

Social Network Analysis **SNA**



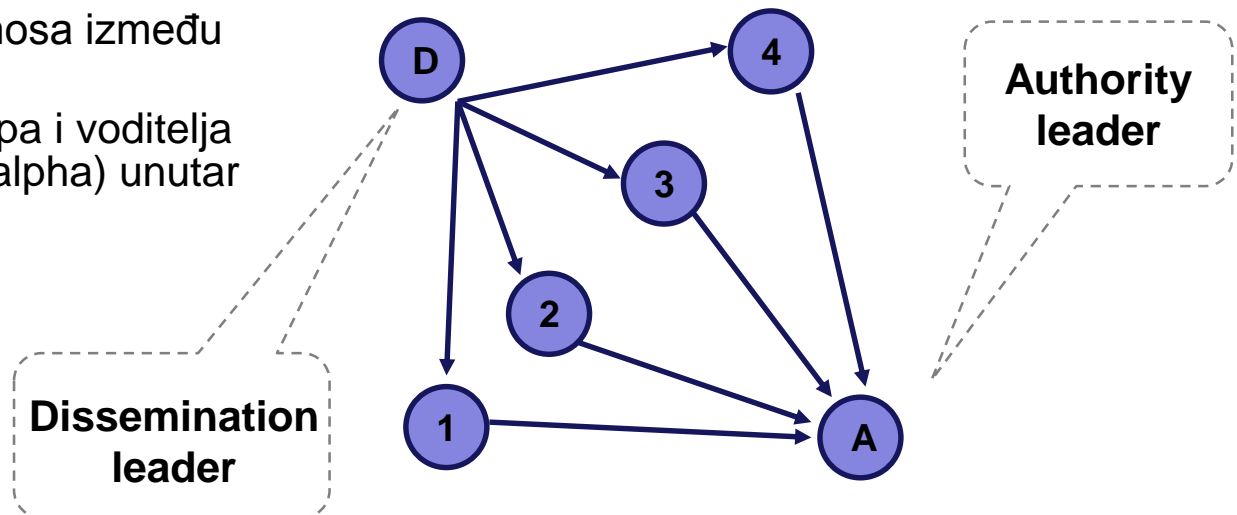
Zagreb, 12.04.2016

SNA - Općenito

- Analiza grupe
- Difuzijska analiza
- Obje analize zahtijevaju zapise o interakcijama
- Difuzijska analiza dodatno zahtijeva listu korisnika s analiziranom aktivnošću

Analiza grupe:

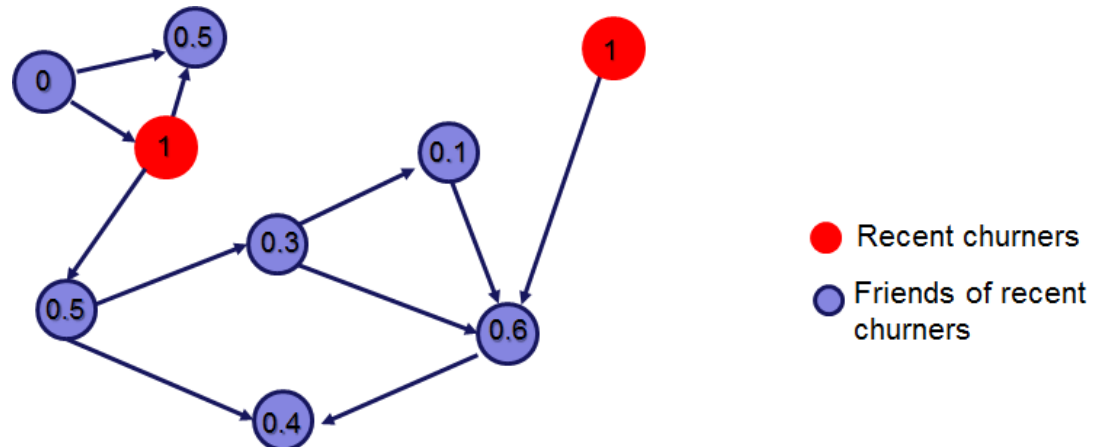
- Identifikacija odnosa između pojedinaca
- Identifikacija grupa i voditelja grupa (leadera, alpha) unutar grupa



SNA - Općenito

Difuzijska analiza

- Predviđa utjecaj koji korisnici specifične aktivnosti imaju na svoje 'mrežne susjede'
- Utjecaj korisnika specifične aktivnosti računa se prijenosom odnosno difuzijom energije s jednog korisnika na drugog
- Lista izvora za difuziju (diffusion seed list): identificira mrežne node (korisnike) koji šire difuzijsku energiju



SNA - Groupna Analiza

Grupna Analiza ključni koraci:

- Odrediti jačine veza (odražava socijalnu povezanost (social proximity))
- Podijeliti mrežu u grupe (radi se na bazi jačine veza, uz poštovanje ograničenja)
- Profiliranje grupa i korisnika (grupna statistika, vođe grupa, statistika pojedinca)
- Ograničavajući parametri za definiciju grupa unutar mreže:
 - Prag pokrivenosti (Coverage treshold) (% najjačih veza)
 - Minimalna veličina grupe (npr. 2)
 - Maximalna veličina grupe (npr. 100)



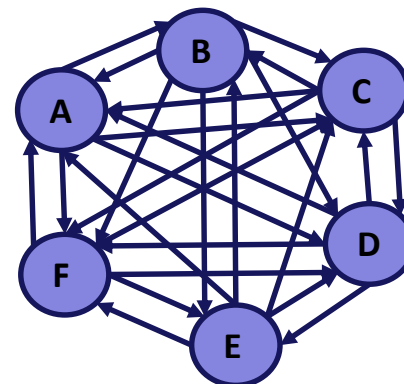
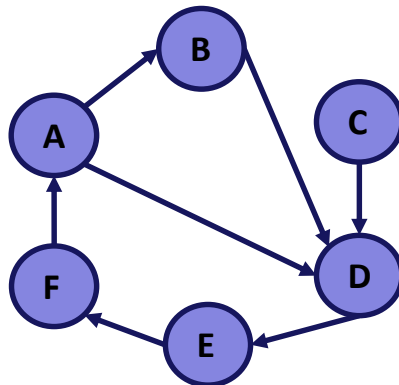
SNA - Group Analysis

Opis grupe:

- Gustoća grupe (prosječni udio direktnih veza između noda u grupi)
- In degree (prosječni broj ulaznih veza)
- Out degree (prosječni broj izlaznih veza)

Opis čvorova

- **Authority score** (mjera tendencije drugih čvorova da se povežu s danim čvorom)
 - Authority leader (čvor s najvišim authority scoreom)
- **Dissemination score** (mjera tendencije čvora da se povezuje s drugim čvorovima)
 - Dissemination leader (čvor s najvišim dissemination scoreom)



SNA - Grupna analiza – statistika na nivou mreže



- **Ukupan broj noda u grupama**
- **Broj veza u grupama**
- **Ukupan broj grupa** – Broj grupa identificiranih u mreži
- **Prosječna veličina grupe** – Prosječni broj čvorova u grupi
- **Prosječna gustoća grupe** – Prosječni udio direktnih veza između čvorova u grupi u odnosu na mogući broj veza u grupi
- **Prosječni udio ključnih članova grupe** – Prosječni udio čvorova koji predstavljaju ključne članove grupe u odnosu na ukupni broj članova grupe
- **Prosječna gustoća veza između ključnih članova grupe** – prosječni udio direktnih veza između ključnih članova grupe
- **Prosječni broj ulaznih veza (InDegree)**
- **Prosječni broj izlaznih veza (OutDegree)**

SNA - Grupna analiza – statistika na nivou grupe

- **Broj grupe** (ID)
- **Broj čvorova u grupi**
- **Gustoća grupe** (udio direktnih veza između čvorova u grupi u odnosu na mogući broj veza u grupi)
- **Gustoća jezgre grupe** (udio direktnih veza između ključnih članova grupe)
- **Udio ključnih članova u grupi**
- **Maksimalni rang za authority score**
- **Minimalni rang za authority score**
- **Maksimalni rang za dissemination score**
- **Minimalni rang za dissemination score**



SNA - Grupna analiza – statistika na nivou člana mreže



- **Identifikator člana grupe**
- **Identifikator je li čvor ključan ili ne**
- **Authority score**
- **Rang u grupi bazirano na authority scoreu**
- **Dissemination score**
- **Rang u grupi bazirano na dissemination scoreu**
- **Broj veza u kojima je čvor izvor veze**
- **Broj veza u kojima je čvor cilj veze**
- **Indikator je li čvor Authority leader**
- **Vjerojatnost da je čvor Authority leader**
- **Indikator je li čvor Dissemination leader**
- **Vjerojatnost da je čvor Dissemination leader**

SNA - Group Analysis Results

Jedan uzorak, 4 modela:

- **Prag pokrivenosti: 99%, 10%**
- **Težinski faktori: yes, no**

Threshold: 99%
 Weight : no

Name	Value
Total Nodes in Groups	10
Total Links in Groups	12
Total Number of Groups	3
<u>Mean Group Size</u>	3,33
<u>Mean Group Density</u>	0,76
<u>Mean Fraction of Core Members</u>	0,94
<u>Mean Density of Core Group</u>	0,80
<u>Mean In-Degree</u>	1,20
<u>Mean Out-Degree</u>	1,20

Threshold: 10%
 Weight : no

Name	Value
Total Nodes in Groups	8
Total Links in Groups	9
Total Number of Groups	2
<u>Mean Group Size</u>	4,00
<u>Mean Group Density</u>	0,40
<u>Mean Fraction of Core Members</u>	0,63
<u>Mean Density of Core Group</u>	0,83
<u>Mean In-Degree</u>	1,12
<u>Mean Out-Degree</u>	1,12

Threshold: 99%
 Weight : yes

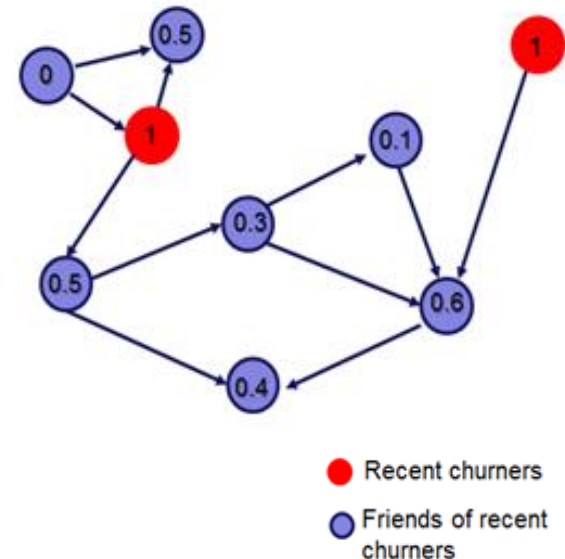
Name	Value
Total Nodes in Groups	10
Total Links in Groups	13
Total Number of Groups	1
<u>Mean Group Size</u>	10,00
<u>Mean Group Density</u>	0,14
<u>Mean Fraction of Core Members</u>	1,00
<u>Mean Density of Core Group</u>	0,20
<u>Mean In-Degree</u>	1,30
<u>Mean Out-Degree</u>	1,30

Threshold: 10%
 Weight : yes

Name	Value
Total Nodes in Groups	5
Total Links in Groups	6
Total Number of Groups	1
<u>Mean Group Size</u>	5,00
<u>Mean Group Density</u>	0,30
<u>Mean Fraction of Core Members</u>	0,40
<u>Mean Density of Core Group</u>	1,00
<u>Mean In-Degree</u>	1,20
<u>Mean Out-Degree</u>	1,20

SNA - Analiza Difuzije

- Identifikacija korisnika (čvorova) koji su pod najvećim utjecajem drugih korisnika (čvorova) u mreži
- **Difuzijski parametri:**
 - Faktor širenja
 - Najveći broj iteracija
 - Prag točnosti
- **Difuzijski proces će biti zaustavljen:**
 - Aktivirane node nisu izvor direktne veze
 - Količina energije koja se transferira je manja od specificiranog praga točnosti
 - Broj iteracija je dosegao gornju granicu
- **Difuzijska lista izvora energije:**
 - Identificira čvorove u mreži koji su izvor energije za difuziju
 - Primjena: Churn prediction
Early adopters



SNA - Difuzijska analiza – Statistika

Statistika na novu mreže:

- Broj noda u mreži
- Broj veza u mreži
- Broj noda koji su izvori difuzijske energije
- Prosječni utjecaj - prosječna vrijednost difuzijske energije po čvoru
- Prosječni broj veza u kojima je čvor cilj veze
- Prosječni broj veza u kojima je čvor izvor veze

Statistika na nivou čvora

- Identifikator čvora
- Količina difuzne energije za čvor
- Broj veza u kojima je čvor cilj veze (može biti s težinama)
- Broj veza u kojima je čvor izvor veze (može biti s težinama)



SNA - Diffusion Analysis Example 1

Source	Target	Weight
21	22	100
21	22	150
21	22	120
22	21	200
22	21	110
22	21	130
23	24	200
24	23	150
24	21	120
25	23	100
23	25	99
21	82	90
81	22	80
81	82	100
81	82	150
81	82	120
82	81	200
82	81	110
82	81	130
83	84	200
84	83	150
84	81	120
85	83	100
83	85	99

Diffusion factor: 30%

	DA_NodeNumber	DA_DiffusedEnergy	DA_InDegree	DA_OutDegree	DA_WeightedInDegree	DA_WeightedOutDegree
1	21	0.979	2	2	4.000	4.000
2	22	0.932	2	1	4.000	3.000
3	85	0.752	1	1	1.000	1.000
4	84	0.752	1	2	1.000	2.000
5	83	0.336	2	2	2.000	2.000
6	81	0.143	2	2	4.000	4.000
7	82	0.107	2	1	4.000	3.000
8	25	0.000	1	1	1.000	1.000
9	24	0.000	1	2	1.000	2.000
10	23	0.000	2	2	2.000	2.000

Diffusion factor: 70%

	DA_NodeNumber	DA_DiffusedEnergy	DA_InDegree	DA_OutDegree	DA_WeightedInDegree	DA_WeightedOutDegree
1	21	0.893	2	2	4.000	4.000
2	22	0.840	2	1	4.000	3.000
3	83	0.499	2	2	2.000	2.000
4	85	0.474	1	1	1.000	1.000
5	84	0.474	1	2	1.000	2.000
6	81	0.439	2	2	4.000	4.000
7	82	0.380	2	1	4.000	3.000
8	25	0.000	1	1	1.000	1.000
9	24	0.000	1	2	1.000	2.000
10	23	0.000	2	2	2.000	2.000

SNA - Diffusion Analysis Example 2

Source	Target	Weight
21	22	100
21	22	150
21	22	120
22	21	200
22	21	110
22	21	130
23	24	200
24	23	150
24	21	120
25	23	100
23	25	99
21	82	90
81	22	80
81	82	100
81	82	150
81	82	120
82	81	200
82	81	110
82	81	130
83	84	200
84	83	150
84	81	120
85	83	100
83	85	99

Diffusion factor: 30%

	DA_NodeNumber	DA_DiffusedEnergy	DA_InDegree	DA_OutDegree	DA_WeightedInDegree	DA_WeightedOutDegree
1	23	1.088	2	2	2.000	2.000
2	25	0.864	1	1	1.000	1.000
3	24	0.864	1	2	1.000	2.000
4	21	0.138	2	2	4.000	4.000
5	22	0.032	2	1	4.000	3.000
6	82	0.011	2	1	4.000	3.000
7	81	0.003	2	2	4.000	4.000
8	85	0.000	1	1	1.000	1.000
9	84	0.000	1	2	1.000	2.000
10	83	0.000	2	2	2.000	2.000

Diffusion factor: 70%

	DA_NodeNumber	DA_DiffusedEnergy	DA_InDegree	DA_OutDegree	DA_WeightedInDegree	DA_WeightedOutDegree
1	23	0.975	2	2	2.000	2.000
2	25	0.640	1	1	1.000	1.000
3	24	0.640	1	2	1.000	2.000
4	21	0.372	2	2	4.000	4.000
5	22	0.202	2	1	4.000	3.000
6	82	0.098	2	1	4.000	3.000
7	81	0.073	2	2	4.000	4.000
8	85	0.000	1	1	1.000	1.000
9	84	0.000	1	2	1.000	2.000
10	83	0.000	2	2	2.000	2.000

SNA - Ostali izrazi

- Centrality** – odnosi se na grupu mjera kojima se opisuje važnost ili utjecajnost čvora u grupi
- **Degree Centrality** – broj direktnih izlaznih veza danog čvora s ostalim čvorovima u grupi
 - **Betweenness Centrality** – kvantificira koliko puta se čvor nalazi na najkraćem putu između druge dvije node u grupi
 - **Closeness Centrality** – kolika je udaljenost node do svih ostalih noda u grupi
 - **Eigenvector Centrality** – mjeri utjecaj node u smislu s koliko ključnih nodi je povezana (ako je povezana s ključnim nodama više doprinosi)
 - **Alpha Centrality** – dodatak na Eigenvector Centrality gdje je važnost node određena vanjskim faktorom

Modularity – razbijanje mreže na klastere u kojima je povezanost među čvorovima mnogo jača nego prema čvorovima van klastera

Page rank – određivanje važnosti node u grupi na temelju veza prema drugim nodama

Prestige – važnost node izračunata na temelju veza s drugim čvorovima, ali se u obzir uzimaju i ulazne i izlazne veze (za razliku od centralnosti gdje se gledaju samo izlazne veze)

