

Analiza socijalnih mreža

Osnovna svojstva socijalnih mreža

Marko Mišić, Jelica Protić

13M111ASM

2019/2020.

Osnovne mrežne metrike (1)

- Metrike bazirane na teoriji grafova
 - Često zasnovane na statističkim metodima
 - Opisuju osnovnu strukturu grafa
- Često korišćene metrike:
 - Veličina mreže, gustina mreže, fragmentacija
 - Metrike u vezi stepena čvora
 - Minimalni, maksimalni, prosečan stepen čvora
 - Metrike u vezi distanci
 - Povezanost, geodezijska distanca, eksecntričnost čvora, dijametar, protoci

Osnovne mrežne metrike (2)

- Računaju se na nivou pojedinačnih čvorova ili na nivou cele mreže
- Manje razlike za usmerene i neusmerene mreže
- Neke uzimaju u obzir težine grana, a neke ne
 - Potreba za dihotomizacijom mreže
 - Stvaranje binarnog grafa od težinskog
 - Zasnovana na srednjoj, medijalnoj ili nekoj određenoj vrednosti grana u mreži

Metrike u vezi relacija (1)

- Stepen čvora se odnosi na broj grana koje su incidentne (susedne) posmatranom čvoru
 - U direktnoj je vezi sa centralnošću po stepenu
 - Od interesa utvrditi minimalan, maksimalan i prosečan stepen čvorova u mreži, kao i standardnu devijaciju
 - Govori o potencijalju za saradnju
- Razlikuje se ulazni i izlazni stepen kod usmerenih mreža
- Često se određuje distribucija čvorova po stepenu

Metrike u vezi relacija (2)

- Dostižnost čvora u mreži pokazuje da li se iz nekog čvora mreže mogu dosegnuti preostali čvorovi
 - Ukoliko neki čvorovi nisu dostižni u mreži, to može ukazivati na određene podele unutar nje ili postojanje izolovanih čvorova ili skupina čvorova
- Povezanost čvorova definiše broj čvorova koji treba ukloniti na putanji između dva posmatrana čvora da bi oni postali nepovezani
 - Ukoliko postoji veći broj puteva između dva aktera u mreži, tada su oni povezani

Metrike u vezi relacija (3)

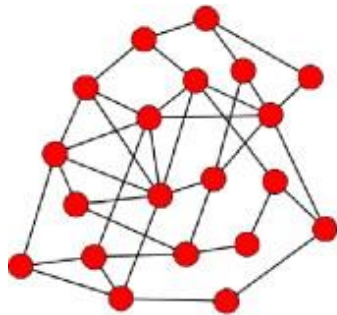
- Veličina mreže se odnosi na ukupan broj čvorova u mreži
- Na osnovu nje se može definisati maksimalni mogući stepen svakog čvora
 - To je $n-1$, gde je n broj čvorova u mreži
 - Gornja granica broja konekcija u mreži (stepena čvora)
- Veličina mreže određuje broj mogućih relacija u mreži
 - Kod usmerenih mreža iznosi $n*(n-1)$
 - Kod neusmerenih mreža iznosi $n*(n-1)/2$
- Kompleksnost mreže raste eksponencijalno sa njenom veličinom

Metrike u vezi relacija (4)

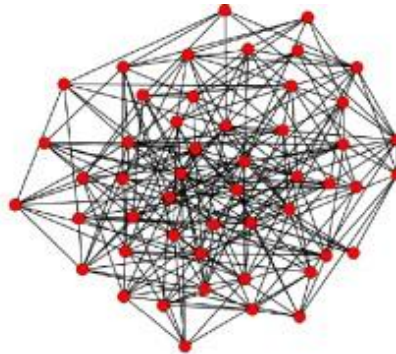
- Gustina mreže se računa kao količnik ukupnog broja grana i maksimalnog mogućeg broja grana u mreži
 - Ukazuje na stepen povezanosti mreže
 - U gušćim mrežama, postoji znatno više veza među čvorovima, čime se omogućava bolji protok informacija
 - U retkim mrežama, bolje pozicionirani pojedinci mogu da kontrolišu širenje informacija i sl.
- Fragmentacija mreže se odnosi na proporciju parova čvorova u mreži koji nisu dostižni u odnosu na ukupan broj parova

Metrike u vezi relacija (5)

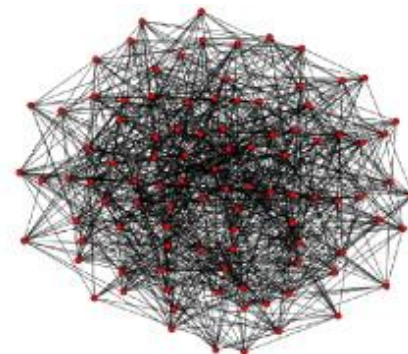
- Gustina mreže može dovesti do pogrešnih zaključaka, ukoliko se porede mreže različite veličine
- Alternativna metrika za poređenje je links per node
 - Predstavlja odnos broja konekcija i broja čvorova u mreži
 - Može biti značajno različit za mreže iste gustine!



20 nodes, 38 links
Density = 0.20
Links per node = 1.9



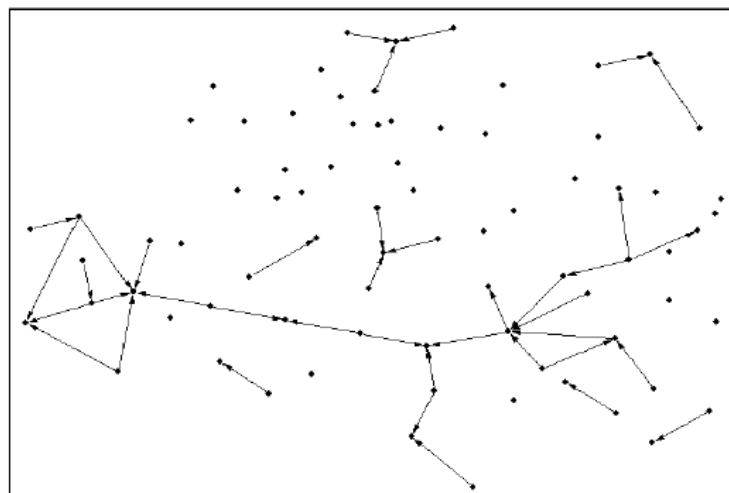
50 nodes, 245 links
Density = 0.20
Links per node = 4.9



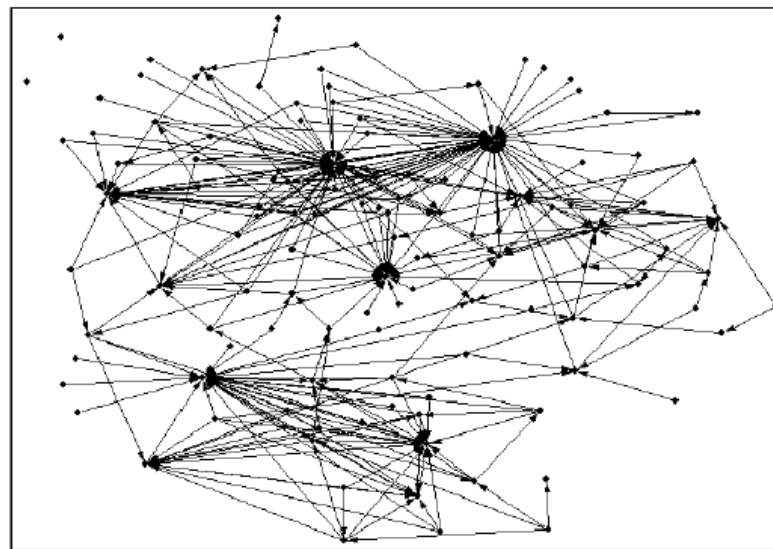
100 nodes, 990 links
Density = 0.20
Links per node = 9.9

Metrike u vezi relacija (6)

- U gušćim mrežama postoji bolji potencijal za saradnju
 - Mera raspoloživih socijalnog kapitala
 - *Help with the rice harvest* primer (*Entwistle et al.*)



Village 1



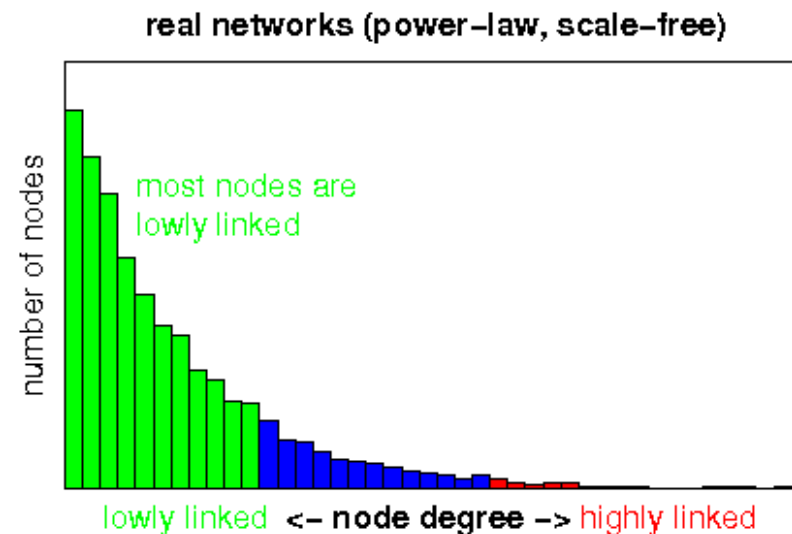
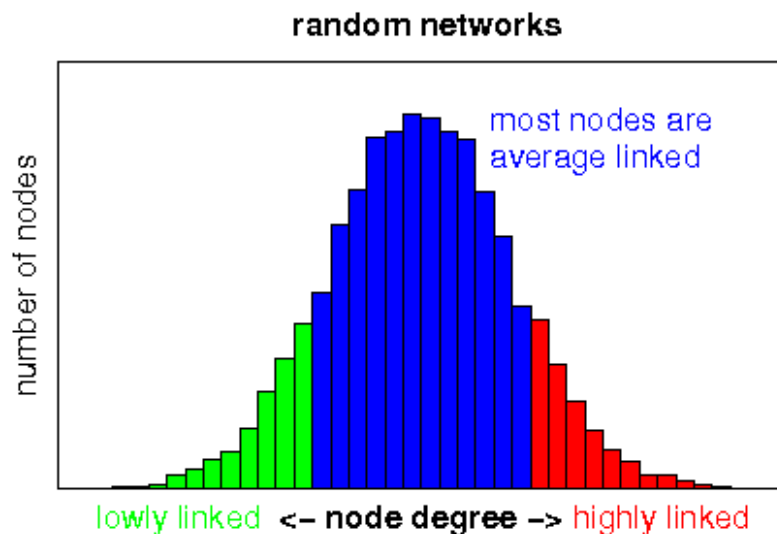
Village 2

Distribucija čvorova po stepenu (1)

- U tradicionalnim, *random* mrežama, većina čvorova ima stepen čvora blizak proseku
 - Stepeni čvorova su distribuirani oko proseka
- U realnim (društvenim) mrežama, većina čvorova ima nekoliko veza, ali postoje čvorovi sa veoma velikim brojem veza
 - Takve mreže prate *power-law* (*scale-free*) raspodelu
- Neka $P(k)$ predstavlja procenat čvorova mreže kod kojih čvor ima stepen k
- *Power law* raspodela se može opisati kao:
$$P(k) \sim k^{-\gamma}, \text{ gde } \gamma \text{ je konstanta } (\sim 2 < \gamma < 3)$$

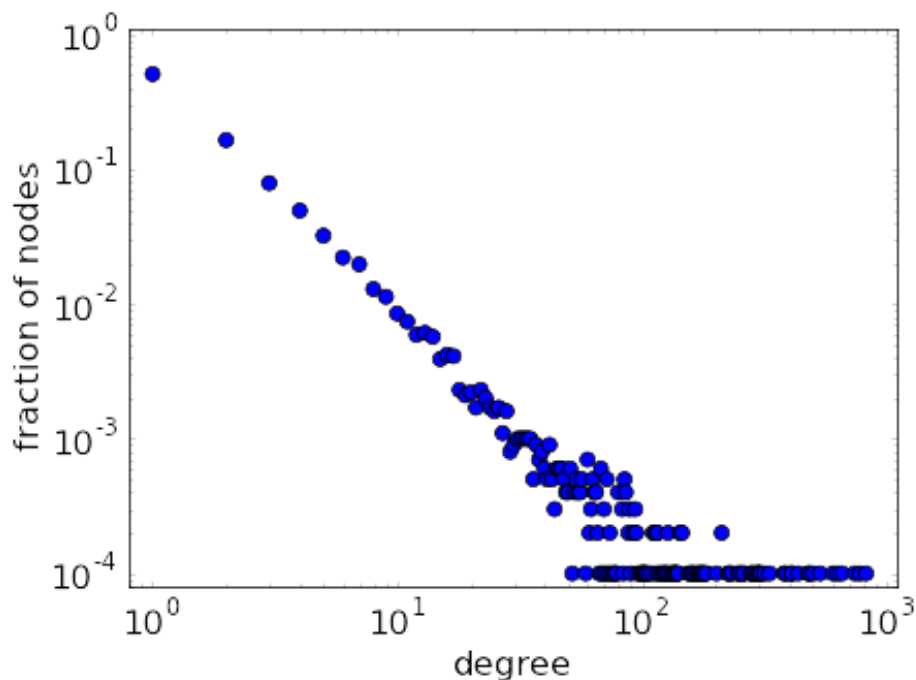
Distribucija čvorova po stepenu (2)

- *Power law* raspodela je karakteristika tzv. *scale free* mreža
- Realne mreže asimptotski prate ovu raspodelu
 - Izraženije ukoliko su mreže retke

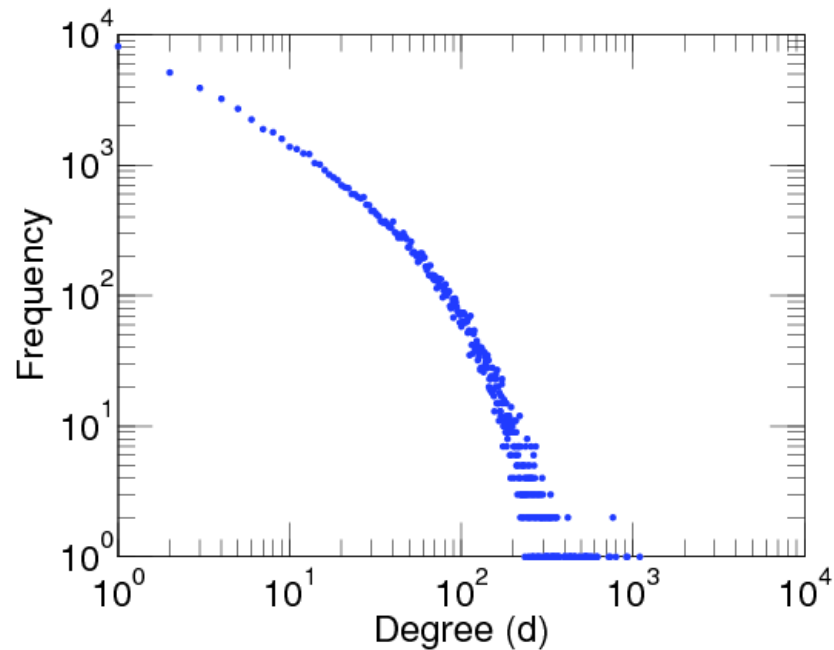


Distribucija čvorova po stepenu (3)

- Istraživanja pokazuju da raspodela opstaje tokom vremena u realnim mrežama



Generisana mreža,
 $N=10,000$, $\gamma = 2$



Facebook,
63,731 aktera, 817,035 veza

Distance u mrežama (1)

- Koncept distanci se realizuje kroz nekoliko metrika
 - Govore o načinu širenja i dosegu informacija
 - U mrežama sa velikim distancama, širenje informacija može biti usporeno
 - Opisuju način na koji je akter ugrađen u mrežu
- Prost put (*walk*) u mreži se definiše kao niz aktera i relacija koji počinju i završavaju se akterom
 - Na složenom putu su dozvoljena ponavljanja aktera
 - Često od interesa samo putevi određene dužine
- Ciklus je prost put koji počinje i završava se istim čvorom

Distance u mrežama (2)

- Geodezijska distanca predstavlja broj grana na najkraćoj putanji između dva aktera u mreži
 - Putanja u širinskom (BFS) obuhvatnom stablu grafa
 - Ukoliko su dva aktera susedi unutar mreže, onda ova distanca ima vrednost jedan
- Ponekad se geodezijska distanca transformiše u meru bliskosti
 - *Nearness* transformacija

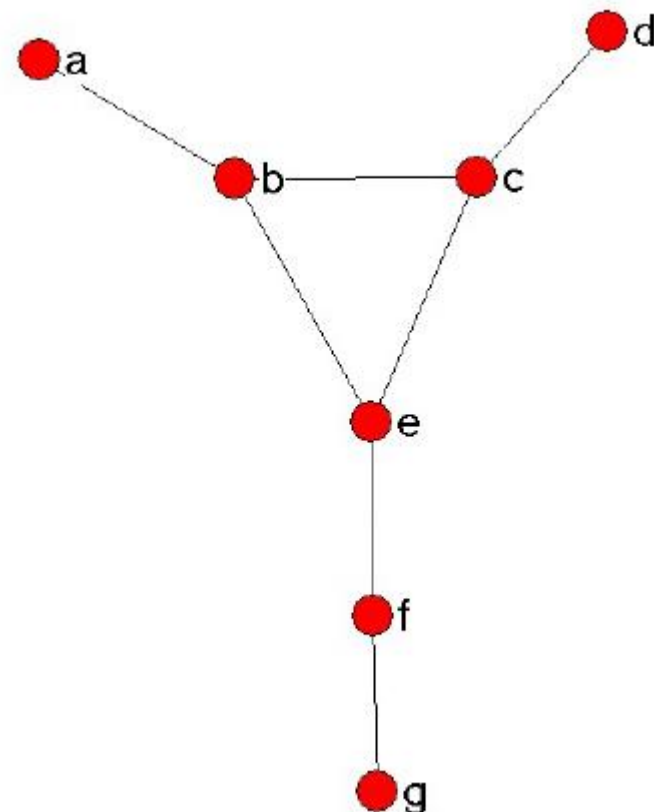
Distance u mrežama (3)

- Kod težinskih grafova, geodezijska distanca se može računati i na drugačiji način:
 - Suma težina na putu između A i B (*cost*)
 - Minimum težina svih grana na putu između A i B (*strength, maximum flow*)
 - Proizvod težina na putu između A i B (*probability*)
- Floyd-Warhshall, Dijkstra algoritmi se koriste za određivanje ovako definisanih metrika

Distance u mrežama (4)

- Matrica geodezijskih distanci se često koristi za pamćenje rastojanja

	a	b	c	d	e	f	g
a	0	1	2	3	2	3	4
b	1	0	1	2	1	2	3
c	2	1	0	1	1	2	3
d	3	2	1	0	2	3	4
e	2	1	1	2	0	1	2
f	3	2	2	3	1	0	1
g	4	3	3	4	2	1	0



Distance u mrežama (5)

- Ekscentričnost čvora predstavlja najveću geodezijsku distancu čvora
 - Govori o udaljenosti u odnosu na najdalji čvor sa kojim je on povezan
- Dijametar mreže je jednak najdužoj od svih najkraćih putanja u mreži između parova čvorova
 - Služi kao indikator udaljenosti čvorova, ali i veličine mreže
- Prosečna dužina putanja u mreži ponekad bolji indikator od dijametra

Literatura

- Hanneman, Robert A. and Mark Riddle, Introduction to social network methods, University of California, Riverside, 2005.
- J. Jovanović, Softverska analiza društvenih mreža, FON, 2017.
- M. Mišić, Unapređenja sistema za detekciju plagijarizma u izvornom programskom kodu – Primena metoda za analizu socijalnih mreža u detekciji plagijarizma, PhD disertacija, ETF 2017.
- B. Hoppe, C. Reinelt, Social Network Analysis and the Evaluation of Leadership Networks, Leadership Quarterly, 21, 2010. pp. 600-619.
- <http://www.network-science.org/>