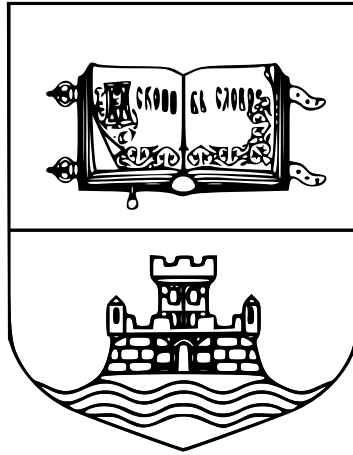




UNIVERZITET U BEOGRADU

Master akademske studije pri Univerzitetu

Računarstvo u društvenim naukama

MASTER RAD

Primena programskog jezika R u računarskoj analizi društvenih mreža

Primer izbora u Srbiji 2016

Autor:

msr Marko Galjak

Mentor:

prof. dr Jelena Jovanović

Beograd, januar 2017. godine

SADRŽAJ

1	Uvod.....	5
1.1	Definicija predmeta rada i cilja istraživanja.....	5
1.2	Pregled rada po poglavljima.....	5
1.3	Pregled relevantnih oblasti.....	6
1.3.1	Analiza društvenih mreža.....	6
1.3.1.1	Osnovni pojmovi u analizi društvenih mreža.....	8
1.3.2	Programski jezik R.....	12
1.3.2.1	R i podrška za analizu društvenih mreža.....	13
1.4	Motivacija.....	16
1.4.1	Izbori u Srbiji.....	16
1.4.2	Analiza društvenih mreža na Tviteru.....	16
1.4.2.1	Dosadašnja istraživanja u Srbiji.....	16
1.4.2.2	Dosadašnja istraživanja u svetu.....	17
1.5	Izbor softvera.....	17
1.6	Polazne hipoteze.....	20
2	Podaci i metode.....	21
2.1	Podaci.....	21
2.2	Metode.....	22
2.3	R paketi i replikabilnost.....	24
3	Postupak analize u R-u i rezultati.....	26
3.1	Akvizicija podataka.....	26
3.1.1	R i Tviter API.....	26
3.1.2	Akvizicija podataka osnovnog uzorka.....	27
3.1.3	Kreiranje proširenog uzorka.....	29
3.1.4	Akvizicija dodatnih podataka o Tviter naložima.....	30

3.2	Sređivanje podataka.....	31
3.3	Kreiranje grafova.....	33
3.4	Vizualizacija grafova.....	34
3.5	Izračunavanje mrežnih metrika.....	40
3.5.1	Globalne metrike mreža.....	40
3.5.2	Metrike centralnosti čvorova.....	42
3.5.3	Homofilija.....	46
3.6	Klaster analiza.....	50
3.6.1	Izbor klaster algoritma.....	50
3.6.2	Vizualizacija klastera.....	54
4	Diskusija.....	57
4.1	Uzorak.....	57
4.2	Rezultati analize.....	57
4.2.1	Karakteristike analiziranih mreža.....	57
4.2.2	Centralnost čvorova i centralizacija u mrežama.....	59
4.2.3	Međuzavisnost pokazatelja i kompozitni pokazatelj centralnosti.....	61
4.2.4	Homofilija i klaster analiza.....	63
4.3	Vizualizacija mreža.....	65
4.4	Primena R-a.....	67
4.4.1	Podrška za analizu društvenih mreža.....	67
4.4.2	Performanse.....	68
4.4.3	Replikabilnost koda.....	68
4.5	Etička razmatranja.....	69
4.5.1	Privatnost.....	69
4.5.2	Klasifikacija naloga.....	69
5	Zaključak.....	70
5.1	Analiza izbora na Tviteru.....	70
5.2	Preporuke analiziranim strankama.....	71

5.3	Mogući pravci budućih istraživanja.....	73
5.3.1	Ponovljena analiza.....	73
5.3.2	Analiza retvitova i detekcija 'botova'.....	73
5.3.3	Analiza sadržaja prikupljenih tvitova.....	74
6	Literatura.....	75
7	Prilozi.....	85
7.1	Prilog I.....	85
7.2	Prilog II.....	86
7.3	Prilog III.....	87
7.4	Prilog IV.....	89

1 UVOD

1.1 Definicija predmeta rada i cilja istraživanja

Predmet rada je društvena mreža koju su činili politički akteri na Tviteru (*Twitter*) tokom kampanje za parlamentarne izbore u Srbiji 2016. Ta mreža je matematički apstrahovana na sledeći način: Tviter nalozi političara i samih stranaka predstavljaju čvorove, dok su ivice mreže njihove međusobne interakcije na Tviteru. Tako definisana društvena mreža je predmet svih analiza u radu, ali i sam proces analize je takođe predmet rada. Analiza je sprovedena u statističkom programskom jeziku R. Razlog za dvojakost predmeta rada proističe iz činjenice da je i cilj rada dvojak:

- sociopolitikološki, da se analizira interakcija političkih aktera na društvenoj mreži Tviter u periodu koji prethodi izborima 2016. u Srbiji.
- metodološko–didaktički, da demonstrira upotrebu programskog jezika R u računarskoj analizi društvenih mreža i omogućiti replikabilnost analize u drugim kontekstima.

Analiza interakcija političara na Tviteru je važna zbog rastuće važnosti ove društvene mreže u javnom diskursu. Konkretno, mrežna analiza može da ponudi nove uvide, koji se ne mogu dobiti analizom sadržaja ili primenom konvencionalnih metrika karakterističnih za društvenu mrežu Tviter (npr. broj pratilaca i ritvitova (*retweets*)). Saznanja dobijena ovom analizom mogu biti od praktične koristi političkim akterima, kako bi otkrili nedostatke dosadašnje komunikacije na Tviteru i eventualno je popravili, ali i naučnoj i široj javnosti.

Drugi cilj je da se istraživačima u oblasti društvenih nauka približi primena najnovijih softverskih alata za analizu društvenih mreža. Uzevši u obzir da se trenutno veliki pomaci u mnogim društvenim naukama prave zahvaljujući sve većem obimu dostupnih podataka i njihovoj analizi, postoji i rastuća potreba za upotrebom informatičkih alata. Finalni rezultat master rada biće softver, ali i praktične smernice za korišćenje jezika R koje će omogućiti primenu predstavljene metodologije za analizu društvenih mreža u različitim kontekstima.

1.2 Pregled rada po poglavljima

Master rad ima pet poglavlja. Prvo, uvodno poglavlje sadrži definicije predmeta i cilja istraživanja, pregled oblasti relevantnih sa samo istraživanje. Osim toga u uvodnom poglavlju

data je i motivacija za izbor teme rada kao i pregled domaće i strane literature a dato je i objašnjenje razloga za izbor programskog jezika R kao softverskog alata za analizu društvenih mreža. Na kraju uvodnog poglavlja date su polazne hipoteza istraživanja. Drugo poglavlje objašnjava podatke i pokazatelje korišćene u radu. U istom poglavlju date su smernice za reprodukciju istraživanja, uz pomoć suplementarnih datoteka koje dolaze uz elektronsku verziju rada. U trećem poglavlju demonstrira se upotreba R-a za akviziciju podataka, njihovu analizu i vizualizaciju. U istom poglavlju paralelno su dati rezultati analiza. Diskusija na temu dobijenih rezultata data je u zasebnom, četvrtom poglavlju. Osim diskusije o rezultatima u istom poglavlju data je i kratka diskusija o primeni R-a. Zaključno, peto poglavlje sadrži tri dela: *i)* zaključna razmatranja same analize, *ii)* praktične preporuke analiziranim akterima izvedene iz rezultata istraživanja, u cilju maksimizacije uticaja na online društvenoj mreži Tviter, i *iii)* moguće pravce budućih istraživanja.

1.3 Pregled relevantnih oblasti

1.3.1 Analiza društvenih mreža

Analiza društvenih mreža je pristup proučavanju društvenih struktura upotrebom teorija grafova i mreža (Otte & Rousseau, 2002). U njoj se matematičkom analizom grafova, koji reprezentuju neku društvu mrežu, izvlače zaključci o društvenoj mreži koju graf apstrahuje. Graf je način predstavljanja odnosa između skupa entiteta koji se sastoji se iz čvorova koji predstavljaju entitete, i ivica koje povezuju čvorove i predstavljaju odnos između entiteta (Easley & Kleinberg, 2010).

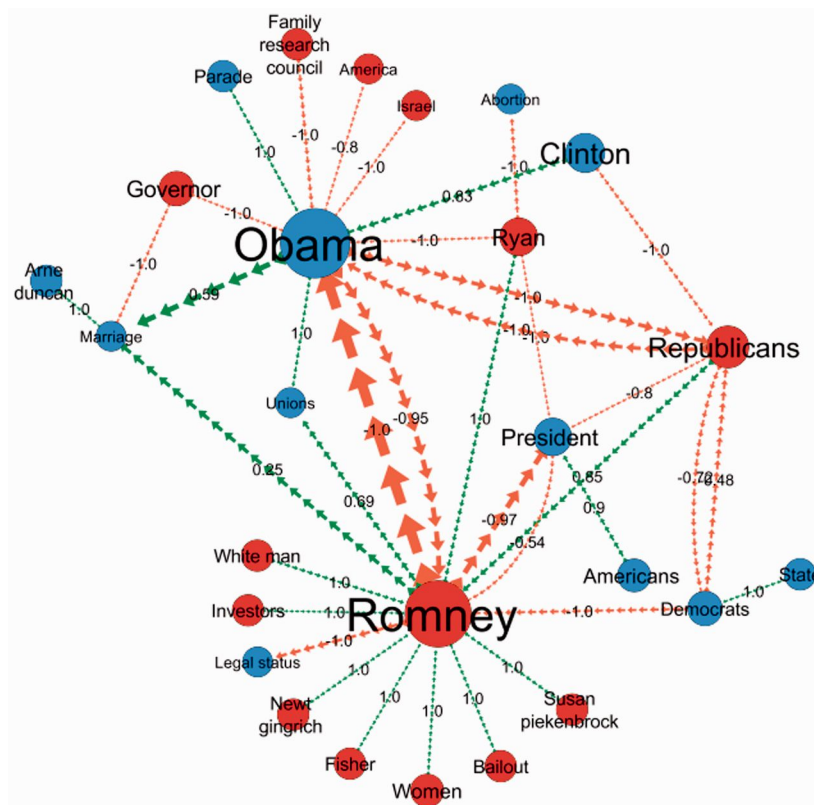
Istorija istraživanja i razmišljanja na temu društvenih mreža je relativno nova. Prva razmišljanja o društvenom životu iz strukturalne perspektive predložio je u XIX veku Ogist Kont (Freeman, 2004). Prema Frimenu, teorijski okvir nastavlja da se razvija preko Dirkema, le Bona, Zimela i fon Vizea. Kada su u pitanju sistematski podaci društvenih struktura, Frimen ističe da su prirodnjaci još pre Konta, prikupljali i obrađivali takve podatke. On navodi da je Pijer Uber, švajcarski entomolog rođen 1777, prikupljao podatke od društvenoj strukturi kolonija pčela i mrava da bi prikupljanje takvih podataka o ljudskim strukturama otpočelo tek pola veka kasnije sa Luisom Henrijem Morganom i da je do danas nastavljalo da se razvija, preko Hobsona, Almaka, Morena, Velmana, Bota i Hagmana. Frimen takođe ističe da, iako su razmišlja na ovu temu relativno nova pojava, korišćenje grafičkog prikaza pomoću čvorova i veza počinje da se koristi još u doba Rimljana za potrebe genealogije.

Osnove teorije grafa u matematici postavio je Leonard Ojler 1736. god. pokušavajući da reši logički problem „Kenizberških mostova“ (Alexanderson, 2006). Od tada do danas mnogi slavni matematičari bavili su se ovom problematikom, a među njima su: Ogisten Luj Koši, Deneš Kening, Pal Erdeš, Alfred Renji i dr.

Danas analiza društvenih mreža nalazi primenu u mnogim naukama. Ova oblast, dakle, nije u vezi samo sa informatikom, biologijom, sociologijom, psihologijom, antropologijom, političkim naukama već je našla primenu i u ekonomiji i menadžmentu, geografiji, izučavanju životne sredine itd. Metodologija analize društvenih mreža pripada širem pojmu društvene fizike, koja je tokom poslednjih godina dobila na popularnosti, zahvaljujući sve većoj dostupnosti podataka i rastućim kapacitetima za njihovu obradu (Pentland, 2014).

Oblast analize društvenih mreža ima mnoge praktične primene jer se mnogi sistemi mogu modelovati kao društvene mreže i tako analizirati. Osim primene u mnogim naukama, sve je učestalija primena od strane vladinih i nevladinih agencija za rešavanje različitih tipova problema. Od lokalnih vlada, i dolaženja do saznanja koja se koriste za urbano planiranje (Dempwolf & Lyles, 2012), do kriminalističke analize društvenih mreža osoba koje se sumnjiče za krivična dela i sprečavanje terorizma (Perliger & Pedahzur, 2011). U menadžmentu, razvojem poslovne inteligencije, i spoznajom vrednosti informacija od strane kompanija, analiza društvenih mreža je često deo sistema korporativnog odlučivanja (Hoppe & Reinelt, 2010).

Vizualizacija društvenih mreža je važan deo analize koji nam pomaže da intuitivno primetimo ono što nam metrike neke mreže kvantitativno pokazuju. Takva vizualizacija se naziva sociogram (Slika 1), ali se u praksi za vizualizacije često koristi termin graf, iako pojam grafa, iz diskretne matematike, predstavlja apstrakciju, a ne samo vizualizaciju.



Slika 1 Primer vizualizacije društvene mreže, predsedničke kampanje SAD 2012. (Sudhahar, Veltri, & Cristianini, 2015). Plavi čvorovi su pripadnici Demokratske stranke a crveni Republikanske. Pozitivna komunikacija označena je zelenom bojom, a negativna narandžastom.

1.3.1.1 Osnovni pojmovi u analizi društvenih mreža

Svaka društvena mreža je sačinjena od **čvorova** (*nodes*), koji predstavljaju subjekte mreže i **ivica** (*edges*), tj. veza koje te subjekte povezuju. Svaki čvor ima svoj **stepen** (*degree*) – broj čvorova sa kojima deli ivicu. Na primer, društvena mreža Fejsbuka (*Facebook*) se može apstrahovati na sledeći način: osobe su čvorovi, prijateljstva su ivice, a stepen je broj prijatelja koje neka osoba ima. Fejsbuk prijateljstva su primer **neusmerene mreže**, jer je prijateljstvo međusobno. Postoje mreže kod kojih veza između dva čvora može imati smer. Primer **usmerene mreže** u okviru online društvenih mreža bio bi sistem pratilaca na Tviteru u kom osoba A može pratiti osobu B, ali osoba B ne mora pratiti osobu A.

Putanja (*path*) je niz čvorova u mreži u kojem je svaki uzastopni par (čvorova u tom nizu) povezan ivicom (Easley & Kleinberg, 2010). U primeru društvene mreže na Fejsbuku, putanja između osobe A i osobe D koje međusobno ne dele ivicu (nisu prijatelji), bi išla preko B i C koji su međusobno prijatelji, B je prijatelj sa A, a C je prijatelj sa D (A–B–C–D). Važno je napomenuti da se u jednoj putanji ne mogu ponavljati čvorovi i ivice. Svaka putanja ima svoju **dužinu** koja se izražava u broju ivica koje putanja sadrži (kod nepovezanih čvorova dužina putanje je beskonačna). U slučaju mreže sa Fejsbuka dužina putanje (A–B–C–D) je 3 i to je

jedina moguća putanja između čvorova A i D u tako definisanoj mreži, međutim u realnim mrežama može da postoji mnogo putanja različitih dužina koje povezuju dva čvora mreže. Zato je razvijen koncept **geodezika** (*geodesic*) tj. najkraće putanje ili, putanje između dva čvora sa najmanjom dužinom.

Najmanja mrežna struktura u kojoj čvor može da učestvuje je **dijada** (jedan par povezanih čvorova). Struktura koju čine tri čvora naziva se **trijada**. Posebno mesto u izučavanju društvenih mreža zauzima koncept **trijadskog zatvaranja** (*triadic closure*), koji predstavlja sklonost čvora A koji ima jake veze sa čvorom B, da se poveže sa čvorom C koji takođe ima jake veze sa čvorom B (Simmel & Wolff, 1964).

Da bi se različite društvene mreže međusobno mogle porediti i klasifikovati prema karakteristikama koje ispoljavaju, razvijeni su brojni sumarni statistički pokazatelji koji nude značajan uvid u strukturu mreže (Jackson, 2008). Prema Jackson-u (2008), mreže se sumarno mogu opisati pomoću pokazatelja:

- distribucije stepena
- prečnika, prosečne dužine putanje i gustine
- sklonosti stvaranja klika, koheziji i grupisanju
- centralnosti

Distribucija stepena (*degree distribution*) je distribucija učestalosti čvorova sa različitom vrednošću stepena (Jackson, 2008), odnosno u pitanju je raspodela verovatnoća čvorova sa različitim vrednošću stepena. Nasumično generisani grafovi (*random graphs*) prate Puasonovu raspodelu, dok mreže u realnom svetu (npr. online društvene mreže) prate *power law* raspodelu (Paretovu raspodelu) (Albert & Barabási, 2002). Ovu raspodelu karakteriše dugačak 'rep', tako da mali broj čvorova ima (izrazito) veliku vrednost stepena, dok veliki broj čvorova ima mali stepen. Mnoge prirodne pojave prate ovu distribuciju uključujući gradove po broju stanovnika, učestalost reči u nekom jeziku, jačine zemljotresa, itd.

Prečnik ili dijametar (*diameter*) jedne mreže predstavlja najdužu putanju u toj mreži. **Prosečna dužina putanje**, poznata i pod nazivom karakteristična dužina putanje, predstavlja prosečnu dužinu geodezika između svaka dva čvora u mreži. Mala vrednost ove metrike je jedna od ključnih karakteristika fenomena malog sveta (Watts & Strogatz, 1998). Na primer, prosečna dužina putanje na Fejsbuku je 3,57 (Edunov, Diuk, Filiz, Bhagat, & Burke, 2016).

Gustina (*density*) mreže je relativni udeo prisutnih ivica mreže, i izračunava se kao količnik broja postojećih ivica i broja svih potencijalnih ivica mreže. U primeru mreže A–B–C–D postoje 4 ivice, a broj potencijalnih ivica je 6 (dobijeno formulom $\frac{n*(n-1)}{2}$), pa je gustina mreže 0,67, što znači da je ta mreža prilično gusta. Online društvene mreže karakteriše mala gustina (*sparseness*), npr. prosečna gustina mreže prijatelja na Fejsbuku je 0,12, dok je u stvarnosti gustina kod *offline* prijateljstava 0,36 (Hampton, Goulet, Marlow, & Rainie, 2012).

Povezanost (*connectivity*) ili kohezija mreže je mera koja predstavlja najmanji broj čvorova koje je potrebno ukloniti iz mreže da bi mreža postala nepovezana (Csardi & Nepusz, 2006; White & Harary, 2001). Kompletно povezana mreža je ona kod koje su svi čvorovi koji je čine međusobno povezani. Nepovezana mreža je ona kod koje ne postoji putanja između bilo koja dva čvora u mreži. Kod nepovezanih grafova vrednost mere povezanosti je 0.

Česta zajednička osobina društvenih mreža je pojava klika unutar mreže (Albert & Barabási, 2002). **Klika** je podskup čvorova mreže (tj. podgraf) u kojem su svi čvorovi međusobno povezani, uz uslov da klike moraju da sadrže barem tri čvora da bi se uopšte smatrale klikama, a jedan čvor može da pripada većem broju klika (Easley & Kleinberg, 2010). Jackson (2008) navodi da je najjednostavniji način za ocenu sklonosti stvaranja klika (*cliquishness*) u jednoj mreži prebrojavanje klika i čvorova koji ih čine (što može biti izraženo relativnim ili apsolutnim vrednostima). On ističe još jedan način za ocenu, barem nekih od aspekata, ove sklonosti preko **tranzitivnosti** trojki ili koeficijenta **grupisanja**, koji predstavlja verovatnoću da su čvorovi, povezani sa jednim zajedničkim čvorom, takođe međusobno povezani (Csardi & Nepusz, 2006).

Homofilija je pojava koja se odnosi na sklonost ljudi da održavaju veze sa ljudima koji su im slični (Jackson, 2008). Ova pojava ima za posledicu to da skup prijatelja jedne osobe ne oslikava dobro slučajan uzorak populacije (Easley & Kleinberg, 2010). Ova pojava se odražava u izreci „svaka ptica svome jatu leti“ (*birds of a feather flock together*). Homofilija se može meriti različitim metrikama, a jedna od njih je i **koeficijent asortativnosti** (*assortativity*) baziran na dodatnim atributima ili vrednostima koje dodeljujemo čvorovima. Ako je koeficijent visok, to znači da povezani čvorovi pretežno dele iste attribute ili vrednosti koje su im dodeljene (Csardi & Nepusz, 2006). U praksi se najčešće koristi koeficijent asortativnosti prema stepenu čvorova kada on predstavlja Pirsonov koeficijent korelacije između stepena povezanih čvorova (Newman, 2002). Međutim koeficijent asortativnosti se može računati i za kategorijske promenljive. Za mreže sa koeficijentom asortativnosti 1 se može reći da su u potpunosti

homofilične, a kada je ta vrednost -1 mreža je u potpunosti heterofilična (čvorovi se povezuju isključivo sa drugačijim čvorovima).

Veliki deo istraživanja mreža posvećen je davanju odgovora na pitanje: „Koji su čvorovi najvažniji ili najcentralniji u mreži?“ (Newman, 2010). Odgovor na ovo pitanje daju mere **centralnosti**. Osim lokalne mere centralnosti, koja kvantifikuje centralnost svakog čvora u mreži, postoje i mere globalne centralnosti ili **centralizacije** (Scott, 2013). Centralizacija odgovara na pitanje: „U kojoj meri je struktura mreže centralizovana?“. Opšta procedura za izračunavanje metrika centralizacije mreže svodi se na pronalaženje razlike između centralnosti najcentralnije tačke i centralnosti svih drugih tačaka; centralizacija je zapravo proporcija sume tih razlika i maksimalne moguće sume razlika (Scott, 2013). Mreža sa metrikom centralizacije 1 je u potpunosti centralizovana dok je ona sa metrikom 0 u potpunosti decentralizovana. Centralnost pa samim tim i centralizacija, mogu se operacionalizovati na više načina, a u zavisnosti od potrebe analize. Najčešće mere centralnosti su centralnosti: stepena, bliskosti, intermedijarnosti i svojstvenog vektora.

Centralnost stepena (*degree centrality*) je najjednostavniji način za procenu centralnosti čvorova u mreži, i određuje se preko stepena, tj. broja veza koje čvorovi uspostavljaju. Često se izražava normalizovano, kao količnik stepena čvora i ukupnog broja mogućih veza ($n-1$) i tada se kreće se u rasponu od 0 do 1. Kod usmerenih mreža, osim ukupnog stepena, može se izračunati stepen samo dolazećih ili samo odlazećih veza (*indegree* i *outdegree*). Najveći nedostatak ove metrike centralnosti je zanemarivanje strukture mreže, tako da npr. neki čvor koji ima centralno mesto u mreži, bez kog bi se mreža raspala, može imati veoma mali stepen (Jackson, 2008).

Centralnost bliskosti (*closeness centrality*) meri koliko je neki čvor blizu ostalim čvorovima, tj. kolika je dužina putanje između njega i svakog drugog čvora u mreži. Jedan od očiglednijih načina za izračunavanje bliskosti je recipročna vrednost prosečne udaljenosti između tog čvora i svih drugih čvorova (postoje različite konvencije za izračunavanje bliskosti nepovezanih čvorova)(Jackson, 2008). Visoke vrednosti bliskosti su tipične za čvorove koji se nalaze blizu centara lokalnih grupa u većoj mreži. Kao i kod centralnosti stepena, kod usmerenih mreža postoje dolazeća i odlazeća varijanta (*in-closeness* i *out-closeness*).

Centralnost intermedijarnosti (*betweenness centrality*) definiše centralnost kao meru u kojoj se čvor nalazi na najkraćoj putanji između bilo koja druga dva čvora mreže (Freeman, 1977). Za dva proizvoljna čvora mreže, određuje se broj geodezika koji prolaze kroz čvor za koji se ova

centralnost računa i deli se sa ukupnim brojem geodezika između ta dva čvora. Tako dobijene vrednosti se sumiraju za sve parove čvorova u mreži. Ova mera omogućava identifikaciju čvorova koji posreduju u komunikaciji između drugih čvorova u mreži. Ti čvorovi često nemaju centralnu poziciju u lokalnim grupama kojima pripadaju unutar mreže, već se nalaze na periferiji više grupa, jer tako povezuju različite grupe, koje bi možda bile nepovezane kada bi se čvorovi sa visokom centralnošću intermedijarnosti uklonili.

Centralnost svojstvenog vektora (*eigenvector centrality*) je mera centralnosti kod koje je jedinica centralnosti zbir ivica ka drugim čvorovima, ponderisan centralnostima tih čvorova (Bonacich, 1987). Drugim rečima, ova mera centralnosti nam omogućuje da odredimo da li je neki čvor u mreži povezan sa uticajnim članovima mreže. Newman (2010) ističe da činjenica da je zbir centralnosti jednog čvora proporcionalan centralnosti njemu susednih čvorova, daje centralnosti svojstvenog vektora dobru karakteristiku da čvor može imati visoku vrednost ili zbog toga što je povezan sa mnogo čvorova ili zato što je povezan sa važnim čvorovima (ili i jedno i drugo). Newman dalje objašnjava da osoba u društvenoj mreži može, na primer, biti važna zato što poznaje mnogo ljudi (iako možda ti ljudi nisu previše važni u mreži) ili zato što poznaje malo ljudi sa velikim uticajem. Interesantno je da je ovaj princip centralnosti našao praktičnu primenu u rangiranju rezultata veb pretraživača: veb strana ima viši *PageRank* ako se na nju linkuju druge stranice sa visokim *PageRankom* (Brin & Page, 1998).

Matrica povezanosti je matematički način predstavljanja konačnog grafa. U pitanju je kvadratna matrica (broj kolona jednak broju redova) u kojoj redni broj kolone i reda odgovara jednom čvoru u mreži. Element matrice može biti 1 ili 0 kada je u pitanju Bulova matrica, u kom slučaju se samo označava da postoji ivica između dva čvora, ili neki drugi broj koji označava težinu ivice između dva čvora. Kod usmerenih grafova, matrica povezanosti je simetrična. Mnogi softverski alati kao ulazne podatke primaju upravo matricu povezanosti.

1.3.2 Programski jezik R

R (izgovara se „ar“) je programski jezik koji se koristi za statističke proračune i kreaciju grafičkih prikaza podataka (Ihaka & Gentleman, 1996). R je zapravo sistem koji u sebi sadrži programski jezik, izvršno okruženje sa grafičkom podrškom, debugger, pristup određenim sistemskim funkcijama i sposobnost da pokreće programe sačuvane u skriptima (R Core Team, 2016). R je nastao 1993, kada su ga Ros Ihaka i Robert Džentlmen osmisli kao naslednika jezika S (Morandat, Hill, Osvald, & Vitek, 2012). Jezik S je napravljen u okviru Bell Labs kompanije još sredinom sedamdesetih i dobio je ime po uzoru na druge programske jezike koji su u isto

vreme nastajali (kao npr. C), s obzirom da je u pitanju statistički programski jezik izabrano je ime S. Ime R je verovatno izabrano jer je to početno slovo imena autora. Postoji dosta sličnosti između programskih jezika S i R, pa je kod napisan za jedan programski jezik u velikoj meri kompatibilan sa drugim (R Core Team, 2016). Za razliku od programskog jezika S, koji je intelektualno vlasništvo kompanije koja ga je razvila, R je izdat pod GNU GPL licencom (General Public License), te je stoga potpuno otvoren i besplatan za nekomercijalnu i komercijalnu upotrebu.

Ovaj objektno orijentisan programski jezik je od svog nastanka do danas postao značajan alat u mnogim naukama, pre svega zahvaljujući svojoj otvorenosti, stalnom unapređivanju, kako osnovnog koda, tako i brojnih paketa. Danas ogromna zajednica R korisnika broji milione članova, a trenutni broj paketa koji su deo zvaničnog CRAN-a (The Comprehensive R Archive Network) je 9437 (CRAN, 2016). Među tim paketima su oni koji se koriste za opštu statističku analizu podataka i *data mining*, oni koji se koriste u matematici, fizici, biologiji (naročito biostatistici i genomics), hemiji i drugim srodnim disciplinama. Takođe, postoji veliki broj paketa namenjenih društvenim naukama kao što su psihologija, sociologija, demografija, lingvistika i dr.

1.3.2.1 R i podrška za analizu društvenih mreža

Moć R jezika se ogleda u mnogobrojnim paketima, tj. bibliotekama, koje umnogome olakšavaju različite vrste analiza podataka, među kojima i analizu društvenih mreža. Trenutno postoji više R paketa koji su namenjeni analizi i vizualizaciji društvenih mreža, među kojima je i *igraph* paket koji će se najviše koristiti u ovom radu.

U daljem tekstu su predstavljeni neki od R paketa namenjenih analizi društvenih mreža koji su deo zvaničnog CRAN repozitorijuma. Osim njih postoji mnoštvo različitih implementacija koje nisu deo zvaničnog CRAN repozitorijuma, a mogu se naći u repozitorijumima zajednica kao što je GitHub.

igraph je R paket i biblioteka (takođe dostupna za Python i C) čiji je glavni cilj da pruži skup tipova podataka i funkcija za: 1) jednostavnu implementaciju algoritama za grafove, 2) brzo rukovanje velikim grafovima, odnosno grafovima sa milionima čvorova i ivica, 3) brzo kreiranje prototipa (Csardi & Nepusz, 2006).

sna je R paket koji sadrži brojne alate za analizu društvenih mreža, uključujući i mogućnost indeksiranja na nivou čvora ili grafa, metode strukturalnog odstojanja i metode kovarijanse,

detekciju strukturalnih jednakosti, mrežnu regresiju, nasumično generisanje grafova, kao i 2D i 3D vizualizaciju (Butts, 2016).

statnet (<http://www.statnet.org/>) uključuje mnoštvo različitih paketa koji zajedno pružaju mogućnost ne samo analize društvenih mreža, već i stvaranje nasumičnih grafova i brojnih simulacija. Analitički okvir baziran je na eksponencijalnoj porodici nasumičnih modela grafova (*Exponential family Random Graph Models – ergm*). Statnet pruža sveobuhvatni okvir za modeliranje mreža baziran na *ergm* modelima, uključujući modele za procenu, simulaciju i vizualizaciju mreža. Paket koristi Monte Karlo algoritam zasnovan na Markovljevim lancima (*Markov chain Monte Carlo – MCMC*), koji pokreće njegovu široku funkcionalnost (Handcock et al., 2016; Handcock, Hunter, Butts, Goodreau, & Morris, 2008)

network paket predstavlja skup alata za kreiranje i izmenu mrežnih objekata. Mrežnom klasom se može predstaviti širok spektar relacionih tipova podataka, a podržani su i arbitrarni atributi čvorova, ivica i grafova (Butts, 2008a, 2015)

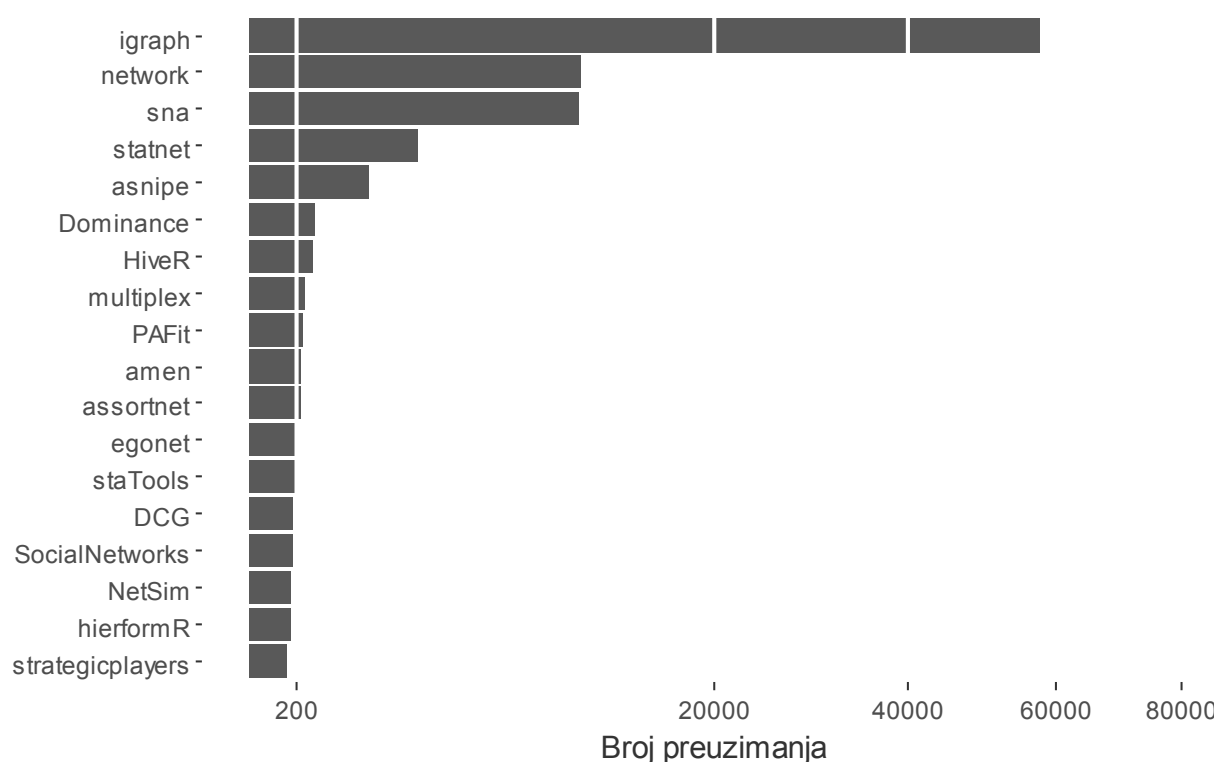
egonet je mala alatka za analizu društvenih mreža, koja služi za izračunavanje metrika egocentričnih mreža, uključujući Burtovu efektivnu veličinu (*Burt's effective size*) i agregatno ograničenje (*aggregate constraint*) uz kod za importovanje koji je koristan za obrađivanje velikog broja matrica povezanosti (Sciandra, Gioachin, & Finos, 2012)

HiveR (<http://www.hiveplot.net/>) služi za vizualizaciju 2D ili 3D grafikona košnice. Grafikoni košnice su jedinstveni metod prikaza mreža različitih tipova u kojima su osobine čvora mapirane na ose korišćenjem razumljivih osobina, a ne arbitrarnog pozicioniranja. Koncept grafikona košnice smislio je Martin Krzywinski (Hanson, 2016).

Ostali R paketi koji se koriste u analizi društvenih mreža uključuju: **NetSim** (Stadtfeld, 2013), **amen** (Hoff, Fosdick, Volfovsky, & He, 2015), **staTools** (Bessi, 2015), **SocialNetworks** (Nightingale & Nightingale, 2014), **assortnet** (D. Farine, 2016), **Dominance** (Krueger & Krueger, 2016), **DCG** (Chen, Jin, Beisner, McCowan, & Fushing, 2016), **hierformR** (Curley & Chase, 2016), **strategicplayers** (Ott, 2016), **PAFit** (Pham, Sheridan, & Shimodaira, 2015a, 2015b, 2016; Pham, Thong, et al., 2016), **multiplex** (Ostoic, 2016), **asnipe** (D. R. Farine, 2016).

Među svim ovim paketima, ima dosta preklapanja kada je reč o funkcionalnosti, pa je jednu stvar moguće uraditi na više načina korišćenjem različitih paketa. Zajedno svi navedeni paketi čine R veoma moćnim oruđem za analizu društvenih mreža. Međutim, važno je istaći da nisu svi paketi podjednako popularni. Naime, CRAN vodi statistiku preuzimanja paketa sa svojih

servera, tako da je moguće videti koji paket je koliko puta preuzet u nekom vremenskom periodu (Csardi, 2015). Prema tim podacima daleko najpopularniji paket je *igraph*. U periodu od 29. novembra do 29. decembra ovaj paket je preuzet sa CRAN repozitorijuma dva puta češće nego svi ostali paketi za analizu društvenih mreža zajedno (Slika 2). Drugo mesto, prema broju preuzimanja, zauzima paket *network*. Razlog za to je što se mnogi drugi paketi pozivaju na njega (kao osnovi alat za kreaciju i izmenu mrežnih objekata). Sledeći po popularnosti su paketi *sna* i *statnet*, koji kao i *igraph* imaju relativno širok spektar mogućnosti, dok njih slede ostali paketi koju su uže fokusirani.



Slika 2 R paketi za analizu društvenih mreža prema broju preuzimanja sa CRAN servera u periodu od 29. novembra do 29. decembra 2016. god.

Izvor podataka: CRAN (Csardi, 2015)

1.4 Motivacija

1.4.1 Izbori u Srbiji

U Srbiji su se u proleće 2016. održali parlamentarni, pokrajinski i u mnogim opštinama lokalni izbori. Iako rezultati izbora nisu bili neočekivani, uticaj društvenih mreža je, u skladu sa duhom vremena, bio itekako primetan, kao što je to bio slučaj sa nedavnim izborima u drugim državama (Ahmed, Jaidka, & Cho, 2016; Anstead & O'Loughlin, 2015; Huberty, 2015). Porast penetracije interneta, koja je u Srbiji već 2013. god. prešla 50% (World Bank, 2016), je učinio društvene mreže relevantnim za ishod izbora. Vrlo loša situacija u domaćim medijima, u smislu cenzure, autocenzure, neprofesionalnosti, pristrasnosti i korupcije (ANEM, 2016; Freedom House, 2016; Reporters sans frontières, 2016), samo pojačava značaj društvenih mreža na internetu.

1.4.2 Analiza društvenih mreža na Tviteru

Iako Tviter nije mreža sa velikim brojem korisnika u Srbiji, ona je veoma značajna jer je koristi veliki broj političara, novinara i intelektualca, što značajno pomaže u premeštanju javnog diskursa sa ugušenih *mainstream* medija na internet. Samo zbog ovoga analiza interakcija na Tviteru i identifikacija ključnih aktera u mreži je uputna. Još jedan važan faktor izbora baš ove društvene mreže za analizu je njena otvorenost. Naime, po svojoj prirodi svaki tweet je javan, dostupan svima. Za razliku od drugih društvenih mreža, gde sadržaj mogu videti samo prijatelji (ili pratioci), kod Tvitera je komunikacija dostupna svima, čak i korisnicima koji nemaju registrovan nalog na toj društvenoj mreži. Iako je postavke privatnosti moguće prilagoditi za svaki pojedinačni nalog da bude potpuno privatn ili potpuno javan, na Tviteru je standardna opcija potpuna javnost. Zbog tih karakteristika, Tviter važi za otvorenu online društvenu mrežu, koja okuplja ljude koji žele da podele svoje mišljenje sa celim svetom, pa je broj ljudi koji biraju da zaštite svoje tweetove od javnosti veoma mali. To znači da je komunikacija transparentna ne samo u smeru političari – ciljane javnosti, već i obratno, ali transparentnost karakteriše i međusobnu interakciju među političarima.

1.4.2.1 Dosadašnja istraživanja u Srbiji

Prema saznanjima autora do sada nije bilo socioloških i politikoloških istraživanja bilo kojih izbora u Srbiji primenom analize društvenih mreža koju formiraju političari. Jedan razlog za to je relativna novina značajne uloge *online* društvenih mreža na politički diskurs. Drugi razlog je taj što je i analiza društvenih mreža relativno nova oblast koja renesansu doživljava zahvaljujući rastućim računarskim kapacitetima. Iako je bilo moguće apstrahovati i konvencionalnu interakciju među političarima pre popularizacije *online* društvenih mreža, sama oblast analize

društvenih mreža nije bila dovoljno zastupljena, u međunarodnim, a naročito u domaćim okvirima. Ovaj nedostatak primene analize društvenih mreža je prisutan ne samo kada je reč o politikološkoj analizi izbora, već i domaćoj politikologiji, ali i sociologiji. Naime, čak i u slučajevima kada se neki od domaćih sociologa bavio analizom društvenih mreža, to je bilo na nivou preglednih članaka (Babovic, 2005; Vitkovic, 2013). Nedostatak primene metoda analize društvenih mreža u Srbiji je zabrinjavajuć uzevši u obzir da se trenutno najveći napreci u sociologiji prave baš u slučajevima krosfertilizacije sa drugim naučnim disciplinama, a naročito u slučaju računarske sociologije (Squazzoni, 2012).

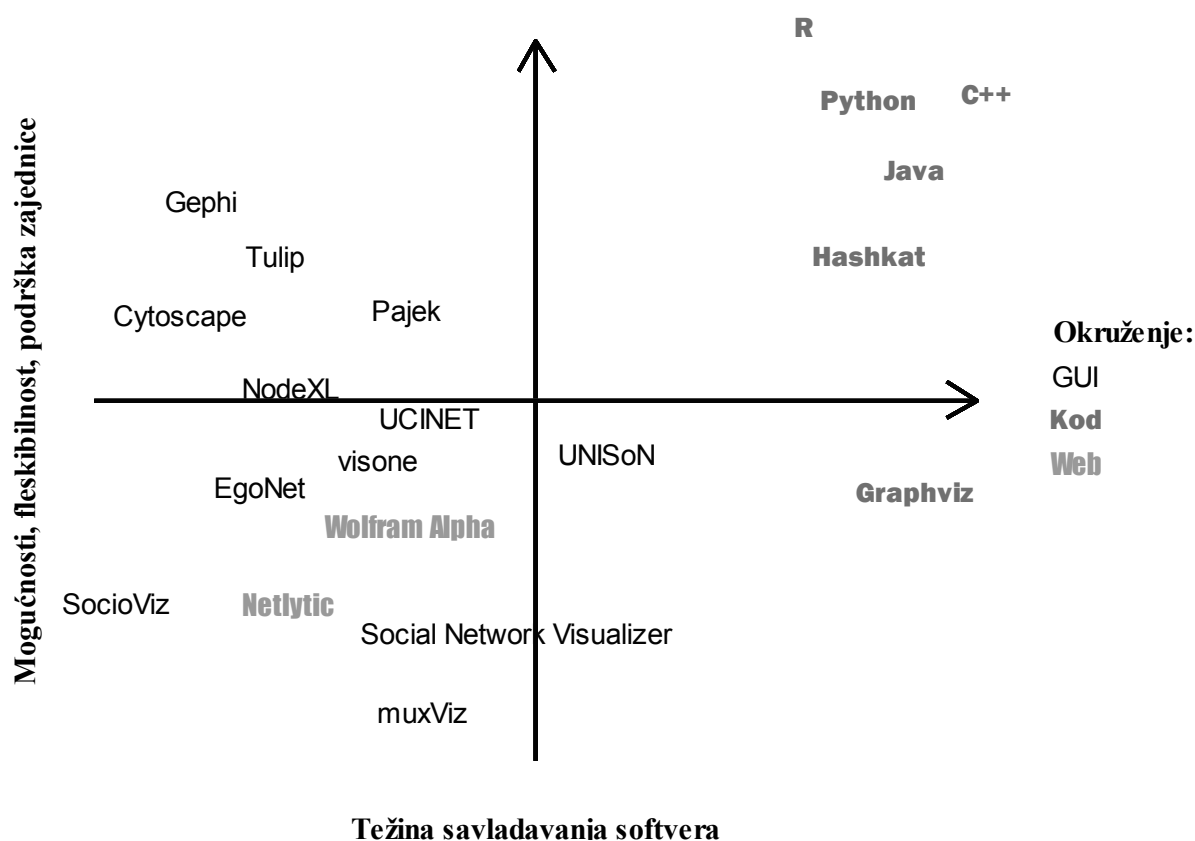
1.4.2.2 Dosadašnja istraživanja u svetu

Online društvene mreže kao medij značajan u izbornom procesu postaju zapažene u svetu od izbora u SAD 2008. i njihovog često isticanog značaja, a naročito Tvitera, u kampanji Baraka Obame. Analiza sadržaja te kampanje ukazala je na veliki potencijal za istraživanje sadržaja poruka koje se razmenjuju na *online* mrežama poput Tvitera (Zappavigna, 2011). Veći broj novijih istraživanja ukazuje na rastući značaj *online* društvenih mreža u političkom diskursu (Bacallao-Pino, 2016; Chan, 2016; Haleva-Amir, 2016; Heo, Park, Kim, & Park, 2016; Steinfeld, 2016; Theocharis & Quintelier, 2016). Iako je uticaj društvenih mreža na izborni proces prvi put registrovan u državama sa najrazvijenijom demokratijom, njihov značaj je možda još veći u nedemokratskim državama. Uloga Tvitera u Iranu 2009 (Evgeny Morozov, 2009) i u različitim uspešnim i neuspešnim prevratima koji se kolektivno nazivaju *Arapsko proleće* (Khondker, 2011; Lotan et al., 2011), je ne samo očigledna, već ključna. Situacija u drugim autoritarnim državama poput Turske i Rusije je po pitanju *mainstream* medija takođe veoma teška (Gökçe, Hatipoğlu, Göktürk, Luetgert, & Saygin, 2014; Reuter & Szakonyi, 2015), a *online* društvene mreže, kao bastion slobodnog izražavanja mišljenja, su sve češće na udaru države. Nedavno je grupa autora iz Singapura (Ahmed et al., 2016) istraživala uticaj komunikacije na Tviteru na izborni proces u Indiji. Njihovi rezultati pokazuju da je prisustvo na Tviteru mnogo intenzivnije kod novih političkih opcija, dok se političke stranke koje su postale deo establišmenta, drže *mainstream* medija. Na primeru Španije, tj. Katalonije, Meri Lopez (2015), pokazuje da se analiza komunikacije na Tviteru može koristiti kao 'termometar' javnog mnjenja, za identifikaciju trendova u politici, ali čak i za predikciju rezultata izbora.

1.5 Izbor softvera

Poslednjih godina postoji veliko interesovanje za primenu metoda i tehnika analize društvenih mreža u društvenim naukama, pogotovo u sociologiji. Sama SNA (*Social Network Analysis* –

analiza društvenih mreža) oblast pripada široj oblasti informatičke sociologije (*Computational Sociology*), koja ima veliki značaj jer se u njoj vidi mogućnost za validacijom sociologije kao nauke, koja je veoma kasno prihvatila novonastalu metodologiju (Macy & Willer, 2002), a zasnovana je na kvalitativnim istraživanjima. Upravo taj fokus na kvalitativnom i dugogodišnje obrazovanje sociologa u tom smeru, čini klasično obrazovane sociologe i druge društvene naučnike nedovoljno opremljenim za kvantitativnu analizu. To je jedan od razloga za postojanje brojnih softverskih alata koji u ogromnoj meri olakšavaju analizu društvenih mreža. Među njima ima onih koji su veoma laki za korišćenje do onih koje je teže savladati (Slika 3).



Slika 3 Softverski alati za analizu društvenih mreža prema: okruženju, težini savladavanja i mogućnostima

Napomena: UCINET je komercijalni alat. NodeXL i Wolfram Alpha nude ograničenu funkcionalnost u svojim besplatnim verzijama. Ostali prikazani alati su besplatni.

Softverski alati za analizu društvenih mreža se prema obliku interakcije sa korisnikom, a time i nivou programerskih znanja i veština koji zahtevaju od korisnika, mogu podeliti u tri grupe. Prva obuhvata programske jezike i njihove pakete. Ovi alati nude najveće mogućnosti za analizu, ali njihov nedostatak je programersko znanje koje je neophodno, pa su često nedostupni istraživačima u domenu društvenih nauka, koji tokom svog obrazovanja najčešće ne stiču programerske veštine. Iako analiza društvenih mreža može da se realizuje u skoro svakom programskom jeziku, trud koji bi se tom prilikom uložio u zavisnosti od specifičnosti sintakse i

već postojećeg koda, varira. Jezici koji se najviše koriste za analizu društvenih mreža su R, Pajton (*Python*), Java i C++. Biblioteke koje u velikoj meri olakšavaju analizu postoje za sva četiri programska jezika. Važno je istaći da se C++ najčešće koristi kao dopuna postojećim jezicima, kada se komputaciono intenzivni delovi softvera implementiraju u njemu radi bržeg izvršavanja, kao i za razvoj aplikacija, a manje za samu analizu. Brojne aplikacije napisane su u Javi, ali i u njoj se sama analiza takođe retko obavlja. Za direktnu analizu najčešće se koriste programski jezici R i Python. Ovoj grupi softvera pripadaju i alati koji nemaju grafički interfejs, već se izvršavanjem koda (tipično zasnovanog na veoma uprošćenoj sintaksi) vrši analiza. Ovaj tip softverskih alata je često veoma specijalizovan, i veoma su dobri u toj stvari za koju su specijalizovani, ali ih je nemoguće primeniti u drugim okolnostima. Primeri uključuju Hešket (*Hashkat*) i Grafviz (*Graphviz*).

Drugu grupu softverskih alata za analizu društvenih mreža čine alati sa grafičkim korisničkim interfejsom. Ovi alati poput Gefija (*Gephi*), Tulipa i Pejaka korisnicima koji nemaju iskustva sa programiranjem daju velike mogućnosti za analizu i vizualizaciju društvenih mreža. Neki od njih imaju mogućnost dodavanja priključaka (*plugins*) kojima se dodatno proširuju njihove mogućnosti.

Treću grupu alata čine Veb servisi. Ova vrsta alata je često namenjena komercijalnoj upotrebi. Vrlo su specijalizovani, a oni koji su dostupni besplatno su još više ograničeni. Njihova prednost je to što sto su veoma laki za korišćenje. U njima je često integrisana i funkcija akvizicije podataka, što je često i najveća prepreka u bilo kojoj analizi društvenih mreža. Primeri uključuju Netlitik (*Netlytic*) i Vulfram Alfa (*Wolfram Alpha*).

Izbor R-a, kao softvera pomoću kojeg je sprovedena analiza demonstrirana u ovom radu, zasnovan je na više razloga. Prvi razlog je raspoloživost programskih paketa koji u ogromnoj meri proširuju osnovne mogućnosti R-a i olakšavaju analizu. U vezi sa tim je i jaka podrška zajednice. R zajednica je živa, dok su zajednice koje su nekada i postojale oko drugih softverskih alata, trenutno uglavnom neaktivne. Alternativa R-u, bio bi programski jezik Python, za koji postoje identični tipovi paketa, ali ga je teže savladati i njegova zajednica, iako životna, okuplja mahom programere, dok je R mnogo zastupljeniji u akademskoj zajednici i generalno ima dosta kvalitetniju dokumentaciju. R je tokom godina postao *lingua franca* statistike, pa se sve češće upotrebljava u svim naukama koje koriste statistiku. Istraživači koji treba da se opredele za alat koji će koristiti za analizu društvenih mreža, treba da imaju na umu i činjenicu da je za komplementarne statističke i kvantitativne analize koje budu radili uz analizu društvenih

mreža, R generalno bolji izbor od Pajtona. Još jedna prednost R-a su vizualizacije, koji ima mnogo jaču podršku (u odnosu na Pajton), jer je to i jedna od glavnih svrha ovog jezika.

Grafički alati poput Gefija, zahtevaju već preuzete i srede podatke, dok kod programskih jezika postoji mogućnost dobijanja podataka preko aplikacionog programskog interfejsa (API) raznih *online* društvenih mreža. Postoji mogućnost razvoja prilagođenih priključaka za alate poput Gefija, ali to je znatno izazovnije od rada u R-u ili Python-u. Kada bi se problem prikupljanja podataka zanemario, grafički alati ne nude dovoljno fleksibilnosti za izvođenje svih analiza koje su neophodne u ovom istraživanju.

1.6 Polazne hipoteze

1. Pojedini akteri i grupe su značajno uticajni od ostalih u analiziranoj društvenoj mreži.
2. Kompozitni pokazatelj formiran na osnovu različitih mera centralnosti čvorova u analiziranoj društvenoj mreži nema jaku korelaciju ranga sa samo jednom merom centralnosti, te je njegova upotreba za rangiranje centralnosti čvorova u razmatranoj mreži opravdana.
3. Homofilija je neravnomerno zastupljena među analiziranim grupama.
4. Broj zatvorenih trijada u analiziranoj društvenoj mreži je neravnomerno zastupljen među političkim strankama.
5. Grupe identifikovane klaster analizom u analiziranoj društvenoj mreži se preklapaju sa formalnim grupama kojima akteri pripadaju.

2 PODACI I METODE

U ovom radu uloga koju je komunikacija na Tviteru imala na izbore analizirana je na dva načina. Prvi način podrazumeva kvantitativnu analizu osnovnih metrika popularnosti (broj pratioca, broj retvitova (retweets) i sl.), svakog naloga na Tviteru koji je od interesa za analizu. Drugi način podrazumeva analizu društvene mreže koju političari formiraju na Tviteru, konkretno, analizu njene topologije i karakteristika čvorova koji je sačinjavaju. Ukrštanjem rezultata ova dva oblika analize - analize društvenih mreža (SNA - *Social Network Analysis*) i analize metrika popularnosti specifičnih za Tviter - ostvaren je detalj uvid u društvenu mrežu političkih aktera na Tviteru.

Sama SNA izvršena je pomoću više različitih pristupa. Ocenom centralnosti pojedinačnih Tviter naloga utvrđeno je ko su bili ključni akteri na Tviteru u kampanji svake stranke. Takođe, upotrebom različitih metrika centralnosti u mreži i statističke analize ocenjene su performanse čitavih stranaka koje su učestvovala u kampanji, u smislu zbirne centralnosti čvorova prema stranačkoj pripadnost. Ispitani su odnosi između stranaka, tj. koliko su pojedinci u strankama spremni da komuniciraju sa pripadnicima drugih stranaka, i da li se, i u kojoj meri, grupisanje čvorova u mreži, poklapa sa stranačkom pripadnošću čvorova.

Na kraju rada, data su preporuke, za sve stranke koje su bile predmet analize u ovom radu, bazirane na rezultatima analize.

2.1 Podaci

Podaci su prikupljeni za period od 04.03.2016, kada su izbori raspisani, do 05.05.2016, kada su se ponovljeni izbori završili. Uzorak obuhvata Tviter naloge vezane za sve stranke koje su ušle u parlament. Podaci za svaki uzorkovan Tviter nalog obuhvataju sve tvitove tog naloga i njihove metapodatke za navedeni vremenski period. Uzorkovanje je započeto tako što su prvo prikupljeni podaci zvaničnih naloga stranaka i njihovih istaknutih članova na Tviteru (N=40, videti Prilog I za listu političara, odnosno Tviter naloga uključenih u osnovni uzorak), da bi se kasnije uzorak proširio *snowball* metodom na naloge koji su najviše interagovali sa naložima iz osnovnog uzorka. Prikupljeni podaci za svaki nalog obuhvataju i osnovne Tviter pokazatelje za svaki uzorkovani nalog, kao što su: broj pratioca, prijatelja, tvitova, retvitova, omiljenih tvitova i sl. (Prilog II). Svi podaci su preuzeti preko Twitter REST API-ija, pomoću R paketa *twitteR* (Gentry, 2015), a detaljan proces prikupljanja i sređivanja podataka opisan je u poglavlju 3.

2.2 Metode

Nakon prikupljanja podataka, napravljeni su grafovi u kojima uzorkovani Tviter nalozi predstavljaju čvorove, a interakcija između tih naloga su ivice grafa. Interakcija se u slučaju ovih grafova definiše kao pominjanje (*mention*) nečijeg korisničkog imena (*handle-a*) u tvitu, bez obzira da li se radi o originalnom tvitu ili o retvitu. Ivice u grafu imaju smer, kao i težinu, koja zavisi od toga koliko često je jedan nalog pomenuo drugi nalog, a na grafu je predstavljena debljinom ivica.

U radu su ispitane 3 grupe pokazatelja za naloge iz osnovnog uzorka političara.

- Prva grupa obuhvata mere centralnost čvorova u mreži:
 - Centralnost intermedijarnosti
 - Centralnost stepena (dolaznog i odlaznog)
 - Centralnost bliskosti (dolazne i odlazne)
 - Centralnost svojstvenog vektora
 - Kompozitni pokazatelj koji objedinjuje sve mere centralnosti
- Druga grupa pokazatelja se odnosi na podatke o nalogima:
 - Broj pratioca (*followers*)
 - Broj prijatelja (*friends*)
 - Ukupan broj tvitova (*tweets*)
 - Broj tvitova koje nalog označio da mu se dopadaju (*favorited*)
 - Starost naloga (pre koliko vremena je nalog otvoren)
- Treća grupa pokazatelja se odnosi na tvitove koje su nalozi tviitali tokom kampanje:
 - Broj tvitova za vreme kampanje
 - Broj puta koliko su tvitovi naloga označeni kao omiljeni (*favorties*)
 - Broj retvitovanja tvitova određenog naloga (*retweets*)
 - Prosečan broj retvitova po tvitu (odnos prve i treće promenljive iz ove grupe)

Interpretacija izabranih mera centralnosti čvorova u grafu nije univerzalna, tj. interpretacija umnogome zavisi od topologije mreže i u različitim mrežama mere centralnosti se mogu interpretirati na različite načine (Borgatti & Everett, 2006). Dodatno, svi ovi pokazatelji najčešće različito rangiraju čvorove u analiziranom grafu pa se javlja problem: kako izabrati pokazatelj

koji najbolje oslikava uticaj čvorova? Zato je na osnovu svih izabranih mera centralnosti izračunat nov, kompozitni pokazatelj.

Metod koji je korišćen za izračunavanje kompozitnog pokazatelja je P_2 odstojanje. Ovaj pokazatelj baziran je na Ivanovićevom odstojanju (Ivanović, 1974; Pena, 1977) i omogućava izvlačenje varijabiliteta svih ulaznih pokazatelja, bez dupliciteta varijabiliteta između srodnih pokazatelja (Dobrota, 2015; Jeremić, 2012). Za P_2 odstojanje, koje je veoma sličan pokazatelj, postoji R implementacija u vidu paketa *p2distance* (Perez-Perez & Bonet, 2012). Prema Perez-Perezu i Bonetu (2012) P_2 odstojanje se izračunava na sledeći način:

$$DP_2 = \sum_{i=1}^n \left\{ \left(\frac{d_i}{\sigma_i} \right) \left(1 - R_{i,i-1,i-2,\dots,1}^2 \right) \right\}$$

sa $R_1^2=0$; gde je $d_i = |x_{ri} - x_{*i}|$ za skup obeležja $X_* = (x_{*1}, x_{*2}, \dots, x_{*n})$ gde je:

- n broj promenljivih
- x_{ri} vrednost promenljive i za entitet r
- σ_i standardna devijacija promenljive i
- $R_{i,i-1,\dots,1}^2$ koeficijent determinacije u regresionom modelu sa X_i kao zavisnom promenljivom i $X_{i-1}, X_{i-2}, \dots, X_1$ kao nezavisnim promenljivima.

P_2 odstojanje se razlikuje od Ivanovićevog odstojanja po tome što koristi koeficijent determinacije, dok se kod Ivanovićevog odstojanja koristi koeficijent parcijalnih korelacija. I jedan i drugi pokazatelj se formiraju postupno. Prema Jeremiću (2012):

počinje se integracijom celokupnog diskriminacionog efekta promenljive X_1 , tj. promenljive koja sadrži najveću količinu informacije o entitetu koji se rangira (što se postiže izračunavanjem zbira apsolutnih vrednosti korelacija matrice). Zatim se dodaje onaj deo diskriminacionog efekta druge (po rangu) promenljive, koji nije bio uključen u diskriminacionom efektu prve promenljive, pa onaj deo diskriminacionog efekta treće promenljive koji nije bio uključen u diskriminacionom efektu prve dve promenljive itd. (Ivanović, 1977). (str. 83)

Još jedna razlika između P_2 i Ivanovićevog odstojanja je način na koji se određuje prvobitni redosled promenljivih prema količini informacija koju nose. Kod Ivanovićevog odstojanja se za

rangiranje koristi suma apsolutnih koeficijenata korelacije, dok se kod P_2 odstojanja koristi Frešeovo odstojanje.

Mrežne metrike dobijene SNA analizom su onda upoređene sa pokazateljima iz druge i treće grupe, tj. konvencionalnim metrikama kojima se najčešće meri popularnost na Tviter mreži (Romero, Galuba, Asur, & Huberman, 2011). Za sve tri grupe metrika tj. pokazatelja izračunata je korelaciona matrica, kako bi se ispitalo prisustvo i smer korelacija između razmatranih pokazatelja. Metod korišćen pri izradi korelacione matrice je Spirmanova ρ korelacija ranga. Ovaj neparametarski metod je izabran jer distribucija za najveći broj promenljivih (tj. pokazatelja) nije normalna, a neprekidnost promenljivih je isključila upotrebu Kendalovoe τ korelacije ranga.

Homofilija, tj. sklonost aktera da više komuniciraju sa akterima koji pripadaju istoj stranci, ispitana je koeficijentom asortativnosti, izračunatim za podgrafove pojedinačnih stranaka. Podgrafovi su formirani tako da uključuju čvorove pripadnika stranaka i čvorove sa kojima su oni ostvarili komunikaciju u izbornom periodu, s tim da se posmatra samo obostrana komunikacija, a ne samo obično pominjanje. Učestalost zatvorenih trijada u istim podgrafovima testirana je pomoću mere tranzitivnosti, koja predstavlja odnos broja zatvorenih trijada i povezanih trojki. Za klaster analizu razmotreno je više različitih klaster algoritama, a najpogodnijom se pokazala metoda mapa slučajnih šetnji, kojom se po principu toka informacija uzima u obzir i smer i težina ivice (Rosvall, Axelsson, & Bergstrom, 2009).

2.3 R paketi i replikabilnost

Osim samog R-a (R Core Team, 2016), u radu su korišćeni i brojni paketi kojima je proširena osnovna funkcionalnost R-a. Za akviziciju podataka korišćen je paket *twitteR* (Gentry, 2015). Za analizu, izračunavanje metrika i vizualizaciju društvenih mreža korišćen je paket *igraph* (Csardi & Nepusz, 2006). Grafikoni košnice vizualizovani su pomoću paketa *HiveR* (Hanson, 2016). Za ostale vizualizacije u radu korišćen je paket *ggplot2* (Wickham, 2009). P_2 odstojanje pokazatelja centralnosti izračunato je pomoću *p2distance* paketa (Perez-Perez & Bonet, 2012).

Sva analiza praćena je R kodom koji je korišćen za njeno izvođenje. Rad je strukturiran tako da je u narednom poglavlju demonstrirana upotreba R-a i predstavljeni dobijeni rezultati te analize. Diskusija rezultata data je u posebnom poglavlju. Uz master rad dolazi i CD, na kojem se osim elektronske verzije rada, nalazi i celokupan R kod u datotekama sa .R ekstenzijom. Data su dve R datoteke: prva po imenu „kod.R“, koja sadrži kod koji se nalazi u svim Listinzima

predstavljenim u radu i druga datoteke po imenu „grafikoni.R“, u kom je dat kod za vizualizaciju svih grafikona koji korišćeni u radu, a koji nemaju direktno veze sa analizom društvenih mreža, zbog čega kod za njih nije predstavljan u listinzima. Osim elektronske verzije i koda deo elektronske verzije su i podaci preuzeti sa Tviter servera, kao i dodatna CSV datoteka koja sadrži attribute koji su korišćeni u analizi. Kako bi se osigurala replikabilnost, data je i datoteka „paketi.txt“ sa izlistanim paketima koji su bili korišćeni u radu, uz njihove verzije. Celokupan kod, podaci, spisak korišćenih paketa i elektronska verzija teksta može se naći i na internetu (<http://github.com/gljk>).

3 POSTUPAK ANALIZE U R-U I REZULTATI

3.1 Akvizicija podataka

3.1.1 R i Tvider API

Za prikupljanje podataka korišćen je `twitterR` paket, koji koristi Tviderov aplikacioni programski interfejs (API) za pristup podacima sa Tvider servera. Da bi se Tviderov API uopšte mogao koristiti potrebno je kreirati Tvider aplikaciju (na adresi <http://apps.twitter.com>). Moguće je koristiti Twitter API i preko korisničkog naloga, ali je u tom slučaju dozvoljena količina podataka koji se mogu povući preko API-a u određenom vremenskom periodu značajno manja. Važno je napomenuti da postoji ograničenje su samom API-u koje ne dozvoljava da se povuče više od poslednjih 3200 tvitova koje je neki nalog tvitovao (npr. ako je neko od otvaranja naloga do danas tvitovao 6000 tvitova, prvih 2800 tvitova će biti nedostupno preko Tvider API-a). Ovo može da predstavlja problem ako nalozi za koje se podaci preuzimaju učestalo tvituju, pa je nemoguće doći do starijih tvitova. Zbog ovog ograničenja preporučljivo je da se akvizicija podataka obavi što ranije (odmah nakon posmatranog perioda) ili kontinuirano, kako bi se osigurala kompletnost uzorka.

Nakon registracije aplikacije, potrebno je iz postavki kreirane aplikacije u sekciji *Keys and Access Tokens* prekopirati sledeće ključeve: *Consumer Key*, *Consumer Secret*, *Access Token* i *Access Token Secret*. Ti ključevi predstavljaju podatke potrebne za autentikaciju, tj. za identifikaciju svakog zahteva koji će R sesija praviti prema Tvider serverima. Kod za autentikaciju na Tvider API prikazan je u Listingu 1.

```
require(twitterR)

#Unošenje ključeva u nove varijable

consumer_key <- " Niz dugačak 25 karaktera "
consumer_secret <- " Niz dugačak 50 karaktera "
access_token <- " Niz dugačak 45 karaktera "
access_secret <- " Niz dugačak 50 karaktera "

#Samo podešavanje autentikacije vrši se pomoću sledeće funkcije

setup_twitter_oauth(consumer_key, consumer_secret, access_token,
access_secret)
```

Listing 1 Autentikacija na Tvider API koristeći R paket *twitterR*

Nakon podešavanja autentikacije sam proces prikupljanja podataka može početi. Celokupni Tvider API nam omogućava sakupljanje raznih podataka, od pretraživanja najskorijih tvitova, do

imena naloga koje određen korisnik prati, ili koji njega prate. U ovom radu podaci koji su zahtevani od Tvitera su podaci o tvitovima koje su selektovani korisnici tviitali. Funkcija u okviru paketa *twitter* koja je korišćena za prikupljanje podataka prikazana je u Listingu 2.

```

tvitoviNaloga <- userTimeline("Nalog", includeRts=TRUE, n=3200)

#Argumenti n i includeRts označavaju broj tvitova i inkluziju retvitova.
#Konverzija dobijenog objekta u R dataframe:

tvitoviNaloga <- twListToDF(tvitoviNaloga)

```

Listing 2 Demonstracija prikupljanja tvitova za jedan nalog pomoću R paketa *twitter*

3.1.2 Akvizicija podataka osnovnog uzorka

Opisani postupak akvizicije podataka potrebno je primeniti na veći broj naloga. Za te potrebe unapred je sačinjena jednostavna CSV datoteka (*političari.csv*) u kojoj su sadržani osnovni podaci o nalogima za koje će se podaci prikupljati, a koja ima format prikazan u Tabela 1. Ta datoteka sadrži podatke o 40 naloga, koji su izabrani za početnu analizu. U narednom koraku biće nam potrebni samo podaci iz kolone Nalog, a ostali podaci će biti korišćeni u narednim koracima, kao što je npr. vizualizacija, za čije potrebe postoji kolona Boja, koja predstavlja boju kojom će članovi date stranke biti reprezentovani.

Tabela 1 Format pripremljenih podataka o nalogima koji su predmet analize (*političari.csv*)

Ime	Prezime	Stranka	Nalog	Boja	Opozicija
Saša	Radulović	DJB	SasaRadulovich	#FF7B00	1
Bojan	Pajtić	DS	PajticBojan	#FDF003	1
	DS	DS	Demokrate	#FDF003	1
Aleksandar	Vučić	SNS	avucic	#152755	2
Vojislav	Šešelj	SRS	seselj_vojislav	#024176	1

Uz pripremljene podatke, i osnovne principe dolaženja do podataka o pojedinačnom nalogu, moguće je preuzeti podatke manuelno jedan po jedan, za svaki uzorkovani nalog, ali prednost upotrebe programskoj jezika je u tome što se ovakvi procesi mogu automatizovati, kao što je prikazano u Listingu 3.

#Funkcija ekstrakcija() kao ulaz prima spisak naloga, i dva vremenska perioda: početak i kraj vremenskog intervala za koji želimo tvitove

```

ekstrakcija <- function(spisak, od = NULL, do = NULL) {
  tvitovi <- data.frame()
  for (i in 1:length(spisak)) {
    a <- userTimeline(spisak[i], n = 3200, includeRts = TRUE)
    a <- twListToDF(a)
    if (!is.null(od)) {

```

```

    a <- a[a$created > od, ]
  }
  if (!is.null(do)) {
    a <- a[a$created < do, ]
  }
  tvitovi <- rbind(tvitovi, a)
  if (as.numeric(getCurRateLimitInfo()[39, 3]) < 5) {
    Sys.sleep(60 * as.numeric(getCurRateLimitInfo()[39, 4] - Sys.time()) +
10)
  }
}
return(tvitovi)
}

```

Listing 3 Funkcija za prikupljanje tvitova sa više naloga, tvitovanih u izabranom vremenskom periodu

Demonstrirana funkcija (Listing 3) pomoću petlje ponavlja proces prikupljanja podataka onoliko puta koliko ima naloga u spisku koji je dat funkciji. Osim prikupljanja podataka funkcija pravi podskup podataka, tj. zadržava samo one tvitove koji su tvitovani u vremenskom intervalu između vremena datog argumentima od i do. Dodatno funkcija ima funkcionalnost da u slučaju da se približi limitu dozvoljenih upita prema Tviter serverima, pauzira prikupljanje podataka do vremena kada se kvota resetuje.

Sledeći korak, prikazan Listingom 4, je primena prikazane funkcije za akviziciju tvitova srpskih političara za vreme izborne kampanje za parlamentarne izbore 2016.

```

#Učitavamo unapred pripremljenu datoteku sa podacima: političari.csv
političari <- read.csv("političari.csv", header = TRUE, encoding = "UTF-8")

#Primenjujemo funkciju na kolonu Nalog, datoteke političari.csv za vremenski
interval od raspisivanja izbora do završetka ponovljenih izbora 2016.

tvitovi <- ekstrakcija(spisak = političari$Nalog,
                      od="2016-03-04 00:00:00 CET",
                      do="2016-05-05 00:00:00 CET")

```

Listing 4 Demonstracija upotrebe funkcije za akviziciju podataka iz Listing 3

Trajanje izvršavanja funkcije ekstrakcija() zavisi od više faktora, a pre svega od broja naloga koji se nalaze na spisku. Brzina sistema na kome se kod izvršava takođe je jedan od važnih faktora, kao i propusni opseg internet konekcije, ali u znatno manjoj meri. Ako je nalog tvitovao 3200 tvitova (maksimum koji preko API-a može da se povuče), očekivano vreme trajanja izvršenja funkcije ekstrakcija() za taj jedan nalog je 79,29 sekundi. Za spisak naloga političara i političkih stranaka koji su povučeni za potrebe ovog rada (n=40), vreme izvršenja funkcije trajalo je oko 50 minuta. Ako spisak sadrži 1000 naloga, izvršenje funkcije bi trajalo oko 22 sata.

Dobijeni *data.frame* tvitovi sadrži 22954 tvitova i njihove metapodatke. To znači da je prosečan broj tvitova po nalogu dnevno bio 9,49. Dostupni metapodaci uključuju: jedinstveni identifikacioni broj tvita, vreme kada je tvit tvitovan, broj retvitova, broj favorita (*favorites*), da li je tvit samostalan ili je dat kao odgovor na neki drugi tvit, u kom slučaju je identifikovan broj tvita i korisnički nalog kome je odgovoreno. Podaci o lokaciji (geografska širina i dužina) su takođe dostupni, ali je procenat korisnika koji bira da deli ove podatke zanemarljivo mali (od 39 naloga, samo jedan je birao da podeli svoju lokaciju).

3.1.3 Kreiranje proširenog uzoraka

Kada su podaci preuzeti sa Tviter servera, dobijeni su ne samo podaci o međusobnoj komunikaciji između političara, već o celokupnoj njihovoj javnoj komunikaciji na Tviteru. To nam omogućava da originalni uzorak političara proširimo tako što ćemo ustanoviti koji su nalozi sa kojima političari iz osnovnog uzorka najviše komuniciraju. Proces je moguće ponoviti više puta, ali zbog fenomena malog sveta, jako brzo bi se došlo do gotovo svih korisnika Tvitera, pa je u ovom radu uzorak proširen samo za taj jedan korak. Ova metoda se naziva i *snowball* (grudva snega) metoda, jer se počinje sa malim uzorkom, da bi se postepeno uvećavao tako što se u uzorak uključuju čvorovi koji su povezani sa onima koji su već deo uzorka. Postoje najmanje dva načina na koja je to moguće uraditi u slučaju prikupljenih podataka sa Tvitera. Preko metapodataka, brojanjem odgovora upućenih nalogima koji nisu uključeni u originalni uzorak, i analizom sadržaja Tvitova pronalaženjem niza karaktera posle karaktera „@“ koji predstavlja prefiks svakog naloga na Tviteru. Primer širenja uzorka dat u Listingu 5 je zasnovan na prvom metodi, dakle radi se o već dostupnim metapodacima.

#Funkcija snowball() uzima podatke sa Tvitera i kao rezultat daje imena naloga osnovnog uzorka, i sve one naloge sa kojima su nalozi iz osnovnog uzorka natprosečno često komunicirali

```
snowball <- function(tvitovi) {  
  x <- tvitovi  
  sb <-  
    names(sort(table(x$replyToSN)[table(x$replyToSN) >  
mean(table(x$replyToSN))], decreasing = T))  
  sb <- sb[!(sb %in% unique(x$screenName))]  
  return(c(unique(x$screenName), sb))  
}
```

Listing 5 Funkcija za primenu snowball metode za širenje uzorka

Rezultat funkcije snowball() predstavlja samo one naloge koji su natprosečno često komunicirali sa političarima. Dakle neko kome je jedan političar odgovorio jednim tvitom neće ući u prošireni uzorak, taj nalog bi imao samo jednu zabeleženu „komunikaciju“. U skupu naloga kome su

političari nešto odgovorili u toku kampanje prosečan broj takvih komunikacija je bio 4,63, tako da su u prošireni uzorak uključeni samo nalozi koji imaju pet ili više interakcija sa političarima. Sledeći korak je ponovno prikupljanje podataka za sada prošireni uzorak, što je prikazano u Listingu 6.

```
prošireni_uzorak <- snowball(tvitovi)
tvitovi.p.u <- ekstrakcija(prošireni_uzorak)
```

Listing 6 Primena funkcije snowball() i prikupljanje podataka za prošireni uzorak

Ukupan broj naloga proširenog uzorka je 411, međutim funkcija ekstrakcija je uspela da preuzme podatke za 355 naloga. Među 56 naloga za koje nisu prikupljeni podaci, su oni nalozi koji su u međuvremenu ugašeni, oni koji imaju podešenu opciju privatnosti tako da samo odobreni nalozi mogu da čitaju njihove tvitove, i oni koji su u periodu od izbora do vremena prikupljanja podataka tviitovali više od 3200 tvitova (Tviter API limit).

3.1.4 Akvizicija dodatnih podataka o Tviter naložima

Osim sadržaja tvitova izabranih korisnika postoje i podaci naloga koji takođe mogu biti interesantni za analizu. Proces akvizicije tih podataka dat je u Listingu 7.

```
naloziMeta <- function(spisak) {
  poltab <- matrix(nrow = length(spisak), ncol = 6)
  for (i in 1:length(spisak)) {
    a <- getUser(spisak[i])
    poltab[i, 1] <- a$screenName
    poltab[i, 2] <- a$followersCount
    poltab[i, 3] <- a$friendsCount
    poltab[i, 4] <- a$statusesCount
    poltab[i, 5] <- a$favoritesCount
    poltab[i, 6] <- as.character(a$created)
  }
  poltab <- as.data.frame(poltab)
  colnames(poltab) <-
    c("Nalog",
      "Pratioci",
      "Priјateljji",
      "Tvitovi",
      "Favorisao",
      "DatumOtvaranja")
  poltab$Pratioci <- as.numeric(as.character(poltab$Pratioci))
  poltab$Priјateljji <- as.numeric(as.character(poltab$Priјateljji))
  poltab$Tvitovi <- as.numeric(as.character(poltab$Tvitovi))
  poltab$Favorisao <- as.numeric(as.character(poltab$Favorisao))
  poltab$DatumOtvaranja <- as.Date(poltab$DatumOtvaranja)
  return(poltab)
}
```

Listing 7 Preuzimanje podataka o naložima sa Tviter servera

Preuzeti podaci o korisničkim nalogima uključuju broj pratioca datog Tviter naloga (*Followers*), broj naloga koje taj nalog prati na Tviteru (*Friends*), broj tvitova koje je nalog tvitovao, broj tvitova koje taj nalog označio da mu se dopadaju (*Favorites*) i datum kreiranja naloga. Rezultati ove funkcije prikazani su u Prilog II. Važno je napomenuti da se ovi metapodaci odnose na celokupan period od kada je pojedinačni nalog otvoren do dana preuzimanja podataka, a ne postoji način da se direktno dobiju navedeni metapodaci za određeni vremenski period. Ipak, u narednom poglavlju će se do tih podataka doći analizom celokupnog korpusa preuzetih podataka o tvitovima koji su preuzeti funkcijom `ekstrakcija()` prikazanom ranije u ovom poglavlju.

3.2 Sređivanje podataka

Da bi podaci uopšte mogli da se koriste za dalje analize potrebno je da budu u obliku koji je prihvatljiv za softverske paket koji će se koristiti u analizi. Najuobičajeniji format za predstavljanje društvene mreže, koji podržavaju gotovo svi alati jeste matrica povezanosti (*adjacency matrix*). Matricu povezanosti je moguće koristiti i u većini softverskih alata sa grafičkim interfejsom koji se koriste za analizu društvenih mreža (kao npr. Gefi).

Jedna od karakteristika Tvitera je mogućnost pominjanja drugih Tviter naloga u svakom tvitu. Pominjanje se razlikuje od odgovora, jer u odgovoru na neki tvit može ali i ne mora da bude pomenut nalog kome se odgovara. Naime kada korisnik na Tviteru klikne da odgovori, generiše se metapodatak uz taj tvit i on označava kome je upućen tvit, i taj tvit automatski počinje sa „@nalog“. Međutim, korisnik može da obriše početak tog odgovora „@nalog“, ali metapodatak, koji je generisan čim je započet odgovor, ostaje. Takođe, moguće je da korisnik u jednom tvitu pomene više različitih naloga. Tako da postoje dva načina za mapiranje relacija, jedan je korišćenjem metapodataka o odgovoru, a drugi je analizom sadržaja tvita i pronalaženjem niza karaktera koji čine nalog, tj. pominjanje. Ako se u nekom tvitu jednog naloga (čvora mreže) nalazi niz karaktera koji označava neki drugi nalog onda će to biti brojano kao jedna komunikacija. Funkcija `data` u Listingu 8 funkcioniše na sledeći način: prvo se konstruiše prazna matrica koja za imena kolona i redova ima naloge u uzorku. Potom se za svaki red, tj. nalog, broji koliko puta se u tvitovima tog naloga pojavljuju imena ostalih naloga.

#Funkcija matrica_povezanosti kao ulaz ima rezultat funkcije ekstrakcija, a kao rezultat daje matricu povezanosti.

```
matrica_povezanosti <- function(x) {  
  sp <- unique(x$screenName)  
  lsp <- length(sp)  
  mp <-  
    matrix(nrow = lsp, ncol = lsp)
```

```

colnames(mp) <- sp
row.names(mp) <- sp
for (i in 1:lsp) {
  l <- x[x$screenName == sp[i], ]
  for (j in 1:lsp) {
    if (j == i) {
      j <- j + 1
    }
    if (j > lsp) {
      break
    }
    k <- x[x$screenName == sp[j], ]
    if (any(grepl(k$screenName[1], l$text))) {
      mp[i, j] <- table(grepl(k$screenName[1], l$text))["TRUE"]
    }
  }
}
mp[is.na(mp)] <- 0
return(mp)
}

```

Listing 8 Funkcija za konstrukciju matrice povezanosti

Funkcija se dalje primenjuje na konstrukciju matrica povezanosti za uzorak političara i prošireni uzorak (Listing 9), a rezultat su dve matrice od kojih će se u narednom poglavlju formirati zasebni grafovi.

```

mat.pol <- matrica_povezanosti(tvitovi)
mat.pol.pu <- matrica_povezanosti(tvitovi.p.u)

```

Listing 9 Primena funkcije matrica_povezanosti na osnovni i prošireni uzorak

Vreme izvršavanja funkcije matrica_povezanosti zavisi od broja naloga u matrici, broja tvitova i naravno od brzine sistema. U slučaju matrice povezanosti proširenog uzorka u ovom radu (sa 355 naloga), izvršenje funkcije je trajalo 38 minuta.

Dobijena matrica je asimetrična, i prikazuje vrednosti, tj. ne radi se o Bulovoj matrici (binarnoj, logičkoj matrici). Pošto za prikaz čitave matrice nema prostora na papiru, u Tabela 2 je prikazan primer matrice povezanosti nekoliko naloga koji su ranije prikazani u Tabela 1.

Tabela 2 Segment matrice povezanosti dobijene primenom funkcije matrica_povezanosti() na podacima prikupljenim o nalogima iz Tabele 1.

	SasaRadulovich	PajticBojan	Demokrate	avucic	seselj_vojislav
SasaRadulovich	0	3	7	1	0
PajticBojan	1	0	19	3	0
Demokrate	0	139	0	0	0
avucic	0	0	0	0	0
seselj_vojislav	0	0	0	2	0

U prethodnom poglavlju pokazano je kako se preuzimaju metapodaci o samim nalogima, koji važe od otvaranja naloga do dana preuzimanja. Međutim ako nas zanima koliko tvitova, retvitova i omiljenih tvitova je bilo za vreme kampanje, to se mora uraditi obradom preuzetih podataka o svim tvitovima koje su nalozi tvitovali u tom periodu. Funkcija kojom je urađena agregacija ovih podataka nalazi se u Listingu 10, a rezultati su dati u Prilogu II.

```
retfav <- function(tvitovi) {
  sp <- unique(tvitovi$screenName)
  retfav <- matrix(ncol = 4, nrow = length(sp))
  for (i in 1:length(sp)) {
    l <- tvitovi[tvitovi$screenName == sp[i], ]
    retfav[i, 1] <- l$screenName[1]
    retfav[i, 2] <- sum(l[l$isRetweet == F, 3])
    retfav[i, 3] <- sum(l[l$isRetweet == F, 12])
    retfav[i, 4] <- length(l[l$isRetweet == F, 12])
  }
  retfav <- as.data.frame(retfav)
  colnames(retfav) <- c("Nalog", "Favorisan", "Retvitovan", "Tvitovi")
  retfav$Favorisan <- as.numeric(as.character(retfav$Favorisan))
  retfav$Retvitovan <- as.numeric(as.character(retfav$Retvitovan))
  retfav$Tvitovi <- as.numeric(as.character(retfav$Tvitovi))
  retfav$Pros.Ret <- retfav$Retvitovan / retfav$Tvitovi
  return(retfav)
}
```

Listing 10 Prebrojavanje broja tvitova i broja njihovih retvitova i favorite za svaki nalog pojedinačno

3.3 Kreiranje grafova

Kreiranje grafova kada posedujemo matricu povezanosti je veoma lako. U okviru paketa *igraph* postoji funkcija koja od matrice povezanosti pravi tzv. graf objekat (Listing 11) .

```
require(igraph)

mreza.pol <- graph.adjacency(mat.pol, mode="directed", weighted = TRUE)
mreza.pol.pu <- graph.adjacency(mat.pol.pu, mode="directed", weighted = TRUE)
```

Listing 11 Kreiranje grafova uz pomoć *igraph* paketa

Dostupni podaci nam omogućavaju da kreiramo usmereni graf sa ponderisanim ivicama, što se u funkciji `graph.adjacency()` kontroliše argumentima *mode* i *weighted*.

Paket *igraph* nam nudi mogućnost dodeljivanja atributa čvorovima i ivicama. Pošto posedujemo dodatne podatke o političarima u osnovnom uzorku, možemo neke od tih atributa dodeliti čvorovima u novostvorenoj mreži. Stranačka pripadnost je atribut koji će biti dodeljen

čvorovima. Moguće je dodeliti attribute i ivicama; ponderi, odnosno težina ivica, su primer atributa ivica (Listing 12).

```
V(mreza.pol) #funkcijom V() se vrši selekcija čvorova
E(mreza.pol) #funkcijom E() se vrši selekcija ivica

#atributi se mogu dodeliti pomoću gore navedenih funkcija

V(mreza.pol)$stranka <- as.factor(političari$Stranka)
V(mreza.pol)$prezime <- političari$Prezime

#dodeljivanje atributa se takođe može postići i ugrađenom funkcijom:
set_vertex_attr(mreza.pol, name="boja", value = političari$Boja)
```

Listing 12 Dodeljivanje stranačke pripadnosti kao atributa čvorova

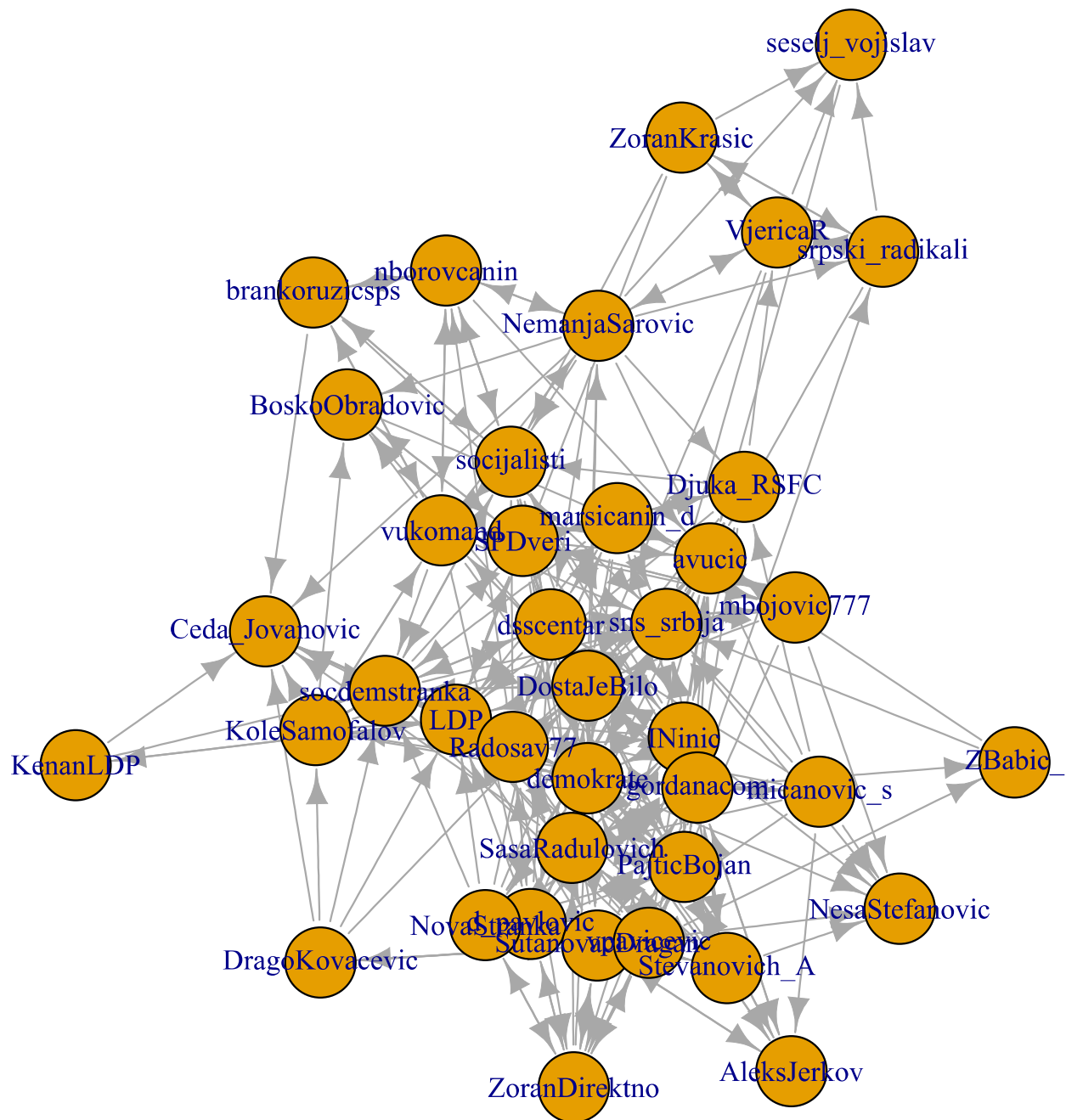
3.4 Vizualizacija grafova

Dobra vizualizacija može da komunicira informacije efikasnije od bilo kog teksta ili tabele sa metrikama koje opisuju mrežu. Zato je pri svakoj analizi društvenih mreža potrebno posvetiti pažnju baš vizualizaciji. Važno je vizualizovati graf tako da se intuitivno može gotovo trenutno protumačiti. Ipak, vizualizacija nije uvek uputna kada se radi o velikim i kompleksnim grafovima, kada je jedini način za analizu oslanjanje na metrike. Jedan od najvažnijih elemenata vizualizacije grafa je raspored (*layout*) čvorova njemu. Pošto je graf apstrakcija, raspored čvorova u 2D ili 3D prikazu je potpuno arbitraran. Vizualizacija se može konstruisati tako da se svaki čvor manuelno postavi na željenu poziciju. Takav pristup može biti čak i poželjan kada su u pitanju manje, proste mreže, ali kada se radi o iole većim mrežama potrebno je automatsko pozicioniranje čvorova u mreži. Za te potrebe su razvijeni brojni algoritmi čiji je zadatak da daju optimalan raspored, a među njima su najzastupljeniji tzv. silom usmereni (*force-directed graph drawing*) algoritmi.

Osim rasporeda čvorova u mreži, postoji i dodatni elementi vizualizacije grafa, koji mogu učiniti vizualizaciju efikasnijim sredstvom za komunikaciju informacija o mreži koju predstavlja. Ti elementi mogu biti vizualne karakteristike čvorova i ivica, čiji izgled može zavisiti od određene promenljive, koja može biti ili neka od mrežnih metrika, ili čak neka eksterna promenljiva. Primer vizualizacije mreže dat je u Listingu 13 i Slici 4

```
plot.igraph(mreza.pol)
```

Listing 13 Vizualizacija mreže političara bez dodatnih parametara vizualizacije (Slika 4)



Slika 4 Vizualizacija mreže bez dodatnih parametara vizualizacije (Listing 13)

Primer vizualizovanja mreže osnovnog uzorka sa kontrolom dodatnih vizuelnih elemenata dat je u Listingu 14.

```
require(scales)

#Za raspored čvorova primenjuje se force-directed Fruhtermann-Rajngold
algoritam

frl <- layout.fruchterman.reingold(mreza.pol, niter = 100000)

plot.igraph(
  mreza.pol,
```

```

layout = fr1, #dodeljivanje izračunatog rasporeda čvorova
edge.curved = FALSE, #ivice su prave linije
edge.width = log(E(mreza.pol)$weight),
edge.color = alpha("gray", rescale(
  x = log(E(mreza.pol)$weight), to = c(0.3, 1))),
edge.arrow.size = 0.2,
vertex.color = V(mreza.pol)$boja,
vertex.label.color = "black",
vertex.label.family = "Arial",
vertex.frame.color = NA,
vertex.label = V(mreza.pol)$prezime,
vertex.size = rescale(degree(mreza.pol), to=c(5,15)),
vertex.label.cex = 0.7,
vertex.label.dist = 0.5
)

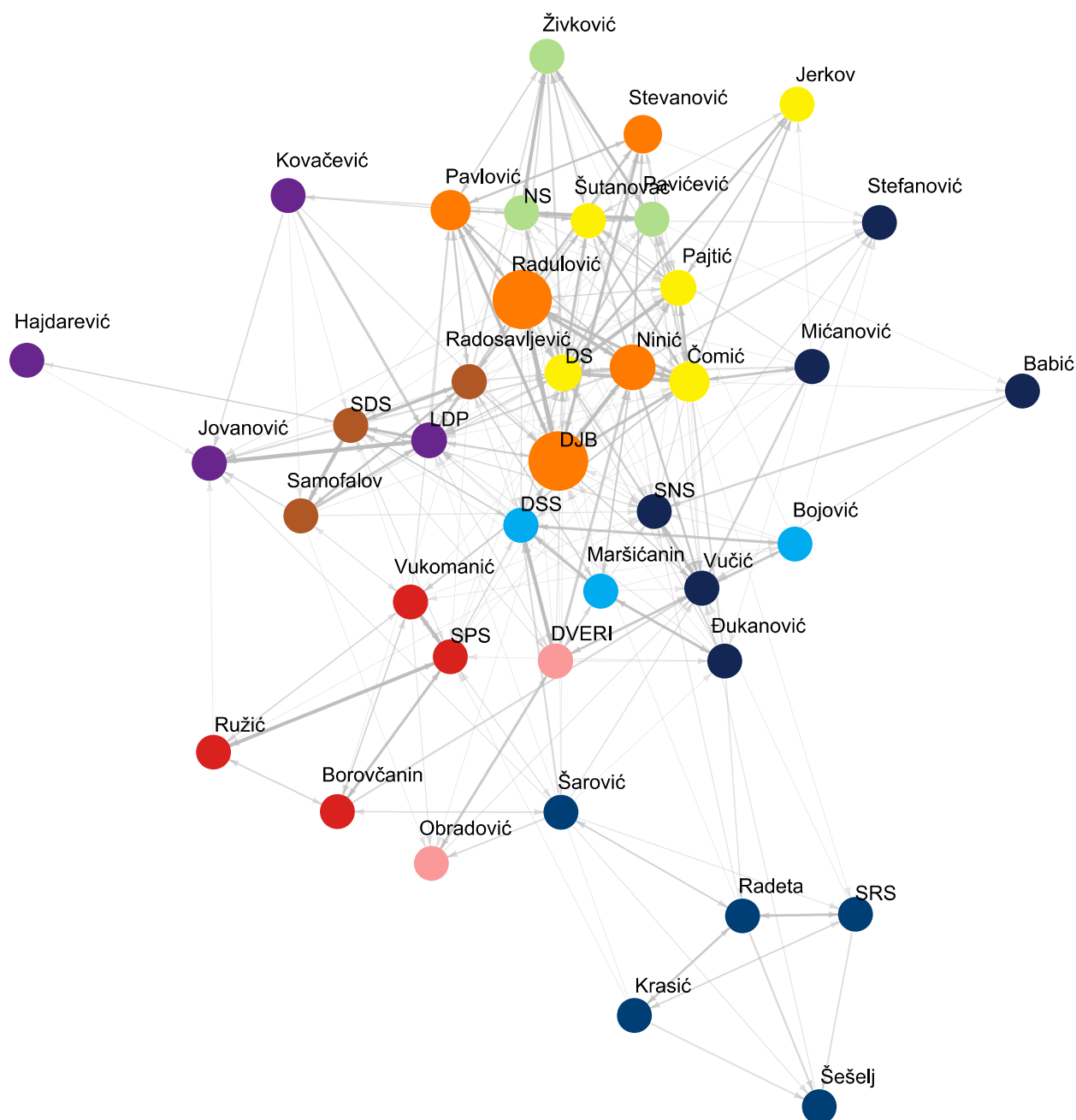
```

Listing 14 Vizualizacija mreže sa dodatnim parametrima vizualizacije (Slika 5)

Vizualizacija sa dodatnim parametrima nudi mnogo više informacija. Čvorovi su raspoređeni korišćenjem Fruchterman–Rajngold *force-directed* algoritma (Fruchterman & Reingold, 1991).

Debljina ivica kao i nivo njihove transparentnosti zavisi od težine, tj. broja pominjanja. Veličina čvora zavisi od stepena čvora. Umesto nasumičnog dodeljivanja boja čvorovima različitih stranaka, iskorišćeni su podaci iz unapred pripremljene datoteke političari.csv, koji su u Listingu 12 dodeljeni grafu. Još jedan način na koji je ista stav mogla da se uradi je da se umesto atributa grafa, koji je sastavni deo grafa, da bilo koji niz tekstualnih vrednosti, kao što je npr. celokupna kolona boja datoteke političari.csv, što bi u tom slučaju radilo jer su broj čvorova u grafu i njihov redosled identični onim u datoteci političari.csv. Da je npr. formiran podskup čvorova istog grafa tako nešto ne bi radilo, pa je zato bolje koristiti atribut grafa. Isto važi i za atribut prezime, koji je u vizualizaciji grafa iskorišćen za nazive čvorova, kao zamena za podrazumevane (*default*) vrednosti imena čvorova koja predstavljaju imena naloga na Tviteru.

Iako postoje dokazi da grafovi sa zakrivljenim ivicama efektivnije komuniciraju informacije u grafu (Xu, Rooney, Passmore, Ham, & Nguyen, 2012), u slučaju usmerenih mreža veće gustine, kao što je to slučaj sa grafom na Slici 5 broj ivica je prevelik pa prave ivice nude svojevrsnu simplifikaciju grafa, tako što obostrane ivice imaju dve strelice u suprotnim smerovima. U trenutnoj verziji *igraph* (1,01) paketa ne postoji mogućnost iscertavanja dvosmerne krive ivice. Zbog toga, kada bi se na isti graf primenile krive ivice, graf bi izgledao mnogo 'zamršenije', jer bi se tako udvostručio broj ivica između čvorova koji su obosmerno povezani.



Slika 5 Vizualizacija grafa sa dodatnim parametrima (Listing 14)

Poseban izazov predstavlja vizualizacija velikih mreža. Kada bi se mreža dobijena na osnovu proširenog uzorka sa 355 čvorova (i njihovih 13597 ivica) predstavila metodom koja je korišćena za predstavljanje 40 čvorova (sa 320 ivica), ona ne bi efikasno komunicirala informacije. Jedan od načina za vizualizaciju tako velikih grafova je grafikon košnice (*hive plot*). Kod grafikona košnice ne postoji raspored čvorova kao kod običnih sociograma. Naime, čvorovi se nalaze na osama, a pozicija na osi se određuje pomoću metrika. To znači da je grafikon košnice determinističan, tj. da je od istih podataka uvek moguće kreirati isti grafikon, jer se ne zasniva na nasumičnom generisanju rasporeda kao što je to slučaj sa konvencionalnim sociogramima (Krzywinski, Birol, Jones, & Marra, 2012). Ovaj metod je nastao za potrebe vizualizacije podataka iz oblasti genomike.

Tvorci grafikona košnice su upravo bili motivisani teškoćom da se interpretiraju veći sociogrami, koji, prema njihovim rečima, podsećaju na 'klupko dlaka', koje je nemoguće interpretirati. Sa druge strane, grafikon košnice je veoma lako interpretirati: date su ose, pozicija čvorova na osama zavisi od izabrane metrike i ti čvorovi su međusobno povezani (dobra vizualna reprezentacija prednosti grafikona košnice prikazana je na veb sajtu tvorca ovog načina vizualizacije <http://hiveplot.net>). Primer vizualizacije grafikona košnice u R-u dat je u Listingu 15.

```
require(HiveR)

#Paket HiveR zahteva matricu povezanosti kao ulaz
hplot <- adj2HPD(mat.pol.pu)
hplot <- mineHPD(hplot, option = "rad <- tot.edge.count") #stepen (na osama)

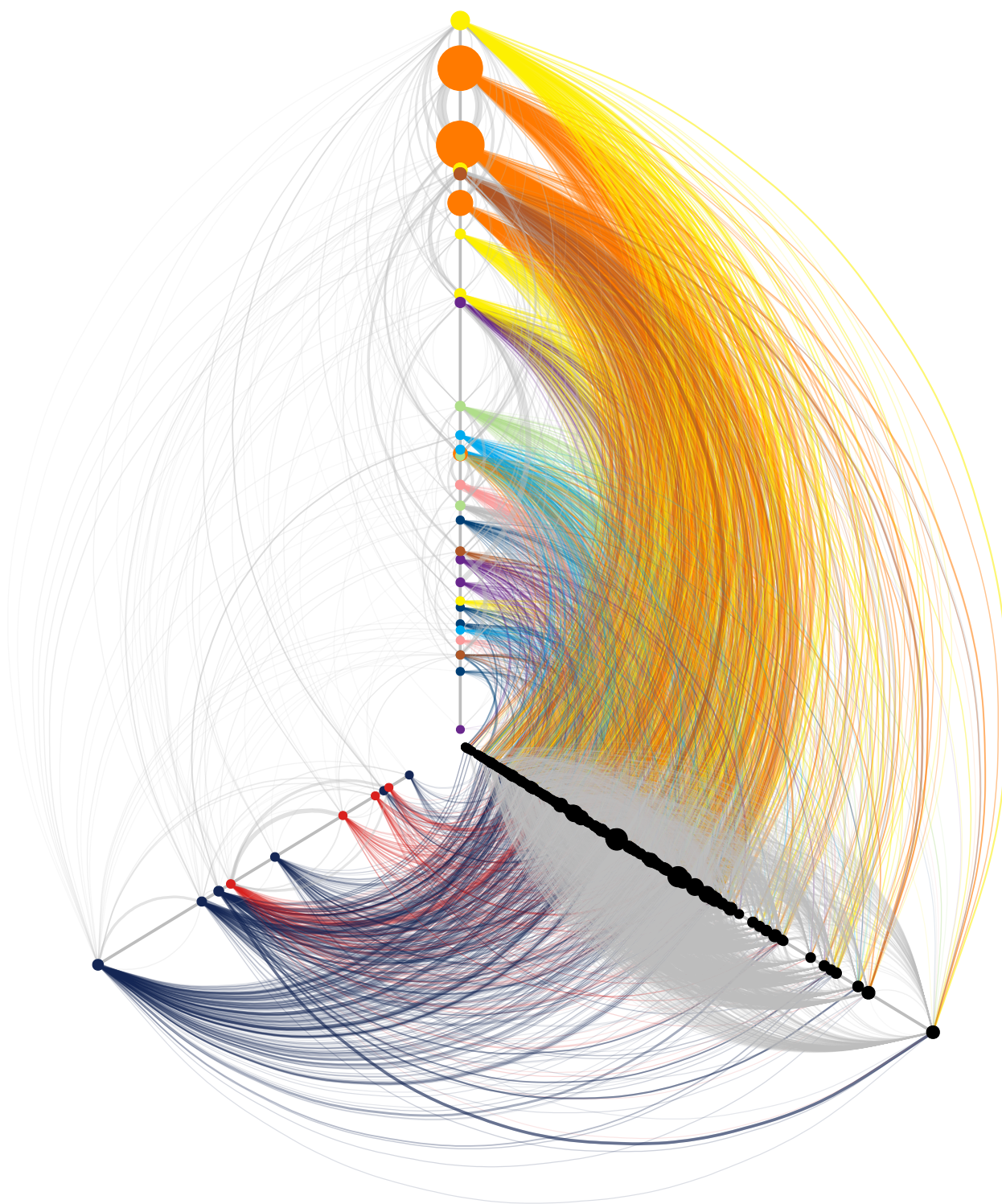
#Formiranje osa na osnovu statusa: opozicija(1), vlast(2), ostali(3)
hplot$nodes$axis[1:40] <- as.integer(političari$Opozicija)#opozicija i vlast
hplot$nodes$axis[41:355] <- as.integer(3) #ostali tviteraši
hplot$nodes$color[1:40] <- političari$Boja #boja čvorova
hplot$axis.cols <- c(rep("#bdbdbd", 3)) #boja osa

#Dodeljivanje boja ivicama na osnovu stranačke pripadnosti. Svaka ivica koja
je između čvora koji pripada političkoj stranci i nepoznatog čvora (dodatog
u proširenom uzorku) biće obojena bojom političke stranke kojoj čvor koji je
u stranci pripada. Ostale ivice biće sive boje.

hplot$edges$color <- "#bdbdbd"
for (i in 1:length(unique(političari$Stranka))) {
  hplot$edges$color[E(mreza.pol.pu)[V(mreza.pol.pu)[V(mreza.pol.pu)$name %in%
političari$Nalog[političari$Stranka == unique(političari$Stranka)[i]]] %--%
V(mreza.pol.pu)[V(mreza.pol.pu)$opozicija == 3]]] <-
unique(političari$Boja[političari$Stranka == unique(političari$Stranka)[i]])
}
#Dodavanje transparentosti ivicama
hplot$edges$color <-
  paste0(hplot$edges$color, substring(alpha("gray", rescale(
    x = log(E(mreza.pol.pu)$weight), to = c(0.1, 0.5)
  )), 8, 9))

#uklanjanje ivica koje počinju i završavaju se na istom mestu na jednoj osi.
U ovom konkretnom slučaju ne radi se o čvorovima koju su pominjali sami sebe
već su to oni čvorovi koji imaju isti stepen i nalaze se na istoj osi.
Trenutna implementacija paketa HiveR (0.2.55) ne radi ukoliko se prethodno
ne uklone ivice između takvih čvorova, pa je sledeći korak neophodan.
hplot <- mineHPD(hplot, option = "remove zero edge")
#Veličina čvora zavisi od centralnosti svojstvenog vektora
hplot$nodes$size <-
  rescale(x = as.numeric(eigen_centrality(mreza.pol.pu)$vector), to = c(1,
6))
hplot$edges$weight <- rescale(hplot$edges$weight, to = c(1, 10))
plotHive(hplot, bkgnd = "white")
```

Listing 15 Implementacija grafikona košnice (*hive plot*) u R-u pomoću paketa *HiveR* (Slika 6)



Slika 6 Vizualizacija grafa proširenog uzorka pomoću grafikona košnice (*hive plot*) (Listing 15). Ose: leva – vlast, srednja – opozicija, desna – nalozi dodati u proširenom uzorku. Pozicija čvorova na osi zavisi od stepena; veličina čvorova zavisi od centralnosti svojstvenog vektora (*Eigenvector centrality*).

U grafikonu košnice proširenog uzorka čvorovi su podeljeni u tri vrste: one koji pripadaju vladajućoj većini (koja se nije promenila ni posle izbora), strankama opozicije (one koje zvanično nisu deo vladajuće većine) i čvorove čija stranačka afilijacija nije poznata (one čvorove koji se ne nalaze u osnovnom uzorku). Pozicija čvorova na osama (praktično na poluprečnicima 'kružnica') zavisi od stepena, tj. broja ivica svakog čvora. Tako se čvorovi sa najvišim stepenom, kao što su Aleksandar Vučić kod vlasti i Gordana Čomić kod opozicije, nalaze na krajevima svojih osa, tj. imaju najveći poluprečnik. Veličina čvora zavisi od centralnosti svojstvenog vektora. Debljina ivica, kao i nivo njihove transparentnosti zavisi od težine ivica, tj. učestalosti komunikacije, a boja od stranačke pripadnosti jednog od čvorova. Međusobna komunikacija između članova političkih stranaka, kao i međusobna komunikacija čvorova koji nisu pripadali osnovnom uzorku je obojena sivom bojom.

Primer vizualizacije mreže proširenog uzorka konvencionalnim sociogramom, sa sličnim parametrima kao grafikonu košnice, dat je u Prilogu IV.

3.5 Izračunavanje mrežnih metrika

Samo izračunavanje osnovnih mrežnih metrika je relativno lako, u poređenju sa prethodnim koracima koji se odnose na sakupljanje i sređivanje podataka. Proces izračunavanja se svodi na pozivanje funkcija izabranog R paketa. U ovom radu korišćen je *igraph* paket za izračunavanje svih mrežnih metrika, ali je mnoge metrike moguće izračunati pomoću drugih paketa pomenutih u poglavlju 1.3.2.1.

3.5.1 Globalne metrike mreža

Globalni pokazatelji grafa su neophodni kako bi se mreža u potpunosti sagledala. Vizualizacija može da pomogne u intuitivnom razumevanju mreže, ali zbog svoje nasumičnosti i arbitrarnosti može takođe dosta toga da sakrije. U Listingu 16 data je funkcija za izračunavanje metrika celog grafa.

```
met_graf <- function(graf, atribut=NULL) {  
  met_graf <- list(  
    "Broj čvorova" = vcount(graf),  
    "Broj ivica" = ecount(graf),  
    "Gustina" = graph.density(graf),  
    "Gustina neusmereni" = graph.density(as.undirected(graf, mode =  
"mutual")),  
    "Povezanost" = vertex.connectivity(graf),  
    "Povezanost neusmereni" = vertex.connectivity(as.undirected(graf, mode =  
"collapse")),  
    "Dijametar" = diameter(graf, weights = NA),
```

```

"Najudaljeniji" = paste(
  farthest_vertices(graf, weights = NA)$vertices$name[1],
  farthest_vertices(graf, weights = NA)$vertices$name[2],
  sep = " -- "),
"Prosečna dužina putanje" = average.path.length(graf),
"Tranzitivnost" = transitivity(graf),
"Asortativnost stepena" = assortativity.degree(graf),
"Asortativnost nom" = if(!is.null(atribut))
{assortativity.nominal(graf, get.vertex.attribute(graf, atribut))},
"Centralizacija intermedijanosti" = centralization.betweenness(graf,
directed = TRUE, normalized = TRUE)$centralization,
"Centralizacija stepena" = centralization.degree(graf, normalized = TRUE,
mode = "all")$centralization,
"Centralizacija bliskosti" = centralization.closeness(graf, mode = "all")
$centralization,
"Centralizacija svojstvenog vektora" = centralization.evcent(graf,
directed = TRUE, normalized = TRUE)$centralization
)
return(met_graf)
}

```

Listing 16 Izračunavanje metrika celog grafa

Kada je u pitanju izračunavanje globalnih metrika mreže, a naročito njihova interpretacija, veoma važno je voditi računa o tipu mreže za koju se metrike računaju. Jedna metrika može imati mnogo drugačiju interpretaciju u zavisnosti od toga da li je mreža usmerena ili neusmerena, kao i od toga da li se uzimaju u obzir težine ivica. Mrežu je uvek moguće učiniti jednostavnijom, tj. konvertovati iz oblika koji nosi više informacija u oblik koji nosi manje informacija. Tako npr. usmerenu mrežu je moguće konvertovati u neusmerenu, ali ne i obrnuto. Zato je bolje, pri svakoj analizi konstruisati mrežu koja nosi što više informacija, koliko to podaci dozvoljavaju, jer je simplifikacija uvek moguća. U Listingu 16, konverzija grafa u jednosmerni je korišćena za one metrike kod kojih interpretacija ima više smisla kod neusmerenih grafova. Korišćena su dva tipa konverzije u neusmerenu mrežu: za pokazatelj gustina mreža je konvertovana tako da su sačuvane samo ivice kada je bilo obostrane komunikacije između dva čvora, dok je za pokazatelj povezanosti i dijametra korišćena metoda koja je sačuvala po jednu ivicu za svaka dva povezana čvora. Isti metod je primenjen i za izračunavanje metrika centralizacije stepena i bliskosti. Metrike gustine je izračunata i za usmerenu i za neusmerenu verziju grafa zbog poređenja sa neusmerenim društvenim mrežama. Kod metrike povezanosti je pristup simplifikacije izabran zato što je vrednost ove metrike 0 za nepovezane grafove, a oba grafa na kojima je ova metrika primenjena su u svojim usmerenim varijantama nepovezani. Za izračunavanje dijametra i za pronalaženje najudaljenijih čvorova u mreži zanemarene su težine ivica. Metrika nominalne asortativnosti (atributa) zahteva kategorijsku varijablu, koja treba da bude atribut čvora da bi funkcija radila, ako atribut nije dat,

funkcije neće izračunati nominalnu asortativnost, već samo asortativnost stepena. Rezultati primene funkcije iz Listinga 16 na mreže osnovnog i proširenog uzorka dati su u Tabela 3.

Tabela 3 Rezultati globalnih metrika funkcije iz Listinga 16 primenjene na obe analizirane mreže

Metrika	Osnovni uzorak	Prošireni uzorak
Broj čvorova	40	355
Broj ivica	320	13597
Gustina	0,21	0,11
Gustina (kao neusmerena mreža ¹)	0,10	0,04
Povezanost	0	0
Povezanost (kao neusmerena mreža ²)	3	1
Dijametar *	6	5
Najudaljeniji čvorovi*	seselj_vojislav -- NemanjaSarovic	avucic -- antarizana
Prosečna dužina putanje	2,28	2,09
Tranzitivnost	0,47	0,35
Asortativnost stepena	-0,058	-0,166
Asortativnost atributa	(atr. stranka) 0,226	(atr. opozicija) -0,039
Centralizacija intermedijarnosti	0,13	0,08
Centralizacija stepena (kao neusmerena mreža ²)	0,24	0,38
Centralizacija bliskosti (kao neusmerena mreža ²)	0,33	0,43
Centralizacija svojstvenog vektora	0,73	0,82

1 Ivice postoje između dva čvora samo ako su se međusobno pominjali

2 Ivice postoje između dva čvora ako je barem jedan od čvorova pomenuo drugi čvor

* Bez uzimanja u obzir težina ivica (pondera)

3.5.2 Metrike centralnosti čvorova

Umesto zasebnog izračunavanja metrika, u Listingu 17 prikazana je funkcija koja za ulaz uzima mrežu, a kao rezultat daje tabelu sa metrikama centralnosti i kompozitnim pokazateljem. Rezultati te funkcije za mrežu osnovnog uzorka dati su u Tabeli 4, dok su rezultati za prošireni uzorak dati u Prilogu III (samo 100 najcentralnijih čvorova, potpuna lista je dostupna u pratećim datotekama elektronske verzije rada).

```
require(p2distance)

met_cent <- function(graf) {
  cent <- as.matrix(
    cbind(
      betweenness(graf),
      degree(graf, mode = "in"),
      degree(graf, mode = "out"),
      closeness(graf, mode = "in"),
      closeness(graf, mode = "out"),
      eigen_centrality(graf)$vector
    )
  )
}
```

```

)
colnames(cent) <-
  c(
    "intermedijarnosti",
    "dolaznog stepena",
    "odlaznog stepena",
    "dolazne bliskosti",
    "odlazne bliskosti",
    "svojstvenog vektora"
  )
p2d <- p2distance::p2distance(cent)
cent <- as.data.frame(cent)
cent$P2odstojanje <- as.vector(p2d$p2distance)
print(paste("Redosled pokazatalja: ", p2d$variables_sort))
return(cent[order(-cent$P2odstojanje), ])
}

```

Listing 17 Funkcija za izračunavanje mera centralnosti

Prilikom izračunavanja svih metrika centralnosti čvorova korišćeni su podaci o smeru ivica, kao i njihove težine. Pošto sve mere centralnosti oslikavaju drugačiji način gledanja ne centralnost u mreži, napravljen je kompozitni pokazatelj (P_2 odstojanje), koji objedinjuje informacije svih mera centralnosti (vidi poglavlje 2.2).

Tabela 4 Mere centralnosti za graf osnovnog uzorka sa Slika 5

Tviter nalog	Centralnost						P_2 odstojanje
	intermedijarnosti	dolaznog stepena	odlaznog stepena	dolazne bliskosti	odlazne bliskosti	svojstvenog vektora	
DostaJeBilo	238,92	16	13	0,006	0,006	1,000	12,88
gordanacom	218,61	13	18	0,005	0,006	0,228	10,30
PajticBojan	325,57	16	8	0,006	0,006	0,045	9,73
SasaRadulovich	36,97	10	11	0,006	0,005	0,984	9,27
vpavicevic	291,50	10	11	0,006	0,007	0,009	8,93
demokrate	122,68	25	9	0,006	0,006	0,084	8,59
LDP	213,61	22	5	0,007	0,005	0,037	8,54
SPDveri	213,44	6	10	0,006	0,006	0,020	7,34
socijalisti	180,45	12	7	0,006	0,006	0,003	7,21
Radosav77	71,19	6	19	0,004	0,007	0,018	7,17
SutanovacDraga n	122,99	10	14	0,004	0,007	0,017	7,17
INinic	13,73	9	14	0,004	0,005	0,439	7,10
sns_srbija	155,44	17	4	0,007	0,005	0,007	6,94
dsscentar	133,40	11	9	0,006	0,006	0,012	6,93
NemanjaSarovic	171,44	4	11	0,005	0,006	0,002	6,55
sodemstranka	58,99	12	10	0,006	0,005	0,012	6,36
NovaStranka	100,02	6	12	0,005	0,007	0,010	6,33
marsicanin_d	102,31	5	12	0,005	0,007	0,008	6,13
vukomand	119,10	6	11	0,004	0,007	0,004	6,01

Tviter nalog	Centralnost						P ₂ odstojanje
	intermedijarnosti	dolaznog stepena	odlaznog stepena	dolazne bliskosti	odlazne bliskosti	svojstvenog vektora	
d_pavlovic	13,37	8	12	0,005	0,006	0,215	5,99
avucic	90,67	20	1	0,007	0,003	0,010	5,66
mbojovic777	111,39	3	8	0,004	0,007	0,002	5,05
Djuka_RSFC	44,96	4	9	0,005	0,006	0,003	4,61
Stevanovich_A	8,45	5	9	0,004	0,005	0,152	4,54
KoleSamofalov	46,15	4	9	0,004	0,007	0,006	4,37
ZoranDirektno	6,46	7	7	0,005	0,004	0,013	3,90
Ceda_Jovanovic	23,15	9	2	0,005	0,005	0,015	3,61
nborovcanin	19,85	5	5	0,004	0,005	0,000	3,32
srpski_radikali	39,00	4	4	0,005	0,003	0,000	3,21
micanovic_s	0,00	1	11	0,001	0,008	0,011	3,15
DragoKovacevic	14,52	1	7	0,003	0,006	0,005	3,12
VjericaR	0,00	4	6	0,004	0,005	0,001	3,09
KenanLDP	21,83	1	3	0,005	0,005	0,000	2,94
AleksJerkov	1,67	5	4	0,004	0,004	0,009	2,73
BoskoObradovic	1,83	5	2	0,005	0,004	0,002	2,70
ZoranKrasic	21,50	2	5	0,003	0,006	0,000	2,58
brankoruzicsps	0,00	3	5	0,003	0,006	0,001	2,56
NesaStefanovic	0,00	7	0	0,006	0,001	0,005	2,39
ZBabic_	0,00	2	2	0,005	0,002	0,000	2,16
seselj_vojislav	0,00	4	1	0,004	0,002	0,000	1,79

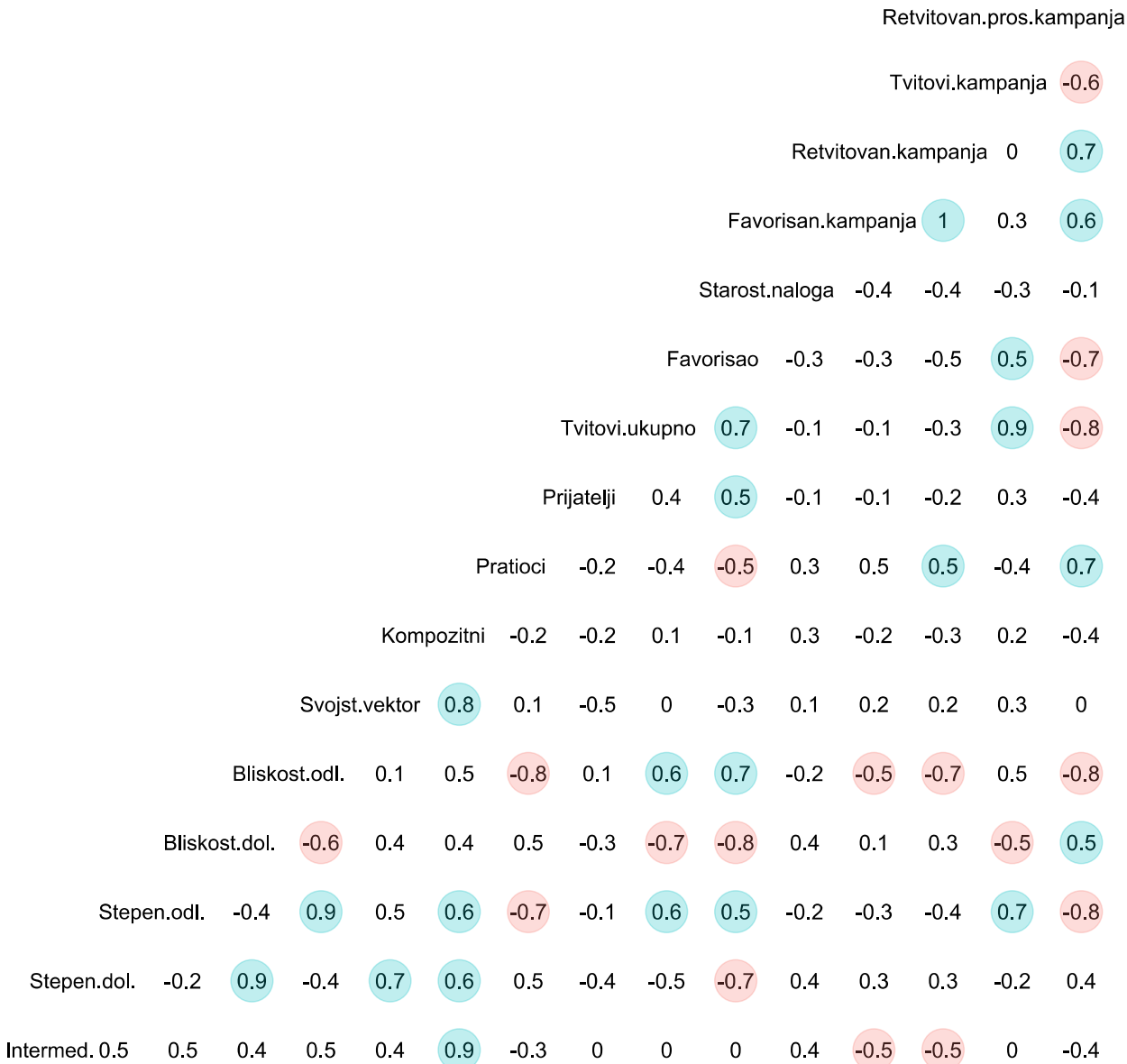
Kao što je navedeno u poglavlju o metodama (2.2) proces izračunavanja P₂ odstojanja je postupan tako da se prvo uzima u obzir varijabilitet promenljive sa najvećom količinom informacija, zatim varijabilitet druge po redu promenljive, a koji nije bio sadržan u prvoj promenljivoj itd. Redosled promenljivih, prema količini informacija koje nose, za formiranje P₂ odstojanja kada je u pitanju osnovni uzorak, je sledeći: (centralnosti) intermedijarnosti, dolaznog stepena, odlaznog stepena, svojstvenog vektora, dolazne bliskosti i na kraju odlazne bliskosti. Korelacija kompozitnog pokazatelja sa merama centralnosti, ali i sa podacima o nalogima koji su sakupljeni, kao i sa podacima o broju tvitova, retvitova i sl. dobijenih obradom korpusa tvitova tvitovanih za vreme kampanje izračunata je na način prikazan u Listingu 18.

```
svipokazatelji <-
  cbind(met_cent(mreza.pol), #metrike centralnosti grafa
        nalozMeta(politiciari$Nalog)[-1], #metrike naloga na Tviteru
        retfav(tvitovi)[-1]) #metrika naloga za vreme kampanje

#Izračunavanje starosti naloga u nedeljama
svipokazatelji$DatumOtvaranja <-
  as.numeric(abs(
    difftime(svipokazatelji$DatumOtvaranja, Sys.time()), units = "weeks"))
```

```
))
cor(svipokazatelji, method = "spearman")
```

Listing 18 Korelaciona matrica svih dostupnih pokazatelja naloga u mreži političara (Slika 7)



Slika 7 Korelaciona matrica pokazatelja za Tviter naloge u osnovnom uzorku

Nakon izračunatih vrednosti centralnosti za sve čvorove, postavlja se pitanje kakva je zbirna ocena centralnosti čvorova prema stranačkoj pripadnosti. Dakle koja stranka je zbirno imala najcentralniju poziciju? Agregacija je izvršena na način prikazan u Listingu 19

```
V(mreza.pol)$stranka <- as.character(politiciari$Stranka)

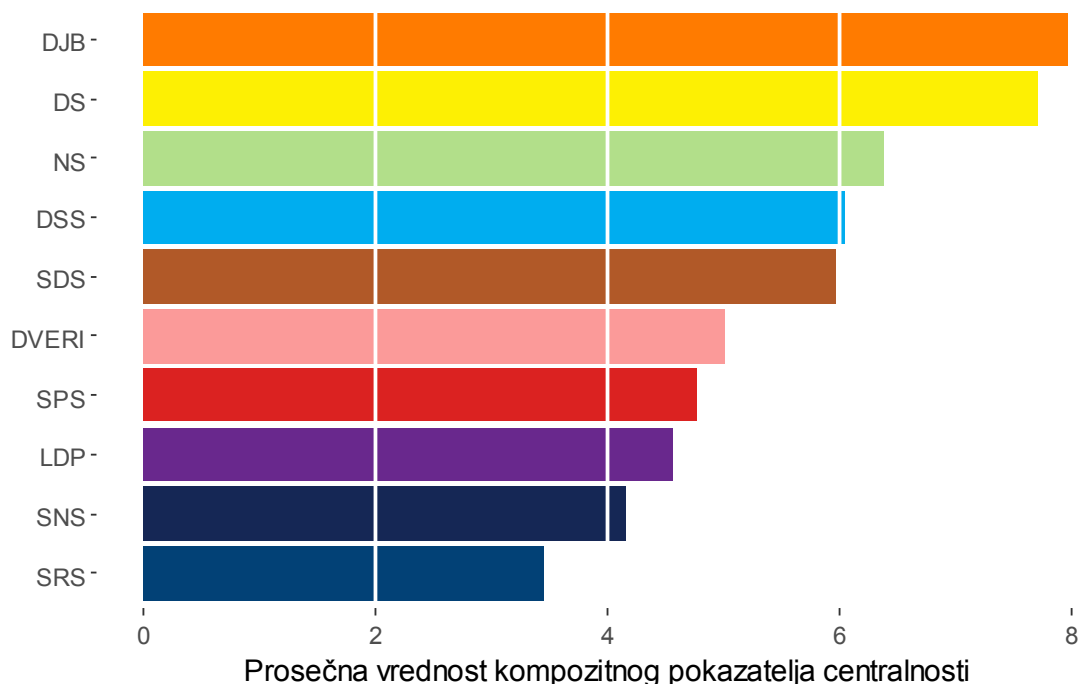
g <- aggregate(
```

```
gg <- aggregate(
```

```
x = met_cent(mreza.pol)$P2odstojanje,
FUN = sum,
by = list(
  get.vertex.attribute(mreza.pol, "boja"),
  get.vertex.attribute(mreza.pol, "stranka")
))
g$Broj <- as.numeric(table(get.vertex.attribute(mreza.pol, "stranka")))
colnames(g1) <- c("Boja", "Stranka", "Centralnost", "Broj")
g$CentStand <- g$Centralnost / g$Broj
```

Listing 19 Agregacija centralnosti po stranačkoj pripadnosti

Agregacija je izvršena po atributu stranke (i pratećim atributom boje, zbog grafičkog predstavljanja). Sumiran je kompozitni pokazatelj centralnosti, koji je nakon toga standardizovan brojem čvorova. Naime, pošto su neke stranke u uzorku bile reprezentovane sa manje čvorova a neke sa više, potrebno je bilo da se izračuna prosečna vrednost kompozitnog pokazatelja centralnosti čvorova po strankama. Rezultati agregacije iz Listing 19 dati su na Slika 8.



Slika 8 Prosečna vrednost kompozitnog pokazatelja centralnosti čvorova po strankama u mreži osnovnog uzorka

3.5.3 Homofilija

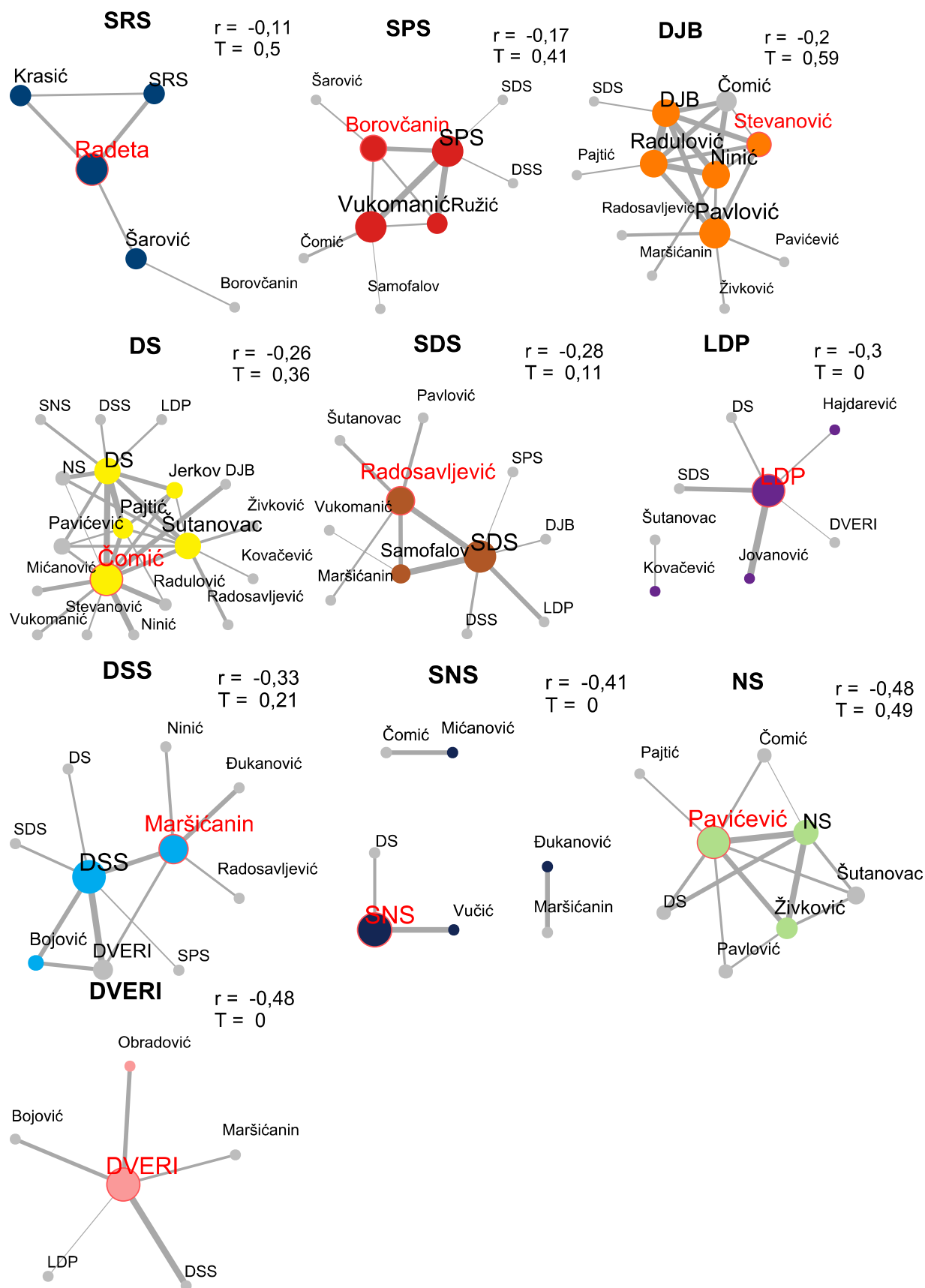
Na osnovu već izračunatih metrika mogu se izvući zaključci o homofilčnosti čvorova u ispitanim mrežama, ali i posle te analize ostaje pitanje: da li su neke stranke homofilnije od drugih? Da bi se odgovorilo na ovo pitanje, izračunavanje mera homofilije po strankama za osnovni uzorak izvršeno je tako što su napravljeni neusmereni podgrafovi koji obuhvataju sve članove jedne stranke i sve ostale čvorove koji su sa njima imali obostranu komunikaciju. Potom

se nad takvim podgrafovima izračunavaju metrike asortativnosti i tranzitivnosti, i identifikuje se čvor sa najvećom vrednošću intermedijarne centralnosti. Analiza podgrafova kao i njihova vizualizacija dati su u Listingu 20 i Slici 9.

```
homofilija <- vector()
tranzitivnost <- vector()
V(mreza.pol)$stranka <- političari$Stranka
for (i in 1:length(unique(V(mreza.pol)$stranka))) {
  x <-
    subgraph.edges(mreza.pol, as.vector(E(mreza.pol)[V(mreza.pol)
[V(mreza.pol)$stranka == unique(V(mreza.pol)$stranka)[i]] %--%
1:length(V(mreza.pol))]))
  x <- as.undirected(x, mode = "mutual")
  x <- delete.vertices(x, v=degree(x)==0)
  V(x)$stranka <- as.factor(V(x)$stranka)
  homofilija[i] <- assortativity.nominal(x, types = V(x)$stranka)
  tranzitivnost[i] <- transitivity(x)
  intermed <- rep(NA, vcount(x))
  intermed[which(betweenness(x) == max(betweenness(x)))] <- "#ff5658"
  svg(
    paste0(unique(V(mreza.pol)$stranka)[i], ".svg"),
    height = 4.13,
    family = "Arial"
  )
  plot.igraph(
    x,
    edge.curved = F,
    layout = layout.davidson.harel(
      x,
      weight.node.dist = 70,
      weight.node.edge.dist = 70
    ),
    edge.width = log(E(x)$weight),
    vertex.color = ifelse(V(x)$boja==unique(V(mreza.pol)$boja)[i],
unique(V(mreza.pol)$boja)[i], "#bdbdbd"),
    vertex.label.color = ifelse(is.na(intermed), "black", "red"),
    vertex.label.dist = 1.3,
    vertex.frame.color = intermed,
    vertex.label = V(x)$prezime,
    vertex.size = rescale(x=degree(x), to=c(10,30)),
    vertex.label.cex = rescale(x=degree(x), to=c(0.8,1.2)),
    vertex.label.family = "Arial",
    main = unique(V(mreza.pol)$stranka)[i]
  )
  text(0.7, 1.6, paste("r = ", gsub("*\\.", "", as.character(
    round(homofilija[i], 2)
))), pos = 4)
  text(0.7, 1.4, paste("T = ", gsub("*\\.", "", as.character(
    round(tranzitivnost[i], 2)
))), pos = 4)
  dev.off()
}
```

Listing 20 Asortativnost i tranzitivnost vizualizovanih podgrafova pojedinačnih stranaka u osnovnom uzorku (Slika 9)

Izabrani pristup mapiranja podgrafova tako da se u obzir uzima samo obostrana komunikacija je izabran zato što bi u suprotnom mogli da budu mapirani odnosi koji zbirno ne bi dovodili do sociološkog fenomena homofilije. Uzimo npr. čvor Aleksandar Vučić, koji je meta mnogih opaski i prozivki na Tviteru, na koje ne odgovara, tj. ima mnogo tvitova u kojima je on pominjan, a malo ili nijedan tvit u kojima on pominje druge. Takav čvor bi smanjio vrednost koeficijenta asortativnosti podgraфа svoje stranke, jer je u 'komunikaciji' sa mnogim čvorovima drugih stranki, dok je realna situacija potpuno obrnuta. Isto važi i na obrnutom primeru čvora Slaviše Mićanovića koji pominje mnogo članova drugih stranaka, ali mu niko ne odgovara (Tabela 4). Osim koeficijenata asortativnosti i tranzitivnosti na Slika 9 je za svaki podgraf identifikovan i čvor za najvećom vrednošću intermedijarne centralnosti. Drugim rečima identifikovan je čvor koji je najzaslužniji za povezanost čvorova u svakom podgrafu. Na Slici 9 ti čvorovi se izdvajaju tako što im je natpis i okvir čvora crvene boje.



Slika 9 Podgrafovi mreže osnovnog uzorka po strankama sa izračunatim koeficijentom asortativnosti (r) i tranzitivnosti (T), čvorovi sa najvećim vrednostima intermedijarnosti su uokvireni i njihovi nazivi su ispisani crvenom bojom

3.6 Klaster analiza

3.6.1 Izbor klaster algoritma

Zadatak klaster analize u ovom radu je da utvrdi da li se pripadnost čvorova strankama može utvrditi algoritmom. Većina algoritama za detekciju zajednica u mreži ne podržava usmerene mreže, pa je simplifikacija usmerene mreže standardna praksa za pronalaženje zajednica (Missaoui & Sarr, 2014). Ipak, postoje algoritmi koji podržavaju i usmerene mreže, a *igraph* R paket implementira 5 takvih algoritma:

- **Algoritam intermedijarnosti ivica** (*Edge Betweenness*) iskorišćava činjenicu da ivice koje povezuju zasebne zajednice imaju više vrednosti intermedijarnosti jer sve najkraće putanje između različitih zajednica sigurno prolaze kroz njih (Newman & Girvan, 2004).
- **Algoritam širenja oznaka** (*Label Propagation*) se ostvaruje u više koraka; u prvom se svim čvorovima dodele jedinstvene oznake i u svim narednim koracima svaki čvor preuzima oznaku koju većina njegovih 'komšija' trenutno ima (u slučaju nerešenog rezultata nasumično se odabira), a iteracijom se dolazi do konačnog rešenja (Raghavan, Albert, & Kumara, 2007).
- **Algoritam spinsko staklo** (Spin-glass) je koncept preuzet iz oblasti statističke fizike i baziran je na tzv. Potsovom modelu, tako da svaka čestica (čvor) može biti u jednom od z spin stanja (broja zajednica), a interakcije između čestica (tj. ivice grafa) određuju koji par čvorova bi preferirao da ostane u istom spin stanju, a koji bi preferirao drugačije spin stanje, proces se ponavlja određeni broj puta, a finalno stanje spina određuje zajednice (Nepusz, 2012; Reichardt & Bornholdt, 2006).
- **Algoritam walktrap** (*walk* – šetnja, *trap* – zamka) algoritam, koji pokušava da pronađe gusto povezane podgrafove (zajednice) koristeći kratke slučajne šetnje, koje imaju tendenciju da ostanu u istoj zajednici, pa je tako dobio ime (Pons & Latapy, 2005). Postoje dokazi da ovaj algoritam generalno najpreciznije detektuje zajednice u mreži (Sousa & Zhao, 2014).
- **Algoritam infomap** je relativno nov algoritam, koji je takođe baziran na slučajnim šetnjama, ali radi tako što koristi verovatnoću toka slučajnih šetnji kroz mrežu, kao zamenu za tok informacija u realnom sistemu (Rosvall & Bergstrom, 2008). Mreža se dekomponuje u module kompresijom opisa verovatnoće toka, a kao rezultat se dobija mapa koja ujedno simplifikuje i ističe parterne u strukturi i njihove veze (Rosvall et al., 2009).

Primena algoritama u ovom radu nije ograničena na 'slepo' otkrivanje zajednica. Naime, stranačka pripadnost čvorova u mreži osnovnog uzorka je u potpunosti poznata, dok je u drugoj mreži, proširenog uzorka, poznata politička pripadnost barem 11.26% od ukupnog broja čvorova (udeo koji čvorovi iz originalne mreže uzimaju u ukupnom broju čvorova druge mreže). Poznata pripadnost čvorova strankama nam omogućava da ocenimo koliko dobro se rezultati svakog algoritma za detekciju zajednica poklapaju sa realnom stranačkom pripadnošću. U Listingu 21 je data funkcija koja broji 'pogrešno' grupisane čvorove, tj. one čvorove koje su analizom grupisanja svrstani u grupu kojoj prema svom atributu ne pripadaju.

```
klast_eval <- function(graf, atribut, metod) {
  klaster <- eval(parse(text = metod))(graf) #primena klaster metoda na mrežu
  l <- unique(na.omit(get.vertex.attribute(graf, atribut)))#poznati atributi
  y <- vector()
  x <- matrix(ncol = max(klaster$membership), nrow = length(l))
  for (i in 1:max(klaster$membership)) { #za svaki klaster
    for (j in 1:length(l)) { #za svaku vrednost atributa (stranku)
      y[j] <-
        table(unlist(communities(klaster)[i]) %in% V(graf)
$name[get.vertex.attribute(graf, atribut) == l[j]])["TRUE"] #br. pripadnika j atributa u i klas. }
      x[, i] <- y
    }
    x[is.na(x)] <- 0
    return(sum(apply(x, 2, sum) - apply(x, 2, max)))
#izlaz je razlika br. čvorova klastera i br. čvorova najbrojnije grupacije klastera
  }
```

Listing 21 Funkcija za evaluaciju podobnosti različitih metoda klasterovanja

Funkcija za evaluaciju metoda grupisanja iz Listingu 21 je potom primenjena na mreže osnovnog i proširenog uzorka, a uzeti su obzir svi dostupni algoritmi za detekciju zajednica koji podržavaju usmerene mreže (Listing 22).

```
metode <-
  c(
    "cluster_edge_betweenness",
    "cluster_label_prop",
    "cluster_springlass",
    "cluster_walktrap",
    "cluster_infomap"
  )
  for (i in 1:length(metode)) {
    print(paste(metode[i], klast_eval(mreza.pol, "stranka", metode[i])))
  }
```

Listing 22 Primena funkcije klast_eval() na mrežu osnovnog uzorka i dostupne klaster metode

Rezultati evaluacije za mreže osnovnog i proširenog uzorka dati su u Tabeli 5. Klaster metoda čije grupisanje najviše odgovara stranačkoj pripadnosti čvorova je *Infomap*.

Tabela 5 Evaluacija klaster metoda na mrežama osnovnog i proširenog uzorka

Klaster metod	Boj čvorova koji nisu grupisani u stranku kojoj pripadaju	
	Mreža osnovnog uzorka (N=40)	Mreža proširenog uzorka (N=355)
<i>Edge Betweenness</i>	10	8
<i>Label</i>	14	34
<i>Propagation</i>	11	8
<i>Spin-Glass</i>	14	12
<i>Walktrap</i>	7	7
<i>Infomap</i>		

Pošto se u Tabeli 5 *Infomap* algoritam pokazao kao metod grupisanja koji najbliže prati stranačku pripadnost, upravo taj algoritam je izabran za grupisanje čvorova u ovom radu (Listing 23).

```
klast.pol <- cluster_infomap(mreza.pol)
klast.pol.pu <- cluster_infomap(mreza.pol.pu)
```

Listing 23 Primena *Infomap* klaster algoritma na mrežu osnovnog i proširenog uzorka

Rezultati grupisanja ovom metodom za mrežu osnovnog uzorka dati su u Tabeli 6, a za mrežu proširenog uzorka u Tabeli 7.

Tabela 6 Čvorovi mreže osnovnog uzorka grupisani *Infomap* algoritmom. U tabeli su označeni (boldovani) čvorovi koji su grupisani van svoje izborne liste.

Klaster (Stranka)	Čvorovi
DS	S. Mićanović , B. Pajtić, G. Čomić, DS, D. Šutanovac
LDP – SDS	Č. Jovanović, D. Kovačević, K. Hajdarević, LDP, G. Radosavljević, SDS, K. Samofalov
SNS	A. Vučić, Zoran Babić, SNS
DJB	S. Radulović, D. Petrović, A. Stevanović, I. Ninić, DJB
DSS – Dveri	V. Đukanović , D. Maršićanin, M. Bojović, DSS, B. Obradović, DVERI
SPS	B. Ružić, N. Borovčanin, D. Vukomanić, SPS
NS	Z. Živković, V. Pavićević, NS
SRS	V. Šešelj, V. Radeta, N. Šarović, Z. Krasić, SRS
SNS	N. Stefanović

Tabela 7 Čvorovi mreže proširenog uzorka grupisani *Infomap* algoritmom. Označeni (boldovani) su čvorovi koji se nalaze u osnovnom uzorku.

Klaster (Stranka)	Čvorovi
1. DS – SNS	NesaStefanovic, Djuka_RSFC, micanovic_s, PajticBojan, gordanacom, AleksJerkov, SutanovacDragan, demokrate, dusanmasic, Vukasin90, Aledar989, BiljanaLuki, jelenak021, stanisasms, DaniloKoprivica, PeckoPivo, nebojsarako, dedadjinaj, Budza_Sremacka, izivkov, BogdanTatic, milorad984, Djukela_SNS, micin_taja, momcilocrvak, amir_bisljimi, cipolla12, SashaPavlovic, BorisGolubovic, Stefanivanovic2, wujisicm, holzwege11, bborovic3, UdbiEtOrbi, biljana0001, dolazi_cirkus, ljubicapuric, plavozuto, draganns79, jevticdalibor, kejsiberger, oknarbnb, bbugarski, comic_marko, DraganPopovic11, ElGuapoGigaMega, HippoNada, nszoranans, BaneKuzmanovic, BojanaMaljevic, CKokanovic, Gunjo1996, nijekodeks, Rombolad, RooneySeBuni, StojcevicNikola, Tremi028, BrankaVookovic, SaobrKamera, djoca, lazicradosav, Gaia_Foks, intersrbija10, IvoJovanovic, mancekg, PekaLatinovic, radomir_martin, aleksandarmorga, barski6, Bivsi_Direktor, DelBoske021, sashagracin, StojanovR, ZarkoJokanovic, Hamiltonman, hrabri_markuze, mara_jovicic, ZiziDragan, antioksidantAO, GIZUJP, UkiPajp, zeljkoividakovic, BJCalamity, bojanstojanovi4, cheetah75, DenisKolundzija, Gordana_Burazer, PobesneliZika99, PsetoFromGhetto, pulsar1, Dragan38380610, DraganBozic14, FraBrne, igorigor1977, mad_you_team, marymerlak, milijaktitor, RadojevGoran, RFilipovic, scepticalbg, vestoma_Tomy, ZeljkoPa, Architecta, BlagicaKostic, Braduljica, freehandmeister, gordix401, igorkostickg, ivancosic, izeljko77, jelenamit1960, jucedanas, karlosboki, krajan51, Marija28, milen_kakuca, odlicno01, ODragicevic, otsoa373, petko962gmailco, RadomirKukobat, SrdjanNovakovic, n, tisov_sumarak, Uros_Vozdovac, AcoMDR, aconibg, aleksrado, Bilja_Pantic11, DejoGenije, DottieRParker, DozaNormalnosti, etiketirani, friezeoflife, gradjaninpokorn, JoanaLegend, lukazb, lyooba, MaryPiglet, md_nest, mirsto100, mumerah, necha64, pblagojevic, rumirolava, shpira7, simon_w_helper, srdjandrigo, suslov_tatjana, svetislava80, TooLateBloomer, V Obradovic, zarkobns
2. DJB	SasaRadulovich, d_pavlovic, Stevanovich_A, INinic, DostaJeBilo, antarizana, nedeljnik, MarkoDimic2, zivstapa, LicniStav, Lazariques, masa_meda76, leteciGasa, tkovic, dusanlj, DraganVujevic, sneza_kg, vbjelak, Antonije_S, bojan_007, VpopNS, mikicmiki77, MrDjibrilo, vladovojanovi, Mittwoch22, zivdus, Djekara1, galjakv, godza, msimic92, Slobelix_, DukeOfKarijes, PolicMusic, stffury, BrankaStamenkov, cheja88, garginis, nbozic, sigurnost_ng, zoranivanovic01, DeanovicV, Gacone85, I_Cosic, NickVanExelName, Nunciije, QuartusVir, yooyo3d, IvanMiljkovic60, jankovicmilos, mihizovac, Milica2704, ummagumma9449, BobanaMacanovic, BorkoRadojkovic, bucalapucala, Dragana_Nedin, edislman, mirkojankov, OgnjanStanisha, Ristatraktorist, SilaPrirode, toxonics, tvitdjole
3. SNS	avucic, ZBabic_, sns_srbija, Nietzscheanac, Askorbinska_kis
4.	Radosav77, trzisnoresenje, TomaMomirovic, somboracmarko, severnojuzni, VladimirTodoric, dukimadjar, a_ono_medjutim, pavlemihajlovic, tatanadete, timocanin, miloskg, ilegalista, VeraMarjanovic, nikolatanasic, paraliza_mozga, ivankovic_milos, zolezorule, Voja Stevanovic, erosonogsveta, Patuljcica Jo, vojkanglavinic
5. LDP – SDS	Ceda_Jovanovic, KenanLDP, LDP, KoleSamofalov, socdemstranka
6.	RomicSandra, mrburns0102, DavidDBabic, bonnie_wynand, blagojepantelic, bulevar38, djuracvorovic, lidermg1, bogdan313, PavloviMilo, Milos G
7.	Debela_zvezda, AlexaTodorovic, good_neighbor, harberger_, parun_, Ikan999, ojinjic, trikosh, beli_chovek, vanjagator, Dexonomics, Rookie2005, DejanBursac, ninoslav
8. NS	ZoranDirektno, vpavicevic, NovaStranka, ArisMovsesijan, _Trillian_
9. SPS	brankoruzicsps, nborovcanin, socijalisti, vukomand, _vuk, redinmind, VelikiMo
10. DSS – Dveri	mbojovic777, dsscentar, BoskoObradovic, SPDveri, malagurski, ostojasimetic, dusanjecar, majstorovic56, Riddick82, demossersbos, DrTrifkovic
11.	RevoltinCook, BojanDjuric78, IvaCvrchak317, AlexJovanovich, Solisio, cika_crni_
12.	marsicanin_d, Interceptor_MR, BanTomN, DjuraKrompic, mapkoKraljevic, MoravacPZ, kamperelic, istorijska, gjCGQPMsX1a4RBU, svezarock, bela_sova, el_pinturicchio, vesna_zvkvc
13. SRS	seselj_vojislav, VjericaR, ZoranKrasic, srpski_radikali, PuricLjubica, seksanoreksik,

	CosaNostra1991, MiciD77, nidje_vezeee, ivodragicevic53
14.	nenadcicko, banegrkovic, grujša, KatarinaTadic_, npesic82, Gojgic, teskicentar
15.	AdvVlahovic, Goran11000, Snezana66855918, B_atV
16.	DragoKovacevic , zejo, GISadzija, aleksandarbojic, bubahana, Djora007, dusan_rg11, pobednick
17.	NemanjaSarovic , IFAPy1R5AgaVzYN, SrbskaBorba, PrDrMr, KlopkaOlja
18.	neriljano, njodac, MikaSolid, sretenzivkovic2
19.	Vladan_KV, TeleprompterRS

3.6.2 Vizualizacija klastera

Postoji više načina za vizualizaciju grupisanih čvorova. Čvorovi u mreži koji pripadaju istoj grupi mogu deliti vizualne elemente koji se koriste za vizualizaciju mreža (npr. boja čvora, boja oznake, font oznake itd.), ali je takođe moguće vizuelno predstaviti klaster kao površinu oko čvorova koji mu pripadaju, a te površine mogu biti različitih boja. Još jedan način da se grupisani čvorovi vizuelno diferenciraju je promenom rasporeda čvorova. U Listingu 24 dat je primer koda za generisanje rasporeda na osnovu klaster analize, koji se može koristiti za vizualizaciju.

```
layout.modular <- function(graf, klaster) {
  graf$layout <- layout.auto(graf) #inicijalizacija rasporeda
  nm <- length(levels(as.factor(klaster$membership))) #broja klastera
  gr <- 2

  while (gr ^ 2 < nm) { #u koliko redova će klasteri biti prikazani
    gr <- gr + 1
  }
  i <- j <- 0
  for (cc in levels(as.factor(klaster$membership))) { #za svaki klaster
    F <- delete.vertices(graf, klaster$membership != cc) #samo čvorovi
    klast.
    F$layout <- layout.kamada.kawai(F)
    F$layout <- layout.norm(F$layout, i, i + 0.5, j, j + 0.5) #skaliran
    raspored unutar klastera
    graf$layout[klaster$membership == cc, ] <- F$layout
    if (i == gr) { #brojači za položaj klastera u rasporedu
      i <- 0
      if (j == gr) {
        j <- 0
      } else{
        j <- j + 1
      }
    } else{
      i <- i + 1
    }
  }
}
```

```

    return(graf$layout)
}

```

Listing 24 Primer koda za generisanje rasporeda na osnovu klastera

Izvor: (Türei, 2013)

Tako generisani raspored se može primeniti za vizualizaciju prikazanu ranije u Listingu 14, a dodavanjem opcije za markiranje grupa daje obojenu površinu za svaki klaster.

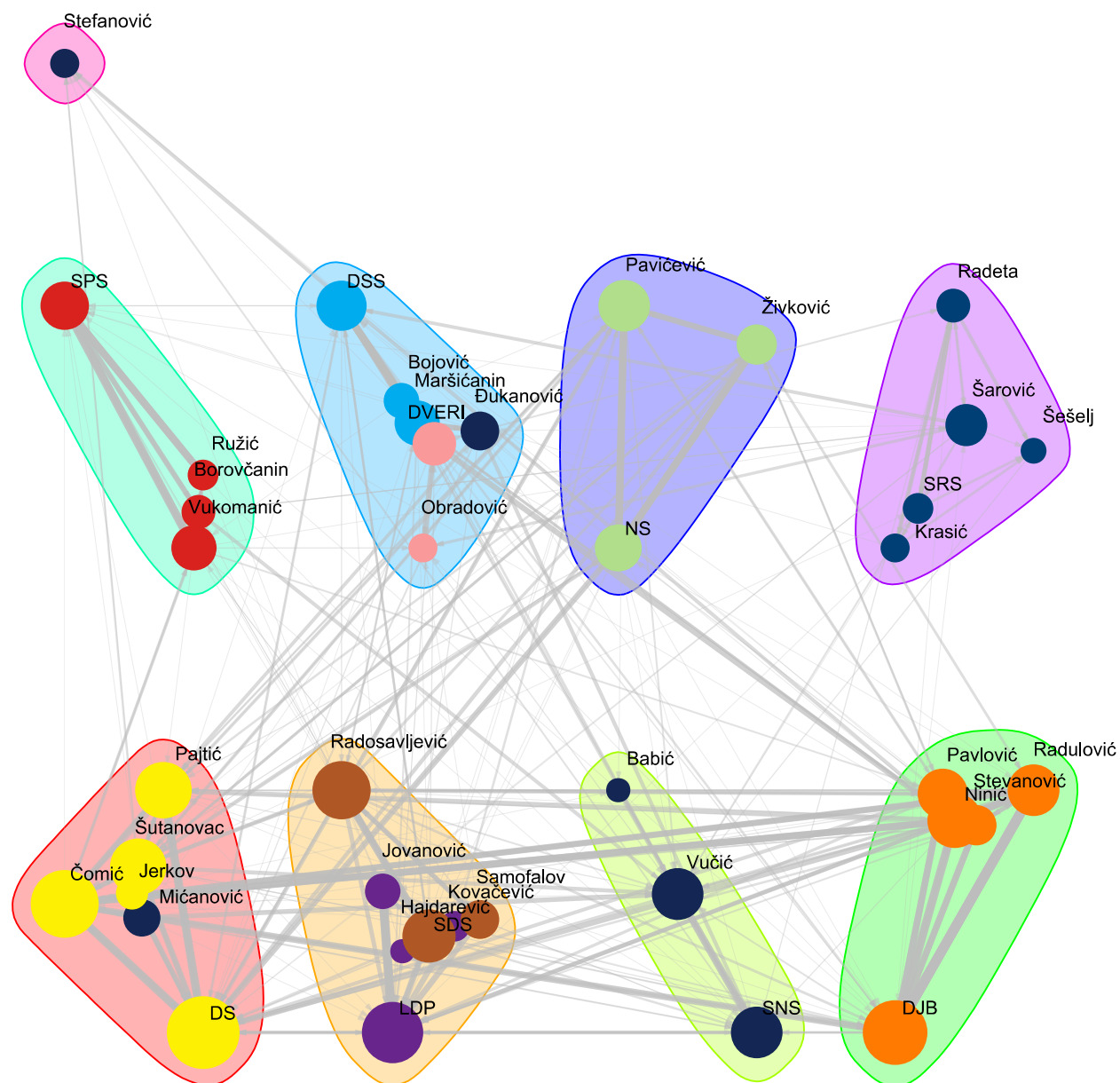
```

graspored <- layout_modular(mreza.pol, klast.pol) #generisanje rasporeda

plot.igraph(mreza.pol,
             layout = graspored,
             edge.width = log(E(mreza.pol)$weight),
             edge.curved = FALSE,
             edge.color = alpha("gray", rescale(x=log(E(mreza.pol)$weight),
to=c(0.3,1))),
             edge.arrow.size = 0.1,
             edge.lty = 1,
             vertex.color = V(mreza.pol)$boja,
             vertex.label.color = "black",
             vertex.label.family = "Arial",
             vertex.frame.color = NA,
             vertex.label = političari$Prezime,
             vertex.size = rescale(degree(mreza.pol), to=c(5,15)),
             vertex.label.cex = 0.7,
             vertex.label.dist = 0.5,
             mark.groups = klast.pol #označavanje grupa bojom
)

```

Listing 25 Vizualizacija mreže osnovnog uzorka sa klasterima (Slika 10)



Slika 10 Vizualizacija mreže osnovnog uzorka sa izdvojenim klasterima

4 DISKUSIJA

4.1 Uzorak

Snowball metoda za formiranje drugog, proširenog uzorka pokazala se prilično dobrom za obuhvatanje važnih političkih aktera na Tviteru tokom izbora. Primenom te metode u uzorak su uključeni mnogi važni akteri i to ne samo članovi političkih stranka, koji nisu bili deo osnovnog uzorka. Uzorak takođe uključuje i bivše političare, koji nisu učestvovali na izborima 2016. godine, ali i novinare, književnike, glumce i druge ljude sa javne scene.

Nedostatak uzorka je u tome što u njemu nedostaju nalozi koji su od vremena raspisivanja izbora do vremena sakupljanja podataka (oko 4 meseca), titovali više od 3200 tvitova. Kao što je navedeno ranije, postoji ograničenje Tviter API-a prema kome je moguće preuzeti samo 3200 najnovijih tvitova. Zbog toga postoje čvorovi, koji bi potencijalno bili važni za strukturu razmatranih mreža, ali nisu bili deo ni osnovnog ni proširenog uzorka. Ovi čvorovi bi mogli imati intenzivnu interakciju sa drugim čvorovima razmatrane mreže, međutim, njihov uticaj, zbog nemogućnosti pristupa podacima, u ovom radu nije mogao biti analiziran. Njih je bilo moguće predstaviti na grafu tako što bi bile uzete u obzir samo dolazne ivice (kada su ih drugi čvorovi pominjali), ali postoji mogućnost da bi takav pristup iskrivio rezultate metrika, i samim tim doneo više problema nego što je to slučaj kod jednostavnog odstranjivanja naloga iz uzorka. Na sreću, takvih naloga nema mnogo, a među značajnim su Goran Ješić iz Demokratske stranke i Nebojša Krstić nekadašnji savetnik predsednika Republike.

4.2 Rezultati analize

4.2.1 Karakteristike analiziranih mreža

Mreža osnovnog uzorka političara je mreža male gustine (0,21). Ova mala gustina, svojstvena online društvenim mrežama, bi bila još manja ako bismo uzeli u obzir samo veze između čvorova kod kojih je postojala međusobna komunikacija (0,10, što je gustina karakteristična za neusmerene online društvene mreže, npr. Fejsbuk). Mreža proširenog uzorka je skoro dvostruko 'razređenija', (0,11 za usmerenu varijantu, odnosno 0,04 za neusmerenu varijantu). Razlika u gustini između dve analizirane mreže leži u selekciji čvorova koji ih sačinjavaju. Mreža osnovnog uzorka sadrži čvorove koji u političkoj borbi moraju da sukobljavaju mišljenja sa drugim političarima. U drugoj mreži (proširenog uzorka) čvor koji interaguje sa političarem sa jednog

dela političkog spektra i drugi čvor koji interaguje sa političarem sa drugog kraja političkog spektara, možda nikada neće imati priliku da međusobno interaguju.

Obe analizirane mreže su nepovezane u svom usmerenom obliku, zbog pojave da neki čvorovi nemaju komunikaciju u oba smera, što znači da postoje čvorovi koji nikada nisu pomenuli nijedan drugi čvor u mreži, pa je iz tog razloga usmereni graf nepovezan. Kada se smer komunikacije zanemari, obe mreže su povezane. Broj čvorova koje treba ukloniti da bi mreža osnovnog uzorka postala nepovezana je 3, a u slučaju mreže proširenog uzorka, koja ima mnogo manju gustinu, dovoljno je ukloniti jedan čvor da bi mreža postala nepovezana.

Dužina najdužeg geodezika u mreži osnovnog uzorka, tj. njen dijametar je 6 (bez uzimanja u obzir težina ivica). Najduži geodezik je putanja između članova iste stranke: Vojislava Šešelja i Nemanje Šarovića iz Srpske radikalne stranke, što je posledica prirode interakcije između razmatranih aktera (putanja između ta dva čvora bi bila mnogo kraća da se radi o neusmerenoj mreži). Prosečna dužina putanje u mreži osnovnog uzorka je 2,28. Kod mreže proširenog uzorka dijametar je 5, a prosečna dužina putanje 2,09. Značajna razlika između dužine dijametra i prosečne dužine putanje kod obe mreže je posledica velikog broja najkraćih putanja dužine 2, pa duži geodezici malu utiču na prosečnu dužinu putanje. Samo na osnovu činjenice da je dijametar mreže proširenog uzorka manji od dijametra mreže osnovnog uzorka, može da se zaključi da, ipak, postoji komunikacija između čvorova dodatih u prošireni uzorak, jer da ne postoji dijametar mreže bi bio veći ili jednak onome iz osnovne mreže, a ne manji.

Tranzitivnost u mreži proširenog uzorka je niža nego u mreži osnovnog uzorka. Ovakav rezultat ne iznenađuje s obzirom da je mreža osnovnog uzorka, mreža ljudi koji se poznaju i u stvarnom životu, a ne samo preko online interakcija. Treba se uzeti u obzir i prisustvo stranačke homofilije, koja se potencijalno manifestuje i u mreži proširenog uzorka, ali u mnogo manjoj meri jer je procenat stranačkih saboraca mali u odnosu na ukupan broj čvorova, za razliku od mreže osnovnog uzorka.

Niska asortativnost stepena demonstrirana, posebno u proširenom uzorku ($-0,166$), govori o demokratičnosti Tvitera kao online društvene mreže. Naime, ne postoji tendencija da popularni čvorovi (sa visokim stepenom), komuniciraju mahom međusobno, već komuniciraju i sa čvorovima nižeg stepena. Ovo govori u prilog toga da je Twitter dobro mesto za razmenu mišljenja, ali možda reflektuje i svesni napor političara da u toku kampanje što više komuniciraju sa 'plebsom'. Postoji mogućnost da je asortativnost stepena drugačija van kampanje.

Metrika asortativnosti atributa (nominalna asortativnost) prema stranačkoj pripadnosti u mreži osnovnog uzorka ima vrednost 0,222 što u drugim kontekstima može izgledati nisko, ali u kontekstu nominalne asortativnosti realnih društvenih mreža, taj stepen korelacije je zapravo prilično visok. Isti zaključak na osnovu stranačkog atributa nije moguće izvesti kada je u pitanju mreža proširenog uzorka, jer podaci o stranačkoj pripadnosti 315 od 355 čvorova nedostaju, tj. poznati su samo za 40 čvorova iz osnovnog uzorka. Ipak, taj skup od 355 čvorova je moguće podeliti u 3 grupe: opozicija i vlast (za 40 čvorova za koje je poznata stranačka pripadnost) i nepoznati (za čvorove dodate u proširenom uzorku). Tendencija grupisanja prema pripadnosti tako definisanim velikim grupama nije prisutna ($r=-0,039$). To znači da u mreži proširenog uzorka ne postoji sklonost da se čvorovi drže unutar tako široko definisanih grupa. Sa jedne strane to je i normalno, jer je cilj većine političara (bez obzira da li se radi o vlasti ili opoziciji) da što više interaguju sa uticajnim tviterašima koji nisu deo njihove stranke, kako bi dopreli do što više glasača. Sa druge strane uzorak je pristrastan (*biased*), jer je mreža proširenog uzorka formirana tako što je selekcionirani kriterijum bio komunikacija sa političarima, a ne slučajnim uzorkom iz skupa domaćih tviteraša.

4.2.2 Centralnost čvorova i centralizacija u mrežama

Rezultati koji se odnose na centralnost čvorova i distribuciju centralnosti u posmatranim mrežama, kroz metriku centralizacije, dati su u poglavlju 5 u Tabelama 4 i 3.

Centralizacija intermedijarnosti je veoma niska kod obe posmatrane mreže (0,13 i 0,08), što ukazuje na to da ne postoji mala grupa čvorova od koje zavisi povezanost ostalih čvorova grafa, tj. nema tzv. čuvara kapija (*gatekeepers*) u komunikaciji na Tviteru, što dodatno govori u prilog demokracije ove društvene mreže. Čvorovi sa najvišim vrednostima intermedijarne centralnosti u mreži osnovnog uzorka su Bojan Pajtić, predsednik Demokratske stranke za vreme kampanje 2016, i Vladimir Pavićević iz Nove stranke. U istoj mreži postoje čak četiri čvora koji imaju vrednost 0 za intermedijarnu centralnost: Vojislav Šešelj (SRS), Zoran Babić (SNS), Nebojša Stefanović (SNS) i Branko Ružić (SPS). To su oni čvorovi koji bi mogli da se uklone bez ugrožavanja povezanosti mreže. Kod mreže proširenog uzorka najveću intermedijarnu centralnost imaju Saša Radulović, predsednik pokreta „Dosta je bilo“ i Gordana Čomić iz Demokratske stranke.

Centralizacija stepena je slabo prisutna kod mreže osnovnog (0,24), a više u mreži proširenog uzorka (0,37), zbog prisustva čvorova sa ogromnim stepenom. Centralizacija bi imala vrednost 0 kada bi svi čvorovi imali isti stepen, tj. tada bi se moglo reći da je mreža u potpunosti

decentralizovana. Najviši stepen u mreži osnovnog uzorka ima nalog Demokratske stranke (34 tj. povezan je sa 87,3% drugih čvorova). Čvor sa najvećim stepenom u mreži proširenog uzorka je Gordana Čomić, koja je povezana sa 348 (98,3%) drugih čvorova. Nju slede Saša Radulović 325 (91,8%) i Dušan Pavlović (DJB) 289 (81,6%). U mreži osnovnog uzorka čvor koji je najčešće pominjan je nalog Demokratske stranke, a čvor koji je najviše pominjao druge je Goran Radosavljević iz Socijaldemokratske stranke. Kombinacija ova dva pokazatelja ukazuje na zanimljive čvorove. Čvorovi koji imaju visoke vrednosti odlaznog stepena, a niske vrednosti dolaznog stepena, kao što je čvor Slaviša Mićanović, su oni koji mnogo pominju druge, a njih ne pominje skoro niko. Obrnuta kombinacija, oni koje drugi mnogo pominju, a sami ne pominju mnogo drugih čvorova, otkriva naloge koji nisu lični, već koje vode timovi zaduženi za kampanju na Tviteru, takvi čvorovi su Aleksandar Vučić iz Srpske napredne stranke, Čedomir Jovanović iz Liberalno demokratske Partije, stranački nalozi (LDP, DS, SNS), ali i Bojan Pajtić iz DS-a i Nebojša Stefanović iz SNS-a. Postoji mogućnost da je ovakvo ponašanje deo strategije.

Centralizacija bliskosti je umereno zastupljena kod obe mreže (0,33 i 0,43). Čvor sa najvećom centralnošću bliskosti u mreži osnovnog uzorka je nalog Socijalističke partije Srbije (0,0142), a u mreži proširenog uzorka Saša Radulović (0,0015).

Centralizacija svojstvenog vektora je veoma visoka kod mreže osnovnog (0,73) i kod mreže proširenog uzorka (0,83). To znači da mali broj čvorova ima izuzetno visoke vrednosti centralnosti svojstvenog vektora, dok ostatak čvorova ima niske vrednosti ove centralnosti. Razlika između centralizacije stepena i centralizacije svojstvenog vektora ukazuje na to da je mreža centralizovanja ako se uzmu u obzir indirektne veze između čvorova, nego direktne veze (Tranos, 2013). Prva tri čvora prema centralnosti svojstvenog vektora u mreži osnovnog i proširenog uzorka su čvorovi iz pokreta „Dosta je bilo“: nalog pokreta (osnovni:1; prošireni: 1), Saša Radulović (0,98; 0,93) i Ivan Ninić (0,44; 0,43). Jedno od objašnjenja ovako visokih vrednosti centralnosti svojstvenog vektora je međusobna komunikacija ova tri navedena čvora. Pošto komunikacija ne samo da postoji, nego je i veoma intenzivna (uzimaju se obzir težine ivica), čvorovi koji su dosta uticajni u mreži, učestalom međusobnom komunikacijom podižu vrednosti centralnosti svojstvenog vektora. Kada se centralizacija svojstvenog vektora izračuna bez uzimanja u obzir težine ivica, vrednosti su nešto niže, konkretno, 0,58 za mrežu osnovnog i 0,73 za mrežu proširenog uzorka. U tom slučaju najveću centralnost svojstvenog vektora ima čvor Gordana Čomić, dok su čvorovi iz Dosta je bilo i dalje jako visoko rangirani. Činjenica da su obe mreže centralizovanije kada se posmatraju indirektne ivice nego direktne ivice, stoji, bez obzira na težine ivica. Inače, prema Butts-u (2008b, p. 83) centralnost svojstvenog vektora je

mera centralnosti koja najbolje sumira celokupnu strukturu mreže, jer je usko povezana sa čvorišnim konceptom centralnosti, globalnim odlikama mreže kao što je struktura centara i periferije, strukturnim sumiranjem i smanjenjem broja dimenzija, ali i društvenim procesima kao što su difuzija i uticaj.

4.2.3 Međuzavisnost pokazatelja i kompozitni pokazatelj centralnosti

U poglavlju 5 data je matrica korelacije ranga (Slika 7) različitih pokazatelja popularnosti i centralnosti čvorova u mreži osnovnog uzorka. Matrica ukazuje da postoji jaka korelacija ranga između pokazatelja centralnosti dolaznog stepena i dolazne bliskosti i između odlaznog stepena i odlazne bliskosti, što je karakteristično za stvarne društvene mreže (Li, Li, Van Mieghem, Stanley, & Wang, 2015; Valente, Coronges, Lakon, & Costenbader, 2008). Korelaciona matrica pokazuje da drugi pokazatelji centralnosti čvorova nisu međusobno visoko korelisani, što se može i očekivati s obzirom da se svi odnose na različite stvari, tj. svaki pokazatelj otkriva zasebnu karakteristiku čvora u mreži. Zbog toga, jednostavnim razmatranjem vrednosti ovih pokazatelja ne može se dati odgovor na pitanje koji je to čvor najcentralniji u mreži. Umesto toga, oni daju odgovor na pitanje o specifičnoj centralnosti i od velike su koristi ako nas zanimaju specifični aspekti centralnosti.

Za ocenu ukupne centralnosti napravljen je kompozitni pokazatelj pomoću P_2 odstojanja. Na taj način formiran je jedan pokazatelj koji objedinjuje varijabilitet svih pokazatelja centralnosti. Kompozitni pokazatelj najjače je korelisani sa centralnostima intermedijarnosti i svojstvenog vektora, ali i sa obe centralnosti stepena, dok je korelacija ranga sa centralnostima bliskosti niža. Ne postoji jaka korelacija kompozitnog pokazatelja centralnosti sa testiranim pokazateljima popularnosti na Tviteru (br. pratioca, prosečan br. tvitova itd.), što ukazuje na to da ovakav pokazatelj ukupne centralnosti, pruža drugačiju perspektivu gledanja na uticaj pojedinih aktera na jednoj *online* društvenoj mreži.

Jedan od problema sa ovakvim pristupom primene kompozitnog pokazatelja je izračunavanja istog. Naime, različite mreže će imati različitu strukturu, pa samim tim i potencijalno različite međuzavisnosti između različitih mera centralnosti, što dovodi do toga da će za svaku mrežu kompozitni pokazatelj biti izračunat na različit način, tj. sa drugačijim redosledom varijabli koje ga formiraju. Zbog toga nije moguće poređenje vrednosti pokazatelja, izračunatih P_2 odstojanjem, između mreža različitih topologija. Vrednosti P_2 odstojanja same za sebe nemaju nikakvo značenje, odnosno ne mogu se interpretirati u kontekstu karakteristika grafa za koji su

izračunati. Sve to ograničava razmatrani kompozitni pokazatelj na referentni sistem konkretne mreže za koju je izračunat.

U ovom radu svrha stvaranja kompozitnog pokazatelja bila je rangiranje čvorova i stranka kojima oni pripadaju prema ukupnoj centralnosti. Prema kompozitnom pokazatelju najcentralniji čvorovi u mreži su nalog pokreta „Dosta je bilo“, Gorana Čomić (DS) i Bojan Pajtić (DS). Najperiferniji nalozi prema kompozitnom pokazatelju su Vojislav Šešelj (SRS), Zoran Babić i Nebojša Stefanović (SNS).

Kada se izračuna prosečna vrednost kompozitnog pokazatelja za svaku stranku (Slika 8), pokret „Dosta je bilo“ je na prvom mestu. Na drugom mestu, skoro izjednačena sa pokretom „Dosta je bilo“, je Demokratska stranka, a za njom je Nova stranka. Poslednja tri mesta zauzimaju Srpska radikalna stranka, Srpska napredna stranka i Liberalno demokratska partija. Ovakvi rezultati su oprečni rezultatima koje daje konvencionalna analiza popularnosti na Tviteru. Naime, prema broju pratioca i broju tvitova koji su retvitovani i označeni kao omiljeni, Srpska napredna stranka bi mogla da se proglasi najpopularnijom. Naloge ove stranke karakteriše mali broj tvitova, a veliki broj retvitova. Ova pojava nije iznenađujuća kod naloga Aleksandra Vučića, predsednika vlade, čije retke tvitove mogu retvitovati mnoge novinske agencije (domaće i strane), kao i brojni domaći i strani pratioci. Neobični su veliki brojevi retvitova kod drugih naloga SNS-a, npr. kod naloga Nebojše Stefanovića, stranačkog naloga, ali i najčudnije kod Slaviše Mićanovića Mićana, (poznatim na društvenim mrežama kao vođa „SNS botova“). Slaviša Mićanović primera radi, ima oko 10 puta više retvitova po tvitu od Saše Radulovića iz DJB iako ne zauzima važnu državnu funkciju i van društvenih mreža nije poznat široj javnosti. Zbog navedenog postoji osnovana sumnja da se radi o nekom obliku organizovanog retvitovanja. Potvrda ove teze bi mogla da se ostvari bližom inspekcijom tvitova u pitanju. Naime, ako stalno jedni isti korisnici retvituju tvitove SNS, to bi potvrdilo navedenu tezu. Može se postaviti slučaj da su i tako dobijeni retvitovi važni, jer kroz mrežu propagiraju informacije, ali to je slučaj samo ako su ti korisnici koji organizovano retvituju dobro povezani (tj. ako imaju mnogo pravih pratilaca, a ne samo druge članove svoje stranačke zajednice). Međutim, takva analiza je van okvira ovog rada, a mogla bi da bude interesantna za buduća istraživanja.

Još jedan zadatak potencijalnih budućih istraživanja može da bude ponovljena analiza za period između izbora 2016 i predsedničkih (a potencijalno i opštih) izbora 2017. Poređenjem karakteristika mreže za vreme izborne kampanje i perioda između izbornih kampanja otkrio bi se potencijalni uticaj političke mobilizacije za vreme kampanje.

4.2.4 Homofilija i klaster analiza

Kao što je već konstatovano, homofilija je izražena u mreži osnovnog uzorka. Naime, čvorovi više komuniciraju sa drugim čvorovima koji pripadaju istoj stranci ($r=0,222$). Sproveden test nominalne asortativnosti, koji je dao navedeni rezultat (Tabela 3), nije uzeo u obzir šire grupisanje kao što je pripadnost izbornoj listi. Podgrafovi političkih stranaka (Slika 9), pokazuju da čvorovi koji pripadaju istim izbornim listama međusobno intenzivno komuniciraju. Tako se na podgrafu DS, nalaze čvorovi NS, i obrnuto; na podgrafu DSS, čvorovi DVERI i obrnuto; na podgrafu LDP čvorovi SDS i obrnuto. Kada se koeficijent nominalne asortativnosti izračuna za liste, umesto za stranke, dobija se značajno viša vrednost ($r=0,281$). Kada se stranke gledaju zasebno, homofilija je najprisutnija kod Srpske radikalne stranke i Socijalističke partije Srbije, i pokreta „Dosta je bilo“, a najmanje kod Dveri, Nove stranke i Srpske napredne stranke.

Prema tranzitivnosti (sklonosti formiranja zatvorenih trijada), izdvaja se pokret „Dosta je bilo“, dok je kod Dveri, Liberalno demokratske partije i Srpske napredne stranke tranzitivnost 0, tj. nemaju nijednu zatvorenu trijadu. Treba uzeti u obzir činjenicu da je na podgrafovima mapirana samo obostrana komunikacija. Podgraf Srpske napredne stranke je nepovezan i uopšte nema gigantsku komponentu, najveća komponenta je trijada, koja nije čak ni zatvorena. Ovakva nekoherentnost podgrafa Srpske napredne stranke uticala je i na klaster analizu.

Iz rezultata nominalne asortativnosti za mrežu osnovnog uzorka može se izvući zaključak da postoji tendencija stvaranja zajednica koje čine čvorovi istih političkih stranaka. Takva pretpostavka je testirana u poglavlju 3.5 u kojem je sprovedeno klasterovanje korišćenjem *infomap* algoritma slučajnih šetnji. U mreži osnovnog uzorka klaster analiza je izdvojila 11 klastera. Ako se svakom klasteru dodeli stranačka pripadnost, na osnovu toga kojoj stranci pripada većina čvorova, onda se može konstatovati da je klaster analiza pogrešno kategorisala 7 (17,5%) čvorova. Međutim, kao što je na osnovu analize homofilije u podgrafovima stranaka zaključeno, još jača homofilija zastupljena je kada se pripadnost čvorova posmatra kroz izborne liste. Ako na taj način tumačimo rezultate klasterovanja, onda se može konstatovati da je algoritam pogrešno svrstao samo 3 (7,5%) čvora, označena boldiranim tekstom u Tabeli 6. Zapravo ta tri 'pogrešno' klasifikovana čvora su čvorovi Srpske napredne stranke: Slaviša Mićanović (smešten u klaster Demokratske stranke), Vladimir Đukanović (smešten u klaster Demokratske stranke Srbije i Dveri) i Nebojša Stefanović (kao jedini čvor zasebnog klastera).

Činjenica da je algoritam korišćen za klasterovanje sa velikim nivoom tačnosti klasifikovao čvorove prema političkoj orijentaciji i stranačkoj pripadnosti, potvrđuje da se čvorovi u mreži

osnovnog uzorka zaista grupišu prema stranačkoj pripadnosti. Postavlja se pitanje da li se ista logika može primeniti i na prošireni uzorak. Važno je imati na umu razliku između osnovnog i proširenog uzorka. U osnovnom uzorku su svi čvorovi političari, i to ne bilo koji političari, već najprominentniji političari svojih stranaka koji imaju nalog na Tviteru. U proširenom uzorku se pored čvorova iz osnovnog uzorka, nalaze i ostali članovi političkih stranaka koji aktivno učestvuju u izborima, brojni novinari koji najčešće nisu javno deklarirani simpatizeri neke političke stranke i drugi korisnici Tviter mreže. Iako, su politički čvorovi u centru, u smislu da je selekcija naloga proširenog uzorka vršena na osnovu komunikacije sa njima, međusobna komunikacija ostalih čvorova iz proširenog uzorka ne mora nužno da bude manje zastupljena od njihove komunikacije sa čvorovima iz osnovnog uzorka. To znači da je moguće da u mreži proširenog uzorka postoje zajednice koje ne sadrže nijedan čvor političara iz osnovnog uzorka. Upravo je to i bio slučaj kada se *infomap* algoritam primenio na mrežu proširenog uzorka. Rezultat je 19 klastera, što je mnogo više nego broj stranaka (Tabela 7).

Prvi klaster broji 157 čvorova (43,7%). Fragmentirana struktura podgrafa Srpske napredne stranke, koja je rezultirala 'pogrešnom' klasifikacijom njihovih čvorova u mreži osnovnog uzorka, rezultirala je time da u proširenom uzorku polovina čvorova ove stranke bude grupisana zajedno sa čvorovima Demokratske stranke. Ti čvorovi (Nebojša Stefanović, Vladimir Đukanović i Slaviša Mićanović) komuniciraju više sa čvorovima iz suprotstavljenog političkog tabora i nego sa svojim političkim saborcima. Podgraf Srpske napredne stranke ukazuje na to da je komunikacija navedenih čvorova sa čvorovima iz Demokratske stranke retko bila obostrana. Stoga, verovatno se radi o 'vangardi' koja proziva političke protivnike. Ovakva pojava je primer problematičnosti bilo kog pokušaja klasifikacije i dodeljivanje etikete čvorovima za koje nije poznata stranačka afilijacija. Da li se radi o političkom protivniku koji vatreno raspravlja sa svojim neistomišljenicima, a ne ulazi u rasprave sa svojim istomišljenicima, ili se radi o podržavaocu političkih ideja za koje se zalaže stranka oko koje je klaster otkriven?

Klaster oko pokreta „Dosta je bilo“ broji 63 (17,7%) čvorova od čega je 5 čvorova iz osnovnog uzorka i to su čvorovi istog pokreta. Ostali klasteri sadrže značajno manje čvorova. Klasteri Nove stranke i Socijalističke partije Srbije sadrže samo po dva dodatna čvora koja nisu pripadala osnovnom uzorku. *Infomap* algoritam je otkrio 9 zajednica koje nemaju nijedan čvor iz osnovnog uzorka i još 3 zajednice koje sadrže više čvorova iz proširenog uzorka za koje nije poznata stranačka pripadnost i po jedan zalutali čvor iz osnovnog uzorka. Drugim rečima u mreži proširenog uzorka *infomap* algoritam je izdvojio 8 stranačkih klastera (tj. klastere izbornih lista) kao i kada je primenjen na mrežu osnovnog uzorka.

Infomap algoritam je primenjen na mreže i pronašao je u njima zajednice. Taj algoritam nije uzimao u obzir stranačku pripadnost, ili bilo koji drugi atribut čvorova mreže, već je klasterovanje vršeno samo na osnovu mrežne strukture. Alternativa tome je primena algoritma koji omogućava unapred definisanje pripadnosti klasterima za neke od čvorova. Algoritam koji bi mogao da se upotrebi u tu svrhu je algoritam širenja oznaka (*Label propagation*). Međutim takav pristup se pokazao vrlo nepouzdanim jer se ponovljenom primenom algoritma nad istim podacima dobijaju neuporedivo različiti rezultati.

4.3 Vizualizacija mreža

Vizualizacija društvenih mreža predstavlja zaseban izazov. Iako sama vizualizacija nema direktne veze sa analizom, ona može da nas uputi u kom smeru treba sprovoditi analizu. Sa druge strane, saznanja dobijena u analizi je moguće iskoristi za vizualizaciju u cilju što efikasnije komunikacije saznanja do kojih se došlo analizom. Prilikom vizualizacije mreža, mnogi vizuelni elementi se mogu arbitrarno predstaviti. Za predstavljanje strukture jedne mreže dovoljno da se na vizualizaciji vidi koji je čvor sa kojim čvorom povezan i eventualno kog je smera ta veza u slučaju usmerenih mreža. Svi ostali elementi vizualizacije su arbitrarni i zavise od cilja koji se želi postići vizualizacijom. To uključuje raspored čvorova u mreži, veličinu, boju i oblik čvorova, debljinu, boju i zakrivljenje ivica, brojne tekstualne oznake koje mogu pratiti čvorove ili ivice i sl. Tako da je za efektivnu vizualizaciju mreža potrebno razmišljati o dizajnu više nego to je slučaj kod klasičnih vizualizacija podataka koji nisu relacioni.

Uticaj tih arbitrarnih vizualnih elemenata najvidljiviji je na primeru vizualizacije mreže osnovnog uzorka u slučaju Slike 4 i Slike 5 iz poglavlja 3.3.1. Obe vizualizacije komuniciraju identične relacione podatke (koji čvorovi su međusobno povezani), ali Slika 5 osim tih relacionih podataka sa sobom nosi mnogo više informacija, uključujući pripadnost čvorova strankama, stepen čvorova (izražen kroz veličinu čvora), učestalost relacije (izražena debljinom ivice); takođe, raspored čvorova je formiran na osnovu relacija. Dakle Slika 5 i pored toga što nosi mnogo više informacija je već na prvi pogled je mnogo razumljivija od Slike 4.

Mreža osnovnog uzorka je sa svojih 40 čvorova bila dovoljno mala da bi se efektivno vizualizovala konvencionalnim sociogramom. Isti metod vizualizacije nije efektivan za mrežu proširenog uzorka jer je toliku količinu podatak nemoguće efektivno prikazati na ograničenom prostoru A4 stranice. Iz tog razloga česte su interaktivne vizualizacije velikih, kompleksnih mreža, gde se zanimljivi delovi mogu zumirati. Čak i kada bi se uklonili nazivi čvorova, mreža bi i dalje bila previše kompleksna da bi bila razumljiva (Prilog IV). Zapravo, takva vizualizacija

je u potpunosti beskorisna jer ne komunicira nikakve informacije. Jedini zaključak koji bi mogao da se izuče na osnovu te vizualizacije je odnos stepena između čvorova različitih stranaka (što se može mnogo efikasnije predstaviti običnim grafikonom), a same relacije su potpuno nerazumljive.

Zbog navedenih nedostataka konvencionalnih sociograma za vizualizaciju velikih mreža u radu je primenjen metod vizualizacije grafikonom košnice (Slika 6). To je nov metod, predstavljen u radu koji su objavili Krzywinski et al., 2012. god, koji omogućava vizualizaciju mreža bez nasumičnog generisanja rasporeda čvorova. Umesto klasičnog rasporeda, čvorovi su podeljeni na 3 ose, a lokacija na osi zavisi od zadatog parametra. Podelom čvorova na 3 ose (prema nekoj promenljivoj) omogućava bolju preglednost čvorova. Ta podela u ovom radu je izvršena prema kategorijskoj promenljivoj (pripadnost: opoziciji, vlasti, nepoznato), ali je moguće izvršiti prema numeričkom obeležju, tako što se skala obeležja podeli u 3 intervala, pa čvorovi pripadaju osama u zavisnosti od toga u kom je intervalu vrednost promenljive tog čvora. Lokacija na osi takođe zavisi od neke promenljive (u slučaju grafikona košnice sa Slika 6 to je stepen). Na taj način lokacija čvora u mreži, koja je u sociogramima prepuštena arbitrarnosti i nasumičnom generisanju, u grafikonu košnice ima značaj, tako da po definiciji prenosi više korisnih informacija. Svi ostali grafički elementi koji postoje u sociogramu, postoje i u grafikonu košnice (veličina i boja čvorova i ivica, natpisi itd.)

Na Slici 6 vidi se da postoje 3 grupe čvorova, i odmah se stiče utisak o brojnosti čvorova u svakoj od 3 grupe. Drugi vid informacije koju vizualizacija pruže jeste pozicija čvorova na osama, ali i dužina samih osa. Osa na kojoj se nalaze čvorovi koji pripadaju vladajućim strankama je mnogo kraća od druge dve, što nam odmah daje informaciju da čvorovi na toj osi manje komuniciraju od čvorova na druge dve ose, s obzirom da pozicija čvora na osi održava stepen čvora. Poređenje stranaka prema stepenu takođe postaje jasno na prvi pogled. Svi elementi sociograma i dalje su prisutni, tako da veličina svakog čvora komunicira neki pokazatelj (u slučaju Slike 6 to je centralnost svojstvenog vektora), debljina i nivo transparentnosti ivica zavisi od težine ivica.

Na grafikonu košnice sa Slike 6 najduža je osa opozicije, što znači da čvorovi na kraju te ose imaju najviše vrednosti stepena. Šta više, osa opozicije je dva puta duža od ose opozicije, pa su čvorovi na sredini te ose otprilike sličnog stepena kao čvorovi na krajevima ose vlasti. Na osi opozicije prema stepenu dominiraju čvorovi pokreta „Dosta je bilo“, Demokratske stranke i Socijaldemokratske stranke, dok čvorovi drugih stranka imaju značajno niže vrednosti stepena.

Kada je u pitanju vrednost centralnosti svojstvenog vektora jasna je dominacija čvorova iz pokreta „Dosta je bilo“, koji su mnogo veći od svih ostalih čvorova u mreži. Veliki broj prikazanih čvorova onemogućuje identifikaciju konkretnih čvorova, naročito na osi čvorova čija stranačka pripadnost nije poznata, ali intenzitet komunikacije sa tim čvorovima se može oceniti na osnovu zastupljenosti boja ivica vidljivih između čvorova koji pripadaju strankama i čvorova za koje pripadnost nije poznata (boja ivica odgovara stranačkoj pripadnosti čvora koji komunicira sa čvorom nepoznate afilijacije, vidi Listing 15). Na osnovu toga se može zaključiti da pokret „Dosta je bilo“ i Demokratska stranka imaju najintenzivniju komunikaciju kada je opozicija u pitanju, i da čvorovi Sprske napredne stranke imaju intenzivniju komunikaciju sa čvorovima sa desne ose od Socijalističke partije Srbije.

Nijedan od zaključaka izvedenih analizom vizualizacije sa Slike 6, se ne može izvući analizom konvencionalnog sociograma (Prilog IV), osim možda onih dobijenih poređenjem veličine čvorova, što takođe može biti problem posebno ako se koristi algoritam koji generiše raspored u kome se mnogi čvorovi međusobno preklapaju.

4.4 Primena R-a

4.4.1 Podrška za analizu društvenih mreža

U ovom radu korišćena je funkcionalnost samo dva R paketa (*igraph* i *HiveR*) od ukupno 18 paketa koji se mogu koristiti u analizi društvenih mreža (vidi poglavlje 1.3.2.1). Funkcionalnost samo jednog paketa (*igraph*) bila je dovoljna za sprovođenje svih analiza kao i za skoro sve vizualizacije. Korišćenjem ta dva pomenuta paketa primećeni su neki nedostaci, koji će verovatno biti ispravljani u budućim verzijama paketa.

U trenutnoj verziji paketa *igraph* nemoguće je prikazati dvosmernu ivicu kao jednu krivu liniju sa dva smera. Još jedan nedostatak je odsustvo ugrađenog algoritma za raspored čvorova koji bi uzimao u obzir zajednice izračunate pomoću nekih od *igraph* algoritama za klasterovanje. Takođe, kada je u pitanju vizualizacija, *igraph* paket nema mogućnost kontrolisanja debljine okvira čvorova, što bi moglo da bude korisno kada je u pitanju predstavljanje neke promenljive, ali i sa estetske strane. *HiveR* paket ima brojne nedostatke s obzirom da se nalazi u veoma ranoj fazi razvoja (verzija 0.2.55), što ne iznenađuje s obzirom da je i sam metod grafikona košnice još uvek relativno nov. Veliki deo posla za pripremu vizualizacije mora se uraditi manuelno, tako je za vizualizaciju korišćenjem ovog paketa potrebno više programiranja, pa paket nije najbolje rešenje za početnike. Nažalost trenutno ne postoji alternativni R paket za generisanje grafikona košnice.

4.4.2 Performanse

Izvršenje nekih od funkcija naročito za sređivanje podataka (poglavlje 3.2) može da potraje dugo ako se radi o velikoj količini podataka. U slučajevima kada se radi o analizi mnogo većih uzoraka, preporučljiva je dodatna optimizacija koda jer bi izvršenje moglo da potraje i više sati. Lak način kojim bi vreme izvršenja koda moglo značajno da se ubrza (za oko 10% do 20%) je uz pomoć *data.table* paketa i korišćenjem objekta *data.table* umesto *data.frame* objekta, koji u radu nije korišćen s obzirom da brzina izvršavanja koda nije od primarne važnosti u ovom radu.

Jedan od razloga za tako sporo izvršavanje je i činjenica da R koristi samo jedno jezgro procesora. Podrška u samom jeziku R za višenitno izvršavanje (*multithreading*) ne postoji, ali je moguće dobiti tu funkcionalnost preko paketa. Takođe postoje i paketi koji omogućavaju paralelizaciju i izvršavanje delova koda uz pomoć grafičkih čipova (*GPU*) umesto procesora (McCallum & Weston, 2012). U praksi se, međutim, jako optimizovan kod za te potrebe piše u drugim programskim jezicima koji imaju mnogo bolju podršku za višenitnost i paralelizaciju (npr. Python ili C).

4.4.3 Replikabilnost koda

R je programski jezik koji se veoma brzo razvija. Nove verzije izlaze često (samo u 2016. god. objavljeno je četiri nove verzije), a nove verzije raznih paketa mogu se pojavljivati još češće. Kod pisan za jednu verziju R-a će najverovatnije raditi i na drugoj verziji R-a, ali veći problem predstavljaju promene u paketima. Međutim, CRAN repozitorijumi sadrže sve verzije svih paketa koji su ikada objavljeni na njemu, tako da reprodukcija nije problematična ako je poznat broj aktuelne verzije paketa u vreme pisanja koda. Ovaj problem nije previše prisutan, i većina paketa će raditi sa veoma starim kodom, ali postoje slučajevi kada kod jednostavno neće biti kompatibilan sa novijom verzijom paketa (ili samo nekih funkcija u okviru paketa koje su izmenjene) što može stvoriti probleme. Zato je važno uvek imati na umu verziju za koju je kod pisan i verzije paketa na sistemu gde se kod izvršava. Za kod korišćen u ovom radu data je i datoteka sa verzijama paketa koji su korišćeni i verzija samog programskog jezika R (vidi poglavlje 2.3).

4.5 Etička razmatranja

4.5.1 Privatnost

Čitav rad baziran je na podacima prikupljenim automatizovanim procesom sa Tviter servera. Svi korisnici Tvitera koji u podešavanjima nisu zaštitili svoje tvitove od 'javnosti', de facto i de jure pristaju na javnost svojih tvitova. Zasebno je pitanje osvešćenosti korisnika Tvitera, kada je u

pitanju retenzija podataka od strane Tvitera. Svaki tvit naloga koji je javan, ostaje javan na internetu (i lako dostupan preko internet pretraživača) dok god se ne obriše nalog, konkretan tvit ili ne promene postavke privatnosti. Tviter API ima limit od 3200 poslednjih tvitova koji se mogu povući, ali taj limit ne ograničava internet pretraživače preko kojih se može doći do bilo kog tvita (ili manuelno listanje statusa nekog naloga na sajtu Tvitera). Postoje kompanije koje se bave sakupljanjem podataka sa Tvitera, pa je moguće kupiti podatke, koji uključuju i podatke o obrisanim nalogima i onim nalogima koji su u međuvremenu promenili postavke privatnosti.

Isto važi i za podatke o lokaciji. Iako bi sa istraživačke strane bilo veoma zgodno imati podatke o geolokaciji svakog tvita, kao što je rečeno u poglavlju 3.1, veoma mali broj korisnika ima omogućeno deljenje lokacije.

4.5.2 Klasifikacija naloga

Posebno treba naglasiti da se metodama klasterovanja, primenjenim ne prošireni uzorak, ne etiketiraju korisnici čiji su nalozi tom metodom klasifikovani u klastere pojedinih stranaka. Pošto se ne radi o analizi sadržaja tvitova već samo o mapiranju komunikacija na osnovu metapodataka, ne može se tvrditi da neki čvor pripada klasteru zato što je sa saglasan sa ostalim čvorovima tog klastera ili je sa njima u stalnoj nesuglasici. Čak i da je rađena analiza sadržaja (*natural language processing*) i razdvojena afirmativna od negativne komunikacije i na osnovu tih podataka izvršena klaster analiza, opet bi etiketiranje korisnika bilo problematično. Inače, u analizi komunikacije metapodaci se često smatraju mnogo korisnijim od same sadržine ("Sandy" Pentland, 2014).

Klasifikacija građana na osnovu internet komunikacije je realan problem sa kojim se susreću građani država sa represivnim nedemokratskim režimima. Sa druge strane u zapadnim demokratijama isti metodi se koriste za pronalaženje potencijalnih terorista, a nediskriminacionom analizom metapodataka ugrožava se privatnost građana.

5 ZAKLJUČAK

5.1 Analiza izbora na Tviteru

Merenje uticaja u bilo kojoj društvenoj mreži je po definiciji veoma arbitrarna stvar, jer zavisi od toga kako definišemo taj uticaj i kako ga merimo. Postavlja se pitanje da li u te svrhe koristiti osnovne pokazatelje uticaja na Tviteru (npr. prosečan broj retvita po tvitu) ili rezultate analize društvenih mreža. U ovom radu je pokazano da je neophodno uzeti u obzir i jednu i drugu vrstu metrika u cilju potpunijeg sagledavanja celokupne situacije, jer se na osnovu samo jedne vrste analize mogu steći pogrešni zaključci.

Ako bismo uzeli u obzir samo osnovne Tviter podatke, stekli bismo utisak da je Srpska napredna stranka dominirala kampanjom. Naloge ove stranke karakteriše mali broj tvitova, a veliki broj retvitova. Moguće objašnjenje za ovaj fenomen je pojava tzv. 'botova', tj. da stranački aktivisti organizovano retvituju i označavaju kao omiljeni tvitovi jedne stranke, kako bi se odala slika popularnosti na internetu. Zbog toga konvencionalni pokazatelji popularnosti na *online* društvenim mrežama mogu biti nepouzdati.

Kao kontrast iznetim konvencionalnim pokazateljima popularnosti na Tviteru, stoje rezultati dobijeni analizom društvene mreže političara. Analiza društvene mreže je u ovom kontekstu važna baš zato što uzima u obzir interakciju. Kada političari međusobno interaguju postoji potencijal da njihov auditorijum bude mnogo veći (tvitove mogu videti pratioci političara A i pratioci političara B), za razliku od običnog tvita u kom se političar obraća samo svojim pratiocima, što pomaže u razbijanju filter mehura (*filter bubble*), tj. pojave da glasači prate samo istomišljenike, pa figurativno žive u mehuru u kome mogu da čuju samo svoj eho (Bozdag & van den Hoven, 2015). Zbog toga možemo da zaključimo da su političari, koji zauzimaju centralniju poziciju u mapiranoj mreži interakcije između političara, imali veći uticaj na Tviteru tokom kampanje (barem u smislu dolaženja do što većeg broja korisnika). Mrežne metrike pokazuju potpuno drugačiji rezultat od onog koje pokazuju konvencionalni, gorepomenuti pokazatelji. Pokret „Dosta je bilo“, koji nema jaku stranačku infrastrukturu, je osvojio 6,2% glasova (RIK, 2016) i gotovo utrostručio svoju podršku u odnosu na izbore iz 2014. god. Zanimljivo je da ova lista gotovo da nije imala budžet za kampanju u mainstream medijima, zbog čega je njihova kampanja okarakterisana kao dobar model kampanje oslonjene na internet i računarske društvene mreže (Transparentnost Srbija, 2016). Ukupna centralnost, merena kompozitnim pokazateljem koji uzima u obzir različite mere centralnosti, je upravo izdvojila

pokret „Dosta je bilo“, kao najcentralniju grupaciju u analiziranoj mreži. Srpska napredna stranka je prema istom pokazatelju na trećem mestu, i to uglavnom zahvaljujući dolazećim tvitovima (dakle tvitovima upućenim članovima SNS-a), nego zahvaljujući njihovoj interaktivnosti. Utvrđeno je da je interakcija bila slaba čak i između samih članova te stranke.

5.2 Preporuke analiziranim strankama

Na osnovu rezultata analize mreže osnovnog uzorka moguće je formulisati sledeće preporuke analiziranim strankama.

Srpska napredna stranka – intenzivnija interakcija, na prvom mestu međusobna, ali i sa pripadnicima drugih stranaka. Na potrebu za intenzivnijom interakcijom ukazuje odsustvo gigantske komponente, ali i klaster analiza prema kojoj su dva člana ove stranke zalutala u klastere drugih stranaka (Slika 10). Lokalizacija zalutalih čvorova ove stranke pri klaster analizi je zanimljiva i zahteva dodatnu kvalitativnu analizu kako bi se dao odgovor na pitanje zašto su čvorovi Srpske napredne stranke svrstani u klaster Demokratske stranke. S obzirom da je ova stranka u trenutku raspisivanja izbora bila na svom vrhuncu (kada je reč o podršci), verovatno je strategija bila da ne izgube glasove koje već imaju, a ne da se bore za nove glasače. Postoji mogućnost da je komunikacija na Tviteru bila namerno s vrha ka dnu, a da je slobodna interakcija bila obeshrabrivana.

Srpska radikalna stranka – intenzivnija interakcija sa pripadnicima drugih stranaka, interaktivniji predsednik stranke. SRS se pokazala kao najhomofiličnija stranka, tj. stranka čiji akteri najmanje komuniciraju sa akterima drugih stranaka. Čvor koji povezuje celu stranku je Vjerica Radeta, bez nje ova stranka ne bi imala gigantsku komponentu. Drugi članovi ove stranke imali su sličan broj tvitova u toku kampanje, ali su ređe ulazili u diskusije, pa se za te čvorove preporučuje i učestalija međusobna interakcija. Predsednik stranke bi trebalo da bude interaktivniji sa čvorovima svoje stranke (što je posebno poželjno jer predsednik SRS ima mnogo više pratioca na Tviteru od bilo kog drugog čvora svoje stranke, pa bi ta interakcija potencijalno dovela njegove pratioce do ostalih čvorova stranke), ali i čvorovima drugih stranaka.

Nova Srbija – više istaknutih članova na Tviteru. Iako je veći broj istaknutih članova na Tviteru povoljan za svaku stranku, važno je sagledati taj broj u kontekstu kapaciteta drugih stranaka na Tviteru. Nova Srbija je imala samo dva istaknuta člana na Tviteru. To možda reflektuje i realnost, tj. da Nova Srbija i ima samo dva istaknuta člana. Ova stranka je bila na zajedničkoj

listi sa Demokratskom strankom, pa je i njihova interakcija na Tviteru bila usko isprepletana, toliko da ih je klaster analiza svrstala u jedan klaster. To može da bude i problematično u nekim narednim izborima, ako ove dve stranke izlaze odvojeno na izbore. U tom slučaju bi ovako mali broj istaknutih članova na Tviteru bio još приметniji.

Socijalistička partija Srbije aktiviranje naloga predsednika stranke, dodatni istaknuti članovi na Tviteru, interakcija sa drugim strankama. Neaktivan nalog predsednika stranke je veliki propust. Kada su druge stranke u pitanju obično nalozi predsednika imaju najviše pratioaca (Prilog II). S obzirom da se radi o besplatnom načinu za diseminaciju informacija i promociju, nema opravdanja za neaktivnost. Posle SRS, SPS je druga najhomofiličnija stranka, uz činjenicu da imaju manji broj uticajnih članova na Tviteru od SRS, što se nije odrazilo na spremnost da interaguju sa drugim strankama, čak ni sa Srpskom naprednom strankom sa kojima su do i u toku izbora činili vladajuću većinu.

Demokratska stranka Srbije – otvaranje naloga predsednice stranke, dodatni istaknuti članovi na Tviteru. Za razliku od predsednika SPS-a koji je za vreme izbora posedovao, doista neaktivan, Tviter nalog, predsednica DSS-a ga nije čak ni posedovala. Nalog je otvorila neposredno po završetku izbora, tako da se može reći da je ova preporuka apriori prihvaćena, ali je takođe u međuvremenu i napustila svoju stranku.

Dveri – gotovo ne postojeća kampanja na Tviteru, konkretna preporuka nemoguća.

Liberalno demokratska partija – dodatni istaknuti članovi na Tviteru, intenzivnija međusobna interakcija, interaktivniji predsednik stranke. Sva međusobna komunikacija između članova ove stranke se dešava baš sa stranačkim nalogom, što znači da istaknuti članovi stranke međusobno ne interaguju na Tviteru.

Socijaldemokratska stranka – dodatni istaknuti članovi na Tviteru, uključujući i predsednika stranke, koji i ne poseduje Tviter nalog. Kao stranka sa veoma malim brojem istaknutih članova, poznatih široj javnosti, a čiji je predsednik bivši predsednik Republike Srbije, ova stranka se može okarakterisati kao „*one man show*“, u smislu da je jedino predsednik stranke poznat najširoj javnosti. Činjenica da predsednik ove stranke i ne poseduje Tviter nalog je neobjašnjiva.

Dosta je bilo – intenzivnija interakcija sa drugim strankama. Kod pokreta Dosta je bio prisutan iznenađujuće visok stepen homofilije. Kao najcentralnija stranka na Tviteru tokom izbora 2016. (Slika 8), za očekivati je da taj stepen bude niži. Njihova komunikacija sa drugim strankama se uglavnom svodi na komunikaciju sa Demokratskom strankom, Novom strankom i

Socijaldemokratskom strankom. Da li je to zbog toga što ove stranke dele isto biračko telo pa je u kompeticiji za glasove dolazilo do vatrenih diskusija, ili je u pitanju istinska homofilija, gde je bilo komunikacije između onih koji se najbolje razumeju? U svakom slučaju, intenzivnija interakcija sa ostalim strankama je poželjna, zbog diseminacije informacija i razbijanja „filter bubble-a“.

Demokratska stranka – prema svim metrikama Demokratska stranka je na Tviteru imala komparativno veoma jako prisustvo.

5.3 Mogući pravci budućih istraživanja

5.3.1 Ponovljena analiza

Analiza sprovedena u ovom radu može se ponoviti na istom uzorku za drugi vremenski period. Takva analiza za period između kampanja za izbore 2016. i 2017. god. bi pokazala razlike u ponašanju političkih aktera za vreme kampanje i van kampanje. Analiza bi mogla da se ponovi i za period trajanja kampanje za izbore 2017. god., kako bi se ustanovilo da li je došlo do promena u strategiji komunikacije na *online* društvenim mrežama kod pojedinih političkih partija.

Ista ponovljena analiza mogla bi se sprovesti i na izmenjenom uzorku. Uzorak bi mogao da se dalje proširi praćenjem popularnih haštagova (*hashtag*), što bi moglo da se postigne skriptom koja bi automatski preuzimala *id* svakog naloga koji koristi haštag koji političari upotrebljavaju. Prikupljanje podataka na ovaj način moralo bi se vršiti u veoma kratkim vremenskim intervalima, kontinuirano tokom trajanja kampanje, zbog limita Tviter API-a.

5.3.2 Analiza retvitova i detekcija 'botova'

U poglavlju 4.2.3 izneta je osnovana sumnja da pojedine stranke imaju organizovane aktiviste koji retvituju i označavaju kao omiljene tvitove svoje stranke. Povlačenjem podataka o tvitovima koji imaju neobično mnogo retvitova preko Tviter API-a, mogao bi se dobiti spisak naloga koji su retvitovali takav tvit. Nalozi koji veoma često retvituju tvitove jedne stranke bi bili analizirani, tako što bi se naknadno ispitala njihova sveukupna komunikacija. To bi moglo da se sprovede formiranjem bipartitne mreže gde bi ti sakupljeni nalozi bili primarni čvorovi, a nalozi sa kojima oni interaguju sekundarni. Topologija takve mreže za skup 'botova', koji nemaju izraženu komunikaciju van onoga što organizovano rade, bi bila izraženo drugačija od topologije mreže koju čine nalozi koji 'organski' retvituju tvitove koji im se dopadaju. Poređenje vizualizacija takvih mreža '*retvitera*' za različite stranke bi bila veoma interesantno.

5.3.3 Analiza sadržaja prikupljenih tvitova

Sadržaj preuzetih tvitova za potrebe analize sprovedene u ovom radu se može dalje analizirati. Klasifikacija svih interakcija na pozitivne, negativne i neutralne korišćenjem obrade prirodnog jezika, a potom kreiranje mreže pozitivnih, negativnih i neutralnih interakcija, potencijalno bi otkrila zanimljive paterne. Klasifikacija čvorova nepoznate stranačke afilijacije prema tome kakve su vrste interakcije prema različitim političkim akterima bi takođe bila poboljšana. Rezultujuće vizualizacije, u kojima bi pozitivna i negativna komunikacija bila predstavljena različitim bojama, bi mogle biti vrlo intuitivne i zanimljive.

6 LITERATURA

- Ahmed, S., Jaidka, K., & Cho, J. (2016). The 2014 Indian elections on Twitter: A comparison of campaign strategies of political parties. *Telematics and Informatics*, 33(4), 1071–1087. <https://doi.org/10.1016/j.tele.2016.03.002>
- Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1), 47–97. <https://doi.org/10.1103/RevModPhys.74.47>
- Alexanderson, G. L. (2006). About the cover: Euler and Königsberg's Bridges: A historical view. *Bulletin of the American Mathematical Society*, 43(4), 567–574. <https://doi.org/10.1090/S0273-0979-06-01130-X>
- ANEM. (2016). *Pravni monitoring medijske scene u Srbiji*. Asocijacija nezavisnih elektronskih medija. Preuzeto sa http://www.anem.rs/admin/download/files/_id_2480/MR%2067%20DEC%202015.pdf
- Anstead, N., & O'Loughlin, B. (2015). Social Media Analysis and Public Opinion: The 2010 UK General Election. *Journal of Computer-Mediated Communication*, 20(2), 204–220. <https://doi.org/10.1111/jcc4.12102>
- Babovic, M. (2005). Social networks: Networking of social actors in the sphere of economic activities. *Sociologija*, 47(4), 351–370. <https://doi.org/10.2298/SOC0504351B>
- Bacallao-Pino, L. M. (2016). Redes sociales, acción colectiva y elecciones: los usos de Facebook por el movimiento estudiantil chileno durante la campaña electoral de 2013. *Palabra Clave - Revista de Comunicación*, 19(3), 810–837. <https://doi.org/10.5294/pacla.2016.19.3.6>
- Bessi, A. (2015). *staTools: Statistical Tools for Social Network Analysis*. Preuzeto sa <https://CRAN.R-project.org/package=staTools>
- Bonacich, P. (1987). Power and Centrality: A Family of Measures. *American Journal of Sociology*, 92(5), 1170–1182. <https://doi.org/10.1086/228631>
- Borgatti, S. P., & Everett, M. G. (2006). A Graph-theoretic perspective on centrality. *Social Networks*, 28(4), 466–484. <https://doi.org/10.1016/j.socnet.2005.11.005>

- Bozdag, E., & van den Hoven, J. (2015). Breaking the filter bubble: democracy and design. *Ethics and Information Technology*, 17(4), 249–265. <https://doi.org/10.1007/s10676-015-9380-y>
- Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30(1–7), 107–117. [https://doi.org/10.1016/S0169-7552\(98\)00110-X](https://doi.org/10.1016/S0169-7552(98)00110-X)
- Butts, C. T. (2008a). network: a Package for Managing Relational Data in R. *Journal of Statistical Software*, 24(2). Preuzeto sa <http://www.jstatsoft.org/v24/i02/paper>
- Butts, C. T. (2008b). Social network analysis: A methodological introduction. *Asian Journal Of Social Psychology*, 11(1), 13–41. <https://doi.org/10.1111/j.1467-839X.2007.00241.x>
- Butts, C. T. (2015). *network: Classes for Relational Data*. The Statnet Project (<http://statnet.org>). Preuzeto sa <http://CRAN.R-project.org/package=network>
- Butts, C. T. (2016). *sna: Tools for Social Network Analysis*. Preuzeto sa <https://CRAN.R-project.org/package=sna>
- Chan, M. (2016). Social Network Sites and Political Engagement: Exploring the Impact of Facebook Connections and Uses on Political Protest and Participation. *Mass Communication and Society*, 19(4), 430–451. <https://doi.org/10.1080/15205436.2016.1161803>
- Chen, C., Jin, J., Beisner, B., McCowan, B., & Fushing, H. (2016). *DCG: Data Cloud Geometry (DCG): Using Random Walks to Find Community Structure in Social Network Analysis*. Preuzeto sa <https://CRAN.R-project.org/package=DCG>
- CRAN. (2016). The Comprehensive R Archive Network. Preuzeto sa <https://cran.r-project.org/web/packages/>
- Csardi, G. (2015). *cranlogs: Download Logs from the “RStudio” “CRAN” Mirror*. Preuzeto sa <https://CRAN.R-project.org/package=cranlogs>
- Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, 1695.

- Curley, J. P., & Chase, I. D. (2016). *hierformR: Analysis of Dynamics Hierarchy Formation*. Preuzeto sa <https://CRAN.R-project.org/package=hierformR>
- Dempwolf, C. S., & Lyles, L. W. (2012). The Uses of Social Network Analysis in Planning: A Review of the Literature. *Journal of Planning Literature*, 27(1), 3–21. <https://doi.org/10.1177/0885412211411092>
- Dobrota, M. (2015). *A Statistical Approach To Composite Indicators Construction Based on Ivanović Distance (Unpublished doctoral dissertation)*. Preuzeto sa <http://www.fon.bg.ac.rs/downloads/2014/12/Doktorska-disertacija-Marina-sa-prilogom-Locked.pdf>
- Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets: reasoning about a highly connected world*. New York: Cambridge University Press.
- Edunov, S., Diuk, C., Filiz, I. O., Bhagat, S., & Burke, M. (2016). *Three and a half degrees of separation*. Facebook. Preuzeto sa <https://research.facebook.com/blog/three-and-a-half-degrees-of-separation/>
- Evgeny Morozov. (2009). Iran: Downside to the “Twitter Revolution.” *Dissent*, 56(4), 10–14. <https://doi.org/10.1353/dss.0.0092>
- Farine, D. (2016). *assortnet: Calculate the Assortativity Coefficient of Weighted and Binary Networks*. Preuzeto sa <https://CRAN.R-project.org/package=assortnet>
- Farine, D. R. (2016). *asnipe: Animal Social Network Inference and Permutations for Ecologists*. Preuzeto sa <https://CRAN.R-project.org/package=asnipe>
- Freedom House. (2016). Serbia | Country report | Freedom of the Press | 2015. Preuzeto October 2, 2016, sa <https://freedomhouse.org/report/freedom-press/2015/serbia>
- Freeman, L. C. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry*, 40(1), 35. <https://doi.org/10.2307/3033543>
- Freeman, L. C. (2004). *The development of social network analysis: a study in the sociology of science*. Vancouver, BC : North Charleston, S.C: Empirical Press ; BookSurge.

- Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, 21(11), 1129–1164.
<https://doi.org/10.1002/spe.4380211102>
- Gentry, J. (2015). *twitteR: R Based Twitter Client*. Preuzeto sa <https://CRAN.R-project.org/package=twitteR>
- Gökçe, O. Z., Hatipoğlu, E., Göktürk, G., Luetgert, B., & Saygin, Y. (2014). Twitter and Politics: Identifying Turkish Opinion Leaders in New Social Media. *Turkish Studies*, 15(4), 671–688. <https://doi.org/10.1080/14683849.2014.985425>
- Haleva-Amir, S. (2016). Not all about that Facebook: political campaigns and civic engagement in the 2015 elections. *Israel Affairs*, 22(3–4), 711–726.
<https://doi.org/10.1080/13537121.2016.1174375>
- Hampton, K. N., Goulet, L. S., Marlow, C., & Rainie, L. (2012). *Why most Facebook users get more than they give*. Pew Research Center's Internet & American Life Project. Preuzeto sa <http://www.pewinternet.org/2012/02/03/why-most-facebook-users-get-more-than-they-give/>
- Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., Krivitsky, P. N., Bender-deMoll, S., & Morris, M. (2016). *statnet: Software Tools for the Statistical Analysis of Network Data*. The Statnet Project (<http://www.statnet.org>). Preuzeto sa CRAN.R-project.org/package=statnet
- Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., & Morris, M. (2008). statnet: Software Tools for the Representation, Visualization, Analysis and Simulation of Network Data. *Journal of Statistical Software*, 24(1), 1–11.
- Hanson, B. A. (2016). *HiveR: 2D and 3D Hive Plots for R*.
- Heo, Y.-C., Park, J.-Y., Kim, J.-Y., & Park, H.-W. (2016). The emerging viewertariat in South Korea: The Seoul mayoral TV debate on Twitter, Facebook, and blogs. *Telematics and Informatics*, 33(2), 570–583. <https://doi.org/10.1016/j.tele.2015.08.003>
- Hoff, P., Fosdick, B., Volfovsky, A., & He, Y. (2015). *amen: Additive and Multiplicative Effects Models for Networks and Relational Data*. Preuzeto sa <https://CRAN.R-project.org/package=amen>

- Hoppe, B., & Reinelt, C. (2010). Social network analysis and the evaluation of leadership networks. *The Leadership Quarterly*, 21(4), 600–619.
<https://doi.org/10.1016/j.leaqua.2010.06.004>
- Huberty, M. (2015). Can we vote with our tweet? On the perennial difficulty of election forecasting with social media. *International Journal of Forecasting*, 31(3), 992–1007.
<https://doi.org/10.1016/j.ijforecast.2014.08.005>
- Ihaka, R., & Gentleman, R. (1996). R: A Language for Data Analysis and Graphics. *Journal of Computational and Graphical Statistics*, 5(3), 299. <https://doi.org/10.2307/1390807>
- Ivanović, B. (1974). Comment établir une liste des indicateurs de developpment. *Revue de Statistique Apliquée*, 22(2).
- Jackson, M. O. (2008). *Social and economic networks*. Princeton, NJ: Princeton Univ. Press.
- Jeremić, V. (2012). *Statistički model efikasnosti zasnovan na Ivanovićevom odstojanju (neobjavljena doktorska disertacija)*. Univerzitet u Beogradu, Srbija. Preuzeto sa <http://www.doiserbia.nb.rs/phd/university.aspx?theseid=BG20121001JEREMIC>
- Khondker, H. H. (2011). Role of the New Media in the Arab Spring. *Globalizations*, 8(5), 675–679. <https://doi.org/10.1080/14747731.2011.621287>
- Krueger, K., & Krueger, K. (2016). *Dominance: ADI (Average Dominance Index), Social Network Graphs with Dual Directions, and Music Notation Graph*. Preuzeto sa <https://CRAN.R-project.org/package=Dominance>
- Krzywinski, M., Birol, I., Jones, S. J., & Marra, M. A. (2012). Hive plots--rational approach to visualizing networks. *Briefings in Bioinformatics*, 13(5), 627–644.
<https://doi.org/10.1093/bib/bbr069>
- Li, C., Li, Q., Van Mieghem, P., Stanley, H. E., & Wang, H. (2015). Correlation between centrality metrics and their application to the opinion model. *The European Physical Journal B*, 88(3). <https://doi.org/10.1140/epjb/e2015-50671-y>
- López Meri, A. (2015). Redes sociales y campañas electorales: Twitter como fuente informativa en las elecciones catalanas del #25N. *Comunicació: Revista de Recerca i d'Anàlisi*, (32), 115–137. <https://doi.org/10.2436/20.3008.01.138>

- Lotan, G., Graeff, E., Ananny, M., Gaffney, D., Pearce, I., & Boyd, D. (2011). The Revolutions Were Tweeted: Information Flows During the 2011 Tunisian and Egyptian Revolutions. *INTERNATIONAL JOURNAL OF COMMUNICATION*, 5, 1375–1405.
- Macy, M. W., & Willer, R. (2002). From Factors to Factors: Computational Sociology and Agent-Based Modeling. *Annual Review of Sociology*, 28(1), 143–166.
<https://doi.org/10.1146/annurev.soc.28.110601.141117>
- McCallum, Q. E., & Weston, S. (2012). *Parallel R*. Beijing: O'Reilly.
- Missaoui, R., & Sarr, I. (Eds.). (2014). *Social Network Analysis - Community Detection and Evolution*. Cham: Springer International Publishing. Preuzeto sa
<http://link.springer.com/10.1007/978-3-319-12188-8>
- Morandat, F., Hill, B., Osvald, L., & Vitek, J. (2012). Evaluating the Design of the R Language. In J. Noble (Ed.), *ECOOP 2012 – Object-Oriented Programming* (Vol. 7313, pp. 104–131). Berlin, Heidelberg: Springer Berlin Heidelberg. Preuzeto sa
http://link.springer.com/10.1007/978-3-642-31057-7_6
- Nepusz, T. (2012). What are the differences between community detection algorithms in igraph? Preuzeto sa <http://stackoverflow.com/a/9478989/3967488>
- Newman, M. E. J. (2002). Assortative Mixing in Networks. *Physical Review Letters*, 89(20).
<https://doi.org/10.1103/PhysRevLett.89.208701>
- Newman, M. E. J. (2010). *Networks: an introduction*. Oxford ; New York: Oxford University Press.
- Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, 69(2). <https://doi.org/10.1103/PhysRevE.69.026113>
- Nightingale, G., & Nightingale, P. (2014). *SocialNetworks: Generates social networks based on distance*. Preuzeto sa <https://CRAN.R-project.org/package=SocialNetworks>
- Ostoic, A. R. (2016). *multiplex: Algebraic Tools for the Analysis of Multiple Social Networks*. Preuzeto sa <https://CRAN.R-project.org/package=multiplex>
- Ott, M. (2016). *strategicplayers: Strategic Players*. Preuzeto sa <https://CRAN.R-project.org/package=strategicplayers>

- Otte, E., & Rousseau, R. (2002). Social network analysis: a powerful strategy, also for the information sciences. *Journal of Information Science*, 28(6), 441–453.
<https://doi.org/10.1177/016555150202800601>
- Pena, J. B. (1977). *Problemas de la medición del bienestar y conceptos afines (una aplicación al caso Español)*. Madrid: INE.
- Pentland, A. (2014). *Social physics: how good ideas spread-the lessons from a new science*. New York: The Penguin Press.
- Perez-Perez, A. J. P.-L. R. M. R., & Bonet, F. J. (2012). *p2distance: Welfare's Synthetic Indicator*. Preuzeto sa <https://CRAN.R-project.org/package=p2distance>
- Perliger, A., & Pedahzur, A. (2011). Social Network Analysis in the Study of Terrorism and Political Violence. *PS: Political Science & Politics*, 44(1), 45–50.
<https://doi.org/10.1017/S1049096510001848>
- Pham, T., Sheridan, P., & Shimodaira, H. (2015a). PAFit: A Statistical Method for Measuring Preferential Attachment in Temporal Complex Networks. *PLOS ONE*, (9), e0137796.
<https://doi.org/10.1371/journal.pone.0137796>
- Pham, T., Sheridan, P., & Shimodaira, H. (2015b). *PAFit: Nonparametric Estimation of Preferential Attachment and Node Fitness in Temporal Complex Networks*. Preuzeto sa cran.r-project.org/package=PAFit
- Pham, T., Sheridan, P., & Shimodaira, H. (2016). Joint Estimation of Preferential Attachment and Node Fitness in Growing Complex Networks. *Scientific Reports*, 32558 EP.
<https://doi.org/10.1038/srep32558>
- Pham, Thong, Sheridan, Paul, Shimodaira, & Hidetoshi. (2016). Nonparametric Estimation of the Preferential Attachment Function in Complex Networks: Evidence of Deviations from Log Linearity. In Battiston, Stefano, De Pellegrini, Francesco, Caldarelli, Guido, ... Emanuela (Eds.), *Proceedings of ECCS 2014: European Conference on Complex Systems* (pp. 141–153). Cham: Springer International Publishing. Preuzeto sa http://dx.doi.org/10.1007/978-3-319-29228-1_13
- Pons, P., & Latapy, M. (2005). Computing Communities in Large Networks Using Random Walks. In pInar Yolum, T. Güngör, F. Gürgen, & C. Özturan (Eds.), *Computer and*

- Information Sciences - ISCIS 2005* (Vol. 3733, pp. 284–293). Berlin, Heidelberg: Springer Berlin Heidelberg. Preuzeto sa http://link.springer.com/10.1007/11569596_31
- R Core Team. (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Preuzeto sa <https://www.R-project.org/>
- Raghavan, U. N., Albert, R., & Kumara, S. (2007). Near linear time algorithm to detect community structures in large-scale networks. *Physical Review E*, 76(3). <https://doi.org/10.1103/PhysRevE.76.036106>
- Reichardt, J., & Bornholdt, S. (2006). Statistical mechanics of community detection. *Physical Review E*, 74(1). <https://doi.org/10.1103/PhysRevE.74.016110>
- Reporters sans frontières. (2016). RSF supports journalists' refusal to "kneel" before government | RSF. Preuzeto October 2, 2016, sa <https://rsf.org/en/news/rsf-supports-journalists-refusal-kneel-government>
- Reuter, O. J., & Szakonyi, D. (2015). Online Social Media and Political Awareness in Authoritarian Regimes. *British Journal of Political Science*, 45(1), 29–51. <https://doi.org/10.1017/S0007123413000203>
- RIK. (2016). *Ukupni rezultati izbora 2016*. Preuzeto sa http://www.rik.parlament.gov.rs/doc/izbori-2016/rezultati/1.%20Zbirni_rezultati.pdf
- Romero, D. M., Galuba, W., Asur, S., & Huberman, B. A. (2011). Influence and Passivity in Social Media. In D. Gunopulos, T. Hofmann, D. Malerba, & M. Vazirgiannis (Eds.), *Machine Learning and Knowledge Discovery in Databases* (Vol. 6913, pp. 18–33). Berlin, Heidelberg: Springer Berlin Heidelberg. Preuzeto sa http://link.springer.com/10.1007/978-3-642-23808-6_2
- Rosvall, M., Axelsson, D., & Bergstrom, C. T. (2009). The map equation. *The European Physical Journal Special Topics*, 178(1), 13–23. <https://doi.org/10.1140/epjst/e2010-01179-1>
- Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105(4), 1118–1123. <https://doi.org/10.1073/pnas.0706851105>

- “Sandy” Pentland, A. (2014). Saving Big Data from Itself. *Scientific American*, 311(2), 64–67.
<https://doi.org/10.1038/scientificamerican0814-64>
- Sciandra, A., Gioachin, F., & Finos, L. (2012). *egonet: Tool for ego-centric measures in Social Network Analysis*. Preuzeto sa <https://CRAN.R-project.org/package=egonet>
- Scott, J. (2013). *SOCIAL NETWORK ANALYSIS (THIRD EDITION)*. Los Angeles: SAGE.
- Simmel, G., & Wolff, K. H. (1964). *The sociology of Georg Simmel*. New York: Free Pr. [u.a.].
- Sousa, F. B. de, & Zhao, L. (2014). Evaluating and Comparing the IGraph Community Detection Algorithms (pp. 408–413). IEEE. <https://doi.org/10.1109/BRACIS.2014.79>
- Squazzoni, F. (2012). *Agent-based computational sociology*. Hoboken, N.J: Wiley & Sons.
- Stadtfeld, C. (2013). *NetSim: A Social Networks Simulation Tool in R*. Preuzeto sa <https://CRAN.R-project.org/package=NetSim>
- Steinfeld, N. (2016). The F-campaign: a discourse network analysis of party leaders’ campaign statements on Facebook. *Israel Affairs*, 22(3–4), 743–759.
<https://doi.org/10.1080/13537121.2016.1174385>
- Sudhahar, S., Veltri, G. A., & Cristianini, N. (2015). Automated analysis of the US presidential elections using Big Data and network analysis. *Big Data & Society*, 2(1).
<https://doi.org/10.1177/2053951715572916>
- Theocharis, Y., & Quintelier, E. (2016). Stimulating citizenship or expanding entertainment? The effect of Facebook on adolescent participation. *New Media & Society*, 18(5), 817–836.
<https://doi.org/10.1177/1461444814549006>
- Tranos, E. (2013). *The geography of the Internet: cities, regions and Internet infrastructure in Europe*. Cheltenham, UK ; Northampton, Mass: Edward Elgar.
- Transparentnost Srbija. (2016). Monitoring kampanje za parlamentarne, pokrajinske i lokalne izbore. Preuzeto October 2, 2016, sa <http://www.transparentnost.org.rs/index.php/sr/59-srpski/naslovna/8421-monitoring-kampanje-za-parlamentarne-pokrajinske-i-lokalne-izbore>

- Türe, D. (deenes). (2013). How to make grouped layout in igraph? Preuzeto sa <http://stackoverflow.com/a/16390222/3967488>
- Valente, T. W., Coronges, K., Lakon, C., & Costenbader, E. (2008). How Correlated Are Network Centrality Measures? *Connections (Toronto, Ont.)*, 28(1), 16–26.
- Vitkovic, B. (2013). Internet and democratic public sphere development. *CM - Casopis Za Upravljanje Komuniciranjem*, 8(28), 89–110. <https://doi.org/10.5937/comman1328089V>
- Watts, D., & Strogatz, S. (1998). Collective dynamics of 'small-world' networks. *NATURE*, 393(6684), 440–442. <https://doi.org/10.1038/30918>
- White, D. R., & Harary, F. (2001). The Cohesiveness of Blocks In Social Networks: Node Connectivity and Conditional Density. *Sociological Methodology*, 31(1), 305–359. <https://doi.org/10.1111/0081-1750.00098>
- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. Preuzeto sa <http://ggplot2.org>
- World Bank. (2016). World Development Indicators| World DataBank. Preuzeto October 2, 2016, sa <http://databank.worldbank.org/data/reports.aspx?source=2&series=IT.NET.USER.P2&country=SRB#>
- Xu, K., Rooney, C., Passmore, P., Ham, D.-H., & Nguyen, P. H. (2012). A User Study on Curved Edges in Graph Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 18(12), 2449–2456. <https://doi.org/10.1109/TVCG.2012.189>
- Zappavigna, M. (2011). Ambient affiliation: A linguistic perspective on Twitter. *New Media & Society*, 13(5), 788–806. <https://doi.org/10.1177/1461444810385097>

7 PRILOZI

7.1 Prilog I

Političari uključeni u analizu, podaci o njihovim Tviter nalogima i stranačka pripadnost

Ime i prezime	Stranka	Tviter handle (@)	Datum otvaranja naloga
Saša Radulović	Dosta je bilo	SasaRadulovich	28.3.2012
Dušan Pavlović	Dosta je bilo	d_pavlovic	8.9.2012
Aleksandar Stevanović	Dosta je bilo	Stevanovich_A	24.12.2015
Ivan Ninić	Dosta je bilo	INinic	25.8.2011
[nalog stranke]	Dosta je bilo	DostaJeBilo	2.2.2014
Vojislav Šešelj	Srpska radikalna stranka	seselj_vojislav	22.10.2015
Vjerica Radeta	Srpska radikalna stranka	VjericaR	11.12.2012
Nemanja Šarović	Srpska radikalna stranka	NemanjaSarovic	21.10.2010
Zoran Krsić	Srpska radikalna stranka	ZoranKrasic	11.12.2014
[nalog stranke]	Srpska radikalna stranka	srpski_radikali	28.3.2009
Aleksandar Vučić	Srpska napredna stranka	avucic	16.8.2011
Nebojša Stefanović	Srpska napredna stranka	NesaStefanovic	22.8.2011
Zoran Babić	Srpska napredna stranka	ZBabic_	30.3.2015
Vladimir Đukanović	Srpska napredna stranka	Djuka_RSFC	12.9.2009
Slaviša Mićanović	Srpska napredna stranka	micanovic_s	29.7.2011
[nalog stranke]	Srpska napredna stranka	sns_srbija	31.1.2011
Bojan Pajtić	Demokratska stranka	PajticBojan	4.2.2015
Gordana Čomić	Demokratska stranka	gordanacom	25.5.2010
Aleksandra Jerkov	Demokratska stranka	AleksJerkov	13.6.2011
Dragan Šutanovac	Demokratska stranka	SutanovacDragan	13.2.2011
[nalog stranke]	Demokratska stranka	demokrate	31.7.2009
Zoran Živković	Nova stranka	ZoranDirektno	5.1.2010
Vladimir Pavićević	Nova stranka	vpavicevic	7.2.2010
[nalog stranke]	Nova stranka	NovaStranka	23.11.2012
Čedomir Jovanović	Liberalno demokratska partija	Ceda_Jovanovic	26.12.2008
Drago Kovačević	Liberalno demokratska partija	DragoKovacevic	7.11.2011
Kenan Hajdarević	Liberalno demokratska partija	KenanLDP	22.9.2011
[nalog stranke]	Liberalno demokratska partija	LDP	26.1.2009
Goran Radosavljević	Socijaldemokratska stranka	Radosav77	15.11.2011
Konstantin Samofalov	Socijaldemokratska stranka	KoleSamofalov	2.3.2011
[nalog stranke]	Socijaldemokratska stranka	socdemstranka	30.1.2014
Branko Ružić	Socijalistička partija Srbije	brankoruzicsps	4.3.2016
Nenad Borovčanin	Socijalistička partija Srbije	nborovcanin	3.7.2011
Dijana Vukomanić	Socijalistička partija Srbije	vukomand	11.01.2013
[nalog stranke]	Socijalistička partija Srbije	socijalisti	27.5.2010
Dragan Maršićanin	Demokratska stranka Srbije	marsicanin_d	14.10.2015
Milan Bojović	Demokratska stranka Srbije	mbojovic777	24.10.2014
Milan Lapčević	Demokratska stranka Srbije	dsscentar	9.6.2011
Boško Obradović	Srpski pokret Dveri	BoskoObradovic	17.12.2014
[nalog stranke]	Srpski pokret Dveri	SPDveri	9.12.2010

7.2 Prilog II

Nalozi političara prema konvencionalnim pokazateljima popularnosti na Tviteru

Ime i prezime	Stranka	Ukupan broj (jun 2016.)				Za vreme kampanje			
		Pratioci	Prijatelji	Tvitovi	<i>Favorites</i> <i>d</i>	<i>Favorites</i>	Retvitovi	Tvitovi	RT pros.
Saša Radulović	DJB	38491	275	16257	125	22034	5468	612	8,9
Dušan Pavlović	DJB	13994	564	3194	11806	8564	2868	728	3,9
Aleksandar Stevanović	DJB	2328	103	1448	89	4950	810	408	2,0
Ivan Ninić	DJB	4889	1846	10736	4224	10597	3239	1260	2,6
[nalog stranke]	DJB	15517	1266	7149	1880	27034	9781	803	12,2
Vojislav Šešelj	SRS	7163	1642	2663	398	29644	5193	523	9,9
Vjerica Radeta	SRS	6964	1337	14530	10728	12540	3145	619	5,1
Nemanja Šarović	SRS	4114	252	1301	190	1059	312	104	3,0
Zoran Krasić	SRS	3384	2882	10477	1606	14539	3426	1194	2,9
[nalog stranke]	SRS	7216	2431	5661	1975	477	132	14	9,4
Aleksandar Vučić	SNS	166447	706	1500	580	146415	137219	103	1332,2
Nebojša Stefanović	SNS	12229	1007	552	596	13285	10516	60	175,3
Zoran Babić	SNS	4561	1116	152	457	4681	3674	40	91,9
Vladimir Đukanović	SNS	3267	1286	7125	264	4462	1980	1116	1,8
Slaviša Mićanović	SNS	6263	618	4638	3042	4727	3598	41	87,8
[nalog stranke]	SNS	24248	1262	7921	2595	117032	98059	579	169,4
Bojan Pajić	DS	10673	3081	988	6680	7710	1816	95	19,1
Gordana Čomić	DS	20989	1967	38669	7533	13174	3199	1985	1,6
Aleksandra Jerkov	DS	21624	1499	1901	321	1103	417	22	19,0
Dragan Šutanovac	DS	44546	1294	28370	42108	14680	3481	342	10,2
[nalog stranke]	DS	23004	1438	13364	515	13806	4711	484	9,7
Zoran Živković	NS	25985	604	9290	263	6045	1129	276	4,1
Vladimir Pavićević	NS	19488	4816	3958	4873	7506	1059	135	7,8
[nalog stranke]	NS	6711	3306	3602	1424	3540	1011	180	5,6
Čedomir Jovanović	LDP	28108	431	1529	276	2254	419	62	6,8
Drago Kovačević	LDP	12339	8668	51945	31590	3816	353	757	0,5
Kenan Hajdarević	LDP	1876	732	1417	1432	17	2	6	0,3
[nalog stranke]	LDP	14249	2209	5123	512	3736	943	280	3,4
Goran Radosavljević	SDS	4107	923	27300	15162	4036	829	1341	0,6
Konstantin Samofalov	SDS	4320	953	10003	5879	1143	269	214	1,3
[nalog stranke]	SDS	3018	2094	1726	357	1966	1015	269	3,8
Branko Ružić	SPS	1784	87	151	900	4746	1093	119	9,2
Nenad Borovčanin	SPS	6988	2240	15744	10480	1388	252	277	0,9
Dijana Vukomanić	SPS	6936	261	4773	1944	1250	285	185	1,5
[nalog stranke]	SPS	10682	570	5105	1125	1913	615	260	2,4
Dragan Maršićanin	DSS	3024	1541	4249	1774	6180	1027	572	1,8
Milan Bojović	DSS	606	604	5879	3045	715	159	247	0,6
Milan Lapčević	DSS	5524	1709	3055	1255	5775	2238	312	7,2
Boško Obradović	Dveri	1877	32	205	12	562	180	11	16,4
[nalog stranke]	Dveri	7190	2900	9473	2145	3674	1483	552	2,7

Favorited – broj tvitova koje je korisnik označio kao omiljene;

Favorites (za vreme kampanje) – koliko je ljudi označilo tvit, koji je naznačeni korisnik tvitovao, kao omiljeni.

RT pros – prosečan broj retvita po tvitu

7.3 Prilog III

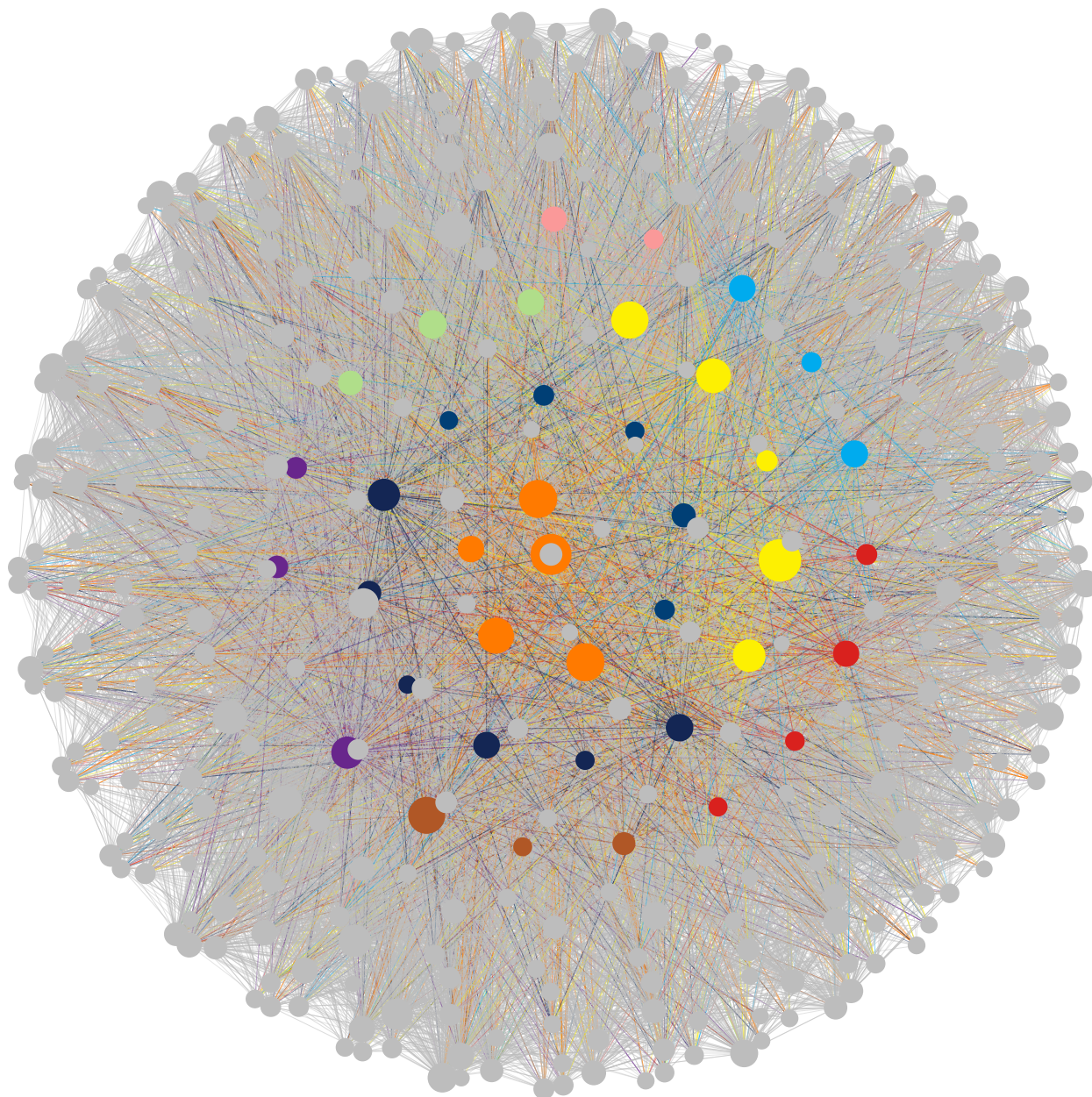
Mere centralnosti Tviter za 100 najcentralnijih naloga u proširenom uzorku, rangiranih prema P_2 odstojanju.

Tviter nalog	Centralnost						P_2 odstojanje
	intermedijarnosti	dolaznog stepena	odlaznog stepena	dolazne bliskosti	odlazne bliskosti	svojtvenog vektora	
SasaRadulovich	7287,87	218	107	0,0007	0,0012	0,93	28,30
DostaJeBilo	3437,76	215	73	0,0007	0,0011	1,00	22,71
gordanacom	4316,88	188	160	0,0006	0,0011	0,27	20,28
d_pavlovic	3247,17	182	107	0,0006	0,0011	0,26	17,44
dusanmasic	3694,98	190	87	0,0007	0,0011	0,12	16,73
INinic	1689,02	159	101	0,0006	0,0011	0,43	16,14
micin_taja	2984,92	144	73	0,0006	0,0011	0,05	14,61
Radosav77	1669,93	139	135	0,0006	0,0011	0,11	14,40
BiljanaLuki	2802,85	179	54	0,0007	0,0011	0,07	14,27
somboracmarko	2450,79	183	56	0,0007	0,0011	0,12	14,16
PajticBojan	2814,47	186	30	0,0007	0,0011	0,08	13,82
SutanovacDragan	1986,49	165	80	0,0006	0,0011	0,05	13,50
demokrate	1799,04	238	38	0,0007	0,0010	0,14	13,33
LicniStav	1746,84	119	65	0,0006	0,0011	0,11	12,97
Budza_Sremacka	2333,21	79	61	0,0007	0,0011	0,07	12,93
SashaPavlovic	1222,94	37	112	0,0006	0,0012	0,15	12,92
Djuka_RSFC	2482,34	81	60	0,0006	0,0010	0,05	12,67
PeckoPivo	1447,46	97	56	0,0006	0,0011	0,07	11,69
SPDveri	2000,85	78	46	0,0006	0,0011	0,04	11,65
Stevanovich_A	1210,82	85	54	0,0006	0,0010	0,15	11,59
AdvVlahovic	1083,23	43	69	0,0006	0,0011	0,10	11,38
izivkov	1198,63	116	55	0,0006	0,0011	0,06	11,37
dsscentar	1880,92	117	24	0,0007	0,0010	0,03	11,37
nebojsarako	1077,91	43	82	0,0006	0,0011	0,06	11,29
vpavicevic	1268,06	136	26	0,0007	0,0010	0,05	10,89
wujisicm	1209,87	126	37	0,0006	0,0011	0,03	10,88
milorad984	744,33	27	85	0,0006	0,0011	0,11	10,86
VjericaR	1724,92	74	33	0,0006	0,0011	0,01	10,76
nenadcicko	1130,66	53	66	0,0006	0,0010	0,05	10,75
biljana0001	1035,72	106	43	0,0006	0,0010	0,04	10,65
MarkoDimic2	771,11	21	68	0,0006	0,0011	0,12	10,60
jelenak021	765,10	43	66	0,0006	0,0011	0,08	10,55
LDP	1011,74	205	7	0,0007	0,0008	0,06	10,48
stanisasms	681,44	44	71	0,0006	0,0011	0,05	10,28
DaniloKoprivica	744,81	36	63	0,0006	0,0011	0,07	10,18
BogdanTatic	955,45	66	45	0,0006	0,0011	0,02	10,17
marsicanin_d	717,85	79	69	0,0006	0,0010	0,02	10,15
VladimirTodoric	718,78	90	50	0,0006	0,0010	0,04	10,13
Djukela_SNS	922,67	44	46	0,0006	0,0010	0,06	9,99
zivstepa	592,28	8	61	0,0006	0,0011	0,10	9,88
UdbiEtOrbi	696,06	36	52	0,0006	0,0011	0,02	9,81
nedeljnik	926,33	68	14	0,0007	0,0009	0,09	9,69
NovaStranka	913,91	90	24	0,0006	0,0010	0,04	9,69
RomicSandra	643,52	50	52	0,0006	0,0011	0,02	9,52
ZoranDirektno	576,87	108	30	0,0006	0,0009	0,03	9,43
cipolla12	454,94	19	68	0,0006	0,0011	0,02	9,42
Aledar989	306,02	67	58	0,0006	0,0010	0,07	9,38
trzisnoresenje	468,92	54	46	0,0006	0,0010	0,07	9,22
socijalisti	657,98	125	9	0,0007	0,0009	0,01	9,17
sns_srbija	753,03	147	4	0,0007	0,0008	0,03	9,11

Tviter nalog	Centralnost						P ₂ odstojanje
	intermedijarnosti	dolaznog stepena	odlaznog stepena	dolazne bliskosti	odlazne bliskosti	svojstvenog vektora	
bborovic3	587,91	49	46	0,0006	0,0010	0,02	9,06
DragoKovacevic	947,04	46	42	0,0006	0,0009	0,02	9,06
momcilocrvak	562,19	24	50	0,0006	0,0011	0,03	9,06
ljubicapuric	392,36	20	60	0,0006	0,0011	0,03	9,04
amir_bislimi	488,68	39	38	0,0006	0,0010	0,02	8,97
TomaMomirovic	245,46	54	46	0,0006	0,0010	0,05	8,74
PuricLjubica	290,52	11	62	0,0006	0,0011	0,02	8,69
draganms79	448,67	91	13	0,0007	0,0009	0,02	8,61
mrburns0102	380,68	28	43	0,0006	0,0010	0,03	8,58
Lazariques	218,22	38	46	0,0006	0,0010	0,05	8,52
ZoranKrasic	1241,19	51	14	0,0006	0,0008	0,00	8,47
tkojovic	311,94	29	28	0,0006	0,0010	0,04	8,35
holzwege11	419,75	11	46	0,0006	0,0011	0,01	8,23
dusanlj	167,91	17	38	0,0006	0,0010	0,05	8,22
avucic	353,00	211	1	0,0007	0,0004	0,07	8,11
mbojovic777	316,06	23	31	0,0006	0,0010	0,01	8,08
oknarbnb	466,34	7	36	0,0006	0,0010	0,02	8,03
kejsiberger	215,23	11	40	0,0006	0,0010	0,02	7,99
AleksJerkov	258,23	55	13	0,0006	0,0009	0,02	7,94
dukimadjar	117,45	32	38	0,0006	0,0010	0,03	7,93
Vukasin90	208,76	20	36	0,0006	0,0010	0,03	7,90
vukomand	350,74	46	22	0,0006	0,0009	0,01	7,88
socdemstranka	142,88	79	13	0,0007	0,0007	0,02	7,76
dedadjinaj	83,41	10	32	0,0006	0,0010	0,07	7,73
dolazi_cirkus	105,64	23	35	0,0006	0,0010	0,02	7,72
masa_meda76	43,24	11	31	0,0006	0,0010	0,06	7,64
Ceda_Jovanovic	411,63	73	4	0,0006	0,0008	0,02	7,64
nborovcanin	352,60	28	21	0,0006	0,0009	0,01	7,64
seselj_vojislav	531,12	52	5	0,0006	0,0009	0,01	7,62
sneza_kg	233,79	18	12	0,0006	0,0009	0,07	7,54
micanovic_s	303,25	26	18	0,0006	0,0010	0,01	7,43
plavozuto	154,81	19	21	0,0006	0,0010	0,01	7,35
Goran11000	118,09	39	14	0,0006	0,0009	0,03	7,34
KoleSamofalov	187,46	19	23	0,0006	0,0009	0,01	7,22
a_ono_medjutim	94,81	9	23	0,0006	0,0010	0,01	7,09
NemanjaSarovic	120,37	24	18	0,0006	0,0009	0,00	7,01
leteciGasa	32,22	7	17	0,0006	0,0009	0,04	6,88
banegrkovic	52,52	16	17	0,0006	0,0009	0,01	6,85
NesaStefanovic	0,00	108	0	0,0009	0,0000	0,02	6,81
brankoruzicsp	103,23	34	7	0,0006	0,0008	0,01	6,65
antarizana	6,81	9	8	0,0006	0,0008	0,05	6,44
BorisGolubovic	5,16	5	32	0,0005	0,0010	0,01	6,32
DraganVujevic	12,51	3	15	0,0005	0,0010	0,02	6,29
severnojuzni	5,47	9	12	0,0005	0,0009	0,03	6,19
srpski_radikali	24,40	28	6	0,0006	0,0007	0,00	6,13
jevticdalibor	1,57	5	7	0,0005	0,0008	0,01	6,00
BoskoObradovic	0,13	47	2	0,0006	0,0005	0,01	5,87
KenanLDP	12,18	3	3	0,0006	0,0006	0,00	5,43
ZBabic_	0,00	27	2	0,0006	0,0003	0,00	4,90
Stefanivanovic2	0,00	3	22	0,0002	0,0010	0,01	4,03

7.4 Prilog IV

Konvencionalni sociogram proširenog uzorka. Korišćen je sferičan raspored, veličina čvora zavisi od stepena boja veze zavisi od stranačke pripadnosti jednog od čvorova.



Potpisivanjem izjavljujem:

- da je rad isključivo rezultat mog sopstvenog istraživačkog rada
- da sam rad i mišljenja drugih autora koje sam koristio u ovom radu naznačio ili citirao u skladu sa Uputstvom
- da su svi radovi i mišljenja drugih autora navedeni u spisku literature/referenci koji su sastavni deo ovog rada i pisani u skladu sa Uputstvom
- da sam dobio sve dozvole za korišćenje autorskog dela koji se u potpunosti/celosti unose u predati rad i da sam to jasno naveo
- da sam svestan da je plagijat korišćenje tuđih radova u bilo kom obliku (kao citata, parafraza, slika, tabela, dijagrama, dizajna, planova, fotografija, filma, muzike, formula, veb sajtova, kompjuterskih programa i sl.) bez navođenja autora ili predstavljanje tuđih autorskih dela kao mojih, kažnjivo po zakonu (Zakon o autorskom i srodnim pravima, Službeni glasnik Republike Srbije, br. 104/2009, 99/2011, 119/2012), kao i drugih zakona i odgovarajućih akata Univerziteta u Beogradu
- da sam svestan da plagijat uključuje i predstavljanje, upotrebu i distribuiranje rada predavača ili drugih studenata kao sopstvenih
- da sam svestan posledica koje kod dokazanog plagijata mogu prouzrokovati na predati master rad i moj status
- da je elektronska verzija master rada identična štampanom primerku i pristajem na njegovo objavljivanje pod uslovima propisanim aktima Univerziteta

Marko Galjak