Kompleksne mreže

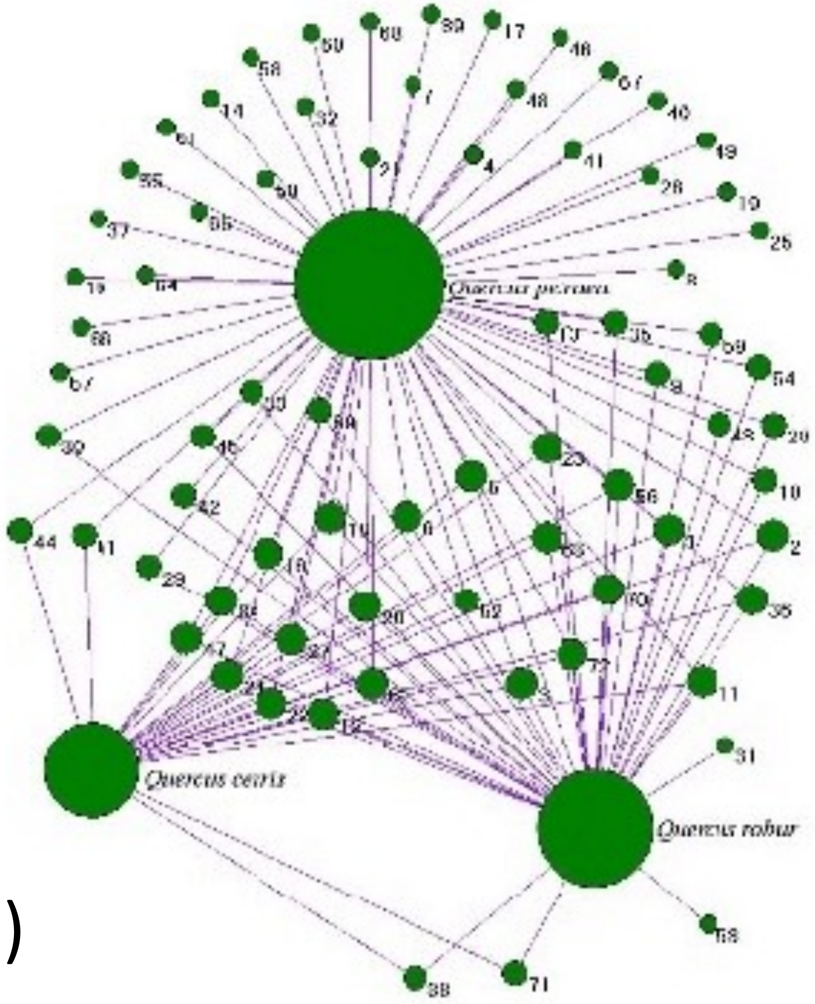
3. predavanje

Što im je zajedničko ?

- Elon Mask
- Khaby Lame
- Christiano Ronaldo
- Blac Chyna
- Bill Gates
- James Stephen Donaldson

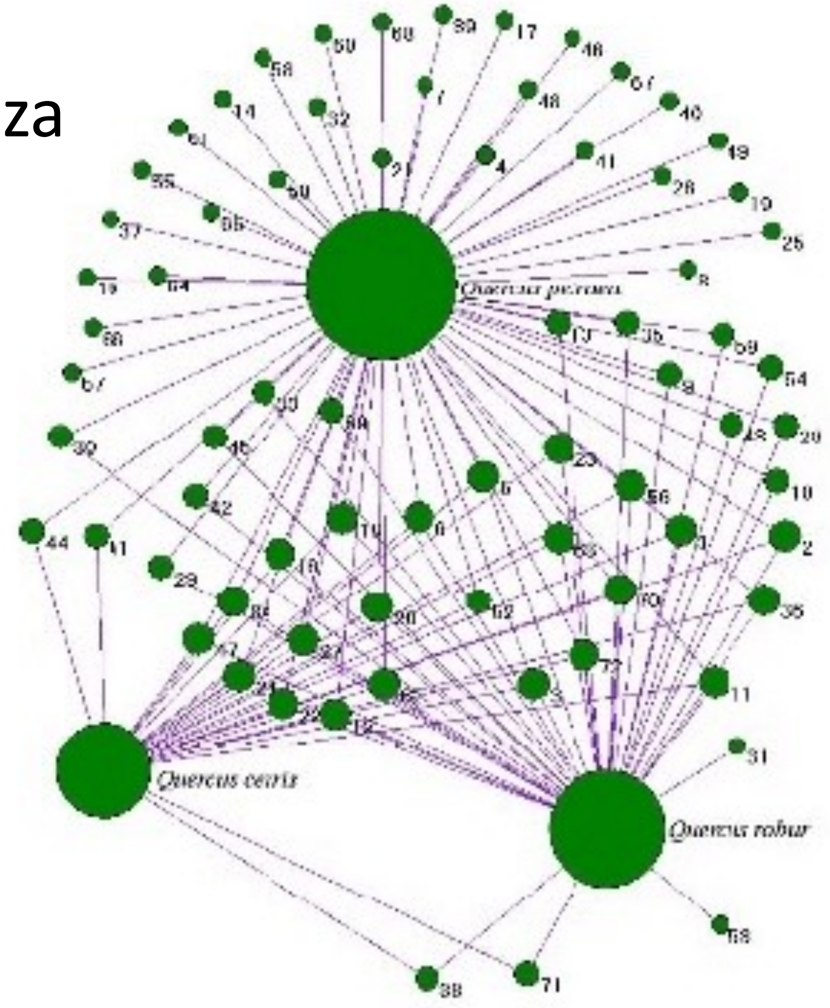
Hubovi

- Primjeri
  - Aerodromi
    - Amsterdam
    - Doha
    - Singapore
    - Istanbul
- Popularne stranice
- Influenceri (X, Instagram, TikTok, LinkedIn)



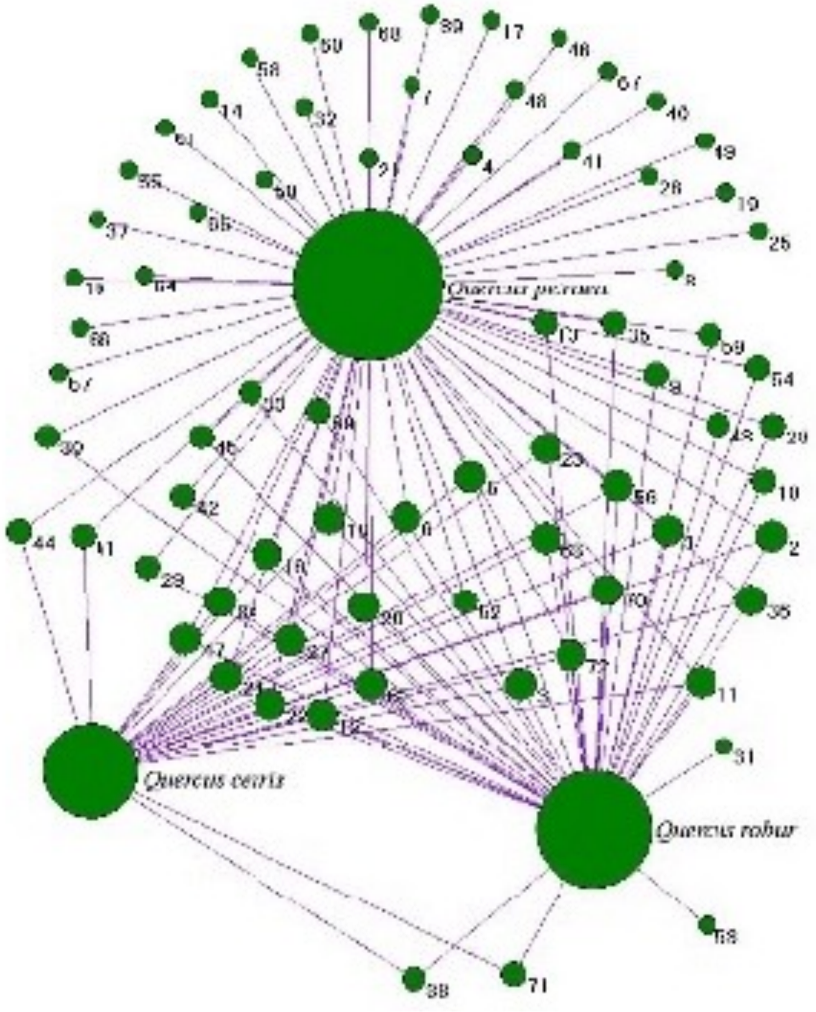
Heterogenost mreža

- Raznolikost u svojstvima i ulogama čvorova i veza
- Izvor heterogenosti – stupanj čvora
- Mjerenje centralnosti
- Hubovi – čvorovi visokog stupnja



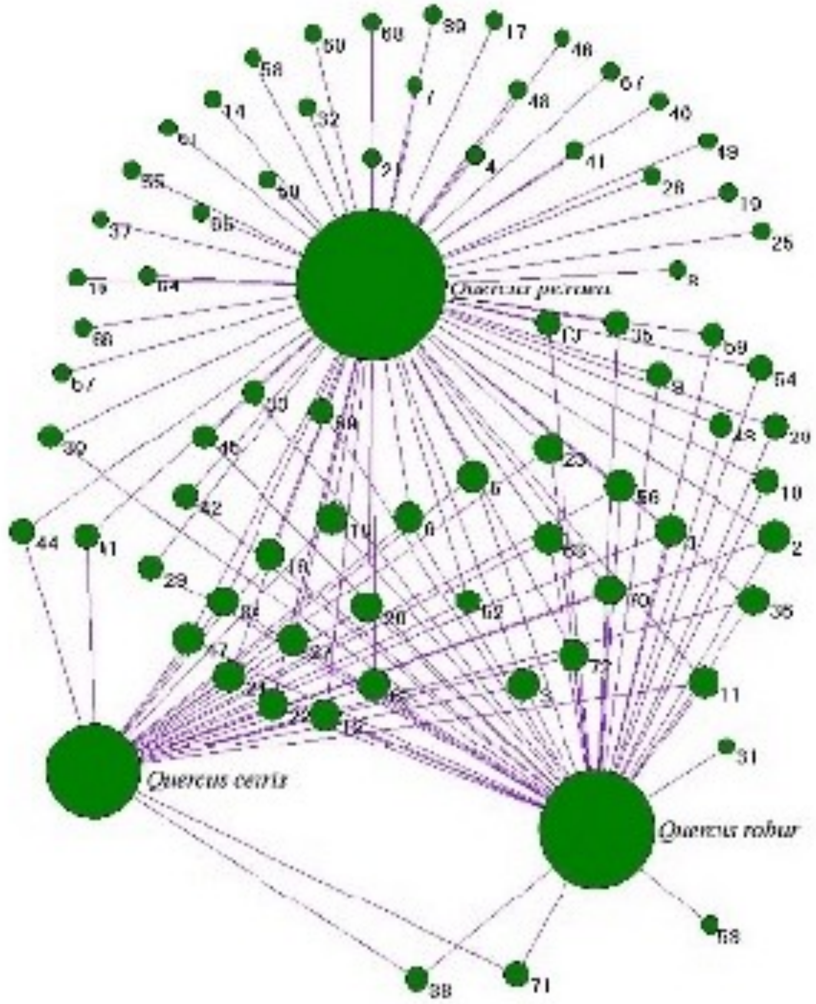
Osnovne mjere centralnosti

- Stupanj (*degree*)
- Bliskost (*closeness*)
- Međupoloženost (*betweenness*)



Stupanj

- Primjeri:
  - Broj direktno dostupnih aerodroma od polaznog
  - Broj socijalnih veza
- Prirodna mjera centralnosti
- Prosječan stupanj – ne odražava heterogenost



Bliskost

- Koliko je pojedini čvor blizak ostalima
- $g_i = \frac{1}{\sum_{j \neq i} l_{ij}}$, gdje je l_{ij} udaljenost između *i* i *j*
- $\tilde{g}_i = (N-1) g_i = \frac{N-1}{\sum_{j \neq i} l_{ij}} = \frac{1}{\sum_{j \neq i} l_{ij}/(N-1)}$ alternativna formulacija
  - Mjera postaje usporediva među različitim mrežama
  - $\sum_{j \neq i} l_{ij}/(N-1)$ je zapravo prosječna udaljenost čvora *i* od ostatka mreže
  - Inverzija prosječne udaljenosti čvora

Međupoloženost

- Difuzni procesi u mreži
  - Prenošenje informacije kroz socijalne mreže
  - Promet dobara kroz luku
  - Širenje epidemije u mreži fizičkih kontakata
- Uključenost čvora u te procese – međupoloženost
  - Računamo koliko najkraćih puteva prolazi čvorom

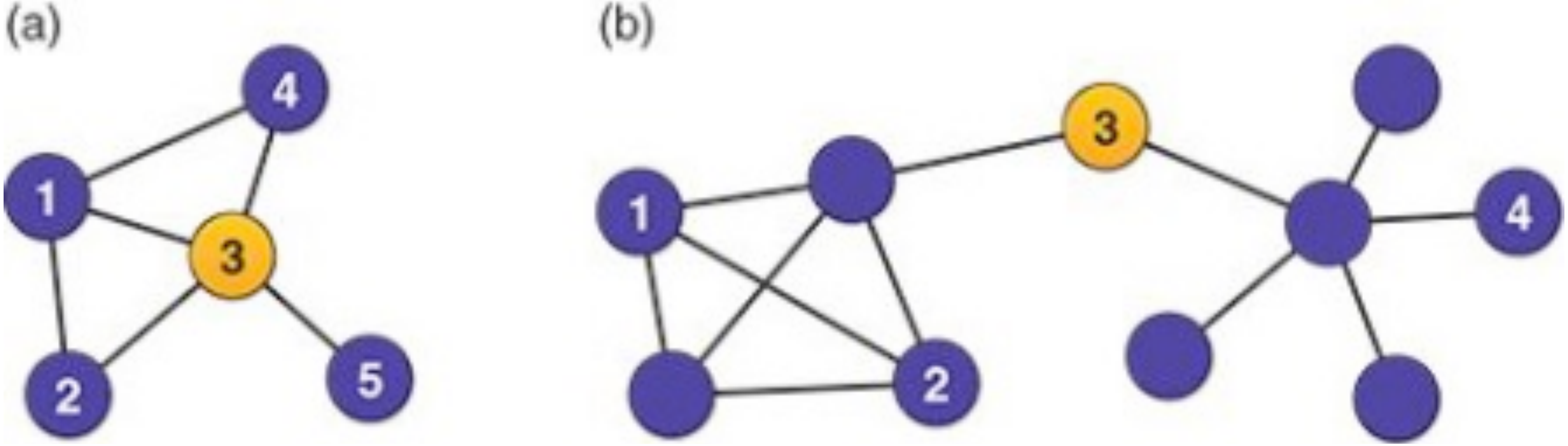
Međupoloženost

- Može postojati više najkraćih puteva između dva čvora u mreži iste duljine
- σ_{hj} ukupan broj najkraćih puteva između čvorova *h* i *j*
- $\sigma_{hj}(i)$ ukupan broj najkraćih puteva koji prolaze čvorom *i*
- međupoloženost je definirana kao

$$b_i = \sum_{h \neq j \neq i} \frac{\sigma_{hj}(i)}{\sigma_{hj}}$$

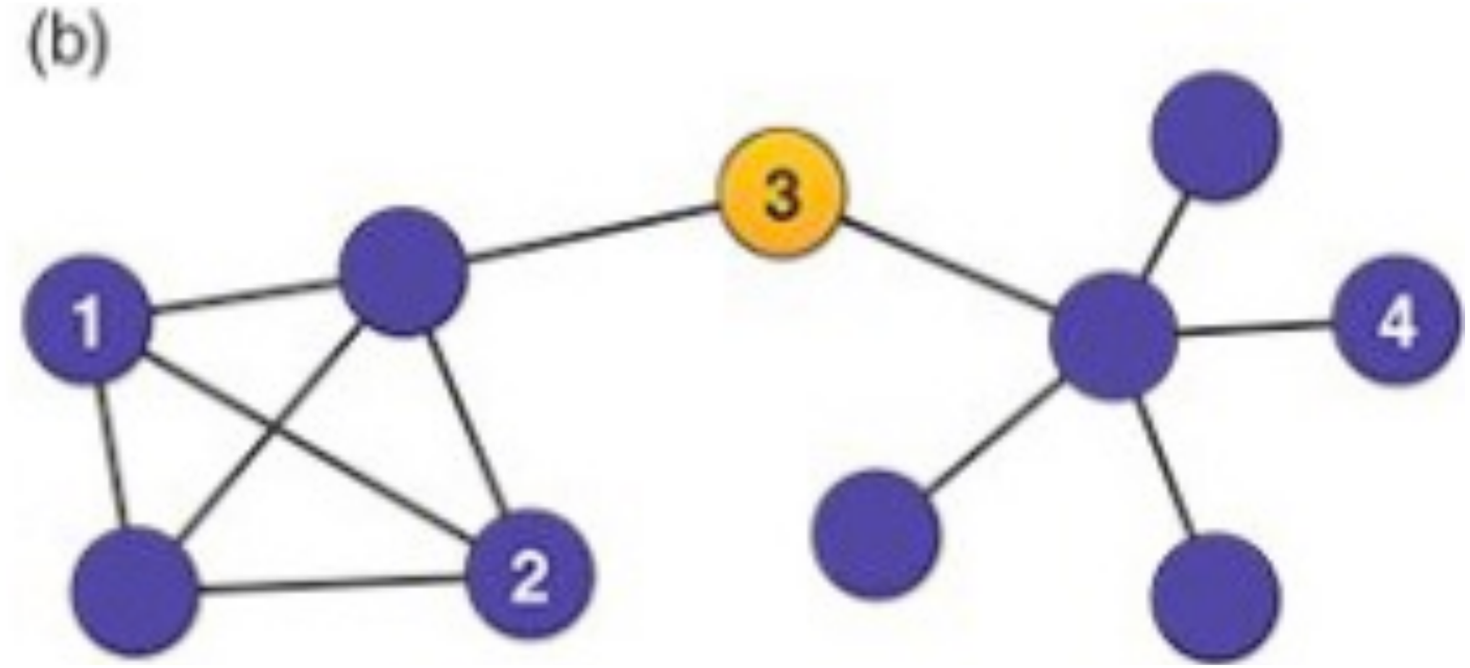
- Normalizacija s $\binom{N-1}{2}$ - ukupan broj parova čvorova bez *i*

Međupoloženost



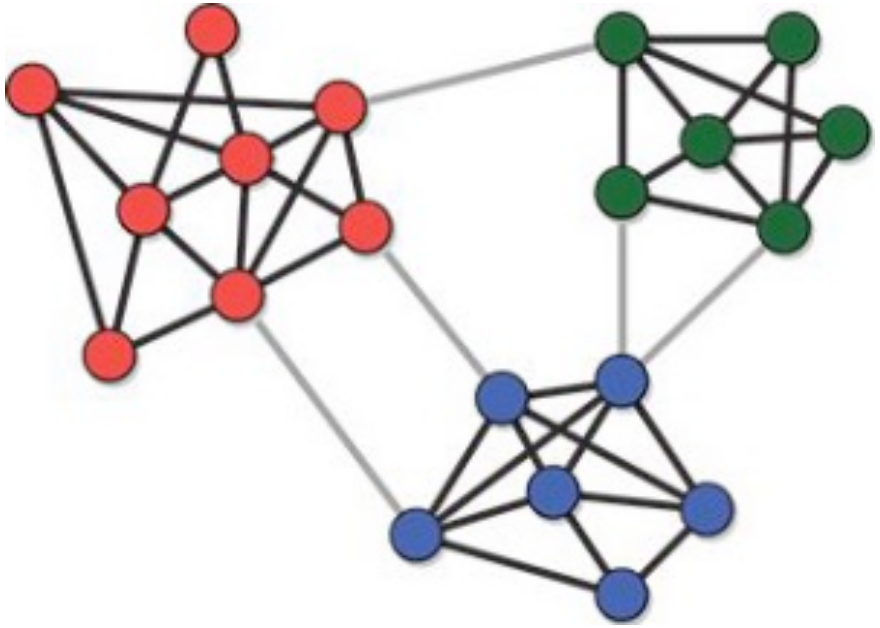
Međupoloženost

- Čvor s velikim iznosom međupoloženosti ne mora imati veliki stupanj
- Čvorovi koji premošćuju regije mreže imaju visoku međupoloženost



Međupoloženost veza

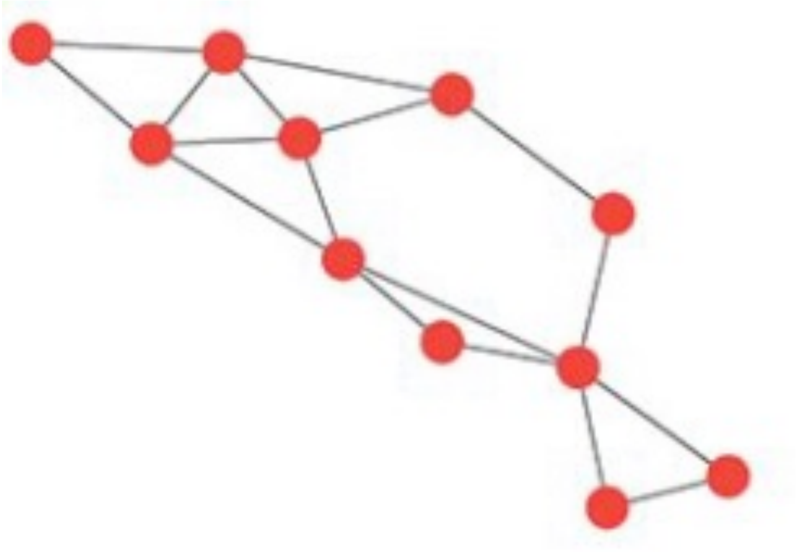
- Dio svih najkraćih putova svih čvorova koji prolaze tom vezom
- Veze s velikom međupoloženošću često povezuju povezane regije – zajednice
- Uklanjanjem veza s visokom međupoloženošću omogućava određivanje zajednica

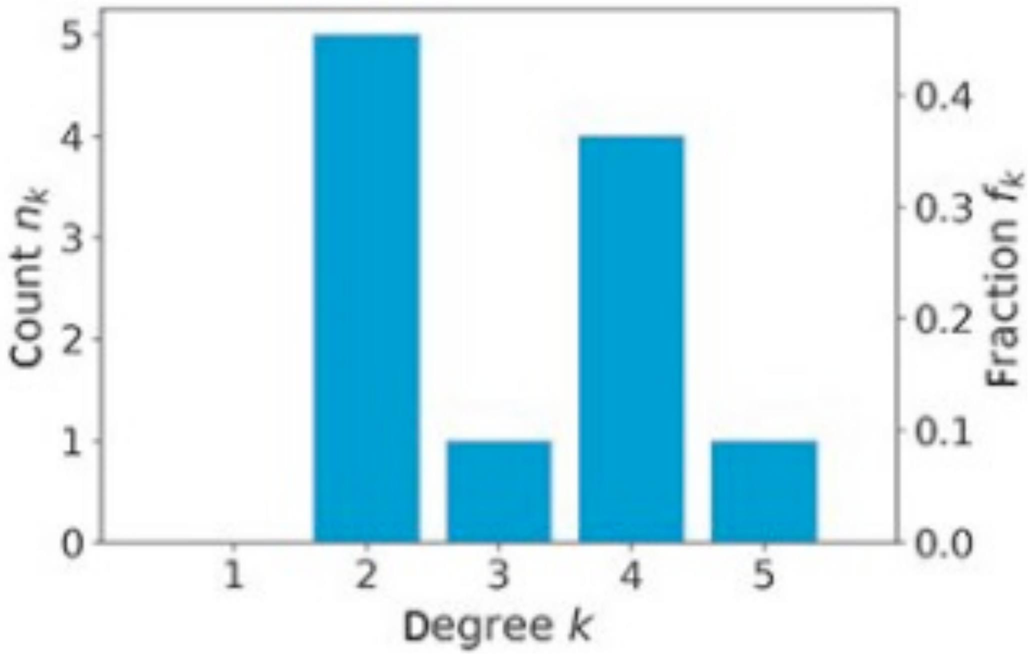


Distribucije centralnosti

- Veliki grafovi – svatko, bez obzira koliko popularan, vezan je uz manji dio mreže
- Statistički pristup:
  - fokus na klase čvorova i veza koje imaju slična svojstva
  - ne pojedini čvor ili veza
- Statistička distribucija – koliko elemenata (čvorova ili veza) ima istu vrijednost
- Utvrđivanje klasa elemenata iz distribucije

Distribucija stupnja za manju mrežu



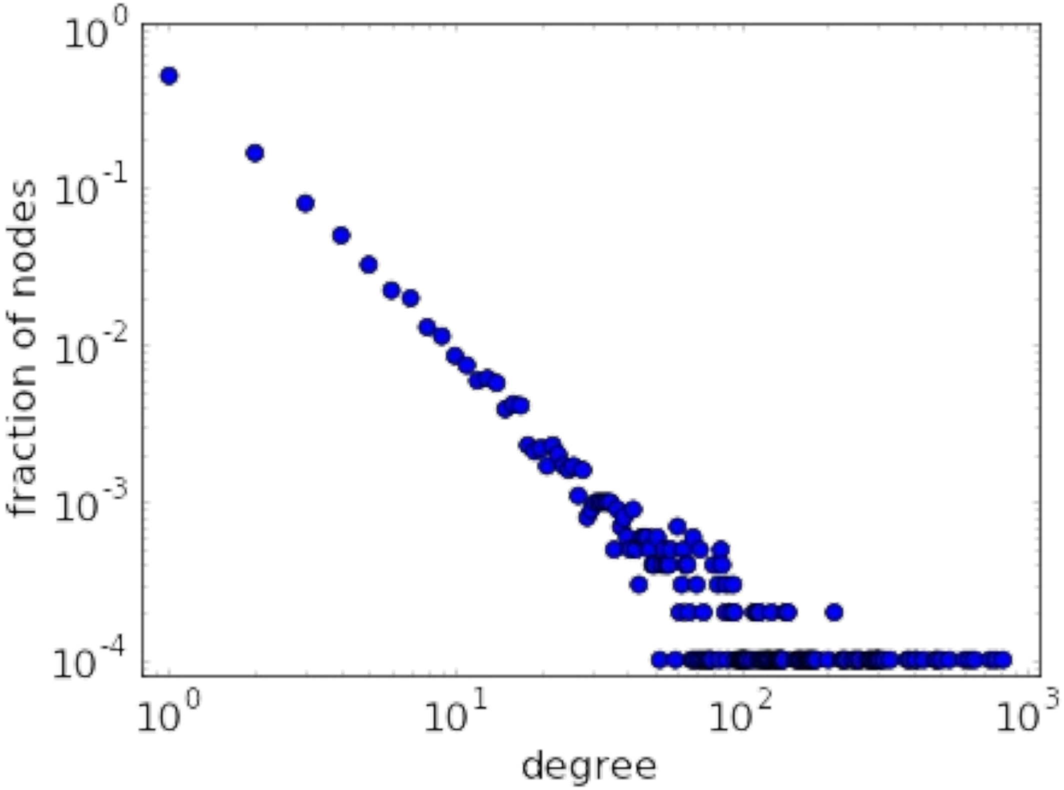


Komplementarna kumulativna distribucija

- *P(x)* - *v*jerojatnost da događaj ima veću vjerojatnost od *x*
- $P(x) = \sum_{v \geq x} f_v$
- Koristi se često kada je raspon varijabilnosti širok
  - Primjer: distribucija stupnja u kompleksnim mrežama
  - Visoke vrijednosti su rijetke -> šum u repu distribucije
  - Kumulativna distribucija to izgladi

Distribucija stupnja čvora

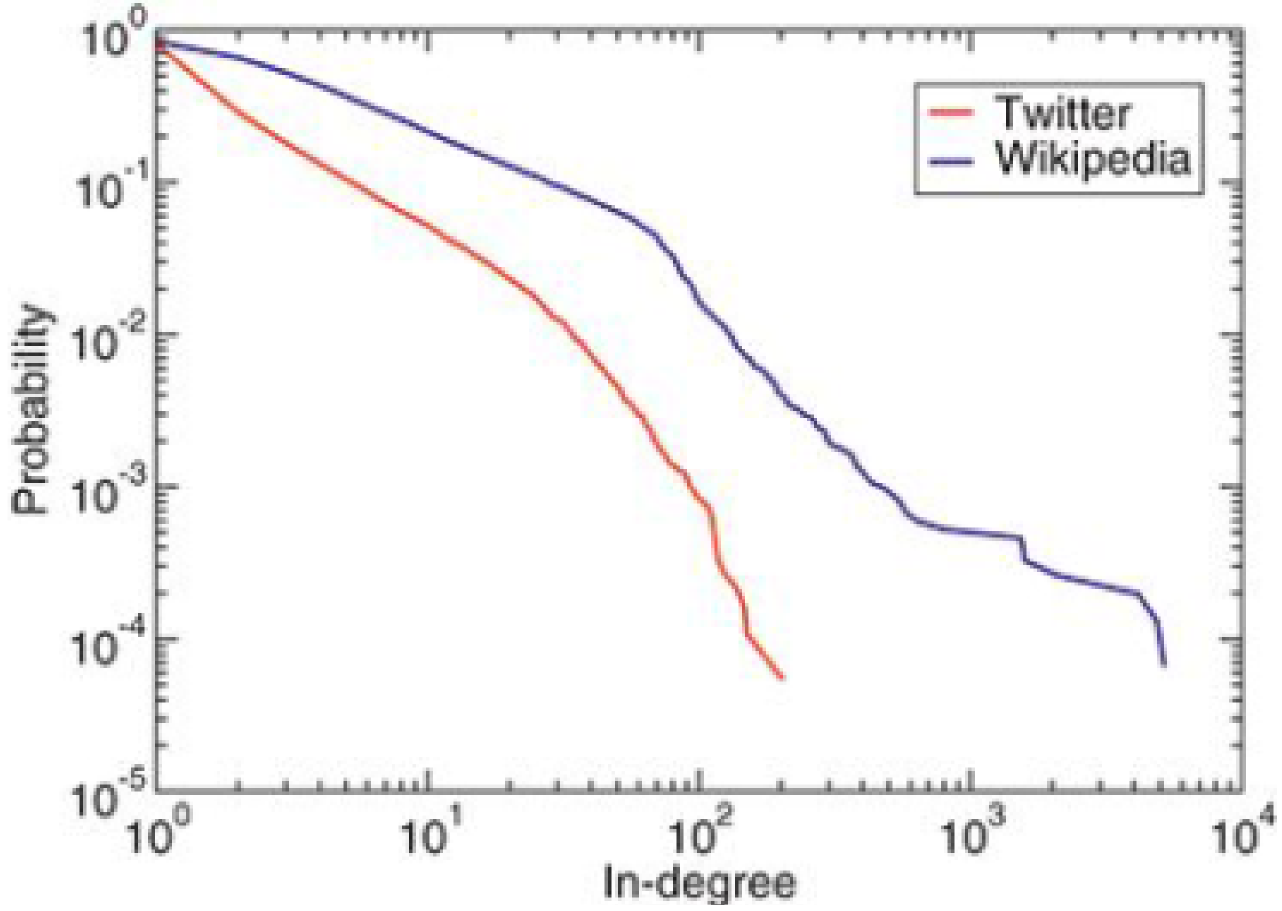
- Protežu se kroz nekoliko redova veličine
- Široke distribucije -> otežani rep -> kumulativna distribucija izgladi
- Distribucije s otežanim repom -> log log skala



Kumulativne distribucije

- Distribucije s otežanim repom pokazuju veliku heterogenost u vrijednostima stupnja čvora.
- Mnogo čvorovi samo nekoliko prijatelja, dio njih velik broj susjeda -> veća uloga u mreži -> takve čvorove nazivamo hubovima
- Mnoge prirodne, društvene, informacijske, ručno kreirane mreže -> otežani rep distribucije s visoko povezanim hubovima

Distribucija stupnja primjer



Heterogenost mreže

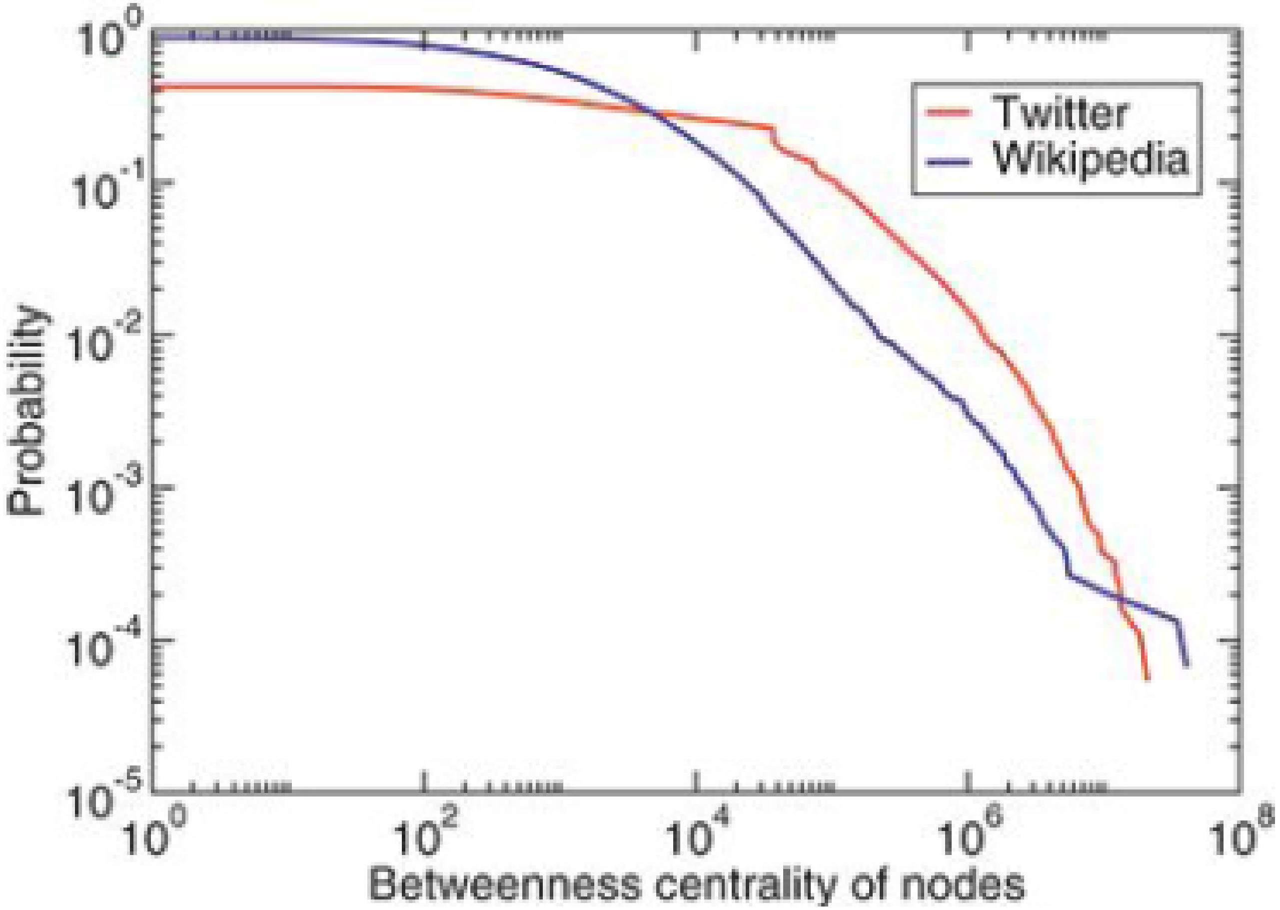
- Parametar heterogenosti $\kappa (kappa)$
- Heterogenost mrežnog stupnja distribucije
- Prosječan kvadratni stupanj $\langle k^2 \rangle = \frac{k_1^2 + k_2^2 + \cdots + k_{N-1}^2 + k_N^2}{N} = \frac{\sum_i k_i^2}{N}$
- $\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle^2}$ - parametar heterogenosti

Heterogenost - primjer

Network	Nodes (N)	Links (L)	Average degree ($\langle k \rangle$)	Maximum degree (k_{max})	Heterogeneity parameter (κ)
Facebook Northwestern Univ.	10,567	488,337	92.4	2,105	1.8
IMDB movies and stars	563,443	921,160	3.3	800	5.4
IMDB co-stars	252,999	1,015,187	8.0	456	4.6
Twitter US politics	18,470	48,365	2.6	204	8.3
Enron email	87,273	321,918	3.7	1,338	17.4
Wikipedia math	15,220	194,103	12.8	5,171	38.2
Internet routers	190,914	607,610	6.4	1,071	6.0
US air transportation	546	2,781	10.2	153	5.3
World air transportation	3,179	18,617	11.7	246	5.5
Yeast protein interactions	1,870	2,277	2.4	56	2.7
C. elegans brain	297	2,345	7.9	134	2.7
Everglades ecological food web	69	916	13.3	63	2.2

Heterogenost izračunata koristeći ulazni stupanj

Distribucija međupoloženosti - primjer



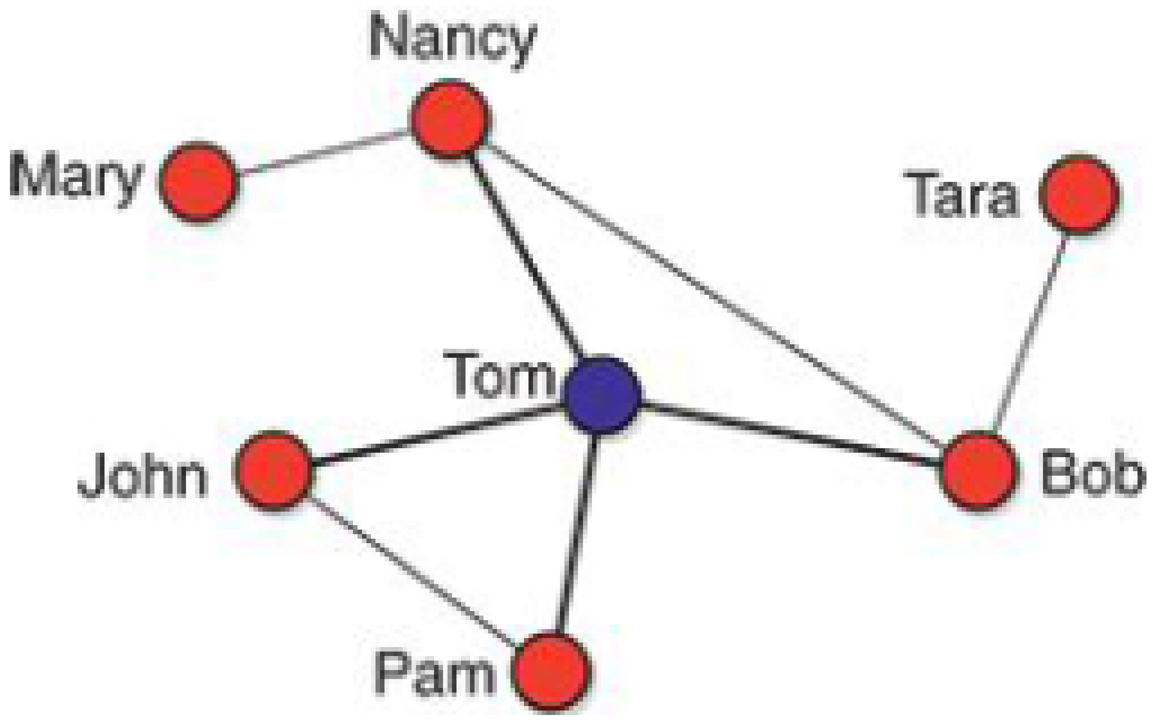
Izračunato samo za najveću komponentu Wikipedie

Paradoks prijateljstva

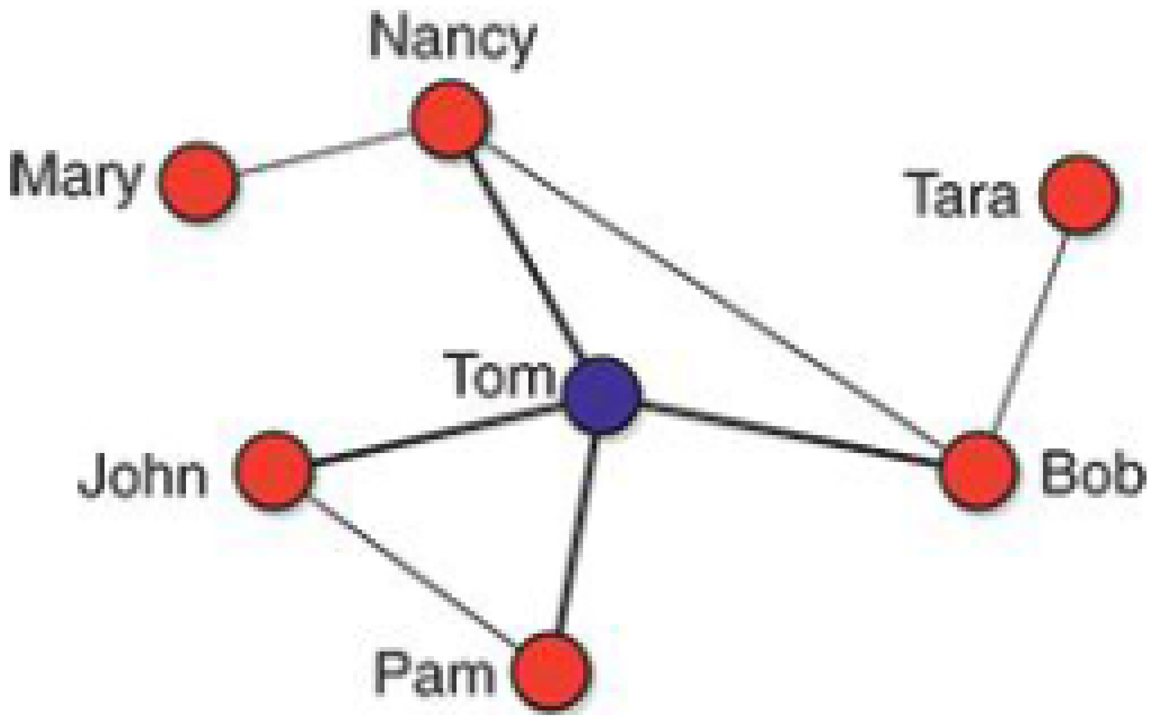
- Većina ljudi ima manje prijatelja nego što to imaju njihovi prijatelji u prosjeku ☺

- Tražimo osobu s najviše prijatelja u grupi od N ljudi
- Imamo samo njihove brojeve telefona i jedan poziv
- Ako slučajnim odabirom nazovemo jednu osobu, vjerojatnost je $1/N$
- Što ako nazovemo jednu osobu i pitamo ju neka navede jednog prijatelja?

Paradoks prijateljstva



Paradoks prijateljstva



Tom je osoba s najviše prijatelja. Isprobamo obje strategije

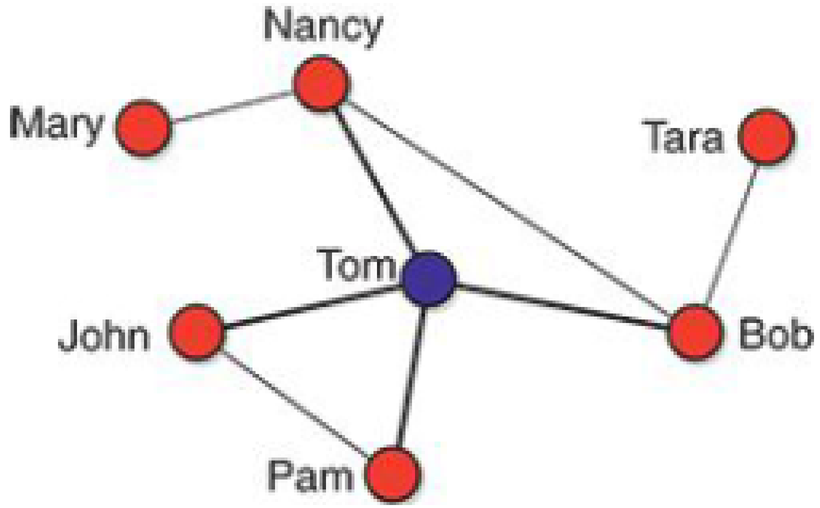
- Slučajno nazovemo jednu osobu – vjerojatnost 1/7 ≈ 0.14
- Nazovemo osobu i pitamo ju za jednog prijatelja 5/21 ≈ 0.24

Paradoks prijateljstva

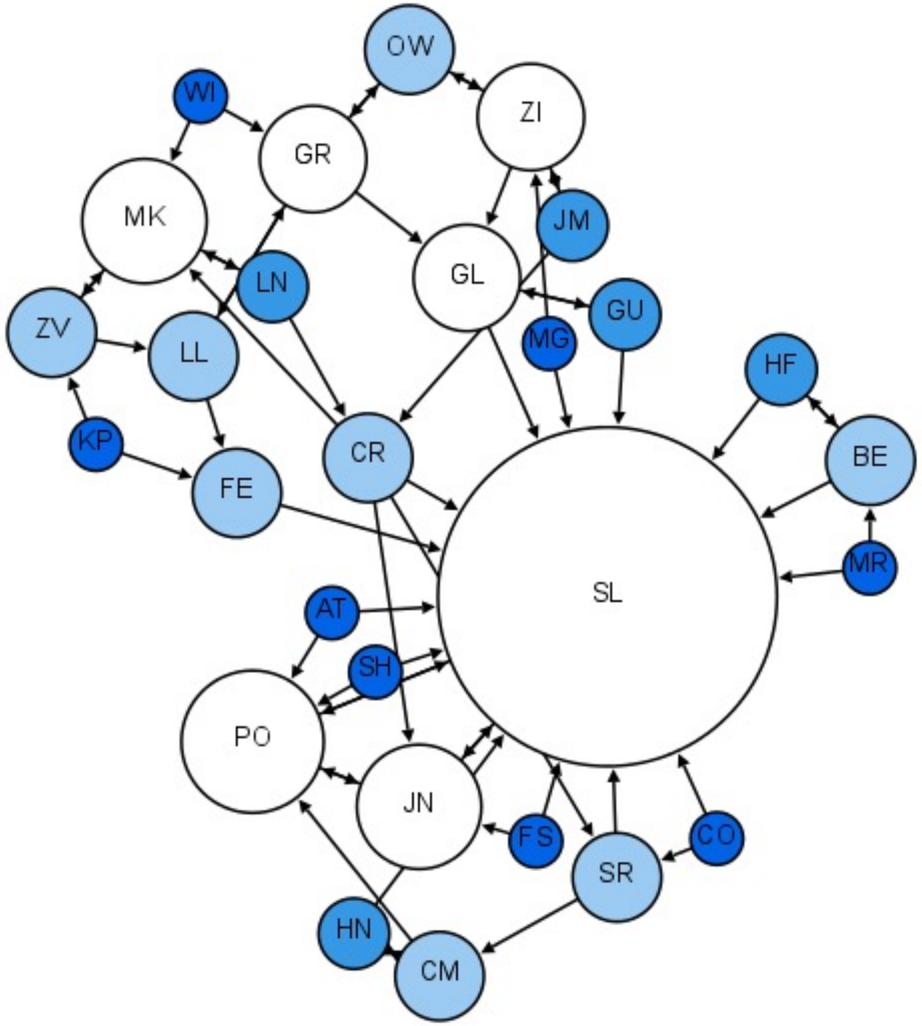
- Biramo veze umjesto čvorova
- Kada biramo čvorove svaki ima jednaku vjerojatnost biti izabran, neovisno o stupnju
- Kada biramo veze, veći broj susjeda -> veća vjerojatnost dohvata
- U našem slučaju – 4 moguća kanal prema Tomu
- Vjerojatnost pronalaska huba raste ako se prebacimo na drugo susjedstvo – raste broj veza

Paradoks prijateljstva

- Prosječan broj susjeda od susjeda čvora je 17/6 = 2.83
- Prosječan stupanj mreže je (1+3+3+1+4+2+2)/7=2.29
- Prosječan stupanj susjeda je veći od prosječnog stupnja čvora
- Šira distribucija stupnja -> jači efekt
- Posebno jak u slučaju distribucije s otežanim repom



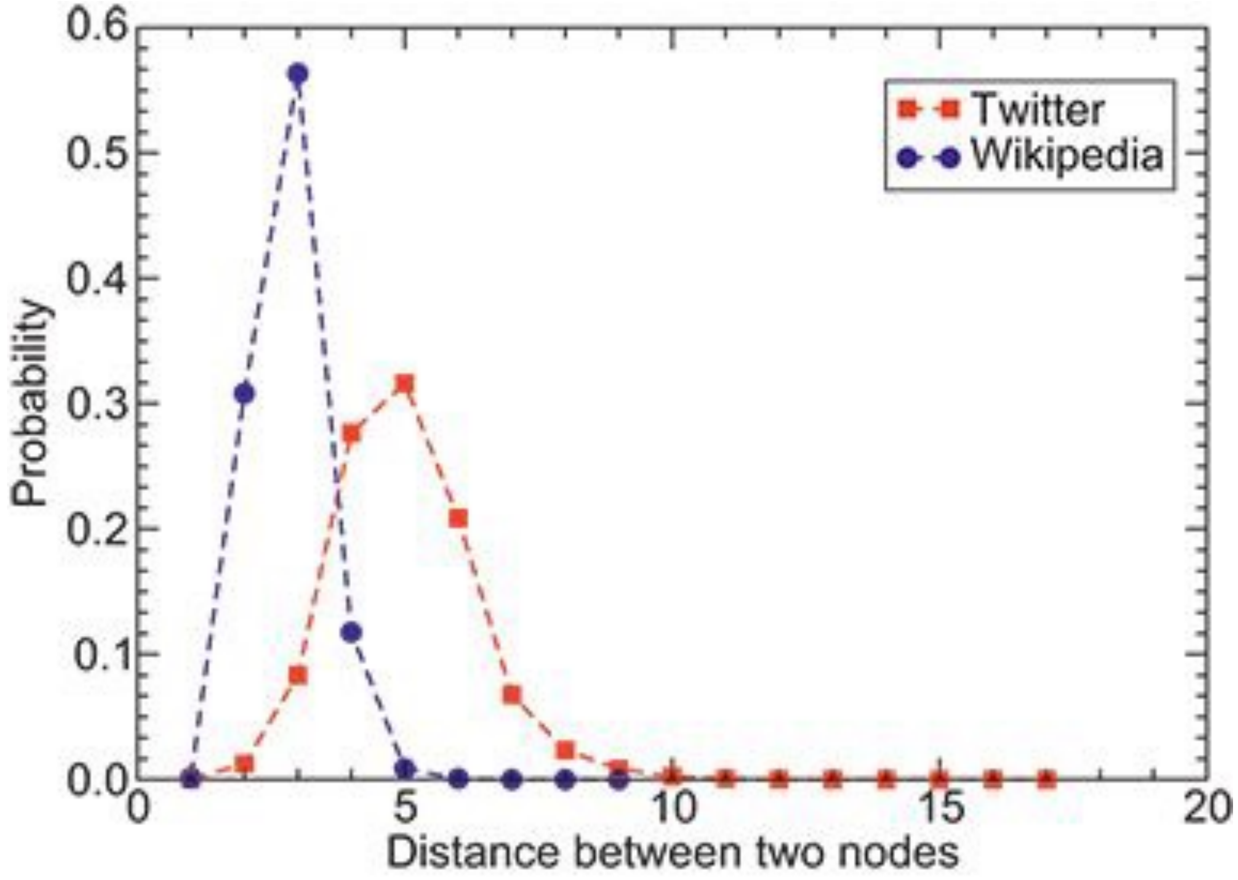
Prvi razred. Dvoje s kojima bi željeli sjediti



Ultra mali svijet

- Hubovi jednostavni za pronaći i postoji velika potražnja za njima
- Npr. Ako želimo letjeti iz A u B. Ako nema direktnog leta -> najbliži hub
- Mreže sa širokom distribucijom stupnja imaju svojstvo ultra malog svijeta
- Ultra-mali svijet – udaljenost između čvorova jako kratke

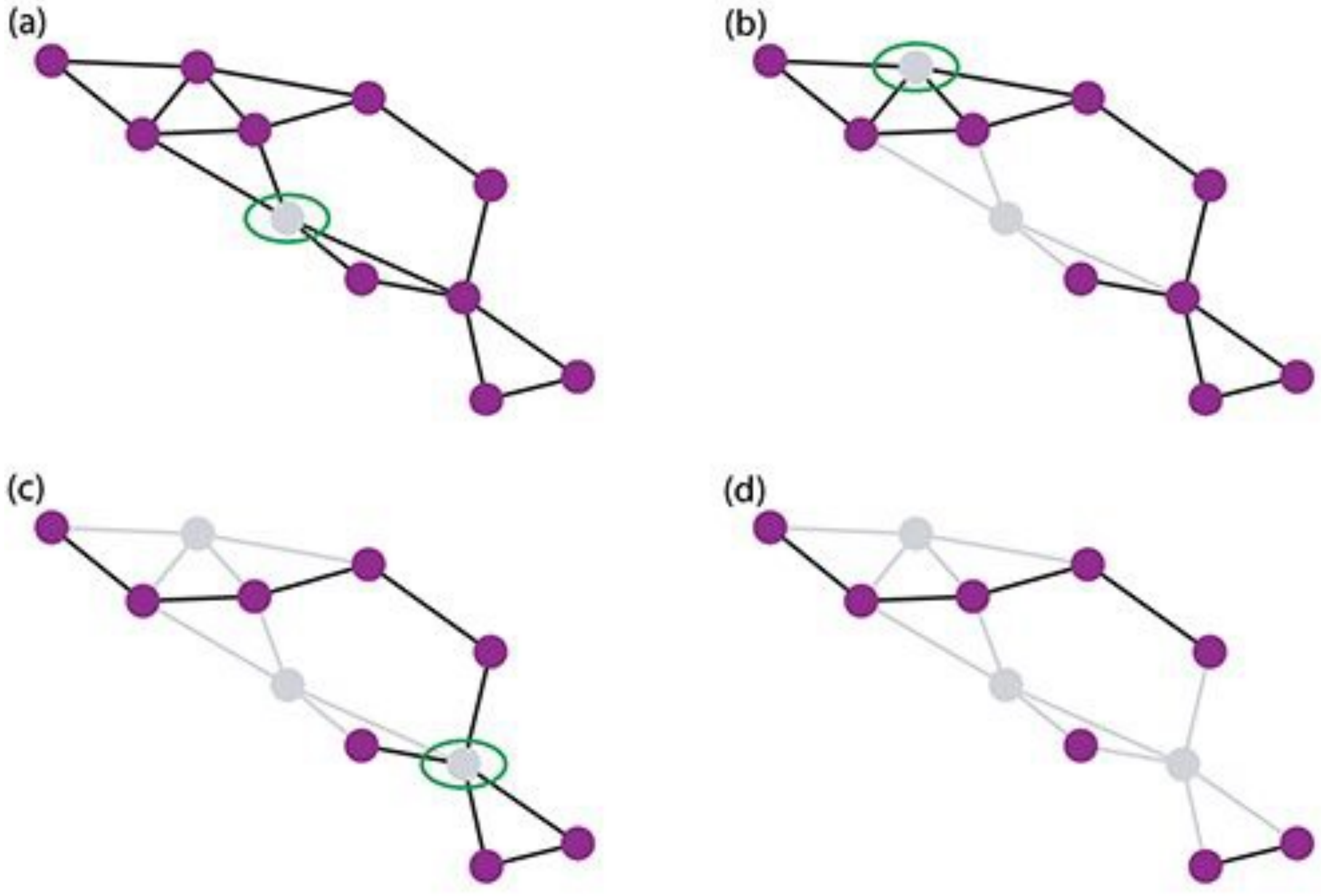
Ultra mali svijet



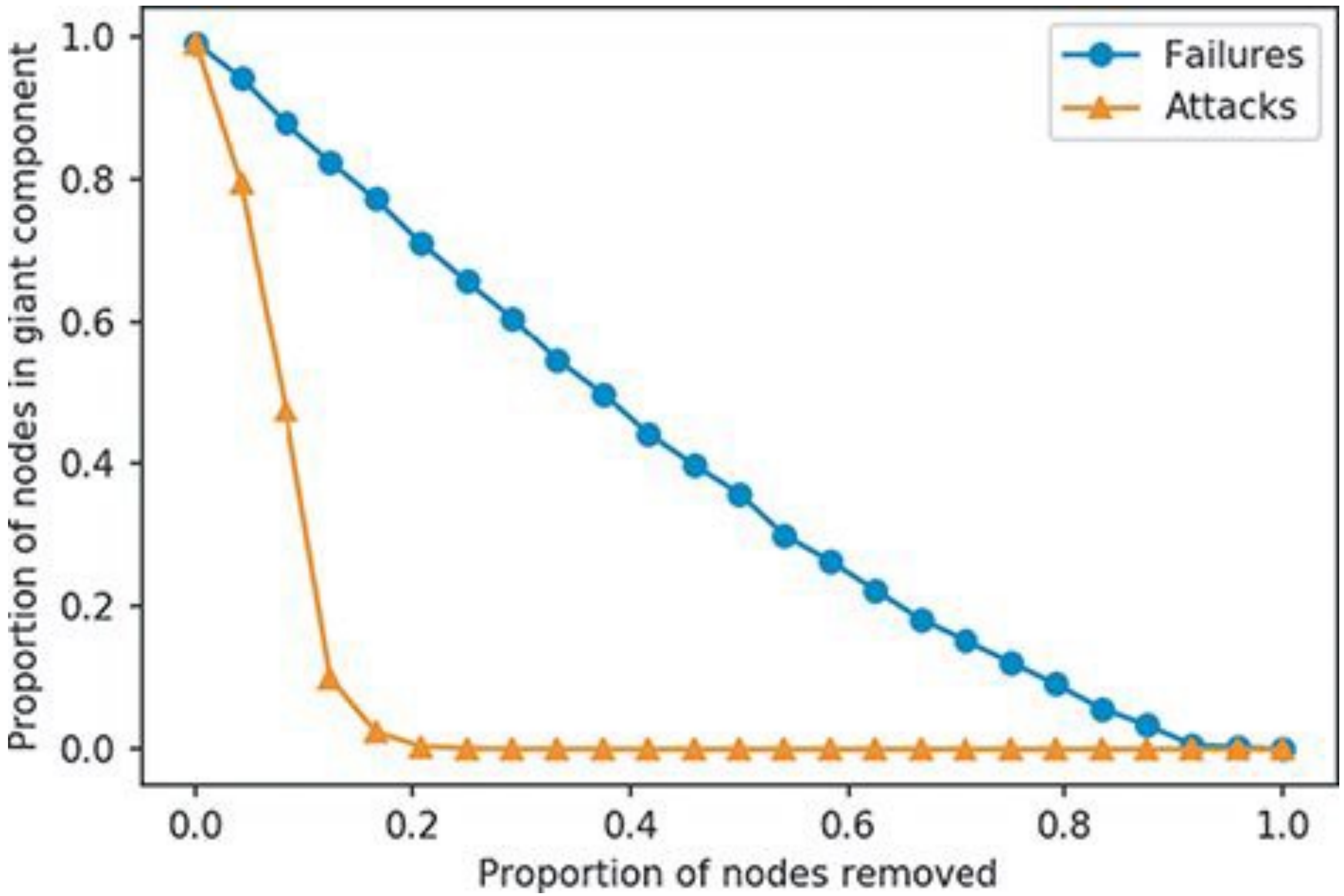
Robusnost

- Kvar na jednoj komponenti ne utječe na funkcionalnost (npr. Avionski motori)
- Definicija robusnosti mreže
  - Povezanost je važno svojstvo
  - Mjerimo utjecaj micanja čvora i njegovih veza na povezanost
  - Slom mreže u nepovezane dijelove signalizira štetu koja utječe na funkcionalnost

Robusnost



Robusnost



Mjerenje broj čvorova u najvećoj komponenti u odnosu na početnu mrežu

Kvar – slučajno uklanjanje
Napad – ciljano uklanjanje (prvo hubovi)

Otpornost na kvar i ranjivost na napad

Dekompozicija jezgre mreže

- Jezgra i periferija mreže
- Razdvajanje mreže u međusobno isključive dijelove (ljuske)
  - Koristimo stupanj
  - Zavisno o poziciji u jezgra-periferija strukturi mreže
  - Vanjske ljuske niskog stupnja su periferija
  - Ljuštimo jednu po jednu – krećemo sa čvorovima sa stupnjem 0
  - Na kraju ostaje jezgra

K-jezgrena dekompozicija

- Počinjemo s k=0
- Iterativno:
  1. Rekurzivno uklanjanje čvorova stupnja *k,* dok niti jedan ne preostane
  2. Uklonjeni čvorovi su *k-ljuska*, preostali čine *k+1 jezgru*
  3. Ukoliko nema više čvorova, završiti; inače, inkrementalno povećati *k*

Dekompozicija jezgre