

Analiza socijalnih mreža

Mere centralnosti

Marko Mišić, Jelica Protić

13M111ASM

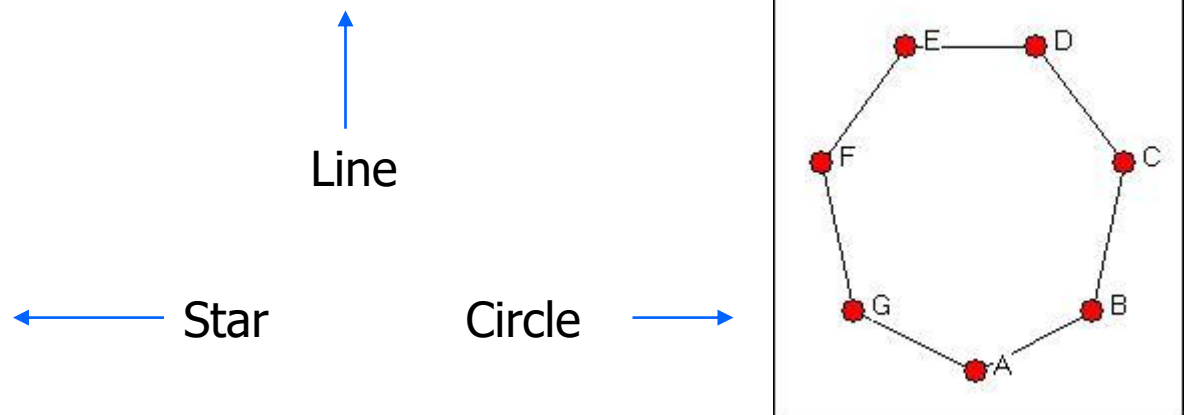
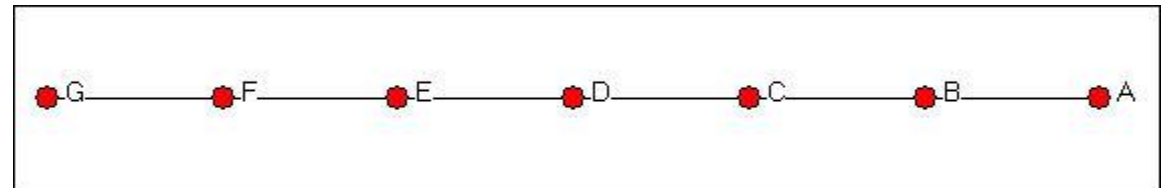
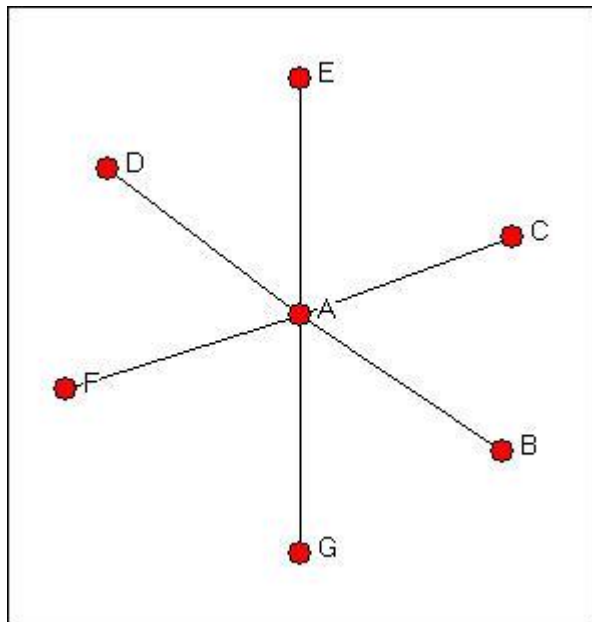
2019/2020.

Motivacija (1)

- Problem raspodele „moći“ u socijalnim strukturama
 - Moć je posledica zasnovanih relacija
 - Akter poseduje moć, jer je u prilici da dominira drugima
 - Ograničavanjem pristupa informacijama, blokadom puteva...
 - Distribucija moći može biti značajno nejednaka
- Pozicija aktera u mreži je odlučujuća za razumevanje njegove moći
 - Manje ograničenja => više moći
- Mere centralnosti mogu pomoći u razumevanju položaja pojedinca u mreži

Motivacija (2)

- Primer različiti topologija mreža
 - Razmotriti položaj čvora A u smislu mogućnosti za saradnju, trgovinu ili razmenu



Motivacija (3)

| | Zvezda | Krug | Linija |
|----------------|--|--|--|
| Broj konekcija | A ima mnogo više mogućnosti od ostalih Ostali akteri ograničeni na A | Svi akteri imaju podjednak broj mogućih partnera | A i G vidno ograničeni Centralniji akteri u povoljnijem položaju |
| Bliskost | A je u proseku bliži svim ostalim akterima A ima geo. distancu 1 do ostalih | Svi akteri ravnopravni | D je u proseku najbliži svima |
| Posredovanje | A leži na putevima između svih ostalih aktera | Svaki akter ima dva moguća puta do nekog drugog | A i G ne leže ni na jednom putu D leži na najviše puteva |

Mere centralnosti (1)

- Otkrivaju koji su čvorovi bitni za održavanje povezanosti mreže
- Daju uvid u položaj pojedinačnih aktera u mreži
 - Kako u odnosu na okruženje, tako i u odnosu na celu mrežu
- Često korišćene metrike u analizi mrežnih grafova
 - Freeman, 1977.
- Mogu se računati za pojedinačne čvorove, ali i za celu mrežu
 - Mera centralizacije mreže po različitim metrikama

Mere centralnosti (2)

- Dva nivoa razmatranja: lokalni i globalni
 - Zavise od veličine mreže
 - Odražava se na mere centralnosti
- Lokalni nivo forsira broj konekcija sa neposrednim okruženjem
 - Čvor je lokalno centralan ako ima veći broj veza sa svojim okruženjem
- Globalni nivo forsira distance
 - Čvor je globalno centralan, ukoliko je u proseku bliži od drugih ostalim čvorovima

Centralnost po stepenu (1)

- *Degree centrality* – lokalna centralnost
- Određuje se kao broj direktnih suseda svakog čvora u mreži
 - Kod neusmerenih grafova, jednaka je broju incidentnih grana tom čvoru
 - Ukoliko se koristi matična reprezentacija grafa sa N čvorova, može se izračunati na sledeći način

$$C_D(p_i) = \sum_{k=1}^N A(p_i, p_k)$$

- Kod usmerenih grafova, razlikuje se centralnost po ulaznom i izlaznom stepenu

Centralnost po stepenu (2)

- Težine grana se mogu uzimati u obzir
 - Na značaju dobijaju čvorovi koji imaju relacije sa većom težinom
- Za poređenje sa drugim mrežama koristi se normalizovana vrednost centralnosti po stepenu
 - Iskazuje se procentualno u odnosu na veličinu mreže umanjenu za jedan (posmatrani čvor)

Centralnost po stepenu (3)

○ Interpretacija

- Akteri u mreži sa većim brojem relacija imaju više alternativnih načina da zadovolje svoje potrebe
 - Manje su zavisni u odnosu na druge aktere
 - Visoka centralnost po ulaznom stepenu može govoriti o popularnosti osobe
 - Visoka centralnost po izlaznom stepenu može govoriti o prijateljski nastrojenoj osobi
- ## ○ Ne uzima u obzir način na koji su susedi povezani
- „Ugrađenost” i pozicija u mreži se ignorišu

Centralnost po bliskosti (1)

- *Closeness centrality* – globalna centralnost
- Definiše kao recipročna vrednost sume najkraćih rastojanja od posmatranog čvora do ostalih sa kojima je povezan

$$C_C(p_i) = \frac{N - 1}{\sum_{k=1}^N d(p_i, p_k)}$$

- $d(p_i, p_k)$ predstavlja najkraće rastojanje između posmatranog čvora i nekog od preostalih u mreži sa kojima je povezan
- N je ukupan broj čvorova u grafu
- Suma je normalizovana veličinom mreže umanjenom za jedan
 - Zbog poređenja mreža različitih dimenzija

Centralnost po bliskosti (2)

- Interpretacija
 - Prosečna udaljenost čvora od svih ostalih čvorova u grafu sa kojima je povezan
 - Čvorovi sa većom vrednošću na centralnijim pozicijama u mreži
 - U proseku potrebno manje koraka da dođu do preostalih čvorova u mreži
 - Razlika kod usmerenih grafova
 - U zavisnosti da li se udaljeno računa od ili do drugih čvorova
 - Odražava brzinu kojom informacije mogu da se šire od nekog čvora
- U slabo povezanim grafovima koristi se varijanta koja se naziva harmonijska centralnost

Relaciona centralnost (1)

- *Betweenness centrality*
- Proporcionalna je broju najkraćih puteva između svih ostalih parova čvorova na kojima se posmatrani čvor nalazi u odnosu na ukupan broj ^(ista) takvih najkraćih puteva

$$C_B(p_i) = \sum_{j=1}^N \sum_{k=1}^{j-1} \frac{g_{jk}(p_i)}{g_{jk}}$$

- g_{jk} predstavlja ukupan broj najkraćih puteva koji povezuju čvorove p_j i p_k
- $g_{jk}(p_i)$ predstavlja broj takvih puteva koji uključuju čvor p_i
- N je ukupan broj čvorova u grafu

Relaciona centralnost (2)

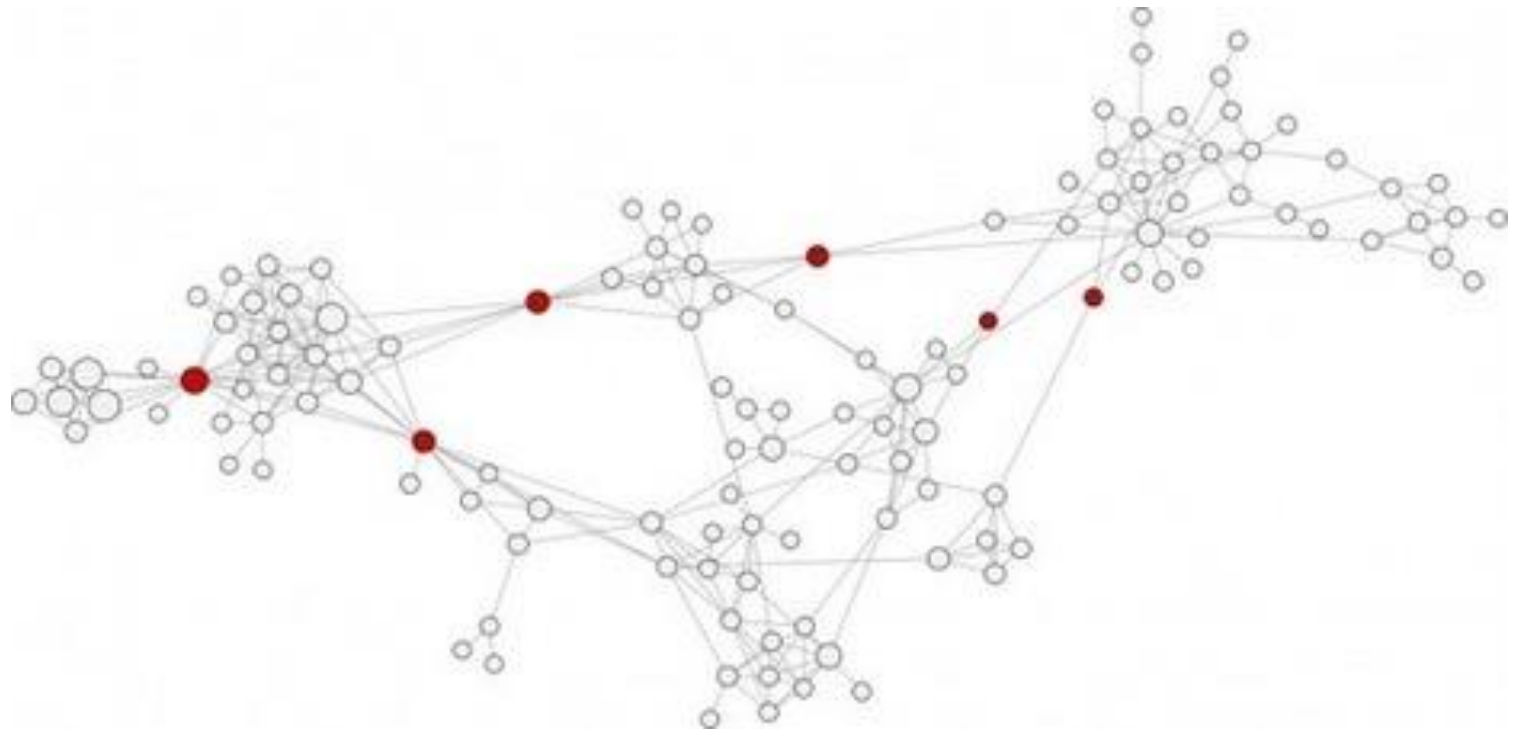
- Algoritam za izračunavanje za neki čvor v u grafu $G(V,E)$ sa N čvorova
 - Za svaki par čvorova (s,t) , odrediti skup najkraćih puteva između njih
 - Za svaki par čvorova (s,t) , odrediti frakciju (broj) puteva koji prolaze kroz čvor v
 - Odrediti ukupan zbir ovih frakcija za sve parove čvorova u mreži
- Normalizovana vrednost se dobija podelom dobijene sume sa ukupnim brojem parova čvorova (s,t) koji ne uključuje v
 - Za usmerene grafove: $(n-1)*(n-2)$
 - Za neusmerene grafove: $(n-1)*(n-2)/2$

Relaciona centralnost (3)

- Kvantifikuje kontrolu komunikacije unutar mreže od strane pojedinih aktera
- Interpretacija
 - Čvorovi sa višim vrednostima relacione centralnosti se nalaze na većem broju najkraćih putanja
 - Ovakvi čvorovi predstavljaju „mostove“ ili posrednike između ostalih čvorova u mreži
 - Označava broj čvorova koji su indirektno povezani u mreži preko direktnih grana
 - Mogu biti i tačka prekida u okviru mreže

Relaciona centralnost (4)

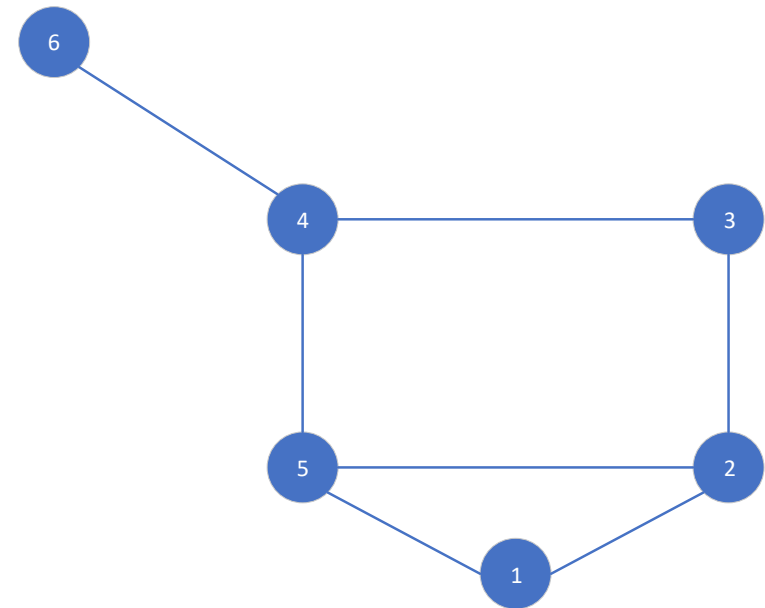
- Čvorovi sa visokom relacionom centralnošću su često na periferiji grupa unutar mreže



Relaciona centralnost (5)

- Primer mreže sa 6 aktera

| Čvor | Relaciona centralnost |
|------|-----------------------|
| 1 | 0 |
| 2 | 1.5 |
| 3 | 1 |
| 4 | 4 |
| 5 | 5 |
| 6 | 0 |



Centralnost po svojstvenom vektoru (1)

- *Eigenvector centrality*
- Varijanta centralnosti po stepenu koja uzima u obzir i susedstvo posmatranog čvora
- Koristi koncepte uticaja i moći
 - Čvor je uticajniji ukoliko njegovi susedi takođe imaju veliki broj suseda
 - Poruka koju on pošalje može brzo da stigne do velikog broja aktera
 - Čvor je moćniji ukoliko njegovi susedi nemaju veliki broj svojih suseda
 - Onda su zavisni od posmatranog čvora

Centralnost po svojstvenom vektoru (2)

- Računa se određivanjem svojstvenog vektora koji sadrži relativne centralnosti x_i za svaki čvor

$$x_i = \frac{1}{\lambda} \sum_{k \in M(i)} x_k = \frac{1}{\lambda} \sum_{k \in G} A(p_i, p_k) x_k$$

- λ je neka konstanta
- A matrica susednosti za posmatranu mrežu
- $M(i)$ skup suseda čvora p_i
- Mera centralnosti se dobija pronalaženjem najveće odgovarajuće sopstvene vrednosti λ , što se vrši dalje iterativnim postupkom

Centralnost po svojstvenom vektoru (3)

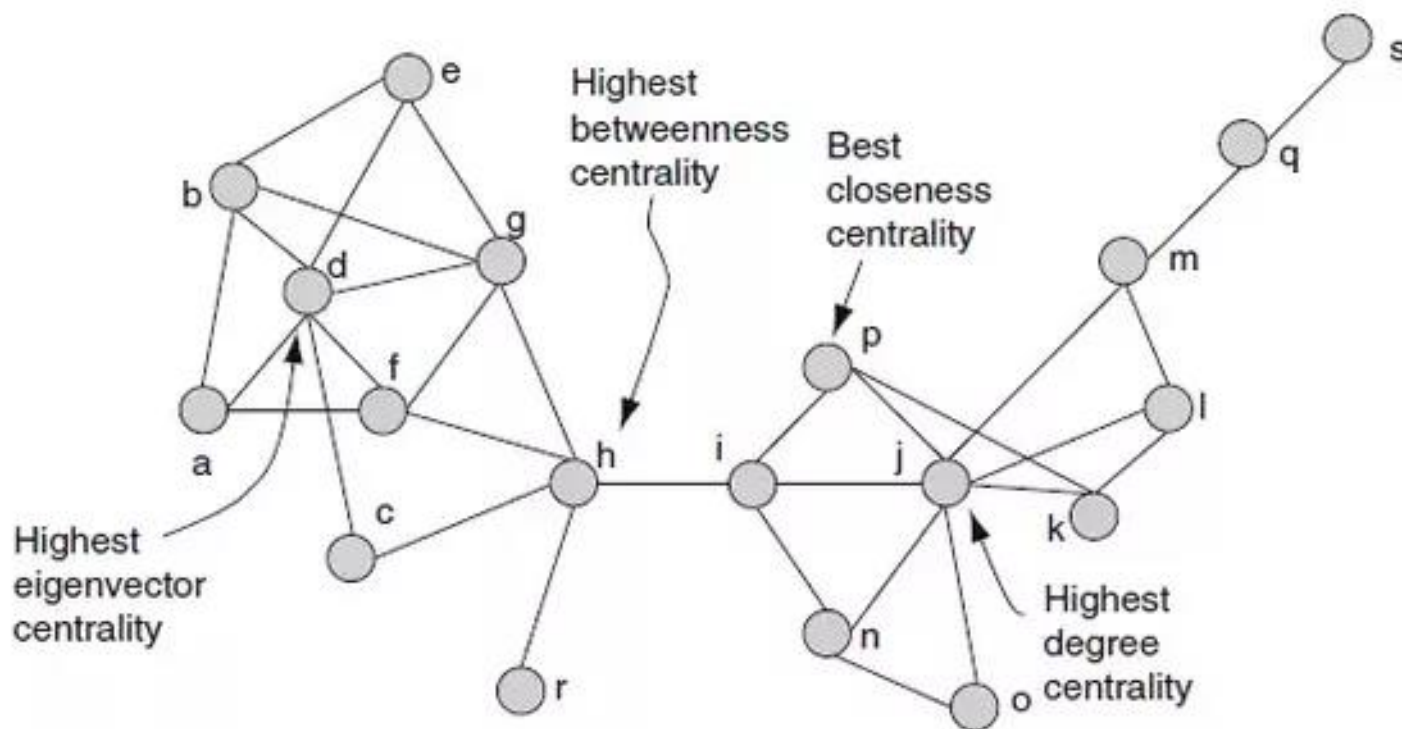
- *Beta (Bonacich) centrality*
 - Varijanta centralnosti po svojstvenom vektoru
- Uvodi dodatan beta parametar za ponderisanje
 - Beta uzima vrednosti u opsegu $[-1, 1]$
 - Pozitivne vrednosti beta parametra ističu čvorove čiji susedi takođe imaju veliki broj suseda
 - Negativne vrednosti parametre ističu čvorove čiji susedi sami nemaju mnogo suseda
- *Influence vs. power*

Druge varijante centralnosti

- *Katz centrality*
 - Varijanta centralnosti po svojstvenom vektoru koja uključuje broj čvorova do kojih se može doći nekim putem
 - Uvodi se faktor slabljenja za udaljene čvorove
- *PageRank*
 - Poznati algoritam za određivanje važnosti stranica na webu
 - Baziran na sličnoj ideji kao centralnosti po svojstvenom vektoru
- *Percolation centrality*
 - Varijanta relacione centralnosti
 - Svakom čvoru se pridružuje i stanje koje uvažava dinamiku mreže i daje mu veću težinu
 - Npr. da li je čvor inficiran
- *Edge betweenness*
 - Varijanta relacione centralnosti koja se računa za grane mreže

Određivanje centralnih čvorova u mreži

- Različite mere ističu različite čvorove



Korelacija među metrikama (1)

- Generalno, centralnost po stepenu, bliskosti i relacionica centralnost su pozitivno korelisane
 - Pokazuju statističke analize
- U slučajevima kada je korelacija niska, to verovatno govori o nekom interesantnom svojstvu posmatrane mreže

Korelacija među metrikama (2)

| | Niska DC | Niska CC | Niska BC |
|-----------|--|--|--|
| Visoka DC | / | Čvor se nalazi u klasteru koji je daleko u odnosu na ostatak mreže | Konekcije čvora su redundantne, komunikacija ga zaobilazi |
| Visoka CC | Bitan čvor, povezan sa važnim drugim akterima | / | Veliki broj putanja u mreži, čvor je blizak drugima, ali su i drugi međusobno bliski |
| Visoka BC | Veze čvora izuzetno bitne za tok informacija u mreži | Čvor monopolizuje veze manjeg broja ljudi ka ostalim čvorovima mreže | / |

Centralizacija na nivou mreže (1)

- Centralizacija mreže se računa na nivou cele mreže
 - Može se izračunati za bilo koju meru centralnosti
- Pokazuje varijansu izračunate mere centralnosti posmatrane mreže procentualno u odnosu na mrežu iste veličine koja ima topologiju zvezde
- Mreža sa topologijom zvezde poseduje jedan centralni čvor sa kojim su povezani svi ostali čvorovi i koji nemaju drugih grana
 - U takvoj mreži, raspodela moći je najviše nejednaka

Centralizacija na nivou mreže (2)

- Centralizacije mreže po stepenu se računa na sledeći način

$$C_D(G) = \frac{\sum_{i=1}^{|V|} [C_D(v_*) - C_D(v_i)]}{\sum_{j=1}^{|Y|} [C_D(y_*) - C_D(y_j)]}$$

- $G(V,E)$ predstavlja posmatranu mrežu
 - $X(Y,Z)$ predstavlja mrežu sa istim brojem čvorova koja ima topologiju zvezde (*star network*)
 - Čvor označen sa * je onaj sa najvećim stepenom u mreži
- Veći procenat centralizacije mreže ukazuje na postojanje čvorova koji su boljoj poziciji u odnosu na ostale

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.
- <http://www.network-science.org/>