

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

ZAVRŠNI RAD br.243

**UTJECAJ ZARAZE NA SVOJSTVA
KOMPLEKSNE MREŽE**

Nino Antulov-Fantulin

Zagreb, lipanj 2008.

Zahvaljujem se dr.sc. Hrvoju Štefančiću, "Institut Ruđer Bošković" i mr.sc. Mili Šikiću, Fakultet elektrotehnike i računarstva za stručno savjetovanje, uspješnu suradnju i njihovu potporu prilikom rada na ovom istraživanju.

Također hvala Alenu Lančiću, studentu treće godine Prirodoslovno-matematičkog fakulteta na stručnim savjetima iz matematike prilikom istraživanja.

Hvala mojoj obitelji Antulov-Fantulin.

Sadržaj

Uvod	1
1. Teorijska podloga	2
1.1. Kompleksne mreže	3
1.1.1. "Mreže bez skale"	4
1.1.2. "Mreže malog svijeta"	4
1.2. Svojstva i mjere centralnosti kompleksnih mreža	5
1.2.1. Distribucija stupnjeva	5
1.2.2. Grupiranje	6
1.2.3. Međupoloženost	7
1.2.4. Asortativnost - korelacije čvorova	8
1.3. Propagacija virusa u kompleksnim mrežama	9
1.3.1. Klasični model epidemiologije – homogene mreže	9
1.3.2. SIR model	9
1.3.3. Propagacija kroz kompleksne mreže	12
2. Podaci	13
2.1. Svojstva promatrane mreže	14
3. Metode	17
3.1. "Complex Networks Package"	20
3.2. "MatlabBGL"	21
3.3. "Matlab Distributed Computing Toolbox"	22
4. Rezultati	23
4.1. Bimodalni proces:	24
4.2. Parametarski prostor (p - q - d)	26
4.2.1. Promjena vjerojatnosti zaraze p	26
4.2.2. Promjena vjerojatnosti oporavka q	35
4.2.3. Promjena vjerojatnosti izbacivanja d	36
Zaključak	39
Literatura	40
Sažetak	42

Uvod

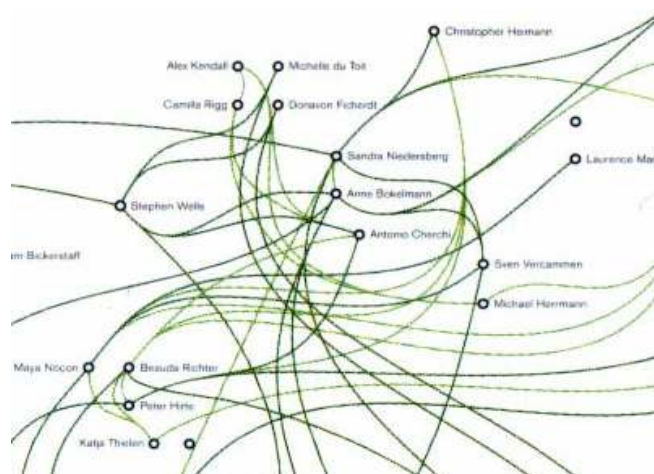
Kompleksne mreže predstavljaju jedinstvenu interdisciplinarnu temu kojom se danas bave fizičari, matematičari, biolozi, sociolozi, informatičari. Svojstva i mjere centralnosti kompleksnih mreža su do sada dobrim dijelom istražena kao i samo područje računalne epidemiologije. No postoji mnogo otvorenih pitanja i tema u području kompleksnih mreža koji još nisu detaljno istraženi. Jedna takva tema je "utjecaj širenja zaraze na svojstva kompleksne mreže".

Cilj ovog rada utvrditi strukturalne promjene kompleksnih mreža uslijed epidemije. Potrebno je računalnim simulacijama modelirati dinamiku procesa epidemije na kompleksnim mrežama i utvrditi njen utjecaj na strukturalne promjene a samim time svojstva i funkciju mreže poslije epidemije.

Poglavlje 1. objašnjava osnovne teorijske koncepte koji su neophodni za razumijevanje i analizu svih simulacija širenja zaraze na kompleksnoj mreži. Osnovne definicije kompleksnih mreža su dane u poglavlju 1.1. Teorijska razmatranja u poglavlju 1.2. uključuju objašnjenje osnovnih svojstava i mjera centralnosti kompleksnih mreža. Poglavlje 1.3. pokazuje osnovne koncepte epidemiologije i SIR modela širenja zaraze. Poglavlje 2. nam opisuje ulazne podatke tj. odabranu mrežu nad kojom su vršene simulacije. Poglavlje 3. nam objašnjava osnovne metode kojima su detektirane promjene svojstava polazne mreže nakon epidemije. Poglavlje 4. i 5. nam predočavaju postignute rezultate simulacija, analizu, potvrde početnih hipoteza i samu diskusiju rezultata.

1. Teorijska podloga

Istraživanje **kompleksnih sustava** i njihovih svojstava zadnjih nekoliko godina zaokuplja pažnju mnogih znanstvenika. Postoji nekoliko definicija što su to kompleksni sustavi, no najčešće se opisuje kao znanost koja proučava kako veze između pojedinih dijelova utječu na kolektivno ponašanje sustava te kako sustav ulazi u interakciju i stvara veze s okolinom.



Slika 1. Vizualizacija jedne kompleksne mreže

Postoji više definicija kompleksnosti, stoga se mnogi prirodni umjetni i apstraktni objekti odnosno mreže mogu smatrati kompleksnim sustavima, a njihovo proučavanje (engl. *complexity science*) je interdisciplinarno. Primjeri kompleksnih sustava su mnogobrojni i uključuju zajednice mrava, ekonomiju, klimu, živčani sustav, stanicu i živa bića uključujući i ljude, kao i moderne energetske ili telekomunikacijske infrastrukture.

Osim činjenice da je struktura svih ovih primjera neka vrsta mreže i da su kompleksni, na prvi pogled se može činiti da imaju malo toga zajedničkog. Bilo kako bilo vjeruje se da mnogi kompleksni sustavi posjeduju ista obilježja ponašanja i istu strukturu, što ih do određene granice ujedinjuje u jedan opći fenomen. Oni se također mogu ujediniti i teoretski jer se svi sustavi mogu u teoriji modelirati nekom matematikom s različitim stupnjem uspješnosti. Zbog navedenog je moguće jasno definirati zajednička svojstva ovih sustava.

1.1. Kompleksne mreže

Statistički opis sustava koji se mogu predstaviti strukturom kompleksne mreže daje uvid u bitna svojstva tih sustava, kao što su: njihova otpornost na slučajne greške (kvarove), osjetljivost na namjerne napade (npr. terorističke), efikasnost njihovog pretraživanja ili podložnost širenju epidemija (npr. računalnih virusa na Internetu ili SARS u ljudskom društvu). Kompleksne mreže predstavljaju jedinstvenu interdisciplinarnu temu kojom se danas bave fizičari, matematičari, biolozi, sociolozi, informatičari. Područje kompleksnih mreža nastoji dati odgovor na pitanje što je zajedničko tako raznorodnim sistemima poput Interneta, www-a, ljudskih socijalnih kontakata, prehrambenih lanaca u ekološkim studijama, interakcije proteina i mnogih drugih, proučavajući njihovu mrežnu strukturu.

U kontekstu „teorije grafova” termin „kompleksne mreže” odnosi se na graf ili mrežu koja ima netrivialna topološka svojstva koja se ne pojavljuju u jednostavnim grafovima. Struktura im je kompleksnija nego u klasičnim slučajnim mrežama (Erdos-Renyi).

Većina socioloških, bioloških i tehnoloških mreža mogu se smatrati kompleksnima s obzirom na svojstvo netrivialne topološke strukture. Ta netrivialna svojstva su npr. distribucije s debelim repom (engl. *fat-tailed degree distribution*), visoki koeficijent grupiranja, korelacija stupnjeva, zajednice (engl. *community*) i hijerarhijska struktura. Suprotno tome, jednostavne mreže nemaju niti jednu od ovih svojstava, i tipično su reprezentirani mrežom kao rešetka (engl. *Lattice*) ili slučajna mreža koja pokazuje veliku sličnost na bilo kojem promatranom djelu mreže [1].

1.1.1. "Mreže bez skale"

Najvažnija karakteristika "mreža bez skale" (engl. *Scale-free network*) je njihova distribucija po zakonu potencije sljedećeg oblika: $P(k) = k^{-\gamma}$ gdje parametar γ predstavlja eksponent distribucije. Takve mreže nemaju prirodnu skalu zbog čega su dobile takav naziv. U "mrežama bez skale" neki čvorovi imaju stupanj za nekoliko redova veličine veći od prosječnog stupnja čvorova u mreži i oni se nazivaju hubovi. "Mreže bez skale" su iznimno robusne na slučajne napade na čvorove (slučajno odstranjivanje čvorova), koje ne utječe znatno na fragmentaciju gigantske komponente. No iste mreže su osjetljive na ciljane napade na čvorove koji veoma lako fragmentiraju mrežu. Čvorovi sa visokim stupnjem su od iznimne važnosti u mnogim procesima na mreži (procesu fragmentacije, širenja zaraze ili informacija) [1].

1.1.2. "Mreže malog svijeta"

Fenomen "malog svijeta" (engl. *Small-world phenomenon*) poznat je već neko vrijeme iz stvarnog svijeta. Najpoznatiji primjer je slavni pokus Stanleya Milgrama iz 1960-tih [2] i njegova demonstracija da su svaka dva čovjeka na svijetu povezana u prosjeku preko najviše šest drugih ljudi. Formalno, efekt "malog svijeta" zahtjeva preciznu definiciju prosječne najkraće udaljenosti između dva čvora mreže.

Očito prosječni broj n najbližih susjeda jednog čvora raste usporedivo sa $\langle k \rangle^n$. Znači prosječan najkraći put \bar{l} ugrubo se aproksimira relacijom $\langle k \rangle^{\bar{l}} : N$. Stoga vrijedi :

$$(1) \quad \bar{l} \approx \frac{\ln N}{\ln \langle k \rangle}$$

Uglavnom, ako $\bar{l}(N)$ raste sporije nego ijedna pozitivna potencija od N za mrežu možemo reći da ima efekt "malog svijeta". [3]

1.2. Svojstva i mjere centralnosti kompleksnih mreža

Teorija grafova i kompleksnih mreža nudi mnoge mjere centralnosti vrhova unutar grafa kako bi uočili relativnu važnost svakog vrha u grafu (npr. koliko je osoba bitna u sociološkoj mreži).

Postoje četiri glavne mjere centralnosti u analizi mreža: stupanj čvora, međupoloženost, bliskost i centralnost prema svojstvenim vektorima. Ovdje su objašnjene sljedeće mjere: stupanj i međupoloženost čvora.

Osim samih mjera centralnosti objašnjene su i neke druge bitne mjere korelacije čvorova u mreži: grupiranje čvorova (engl. *clustering*) i asortativnost (engl. *assortativity*) koje nam mogu dobro opisati svojstva same mreže.

1.2.1. Distribucija stupnjeva

Jedna od najvažnijih karakteristika mreže je njena distribucija stupnjeva (engl. *degree distribution*). Nama najvažniji oblici distribucije su Poissonova, eksponencijalna i polinomna distribucija.

Za svaki čvor možemo definirati njegovu distribuciju stupnjeva $p(k, s, N)$. To je vjerojatnost da će čvor s u mreži veličine N čvorova imati stupanj k . Poznavajući distribuciju svakog stupnja možemo naći ukupnu distribuciju stupnjeva mreže:

$$(2) \quad P(k, N) = \frac{1}{N} \sum_{s=1}^N p(k, s, N)$$

Pretpostavimo li da su svi čvorovi statistički ekvivalentni, kao što je slučaj u klasičnim slučajnim mrežama, onda svaki od njih ima jednaku distribuciju stupnjeva $P(k, N)$. Prvi moment distribucije je zapravo prosječni stupanj cijele mreže $\bar{k} = \sum_k kP(k)$. Moguće je definirati ulazne i izlazne distribucije stupnjeva u usmjerenim mrežama. [4]

1.2.2. Grupiranje

Koeficijent grupiranja (engl. *clustering coefficient*) je mjera koliko je neki čvor zajedno sa svojim susjedima blizak potpunom grafu.

Watts i Strogatz [5] definiraju koeficijent grupiranja C kao omjer između svih y veza koje spajaju najbližih z susjeda čvora i svih mogućih veza koje bi ih mogle spajati. Ista mjera može poslužiti za utvrđivanje pripadnosti mreži malog svijeta. Mjera grupiranja nam kaže koliko "gusto" su povezani prvi susjedi nekog čvora. Dobiveni omjer kod neusmjerene mreže je:

$$(3) \quad C_i = \frac{2y}{z(z-1)}$$

Koeficijent grupiranja je direktno vezan za broj trokuta (ciklusa duljine tri) u mreži. Ujedno je i vrlo dobar pokazatelj postoje li korelacije između čvorova mreži. Neka je $\lambda_G(v)$ broj trokuta u neusmjerenom grafu G , $v \in V(G)$. Drugim riječima $\lambda_G(v)$ je broj podgrafova grafa G sa 3 vrha i 3 brida, od kojih je jedan v . Definiramo $\tau_G(v)$ kao broj tripleta na $v \in V(G)$, točnije broj podgrafova sa 2 brida i 3 vrha, gdje je jedan vrh v incidentan sa ova dva brida. Stoga se koeficijent grupiranja može definirati i ovako:

$$(4) \quad C_i = \frac{\lambda_G(v)}{\tau_G(v)}$$

Koeficijent grupiranja za čitavu mrežu definiramo kao srednju vrijednost svih koeficijenta grupiranja svakog čvora.

$$(5) \quad \bar{C} = \frac{1}{N} \sum_{i=1}^N C_i$$

Mreža se smatra mrežom malog svijeta ako je srednji koeficijent grupiranja \bar{C} znatno veći nego kod slučajnih mreža konstruiranih na istom skupu vrhova i ako mreža ima kratki srednji najkraći put [6].

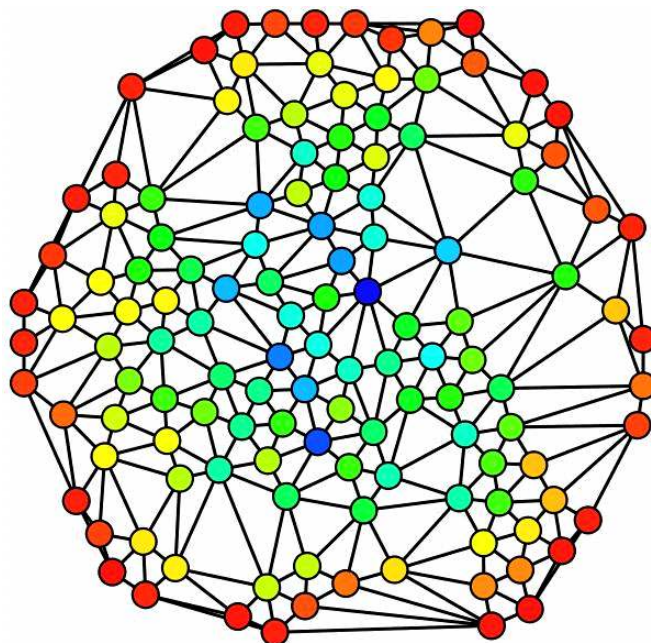
1.2.3. Međupoloženost

Međupoloženost (engl. *betweenness*) je mjera centralnosti čvora unutar mreže. Čvorovi koji leže na više najkraćih putova između drugih čvorova imaju veći koeficijent međupoloženosti. Neka postoji graf $G = (V, E)$ sa n vrhova, međupoloženost $C_B(v)$ za čvor v je definirana kao:

$$(6) \quad C_B(v) = \sum_{s \neq v \neq t \in V, s \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

gdje σ_{st} predstavlja broj najkraćih putova od vrha s do vrha t i $\sigma_{st}(v)$ predstavlja broj najkraćih putova od vrha s do vrha t koji prolaze vrhom v . Ovaj koeficijent međupoloženosti može biti normaliziran dijeljenjem s brojem parova vrhova koji ne uključuju vrh v , a to je $(n-1)(n-2)$ parova.

Računanje međupoloženosti za sve vrhove grafa uključuje pronalazak najkraćih putova između svih parova vrhova u grafu G . Složenost takvog algoritma je $O(V^2)$ ako se koristi Floyd-Warshall algoritam. Ako se koristi Johnsonov algoritam na matrici susjedstva koja nije gusta (engl. *sparse*) složenost je $O(V^2 \log V + VE)$ [7].



Slika 2. Grafička vizualizacija međupoloženosti, od međupoloženost(crvenih čvorova) = 0 pa do međupoloženost(plavih čvorova) = max

1.2.4. Asortativnost - korelacije čvorova

Mjera asortativnosti (engl *assortativity coefficient*) odnosi se na tendenciju spajanja čvorova na druge slične ili pak različite čvorove po nekom svojstvu. Korelacije između čvorova sličnog stupnja često se mogu pronaći u mnogim mrežama. U socijalnim mrežama čvorovi visokog stupnja imaju tendenciju spojiti se s drugim čvorovima visokog stupnja. Tu tendenciju spajanja čvorova često nazivamo "asortativno spajanje" (engl. *assortative mixing*). S druge strane tehnološke i biološke mreže tipično pokazuju svojstvo tendencije spajanja čvorova visokog stupnja s čvorovima niskog stupnja i nazivamo "neasortativno spajanje" (engl. *disassortative mixing*).

Koeficijent asortativnosti je Pearsonov korelacijski koeficijent između parova čvorova. Pozitivne vrijednosti od r ukazuju na korelaciju između čvorova sličnog stupnja dok negativne vrijednosti ukazuju na vezu između čvorova različitog stupnja. Koeficijent r se kreće u rasponu od -1 do 1 [8].

Mreža:	Veličina mreže:	Asortativnost r :
Kolaboracija fizičara	52 909	0.363
Kolaboracija biologa	1 520 251	0.127
Kolaboracija glumaca	449 913	0.208
Internet	10 697	-0.189
WWW	269 504	-0.065
Proteinske interakcije	2 115	-0.156
Slučajna mreža		0
Albert Barabasi		0

Tablica 1. Koeficijent asortativnosti u raznim mrežama

1.3. Propagacija virusa u kompleksnim mrežama

Kako se propagiraju virusi?

Pojavu propagacije virusa proučit ćemo pomoću dva modela:

1. Klasični pogled epidemiologije kroz odjeljke - homogene mreže
2. Propagacija kroz kompleksne mreže

1.3.1. Klasični model epidemiologije – homogene mreže

Populacija se sastoji od velikog broja različitih individua, od kojih su svi drugačiji. Da bismo modelirali razvoj epidemije u takvoj populaciji, ova različitost se mora svesti na nekoliko glavnih karakteristika koje su važne za zarazu koju promatramo. Kao temeljnu pretpostavku koristimo činjenicu da je susret bilo koje dvije jedinice jednako vjerojatan, tj. koristimo hipotezu homogenog miješanja jedinice populacije. Pretpostavlja se da je kontakt između zaraženih i podložnih zarazi (engl. *susceptible*) konstantan i neovisan o mogućoj heterogenosti u sustavu. Također jedinice mogu prelaziti samo između odjeljaka i nove jedinice ne mogu dolaziti.

Kao primjer za objašnjenje modela uzet ćemo obične dječje bolesti koje uzrokuju dugotrajnu imunizaciju. Prirodno je podijeliti populaciju u tri odjeljka (1) podložni zarazi, (2) zaraženi (engl. *infected*) i (3) oporavljeni (engl. *removed*).

1.3.2. SIR model

Standardna konvencija označava ta tri odjeljka S (podložni), I (zaražen), R (oporavljeni). Stoga se taj model naziva SIR model.

Oznake također predstavljaju broj ljudi u svakom odjeljku u nekom trenutku. Da bi označili da se ti brojevi mogu mijenjati tokom vremena čak i kada ukupna populacija ostaje konstantna tokom vremena, oznake predstavljaju funkcije vremena: $S(t)$, $I(t)$ i $R(t)$. Za određenu bolest u određenoj populaciji ove funkcije se mogu razraditi na način da predvide moguće pojave epidemije i dovedu ih pod kontrolu.

1.3.2.1 SIR model – dinamika

Model je dinamičan jer broj jedinki u svakom odjeljku može varirati tokom vremena kao posljedica ovisnosti tih funkcija o vremenu. Važnost ovog dinamičnog aspekta je najočitija u endemičnim bolestima s kratkim vremenom infekcije. Takve bolesti se obično pojavljuju u ciklusima epidemija zbog variranja broja podložnih osoba $S(t)$ tokom vremena. Za vrijeme epidemije, broj podložnih jedinki rapidno pada jer je ih puno postaje zaraženo i tako ulaze u I-odjeljak i R-odjeljak. Zaraza se ne može razbuktati sve dok broj podložnih jedinki ne poraste.



Slika 3. SIR model i parametri prijelaza

Između odjeljka **S** i odjeljka **I** postoji stopa prijelaza (λ). Između odjeljka **I** i odjeljka **R**

postoji stopa prijelaza (δ). Ako se vrijeme oporavka označi s D , onda je ($D = \frac{1}{\delta}$), jer se jedinka oporavi jednom nakon D jedinica vremena.

Matematički opis SIR model-a:

SIR model se može opisati sa sljedećim diferencijalnim jednačbama

$$(7) \frac{dS}{dt} = -\lambda SI \quad (8) \frac{dI}{dt} = \lambda SI - \delta I \quad (9) \frac{dR}{dt} = \delta I$$

Promatrani sustav je nelinearan. Stoga se kod odziva koriste numeričke metode rješavanja.

Označimo efektivni parametar širenja danog sustava s:

$$(10) \quad R_0 = \frac{\lambda}{\delta}$$

Dijeljenjem diferencijalne jednačbe (11) i (12) dobivamo (13)

$$(11) \quad \frac{dS}{dt} = -\lambda SI$$

$$(12) \quad \frac{dR}{dt} = \delta I$$

$$(13) \quad \frac{dS}{dR} = -\frac{\lambda}{\delta} S = -R_0 S / \int$$

$$\ln S(t) = -R_0 R(t) + C$$

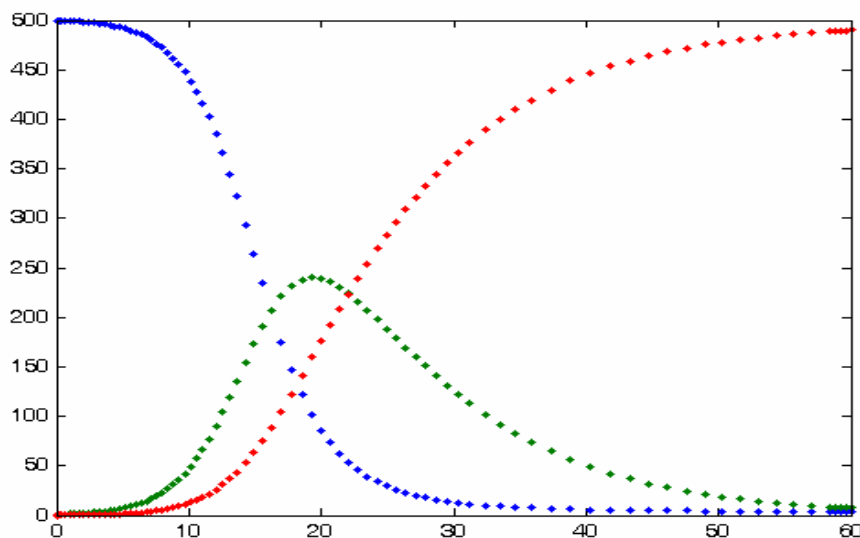
$$S(t) = e^{-R_0 R(t) + C_1} = e^{-R_0 R(t)} C \Rightarrow C = S(0) e^{R_0 R(0)}$$

$$(14) \quad S(t) = e^{-R_0 [R(t) - R(0)]} S(0)$$

U limesu kada vrijeme ide u beskonačno udio oporavljenih jedinki zadovoljava jednačbu:

$$(15) \quad R_\infty = 1 - S(0) e^{-R_0 (R_\infty - R(0))}$$

Razmatranjem ove jednačbe zaključujemo da na kraju epidemije nisu sve jedinke oporavljene, stoga neke moraju ostati podložne. To znači da je kraj epidemije uzrokovan padom broja zaraženih jedinki a ne nedostatkom podložnih [9].



Slika 4. Odziv sustava

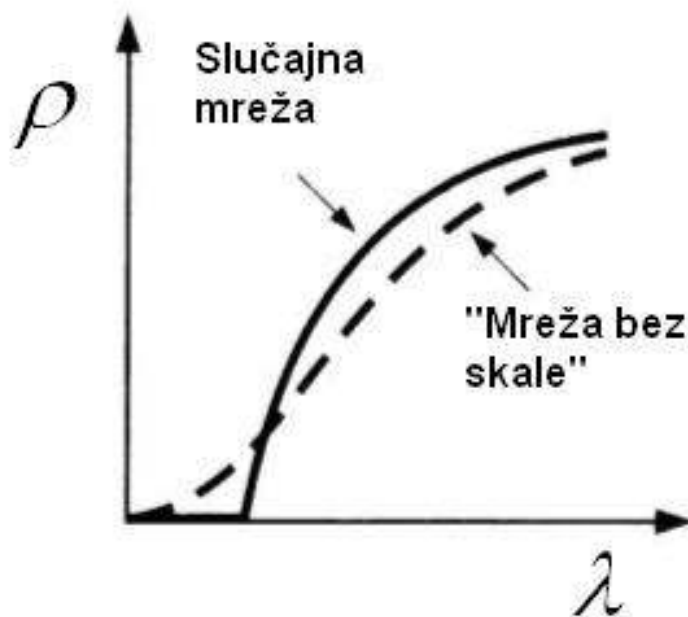
podložni S(t), zaraženi I(t) i oporavljeni R(t)

1.3.3. Propagacija kroz kompleksne mreže

Epidemiološki modeli su jako ovisni o modelu koji opisuje populaciju u kojoj se infektivni agenti šire. Dobar model je mreža, gdje svaka jedinka predstavlja čvor, a međusobne veze predstavljaju potencijalni kontakt preko kojeg se zaraza širi.

Interes promatranja epidemioloških modela u "mrežama bez skale" (engl. *Scale-free networks*) je ogroman, budući da su mreže ljudskih društvenih i seksualnih kontakata *mreže bez skale*. One predstavljaju vrlo zanimljiv slučaj jer imaju distribuciju stupnjeva sljedećeg oblika: $P(k) \approx k^{-\gamma}$.

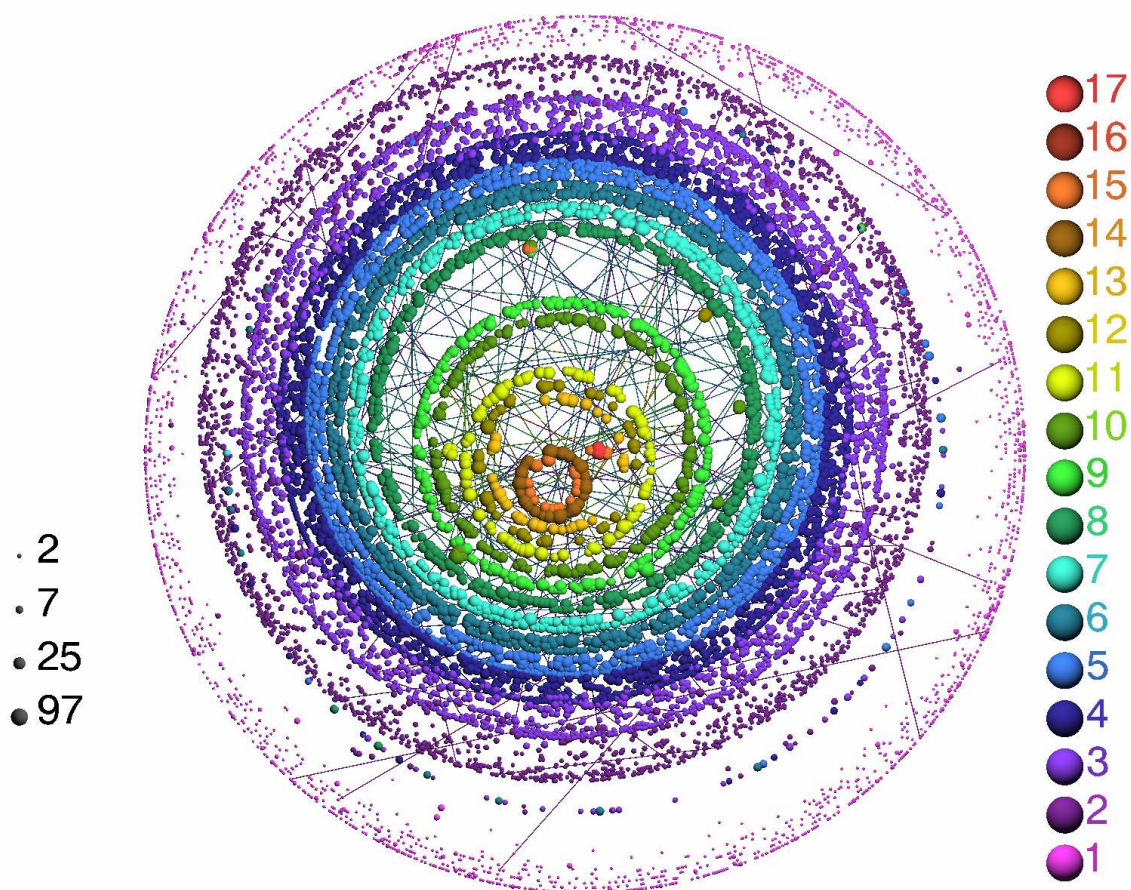
Satorras, R. i Vespignani, A. [10] su 2000. godine pokazali nepostojanje praga epidemije u *mrežama bez skale* točnije prag epidemije $\lambda_c = \frac{\bar{k}}{\langle k^2 \rangle}$ gdje je $\lambda_c = 0$.



Slika 5. Prikaz praga epidemije u slučajnoj mreži i "mreži bez skale"

2. Podaci

Za potrebe simulacija izabrana je jedna realna mreža kolaboracije znanstvenika "Condensed matter collaborations 2003" [11] koja je javno dostupna na internetu [12]. Ta mreža sastoji se od znanstvenika koji su objavljivali radove na "Condensed Matter E-Print Archive" lokaciji između 1. siječnja 1995. godine i 31. prosinca 1999. godine. Najveća komponenta ove mreže ima 27 519 čvorova. Mreža je zapisana u GML formatu [13]. Znanstvenici u toj mreži predstavljaju čvorove dok veza (neusmjerena i netežinska) između dva čvora postoji samo ako su ta dva pripadna znanstvenika bili koautori na barem jednom znanstvenom radu.

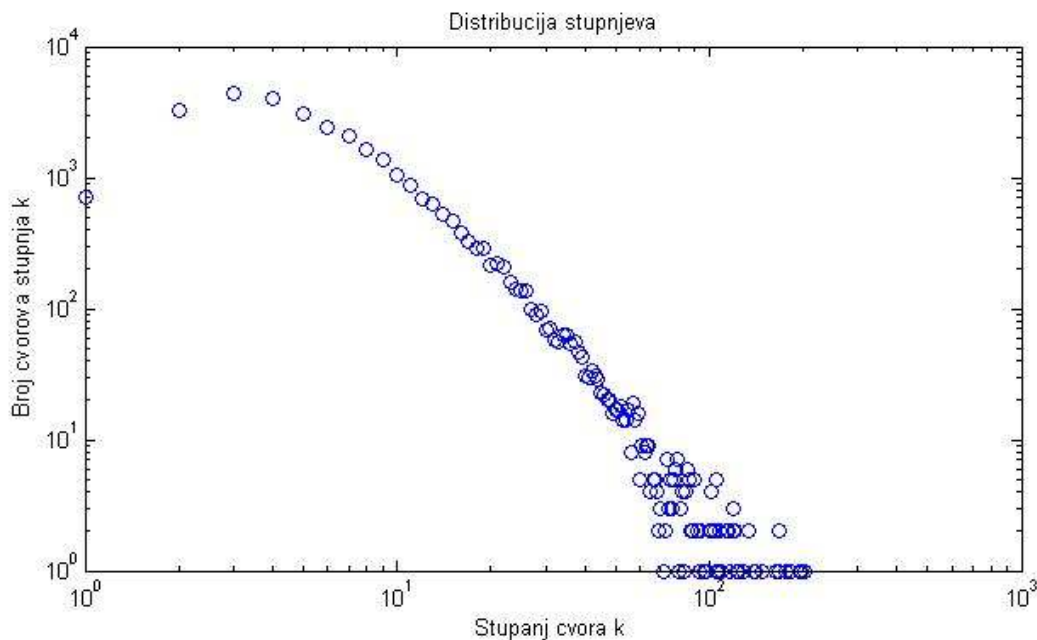


Slika 6. Vizualizacija naše mreže kolaboracije znanstvenika, svaki čvor ovisno o svom stupnju ima drugačiju boju i veličinu [13]

2.1. Svojstva promatrane mreže

Veličina mreže	31 163 čvora
Broj bridova	240 058
Srednji stupanj čvora	7.70
Srednji koeficijent grupiranja	0.723

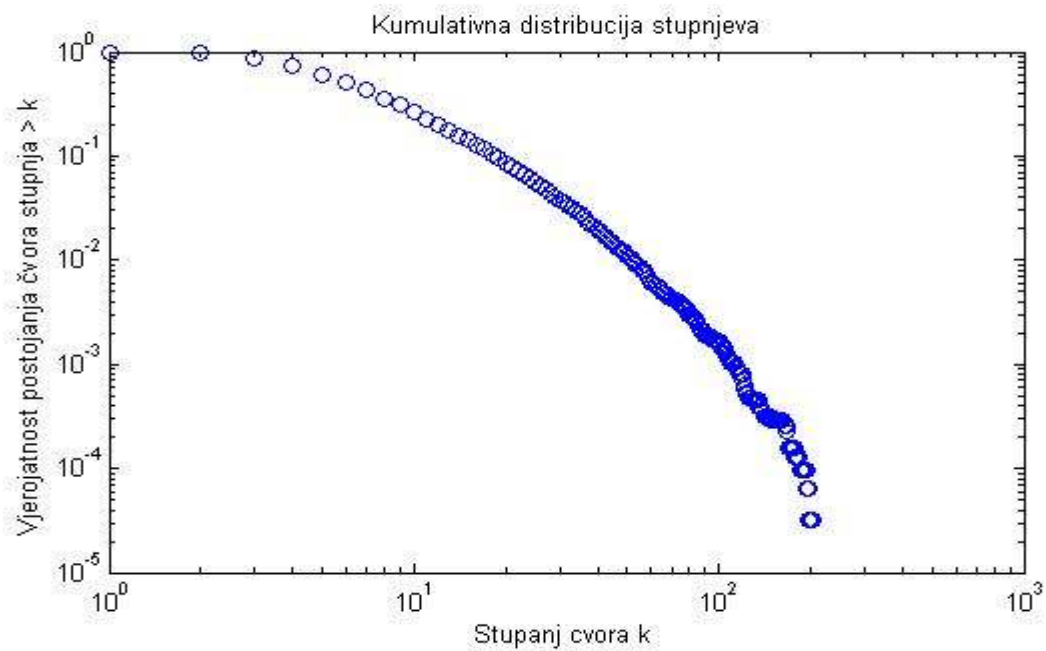
Tablica 2. Osnovne karakteristike mreže



Slika 7.

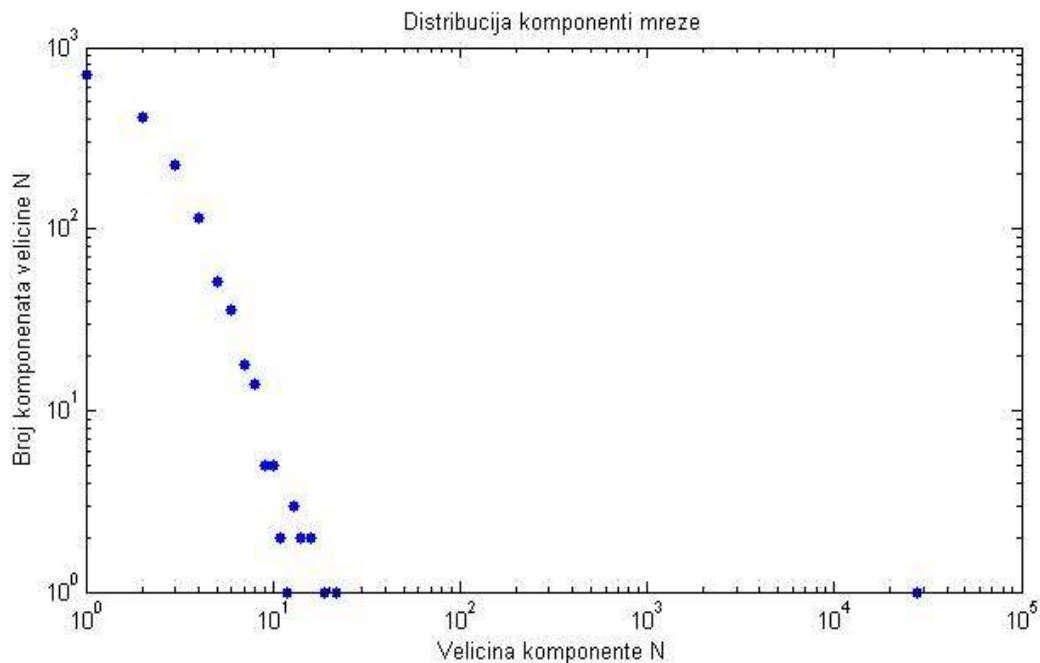
Na temelju prikaza distribucije stupnjeva promatrane mreže može se zaključiti da je mreža kolaboracije znanstvenika raspodijeljena prema $P(k) = k^{-\gamma}$ zakonu, točnije to je mreža bez skale. Također uočavamo efekt "debelog repa" (engl. *heavy-tail*) jer je promatrana mreža konačnih dimenzija tj. postoji veći broj čvorova visokog stupnja kojih ima jednako puno.

Iz odnosa broja vrhova i bridova u mreži vidimo naša mreža nije gusto povezana te smo za strukturu podataka koja reprezentira mrežu izabrali *sparse* matricu susjedstva.



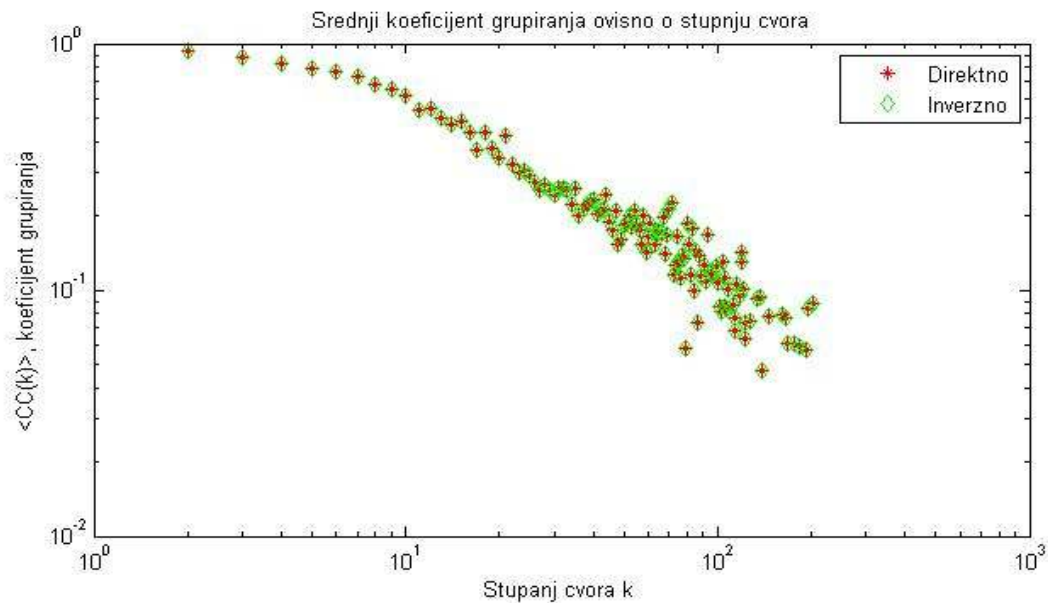
Slika 8.

Osim klasične distribucije stupnjeva moguće je prikazati i kumulativnu distribuciju stupnjeva naše mreže. Kumulativna distribucija stupnjeva nam pokazuje kolika je vjerojatnost postojanja čvora stupnja većeg od k . Kumulativna distribucija nam smanjuje fluktuacije podataka nastalih kod mreža konačnih dimenzija.



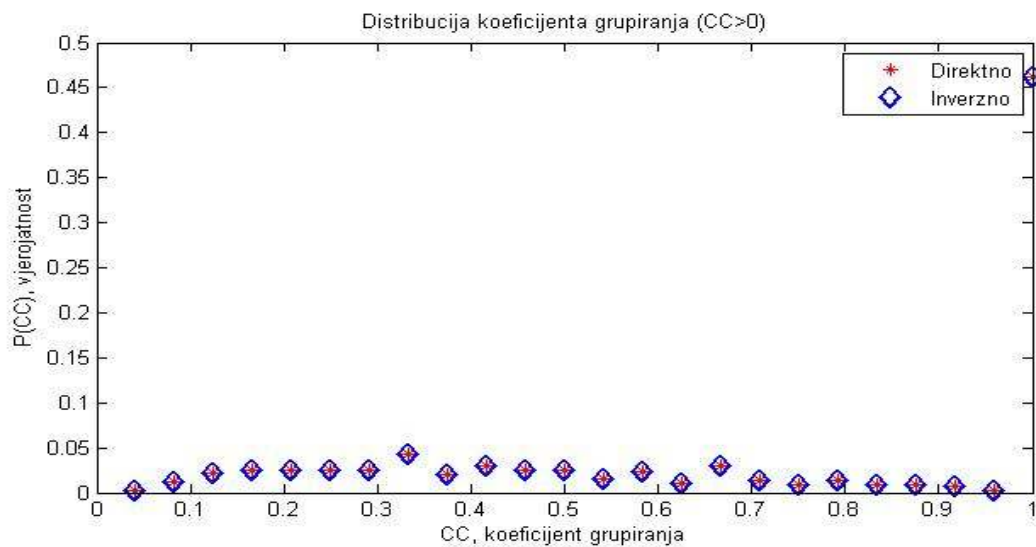
Slika 8.

Gigantska komponenta ove mreže je veličine 27 519 čvorova dok je raspodjela veličina komponenti raspodijeljena prema nekakvom $P(k) = k^{-\gamma}$ zakonu.



Slika 9.

Pomoću programskog paketa "Complex Networks Package" [15] izračunat je koeficijent grupiranja za svaki čvor i distribuciju koeficijenata grupiranja. Na temelju koeficijenata grupiranja svakog čvora može se zaključiti: čvorovi sa manjih indeksom bolje su povezani u svom prvom susjedstvu. Dok srednji koeficijent grupiranja čvorova u mreži iznosi 0.723.



Slika 10.

Na temelju grafa distribucije koeficijenta grupiranja zaključujemo da postoji vjerojatnost 0.45 da slučajno izabrani čvor u mreži posjeduje prve susjede koji su grupirani u potpuni graf. Vjerojatnost pojave koeficijenta grupiranja ostalih vrijednosti je manja od 0.05.

3. Metode

Izabrana je realna mreža kolaboracije znanstvenika (vidi 2. Podaci) kao mreža na kojoj će se vršiti sve simulacije. Uz odabrane početne uvjete simulira se širenje zaraze na mreži i potom se promatra promjena svojstava.

Početni uvjeti svake simulacije predstavljaju 4 parametra:

1. Parametar vjerojatnost širenja p - predstavlja vjerojatnost da se zaraza proširi između susjednih čvorova u jednom diskretnom koraku
2. Parametar vjerojatnost oporavka q - predstavlja vjerojatnost da se zaraženi čvor oporavi u jednom diskretnom vremenskom koraku
3. Parametar vjerojatnost izbacivanja d - predstavlja vjerojatnost da se zaraženi čvor izbacuje iz mreže
4. Vektor zaraženih- predstavlja vektor čvorova koji su zaraženi u početnom trenutku $t=0$

Nakon što su definirani svi početni uvjeti, simulacije se počinju računati. Minimalan broj simulacija određen je dodatnim parametrom dok se gornja granica broja simulacija računa dinamički. Svaka simulacija se sastoji od više modula:

1. Modul "Simulator epidemije": unutar ovog programskog modula simulira se jedan mogući ishod širenja zaraze kroz zadanu mrežu. Ulaz ovog modula su početni uvjeti SIR modela a izlaz je vektor koji nam pokazuje koji su čvorovi bili zaraženi tokom epidemije.
2. Modul "Frekvencije zaraženih": unutar ovog programskog modula pamti se ukupan broj čvorova koji su bili zaraženi tokom trenutne simulacije i rezultat se pohranjuje u poseban vektor. Nakon svih simulacija unutar tog vektora imati ćemo informacije o frekvencijama broja zaraženih tokom svih simulacija.
3. Modul "Izbacivanje čvorova": unutar ovog programskog modula izračunavaju se indeksi čvorova od svih zaraženih koji će biti izbačeni iz mreže na temelju parametra vjerojatnosti izbacivanja d .

4. Modul "Distribucije stupnjeva": unutar ovog programskog modula se na temelju vektora izbačenih čvorova računaju nove distribucije stupnjeva novonastale mreže. Zatim se izračuna kumulativna distribucija novonastale mreže.
5. Modul "Distribucije stupnjeva izbačenih čvorova": unutar ovog programskog modula se na temelju vektora izbačenih čvorova računa distribucija stupnjeva čvorova koji su izbačeni iz mreže na temelju parametra d .
6. Modul "Promjena stupnjeva čvorova": unutar ovog programskog modula se stvara 3D graf koji prikazuje promjene stupnjeva čvorova prilikom izbacivanja nekih čvorova. Nakon što smo iz mreže izbacili neke čvorove za svaki čvor $v \in G$ (početna mreža) vrijedi :
 - a.) ako je $v \in$ vektora *izbačeni* \Rightarrow stupanj poslije (v) = 0
 - b.) ako $v \notin$ vektora *izbačeni* \Rightarrow (ako \exists čvor $y \in$ vektora *izbačeni*) \wedge (ako y incidentan sa v) \Rightarrow (stupanj poslije (v) < stupanj prije (v))
 - c.) inače stupanj poslije (v) = stupanj prije (v)

Apscisa 3D grafa predstavlja stupanj čvora u staroj mreži, ordinata predstavlja stupanj čvora u novoj mreži dok aplikata prikazuje broj takvim promjena tijekom fragmentacije.

7. Modul "Analiza fragmentacije mreže": unutar ovog programskog modula poziva se funkcija za računanje komponentata u novoj mreži. Svaka komponenta je jednoznačno opisana minimalnim rednim brojem čvora maksimalnog stupnja unutar te komponente. Kao izlaz ovog modula imamo vektor distribucije komponenti i vektor promjene veličine komponente. Vektor distribucije komponenti nam pokazuje kako su distribuirane veličine komponenti poslije fragmentacije. Vektor promjene veličine komponenti pokazuje kakve strukturalne promjene su se dogodile unutar zaražene komponente.

Simulacije (svih sedam modula) se računaju zajedno s pripadnim srednjim vrijednostima i standardnim devijacijama svih promatranih veličina sve dok se ne ispuni jedan od sljedeća dva uvjeta:

(1) $\sigma(k, t) \leq \varepsilon \bar{P}(k, t), \forall k$, gdje $\bar{P}(k, t)$ predstavlja srednju vrijednost distribucije stupnjeva u trenutku t za stupanj k , ε - predstavlja granicu okoline unutar koje $\sigma(k, t)$ mora upasti.

Čitav izraz u uvjetu (1) odgovara slučaju kada relativna standardna devijacija $\frac{\sigma(k,t)}{P(k,t)}$ konvergira prema ε okolini nule. Ovaj uvjet se ispunjava u unimodalnom režimu sustava širenja zaraze (vidi 4.1. Bimodalni proces)

(2) $|\sigma(k,t+1) - \sigma(k,t)| \leq \varepsilon \sigma(k,t), \forall k$, gdje $\sigma(k,t+1)$ predstavlja vrijednost standardne devijacije distribucije stupnjeva u trenutku $t+1$ za stupanj k , $\sigma(k,t)$ predstavlja vrijednost standardne devijacije distribucije stupnjeva u trenutku t za stupanj k .

Čitav izraz u uvjetu (2) odgovara trenutku u kojem standardna devijacija konvergira prema nekakvoj konačnoj vrijednosti različitoj od nule. Ovaj uvjet se ispunjava u bimodalnom režimu rada sustava širenja epidemije (vidi Bimodalni proces).

Za sva promatrana svojstva osim srednje vrijednosti podataka računamo i standardnu devijaciju podataka.

$$(16) \quad \sigma = \sqrt{\frac{1}{N} (\sum_{i=1}^N x_i^2 - N \bar{x}^2)}$$

Kako bi se moglo bolje pratiti promjene distribucije stupnjeva prije i poslije fragmentacije, uvodi se nova slučajna varijabla:

$$(17) \quad \Delta(k) = \frac{P(k)_{orig} - P(k)_{sim}}{P(k)_{orig}}$$

Sukladno tome uvodi se i mjera očekivanja slučajne varijable delta:

$$(18) \quad E(\Delta(k)) = \frac{P(k)_{orig} - \overline{P(k)_{sim}}}{P(k)_{orig}}$$

Također, računa se i mjera standardne disperzije slučajne varijable delta:

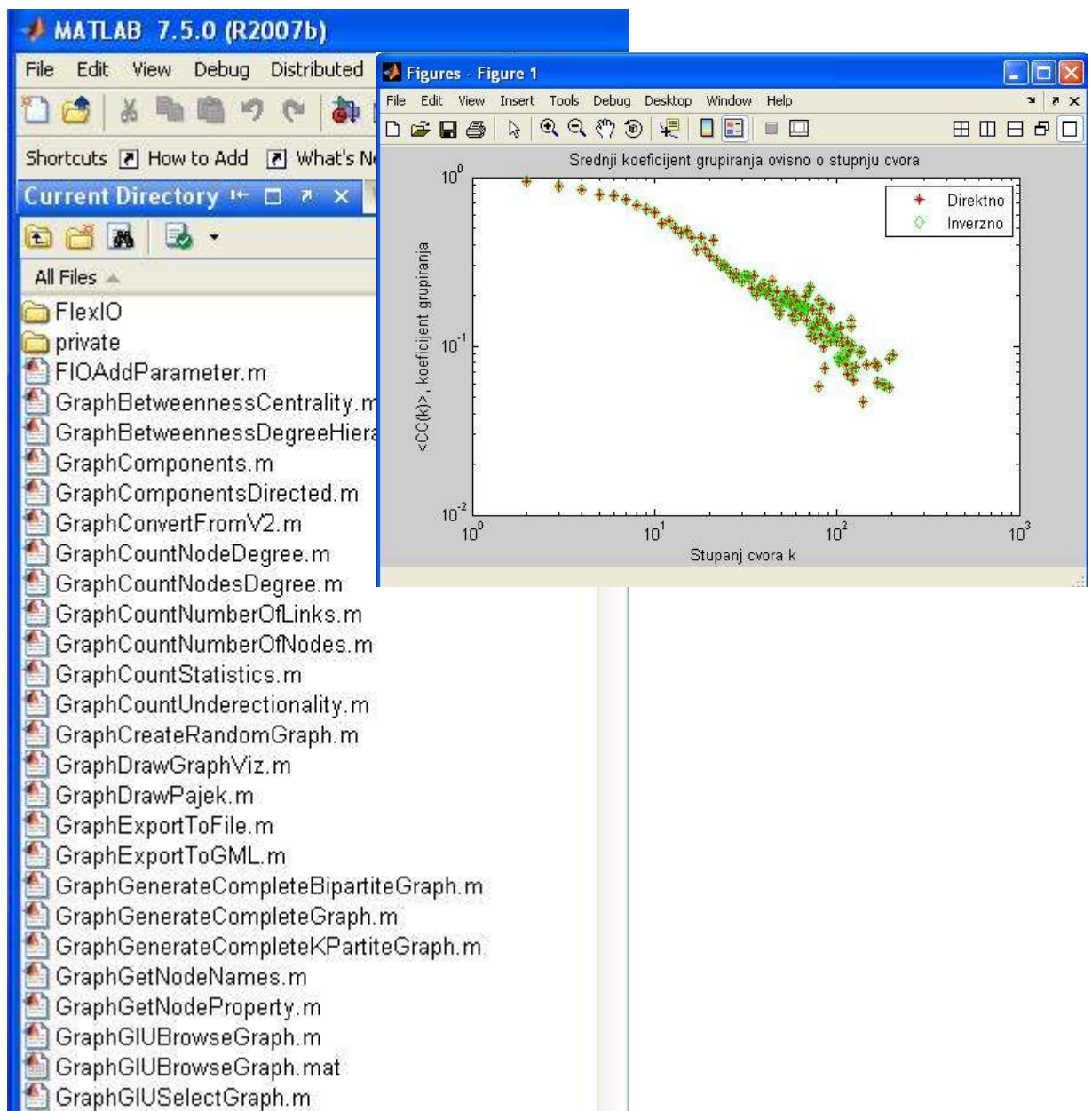
$$(19) \quad \sigma(\Delta(k)) = \frac{\overline{\sigma_{P(k)}(k)_{sim}}}{P(k)_{orig}}$$

Promjenu, tj. razliku između početne i konačne distribucije prikazivat ćemo grafički ovako:

$$(20) \quad \overline{\Delta(k)} \pm \sigma_{\Delta}(k)$$

3.1. "Complex Networks Package"

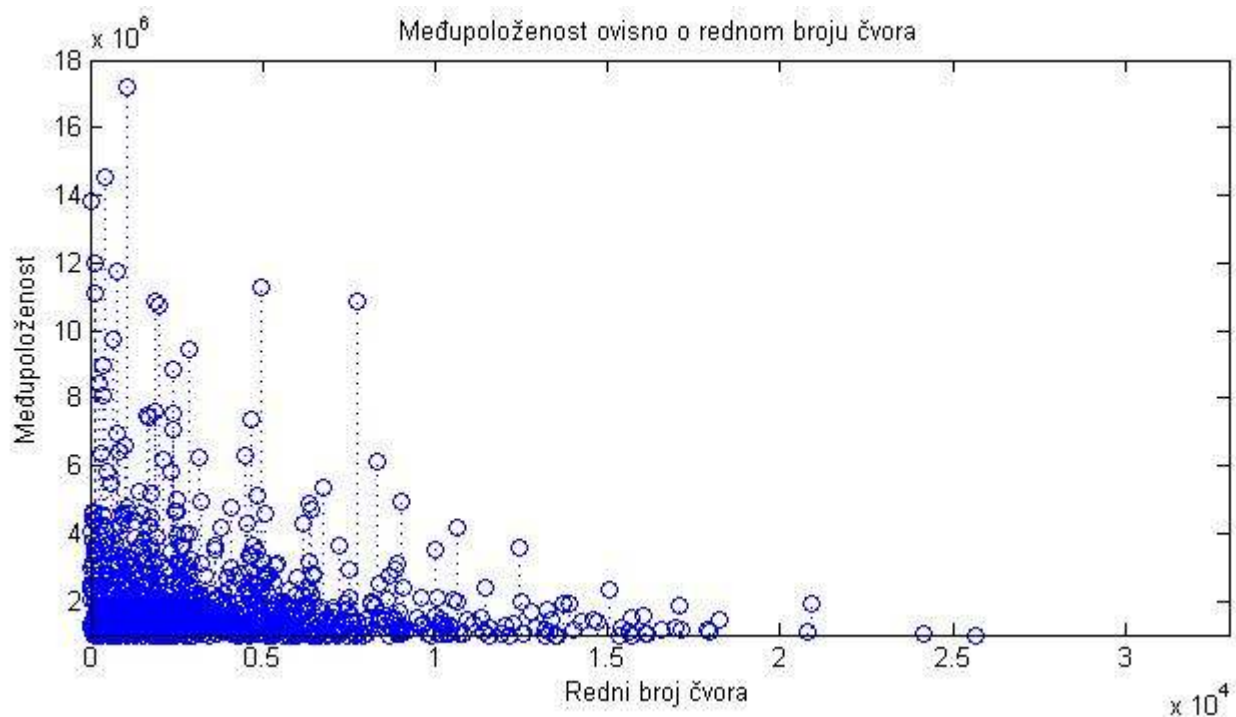
Programski paket "Complex Networks Package" namijenjen je za analizu mreža u Matlabu. Javno je dostupan na [15]. Pomoću njega se računaju svojstva vezana za grupiranje čvorova (graf "Srednji koeficijent grupiranja ovisno o stupnju" i graf "Distribucija koeficijenta grupiranja").



Slika 11. Prikaz mogućih funkcija unutar programskog paketa: "Complex Networks Package"

3.2. "MatlabBGL"

"MatlabBGL" [16] je programski paket za Matlab namijenjen za analizu grafova. Koristi "Boost Graph Library" kako bi efikasno implementirao algoritme za grafove. Namijenjen je za rad s većim grafovima koji nisu "gusto" povezani (engl. *sparse graph*) reda veličine 100 000 čvorova. Ovaj paket je korišten kako bi se izračunala međupoloženost za svaki čvor u mreži kolaboracije znanstvenika.



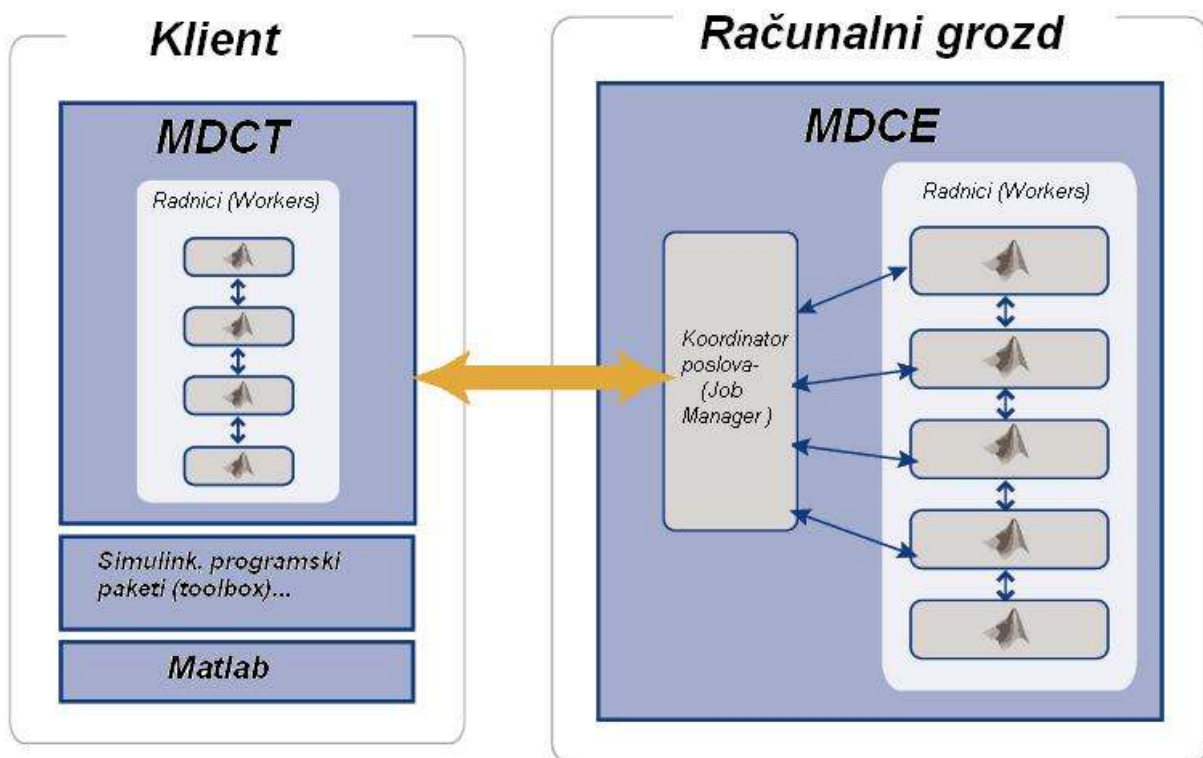
Slika 12.

Na temelju grafa koeficijenta međupoloženosti čvorova može se zaključiti kako postoji skupina čvorova s većim koeficijentom međupoloženosti a to su upravo čvorovi sa manjim rednim brojem. Na temelju grafa koeficijenta grupiranja (Slika 9.) i međupoloženosti (Slika 12.) vidimo da postoji korelacija između visokog koeficijenta međupoloženosti i koeficijenta grupiranja čvorova. Također zaključujemo da su čvorovi sa manjim rednim brojem u našoj mreži u prosjeku bitniji u mnogim procesima na mreži (širenje informacija i zaraze).

3.3. "Matlab Distributed Computing Toolbox"

"Matlab Distributed Computing Toolbox" [17] (MDCT) predstavlja programski paket koji omogućava distribuirano računanje na računalnom grozdu (engl. *cluster*). Svaki program koji je vremenski zahtjevan a moguće ga je raspodijeliti na manje podzadatke (engl. *task*) koji će se paralelno izvoditi zove se zadatak (engl. *job*). Sam autor zadatka bira kako će se njegov veliki zadatak podijeliti na manje podzadatke.

Matlab sjednica (engl. *session*) u kojoj su zadatak i njegovi podzadaci definirani zove se klijentska sjednica (engl. *client session*). Matlab Distributed Computing Engine (MDCE) je servis koji omogućava izvođenje zadatka, podzadataka i vraćanje rezultata u klijentsku sjednicu. Koordinator poslova (engl. *job manager*) je dio MDCE sustava koji koordinira izvođenje podzadataka. Koordinator poslova (engl. *job manager*) distribuira podzadatke radnicima (engl. *worker*). Prikaz navedene sheme rada prikazana je na Slici 13:



Slika 13. Prikaz osnovnih modula unutar Matlab Distributed Computing Toolbox-a

4. Rezultati

Sve simulacije su napravljene na mreži kolaboracije znanstvenika "Condensed matter collaborations 2003" (vidi 2.Podaci) veličine 31163 čvorova. Simulacije su se distribuirano računale na računalnom grozdu pomoću MDCT-a (vidi 3.2. "Matlab Distributed Computing Toolbox").

Simulacije su podijeljene u dva zadataka i više podzadataka:

1. Bimodalnost procesa- utvrditi pod kojim uvjetima (početnim parametrima) je proces širenja zaraze bimodalan a kada unimodalan.
2. Ispitati parametarski prostor ($p-q-d$) te pokušati utvrditi kako se mijenja odziv sustava u različitim točkama parametarskog prostora

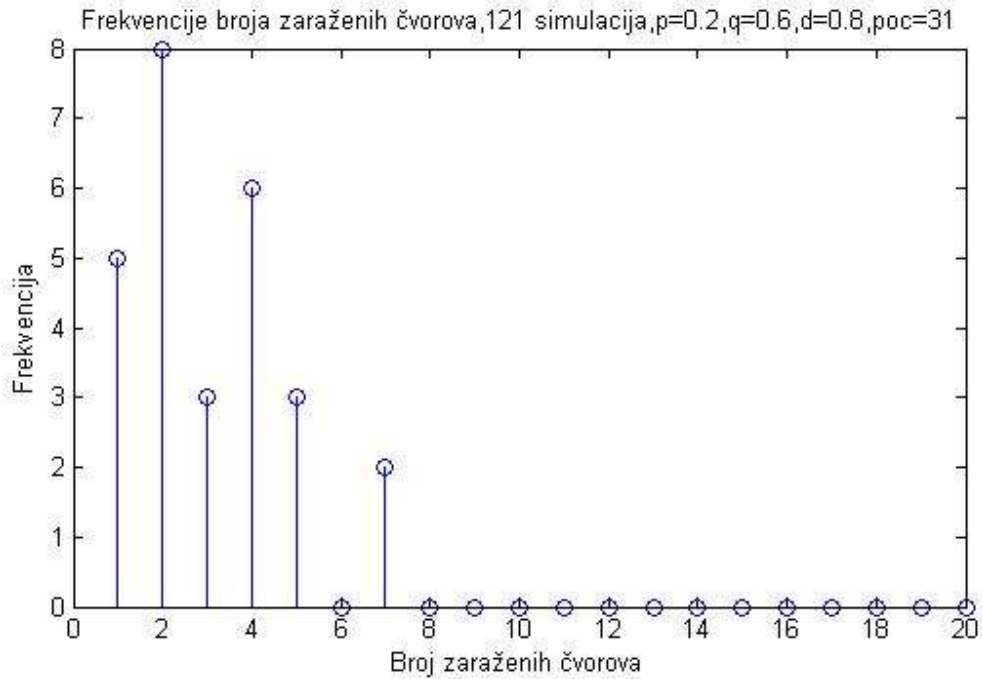
2.1 .Odabrati takav p (vjerojatnost širenja zaraze) da je proces zaraze unimodalan i pokušati sa promjenom parametra q (vjerojatnost oporavka) otići u bimodalno područje. Parametar d fiksirati.

2.2. Odabrati takav p (vjerojatnost širenja zaraze) da je proces zaraze bimodalan i pokušati sa promjenom parametra q (vjerojatnost oporavka) otići u unimodalno područje. Parametar d fiksirati.

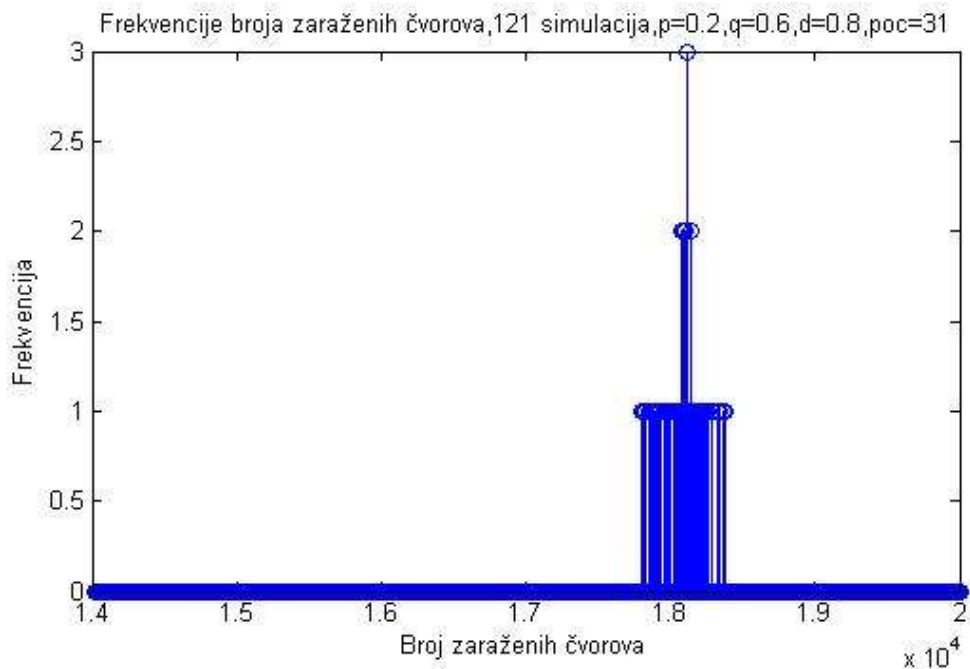
2.3. Fiksirati parametre q i d , mijenjati parametar p i utvrditi karakteristične odzive sustava.

2.4. Fiksirati parametre p i q , mijenjati parametar d i utvrditi karakteristične odzive sustava.

4.1. Bimodalni proces:



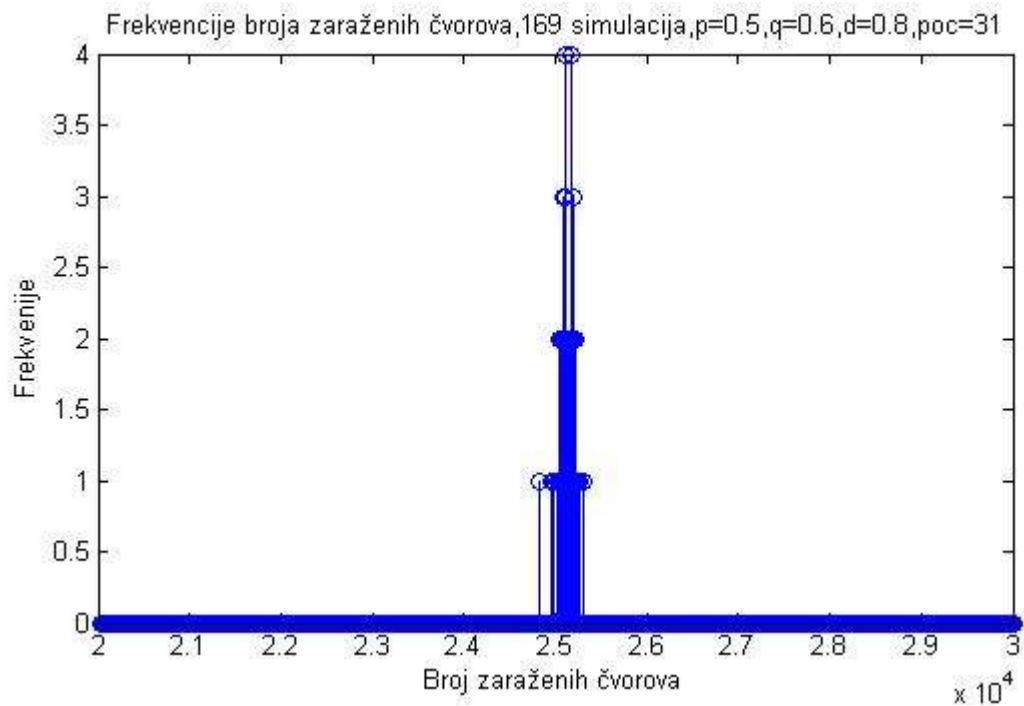
Slika 14.



Slika 15.

Proces širenja zaraze za neke početne parametre (p i q) može biti bimodalan dok za druge unimodalan. Bimodalnost ovog procesa se očituje s dva moguća ishoda širenja zaraze (1)

zaraza se globalno proširila ili (2) zaraza se samo lokalno proširila na nekolicinu susjednih čvorova. Za vrijednost parametra $p = 0.2$ i $q = 0.6$ očituje se bimodalno ponašanje. Na prvom grafu moguće je vidjeti rezultat ishoda kada se zaraza proširi samo lokalno na neke susjede. Dok se na drugom grafu vidi drugi slučaj kada se zaraza proširi globalno. Ukoliko je proces širenja bimodalan konvergencija relativne standardne devijacije kumulativne distribucije stupnjeva prema nuli nije moguća već ona konvergira nekoj konačnoj vrijednosti.



Slika 16.

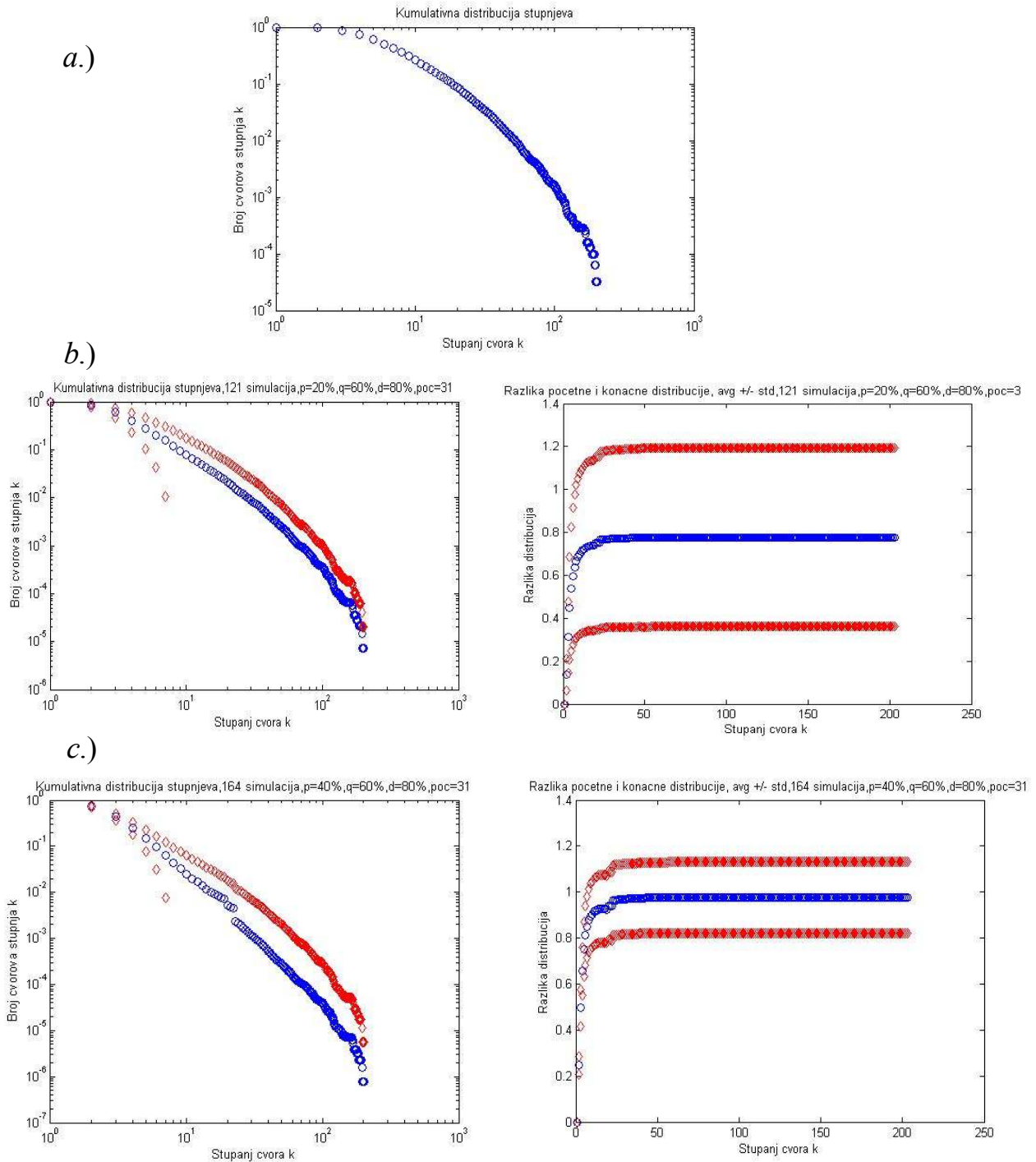
Ukoliko se vrijednost parametra p promjeni sa $p = 0.2$ na $p = 0.5$ bimodalnost nestaje. Proces je unimodalan tj. drugim riječima zaraza se uvijek proširi globalno.

Bimodalnost procesa zaraze je intuitivna jer se za male vrijednosti parametra širenja zaraze i relativno veliku vjerojatnost oporavka može dogoditi slučaj kada se zaraza jednostavno brzo ugasi u lokalnom području oko izbora zaraze.

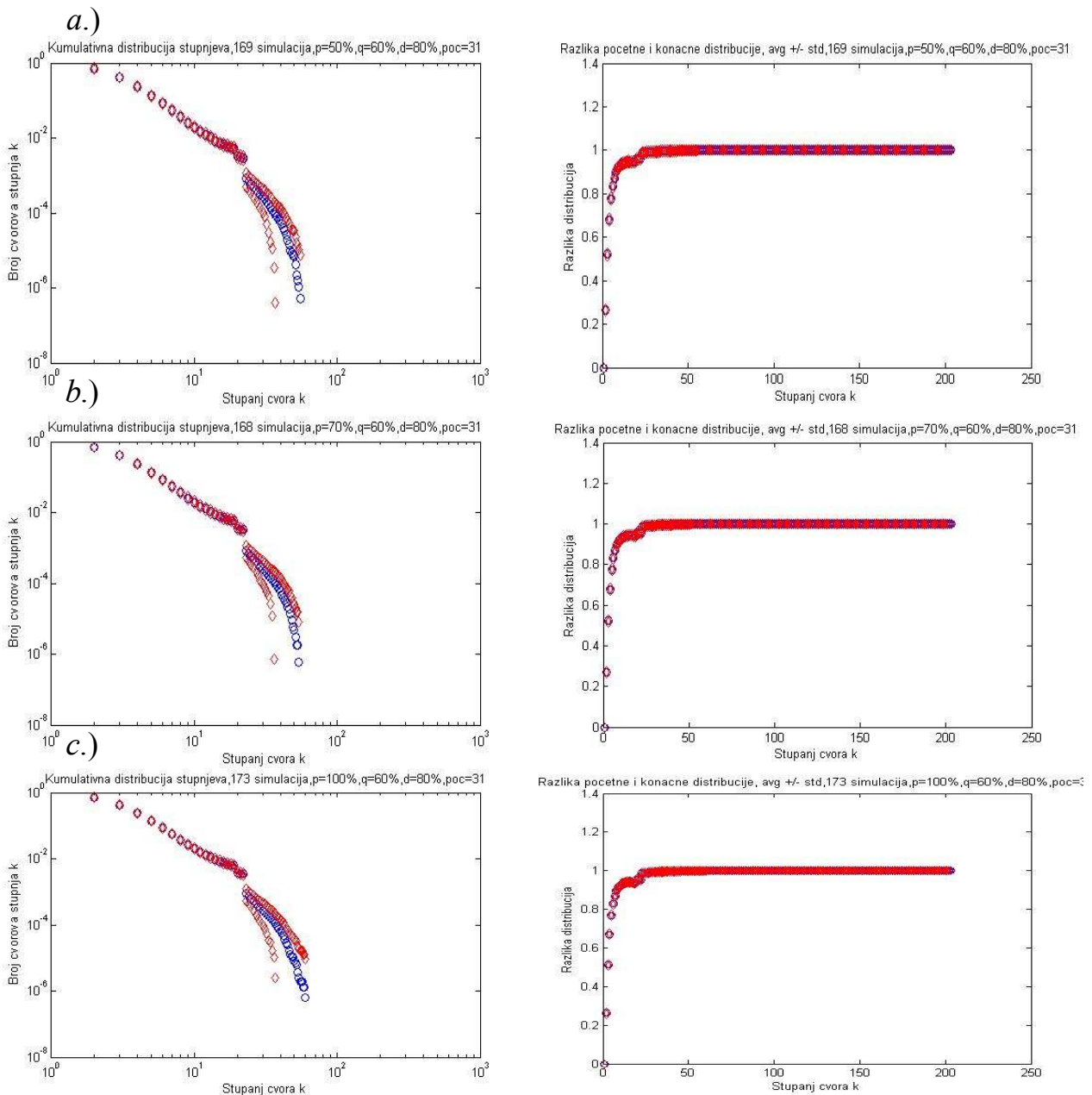
4.2. Parametarski prostor (p - q - d)

4.2.1. Promjena vjerojatnosti zaraze p

4.2.1.1. Distribucija stupnjeva novonastale mreže



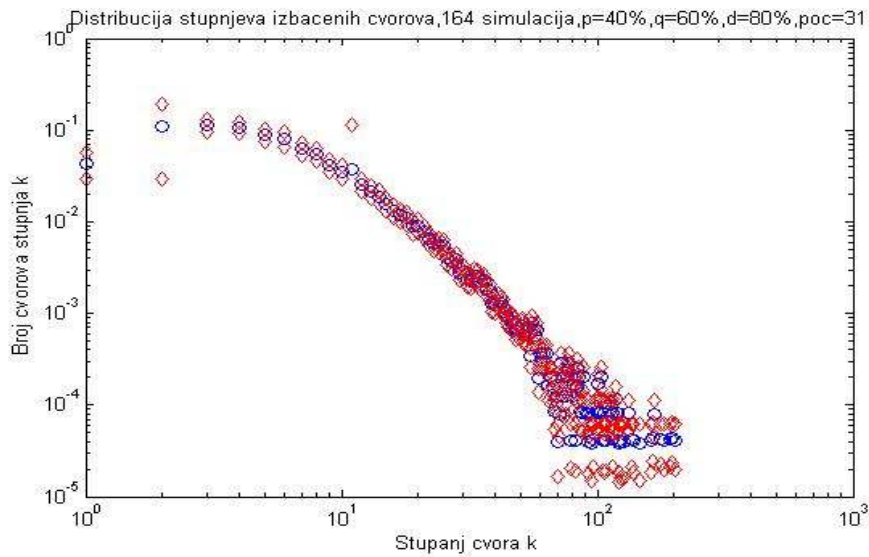
Slika 17. Usporedni prikaz kumulativnih distribucija stupnjeva i razlika distribucija za $q = 0.6$, $d = 0.8$ i različite parametre p : a.) početna kumulativna distribucija, b.) $p = 0.2$, c.) $p = 0.4$



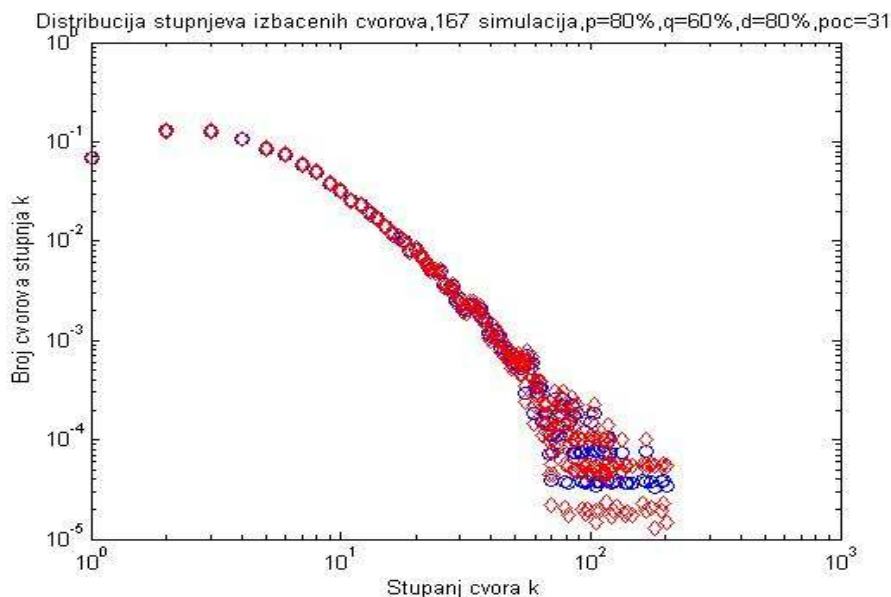
Slika 18. Usporedni prikaz kumulativnih distribucija stupnjeva i razlika distribucija za $q = 0.6$, $d = 0.8$ i različite parametre p : a.) $p = 0.5$, b.) $p = 0.7$, c.) $p = 1$

Na svim grafovima plavi simboli predstavljaju srednje vrijednosti podataka dok crveni simboli predstavljaju $\bar{X} \pm \sigma$ standardno odstupanje od srednje vrijednosti. Prvi graf predstavlja kumulativnu distribuciju stupnjeva početne mreže. Iz oblika kumulativne distribucije možemo zaključiti da je početna mreža "mreža bez skale". Iz navedenih grafova možemo dobiti predodžbu kako vjerojatnost zaraze utječe na distribuciju čvorova novonastale mreže nakon izbacivanja čvorova. Za vrijednosti parametra p od $[0-0.4]$ čitava distribucija stupnjeva ne mijenja znatno svoj oblik. Dok za vrijednosti parametra $p > 0.5$ distribucija stupnjeva mijenja svoj oblik drastično jer dolazi do značajne fragmentacije gigantske komponente.

4.2.1.2 Distribucije stupnjeva izbačenih čvorova



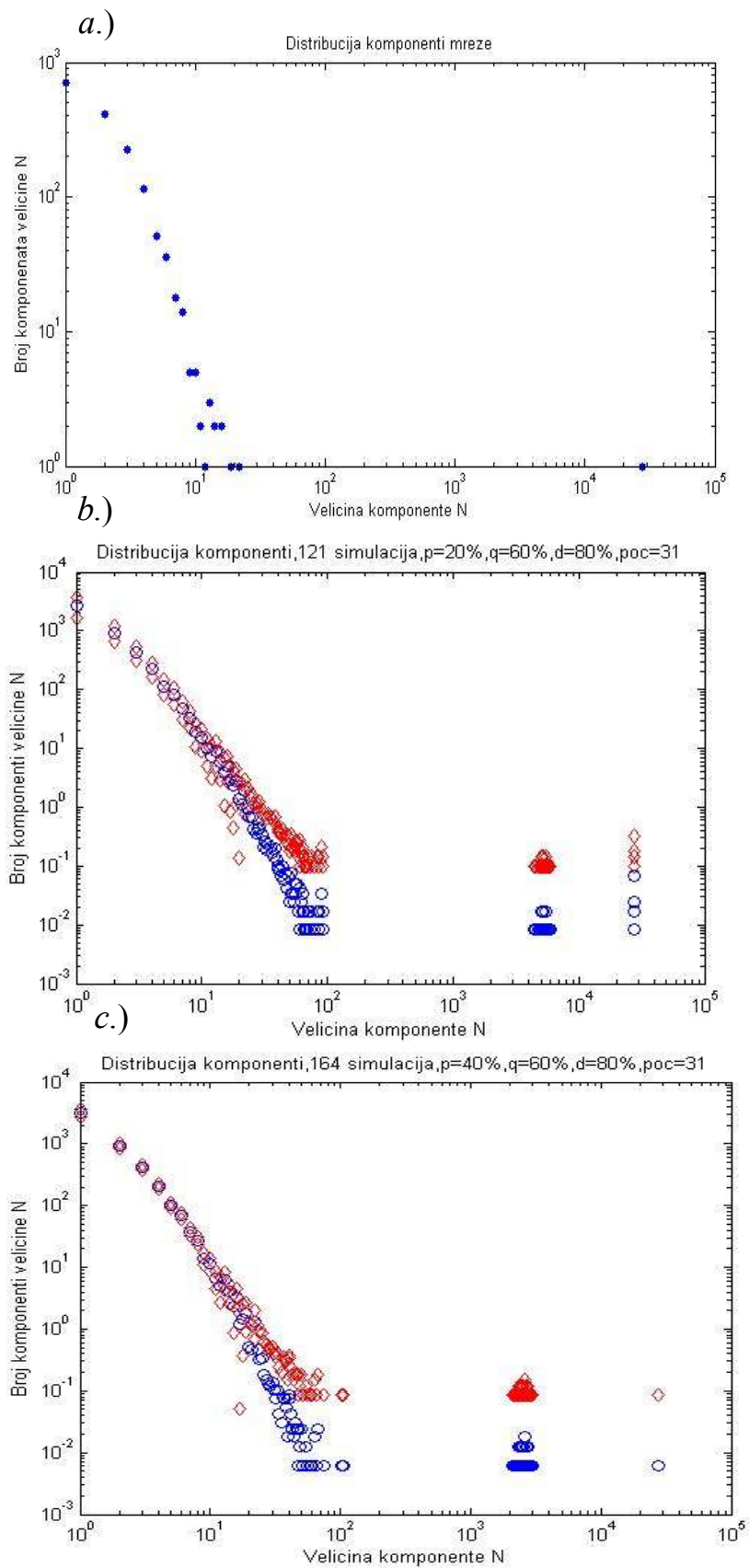
Slika 19.



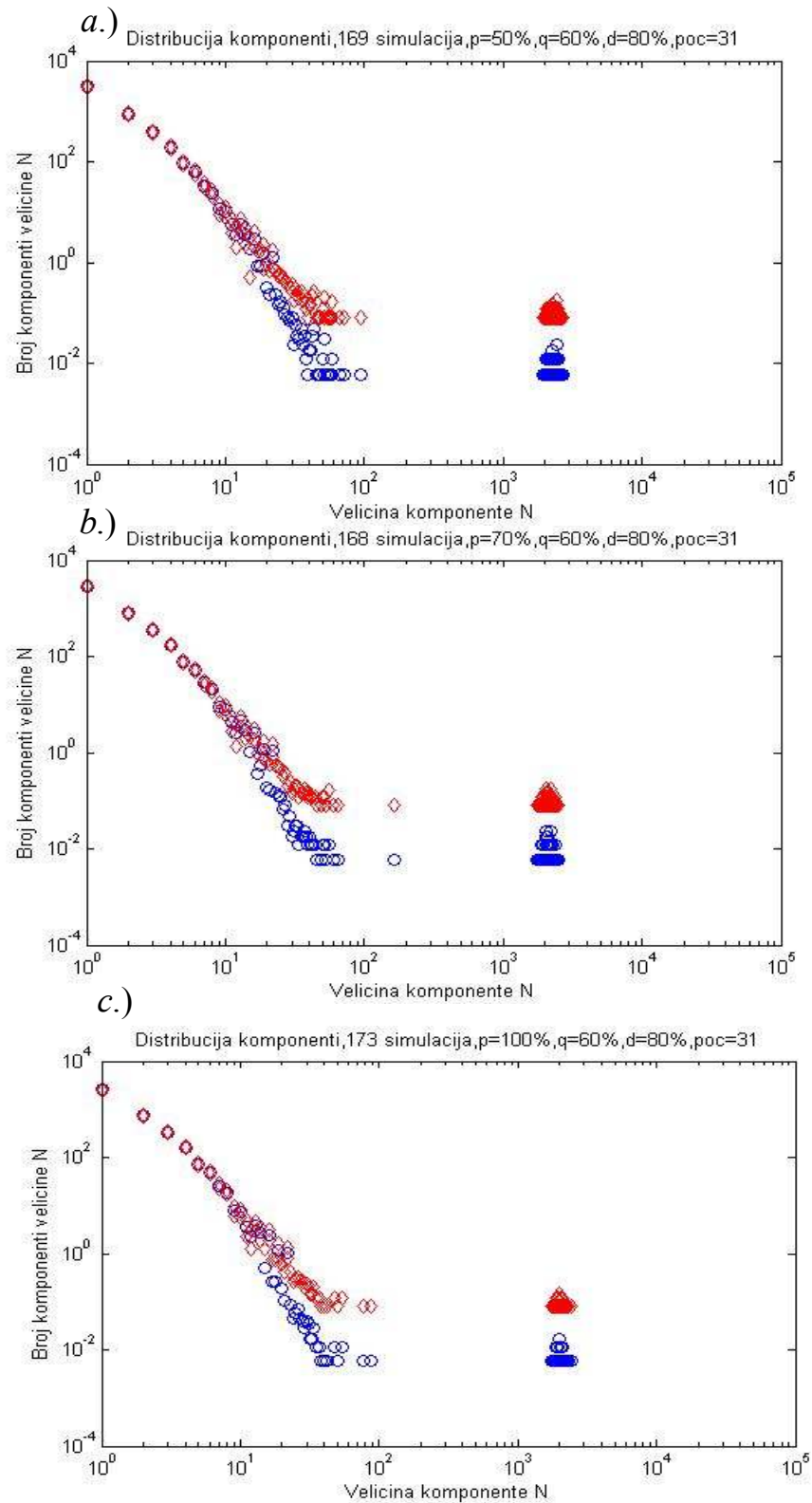
Slika 20.

Nakon prestanka širenja zaraze kroz mrežu dobijemo podgraf G' koji se sastoji od svih čvorova za koje vrijedi: $v \in G \wedge v \in \{skup_zaraženih\}$. Nakon toga uniformno s vjerojatnosti d izabiremo čvorove iz skupa zaraženih koji se izbacuju iz mreže te dobijemo skup izbačenih čvorova. Skup izbačenih čvorova tvori podgraf izbačenih čvorova koji nije nužno povezan. Za razne parametre p distribucija stupnjeva ostaje nepromijenjena jer se vrijednosti normiraju sa brojem izbačenih čvorova. Na grafu plavi simboli predstavljaju srednje vrijednosti podataka dok crveni simboli predstavljaju $\bar{X} \pm \sigma$ standardno odstupanje od srednje vrijednosti.

4.2.1.3 Distribucije komponenti novonastale mreže



Slika 21. Usporedni prikaz distribucija komponenti za $q = 0.6$, $d = 0.8$ i različite parametre p : a.) početna distribucija komponenti, b.) $p = 0.2$, c.) $p = 0.4$



Slika 22. Usporedni prikaz distribucija komponenti za $q = 0.6$, $d = 0.8$ i različite parametre n : a.) $n = 0.5$. b.) $n = 0.7$. c.) $n = 1$

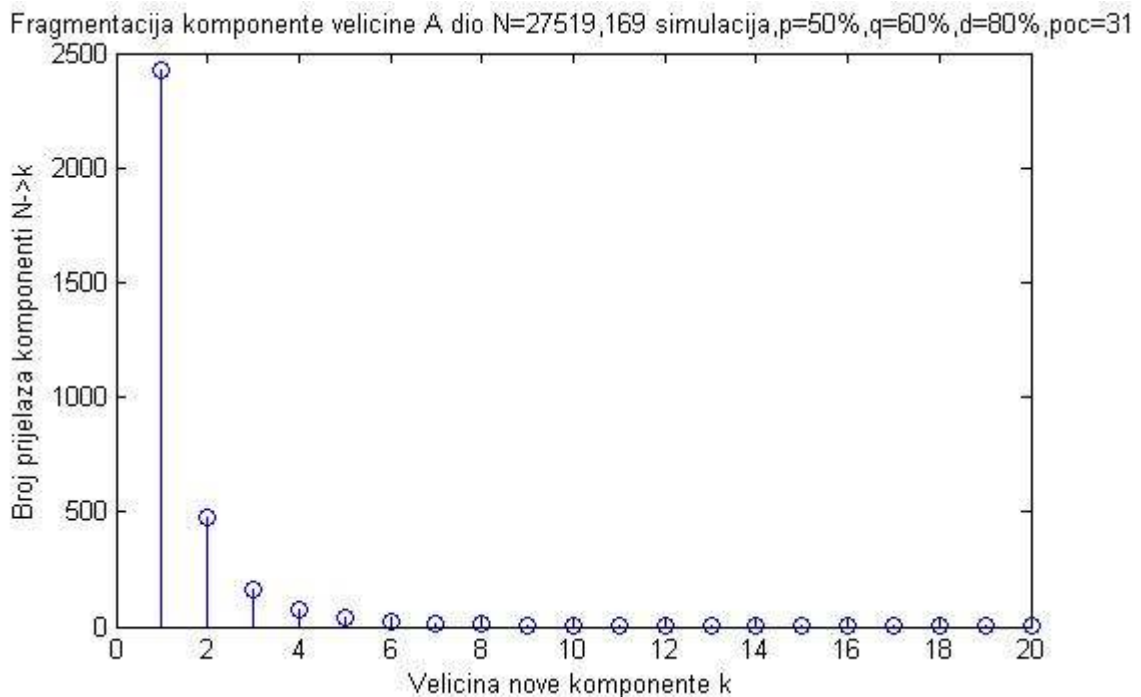
Gigantska komponenta ove mreže je veličine 27 519 čvorova. Dok je raspodjela veličina komponenti raspodijeljena prema nekakvom $P(k) = k^{-\gamma}$ zakonu. Povećanjem parametra p

vidimo kako se gigantska komponenta sve više fragmentira u manje komponente. Za vrijednost parametra $p = 0.5$ gigantska se komponenta fragmentirala u svim simulacijama.

4.2.1.4 Dinamika fragmentacije

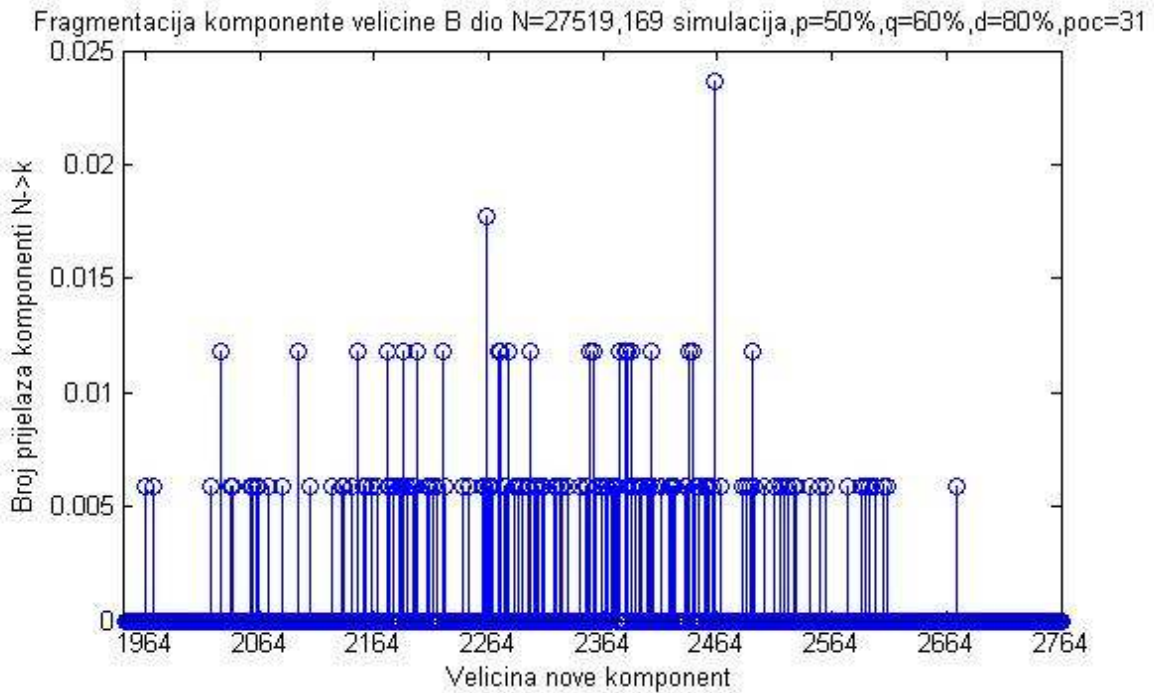
Definiramo graf promjene veličine komponenti tako da na apscisu postavimo vrijednosti veličina novonastalih komponenti, a na ordinatu postavimo broj takvim promjena. Pošto je zaraza ograničena na samo jednu komponentu i to onu komponentu kojoj pripada početno zaraženi čvor sve promjene veličine komponenti se događaju za iste veličine komponente.

Neka se u grafu G nalazi komponenta A veličine 27 519 čvorova i neka se ona raspadne na tri komponente B , C i D redom veličine 15 000, 12000 i 519 čvorova onda možemo napraviti graf promjene veličine komponenti tako da na x-os redom stavimo vrijednosti 15 000, 12000 i 519 i na y-os pripadne vrijednosti broja takvih prijelaza 1, 1 i 1.



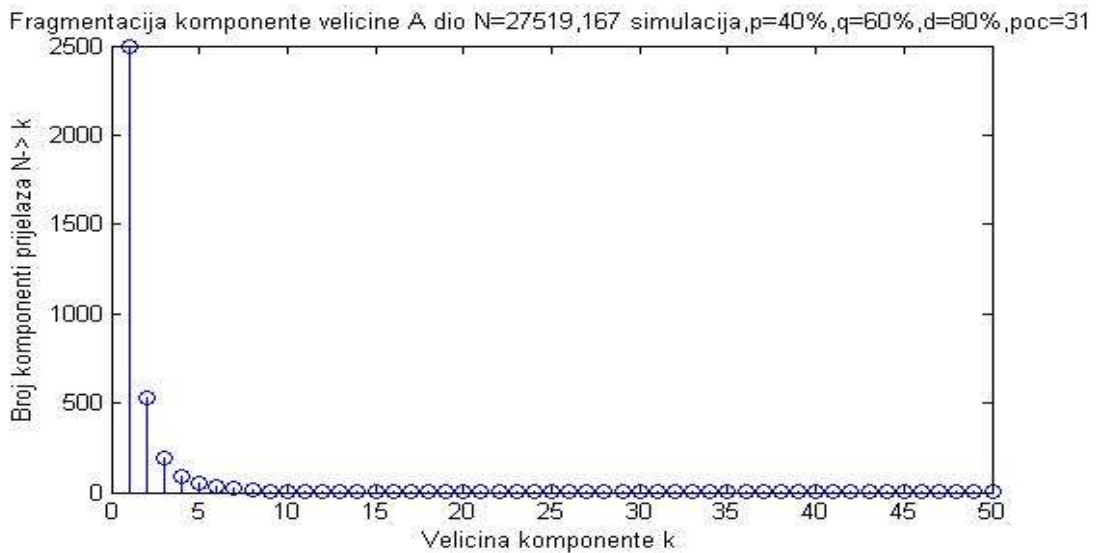
Slika 23.

Ovaj graf prikazuje dio promjena veličine komponenti uslijed fragmentacije s početne gigantske komponente veličine 27 519 čvorova na veličine komponenti do 20 čvorova. Iz ovog grafa zaključujemo da se uslijed fragmentacije gigantske komponente nastaje najviše komponenti veličine redom 1, 2, 3 i 4 što je logično pošto u našoj mreži ima najviše čvorova s malim stupnjem koji su povezani s čvorovima velikog stupnja. Čvorovi velikog stupnja imaju veću vjerojatnost izbacivanja što smo mogli vidjeti iz serije grafova distribucija stupnjeva.



Slika 24.

Ovaj graf prikazuje dio promjena veličine komponenti uslijed fragmentacije s početne gigantske komponente veličine 27 519 čvorova na veličine komponenti iznad 1964 čvora. Iz ovog grafa možemo zaključiti da je najvjerojatniji prijelaz gigantske komponente na veličinu 2464 čvorova.

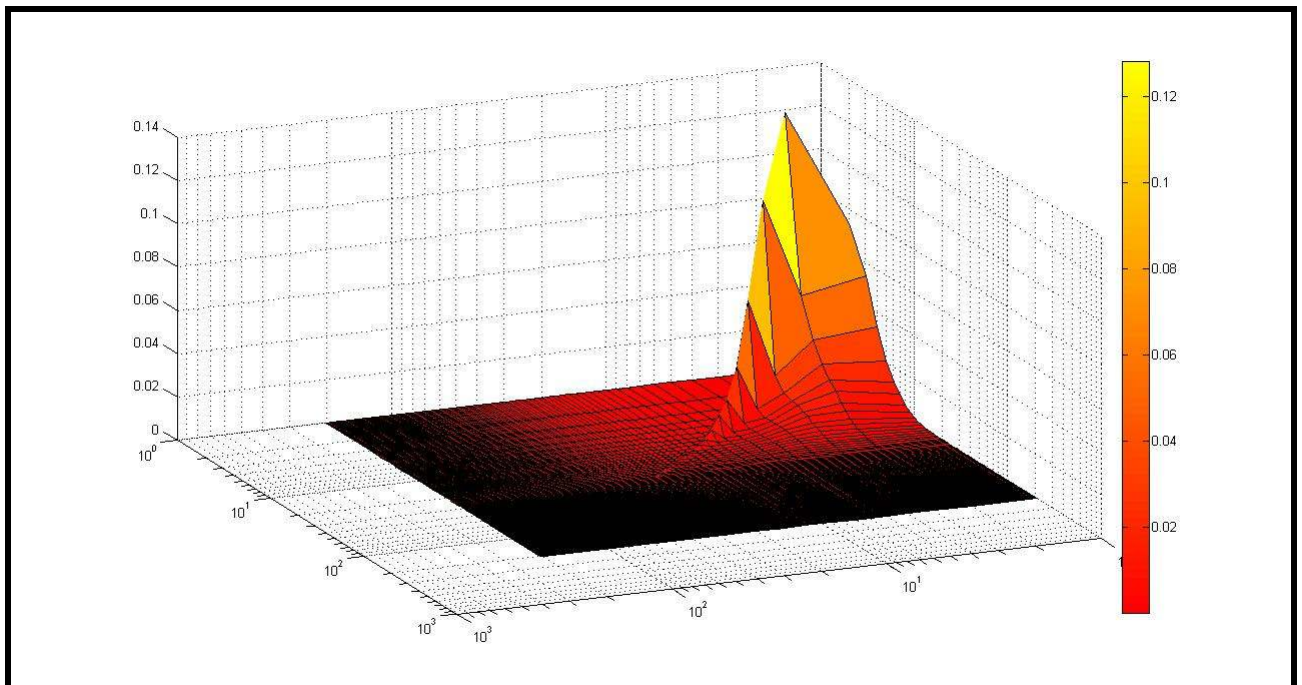


Slika 25.

Također primjećujemo da se za vrijednost parametra $p = 0.4$ gigantska komponenta ne fragmentira na velike komponente. Dok se za vrijednost parametra $p = 0.5$ gigantska komponenta fragmentira što upućuje na postojanje nekakvog praga za fragmentaciju gigantske komponente.

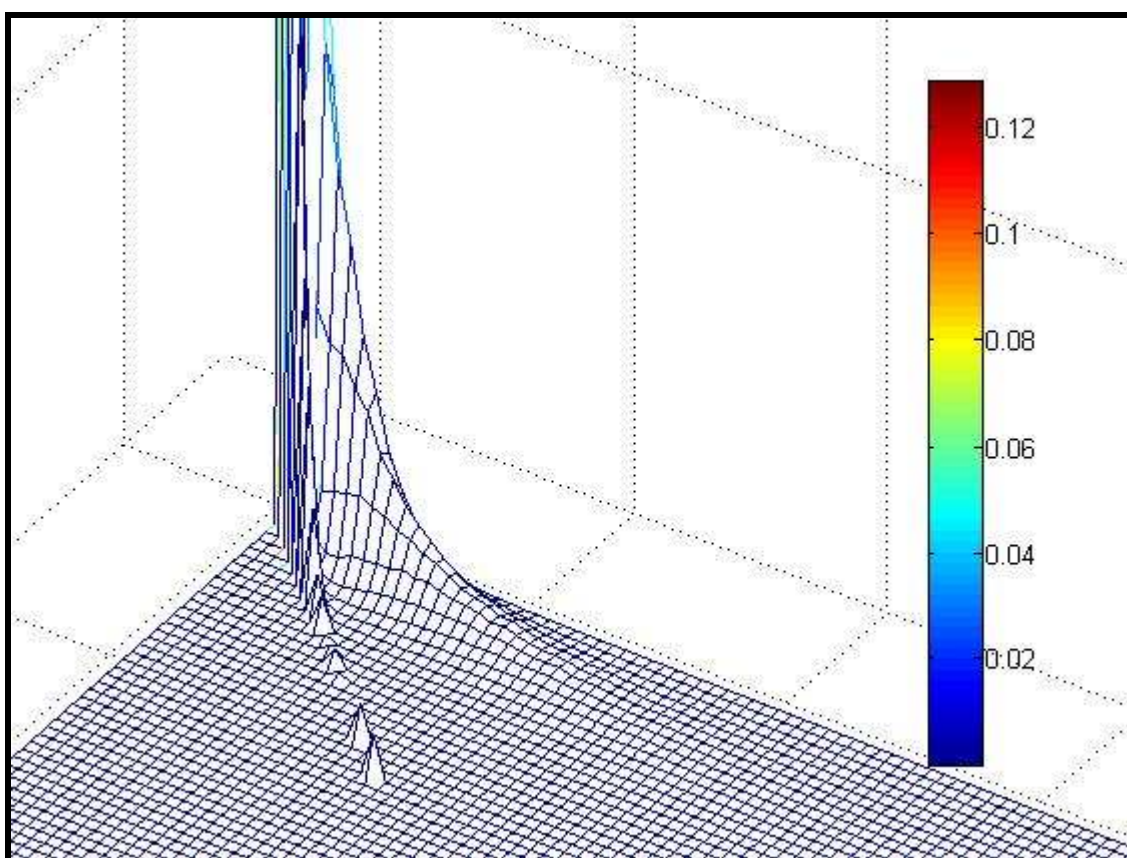
4.2.1.5 Promjena stupnjeva fragmentacijom

Fragmentacijom gigantske komponente mijenjaju se stupnjevi čvorova u mreži. Najbolji način za detekciju tih promjena je stvaranje 3D grafa koji na apscisi ima stupnjeve čvorova u početnoj mreži, na ordinati ima stupnjeve čvorova u novonastaloj mreži i na aplikati vjerojatnost takve promjene kod procesa fragmentacije. Ovisno o početnim parametrima dobiti ćemo drugačiji reljef u tom grafu. Sljedeći graf je napravljen u log-log mjerilu.



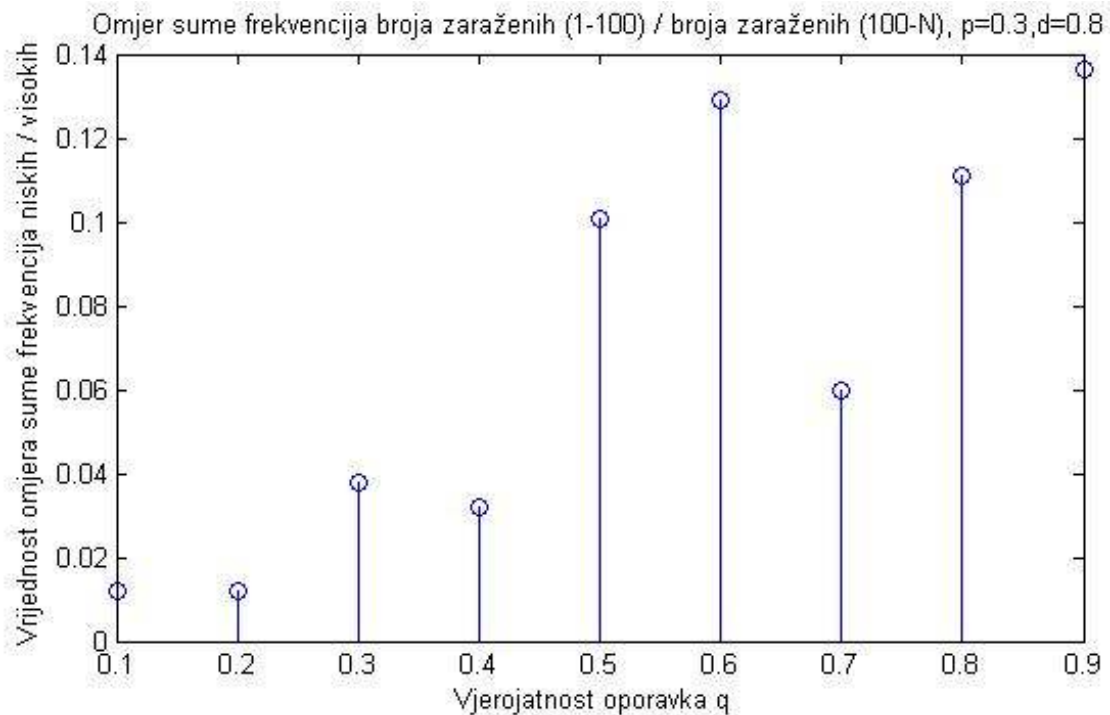
Slika 26. Promjena stupnjeva fragmentacijom mreže, $p = 0.5$, $q = 0.6$ i $d = 0.8$, 167 simulacija, početni čvor id = 31

Na temelju sljedećeg grafa možemo zaključiti da postoji karakteristični hrbat tj. skupina čvorova za koju vrijedi da je promijenila stupanj srednje veličine na mali stupanj, te da postoji karakteristični dijagonalni otoci čvorova koji nisu promijenili stupanj tijekom fragmentacije mreže.



Slika 27. Promjena stupnjeva fragmentacijom mreže, $p = 0.5$, $q = 0.6$ i $d = 0.8$, 167 simulacija, početni čvor $id = 31$

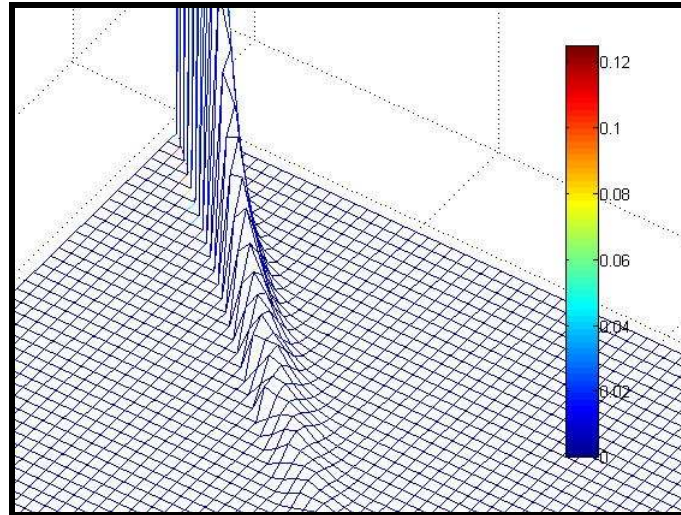
4.2.2. Promjena vjerojatnosti oporavka q



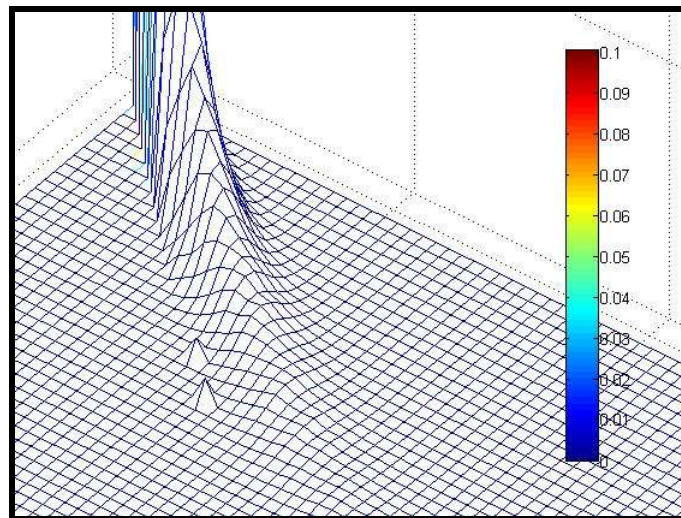
Slika 28.

Ukoliko fiksiramo parametre vjerojatnosti širenja p i vjerojatnosti izbacivanja d na vrijednosti $p = 0.3$ i $d = 0.8$ tada možemo promatrati utjecaj vjerojatnosti oporavka na bimodalnost samog procesa. Parametar q je mijenjan od vrijednosti 0.9 do 0.1 sa razmacima 0.1 i za svaku kombinaciju parametara pokrenuta je simulacija. Mjerali smo omjere suma frekvencija niskog i visokog broja zaraženih. Granicu niskog i visokog broja zaraženih smo proizvoljno odabrali na neki vrijednost koja nam sigurno separira dvije grupe frekvencija kod bimodalnog ponašanja. Na temelju grafa možemo uočiti da ako vrijednost parametra q ide prema nižim vrijednostima omjer frekvencija ima tendenciju pada. Zbog malog broja simulacija (150 simulacija) po svakoj kombinaciji fluktuacije su velike no tendencija pada se sigurno uočava. Vidimo da bimodalni efekt lagano iščezava i pretvara u unimodalni što se približavamo nižim vrijednostima parametra q .

4.2.3. Promjena vjerojatnosti izbacivanja d



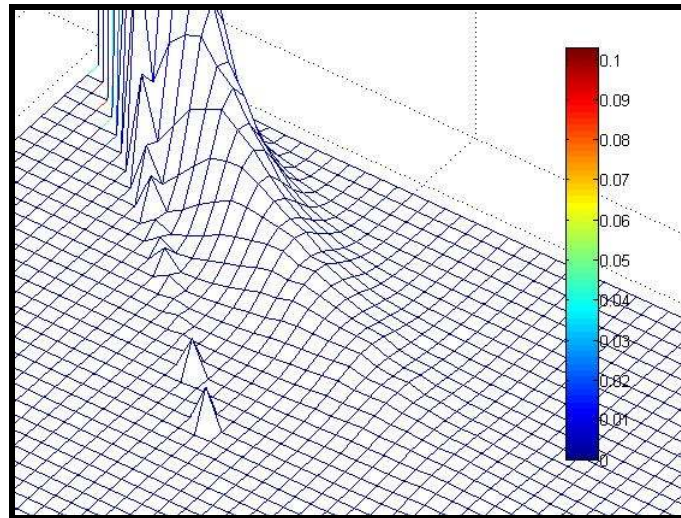
Slika 29. $p = 0.5$, $q = 0.6$, $d = 0.1$, početni čvor $id = 31$



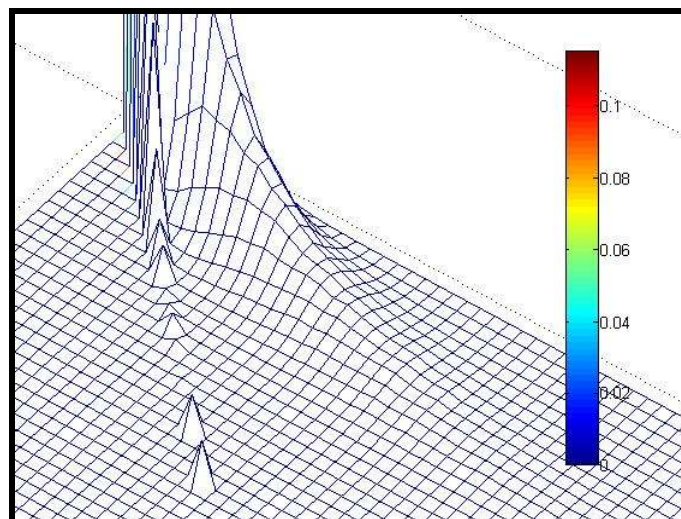
Slika 30. $p = 0.5$, $q = 0.6$, $d = 0.3$, početni čvor $id = 31$

Ukoliko fiksiramo parametre vjerojatnosti širenja zaraze p i vjerojatnosti oporavka q onda dobijemo podgraf G' koji se sastoji od svih čvorova za koje vrijedi : $v \in G \wedge v \in \{skup_zaraženih\}$. Ukoliko pretpostavimo da se skup zaraženih čvorova ne mijenja puno za neovisne simulacije sa istim parametrima p i q onda nam promjena parametra vjerojatnosti izbacivanja d uniformno bira podskup izbačenih čvorova iz skupa zaraženih čvorova. Za vrijednost parametra $d = 0.1$ vidimo da nam ne dolazi do znatne promjene stupnjeva u novonastaloj mreži jer karakteristični hrbat prolazi kroz pravac $f(x) = x$ u x -y

ravnini. Za vrijednost parametra $d = 0.3$ karakteristični hrbat se odmiče od pravca $f(x) = x$ u x - y ravnini što znači da tijekom fragmentacije promjena stupnjeva je mala. Što nam parametar vjerojatnosti izbacivanja čvorova d postaje veći to mi veći podskup čvorova izbacujemo iz mreže i karakteristični hrbat se pomiče od pravca $f(x) = x$ prema pravcu $y = 0$ u x - y ravnini.



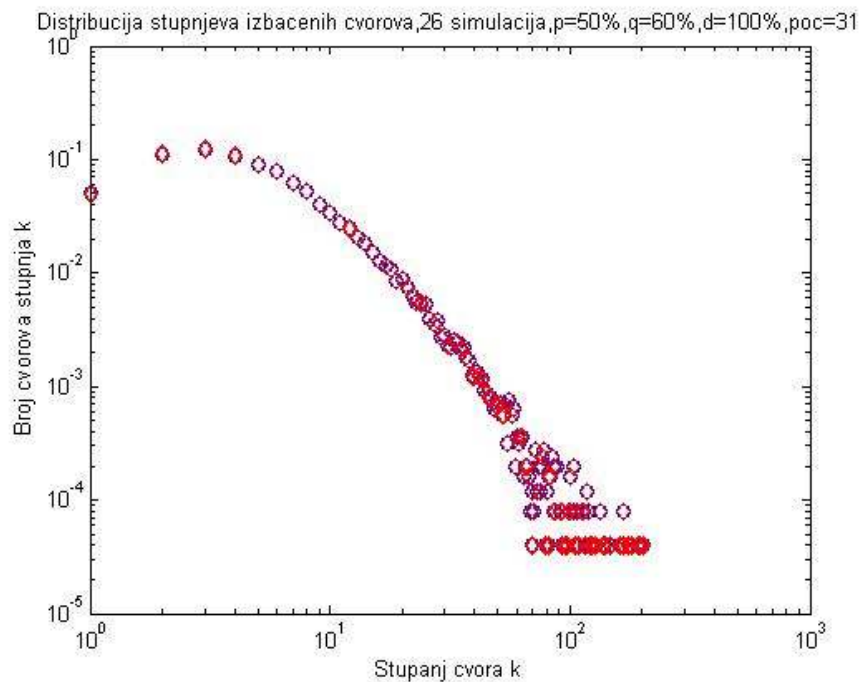
Slika 31. $p = 0.5$, $q = 0.6$, $d = 0.5$, početni čvor $id = 31$



Slika 32. $p = 0.5$, $q = 0.6$, $d = 0.7$, početni čvor $id = 31$

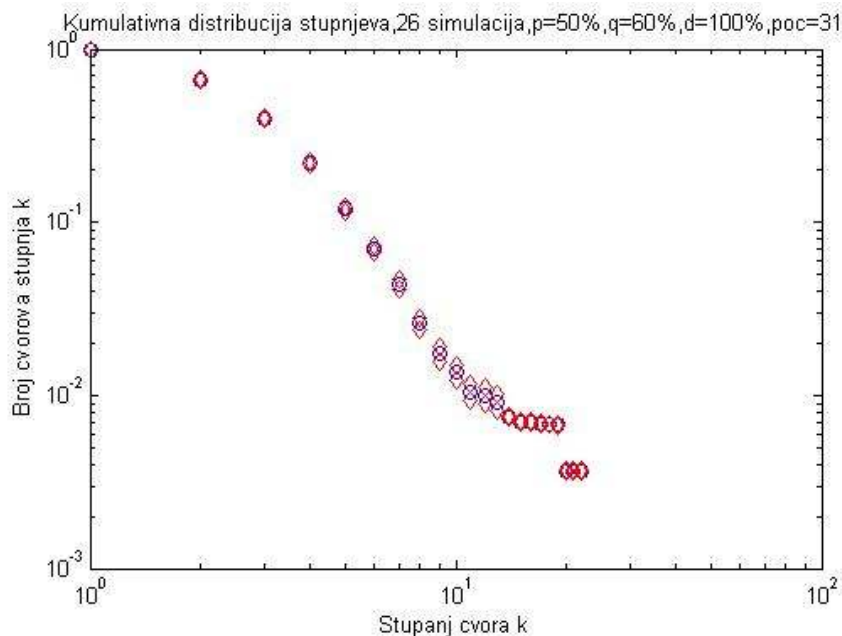
Kako bi mogli objasniti kako parametar vjerojatnost izbacivanja djeluje trebamo pogledati distribuciju zaraženih čvorova drugim riječima gledamo distribuciju izbačenih čvorova za $d = 1$. Distribucija izbačenih čvorova je invarijantna na promjenu parametra d jer on uniformno izbacuje čvorove i distribucija se normira. Dapače distribucija izbačenih čvorova je

invarijantna i na promjenu parametra p i q što znači da nakon bilo kakve epidemije mi znamo kako izgleda distribucija zaraženih čvorova a samim time i distribucija izbačenih čvorova.



Slika 33.

Također je od velike koristi pogledati graf novonastale mreže ako se svi zaraženi izbace iz mreže. Na grafu plavi simboli predstavljaju srednje vrijednosti podataka dok crveni simboli predstavljaju $\bar{X} \pm \sigma$ standardno odstupanje od srednje vrijednosti.



Slika 34.

Zaključak

Temeljem istraživanja utjecaja širenja zaraze na svojstva kompleksne mreže pronađena su jako zanimljiva ponašanja u određenim dijelovima parametarskog prostora. Prvi zanimljiv fenomen je globalno bimodalno ponašanje. Bimodalnost postoji jer postoje dva moguća ishoda proširenja zaraze (1) lokalno u prvom susjedstvu inicijalno zaraženog čvora i (2) globalno proširenje zaraze. Bimodalno ponašanje se mora detaljnije istražiti jer u sadašnjim mjerenjima postoje velike fluktuacije podataka. Fluktuacije podataka su uzrok malog broja simulacija reda veličine 100 simulacija po mjerenju. Kako bi se fluktuacije smanjile potrebno je povećati broj simulacija za barem red veličine. Temeljem sadašnjeg istraživanja utvrđeno je kada nastupa bimodalno a kada unimodalno ponašanje. No granice unimodalnog i bimodalnog ponašanja nisu točno određene. Bimodalnost detektiramo s omjerom frekvencija niskog i visokog broja zaraženih. Višemodalnost procesa širenja epidemije u nekom dijelu parametarskog prostora također nije isključena. Višemodalnost bi se očitovala pojavom više od dva moguća ishoda širenja zaraze. Jedan zamišljeni scenarij bi mogao ovako glasiti: ukoliko bi se na sadašnju gigantsku komponentu sa malim brojem veza spojila nova gigantska komponenta onda bi postojao i treći mogući ishod širenja epidemije. Treći mogući ishod širenja epidemije se događa u onim slučajevima kada se infektivni agent uspije preko jedne nove veze proširiti na novu gigantsku komponentu. Ovakav ishod bi u vremenskoj krivulji broja zaraženih čvorova mogao naglo podići broj zaraženih čvorova što je usko povezano sa valovitim tipom ponašanja nekih epidemija.

Osim bimodalnog ponašanja u određenom dijelu parametarskog prostora (p-q-d) primijećen je prag pri kojem dolazi do značajne fragmentacije gigantske komponente na puno malih komponenti.

Također se na temelju promjene stupnjeva tijekom fragmentacije uočio karakteristični hrbat koji reprezentira određenu skupinu čvorova koji su promijenili svoj stupanj.

Literatura

- [1] Complex Networks, 26.2.2008., http://en.wikipedia.org/wiki/Complex_network, 20.5.2008.
- [2] D. J. Watts and S. H. Strogatz: Collective dynamics of small-world networks, *Nature* 393, 1998.
- [3] Dorogovtsev, S.N., Mendes, J.F.F. The shortest path to complex networks. ArXiv:cond-mat/0404593 v4, 24.7.2004.
- [4] Piškorec, M., Seminar: Kompleksne mreže, Fakultet elektrotehnike i računarstva, 15.5.2007.
- [5] D. J. Watts and S. H. Strogatz: Collective dynamics of small-world networks, *Nature* 393, 1998.
- [6] Clustering coefficient, http://en.wikipedia.org/wiki/Clustering_coefficient, 24.5.2008.
- [7] Betweenness centrality, <http://en.wikipedia.org/wiki/Betweenness#betweenness>, 25.5.2008.
- [8] Assortativity, <http://en.wikipedia.org/wiki/Assortativity>, 25.5.2008.
- [9] Compartmental models in epidemiology, 16.4.2007, http://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology, 17.4.2007.
- [10] Satorras, R., Vespignani, A. Epidemics and immunization in scale-free networks. ArXiv: cond-mat/0205260 v1, 14.5.2002.
- [11] M. E. J. Newman, *Proc. Natl. Acad. Sci. USA* 98, 404-409 (2001).
- [12] Network Data, 13.1.2008., Newman, <http://www-personal.umich.edu/~mejn/netdata/>, 15.2.2008.
- [13] GML: A portable Graph File Format, Michael Himsolt, <http://www.infosun.fim.uni-passau.de/Graphlet/GML/gml-tr.html>, 17.4.2008.
- [14] Large Network Visualization tool, <http://xavier.informatics.indiana.edu/lanet-vi/>, 1.6.2008.],
- [15] Lev Muchnik, Royi Itzhack, Sorin Solomon, and Yoram Louzoun, Self-emergence of knowledge trees: Extraction of the Wikipedia hierarchies, *Phys. Rev. E* 76, 016106 (2007), <http://www.levmuchnik.net/Content/Networks/ComplexNetworksPackage.html>
- [16] David Gleich, Matlab Boost Graph Library, http://www.stanford.edu/~dgleich/programs/matlab_bgl/, 17.5.2008.
- [17] Matlab Distributed Computing Toolbox, http://www.mathworks.com/products/distribtb/index.html?s_cid=HP_FP_ML_ParallelComputingToolbox, 13.3.2008.
- [18] Antulov-Fantulin, Nino, Seminar : Analiza propagacije virusa u mrežama, Fakultet elektrotehnike i računarstva, svibanj. 2007.

- [19] Dorogovtsev, S.N., Mendes, J.F.F. Evolution of Networks: From Biological Nets to the Internet and WWW. Oxford 2003.
- [20] David Smith and Lang Moore, The SIR Model for Spread of Disease, 22.9.2000, <http://www.math.duke.edu/education/ccp/materials/diffcalc/sir/index.html>, 5.2.2008
- [21] Colizza Vittoria, Barrat Alain, Barthelemy Marc, Valleron Alain-Jacques, Vespignani Alessandro. Modeling the Worldwide Spread of Pandemic Influenza: Baseline Case and Containment Interventions. Plos Medicine | Volume 4 | Issue 1 | e13, sječanj 2007
- [22] Guimera, R., Mossa, S., Turtschi, A., Amaral, L. A. N. The world-wide air transportation network: Anomalous centrality, community structure, and cities' global roles. arXiv:cond-mat/0312535 v2, 12.7.2005

Sažetak

Svaka kompleksna mreža ima određena strukturalna svojstva koja odgovaraju njenoj funkciji. Mjere centralnosti čvorova u mreži (stupanj čvora, međupoloženost, koeficijent grupiranja,...) nam pokušavaju opisati relativnu važnost čvora u cjelokupnoj mreži, dok nam neka globalna svojstva kao npr. distribucija stupnjeva, distribucija koeficijenta grupiranja, koeficijent asortativnosti više govore o funkciji cjelokupne mreže. Cilj ovog istraživanja je bio otkriti kako sama epidemija mijenja strukturalna svojstva mreže po kojoj se širi. Kao model širenja zaraze izabran je SIR model koji je proširen s još jednim odjeljkom izbačeni tako da je vjerojatnost prijelaza iz odjeljka oporavljeni u odjeljak izbačeni d .

Simulacije dinamike širenja zaraze na kompleksnoj mreži ostvarene su u Matlabu koristeći računalni grozd (engl. *cluster*) za distribuirano računanje. Kao polaznu mrežu nad kojom su se simulacije računale izabrana je mreža kolaboracije znanstvenika veličine 31 163 čvorova. Mreža je iz GML formata pretvorena u matricu susjedstva koja služi kao osnovna struktura podataka nad kojom se simulacije računaju.

Promjena svojstava promatrana je sa raznih dijelova parametarskog prostora ulaznih parametara. Ulazni parametri predstavljaju uređenu četvorku (p, q, d, poc) , gdje parametar p predstavlja vjerojatnost širenja zaraze, parametar q vjerojatnost oporavka zaraženog čvora, parametar d vjerojatnost izbacivanja čvora, parametar poc predstavlja vektora inicijalno zaraženih čvorova. Prvo smo promatrali promjenu parametra p dok smo parametre q i d fiksirali. Uočen je prag p iznad kojeg dolazni do znatne fragmentacije gigantske komponente na puno manjih komponenti. Također je uočeno bimodalno ponašanje za određene parametre p . Bimodalno ponašanje se očituje kroz dva različita ishoda širenja zaraze: (1) zaraza se globalno proširila ili (2) zaraza se samo lokalno proširila na nekolicinu susjednih čvorova. Ukoliko vrijednost parametra p dovedemo dovoljno visoko bimodalno ponašanje prelazi u unimodalno tj. ostvaruje se prvi ishod. Ukoliko je parametar p dovoljno nizak bimodalno ponašanje se očituje. Zatim smo fiksirali parametre p i d i mijenjali parametar q . Također uočavamo bimodalno ponašanje u određenom dijelu parametarskog prostora.

Abstract

The main goal of this research was to examine how does the spread of the disease change the structural properties of the networks to which it spreads. In order to simulate the spread of the disease on complex networks we have taken basic SIR model and have introduced new parameter d . Parameter d tells us what the probability is of transition from recovered compartment to discarded compartment. Structural properties of networks change only if discarded compartment is not empty due to the fact that all nodes from discarded compartment are removed from the network.

All numerical simulations were done in Matlab using parallel computing on computer cluster. The spread of the disease took place on the network of co-authorships between scientists (31 163 nodes).

Input parameters in our model were: probability of infection p , probability of recover q , probability of discard d and initial infected nodes. These parameters together create a four-dimensional space of input conditions. As different simulations started from different points in our four-dimensional space, different interesting outcomes appeared.

We noticed bimodal behavior in some subspaces of our initial parameter space. Bimodal behavior exhibits through two different outcomes in spread of disease. The first outcome suggests that disease infected just few nodes in local neighborhood and the second one suggests major outbreak of disease.

Also we noticed some thresholds of initial parameters when network fragmentation of the giant component occurs.

Ključne riječi na hrvatskom: kompleksne mreže, simuliranje epidemije, bimodalan proces, fragmentacija mreže.

Ključne riječi na engleskom: *complex networks, epidemic simulation, bimodal process, network fragmentation.*