Kompleksne mreže

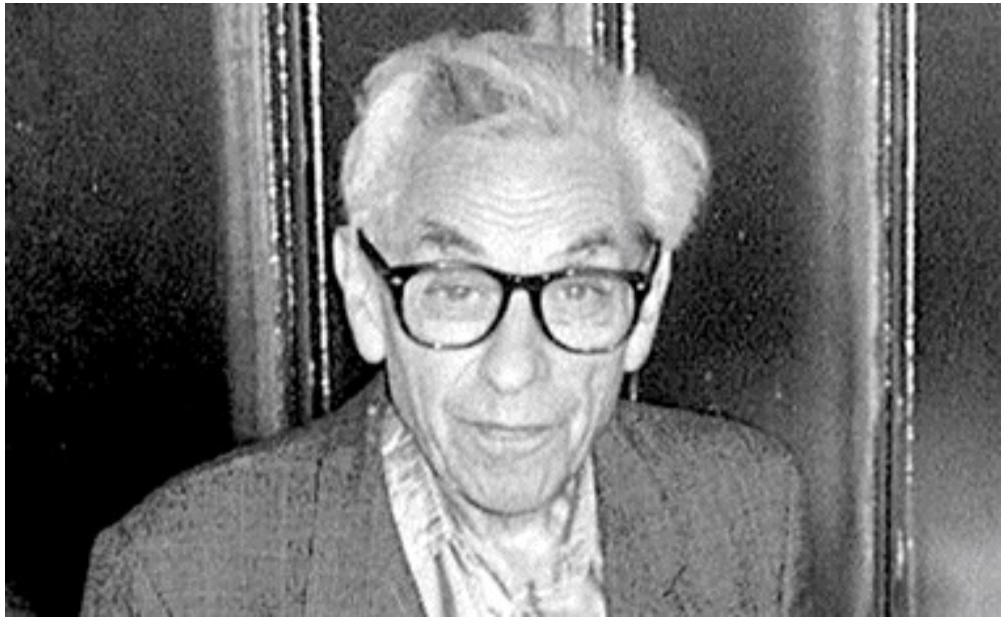
5. Predavanje

Realne mreže

- Kratak najkraći put
- Puno trokuta, što rezultira visokim koeficijentom klasteriranja
- Heterogene distribucije varijabli čvora i veza poput stupnja i težina
- Od kuda nam ta svojstva - MODEL za izgradnju mreža

Slučajne mreže

- Krenemo sa skupinom nepovezanih čvorova
- Dodamo veze između slučajno odabranih parova čvorova
- Slučajna ili Erdos-Renyi mreža



Slučajne mreže

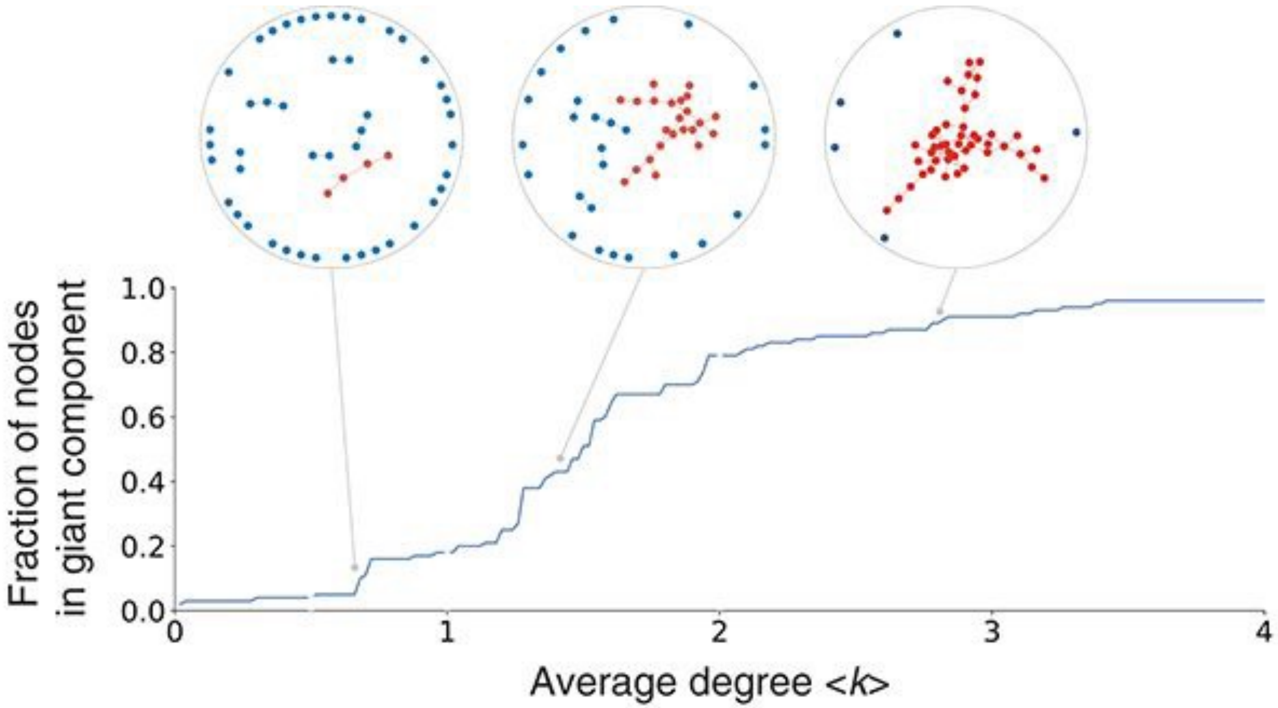
- Gilbertov model nasuprot Erdos-Reny model
  - Erdos-Reny – broj čvorova i veza fiskan
  - Gilbert model – broj čvorova fiksan, broj veza varira
- Gilbertov model:
  - Zadan je broj čvorova N i vjerojatnost veze *p*

1. Odaberemo par čvorova *i* i *j*
2. Generiramo slučajan broj *r* između 0 I 1. Ako je $r<p$, dodamo vezu između *i* i *j*
3. Ponavljamo (1) i (2) za sve parove čvorova

Rast mreže

- Zamislimo velik broj čvorova, bez veza
- Sustav je razlomljen u singletone (izolirane čvorove)
- Dodajemo veze, jednu u svakom vremenskom trenutku
- Sve više čvorova parova čvorova se povezuje
- Nastanak povezanih podmreža
- Mreža postaje povezana
- Tranzicija od puno malih povezanih komponenti u jednu veliku (većina grafa)
- Suprotno očekivanju, tranzicija je nagla – za $\langle k \rangle = 1$

Evolucija Erdos-Renyi grafa za različiti srednji stupanj



Gustoća

- Očekivani broj veza u slučajnoj mreži – proporcionalan vjerojatnosti veze i broju parova čvorova
- Broj mogućih parova čvorova $\binom{N}{2} = \frac{N(N-1)}{2}$
- Očekivani broj veza u slučajnom grafu $\langle L \rangle = p\binom{N}{2} = \frac{pN(N-1)}{2}$
- Očekivani prosječni stupanj $\langle k \rangle = \frac{2\langle L \rangle}{N} = p(N-1)$
- Gustoća d=$\frac{\langle k \rangle}{N-1}$ -> očekivana gustoća $\langle d \rangle = p$
- Realne mreže – rijetke (mali $\langle k \rangle$ u odnosu na ukupan broj čvorova i mala gustoća) -> *p* treba biti mali

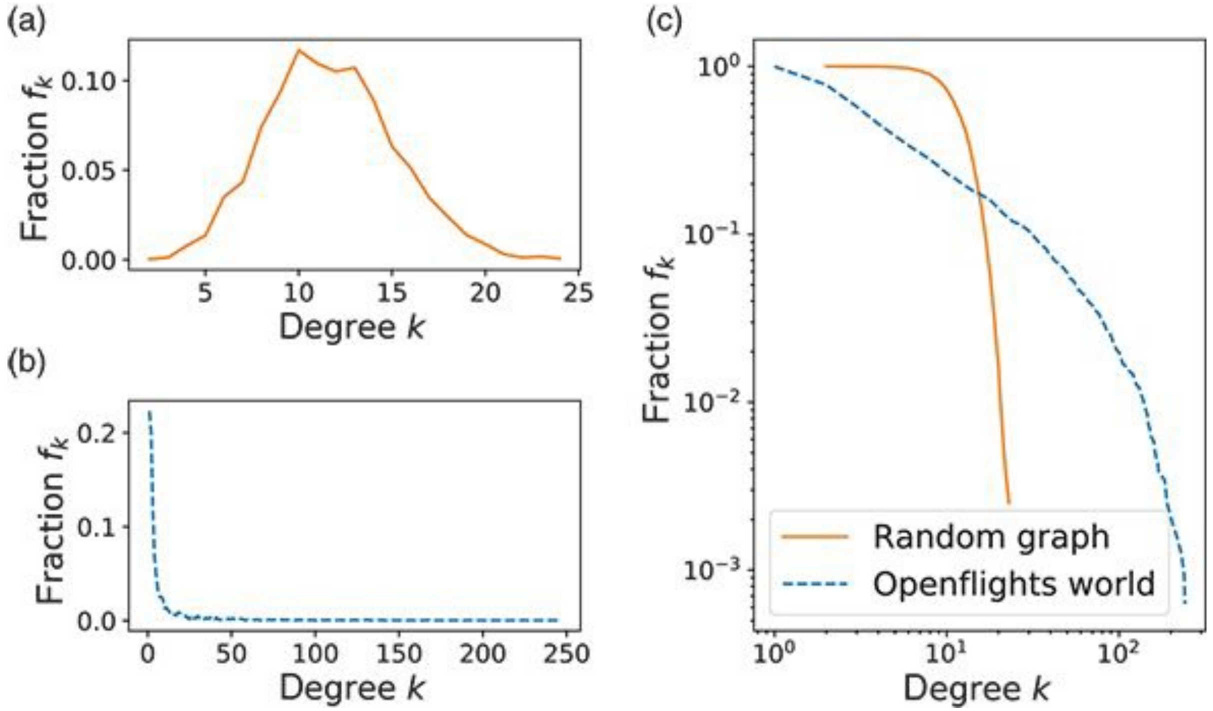
Distribucija stupnja

- Distribucija stupnja slučajne mreže - vjerojatnost da čvor ima *k* susjeda
- Niti jedan čvor nema posebnu ulogu u ovom modelu
- Kolika je vjerojatnost da čvor *i* ima nula, jednog, dva ili više susjeda
- Svaki od preostalih *N-1* čvorova može biti susjed od *i*
- Svaki par koji uključuje *i* ima vjerojatnost *p* da bude spojen, neovisno o ostatku mreže

Distribucija stupnja

- Ekvivalentan problem – broj glava u *N-1* bacanja (vjerojatnost glave *p*)
- Binomna distribucija $P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}$
- Za veliki *N*, $pN \approx \langle k \rangle$ pri čemu binomna distribucija se dobro aproksimira Poissonovom sa srednjom vrijednošću i varijancom $\langle k \rangle$

Distribucija stupnja



- Usporedba Erdos-Renyi slučajnoj grafa s mrežom avionskih letova istoga broja čvorova i veza

Kratki putevi

- Pretpostavimo da je mreža povezana i svi čvorovi imaju stupanj *k*
- Unutar jednog koraka $l = 1$ dosežemo *k* čvorova
- Svaki od njih ima *k-1* susjeda -> nakon dva koraka smo dosegli *k(k-1)* čvorova
- Unutar tri koraka dosižemo $k(k-1)^2$ čvorova
- Možemo zaključiti da na udaljenosti l od korijena možemo naći $k(k-1)^{l-1}$ čvorova
- Za veliki *k* u *l* koraka dosižemo k^l čvorova

Kratki putevi

- $k^{l_{max}} = N$
- $l_{max} = \log_k N = \frac{\log N}{\log k}$
- Aproksimacija diametra mreže ako uzmemo u obzir preklapanja i fluktuacije u stupnju oko $\langle k \rangle$
- udaljenosti su male čak i kada je mreža jako velika

Koeficijent klasteriranja

- Udio trokuta kojima je središte u promatranom čvoru
- Slučajna mreža – vjerojatnost da je par susjeda čvora povezan je *p*
- Koeficijent klasteriranja pojedinog čvora može razlikovati od *p,* no srednja vrijednost kroz sve čvorove se može aproksimirati s *p*
- Zahtjev za realne mreže je mali *p* -> mali očekivani koeficijent klasteriranja

Usporedba realne i slučajne mreže

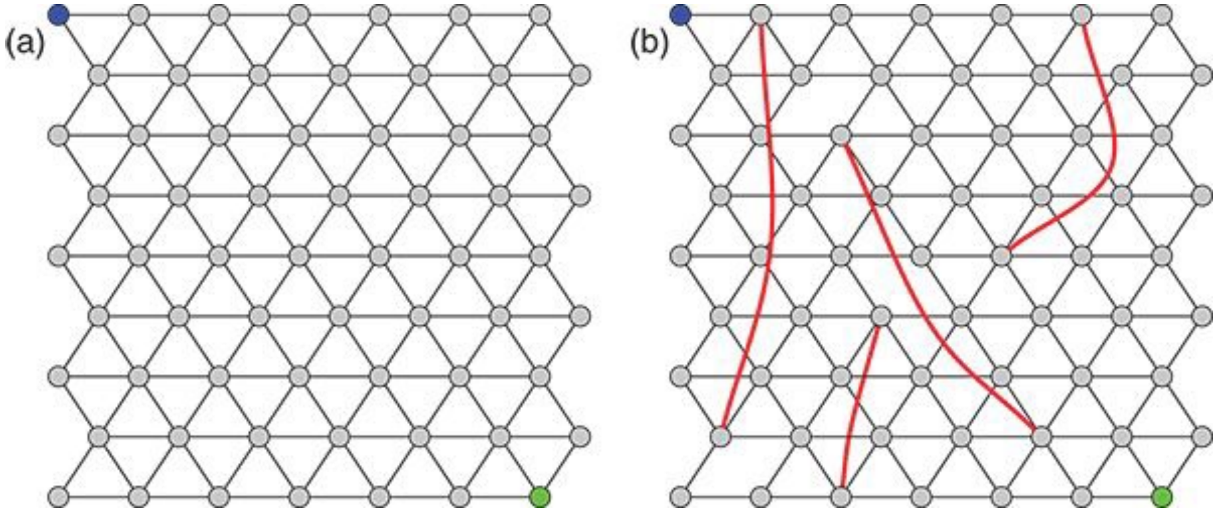
Svojstvo	Realna	Slučajna
Rijetka mreža	DA	DA, ako je *p* mali
Distribucija stupnja	Opadajuća s otežanim repom	Zvonolika
Najkraći srednji put	Kratak	Kratak
Koeficijent grupiranja	Velik	Malen za mali *p*

Mali svijet

- Slučajne mreže različite od realnih
- Watts-Strogatz model (1998)
- Kratak najkraći put i visok koeficijent klasteriranja

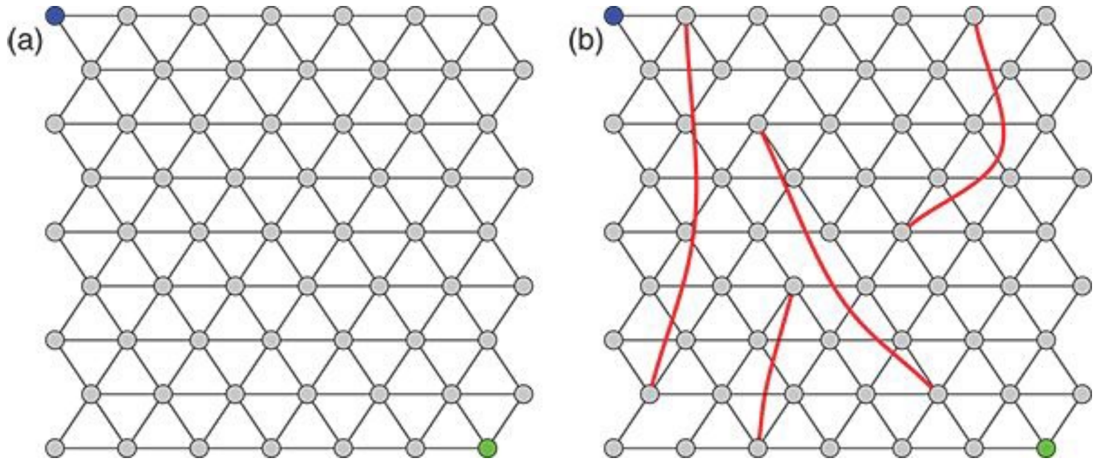
Model malog svijeta

- Krenemo od rešetke (svaki čvor ima isti broj susjeda)
- Visok stupanj klasteriranja – svaki par uzastopnih susjeda povezan
- $C = 6/\binom{6}{2} = \frac{2}{5}$
- za rubne čvorove klastering i veći
- Dugačak najkraći put



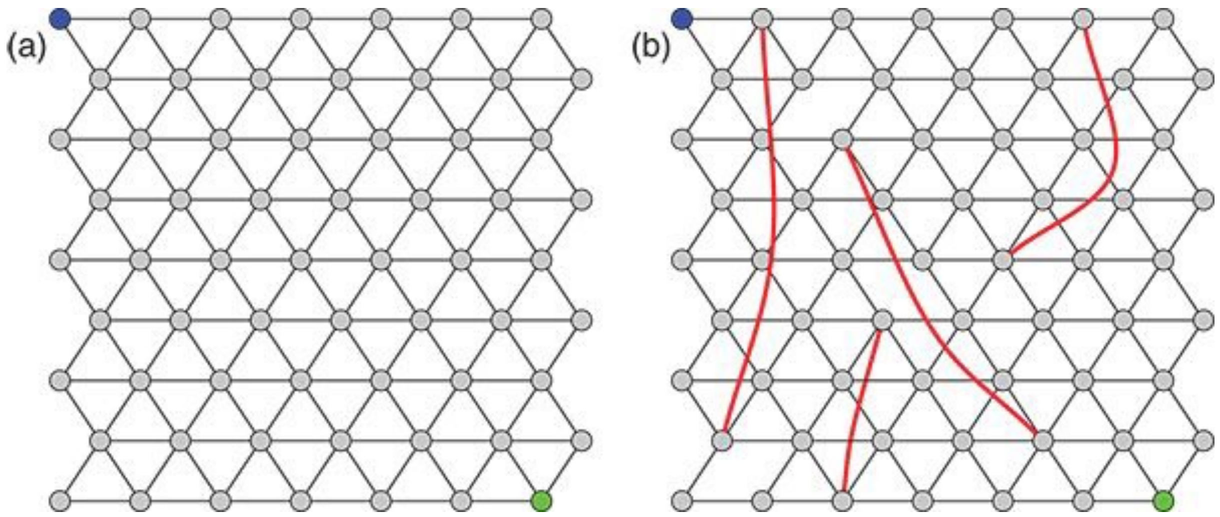
Model malog svijeta

- Smanjenje udaljenosti između čvorova kreiranje prečica među čvorovima
- Odabir početnih veza na slučajan način, sačuvamo jedan čvor, a drugi zamjenom sa slučajno odabranim
- Vjerojatnost premošćivanja pojedine veze *p*
- Broj premoštenih veza proporcionalan vjerojatnosti premošćivanja

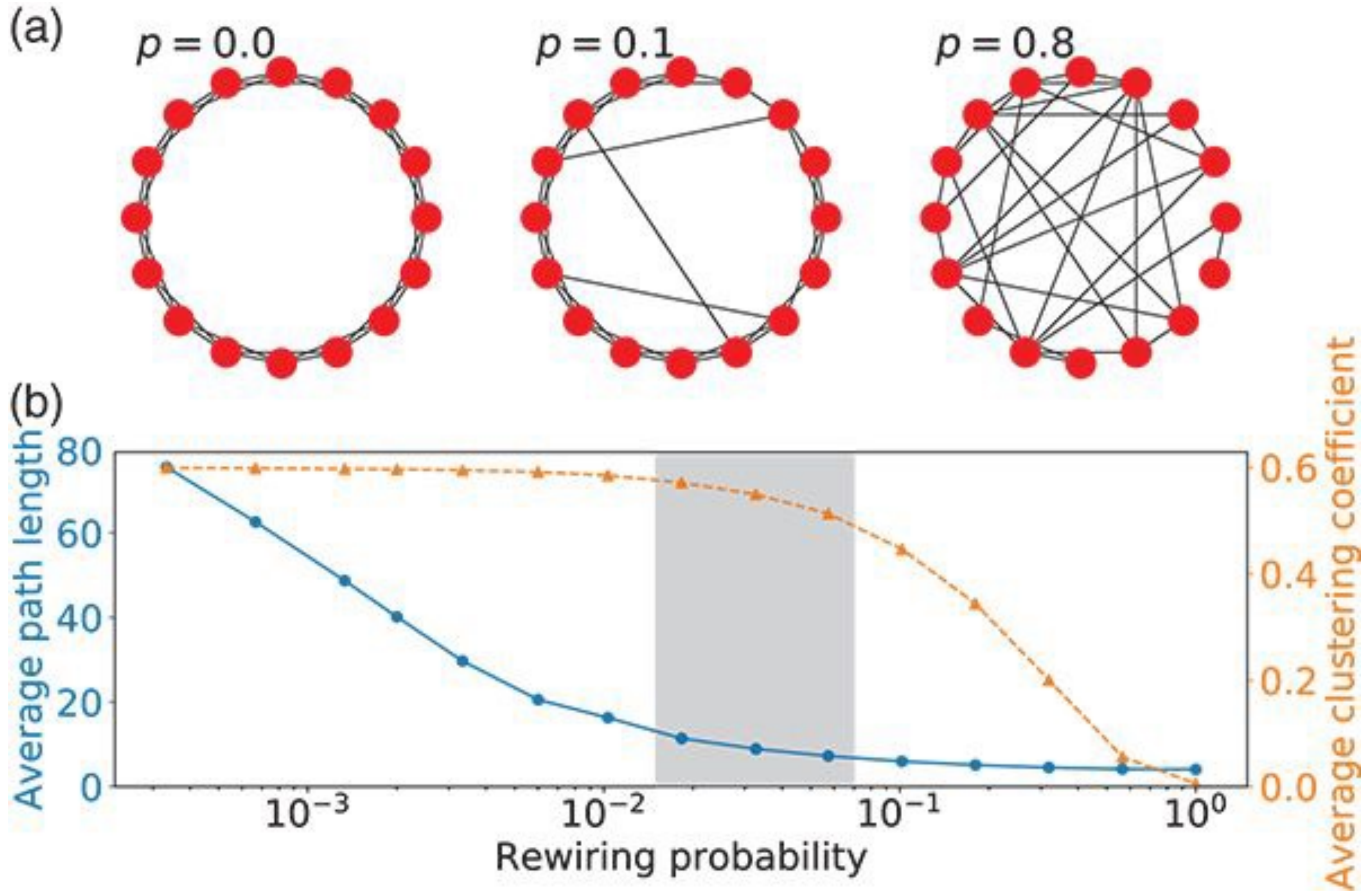


Model malog svijeta

- Premošćivanje rijetko – malo se toga dogodi
- Premošćivanje često – slučajna mreža
- Ako je *p* niti velik ni mali -> moguće postići mali najkraći put sa sačuvanim visokim koeficijentom klasteriranja



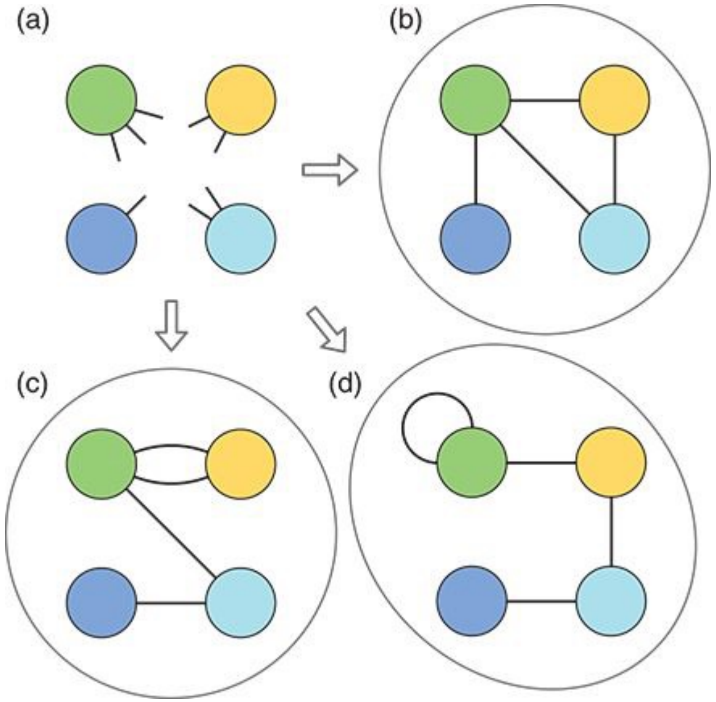
Originalan rad Wattsa i Strogatza



- Kratki putevi i veliki koeficijent klasteriranja
- Model ne može proizvesti hubove
- Za bilo koju vjerojatnost premošćivanja broj čvorova i veza ostaje isti
- Distribucija stupnja zvonolika

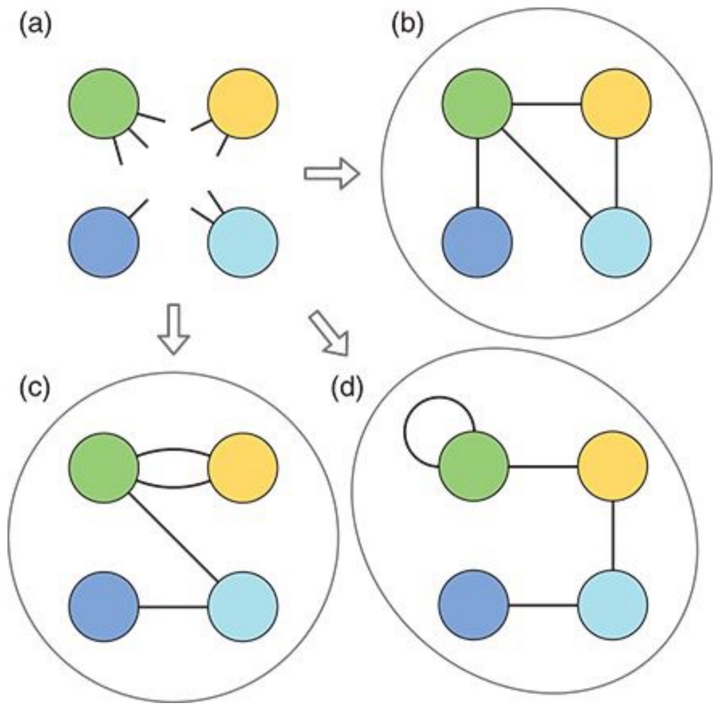
Konfiguracijski model

- Cilj – proizvesti mrežu čiji čvorovi imaju proizvoljan slijed stupnjeva (npr. prvi čvor ima stupanj k_1, drugi čvor ima stupanj k_2 i tako redom)
- Slijed stupnjeva
  - Željena distribucija
  - Iz realne mreže



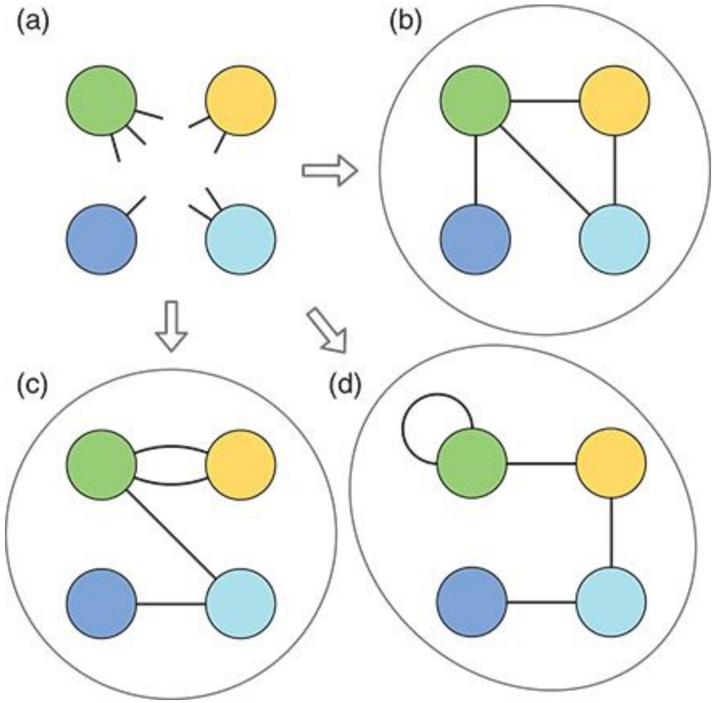
Konfiguracijski model

- Dodjela svakom čvoru izdanka s kojim može kreirati vezu
- Suma izdanaka mora biti paran broj
- Različite mreže kreirane – ovisno o broju kombinacija parova izdanaka
- Pojedini rezultati su neželjeni (npr. Višestruke veze između dva čvora)



Konfiguracijski model

- Za danu mrežu, možemo istražiti je li specifično svojstvo definirano distribucijom stupnja
- Kreiramo po volji nove mreže s istom distribucijom
- Provjerimo jesu li nove mreže zadržale promatrano svojstvo
- Ako ne, postoje drugi faktori u pozadini
- Primjer – koeficijent klasteriranja
- Eksponencijalni slučajni grafovi



Eksponencijalni slučajni modeli

- Proučavanje slučajno proizvedenih mreža koje dijele zajednička kvantitativna svojstva, s razlikom u detaljnoj strukturi
- Potencijalna alternativa specifičnim mrežnim konfiguracijama koje srećom u realnom svijetu
- Omogućeno istraživanje međudjelovanja različitih strukturnih svojstava
- Primjer – koje vrijednosti koeficijenta klasteriranja su kompatibilne specifičnoj vrijednosti gustoće

Eksponencijalni slučajni grafovi

- Klasa slučajnih mreža s ograničenjima
- Definiramo klasu mreža na osnovu skupa *M* mrežnih mjera x_m, *m=1,…,M*
- Unosimo ograničenje za svaku mjeru x_m: srednja vrijednost kroz sve mreže unutar klase mora biti specificirana vrijednost $\langle x_m \rangle = x_m^*$
- Eksponencijalni slučajni grafovi – mreže koje zadovoljavaju ograničenje dok maksimiziraju nasumičnost

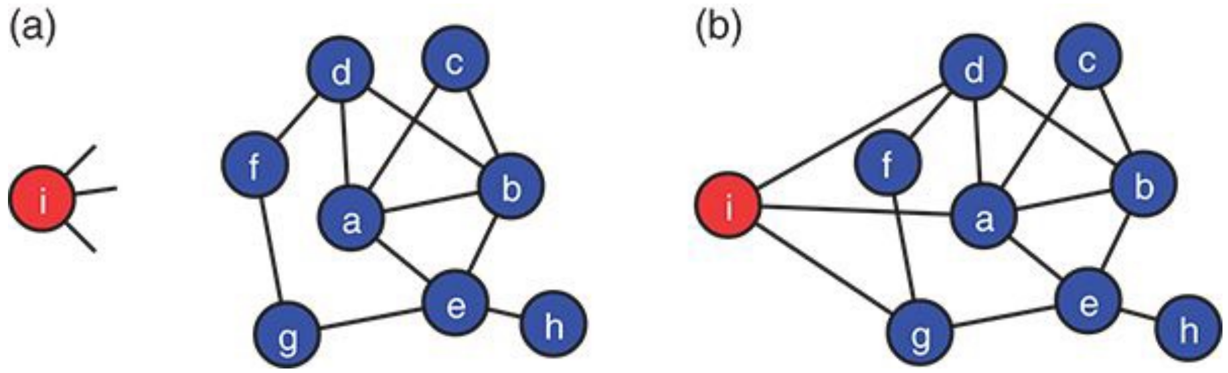
Preferencijalno pridruživanje

- Modeli koje smo do sada istraživali
  - Statični
  - Nema hubova ili znamo veličinu od početka – ne znamo zašto nastaju
- Broj čvorova poznat od početka, dodajemo veze
- Realne mreže su obično dinamičke
- Popularne mreže: Internet, Twitter, LinkedIn
- Veličina stvarnih mreža raste
- Čvorovi mogu nestajati
- Veća vjerojatnost pojave novih čvorova u mreži
- Dinamički modeli rasta

Preferencijalno pridruživanje

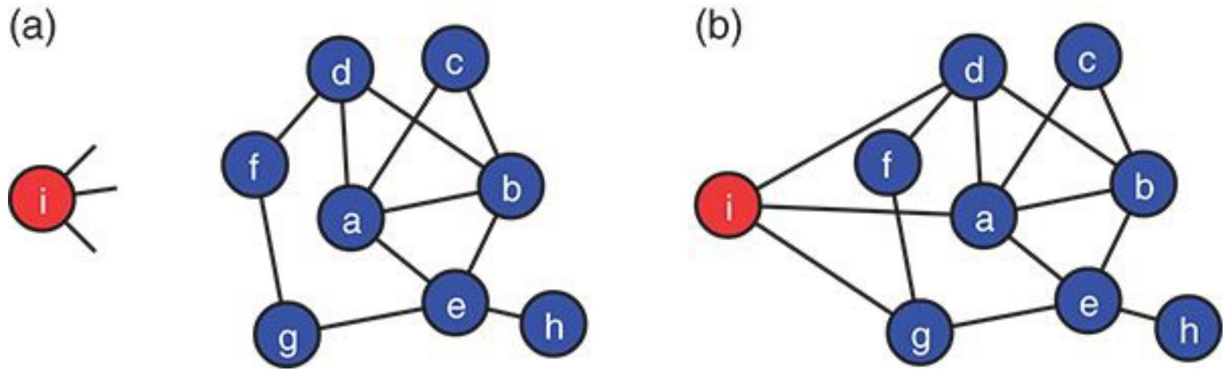
- Krećemo od inicijalne konfiguracije

  (npr. jako mala klika)
- Dodajemo čvor po čvor
- Novi čvor dodajemo na neki broj postojećih čvorova na osnovu pravila – karakteristika modela



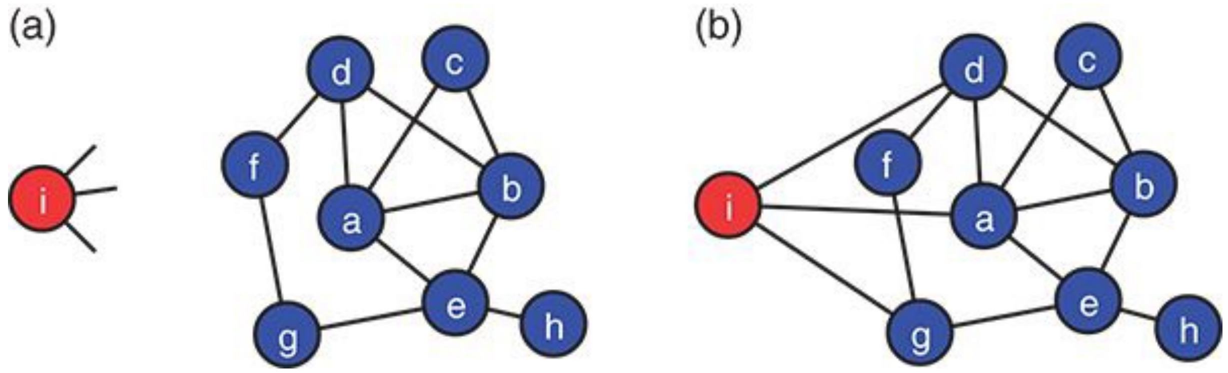
Preferencijalno pridruživanje

- Slučajne mreže i mreže malog svijeta bez hubova
- Jednakost – čvorovi biraju susjede totalno slučajno
- Niti jedan čvor nema prednost pred ostalima
- Mehanizam koji favorizira pojedine čvorove – preferencijalno pridruživanje
- Veći stupanj - > više novih veza



Preferencijalno pridruživanje

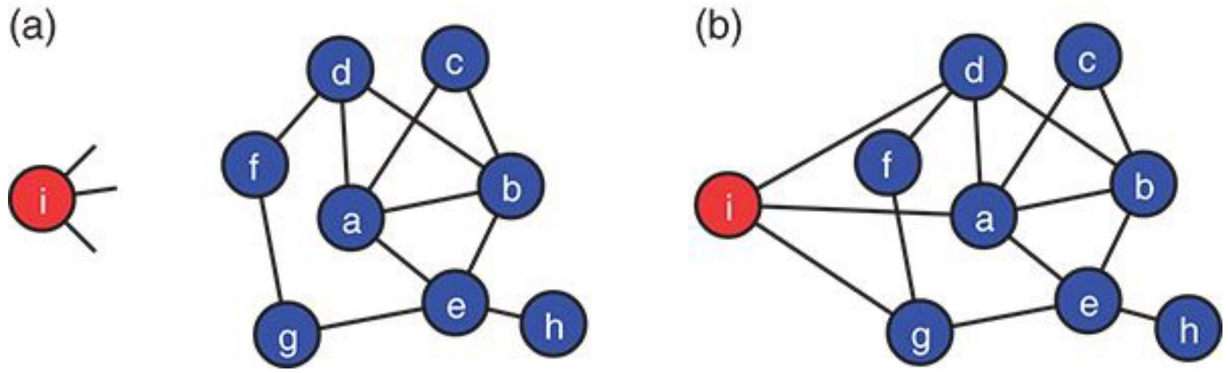
- **Primjer Weba**
- Bilijuni stranica
- Većina stranica kojih smo svjesni – popularne s velikim brojem veza
- Nova stranica – favoriziramo povezivanje na popularne, visoko povezane stranice
- Sličan primjer sa citatima – citiramo one citirane od strane drugih autora
- Čvorovi s visokim stupnjem - veća vjerojatnost povezivanja



Albert – Barabasi model

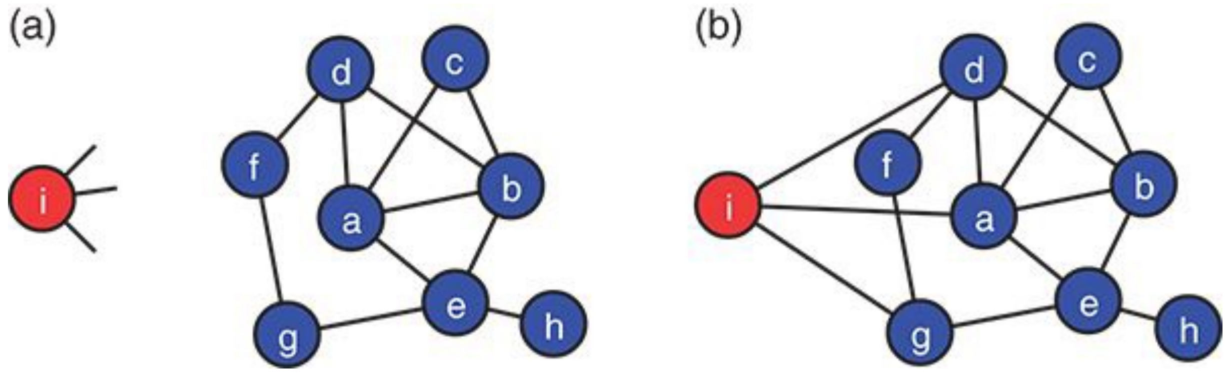
- Model preferencijalnog povezivanja
- Dodavanje novog čvora na postojeći je proporcionalno stupnju postojećeg
- Evanđelje po Mateju (Matthew):

"Tko ima, dat će mu se još
pa će obilovati, a onome tko nema
oduzet će se i ono što ima"



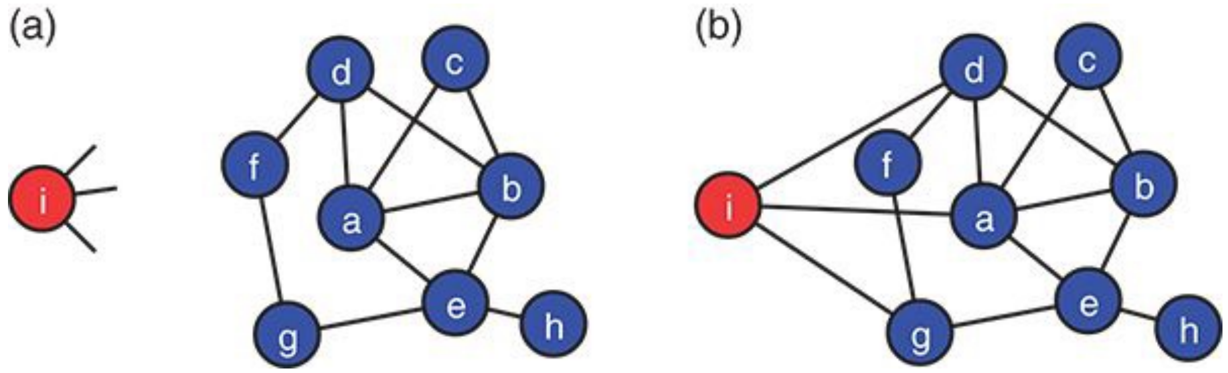
Preferencijalno pridruživanje

- Bogati postaju još bogatiji
- Siromašniji postaju siromašniji
- Dodatni nazivi
  - Matthew efekt
  - Kumulativna prednost



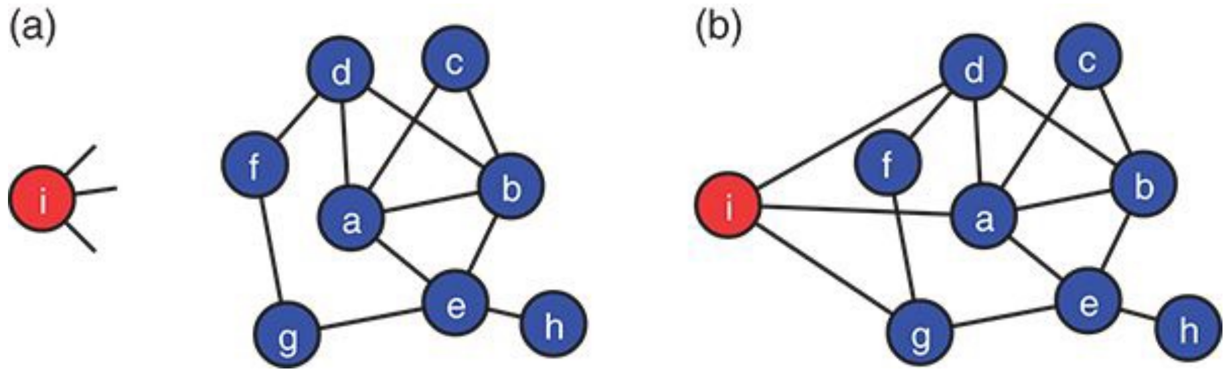
Preferencijalno pridruživanje

- Preferencijalno pridruživanje se koristi za objašnjenje otežanog repa distribucije
  - Broj vrsta biljaka u pojedinom genusu
  - Broj riječi u tekstu
  - Populacija gradova
  - Individualno bogatstvo
  - Znanstvena produkcija
  - Statistika citiranosti



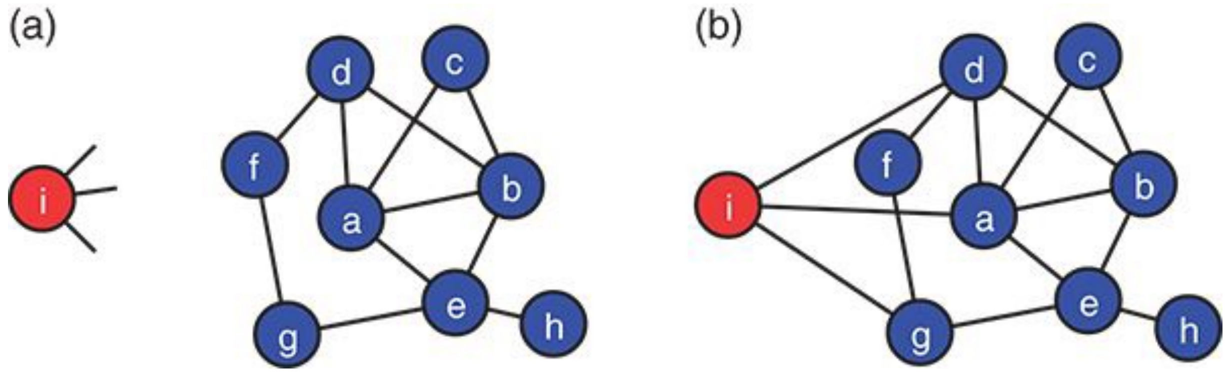
Preferencijalno pridruživanje

- Krećemo s potpunim grafom s m_0 čvorova. Svaka iteracija se sastoji od dva koraka:
    1. Novi čvor se dodaje u mrežu s $m \leq m_0$ novih veza povezanih na njega. Parametar m – prosječan stupanj
    2. Svaka nova veza s povezuje s postojećim čvorom j s vjerojatnošću $\Pi(i \leftrightarrow j) = \frac{k_j}{\sum_l k_l}$
- Ponavljamo dok ne dođemo do željenog broja čvorova

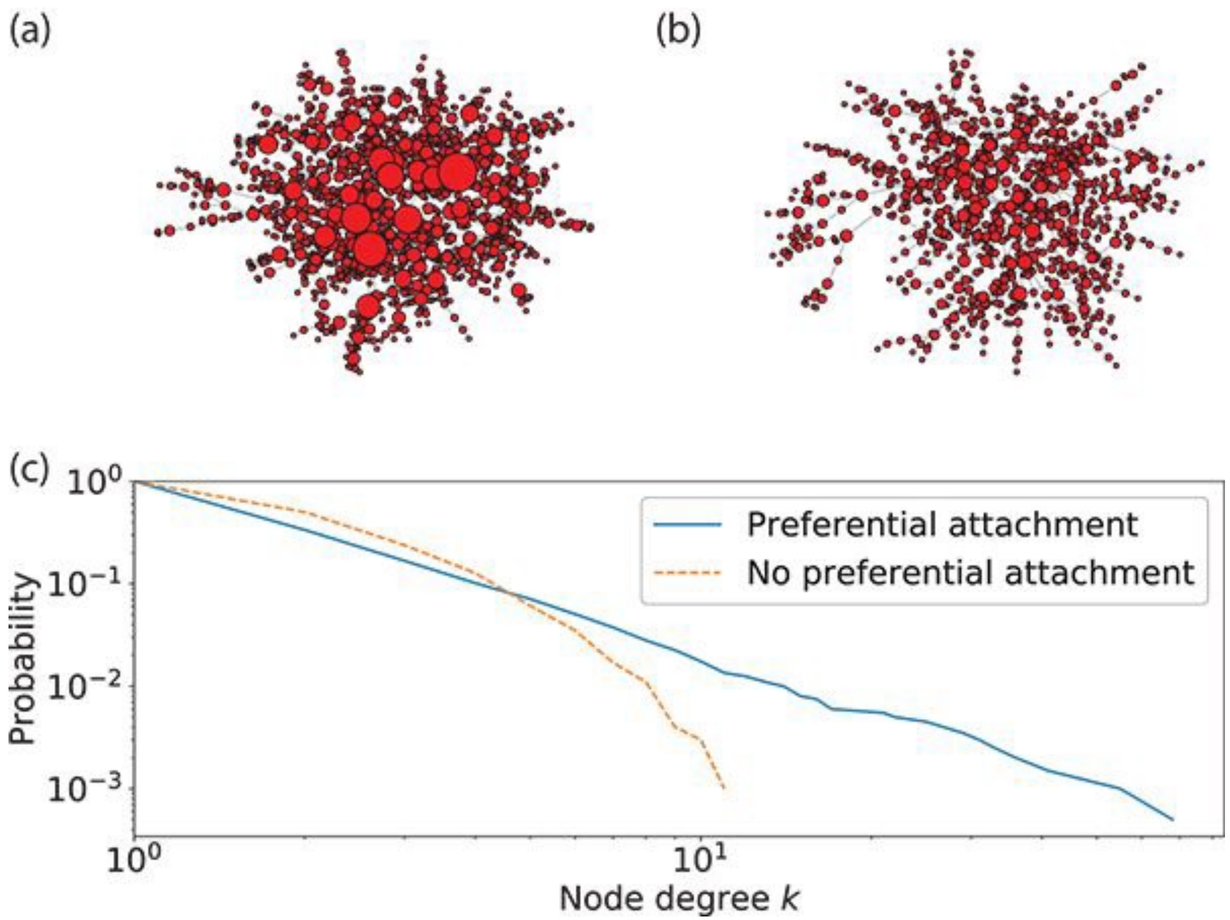


Preferencijalno pridruživanje

- Na početku svi čvorovi imaju isti stupanj
- Dodavanjem novih čvorova i veza raste stupanj čvorova
- Najstariji čvorovi mogu dobiti nove veze u bilo koje trenutku – prednost pred onim koji su kasnije dodani
- Stupanj starijih prelazi onaj novih -> raste vjerojatnost povezivanja
- Bogati postaju bogatiji
- Najstariji čvorovi postaju hubovi



Preferencijalno pridruživanje



- a) Mreža generirana Albert-Barabasi modelom
- b) sličan model rasta sa slučajnim odabirom umjesto preferencijalnog
- c) kumulativne distribucije

Drugi preferencijalni modeli

- Albert-Barabasi model – linearno pridruživanje
- Pridruživanje s potencijom stupnja – nelinearno pridruživanje
- $\Pi_\alpha(i \leftrightarrow j) = \frac{k_j^\alpha}{\sum_l k_l^\alpha}$
- Za $\alpha = 1$ –Albert-Barabasi (AB) model. Za $\alpha \neq 1$, imamo dva slučaja:

1. $\alpha < 1$, vjerojatnost veze raste sporije nego kod AB modela -> manja razlika u veličini čvorova, nema otežanog repa i nestanak hubova
2. $\alpha > 1$, čvorovi visokog stupnja puno brže akumuliraju nove veze nego oni niskoga -> jedan od čvorova će biti povezan s velikim udjelom ostalih. Kada $\alpha > 2$, pobjednik uzima sve – jedan čvor povezan sa svima ostalima koji imaju sličan mali stupanj

- Nelinearno pridruživanje ne stvara hubove koje vidimo u stvarnim mrežama

Ograničenja AB modela

- Mora biti striktno linearno pridruživanje
- Nagib krivulje kumulativne distribucije neovisan o izboru parametara modela – stvarne distribucije mogu opadati brže ili sporije
- Hubovi su najstariji čvorovi, novi ne mogu postići veći stupanj
- Ne proizvodi trokute -> koeficijent klasteriranja manji nego u stvarnim mrežama
- Čvorovi i veze mogu se samo dodati, u stvarnim mrežama mogu nestati
- Jedna povezana komponenta. Stvarna mreža može imati više komponenti

Model privlačnosti

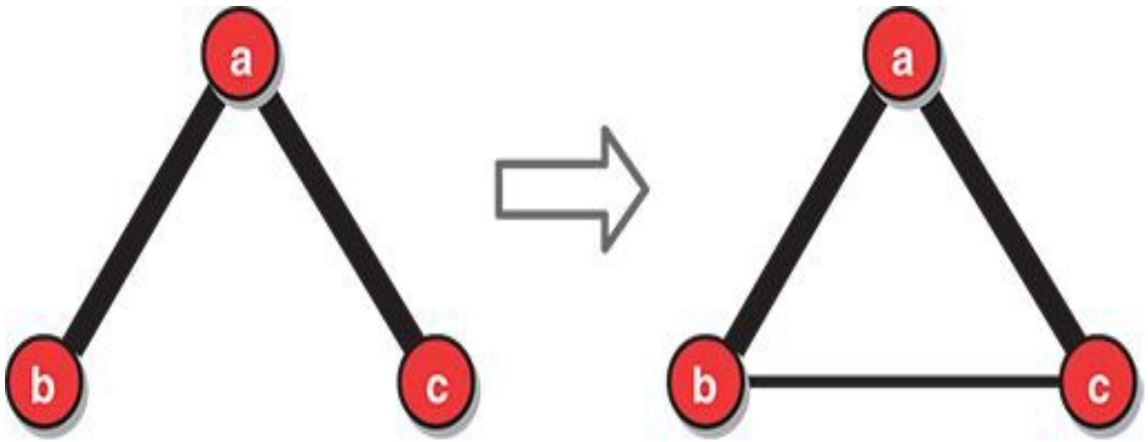
- Preferencijalno pridruživanje – što ako čvor nema susjeda ? -> vjerojatnost dodijele veze je nula!
- Usmjerene mreže, vjerojatnost veze ovisi samo o ulaznom stupnju -> problem su novododani čvorovi koji svi imaju ulazni stupanj nula
- Rješenje
  - Intristična privlačnost
  - Vjerojatnost veze proporcionalna zbroju stupnja i konstantne privlačnosti
  - $\Pi(i \leftrightarrow j) = \frac{A+k_j}{\sum_l (A+k_l)}$
  - Za *A=0* -> AB model

Model sposobnosti

- Primjer
  - Google (1998)
  - Prije njega milijuni stranica
  - Usprkos tome najpopularniji
  - Slično s novim znanstvenima radovima (citati)
- Individualno svojstvo svakog čvora (sposobnost)
- Svakom čvoru *i* dodijelimo sposobnost $\eta_i > 0$ iz distribucije $\rho(\eta)$
- $\Pi(i \leftrightarrow j) = \frac{\eta_j k_j}{\sum_l \eta_j k_l}$, ako su sve sposobnosti iste dolazimo do AB
  - $\rho(\eta)$ po volji velika vrijednosti -> pobjednik uzima sve efekt
  - $\rho(\eta)$ ograničen -> distribucija ima otežani rep
- Pojava više hubova
- Velika sposobnost omogućava novim čvorovima takmičenje bez obzira na starost i status

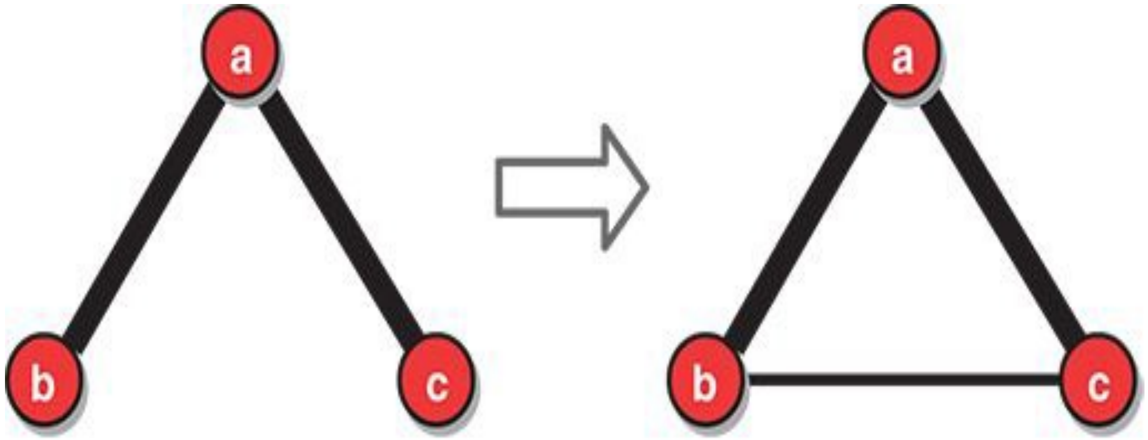
Model slučajne šetnje

- Mreže dobivene AB modelom imaju niski koeficijent klasteriranja
- U AB modelu vjerojatnost da čvor dobije vezu je proporcionalna stupnju bez obzira je li novi par susjeda ima zajedničkog susjeda
- Trokuti se rijetko formiraju
- Potrebno dodati mehanizam koji favorizira kreiranje veza između čvorova sa zajedničkim susjedima



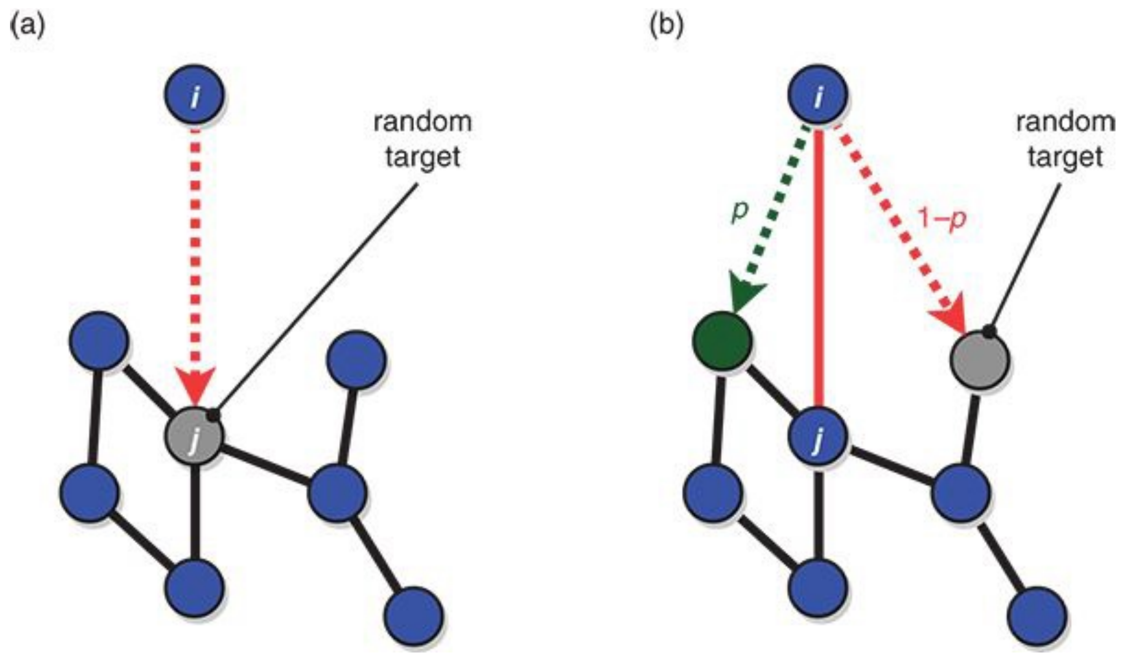
Model slučajne šetnje

- Formiranje trokuta dodavanjem veze se naziva zatvaranje triade
- U praksi mnogo ljudi su upoznati preko zajedničkog poznanika



Model slučajne šetnje

- m > 1 novih veza
- Novi čvor *i* se dodaje na slučajno odabran čvor *j*
- Svaka dodatna veza na *i* se dodaje na susjeda od *j s* vjerojatnošću *p* što vodi kreiranju trokuta
- Inače je povezan sa slučajno odabranim čvorom



Princip jakog zatvaranja triada

- Osoba *a* ima jaku vezu s *b* i *c* -> *b i c* su prijatelji ili će to postati
- Ako *b* i *c* provode puno vremena s *a*, vrlo vjerojatno da će sresti preko *a*
- S obzirom da je *a* dobar prijatelj s oba, *b* i *c* teže da vjeruju jedan drugome
- Ako *b* i *c* ignoriraju jedan drugog to može biti izvor stresa za grupu
- Princip jakog zatvaranja triada propisuje da mora biti veza između *b* i *c*
- *”The strength of weak ties”* Mark S. Granovetter (1973). Bliska veza između trokuta, težina veza i zajednica

Društvene zajednice

- Veze s velikom težinom signaliziraju čvrste veze - unutar zajednice
- Slabe veze - između zajednica
- Slabe veze su kritične za strukturu socijalnih mreža, zbog povezivanja zajednica i omogućuju prijenos informacija kroz mrežu