

Analiza socijalnih mreža

Univerzitet u Beogradu, Elektrotehnički fakultet



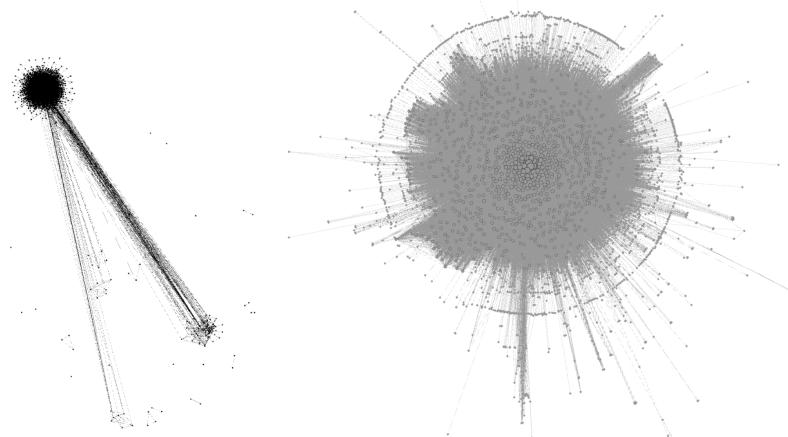
Bogdan Bebić 2022/3051

Marta Avramović 2022/3166

Subreddit Network (SNet)

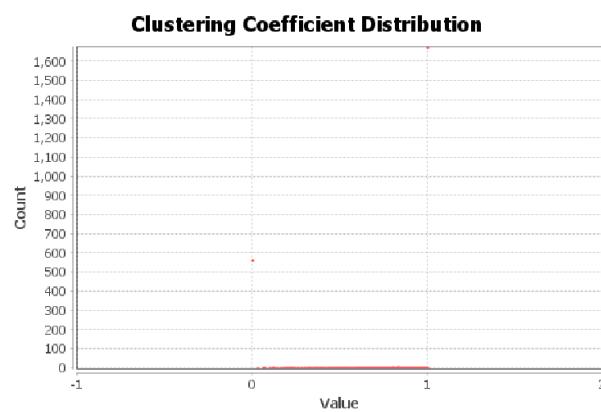
Mreža SNet (od engleskog subreddit network) se dobija na osnovu kompletnih podataka, sadrži sve sabredite i grane formirane na osnovu svih komentara i objava. Ovo je osnovna mreža koja će se koristiti za analizu sistema.

Različiti prikazi mreže (rasporedi *OpenOrd* levo, *Force Atlas 2* desno):

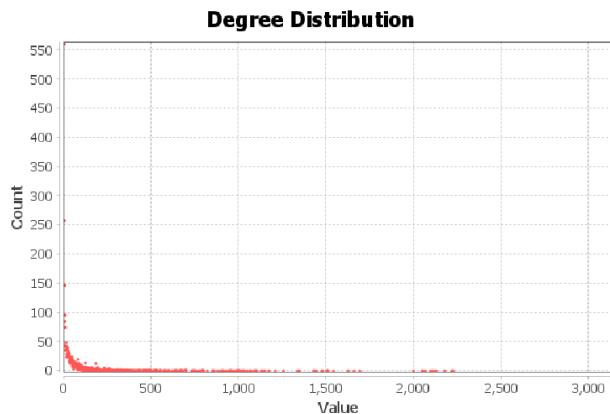


Karakteristike mreže:

- Gustina 0.025
- Prosečna distanca 2.099
- Dijametar 5
- 20 povezanih komponenti od kojih je jedna gigantska i sadrži skoro sve čvorove (98.85%)
- Prosečni koeficijent klasterizacije za odgovarajuću *Erdos-Renyi* mrežu 0.025
- Prosečni koeficijent klasterizacije 0.907 (klasterizacija je izražena), raspodela:



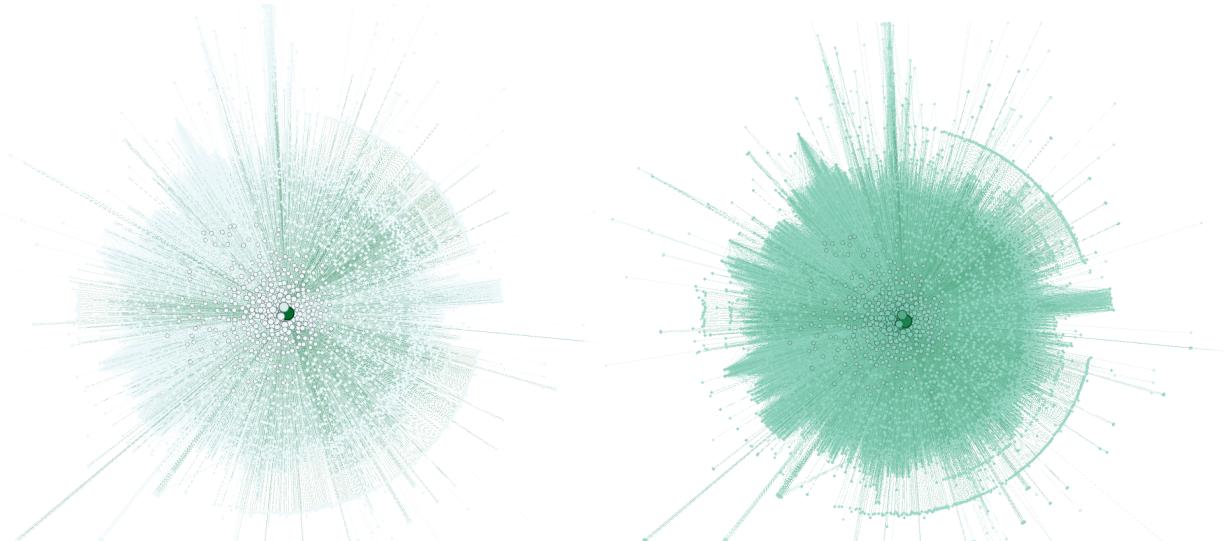
- Prosečni stepen cvora je 88.035, raspodela po stepenu cvorova:

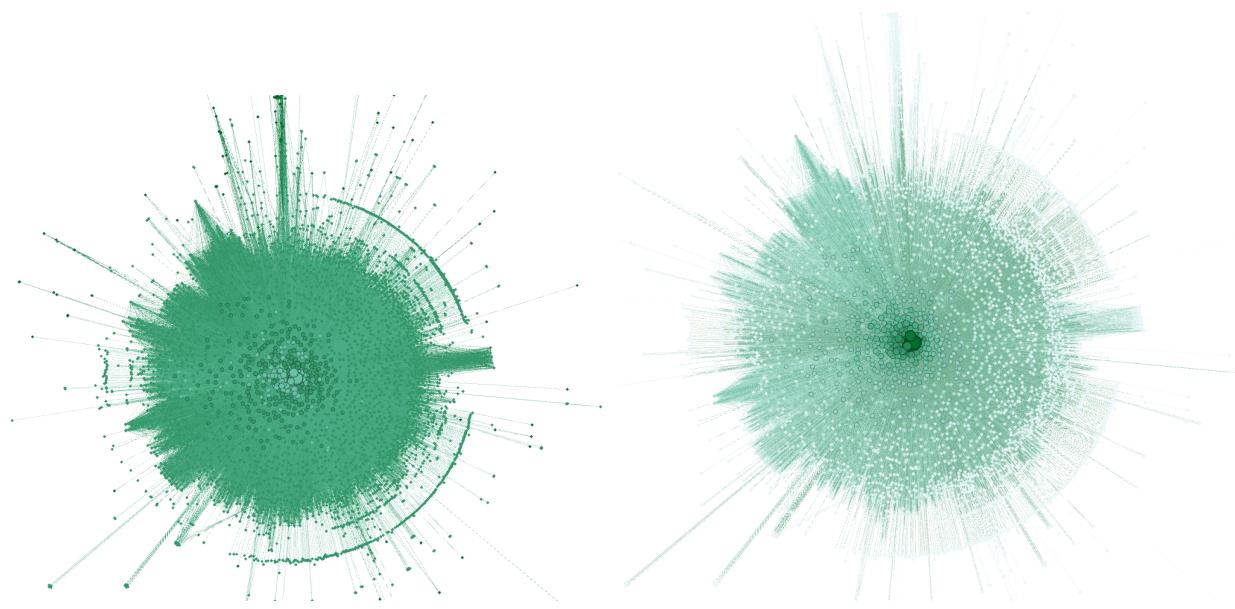


- Čvorovi sa najvećim stepenom su u sredini mreže

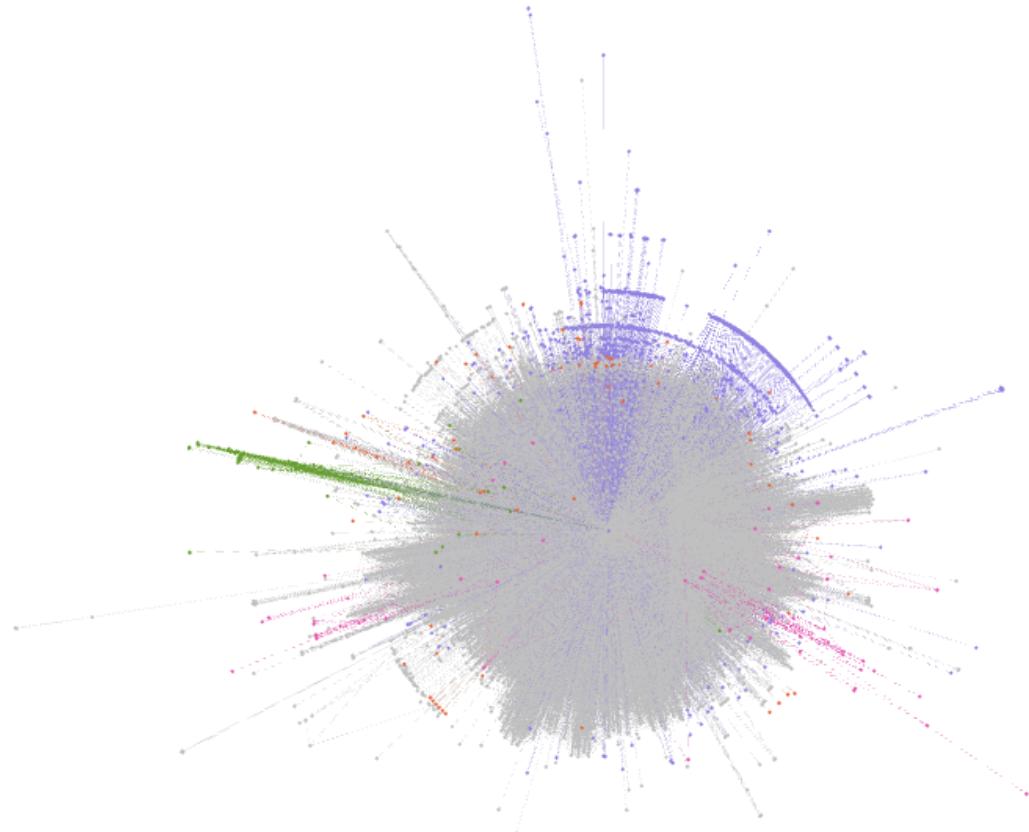
Po svakoj od analiziranih mera centralnosti mreže dobijamo slične rezultate (samim tim, heuristika za analizu centralnosti nije previše bitna jer su rezultati slični, pa možemo npr. pomnožiti dobijene mere). Najvažniji habovi i autoriteti u mreži su ujedno i u centru mreže, jer se po svim merama centralnosti dobiju isti čvorovi kao najvažniji.

Ispod su prikazani redom: *Betweenness Centrality*, *Closeness Centrality*, *Eccentricity*, *Eigenvector Centrality*.



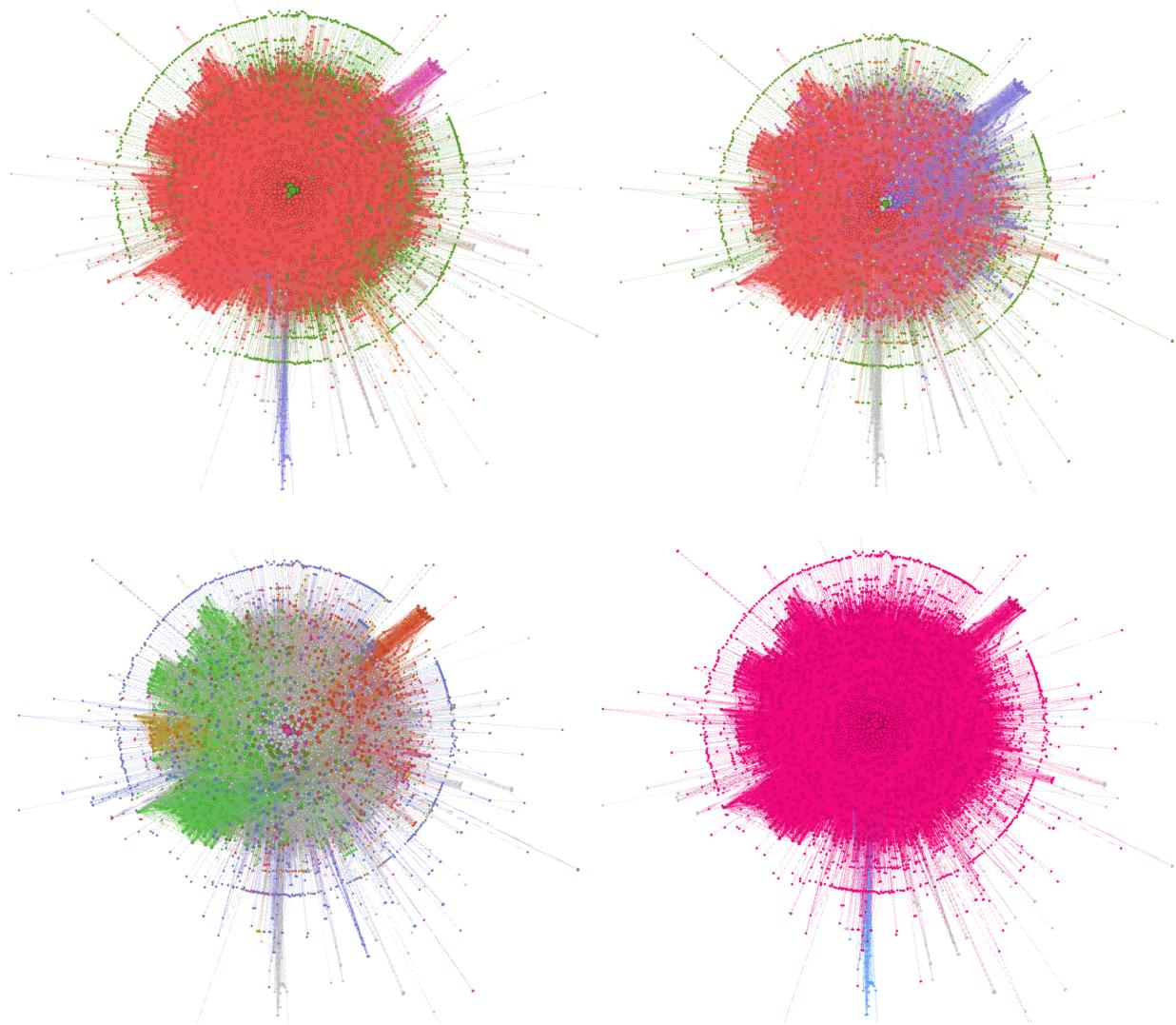


Gledajući *statistical inference clustering* u alatu *Gephi*, dobijamo puno malih komuna (sive) i tri najveće:



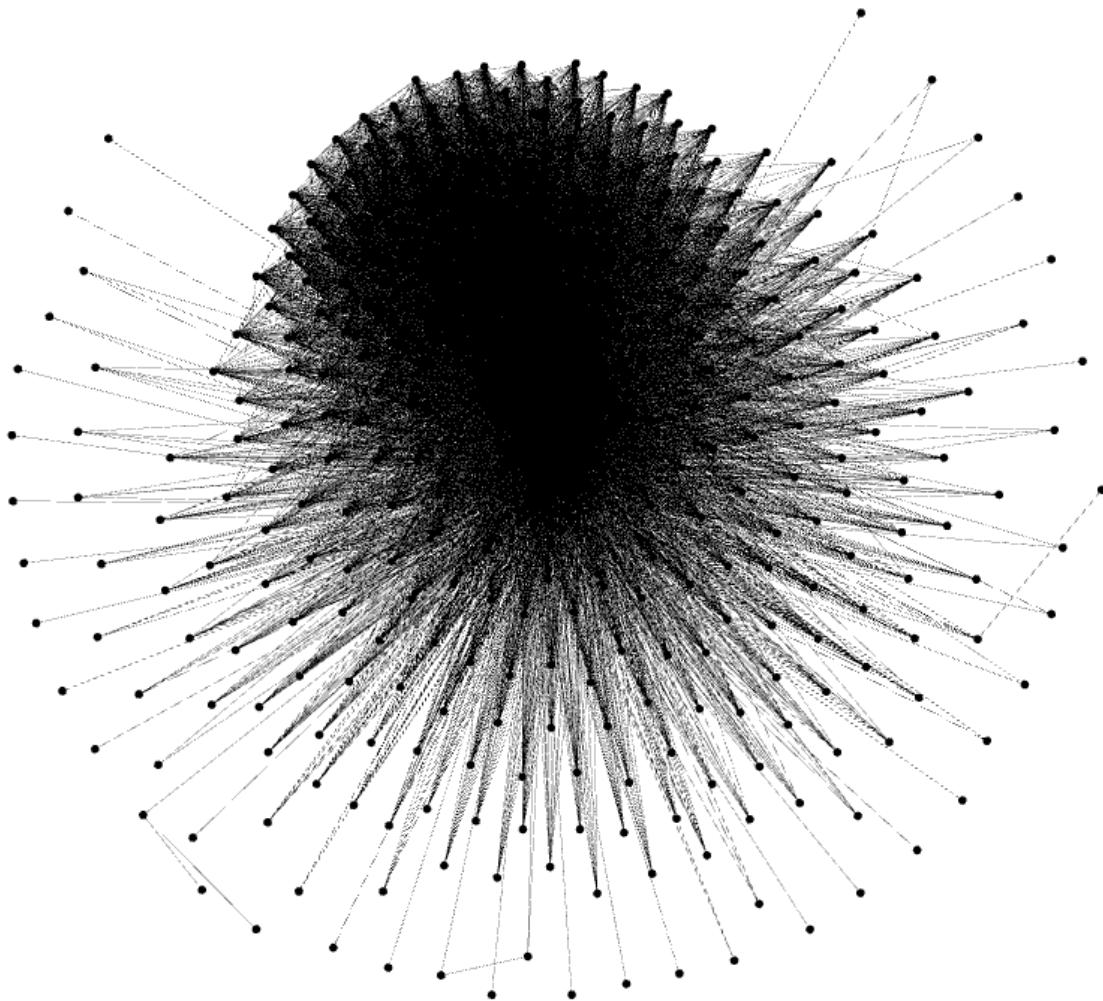
Kada posmatramo analizu komuna u ovoj mreži, dobijamo sledeće rezultate:

- Za rezoluciju 1 se ističu dve najveće komune (prva slika)
- Za rezoluciju 0.7 se ističu tri najveće komune koje su slične prethodnim komunama za rezoluciju 1 (druga slika)
- Za još manje rezolucije se dobija više komuna koje nema previše smisla analizirati jer su dosta male (treća slika)
- Za rezoluciju 1.3 se dobija jedna gigantska komuna i puno malih (četvrta slika)
- Ključni brokeri se nalaze u samom jezgru mreže (predstavljeni većim čvorovima), kroz njih prolazi najviše najkraćih puteva u mreži:



Subreddit Network - Filtered (SNetF)

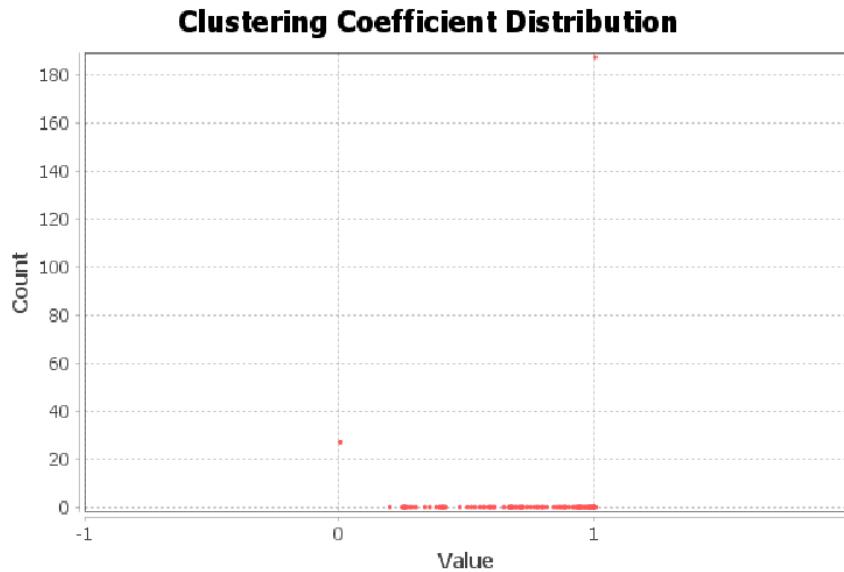
Izvršeno je filtriranje grana male težine i na taj način dobijena filtrirana mreža sabredita SNetF (od engleskog filtered subreddit network). Kao prag za filter je odabrana prosečna težina grana u grafu SNet. Mreža je prikazana kroz *Fruchterman Reingold* raspored:



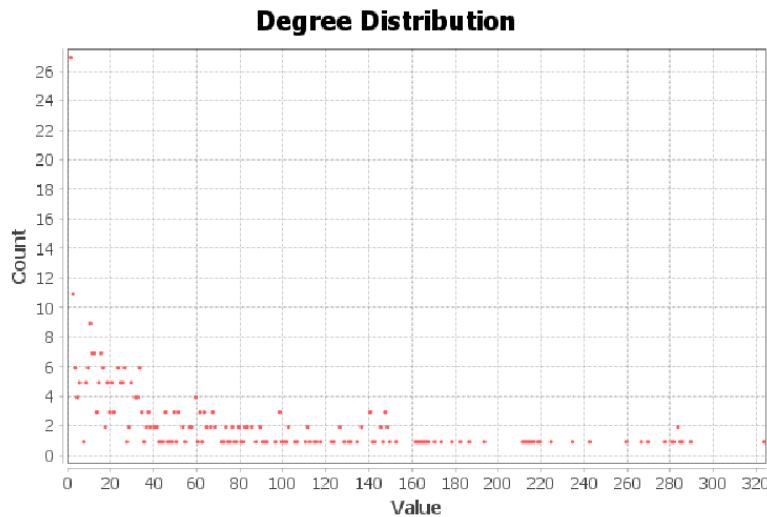
Karakteristike mreže:

- Gustina 0.186
- Prosečna distanca 1.84
- Dijametar 4
- Jedna povezana komponenta koja sadrži sve čvorove
- Prosečni koeficijent klasterizacije za odgovarajuću Erdos-Renyi mrežu 0.183

- Prosečni koeficijent klasterizacije 0.894 (klasterizacija je izražena), raspodela:



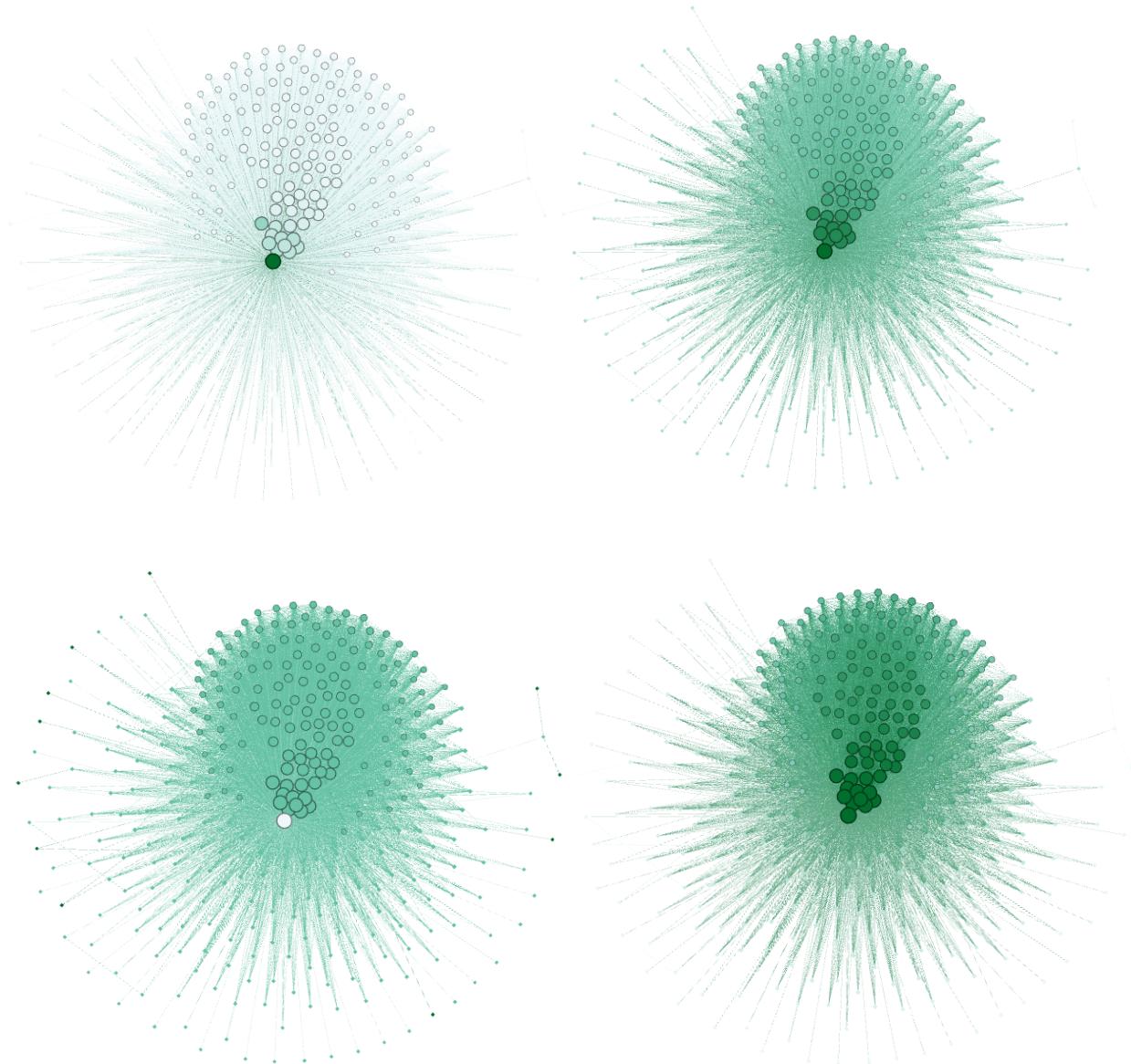
- Kako je koeficijent klasterizacije veliki, a prosečna distanca veća od 1, ova mreža ispoljava osobine malog sveta
- Raspodela po stepenu čvorova - ne prati *power law* raspodelu:



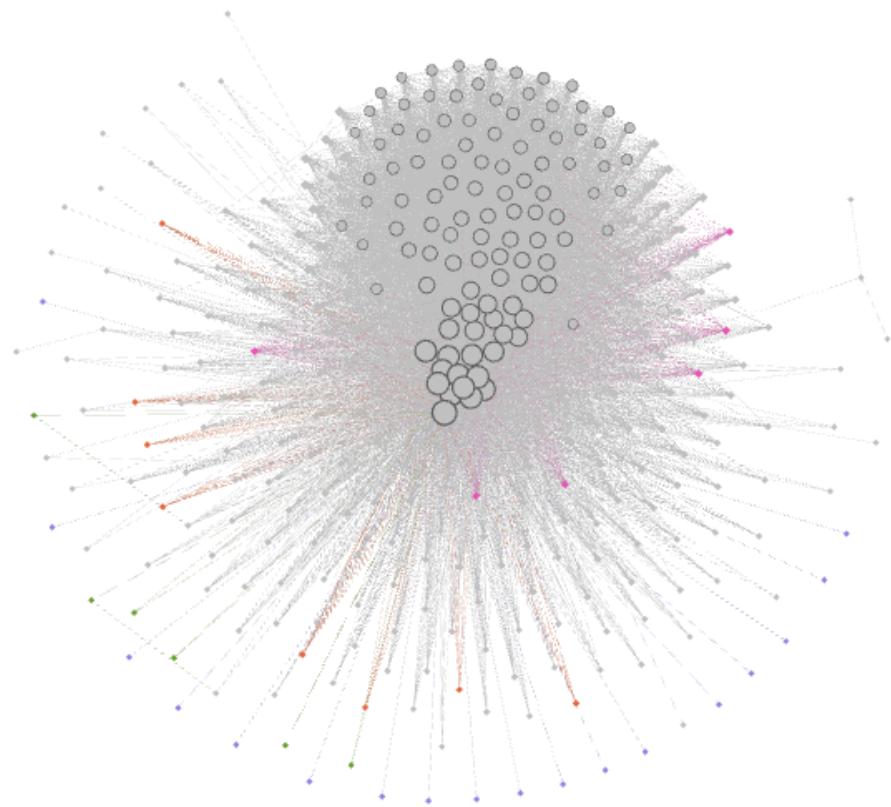
- Čvorovi sa najvećim stepenom su u sredini mreže

Po svakoj od analiziranih mera centralnosti dobijamo slične rezultate (samim tim, heuristika za analizu centralnosti nije previše bitna jer su rezultati slični, pa možemo npr. pomnožiti dobijene mere). Najvažniji habovi i autoriteti u mreži su ujedno i u centru mreže, jer se po svim merama centralnosti dobiju isti čvorovi kao najvažniji.

Ispod su prikazani redom: *Betweenness Centrality*, *Closeness Centrality*, *Eccentricity*, *Eigenvector Centrality*.

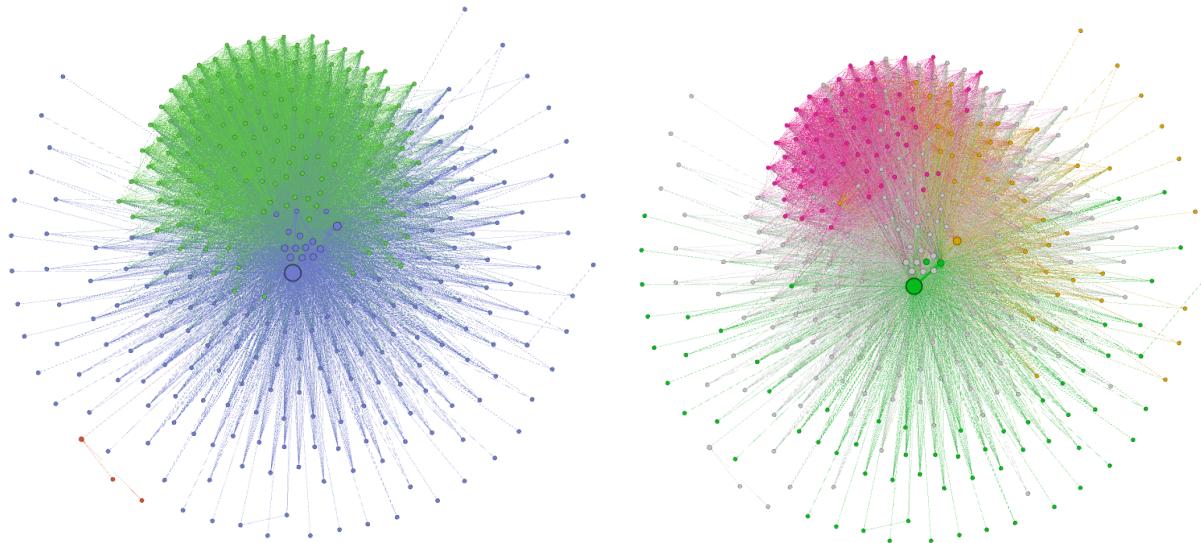


Gledajući *statistical inference clustering* u alatu *Gephi*, dobijamo puno malih komuna (sive) i četiri najveće:



Kada posmatramo analizu komuna u ovoj mreži, dobijamo sledeće rezultate:

- Za rezoluciju 1 se ističu dve najveće komune (prva slika)
- Za rezoluciju 0.5 se ističu tri najveće komune koje su slične prethodnim komunama za rezoluciju 1 (druga slika)
- Za još manje rezolucije se dobija više komuna koje nema previše smisla analizirati jer su dosta male
- Ključni brokeri se nalaze u samom jezgru mreže (predstavljeni većim čvorovima), kroz njih prolazi najviše najkraćih puteva u mreži



Najveći čvor je čvor koji predstavlja sabredit *reddit.com*, posle njega sabrediti *programming, funny, technology, WTF...*

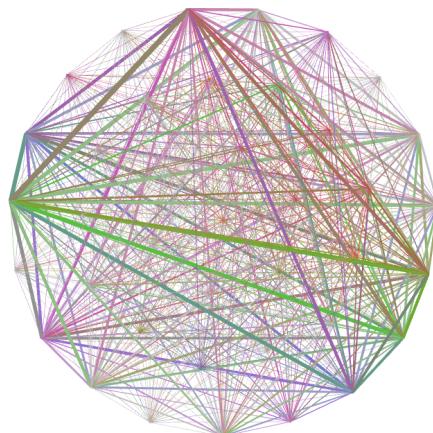
Subreddit Network - Targeted (SNetT)

Mreža SNetT (od engleskog targeted subreddit network) se sastoji od podataka o ciljanoj grupi sabredita čija je tematika bliska temi ekonomske krize. Konstruisana je kao indukovani podgraf SNet mreže čiji su čvorovi sabrediti u skupu: {reddit.com, pics, worldnews, programming, business, politics, obama, science, technology, WTF, AskReddit, netsec, philosophy, videos, offbeat, funny, entertainment, linux, geek, gaming, comics, gadgets, nsfw, news, environment, atheism, canada, math, Economics, scifi, bestof, cogsci, joel, Health, guns, photography, software, history, ideas}.

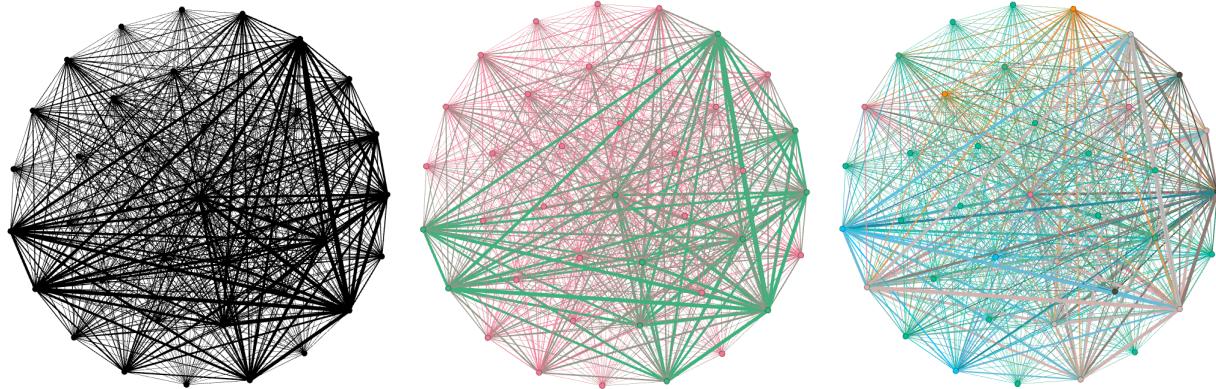
Karakteristike mreže:

- Gustina 1.0
- Prosečna distanca 1
- Dijametar 1
- Jedna povezana komponenta koja sadrži sve čvorove
- Koeficijent klasterizacije 1.0, svi čvorovi imaju koeficijent klasterizacije 1.0
- Svi čvorovi su međusobno povezani
 - Odgovarajuća *Erdos-Renyi* ima isti koeficijent klasterizacije
 - Mreža ima osobine mreže malog sveta u kom su svi čvorovi
 - Mreža ima klub bogatih u kom su svi čvorovi
 - Svi čvorovi imaju isti stepen, distribucija ne prati *power law* raspodelu
 - Mere centralnosti nemaju smisla jer su svi čvorovi povezani podjednako
 - Najvažniji habovi i autoriteti ne postoje jer je svaki čvor podjednako bitan

Gledajući *statistical inference clustering* u alatu *Gephi*, dobijamo 16 malih komuna:



Postoje dve najveće komune koje se okupljaju oko sabredita vezanih za ekonomiju (i krizu) i oko tema vezanih za programiranje. Možemo dobiti i više komuna ako na taj način parametrujemo algoritam za njihovu detekciju.



Na slikama iznad su predstavljeni izgledi mreže, kao i obojene mreže sa detektovanim komunama. Gledajući sa leva na desno, dobijamo jednu komunu, dve komune, devet komuna (za rezoluciju 1.5, 1.0, 0.5).

Poređenje SNet i SNetT mreža

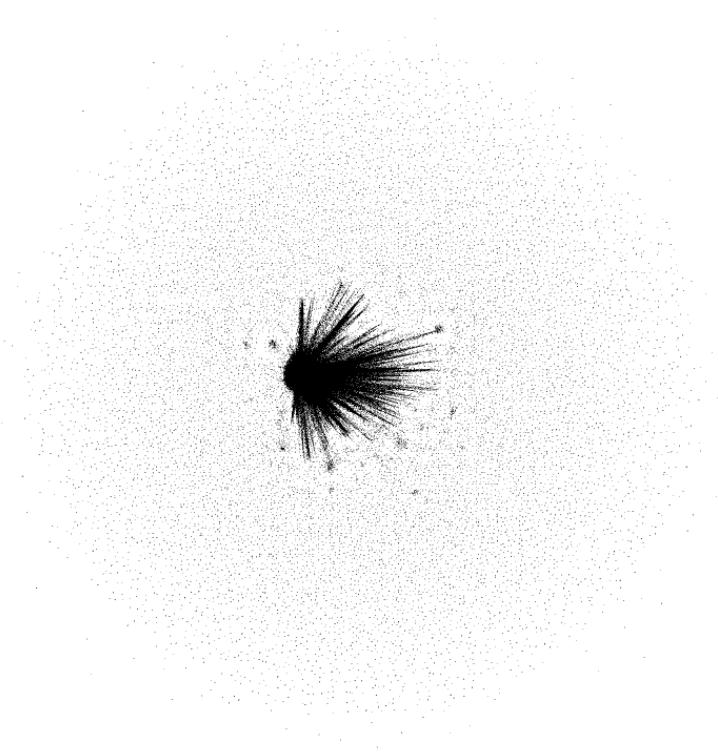
SNetT je odabrana kao targetirana mreža i pokazuje mnogo značajnije povezanosti i kraći put. SNet sadrži nepovezane čvorove, 20 povezanih komponenti i malu gustinu, dok je SNetT potpuno povezana mreža. Iz ovoga se zaključuje da su sabrediti odabrani za SNetT aktivniji i dosta bolje povezani od ostatka mreže SNet. Ako uzimamo u obzir povezanost i stepen čvorova kao meru centralnosti, čvorovi sadržani u SNetT se nalaze u jezgru mreže SNet - jako su dobro povezani sa ostatkom mreže.

User Network (UserNet)

Mreža korisnika modeluje interakcije između korisnika platforme Reddit. Vodi se evidencija o tome ko je komentarisao čiju objavu ili komentar, pri čemu se komentarisanje na objavu i na komentar smatraju istim tipom interakcije.

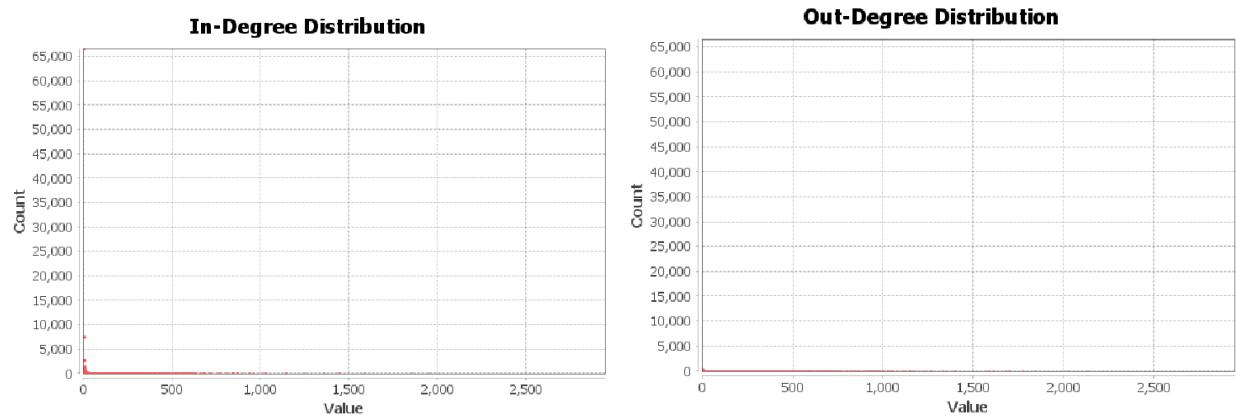
Mreža ima 90927 čvorova, pa mnoge operacije nad njom ne mogu da se izvrše u konačnom vremenu.

Predstava mreže *OpenOrd*:



Karakteristike mreže:

- Gustina 0 (zaokruženo na 0, postoje grane u mreži)
- Prosečna distanca 1.84
- Dijametar se ne može izračunati jer mreža nije povezana
- Zbog veličine mreže, ne može se izračunati prosečna putanja, pa se i dalje analize ne mogu raditi
- Prosečni stepen čvora je 5.678, distribucije po stepenu čvora:



- Broj detektovanih komuna je 12334 za rezoluciju 1.0, ističu se 4 najveće komune (26.39%, 22.79%, 20.32%, 13.63%), ostale su manje od 1.5%. Za rezoluciju 1.5, ističu se komuna od 72.59% i 11.5%, a sve ostale su manje od 0.2%.

