Analiza socijalnih mreža

Mrežni modeli

Marko Mišić, Jelica Protić

13M111ASM

2017/2018.

Motivacija (1)

- Glavni cilj analize mreža je izgradnja modela koji reprodukuju osobine realnih mreža
  - Veoma često je potrebno realne socijalne mreže uklopiti u odgovarajući teorijski model
- Mnoge mreže na prvi pogled izgledaju kao da su nastale na potpuno slučajan način
  - Kakve su karakteristike potpuno slučajnih mreža?
    - Prosečne dužine putanja, distribucija čvorova po stepenu, klasteri...

Motivacija (2)

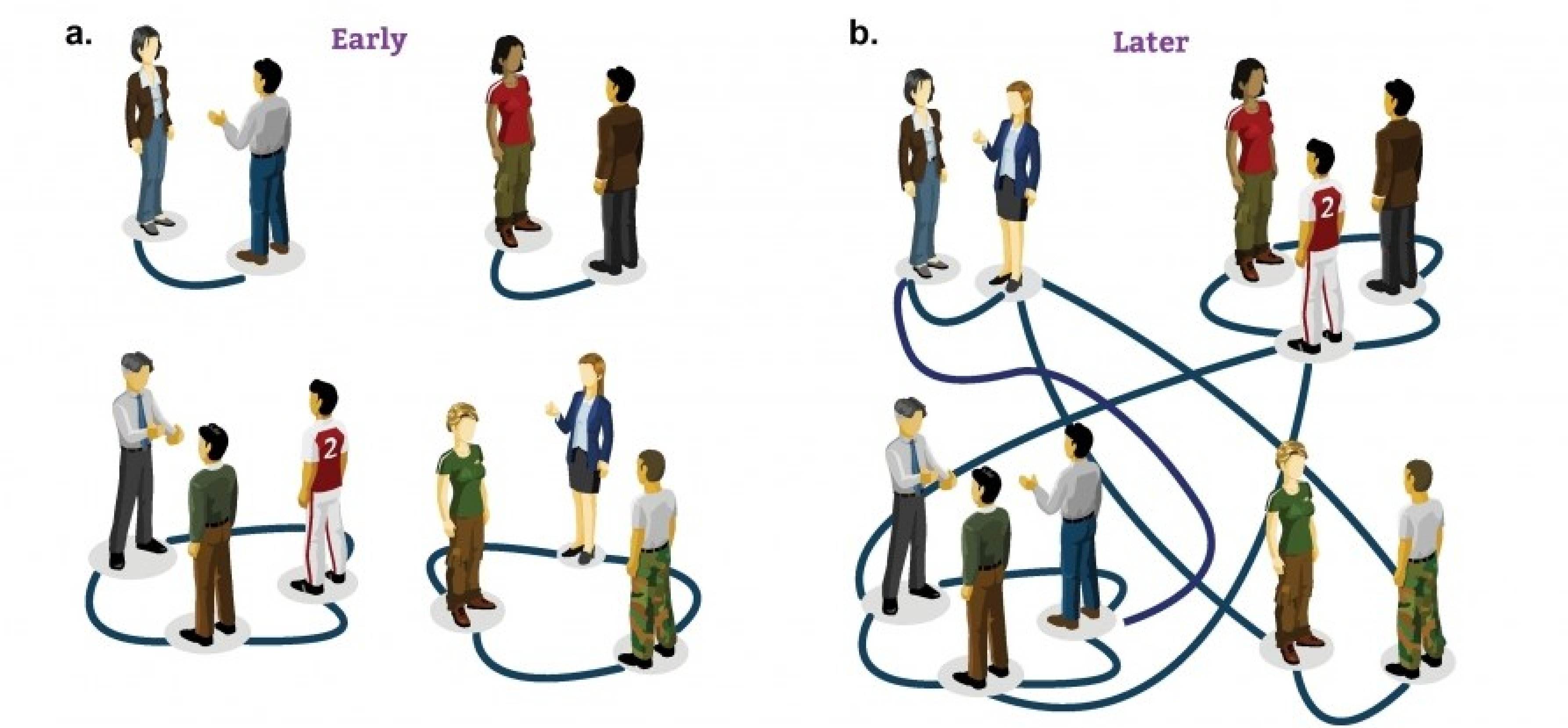
- Postavljaju se pitanja:
  - Da li su realne mreže koje opažamo zaista nastale na potpuno slučajan način?
  - Na koji način se može objasniti klasterizacija u realnim mrežama?
  - Kako se objašnjava postojanje habova?
  - Kakve su karakteristike realnih mreža u odnosu na potpuno slučajne mreže?
- Modelovanjem mreže se omogućava:
  - Bolje razumevanje procesa koji dovodi do formiranja mrežne strukture koju opažamo
  - Predikcija svojstava i ishoda procesa u novim ili nepoznatim mrežama

Motivacija (3)

- Mrežni modeli daju odgovor na pitanje kako se mreža formira i menja kroz vreme
- Daju uvid u proces kreiranja mreže
- Modeliraju ključne strukturne karakteristike mreže, kao što su:
  - Distribucija čvorova po stepenu
  - Prosečne putanje u mreži
  - Dijametar mreže
  - Klastere u mreži
  - Koeficijent klasterizacije
- Ekonomski modeli i modeli teorije igara
  - Daju odgovor zašto se mreža formira na određeni način

Motivacija (4)

- Primer: formiranje mreže poznanstava na zabavi na kojoj se inicijalno niko ne poznaje



Radnom network modeli (1)

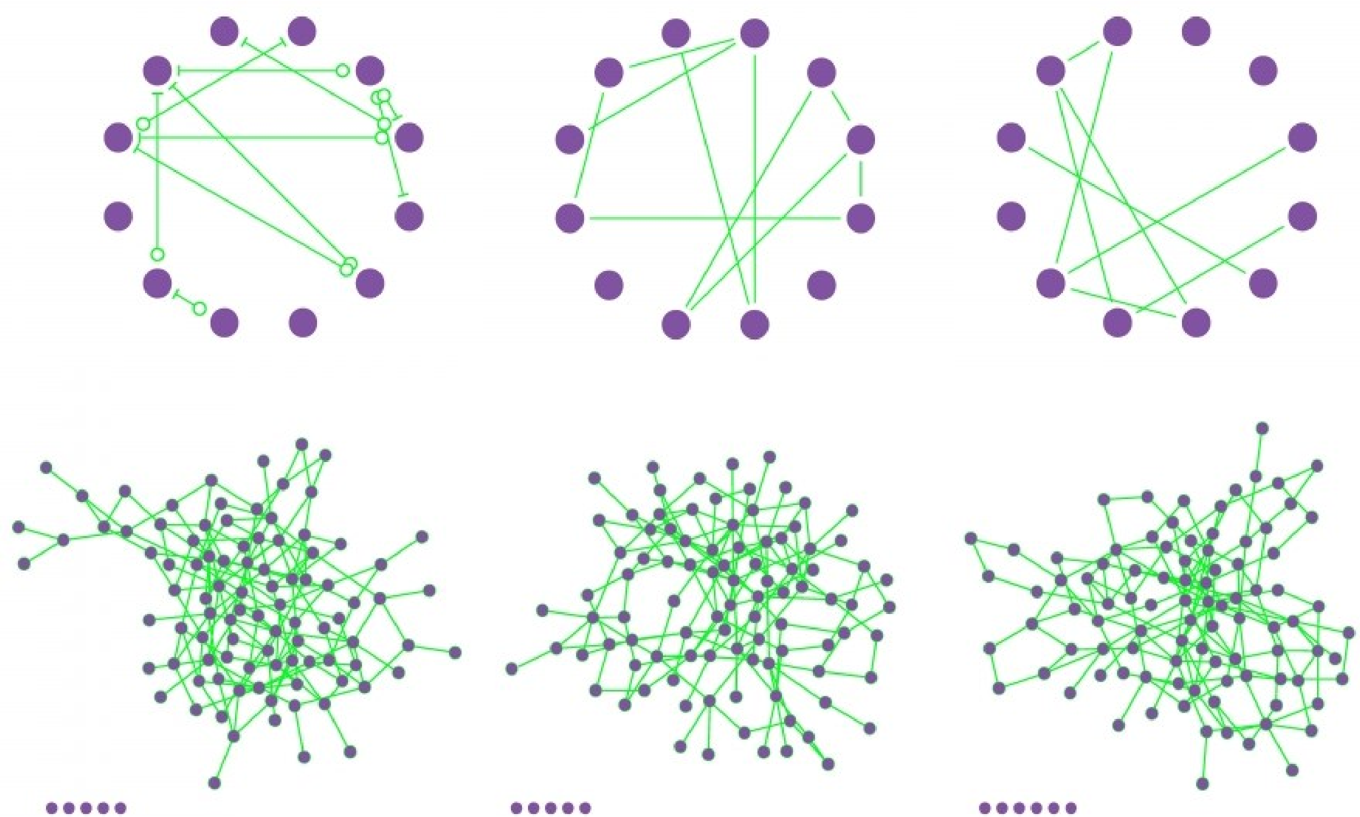
- Kombinuju kombinatoriku i teoriju verovatnoće sa teorijom grafova
- Neka se mreža sastoji od *N* čvorova, a *p* predstavlja verovatnoću povezivanja dva čvora vezom
  - Mreža je neusmerena
  - U početnom trenutku mreža sadrži sve čvorove, a nijednu vezu
  - Veze se formiraju sa uniformnom verovatnoćom
  - Veze se formiraju potpuno nezavisno jedne od drugih
- Pioniri na polju su matematičari:
  - Paul Erdos
    - *Erdos* broj (*Erdos number*)
  - Alfred Renyi
    - *"A mathematician is a device for turning coffee into theorems"*
  - Edgar Nelson Gilbert

Radnom network modeli (2)

- Da bi se izgradila mreža:
  - Započne se sa *N* izolovanih čvorova
  - Izabere se jedan par čvorova
  - Generiše se slučajan broj u opsegu od 0 do 1
  - Ako je generisani broj veći od *p*, izabrani čvorovi se povezuju granom
  - Prethodni postupak se ponavlja za svih *N(N-1)/2* parova čvorova
- *G(N,p)* model (*Erdos-Renyi* mreža, *Gilbert*-ov model)
  - Prazan model, *null* model mreže
- Alternativa – *G(N,L)*
  - *L* predstavlja maksimalan broj na slučajan način raspoređenih veza

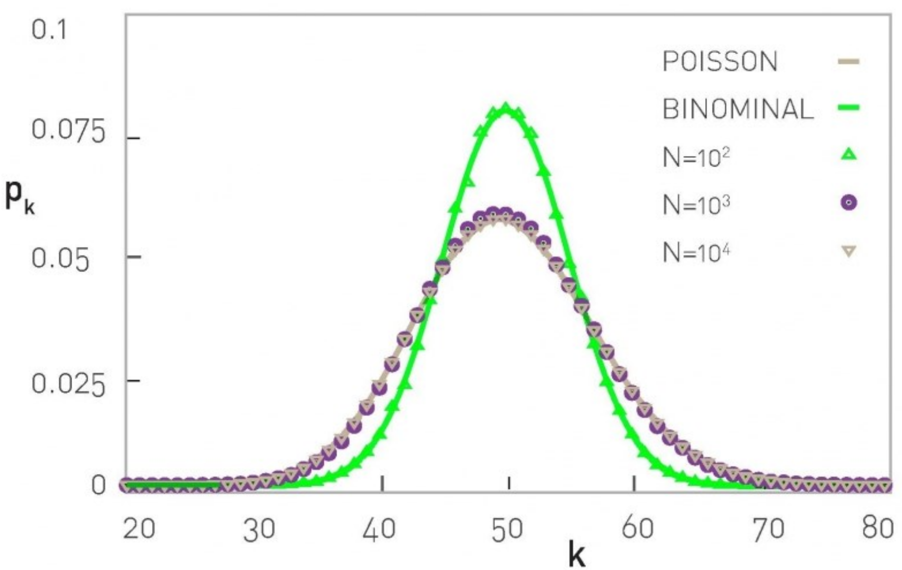
Radnom network modeli (3)

- Primeri ER mreža
  - Gore: *p=1/6* i *N=12*
  - Dole: *p=0.03* i *N=100*



Karakteristike ER modela

- Broj veza između čvorova sledi binomnu raspodelu
  - Često se aproksimira Poasonovom raspodelom kod retkih mreža
  - Postoji nagomilavanje oko prosečnog stepena čvora
  - Nezavisno od veličine mreže
- Prosečan stepen čvora:
  - *p*(N-1)*



Evolucija mreže u ER modelu (1)

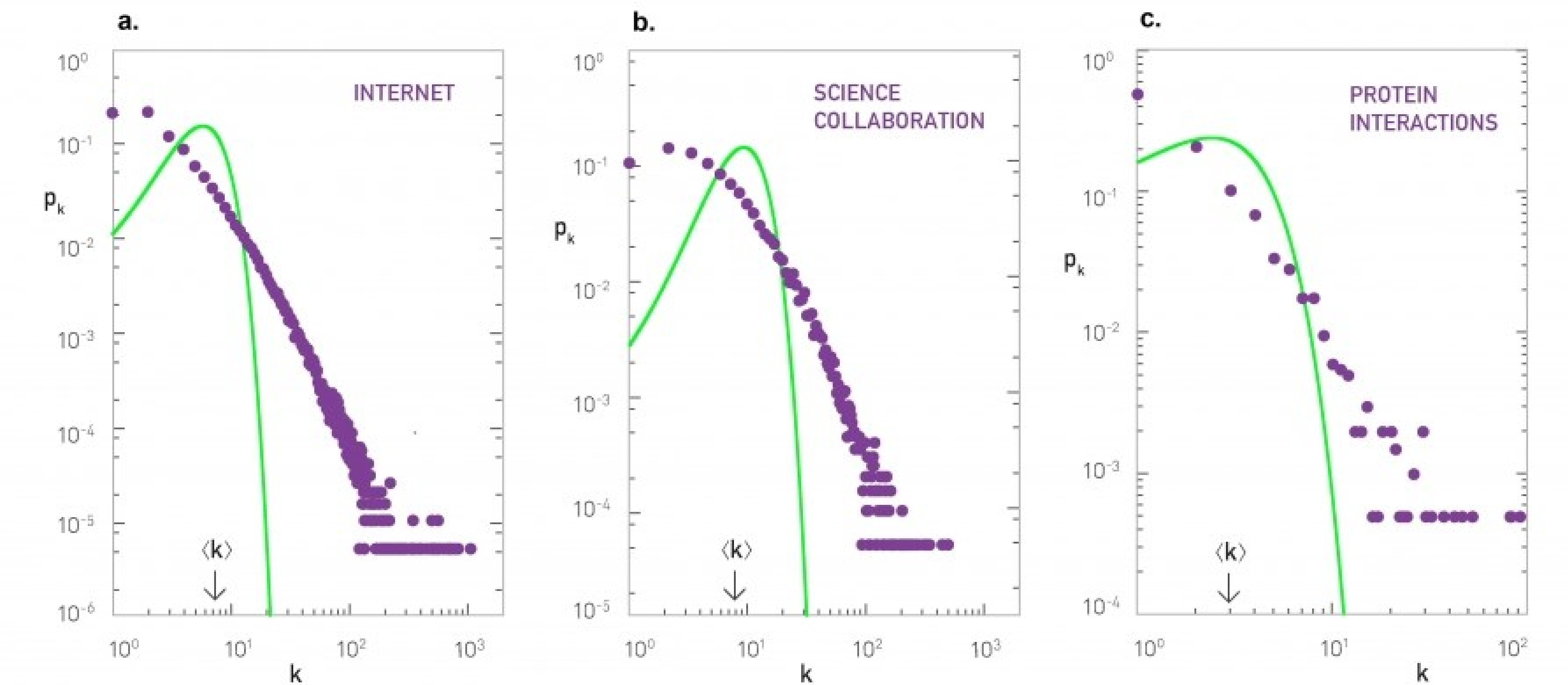
- Za određene vrednosti *p* dolazi do značajnih promena u strukturi ER mreže
  - Pragovi u mreži i faze prelaza (subkritični, super kritični i povezani režim)
- Značajne faze prelaza:
  - Prag za stvaranje veza unutar mreže
    - $p=1/n^2$, prosečan stepen čvora je tada ~ $1/n$
  - Prag za pojavu ciklične putanje i gigantske komponente (kritična tačka)
    - $p=1/n$, prosečan stepen čvora je tada ~ 1
  - Prag za stvaranje povezane mreže
    - $p=ln(n)/n$, prosečan stepe čvora je tada ~ $ln(n)$

Poređenje sa realnim mrežama (1)

- Realne mreže najčešće ne slede Poasonovu raspodelu
  - Postoji značajno veća varijacija u stepenu čvorova
  - Postoje čvorovi koju su izolovani (sa malo veza)
  - Postoje habovi (popularni čvorovi)
- ER model se razmatra u kontekstu velikih mreža
  - Omogućava analize bazirane na verovatnoći i statistici
  - Kada broj čvorova teži beskonačnosti
  - Najbolje opisuje realne mreže kada je $p > 1/N$

Poređenje sa realnim mrežama (2)

- ER model potcenjuje u realnim mrežama:
  - Broj i veličinu čvorova sa visokim stepenom čvora
  - Broj čvorova sa niskim stepenom čvora



Poređenje sa realnim mrežama (3)

- ER model se često koristi
za poređenje sa realnim mrežama
- Na taj način se mogu uočiti svojstva i procesi
u okviru realne mreže
  - Ako postoji dovoljno veliko odstupanje od realne mreže, svojstva i procesi se mogu smatrati sistematskim, a ne nasumičnim
- Postoji značajna zajednička karakteristika
realnih i ER mreža
  - Mali dijametar i mala prosečna dužina putanje
  - Važi čak i kod jako velikih mreža

Alternativni *random* mrežni modeli (1)

- Model upoznavanja (*introduction* model)
  - Nasumično uspostavljanje veza
  - Uspostavljanje veza posedstvom prijatelja
  - Ova dva načina formiranja veza se modeluju verovatnoćama formiranja veze i upoznavanja
- Statični geo-model (*static geographical* model)
  - Raspoređuje čvorove nasumično u pravougaonu mrežu (*lattice*)
  - Pozicija čvora u prostoru određuje način formiranja veza
    - Svaki čvor se povezuje sa zadatim brojem najbližih čvorova
    - Neki čvorovi će imati značajno više čvorova zbog pozicije u prostoru

Alternativni *random* mrežni modeli (2)

- Model slučajnog susreta (*random encounter* model)
  - Čvorovi zauzimaju slučajne pozicije u prostoru
  - Simulira se kretanje čvorova kao čestica u određenom vremenu
  - Pri susretu se uspostavi veza između dva čvora do maksimalnog broja konekcija
- Model rasta (*growth* model)
  - Počinje od malog broja čvorova koji formiraju kliku
  - Dodaju se novi čvorovi jedan po jedan
  - Svaki dodati čvor uspostavlja zadati broj veza na slučajan način

ER model i alternativni modeli

- U odnosu na ER model, alternativni modeli bolje opisuju klasterovanje u realnim mrežama
  - Sadrže više zatvorenih trijada
  - Manji stepen povezanosti komponenti
- Prosečne dužine puta u odnosu na ER model su nešto veće
  - I dalje relativno male u odnosu na veličinu mreže

Fenomen malog sveta (1)

- Realne mreže karakteriše:
  - Mala prosečna udaljenost bilo koja dva čvora u odnosu na veličinu mreže
  - Visok stepen klasterizacije
- Eksperiment koji je sproveo *Milgram*, 1967.
  - Sproveden u USA, u nekoliko saveznih država
    - Početna tačka su bile države srednjeg zapada
    - Krajnja tačka države na istočnoj obali
  - Osobe su zamoljene da proslede pismo drugoj osobi, ukoliko je poznaju
  - Ukoliko je ne poznaju, trebalo je da proslede pismo nekome ko bi mogao da zna više o odredišnoj osobi
- Rezultat – pisma su stizala u proseku u 5.5 do 6 koraka
  - *Six degrees of separation*

Fenomen malog sveta (2)

- Analiza realnih socijalnih mreža (platformi) je pokazala još manje prosečne distance:
  - MS *Messenger* servis 2008. - 6.6 koraka
  - *Facebook* 2008. - 5.28 koraka na 56 miliona korisnika
  - *Facebook* 2011. - 4.74 koraka na 721 miliona korisnika
  - *Facebook* 2016. - 4.57 koraka na 1.59 milijardi korisnika
  - Analize su uključile najkraće moguće putanje u mreži
    - Što nije bio slučaj sa *Milgram*-ovim eksperimentom
- Stepen razdvajanja (*degree of separation*) je za jedan manji od prosečne dužine putanje
  - U 2016. prosečno *three and a half degrees of separation*

Teorema mrežne strukture (1)

- Fenomen malog sveta se može tumačiti kao posledica teoreme mrežne strukture
  - Objašnjava malu vrednost dijametra i prosečne dužine putanje u *random* mrežama
- Teorema mrežne strukture važi za dovoljno velike mreže za koje važi:
  - Velika verovatnoća postojanja puta između dva čvora
    - Mreža je uglavnom povezana
  - Svaki čvor nije direktno povezan sa svakim drugim čvorom mreže

Teorema mrežne strukture (2)

- U dovoljno velikoj *G(N,p)* mreži, za prosečnu dužinu puta l_{av} važi sledeći odnos:

$$l_{av} \sim ln(n) / ln(d)$$

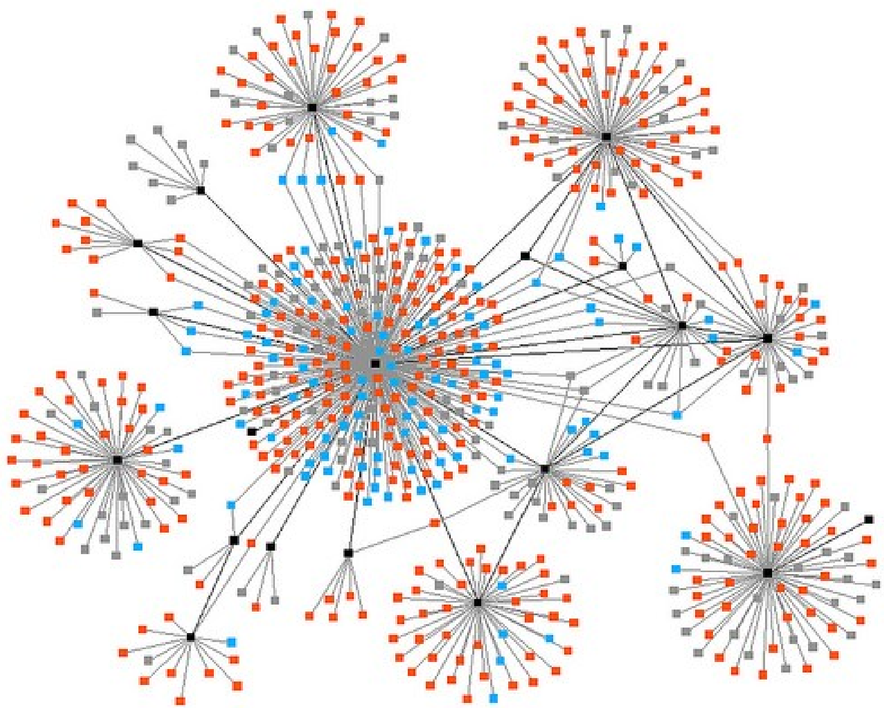
  gde je *d* prosečan stepen čvorova u mreži
- Prosečna dužina puta u mreži raste logaritamski sa porastom veličine mreže
  - Sa značajnim porastom veličine mreže, bilo koja dva čvora mreže će i dalje biti povezana u relativno malo koraka
- *Small-world phenomenon (property)*
  - Društvene mreže, neuroni u mozgu, energetske mreže...
  - Često se opisuje *Watts-Strogatz* modelom

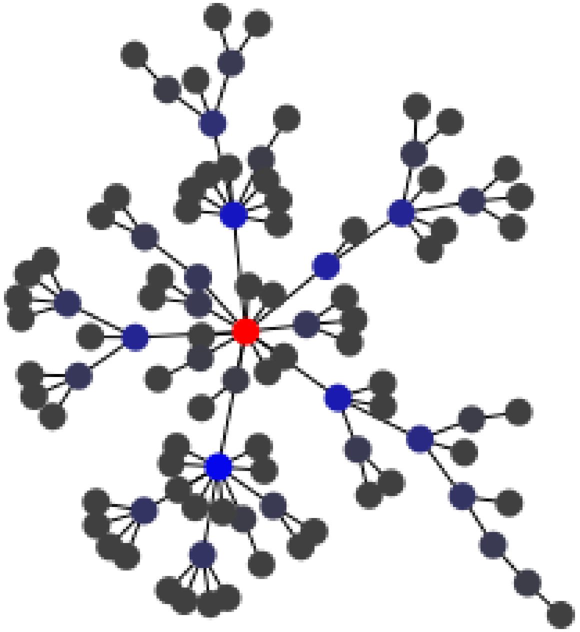
Preferencijalno vezivanje (1)

- Prethodno opisani modeli imaju problem opisivanja realnih mreža koje sadrže habove
  - Habovi su tipični za realne socijalne mreže
- Modeli preferencijalnog vezivanja rešavaju ovaj problem
  - Novi čvorovi pristupaju mreži pretežno vezivanjem sa čvorovima koji već imaju visok stepen čvora
    - Popularne i uticajne ličnosti
- Stepen pojedinih čvorova raste značajno brže nego stepen ostalih čvorova u mreži
  - Rezultat je disproporcija u mreži
  - Mreža sa malim brojem jako dobro povezanih čvorova – habova i velikim brojem čvorova sa malim stepenom

Preferencijalno vezivanje (2)

- Primer mreža sa preferencijalnim vezivanjem
  - *The good get better / the rich get richer*





Preferencijalno vezivanje (3)

- Glavna karakteristika mreža nastalih preferencijalnim vezivanjem je da slede *power law* raspodelu
  - Predstavljaju tzv. *scale-free* mreže
- Novi čvorovi kreiraju veze prema postojećim čvorovima sa verovatnoćom koja je proporcijalna broju veza koje čvor već ima
  - Tako se formira raspodela veza po čvorovima koja je izuzetno naklonjena habovima u mreži
  - Habovi imaju mnogo viši stepen nego što bi bio slučaj da se veze formiraju na slučajan način
- Najpoznatiji *Barabasi-Albert* model

Barabasi-Albert model

- Mreža inicijalno sadrži m međusobno povezanih čvorova
  - U svakoj jedinici vremena pojavi se novi čvor koji uspostavi vezu sa m postojećih čvorova
- U nekom vremenskom trenutku t:
  - Ukupan broj čvorova u mreži je t
  - Broj konekcija je $t \cdot m$
  - Ukupan degree je $2 \cdot t \cdot m$
- Verovatnoća da novi čvor uspostavi vezu sa čvorom i u trenutku t je:
  - $d_i(t)/2 \cdot t \cdot m$, gde je $d_i(t)$ degree čvora i u trenutku t

Simulacija mrežnih modela

- NetLogo simulator mrežnih struktura
  - 2D verzija
  - Odgovarajući modeli u okviru biblioteke simulatora
  - Modeli sa sajta http://www.ladamic.com
- Modeli:
  - ErdosRenyiDegDist.nlogo
  - RandomGraphs.nlogo
  - RAndPrefAttachment.nlogo

Literatura

- Albert-László Barabási, Network science, Cambridge University Press, 2016.
- J. Jovanović, Softverska analiza društvenih mreža, FON, 2017.
- NetLogo simulator, https://ccl.northwestern.edu/netlogo/
- Lada Adamic, NetLearn, http://www.ladamic.com
- http://www.network-science.org/