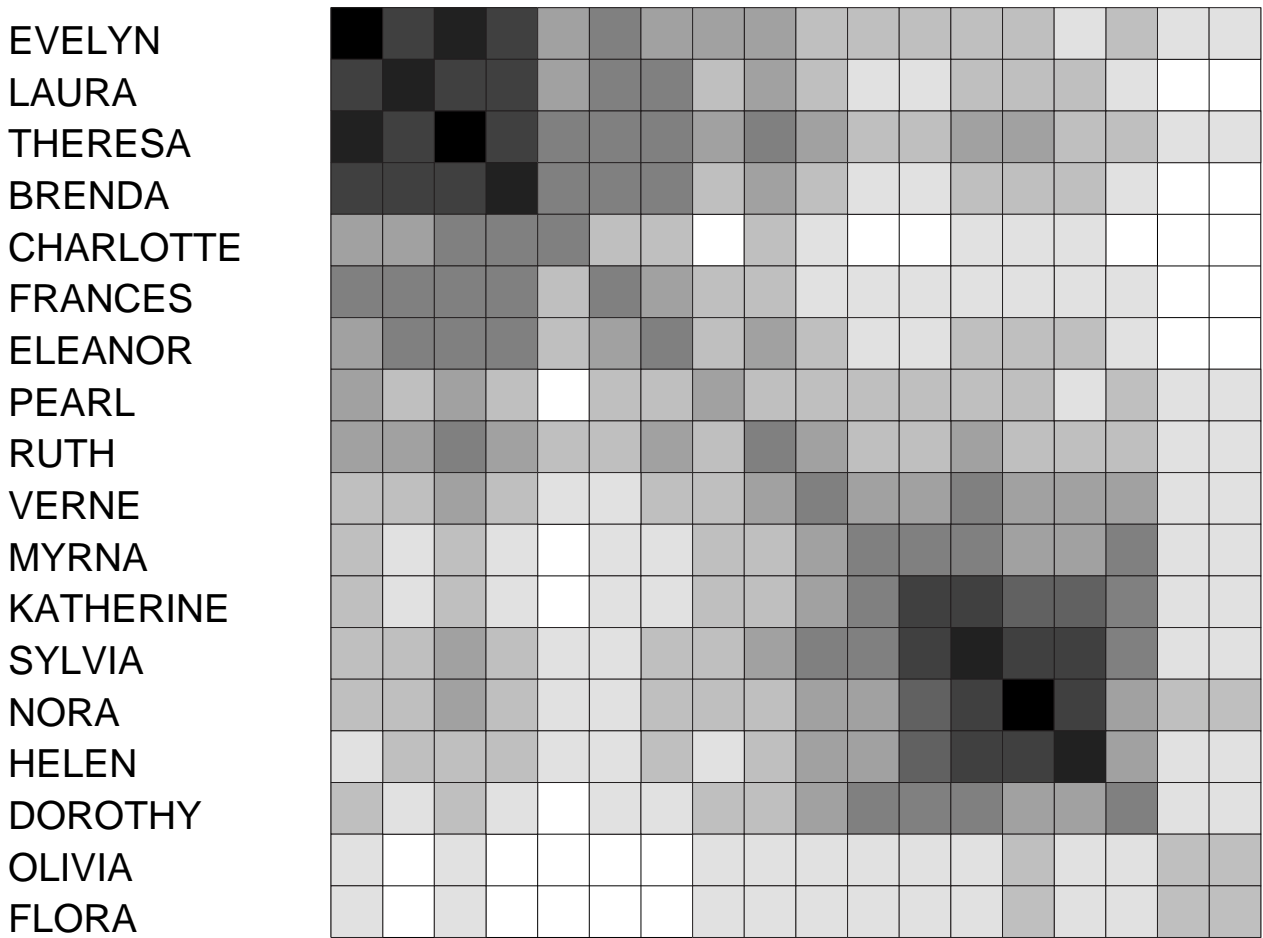
Če pa omrežje pretvorimo v običajno omrežje, kjer so enote dogodki, bosta dva dogodka v relaciji (v ustreznem grafu bo med dvema dogodkoma neusmerjena povezava), če obstaja oseba, ki se je udeležila obeh dogodkov. Vrednost na povezavi med dvema dogodkoma predstavlja število oseb, ki so se udeležile obeh dogodkov. Vrednosti zank predstavljajo skupno število oseb, ki so se udeležile izbranega dogodka.

## Običajno omrežje z osebami in številu skupnih dogodkov (brez zank)



## Matrična predstavitev

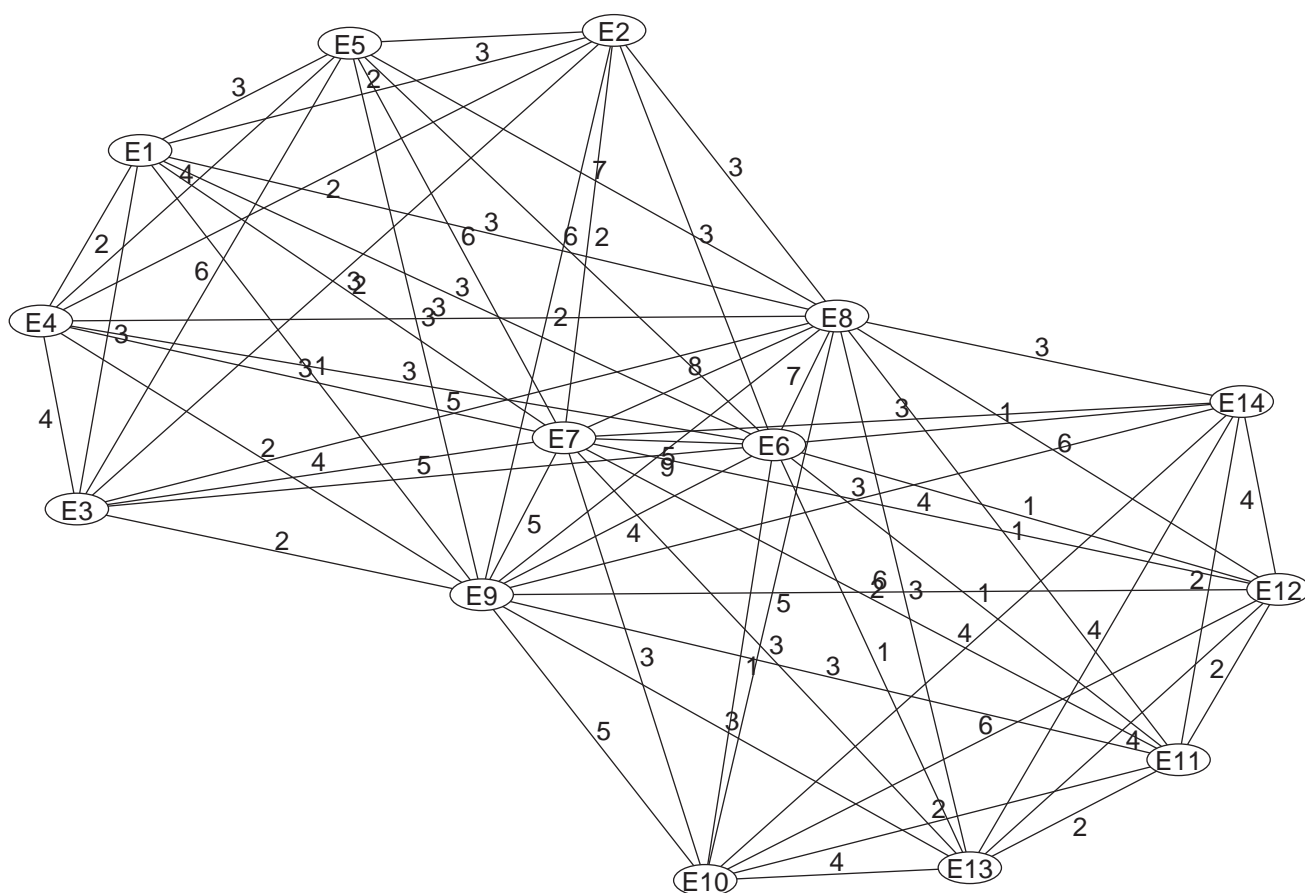
Pajek - shadow [0.00,8.00]



EVELYN  
 LAURA  
 THERESA  
 BRENDA  
 CHARLOTTE  
 FRANCES  
 ELEANOR  
 PEARL  
 RUTH  
 VERNE  
 MYRNA  
 KATHERINE  
 SYLVIA  
 NORA  
 HELEN  
 DOROTHY  
 OLIVIA  
 FLORA

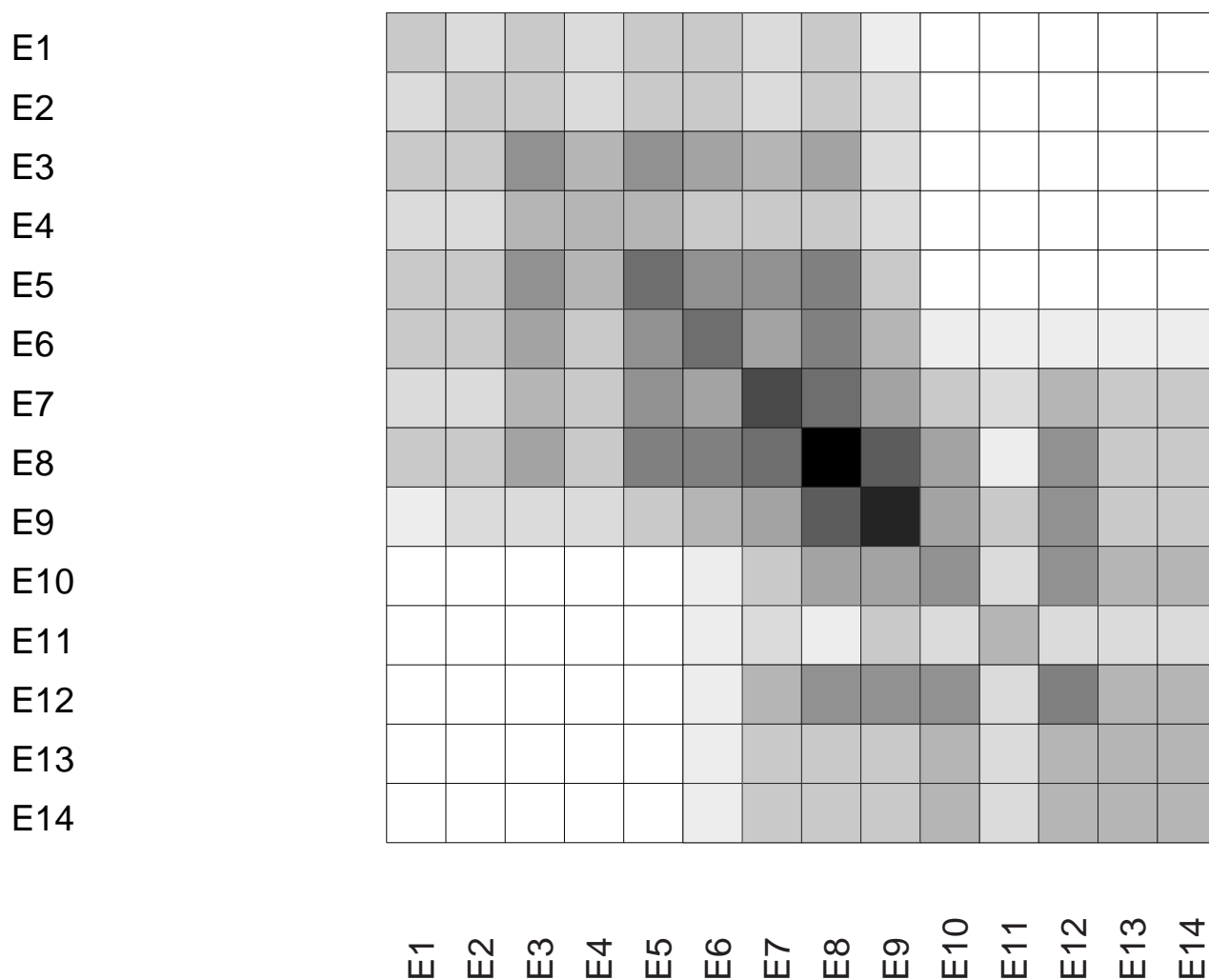


## Običajno omrežje z dogodki in številu skupnih udeleženk (brez zank)



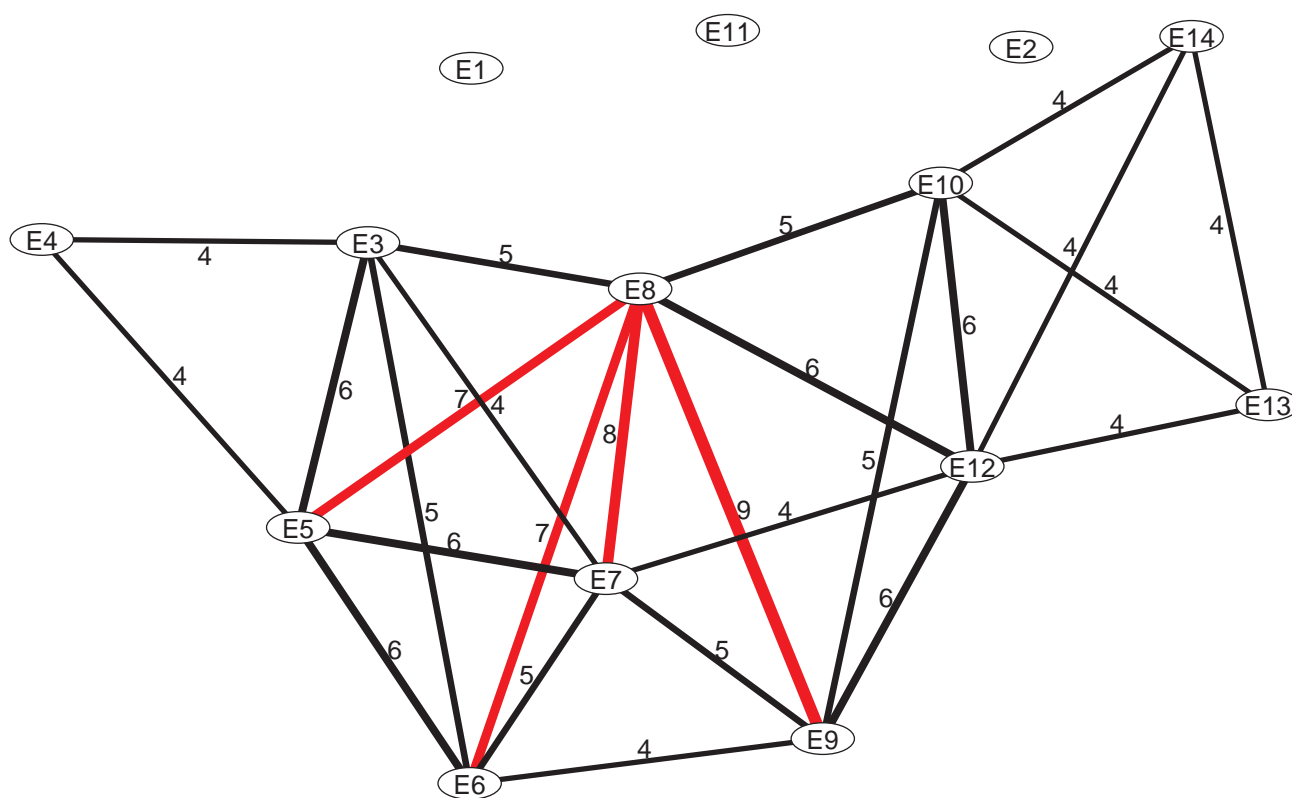
## Matrična predstavitev

Pajek - shadow [0.00,14.00]





## Običajno omrežje z dogodki in številu skupnih udeleženk vsaj 4



# Opis dvovrstnega omrežja na vhodni datoteki

Primer: (Davis.net):

```
*Vertices 32 18
 1 EVELYN          27 E9
 2 LAURA          28 E10
 3 THERESA         29 E11
 4 BRENDA          30 E12
 5 CHARLOTTE       31 E13
 6 FRANCES         32 E14
 7 ELEANOR         *Edgeslist
 8 PEARL           1 19 20 21 22 23 24 26 27
 9 RUTH            2 19 20 21 23 24 25 26
10 VERNE           3 20 21 22 23 24 25 26 27
11 MYRNA           4 19 21 22 23 24 25 26
12 KATHERINE       5 21 22 23 25
13 SYLVIA          6 21 23 24 26
14 NORA            7 23 24 25 26
15 HELEN           8 24 26 27
16 DOROTHY         9 23 25 26 27
17 OLIVIA          10 25 26 27 30
18 FLORA           11 26 27 28 30
19 E1              12 26 27 28 30 31 32
20 E2              13 25 26 27 28 30 31 32
21 E3              14 24 25 27 28 29 30 31 32
22 E4              15 25 26 28 29 30 31 32
23 E5              16 26 27 28 30
24 E6              17 27 29
25 E7              18 27 29
26 E8
```

## Razlaga

Razlika v primerjavi z običajnimi omrežji je le v tem, da moramo pri ukazu \*Vertices poleg števila vseh točk v omrežju (v našem primeru je število oseb in dogodkov 32), navesti še število točk iz prve množice (v našem primeru je število oseb 18). Najprej morajo biti navedene vse točke iz prve množice, potem pa še točke iz druge. Razbitje v obe množici dobimo z ukazom Network/2-Mode Network/Partition into 2 Modes, z vrednostmi 1 za točke prve množice (osebe) in vrednostmi 2 za točke iz druge množice (dogodki).

## Pretvorba dvovrstnih v običajna vrednostna omrežja

Omrežje pretvorimo v običajno omrežje, kjer bodo točke samo točke iz prve množice (v našem primeru osebe) z ukazom Network/2-Mode Network/2-Mode to 1-Mode/Rows.

Če pa želimo dobiti omrežje z elementi iz druge množice, pa uporabimo ukaz

Network/2-Mode Network/2-Mode to 1-Mode/Cols.

Zgeneriramo lahko omrežje z zankami ali brez zank:

Network/2-Mode Network/2-Mode to 1-Mode/Include Loops.

Pravtako lahko zgeneriramo omrežje z vrednostmi na

povezavah (vrednost v našem primeru predstavlja število skupnih dogodkov), ali pa omrežje z večkratnimi povezavami med točkami – za vsak skupni dogodek je med osebama po ena povezava:

### Network/2-Mode Network/2-Mode to 1-Mode/Multiple Lines

Za naše potrebe bomo vedno zgenerirali omrežje z vrednostmi na povezavah (brez večkratnih povezav), medtem ko nam bodo zanke včasih prišle prav včasih pa ne.

## **Prikaz dobljenega omrežja**

Z ukazom Network/Create Vector/Get Loops prenesemo v vektor vrednosti na zankah (npr. skupno število dogodkov, ki se jih je dana oseba udeležila) in le to potem prikažemo z velikostmi točk (Draw/Network + Vector).

Potem ko zgeneriramo omrežje z z vrednostmi na povezavah ga lahko prikažemo na dva načina:

- prikaz v celoti: Narišemo omrežje z energijskim risanjem s tem da pred risanjem izberemo

Options/Values of Lines/Similarities (točke povezane z večjimi vrednostmi bodo bliže skupaj). Vrednosti povezav lahko pokažemo tudi z debelinami povezav (Options/Lines/Different Widths) in/ali sivinami

(Options/Lines/GreyScale).

Nato sliko izvozimo v SVG, kjer postopoma dodajamo povezave glede na njihovo vrednost:

Export/SVG/LineValues/Nested Classes.

- prikaz najpomembnejšega dela omrežja: ohraniti želimo le povezave z dovolj velikimi vrednostmi.

Porazdelitev vrednosti na povezavah pogledamo z

Network/Info/Line Values.

Glede na porazdelitev vrednosti na povezavah, izločimo povezave z nizkimi vrednostmi iz omrežja, tako da uporabimo ukaz

Network/Create New Network/Transform/Remove/...

...lines with value/lower than

ter odtipkamo mejno vrednost.

## Analiza dobljenega omrežja

Za analizo dobljenega omrežja lahko uporabimo pojem  $m$ -rezin ( $m$ -slice): to je podomrežje definirano z vrednostmi na povezavah. Rezine so precej podobne  $k$ -jedrom:

$m$ -rezina je največje podomrežje, ki vsebuje povezave z vrednostjo večjo ali enako  $m$  in vse točke, ki jih imajo te povezave za krajišča.

## *m*-rezine v Pajku

Za izračun karerim *m*-rezinam pripadajo točke, uporabimo ukaz

Network/Create Partition/Valued Core/First Threshold and Step, predhodno pa moramo poskrbeti, da v omrežju ni večkratnih povezav in zank. Pravtako moramo pred izvršitvijo ukaza za izračun rezin poskrbeti, da je izbrana nastavitev

Use max instead of sum.

Ker analiziramo neusmerjeno omrežje, lahko izberemo podizbiro Input ali Output. Kasneje moramo izbrati začetno mejno vrednost in korak s katerim se le ta povečuje. Ponujeni vrednosti (0 in 1) sta pravi za naše potrebe.

Rezultat operacije je novo razbitje, kjer vrednosti pomenijo največjo vrednost *m*-rezine, h kateri točka pripada. Poleg razbitja se nam v oknu *Report* izpiše frekvenčna porazdelitev točk po *m*-rezinah. V frekvenčni porazdelitvi vsebuje vsak razred zgornjo mejo, spodnje pa ne. Še enkrat poudarimo, da v omrežju ne sme biti zank.

Za vizualizacijo *m*-rezin lahko spet uporabimo izbiro Export/SVG/LineValues/Nested Classes, kjer lahko izberemo katere povezave in ustrezne točke naj se prikažejo in katere ne (primer osebe/dogodki).

## Primer: Slovenski dnevniki in revije

V letih 1999 in 2000 je v anketah Cati Centra več kot 100000 anketirancev odgovarjalo na vprašanje, katere dnevnik in revije berejo. Našteli so skupaj 124 različnih revij. Dobljene podatke lahko predstavimo kot dvovrstno omrežje:

	Delo	Dnevnik	Sl.novice	...
Anketiranec1	X		X	...
Anketiranec2		X		...
Anketiranec3	X			...
.....	.....	.....	.....	.....

Iz dvovrstnega omrežja bralec / revija smo zgenerirali običajno omrežje, kjer so enote revije. Neusmerjena povezava z vrednostjo  $a$  med revijama pomeni, da obe reviji hkrati bere  $a$  anketirancev. Zanka pri posamezni reviji pa pomeni skupno število anketirancev, ki berejo to revijo.

Dobimo matriko  $A$ :

	Delo	Dnevnik	Sl.novice	...
Delo	20714	3219	4214	...
Dnevnik		15992	3642	...
Sl. novice			31997	...
.....	.....	.....	.....	.....

Zaradi velikih razlik v številu bralcev posameznih revij je iz osnovnih podatkov težko sklepati na povezanost med revijami, zato je potrebno omrežje najprej normalizirati. Obstaja več smiselnih normalizacij. Na primer:

$$\text{Geo}_{ij} = \frac{a_{ij}}{\sqrt{a_{ii}a_{jj}}}$$

$$\text{Input}_{ij} = \frac{a_{ij}}{a_{jj}}$$

$$\text{Output}_{ij} = \frac{a_{ij}}{a_{ii}}$$

$$\text{Min}_{ij} = \frac{a_{ij}}{\min(a_{ii}, a_{jj})}$$

$$\text{Max}_{ij} = \frac{a_{ij}}{\max(a_{ii}, a_{jj})}$$

$$\text{MinDir}_{ij} = \begin{cases} \frac{a_{ij}}{a_{ii}} & a_{ii} \leq a_{jj} \\ 0 & \text{sicer} \end{cases}$$

$$\text{MaxDir}_{ij} = \begin{cases} \frac{a_{ij}}{a_{jj}} & a_{ii} \leq a_{jj} \\ 0 & \text{sicer} \end{cases}$$

Kot vidimo so vse normalizacije zgrajene tako, da vsak element v matriki delimo z ustreznim(a) diagonalnim(a) elementom(a).

V nadaljevanju si bomo pogledali rezultate za normalizaciji **Geo** in **MinDir**.



Pri normalizaciji **Geo** se vrednosti normalizirajo z geometrijsko sredino obeh diagonalnih elementov. V našem primeru nam ta normalizacija meri povezanost oz. korelacijo med revijami.

Sliki v SVG: revije3.htm, revije4.htm

---

Normalizacija **MinDir** naredi dvoje:

- neusmerjeno omrežje spremeni v usmerjeno - povezave vedno kažejo od revije z manjšim številom bralcev k reviji z večjim številom bralcem
- vrednost na povezavi pove, kolikšen del bralcev, ki bere prvo revijo, bere tudi drugo

V našem primeru nam ta normalizacija meri odvisnost med revijami.

Sliki v SVG: revije5.htm, revije6.htm.

## Normalizacije v Pajku

Pri pretvorbi dvovrstnega omrežja v običajno omrežje moramo imeti vključeno izbiro **include loops** in izključeno **multiple lines**, sicer pretvorbe ne moremo izvršiti.

V Pajku normaliziramo omrežje, ki smo ga predhodno dobili iz dvovrstnega omrežja z ukazom:

Network/2-Mode Network/2-Mode to 1-Mode/Normalize 1-Mode.

Po pretvorbi najprej odstranimo zanke z

Network/Create New Network/Transform/Remove/loops (ker nas ne zanimajo, oziroma so vrednosti na zankah po normalizaciji tako in tako 1).

Dobljeno omrežje je za prikaz veliko pregosto (vsaka točka je povezana z veliko drugimi), zato pogledamo porazdelitev vrednosti na povezavah z Network/Info/Line Values in glede na to odstranimo iz omrežja povezave z majhnimi vrednostmi:

Network/Create New Network/Transform/Remove/...

...lines with value/lower than

Po tej operaciji ostanejo nekatere točke izolirane (brez sosedov), zato jih odstranimo iz omrežja z ukazom

Network/Create New Network/Transform/...

...Reduction/Degree/All

in odtipkamo 1 ter na dodatno vprašanje dogovorimo z Yes.

Rezultat je lahko omrežje iz večih šibko povezanih komponent (kosov), zato se splača izračunati še šibko povezane komponente in narisati omrežje z uporabo le teh Draw/Network + Partition.

Dobljeno sliko lahko izvozimo v SVG: Export/SVG/Line Values/Nested Classes.

## Otoki

Kadarkoli imamo opravka z omrežjem, pri katerem so poznane neke vrednosti na točkah ali povezah, lahko v takem omrežju poiščemo otoke. Če so vrednosti na točkah, se otoki imenujejo *točkovni*, če pa so vrednosti na povezavah, se imenujejo *povezavni otoki*.

Vzemimo npr. omrežje dobljeno iz dvovrstnega omrežja. Kot vemo, ima to omrežje vrednosti na povezavah. V takem omrežju poiščemo povezavne otoke – skupine točk, med katerimi imajo povezave večjo vrednost kot povezave navzven (znotraj otokov imajo povezave večjo vrednost kot med otoki).

Pri tem postopku moramo izbrati še najmanjšo in največjo dovoljeno velikost otoka. V Pajku se povezavni otoki nahajajo v izbiri Network/Create Partition/Islands/Line Weights

Primer: Davis.net – izračunajmo povezavne otoke velikosti 2 do 6 za obe podomrežji dobljeni iz dvovrstnega omrežja.

## Bločno modeliranje

Tudi na dvovrstnem omrežju lahko poženemo bločno modeliranje, s tem da moramo posebej podati število skupin za prvo in drugo množico. Primer Davis.net.

### Vaje

- Analizirajte in narišite dvovrstno omrežje Davis.net.
- Analizirajte dvovrstno omrežje KdoRolaSlovenijo KdoRola.net.
- Analizirajte in narišite omrežje z vrednostmi na povezavah revije.net (število bralcev, ki bere izbrani reviji), ki je dobljeno iz dvovrstnega omrežja. Pomagajte si z ustreznimi normalizacijami in narišite slike v SVG.
- Za vsa omrežja iz prejšnjih nalog izračunajte še povezavne otoke.