

UNIVERZITET U BEOGRADU
FIZIČKI FAKULTET



Master rad

**Obrasci mobilnosti u Gradu Beogradu:
prostorno-vremenska analiza strukture i
dinamike prije i tokom COVID-19 epidemije**

STUDENT

Nikola Stupar

Broj indeksa: 7052/2020

Smjer: Primenjena i kompjuterska fizika

MENTOR

dr Marija Mitrović Dankulov

Beograd, septembar 2021.

Zahvaljujem se svojoj mentorki dr Mariji Mitrović Dankulov na velikoj podršci i savjetima prilikom izrade ove teze.

Takođe se zahvaljujem svojoj porodici i prijateljima na nesebičnoj podršci koju su mi pružili.

Sadržaj

1	Uvod	3
2	Metodologija	5
2.1	Algoritmi mašinskog učenja	5
2.2	K-means klasterovanje	5
2.3	Teorija kompleksnih mreža	7
2.4	Kendalov koeficijent	9
2.5	Ansamblski K-means algoritam	10
3	Podaci o mobilnosti	12
4	Rezultati: razlike i sličnosti izmedju mobilnosti tokom krize i van krize	14
4.1	Raspodjele populacije korisnika	14
4.2	Ukupna mobilnost	16
4.3	Mobilnost unutar ćelija	19
4.3.1	Klasterovanje ćelija pomoću ansamblskog K-means algoritma	19
4.4	Mobilnost izmedju ćelija	23
4.4.1	Klasterovanje veza pomoću ansamblskog K-means algoritma	24
4.5	Poređenje slojeva višeslojne mreže prema rangiranju čvorova i veza	28
4.6	Broj kretanja u funkciji udaljenosti	30
5	Zaključak	33
6	Reference	35

1 Uvod

Fizika se pokazala vrlo korisnom u izučavanju strukture i dinamike višestrukih interagujućih sistema koji tradicionalno ne pripadaju njenom polju interesovanja, kao što su biološki [19] ili socioekonomski sistemi [9]. Glavna karakteristika svih ovih sistema je da se sastoje od velikog broja interagujućih jedinica i da ispoljavaju kolektivno ponašanje. Ovakve sisteme nazivamo kompleksnim sistemima. Fizika u kombinaciji sa teorijom kompleksnih sistema i metodama računarskih nauka omogućava da se kvantitativno opiše struktura interakcija u kompleksnim sistemima, da se prouči kako struktura utiče i određuje dinamiku i funkciju ovakvih sistema, kao i da se pronađu mehanizmi koji su u osnovi nastanka kolektivnog ponašanja u njima.

Istraživanja sprovedena u prethodne dvije decenije na različitim kompleksnim sistemima pokazala su da njihova dinamika i funkcija zavise od strukture mreže interakcija [4, 8]. Teorija kompleksnih mreža omogućava kvantitativni opis i izučavanje evolucije mreže interakcija u kompleksnim sistemima. Kompleksne mreže, kojima se opisuje struktura ovih interakcija, iako opisuju različite sisteme imaju neka svojstva koja su zajednička svim realnim kompleksnim sistemima. Ove mreže imaju široku, vrlo heterogenu, distribuciju vjerovatnoće stepena [4, 8], srednji put između svaka dva čvora koji raste logaritamski sa rastom jedinica u mreži [42], korelisane su [4], i često imaju mezoskopske nehomogenosti. Ova univerzalna svojstva ukazuju na slične mehanizme evolucije ovih kompleksnih sistema i omogućavaju da se okviri teorije kompleksnih mreža i fizike kompleksnih sistema primjene na izučavanje njihove strukture i dinamike.

Fizika socioekonomskih sistema se razvila u prethodne dvije decenije zahvaljujući prije svega velikoj količini podataka o ponašanju ljudi koji su postali dostupni [39]. Razvoj informaciono-komunikacionih tehnologija je omogućio sakupljanje velike količine različitih podataka o svakodnevnom ponašanju ljudi. Fizika, prije svega primjenom numeričkih metoda i u kombinaciji sa metodama računarskih nauka, je kroz analizu ovih podataka uspjela da izvede statističke zakonitosti ponašanja različitih socijalnih sistema. Socijalni sistemi ispoljavaju različite oblike kolektivnog ponašanja od dinamike mišljenja [39], preko kolektivnih emocija [30], do kolektivnih obrazaca mobilnosti [38]. Socijalni sistemi ispoljavaju karakteristične obrasce koji su stabilni u vremenu i prostoru [20], i nezavisni od istorijskih i kulturoloških okolnosti [10]. Ova stabilnost i univerzalnost obrazaca u socijalnim sistemima čini ideju o predvidljivosti ponašanja socijalnih sistema realnom, i daje alate za planiranje i konstruisanje realnih politika i instrumenata sa ciljem da se potpomogne razvoj boljeg i uspješnijeg društva. Razumijevanje mobilnosti ljudi u okviru jednog grada, države ili svijeta je od izuzetnog značaja. Informacija o mobilnosti ljudi u gradu je od ključnog značaja za urbanističko planiranje, planiranje transporta, predviđanje toka epidemija i procjenu doprinosa mobilnosti ljudi zagađenju vazduha [28, 24, 15, 2]. U posljednje vrijeme došlo je do velikog pomaka u istraživanju o mobilnosti ljudi u smislu razumijevanja, kvantitativnog opisa i modeliranja mobilnosti ljudi. Veliki pomak u ovim istraživanjima je najvećim dijelom posljedica pojavljivanja velikih količina podataka o mobilnosti ljudi [46].

Najveći broj istraživanja je zasnovan na proučavanju individualnih trajektorija i proučavanju zasnovanom na teoriji mreža [1]. Kod proučavanja individualnih trajektorija određuju se raspodjele pređene udaljenosti između dvije tačke Δr i vremena čekanja Δt između dva kretanja [37]. Kod ovih rezultata primjećuju se neke statističke regularnosti. Posmatranjem velikih skupova podataka o cirkulaciji novčanica od jednog dolara (uz pretpostavku da vjerodostojno predstavljaju kretanje ljudi koji rukuju njima) primjećeno je da pređena udaljenost Δr pojedinaca može biti opisana raspodjelom $P(\Delta r) \sim (\Delta r)^{(1+\beta)}$, gdje $\beta < 2$ [7]. Ovo je u skladu sa intuicijom: ljudi se najčešće kreću kratkim relacijama, ali povremeno otputuju na neku relativno udaljenu destinaciju.

U istraživanjima mobilnosti ljudi zasnovanim na teorijama mreža, pristup je korištenje mreža koje su smještene u stvarnom prostoru, odnosno čvorovi ovih mreža zauzimaju određen položaj u oblasti u kojoj se mobilnost posmatra. Definiše se mreža sa čvorovima smještenim u stvarnom prostoru tako da je veza između tih čvorova srazmjerna fluksu ljudi između njih [44]. Zbog vremenske rezolucije podataka iz kojih je mreža generisana, topologija mreže se obično mijenja sa vremenom, pa se ove mreže posmatraju kao dinamičke mreže. Dinamičko posmatranje mreža podrazumijeva posmatranje parametara mreže koji se mijenjaju sa vremenom.

Ljudska mobilnost je posmatrana na različitim prostornim skalama. Najveće prostorne skale na kojim je posmatrana mobilnost su kontinenti. To istraživanje je urađeno analizom podataka o vazdušnom saobraćaju [22]. Najmanje skale na kojim je praćena mobilnost su stambene zgrade [26]. Praćenje mobilnosti na tako malim skalama je moguće praćenjem povezivanja bežičnih uređaja svakog pojedinca.

Najveći dio istraživanja obrazaca mobilnosti proučava njihovu stabilnosti i dinamiku tokom normalnih perioda. Rezultati pokazuju da su obrasci mobilnosti stabilni u prostoru i vremenu, što omogućava razvoj različitih modela kojima se modelira mobilnost na različitim prostornim i vremenskim skalama.

Velike krize, kao što su požari, zemljotresi i epidemije neminovno dovode do strukturnih promjena u ponašanju socijalnih sistema [39, 25]. Naše poznavanje kako i u kojoj mjeri različite krize utiču na promjenu mobilnosti ljudi je još uvijek ograničeno. Istraživanja vezana za zemljotrese su otkrila da trajektorije ljudi postaju predvidljivije poslije zemljotresa [25]. Za razliku od zemljotresa, koji izazivaju krize koje traju relativno kratko, krize poput epidemije COVID-19 izazivaju dugotrajne promjene u ponašanju ljudi. Kako ove promjene utiču na promjene obrazaca mobilnosti, i da li su ove promjene univerzalne u odnosu na različite države i kulture je još uvijek otvoreno pitanje. Cilj ove teze je da odgovori na pitanje da li su se i na koji način promijenili obrasci mobilnosti u Gradu Beogradu usljed epidemije COVID-19. U tu svrhu biće iskorištene metode numeričke fizike, teorije kompleksnih mreža i mašinskog učenja za analizu podataka o mobilnosti korisnika Facebooka dobijenih putem Facebook-ovog portala Data for good [48]. Facebook Data for good takođe obezbjeđuje informacije o prostornoj raspodjeli korisnika čiji se obrasci kretanja posmatraju. Biće analizirani podaci o populaciji i mobilnosti korisnika na teritoriji Grada Beograda prije i tokom krize COVID-19. Korištenjem metoda mašinskog učenja pokazuje se sličnost u obrascima kretanja korisnika u raznim oblastima Grada Beograda. Takođe će biti predstavljena dinamička analiza realnih kompleksnih mreža generisanih iz pomenutih podataka o mobilnosti korisnika. Ova teza je organizovana na sljedeći način. U poglavlju 2 su objašnjene metode koje će biti korištene u analizi obrazaca mobilnosti. Poglavlje 3 daje detaljni opis podataka dobijenih sa Facebook-ovog portala Data for good [48]. Rezultati analize su predstavljeni u poglavlju 4, dok je u poglavlju 5 dat zaključak ove teze.

2 Metodologija

2.1 Algoritmi mašinskog učenja

Mašinsko učenje je jedna implementacija vještačke inteligencije. Suština mašinskog učenja je ekstraktovanje znanja iz podataka. Primjene mašinskog učenja koje se sreću svakodnevno su algoritmi za preporuku sadržaja na video platformama, algoritmi za preporuku kupovine određenih proizvoda, algoritmi za preporuku drugih korisnika na društvenim mrežama, prepoznavanje korisnika na fotografijama na društvenim mrežama [31]. Pored komercijalne primjene mašinskog učenja, ono se intenzivno koristi i u istraživanju. Česta primjena je u astronomiji, fizici elementarnih čestica, teoriji kompleksnih mreža i medicini [21, 35, 13, 17]. U zavisnosti od prirode problema koji se rješava primjenom mašinskog učenja, algoritmi mašinskog učenja se dijele na mašinsko učenje sa nadzorom i mašinsko učenje bez nadzora [31].

Kod mašinskog učenja sa nadzorom algoritam mašinskog učenja se obučava pomoću parova ulaza i izlaza. Za dati ulaz unaprijed je određen izlaz. Ovo je određeno ljudskom inteligencijom. Kada je algoritam obučen, za novi nepoznat ulaz on daje izlaz sa određenom tačnošću bez ikakve pomoći ljudske inteligencije. Primjer ovoga je klasifikacija neželjene elektronske pošte [14]. Korisnik prepoznaje neželjenu poštu i označava je kao takvu. U ovom slučaju elektronska pošta je ulaz, a oznaka *željena pošta* ili *neželjena pošta* je izlaz. Algoritam u budućnosti prepoznaje neželjenu poštu sa određenom tačnošću i odlaže je u poseban folder. Sa povećanjem oznaka od strane korisnika tačnost algoritma se povećava.

Kod mašinskog učenja bez nadzora, algoritam ima pristup samo ulaznim podacima, izlazni podaci su nepoznati algoritmu. Ovakvi algoritmi se najčešće koriste u klasterovanju neoznačenih ulaznih podataka [43]. Pridruživanje izlaza ulaznim podacima se vrši bez učešća ljudske inteligencije [31].

Vrlo je važna reprezentacija podataka na kojima se primjenjuju algoritmi mašinskog učenja. Ovi podaci su najčešće prikazani u tabelama u kojima redovi predstavljaju *uzorke*, a kolone predstavljaju *osobine*.

2.2 K-means klasterovanje

Klasterovanje je jedan od najčešće korištenih algoritama mašinskog učenja bez nadzora [43]. Uzorci su predstavljeni kao tačke u N -dimenzionom prostoru gdje je N broj osobina tih uzoraka. Klasterovanje je postupak u kojem se ovi uzorci postavljaju u grupe na osnovu njihovih osobina. Ove grupe se nazivaju klasteri. Kao rezultat ovog algoritma očekuje se da uzorci unutar jednog klastera budu što sličniji drugim uzorcima unutar tog klastera, a uzorci u različitim klasterima da budu što različitiji.

K-means klasterovanje je jedan od najčešće korištenih algoritama za klasterovanje [18]. Ovaj algoritam pokušava odrediti centre klastera koji predstavljaju određene oblasti u N -dimenzionom prostoru u kojem su podaci. Kada se algoritmu za K-means klasterovanje zada broj klastera K , klasterovanje se odvija u tri koraka:

1. Prvi korak je postavljanje K nasumičnih tačaka za centre klastera. Ovim se definišu inicijalni klasteri.
2. Drugi korak je pridruživanje tačaka inicijalnim klasterima na osnovu euklidske udaljenosti od centra inicijalnih klastera.
3. Treći korak je redefinisavanje centra klastera na osnovu srednje vrijednosti klastera.

Drugi i treći korak se ponavljaju sve dok se pripadnost tačaka klasterima ne prestane mijenjati. Tada je algoritam K-means klasterovanja završen.

Podaci mogu da se jave u obliku vremenskih serija. Vremenska serija je skup vrijednosti neke veličine u raznim vremenskim trenucima. Udaljenost između ovih trenutaka je konstantan vremenski interval. Broj trenutaka za koje postoji vrijednost neke veličine je dužina vremenske serije. Pojedine vremenske serije se mogu gledati kao uzorci, a vrijednosti veličina koje ove vremenske serije predstavljaju u raznim trenucima se mogu gledati kao osobine. Dužina vremenske serije je broj osobina N . U ovoj reprezentaciji se vremenske serije mogu posmatrati kao tačke u N -dimenzionom prostoru i na njima se mogu primjeniti algoritmi mašinskog učenja. K-means klasterovanje se može primjeniti na vremenske serije da bi se slične vremenske serije pridružile u klustere. Za ove potrebe je razvijen paket *tslearn* u programskom jeziku *Python* [41].

Ovaj algoritam zahtijeva n vremenskih serija istih dužina N bez nedostajućih podataka kao ulaz:

$$\begin{aligned}
 &x_1(t_1), x_1(t_2), \dots, x_1(t_N) \\
 &x_2(t_1), x_2(t_2), \dots, x_2(t_N) \\
 &\vdots \\
 &x_n(t_1), x_n(t_2), \dots, x_n(t_N)
 \end{aligned} \tag{2.1}$$

Algoritam posmatra vremenske serije kao tačke u N -dimenzionom prostoru:

$$\begin{aligned}
 &x_1(t_1), x_1(t_2), \dots, x_1(t_N) = X_1 \\
 &x_2(t_1), x_2(t_2), \dots, x_2(t_N) = X_2 \\
 &\vdots \\
 &x_n(t_1), x_n(t_2), \dots, x_n(t_N) = X_n
 \end{aligned} \tag{2.2}$$

Vremenske serije se predstavljaju kao tačke u N -dimenzionom prostoru X_l , $l = 1, 2, \dots, n$. Potrebno je zadati broj klastera K i početne uslove na osnovu kojih će algoritam postaviti centre klastera $c_0, c_1, \dots, c_{K-1} = c_j$, $j = 0, 1, \dots, K - 1$ koji predstavljaju inicijalne klustere. Zatim se određuje kvadrat euklidske udaljenosti svake od tačaka u N -dimenzionom prostoru X_l , $l = 1, 2, \dots, n$ od svakog od definisanih centara klastera c_j , $j = 0, 1, \dots, K - 1$:

$$d_{l,j} = \|X_l - c_j\|^2 \tag{2.3}$$

Svaka od tačaka X_l u N -dimenzionom prostoru je pridružena centru klastera od kojeg je varijansa (kvadrat euklidske udaljenosti) minimalna. Tačka pridružena klasteru sa centrom klastera c_j , $j = 0, 1, \dots, K - 1$ se označava sa $X_l^{(j)}$.

Nakon toga, centar klastera c_j , $j = 0, 1, \dots, K - 1$ je redefinisano kao srednja vrijednost klastera $0, 1, \dots, K - 1$. Ponovo se određuje pripadnost tačaka klasterima na osnovu minimuma varijanse. Ovaj proces se nastavlja dok se pripadnost tačaka klasterima u N -dimenzionom prostoru (vremenskih serija dužine N) ne prestane mijenjati. Navedeni postupak se može posmatrati kao problem optimizacije [27]. Opisani proces dodjeljivanja tačaka

klasterima u N -dimenzionom prostoru se može predstaviti minimizacijom funkcije:

$$J = \sum_{j=1}^{K-1} \sum_{i=1}^N \left\| X_i^{(j)} - c_j \right\|^2 \quad (2.4)$$

Kada algoritam završi, svakoj od ulaznih vremenskih serija pridružuje oznaku klastera $0, 1, \dots, K-1$. Za razne početne uslove se dobijaju razne podjele vremenskih serija u klasterne. Ovo je posljedica postavljanja različitih tačaka za centre inicijalnih klastera. Stabilni centri klastera koji se dobijaju po završetku algoritma K -means klasterovanja su tačke u N -dimenzionom prostoru. One se transformišu iz tačaka u N -dimenzionom prostoru u vremenske serije dužine N . Ovo je *centroid* klastera i predstavlja srednju vrijednost veličine predstavljene vremenskim serijama po trenucima unutar vremenskih serija. Centroidi dobijenih K klastera su oblika:

$$\begin{aligned} & c_0(t_1), c_0(t_2), \dots, c_0(t_N) \\ & c_1(t_1), c_2(t_2), \dots, c_2(t_N) \\ & \vdots \\ & c_{K-1}(t_1), c_{K-1}(t_2), \dots, c_{K-1}(t_N) \end{aligned} \quad (2.5)$$

2.3 Teorija kompleksnih mreža

Problemi koji se danas svrstavaju u teoriju kompleksnih mreža su se javili još u 18. vijeku u oblasti diskretne matematike poznatoj pod imenom teorija grafova. Graf je skup čvorova koji su povezani skupom veza. Ako postoji relacija između dva čvora, ta dva čvora su onda povezana vezom. U slučaju da koristimo graf da predstavimo neki realan sistem, radi se o kompleksnoj mreži. Grafovi su matematička reprezentacija kompleksnih mreža. Sa razvojem teorije kompleksnih mreža povećao se i broj sistema koji je opisan pomoću mreže. Neke od mreža koje su posmatrane su transportne mreže, telekomunikacione mreže, Internet, mreže saradnje glumaca, mreže koautorstva u nauci i mreže citiranja u nauci [45, 40, 11, 32, 33, 12]. Neki sistemi u biologiji i medicini se takođe modeliraju pomoću mreža, npr. neuronske mreže, genetske mreže, metaboličke mreže i mreže proteina [4]. Prilikom posmatranja mreže, najčešće od samog interesa nije sama mreža (čvorovi i veze), već dinamički sistem koji ta mreža predstavlja.

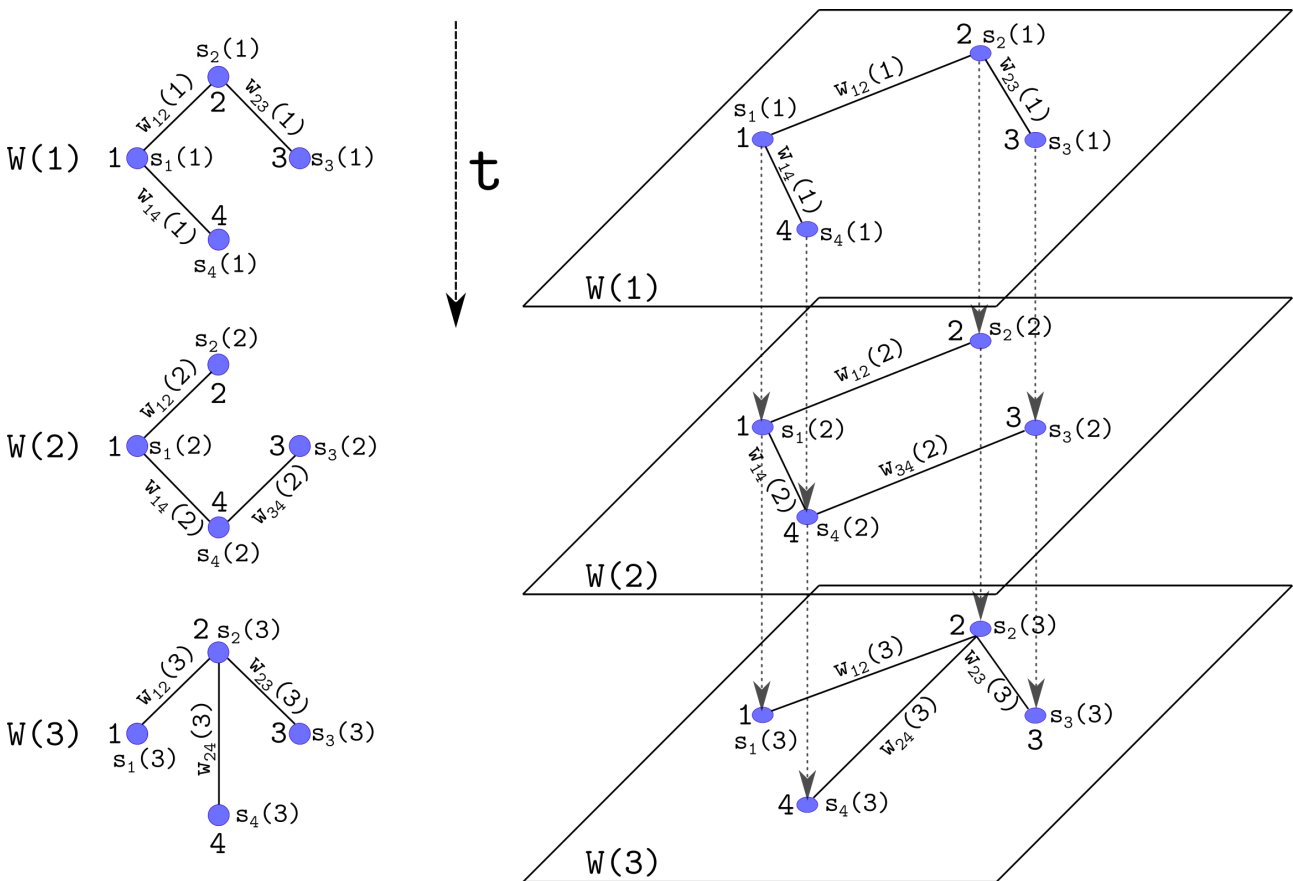
Veličina mreže je definisana brojem čvorova unutar mreže N . Minimalan broj veza između čvorova u mreži je 0, a maksimalan je $N(N-1)/2$. U zavisnosti od toga da li su veze između čvorova usmjerene ili ne, možemo imati usmjerenu ili neusmjerenu mrežu. Mreža od N čvorova se može predstaviti matricom povezanosti A veličine $N \times N$ čiji unosi zavise od toga da li postoji veza između dva čvora. Binarna mreža je mreža predstavljena matricom povezanosti čiji unosi mogu biti ili $a_{ij} = 0$ ili $a_{ij} = 1$. Ako postoji veza između čvorova i i j ($i, j = 1, \dots, N$) unos u matrici je $a_{ij} = 1$. U slučaju da čvorovi i i j nisu povezani unos u matrici je $a_{ij} = 0$. Otežinjene mreže se koriste u slučaju da je potrebno opisati težinu veze između dva čvora. Otežinjene mreže se mogu predstaviti matricom težina gdje unos u matrici može biti bilo koji realan broj $w_{ij} \in \mathbb{R}$. U slučaju da je matrica povezanosti koja opisuje mrežu simetrična matrica, mreža koju ta matrica opisuje je neusmjerena.

U proučavanju kompleksnih mreža, uvedene su veličine za opisivanje strukturnih osobina mreža. Stepennost k_i čvora i je broj veza koje taj čvor ima sa drugim čvorovima [8]. Stepennost čvora i se može odrediti iz unosa unutar matrice povezanosti te mreže $k_i = \sum_{j \in N} a_{ij}$. Osnovni topološki opis mreže je distribucija vjerovatnoće stepena čvora $P(k)$ koji predstavlja vjerovatnoću da nasumično izabrani čvor ima stepen jednak k . Za N čvorova u

mreži i N_k čvorova sa stepenom k , može se procijeniti $P(k)$ pomoću jednačine $P(k) = N_k/N$.

Analogon *stepenu čvora* u slučaju da posmatramo otežinjenu mrežu je veličina *jačina čvora* s . U neusmjerenjenu mrežu, ova veličina se definiše za čvor i kao suma težina svih veza povezanih na taj čvor $s_i = \sum_{j \in N} w_{ij}$, gdje je w_{ij} težina veze između čvorova i i j koja odgovara unosu u matricu povezanosti koja opisuje tu otežinjenu mrežu.

Prilikom modeliranja dinamičkih sistema kompleksnim mrežama javljaju se slučajevi kada veze između čvorova nisu konstantno aktivne. U ovom slučaju koriste se vremenske mreže [23]. Vremenske mreže mogu biti i binarne i otežinjene. Vremenska struktura mreža utiče na dinamiku interagujućih sistema unutar mreže koliko i sama topologija mreže. Vremenske mreže su dobar alat za opisivanje dodatne informacije o vezama između čvorova. Nekad je potrebno opisati i dodatnu informaciju o čvorovima unutar mreže (npr. vrijednost neke veličine vezane za čvor u raznim vremenskim trenucima). Jedna od veličina vezana za čvor je rangiranje čvora u mreži u datom trenutku. U slučaju da je potrebno posmatrati rangiranje čvorova (ili veza) u različitim trenucima, korisno je vremensku mrežu mapirati u višeslojnu mrežu (slika 1). Vremenska mreža je specijalni slučaj višeslojne mreže kada svaki od slojeva predstavlja jedan vremenski trenutak. Prednost predstavljanja dinamičkog sistema pomoću višeslojne mreže je što omogućava čuvanje informacije o vrijednosti neke veličine vezane za čvor.



Slika 1: Lijevo na slici je prikazana otežinjena vremenska mreža u tri vremenska trenutka. Smjer vremena je prikazan strelicom. Desno na slici je ista vremenska mreža mapirana u višeslojnu mrežu. Broj čvorova u mreži je $N = 4$. Označene su jačine svakog čvora $s_i(t)$, $i = 1, 2, 3, 4$, $t = 1, 2, 3$ kao i težine svake veze $w_{ij}(t)$, $i, j = 1, 2, 3, 4$, $t = 1, 2, 3$. U ovakvoj reprezentaciji mreže lako je pratiti dati čvor ili vezu u različitim slojevima mreže (različitim vremenskim trenucima). [5]

2.4 Kendalov koeficijent

Najjednostavniji način da se opiše neka mreža je da se odredi distribucija stepena čvorova u toj mreži. U slučaju višeslojnih mreža može se posmatrati kako se raspodjela stepena čvorova mreže mijenja od jednog sloja do idućeg sloja (u slučaju vremenskih mreža mapiranih u višeslojne mreže to bi odgovaralo posmatranju raspodjele stepena od jednog trenutka do idućeg trenutka). Takođe je korisno posmatrati kako se neka veličina koja opisuje čvor ili vezu u mreži mijenja od jednog do drugog sloja (od jednog do drugog trenutka kod vremenskih mreža mapiranih u višeslojne mreže). Da bi se preciznije odredila korelacija između rangova čvorova ili veza u dva sloja višeslojne mreže, koristi se Kendalov koeficijent korelacije rangova τ [3].

Kendalov koeficijent korelacije ranga predstavlja stepen sličnosti između dva skupa rangova koji su dodijeljeni istom skupu objekata [36]. Rangiranje objekata je izvršeno po nekoj veličini kojom se opisuje svaki od objekata. Kendalov koeficijent korelacije ranga prebrojava broj inverzija parova objekata potrebnih da se jedan skup rangova transformiše u drugi skup rangova. Neka se posmatra skup \mathcal{A} koji sadrži N objekata:

$$\mathcal{A} = \{a_1, a_2, \dots, a_N\} \quad (2.6)$$

Elementi u skupu \mathcal{A} se mogu poredati po rangu koji im je dodijeljen u odnosu na neku veličinu vezanu za svaki objekat. Rangiranje objekata u skupu \mathcal{A} se može predstaviti uređenim skupom \mathcal{O} . Uređen skup \mathcal{O} od N objekata se može razložiti na $N(N-1)/2$ uređenih parova. Uređeni parovi za dati uređen skup objekata \mathcal{O} se označava sa \mathcal{P} .

Kendalov koeficijent upoređuje dva skupa uređenih parova \mathcal{P}_1 i \mathcal{P}_2 (koji predstavljaju dva načina rangiranja objekata u polaznom skupu \mathcal{O}_1 i \mathcal{O}_2) prebrojavanjem različitih parova u dva uređena skupa. Dobijeni broj predstavlja razliku između skupova koja se naziva *simetrična udaljenost razlike*. Simetrična udaljenost razlike između dva skupa uređenih parova \mathcal{P}_1 i \mathcal{P}_2 se označava sa $d_{\Delta}(\mathcal{P}_1, \mathcal{P}_2)$. Da bi se odredio izraz za Kendalov koeficijent korelacije ranga τ , potrebno je normalizovati simetričnu udaljenost razlike $d_{\Delta}(\mathcal{P}_1, \mathcal{P}_2)$ tako da τ uzima vrijednost između -1 i $+1$. Normalizovanje se vrši tako što je poznato da vrijednost $\tau = -1$ treba da odgovara najvećoj mogućoj vrijednosti simetrične udaljenosti razlike. Ovo je slučaj kada je rangiranje predstavljeno skupom \mathcal{O}_1 jednako obrnutom redoslijedu rangiranja koje predstavlja skup \mathcal{O}_2 . Vrijednost $\tau = 1$ treba da odgovara najmanjoj mogućoj vrijednosti simetrične udaljenosti razlike $d_{\Delta}(\mathcal{P}_1, \mathcal{P}_2) = 0$. Ovo je slučaj kada je rangiranje predstavljeno skupom \mathcal{O}_1 isto kao i rangiranje predstavljeno skupom \mathcal{O}_2 . Potrebno je još uzeti u obzir da je maksimalan broj parova za koje se mogu razlikovati dva skupa sa $N(N-1)/2$ elementata jednak $N(N-1)$. Formula po kojoj se računa Kendalov koeficijent korelacije ranga je:

$$\tau = \frac{2 \cdot [d_{\Delta}(\mathcal{P}_1, \mathcal{P}_2)]}{N(N-1)} \quad (2.7)$$

Kako je definicija Kendalovog koeficijenta korelacije ranga zasnovana na prebrojavanju različitih parova koji se pojavljuju u dva uređena skupa, izračunata vrijednost se može tumačiti kao vjerovatnoća. Ako se par objekata uzima nasumično iz jednog uređenog skupa, τ predstavlja razliku vjerovatnoće da ova dva nasumično izabrana objekta imaju isti raspored u drugom uređenom skupu - *P(isti)* i vjerovatnoće da ova dva objekta imaju drugačiji raspored u drugom uređenom skupu - *P(različit)*:

$$\tau = P(\text{isti}) - P(\text{različit}) \quad (2.8)$$

Za svaki od slojeva višeslojne mreže odredi se niz vrijednosti X u kojem je neka veličina vezana za čvor ili vezu u sloju višeslojne mreže. U slučaju da se posmatraju čvorovi ta veličina će biti stepen čvora za binarne (neotežinjene) mreže k_i ili jačina čvora za otežinjene mreže s_i . U slučaju da se posmatraju veze, posmatrana veličina je težina veze w_{ij} čija vrijednost može biti bilo koji realan broj kod otežinjenih mreža i 0 ili 1 kod neotežinjenih (binarnih) mreža. Takođe se određuje niz vrijednosti Y u kojem je ta ista veličina koja opisuje svaki čvora ili vezu u sljedećem sloju višeslojne mreže. U slučaju vremenske mreže mapirane u višeslojnu X predstavlja niz vrijednosti veličine vezane za čvor ili vezu u mreži u jednom trenutku, a Y predstavlja niz vrijednosti veličine vezane za čvor ili vezu u mreži u idućem trenutku. Algoritam koji računa Kendalov koeficijent korelacije rangova uzima vrijednosti X i Y kao ulaz. Ove veličine su rangirane i određuje se Kendalov koeficijent korelacije ranga. Kendalov koeficijent uzima vrijednost $\tau_k(X, Y) = 1$ ako su rangovi dva niza vrijednosti identični, vrijednost $\tau_k(X, Y) = -1$ ako su rangovi niza vrijednosti X jednaki obrnutnom redosljedu rangova niza vrijednosti Y i vrijednost $\tau_k(X, Y) = 0$ u slučaju da rangovi niza vrijednosti X i rangovi niza vrijednosti Y ne zavise jedan od drugog. Ovo nam dozvoljava da pratimo kako se veličina koja opisuje čvorove ili veze u mreži mijenja od jednog do drugog sloja višeslojne mreže (od jednog do drugog trenutka za vremenske mreže mapirane u višeslojne mreže).

2.5 Ansamblski K-means algoritam

Prilikom klasterovanja tačaka u N -dimenzionom prostoru pomoću algoritma K-means klasterovanja dobijaju se različiti rezultati za različite početne uslove. Tačke u N -dimenzionom prostoru su dodijeljene različitim klasterima sa završetkom algoritma K-means klasterovanja za različite početne uslove. Ovo je posljedica definisanja različitih centara klastera za inicijalne klasterne na različite načine sa svakom novom inicijalizacijom algoritma K-means klasterovanja. Rješavanje problema nestabilnosti prilikom K-means klasterovanja je slično rješavanju poznatog problema nestabilnosti rješavanja prilikom detekcije zajednica u mrežama primjenom Leven algoritma [16].

Leven algoritam se koristi za detekciju zajednica u mrežama [6]. Ovaj algoritam je zasnovan na maksimizaciji modularnosti. Algoritam se primjenjuje na matrici težina W koja opisuje otežinjenu mrežu sa N čvorova. Prvo je svaki od čvorova pridružen jednoj zajednici. Zatim se posmatra kako se modularnost mijenja kako se jedan od čvorova u mreži prebaci iz trenutne zajednice u zajednicu njemu susjednog čvora. Čvor ostaje u toj zajednici ako je promjena vrijednosti modularnosti pozitivna. Ovaj proces preraspodjele čvora iz trenutne zajednice u zajednicu njemu susjednog čvora se ponavlja za sve čvorove dok se ne postigne lokalni maksimum vrijednosti modularnosti. U idućem koraku se pravi nova mreža čiji su čvorovi zajednice određene u prethodnom koraku. Postupak sa dodjeljivanjem čvorova u zajednice njemu susjednih čvorova se ponavlja dok se ne postigne maksimalna vrijednost modularnosti čime se algoritam završava. Svaki od čvorova iz polazne mreže će biti dodijeljen jednoj od zajednica.

Standardni Leven algoritam ima problem nestabilnosti rješavanja. Za različite početne uslove (polazne dodjele čvorova u zajednice) dobijaju se različita rješavanja. Ansamblski Leven algoritam se obično koristi za rješavanje problema nestabilnosti rješavanja prilikom primjene standardnog Leven algoritma za različite početne uslove [16]. Algoritam se odvija u nekoliko koraka:

1. Prvi korak je pokretanje standardnog Leven algoritma na mreži nekoliko puta za različite početne uslove.
2. Zatim se pravi nova mreža u kojoj su čvorovi povezani samo ako su se u prethodnom koraku pojavljivali

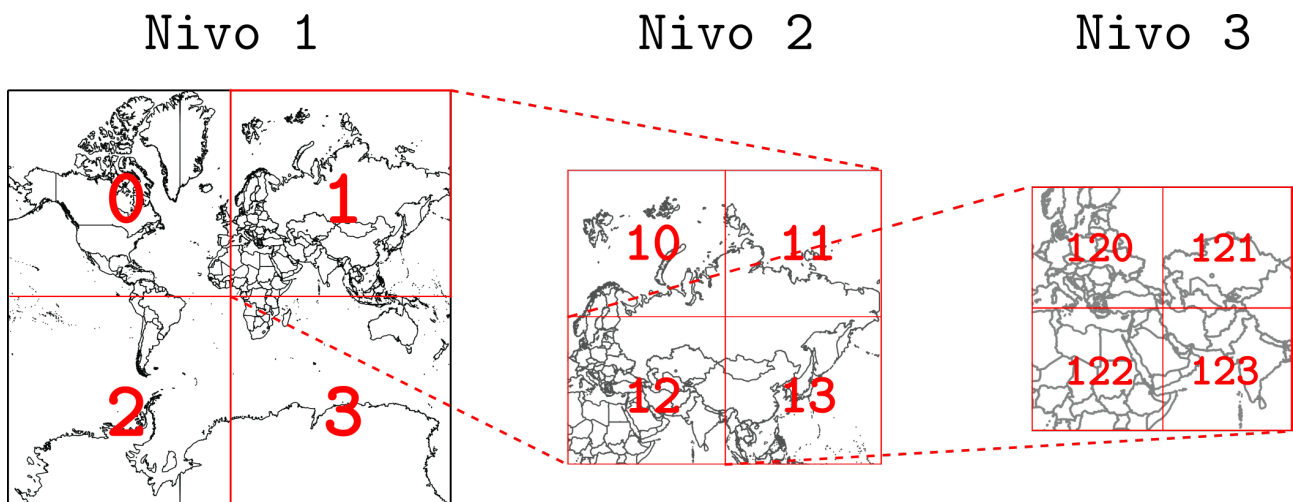
u istoj zajednici iznad određenog praga (obično u 90% slučajeva od ukupnog broja inicijalizacija).

3. Iz napravljene mreže se prepoznaju izolovane zajednice čvorova koje predstavljaju rješenja ansamblskog Leven algoritma.

Problem nestabilnosti rješenja kod pokretanja algoritma K-means klasterovanja za različite početne uslove se može riješiti na sličan način. Algoritam K-means klasterovanja se koristi za klasterovanje N tačaka, svaka označena sa $i = 1, 2, \dots, N$. Algoritam se pokreće n_p puta za n_p različitih početnih vrijednosti. Za n_p inicijalizacija, tačka i i tačka j će se pojaviti u istom klasteru $(n_c)_{ij}$ puta. Definiše se *indeks zajedničkog pojavljivanja* s_{ij} tačke i i tačke j koji je jednak količniku broja ponavljanja tačke i i tačke j u istom klasteru $(n_c)_{ij}$ i broja inicijalizacija n_p : $s_{ij} = \frac{(n_c)_{ij}}{n_p}$. Ovaj indeks može imati vrijednosti između 0 i 1. Može se definisati matrica sličnosti S veličine $N \times N$ čiji su unosi s_{ij} . Ako se pomenuta matrica posmatra kao matrica povezanosti mreže, iz nje se može generisati mreža. Ako se odbace svi unosi u matrici koji imaju vrijednost ispod nekog praga, u matrici ostaju nepreklapajući skupovi koji predstavljaju zajednice u mreži. Ovo bi bilo analogno drugom koraku u ansamblskom Leven algoritmu. Ove zajednice predstavljaju stabilna rješenja algoritma K-means klasterovanja na početnom skupu od N tačaka. Opisani algoritam za dobijanje stabilnih rješenja prilikom K-means klasterovanja ćemo nazvati ansamblski K-means algoritam.

3 Podaci o mobilnosti

Facebook Data for good dijeli korisne anonimizovane podatke o Facebook korisnicima sa istraživačima i humanitarcima [48]. Ovi podaci su korisni za planiranje mjera u slučaju kriza kao što su poplave, zemljotresi i epidemije. Samo korisnici sa uključenim lokacijskim uslugama doprinose setovima podataka na Facebook Data for good. Facebook Data for good obezbjeđuje podatke o populaciji korisnika i podatke o mobilnosti korisnika. Da bi se sačuvala privatnost korisnika, ne mogu se podijeliti podaci visoke rezolucije o populaciji korisnika i precizne trajektorije korisnika. Ovaj problem je riješen predstavljanjem podataka pomoću Sistema ćelija Bing Mapa [47]. Ovaj sistem podrazumijeva niz mreža sačinjenih od ćelija kvadratnog oblika sa jedinstvenim oznakama da bi se označila neka oblast na mapi (slika 2).



Slika 2: Ćelije sa oznakama u prva tri nivoa Sistema ćelija Bing Mapa. Sa povećanjem nivoa, smanjuje se površina ćelije koja predstavlja neku oblast. Najčešće se koriste mreže na nivou 13 ili nivou 14. Površina ćelije na nivou 13 je $3.466 \text{ km} \cdot 3.466 \text{ km} = 12.013 \text{ km}^2$, dok je površina ćelije na nivou 14 jednaka $1.733 \text{ km} \cdot 1.733 \text{ km} = 3.003 \text{ km}^2$. Ćelija na $n - tom$ nivou je označena jedinstvenim stringom dužine n . Da bi se označila ćelija na datom nivou, počinje se sa mapom Zemlje u Merkatorovoj projekciji. Ova mapa se prilagođava da bude kvadratnog oblika. Zatim se mapa dijeli u 4 jednake ćelije kvadratnog oblika. Svako od ćelija se pridruži jedna cifra od 0 do 3 na način prikazan na slici. Ove cifre predstavljaju oznaku ćelija na nivou 1 (dužina stringova je 1). U zavisnosti od oblasti koja treba da bude predstavljena ćelijom, bira se idući kvadrat. Ovim se prelazi na nivo 2. Oblast obuhvaćena kvadratom na nivou 2 se dijeli na četiri jednaka kvadrata. Oznake ćelija na nivou 2 se određuju tako što se na oznaku izabrane ćelije na nivou 1 (u ovom slučaju 1) pridružuje cifra od 0 do 3 na način prikazan na slici. Ovaj proces se nastavlja do nivoa 13 ili nivoa 14 koji se najčešće koriste u Sistemu ćelija Bing Mapa.

Pomoću Sistema ćelija Bing Mapa, populacija korisnika se može predstaviti kao broj korisnika u oblasti koja je obuhvaćena ćelijom, a kretanje korisnika se može predstaviti kao broj kretanja unutar ćelije ili kao fluks korisnika između ćelija.

Da bi se generisali podaci o populaciji korisnika i kretanju korisnika za određenu oblast, administrator na Facebook Data for good bira oblast pravougaonog oblika i od tog dana se generišu podaci. U zavisnosti od površine zahvaćene oblasti, generisani podaci su na nivou 13 ili nivou 14. Generisani podaci su bazirani na prebrojavanju događaja koji se odvijaju u vremenskim intervalima od 8 sati [29]. Brojanje događaja se odvija u tri intervala u danu: interval od 00:00h do 08:00h, interval od 08:00h do 16:00h i interval od 16:00h do 24:00h. Podaci trebaju biti anonimizovani, ali i dalje trebaju da budu korisni istraživačima i humanitarcima. Da bi se sačuvala privatnost korisnika, koristi se nekoliko mehanizama prilikom generisanja podataka. Prvi mehanizam anonimizacije podataka je dodavanje male količine slučajnog šuma u podatke. Ovo onemogućava precizno određivanje broja

korisnika i njihovo kretanje u slabije naseljenim oblastima. Sljedeći mehanizam anonimizacije podataka je prostorno zaglađivanje podataka. Ovo podrazumijeva usrednjavanje broja korisnika (ili broja kretanja) u datoj ćeliji sa istom veličinom u susjednim ćelijama. Koristi se usrednjavanje ponderisano rastojanjem koje daje veću težinu bliskim lokacijama i manju težinu udaljenijim lokacijama. Posljednji korak u anonimizaciji podataka je odbacivanje prebrojanog broja korisnika (ili kretanja) gdje je ta veličina manja od 10.

Prilikom generisanja podataka za vrijeme krize, za svaku od aktivnih ćelija na datom nivou i u datom trenutku se računa predviđeni broj korisnika u slučaju da su normalne okolnosti i predviđeni broj kretanja u ćeliji i između para ćelija u slučaju da su normalne okolnosti. Normalne okolnosti podrazumijevaju period prije početka krize. Sa predviđenim veličinama za slučaj normalnih okolnosti se mogu porediti veličine za vrijeme krize. Predviđeni podaci za slučaj da su normalne okolnosti se generišu na osnovu podataka sakupljenih u periodu prije početka krize. Prilikom generisanja ovih podataka uzimaju se u obzir pravilni dnevni i sedmični obrasci kretanja u periodu prije početka krize. Podaci za normalne okolnosti se određuju za datu ćeliju u datom vremenskom intervalu (uzima se u obzir da dati vremenski interval pripada jednom od tri intervala u danu i da taj vremenski interval pripada jednom danu u sedmici). Ograničavanjem ekstremnih vrijednosti i usrednjavanjem vrijednosti veličine koja se određuje u normalnim okolnostima za taj isti vremenski interval u danu i za taj isti dan u sedmici, dobija se vrijednost veličine u normalnim okolnostima za tu ćeliju ili paru ćelija ako se posmatra kretanje između ćelija. Dobijena vrijednost predstavlja ponašanje u datom trenutku i datoj ćeliji u slučaju da kriza nije nastupila.

Populacija korisnika u krizi je opisana agregiranim brojem korisnika u svakoj od ćelija u periodima od 8 sati. Pored svake vrijednosti broja korisnika za vrijeme krize nalazi se i predviđeni broj korisnika na toj ćeliji kakav bi bio da nije nastupila kriza. U slučaju da se korisnik pojavljivao u više različitih ćelija u periodu od 8 sati za vrijeme krize, biće prebrojan u ćeliji u kojoj je proveo najviše vremena. Svi opisani mehanizmi zaštite privatnosti korisnika se primjenjuju prilikom generisanja ovih podataka.

Mobilnost korisnika je opisana podacima o broju kretanja korisnika unutar ćelije i kretanjem korisnika između parova ćelija u vremenskim intervalima od 8 sati. U podacima je za dati interval prikazana **polazna ćelija, krajnja ćelija, broj kretanja u krizi i predviđeni broj kretanja u slučaju da nije nastupila kriza**. U slučaju da je polazna ćelija ista kao i krajnja ćelija, radi se o kretanju unutar ćelije. Kada su polazna i krajnja ćelija različite, radi se o kretanju između ćelija.

Posmatrani su podaci o populaciji korisnika i kretanju tih korisnika u Gradu Beogradu za period između 13. maja 2020. i 6. marta 2021. sa sajta Facebook Data for good. Dužina navedenog perioda je 298 dana. Kriza zbog koje je započeto sakupljanje podataka je epidemija COVID-19.

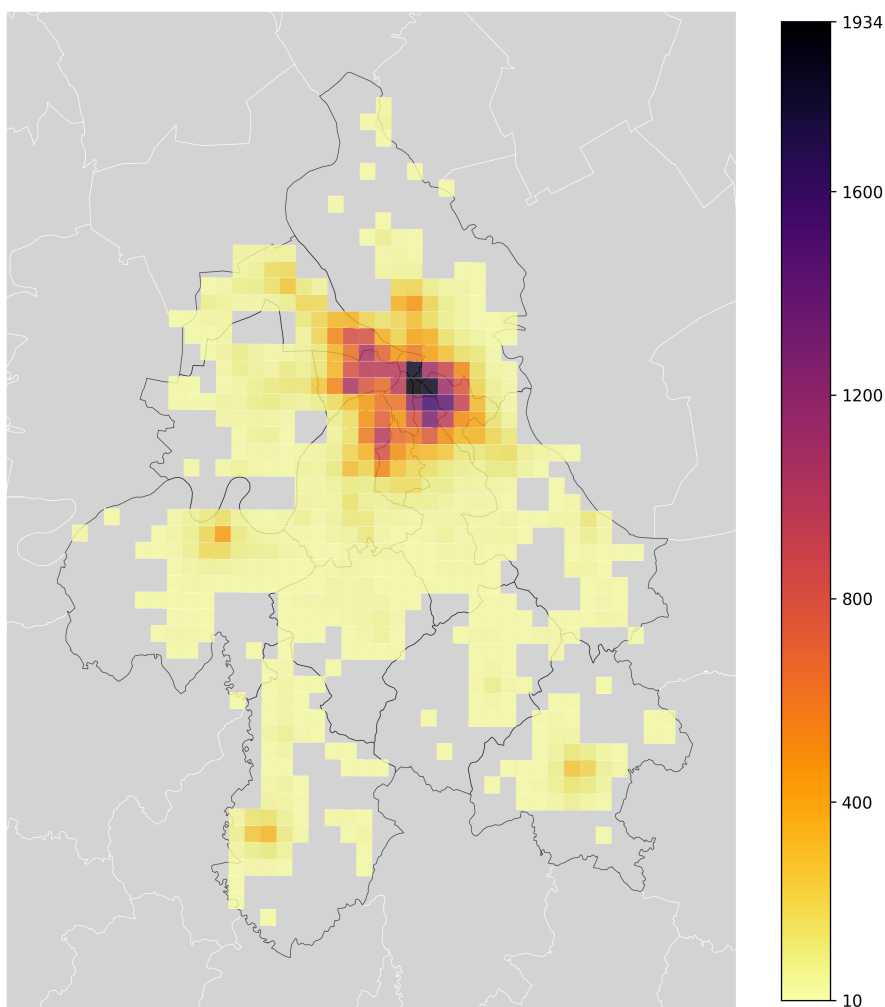
Podaci o populaciji korisnika sa uključenim lokacijskim uslugama na teritoriji Grada Beograda su agregirani na ćelije po Sistemu ćelija Bing Mapa na nivo 14. Na teritoriji Grada Beograda postoje ukupno 603 aktivne ćelije (ćelije koje su bile aktivne bar u jednom trenutku u posmatranom periodu). Podaci o kretanju korisnika sa uključenim lokacijskim uslugama na teritoriji Grada Beograda su agregirani na ćelije po Sistemu ćelija Bing Mapa na nivo 13. Na teritoriji Grada Beograda postoji ukupno 201 aktivna ćelija (ćelije unutar kojih je bilo kretanja i ćelije koje pripadaju parovima ćelija između kojih je bilo kretanja).

Kod podataka o populaciji kao i kod podataka o kretanju korisnika manji dio površine Grada Beograda nije prekriven ćelijama zato što ti dijelovi nisu naseljeni i zbog pomenutih mehanizama zaštite privatnosti korisnika u slabije naseljenim dijelovima.

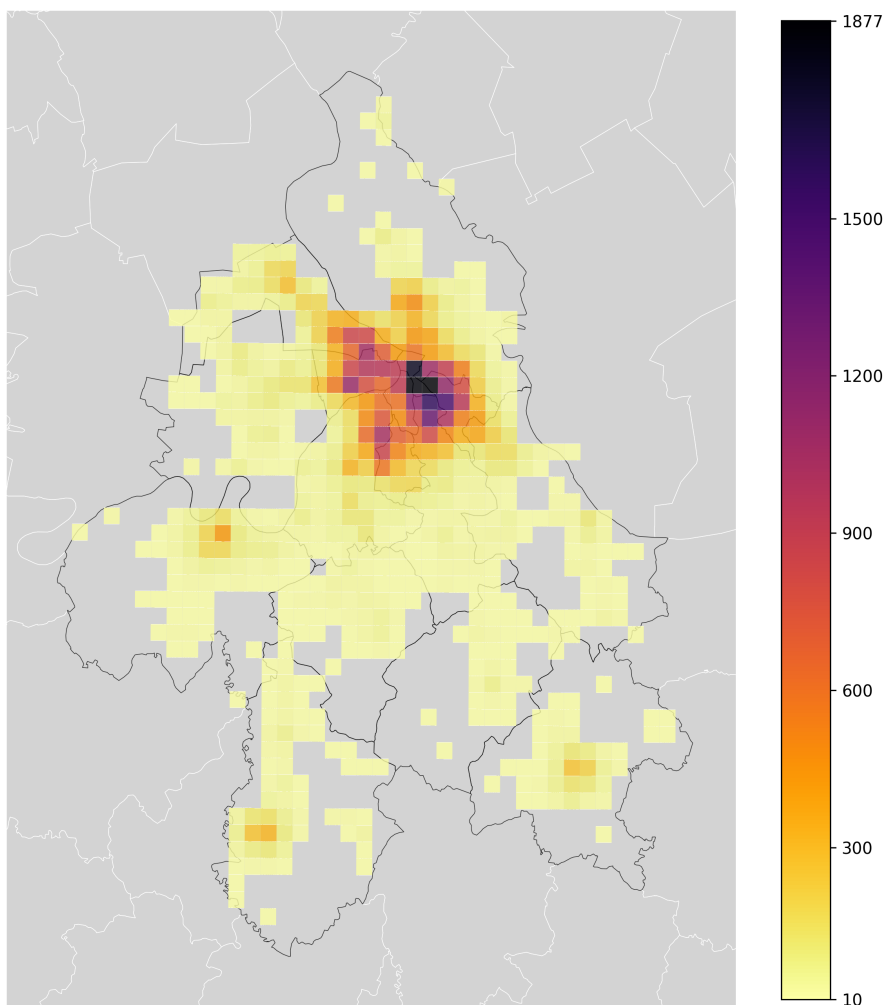
4 Rezultati: razlike i sličnosti izmedju mobilnosti tokom krize i van krize

4.1 Raspodjele populacije korisnika

Kako su podaci o populaciji korisnika i podaci o kretanju korisnika agregirani iz istih podataka o Facebook korisnicima sa uključenim lokacijskim uslugama, korisno je pogledati kakva je prostorna raspodjela ovih korisnika. Za navedeni period postoje podaci o populaciji korisnika agregirani na nivo 14 po Sistemu ćelija Bing Mapa. Površina ćelije na nivou 14 jednaka $1.733 \text{ km} \cdot 1.733 \text{ km} = 3.003 \text{ km}^2$. Za svaku od 603 aktivne ćelije postoje vrijednosti o broju korisnika za vrijeme krize i predviđeni broj korisnika za slučaj da nema krize u raznim vremenskim intervalima. Za svaku od aktivnih ćelija određena je srednja vrijednost broja korisnika za vrijeme krize (epidemija COVID-19) i srednja vrijednost predviđenog broja korisnika u slučaju da nema krize u navedenom vremenskom periodu. Za oba slučaja su iscrtane toplotne mape.



Slika 3: Toplotna mapa raspodjele Facebook korisnika sa uključenim lokacijskim uslugama za vrijeme krize (COVID-19 epidemije) u navedenom vremenskom periodu. Čelije koje predstavljaju naseljenije oblasti su obojene tamnijom bojom. Crna linija predstavlja granice gradskih opština Grada Beograda.

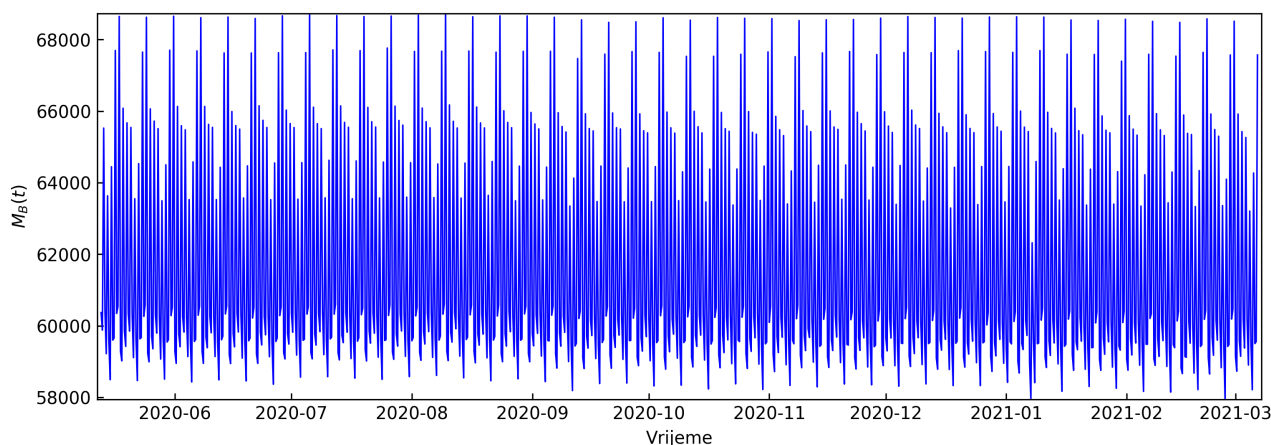


Slika 4: Toplotna mapa predviđene raspodjele Facebook korisnika za slučaj da nema krize (COVID-19 epidemije). Čelije koje predstavljaju naseljenije oblasti su obojene tamnijom bojom. Crna linija predstavlja granice gradskih opština Grada Beograda.

Objekti toplote mape (slika 3 i slika 4) imaju isti broj ćelija. Ovo je očekivano jer su podaci o populaciji i o mobilnosti korisnika za slučaj da nije kriza generisani na osnovu podataka dobijenih za vrijeme krize. Iz prikazanih toplotnih mapa raspodjele Facebook korisnika sa uključenim lokacijskim uslugama za vrijeme epidemije COVID-19 u navedenom vremenskom periodu (slika 3) i prije početka epidemije COVID-19 (slika 4) ne vidi se znatna promjena raspodjele korisnika. Korisnici su skoncentrisani u centralnom dijelu Grada Beograda i u gradskim opštinama Zemun, Obrenovac, Lazarevac i Mladenovac. Čelije sa najvećim brojem korisnika se nalaze na teritorijama opština Stari Grad, Vračar i Zvezdara. Sa legendi toplotnih mapa na slici 3 i slici 4 primjećuje se da se povećao broj korisnika u oblastima predstavljenim ćelijama za vrijeme krize u odnosu na period prije krize, što ukazuje na činjenicu da su korisnici počeli češće da koriste socijalne mreže tokom krize.

4.2 Ukupna mobilnost

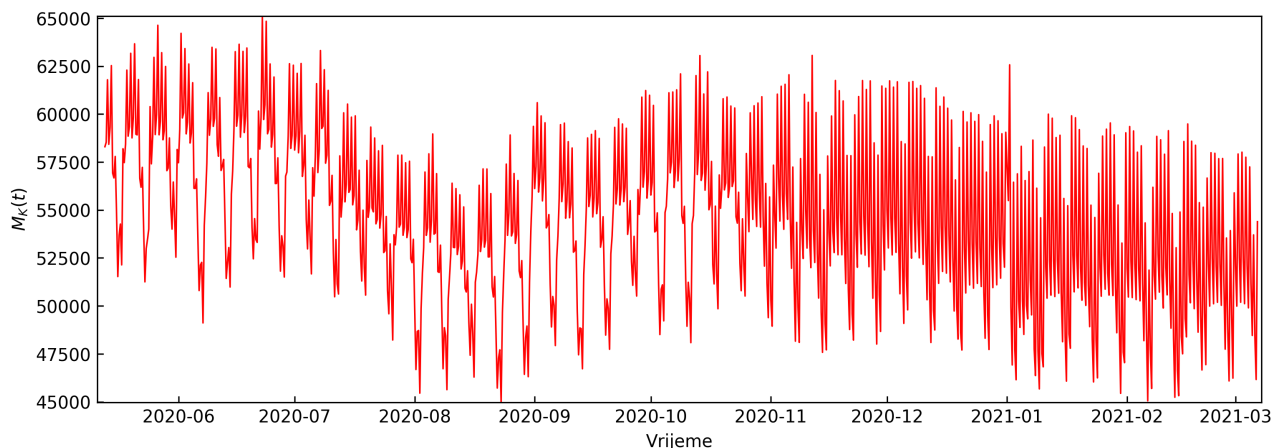
Da bi se posmatralo kako se mijenjaju obrasci kretanja u Gradu Beogradu usljed epidemije COVID-19, analiziran je ukupni broj kretanja korisnika za navedeni period od 298 dana. Posebno je posmatran broj kretanja korisnika za vrijeme krize i predviđeni broj kretanja u slučaju da nema krize. Za navedeni period postoje podaci o mobilnosti korisnika agregirani na nivo 13 po Sistemu ćelija Bing Mapa. Površina jedne ćelije na nivou 13 je $3.466 \text{ km} \cdot 3.466 \text{ km} = 12.013 \text{ km}^2$. Kada se posmatraju podaci o kretanju korisnika, na teritoriji Grada Beograda postoji ukupno 201 aktivna ćelija (ćelije unutar kojih je bilo kretanja i ćelije koje pripadaju parovima ćelija između kojih je bilo kretanja). U ovom poglavlju se posmatraju sva kretanja korisnika bez obzira da li su to kretanja unutar ćelije ili kretanja između ćelija. Podaci su agregirani na periode od 8 sati čime se dobija ukupan broj kretanja (unutar ćelija i između ćelija) za navedeni period za vrijeme krize $M_K(t)$. Ukupni broj kretanja u datom periodu za vrijeme krize je izračunat kao suma svih kretanja koja su se desila u tom periodu za vrijeme krize. Takođe je posmatran i predviđeni broj kretanja u slučaju da nema krize. Za svaki broj kretanja korisnika za vrijeme krize postoji predviđeni broj kretanja korisnika u slučaju da nema krize. Predviđeni broj kretanja korisnika u slučaju da nema krize zavisi od perioda u danu u kojem se desilo kretanje za vrijeme krize i dana u sedmici u kojem se desilo kretanje za vrijeme krize. Predviđeni broj kretanja u slučaju da nema krize je takođe agregiran na periode od 8 sati, čime se dobija ukupan broj kretanja (unutar ćelija i između ćelija) za navedeni period u slučaju da nije kriza $M_B(t)$. Ukupni broj kretanja u datom periodu u slučaju da nema krize je izračunat kao suma svih predviđenih brojeva kretanja koja su se desila u tom periodu u slučaju da nema krize. Kako je dužina vremenskog perioda u kojem su sakupljani podaci 298 dana i u jednom danu postoje tri intervala na koje se podaci agregiraju, postoji ukupno $298 \cdot 3 = 894$ vrijednosti za ukupni broj kretanja za vrijeme krize $M_K(t)$ i predviđeni broj kretanja za slučaj da nema krize $M_B(t)$.



Slika 5: Predviđeni ukupan broj kretanja za svaki od intervala od 8 sati u slučaju da nije kriza. Obrasci kretanja u slučaju da nema krize su stabilni tokom vremena.

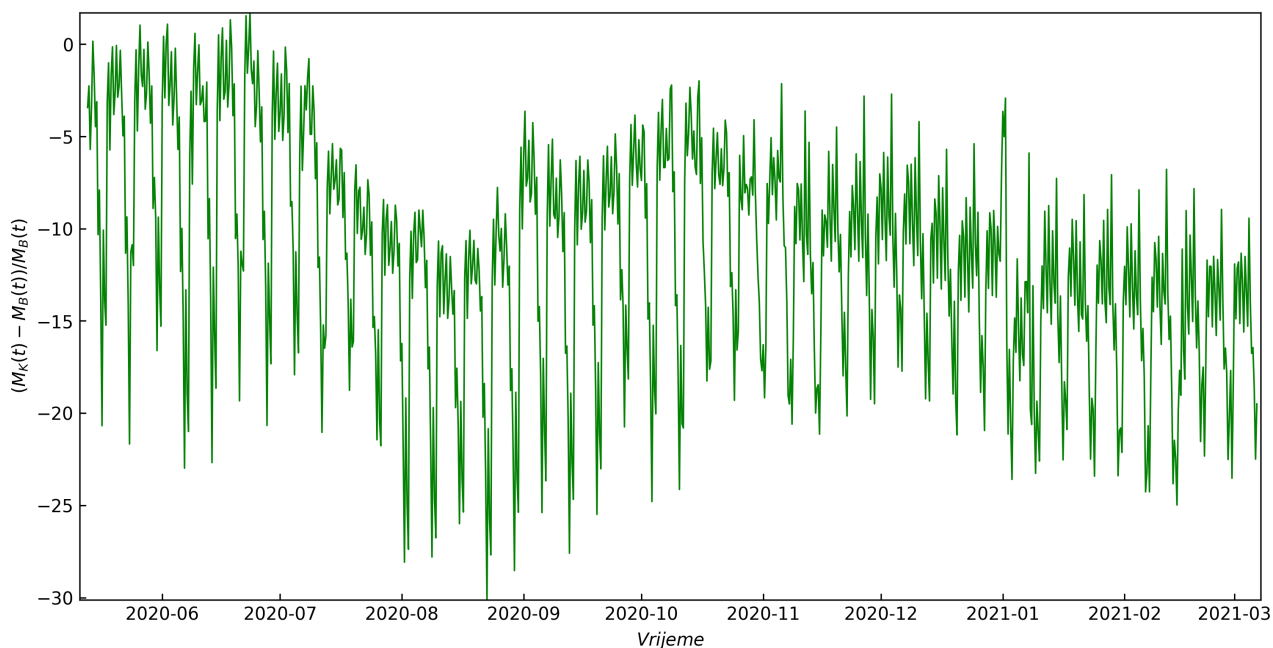
Na osnovu slike 5 može se primjetiti jasna sedmodnevna periodičnost ukupnog broja kretanja korisnika u slučaju da nema krize. Primjećuju se fluktuacije u ukupnom broju kretanja korisnika, što je i za očekivati. Ove fluktuacije su posljedica poslovnih aktivnosti građana Grada Beograda. Maksimalne vrijednosti na slici 5 odgovaraju ukupnom broju kretanja koji se desio tokom subote i nedjelje. Karakteristične maksimalne vrijednosti odgovaraju vremenskim intervalima od 16:00h do 24:00h nedjeljom. Karakteristične minimalne vrijednosti odgovaraju vremenskim intervalima od 00:00h do 08:00h ponedjeljkom. Radnim danima se primjećuje približno

konstantan ukupan broj kretanja za period ponedjeljak-četvrtak, dok se petkom primjećuje pad ukupnog broja putovanja.



Slika 6: Ukupan broj kretanja za svaki od intervala od 8 sati za vrijeme krize. Ukupan broj kretanja se mijenja sa vremenom.

Na osnovu slike 6 primjećuje se približna sedmodnevna periodičnost ukupnog broja kretanja korisnika za vrijeme krize. Primjećuju se fluktuacije u ukupnom broju kretanja korisnika, što je u skladu sa očekivanjima. Za razliku od perioda prije krize, fluktuacije broja kretanja korisnika za vrijeme krize se povećavaju sa vremenom. Na slici 6 ne možemo primjetiti karakteristične maksimalne vrijednosti broja kretanja koje odgovaraju pojedinim intervalima kao što se to jasno vidi u periodu prije krize. Umjesto toga, za vrijeme krize primjećujemo maksimalnu mobilnost za period nedjelja-četvrtak. I ovdje se vide karakteristične minimalne vrijednosti koje odgovaraju periodu petak-subota. U ovom slučaju ne može se odrediti specifičan interval za koji se unutar svake sedmice vidi minimalna vrijednost broja kretanja tokom sedmice.



Slika 7: Relativna promjena $M_R(t)$ (izražena u procentima) ukupnog broja kretanja $M_K(t)$ za vrijeme krize u odnosu na predviđeni ukupni broj kretanja u slučaju da nema krize $M_B(t)$.

Korisno je posmatrati relativnu promjenu ukupnog broja kretanja $M_K(t)$ za vrijeme krize u odnosu na predviđeni ukupni broj kretanja u slučaju da nema krize $M_B(t)$ (slika 7). Za poznate vrijednosti broja kretanja prije krize i predviđenog broja kretanja u slučaju da nema krize, može se odrediti relativna promjena broja kretanja za vrijeme krize u odnosu na predviđeni broj kretanja u slučaju da nema krize po formuli:

$$M_R(t) = \frac{M_K(t) - M_B(t)}{M_B(t)} \quad (4.1)$$

Sa grafika relativne promjene ukupnog broja kretanja (slika 7) vidi se da je kretanje za vrijeme krize u određenim vremenskim intervalima na početku sakupljanja podataka bilo približno isto kao i u slučaju da nema krize. Ovo je period neposredno po ukidanju vanrednog stanja. Pad u mobilnosti poslije ovog perioda se poklapa sa rastom broja zaraženih. Pik drugog talasa koji se desio u periodu jul-avgust, odgovara najvećem relativnom padu mobilnosti. Vidimo da se sličan, ali manji, pad vidi u periodu novembar 2020. - mart 2021., periodu u kojem su se desili 3. i 4. pik epidemije. Maksimalne vrijednosti relativne promjene broja kretanja pripadaju intervalima ponedjeljak-petak. Primjećuju se takođe karakteristične minimamalne vrijednosti koje odgovaraju intervalima u periodu subota-nedjelja.

Ukupna mobilnost se može podijeliti na mobilnost unutar ćelija i mobilnost između ćelija. Ovo dozvoljava posebnu analizu kretanja unutar ćelija i kretanja između ćelija i njihov doprinos ukupnom kretanju.

4.3 Mobilnost unutar ćelija

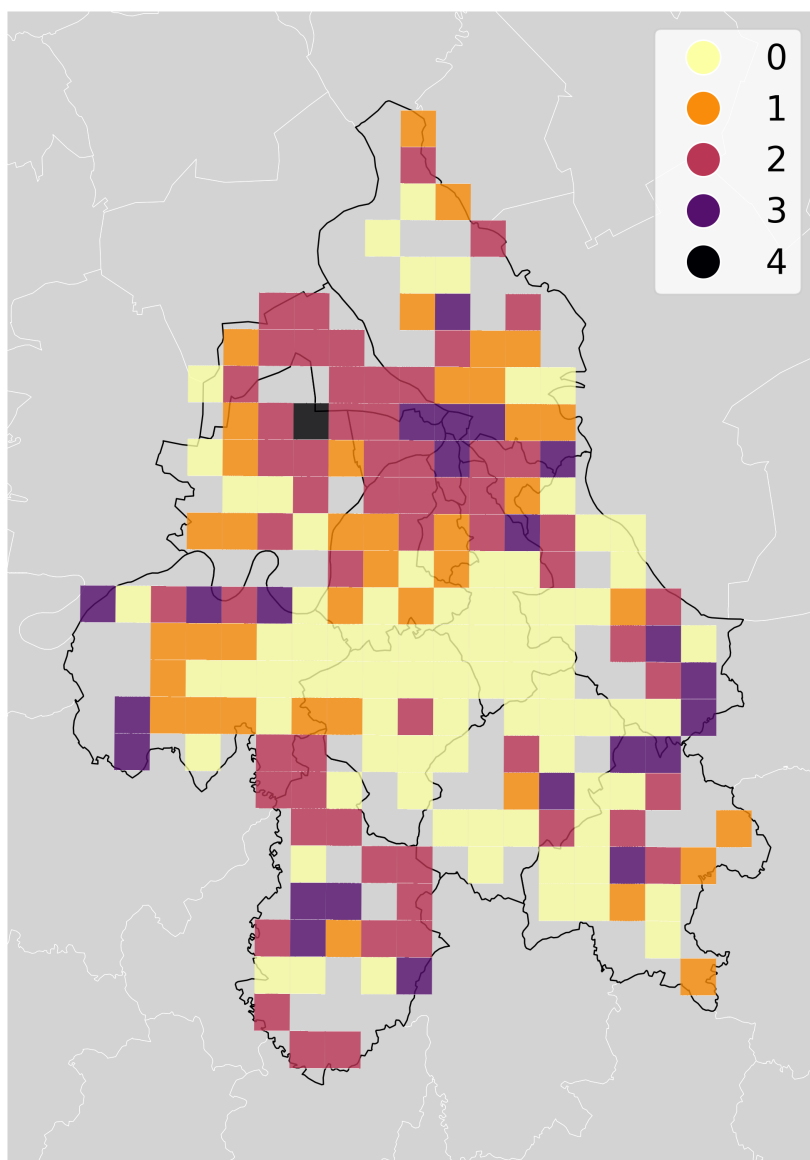
Mobilnost unutar ćelija, kao jedna od dvije komponente posmatranih obrazaca mobilnosti, predstavlja kretanje Facebook korisnika unutar ćelija na teritoriji Grada Beograda. Iz podataka o mobilnosti ekstrahovani su podaci kod kojih je polazna ćelija ista kao i krajnja ćelija. Dobijeni podaci predstavljaju unutrašnje kretanje korisnika na teritoriji Grada Beograda. Za svaku od 201 ćelije na teritoriji Grada Beograda računate su relativne promjene broja kretanja unutar ćelije za svaki od intervala od 8 sati u navedenom vremenskom periodu od 298 dana. U svakom danu postoje tri vremenska intervala od 8 sati, pa je relativna promjena broja kretanja za svaku od ćelija predstavljena kao vremenski niz dužine $298 \cdot 3 = 894$. Vrijednosti relativne promjene broja kretanja unutar ćelije koje nedostaju su izračunate interpolacijom u odnosu na susjedne tačke unutar vremenskih serija. Dobijena je 201 vremenska serija za svaku od ćelija unutar kojih je relativna promjena broja kretanja unutar ćelije. Na dobijene vremenske serije je primjenjen K-means algoritam klasterovanja za razne početne uslove. Za razne početne uslove algoritma K-means klasterovanja dobijana su različita rješenja. Da bi se riješio problem nestabilnosti rješenja koji se javlja kod primjene algoritma K-means klasterovanja, vremenske serije su takođe podijeljene u klustere primjenom ansamblskog K-means algoritma da bi se dobilo stabilno rješenje. Rješenja podrazumijevaju vremenske serije relativne promjene broja kretanja grupisane u klustere. Kako je svaka vremenska serija relativne promjene broja kretanja predstavljena ćelijom, rezultat se prikazuje vizuelno grupisanjem ćelija koje pripadaju istom klasteru na mapi.

4.3.1 Klasterovanje ćelija pomoću ansamblskog K-means algoritma

Za razne vrijednosti početnih uslova K-means algoritma klasterovanja ćelije su raspoređivane na različite načine u klustere. Zavisnost rješenja od proizvoljno izabranih početnih uslova je nepovoljna. Iako se dobijaju različita rješenja za različite početne uslove, ipak su se primjećivale neke pravilnosti u rješenjima. Za razne vrijednosti početnih uslova, od 5 dobijenih klastera, jedna ćelija se uvijek javlja u klasteru sama. Na slici 8 ova ćelija je sama u klasteru 4. Moguće objašnjenje za ovo je što se u oblasti koja je predstavljena ovom ćelijom nalazi Beogradski Aerodrom koji se može smatrati kao tačka iz koje izvire novi korisnici bez informacija o njihovom prijašnjem kretanju. Druga pravilnost koja je primjećena je da se najčešće javlja još jedan klaster koji obuhvata centralni dio Grada Beograda i centre gradskih opština Zemun, Mladenovac, Lazarevac i Obrenovac. Takođe se pojavljuje jedan klaster koji okružuje oblasti zahvaćene ćelijama iz klastera koje obuhvataju centralni dio Grada Beograda i centre gradskih opština Zemun, Mladenovac, Lazarevac i Obrenovac. Posljednja pravilnost koja se primjećuje pri pokretanju K-means algoritma klasterovanja je pojavljivanje klastera ćelija koje obuhvataju slabije naseljene oblasti Grada Beograda.

Da bi se riješio problem nestabilnosti rješenja, korišten je ansamblski K-means algoritam. Očekivanje je da će velik broj inicijalizacija algoritma K-means klasterovanja dati velik broj parova ćelija koje se jako često javljaju zajedno u klasteru za razne realizacije početnih uslova. Takođe je i očekivanje da se javi veliki broj parova ćelija koje se jako rijetko pojavljuju zajedno u klasteru za razne početne uslove. Na osnovu učestalosti zajedničkog pojavljivanja parova ćelija u klasterima, konstruisana je matrica sličnosti. Ovim se eliminiše proizvoljnost odabira početnih uslova i samim tim zavisnost rješenja algoritma od proizvoljno izabranih početnih uslova. U ovom slučaju, broj inicijalizacija algoritma K-means klasterovanja je $n_i = 1200$. Izabrana vrijednost broja inicijalizacija je optimalna vrijednost između računarskog vremena izvršenja algoritma i kvaliteta dobijenih

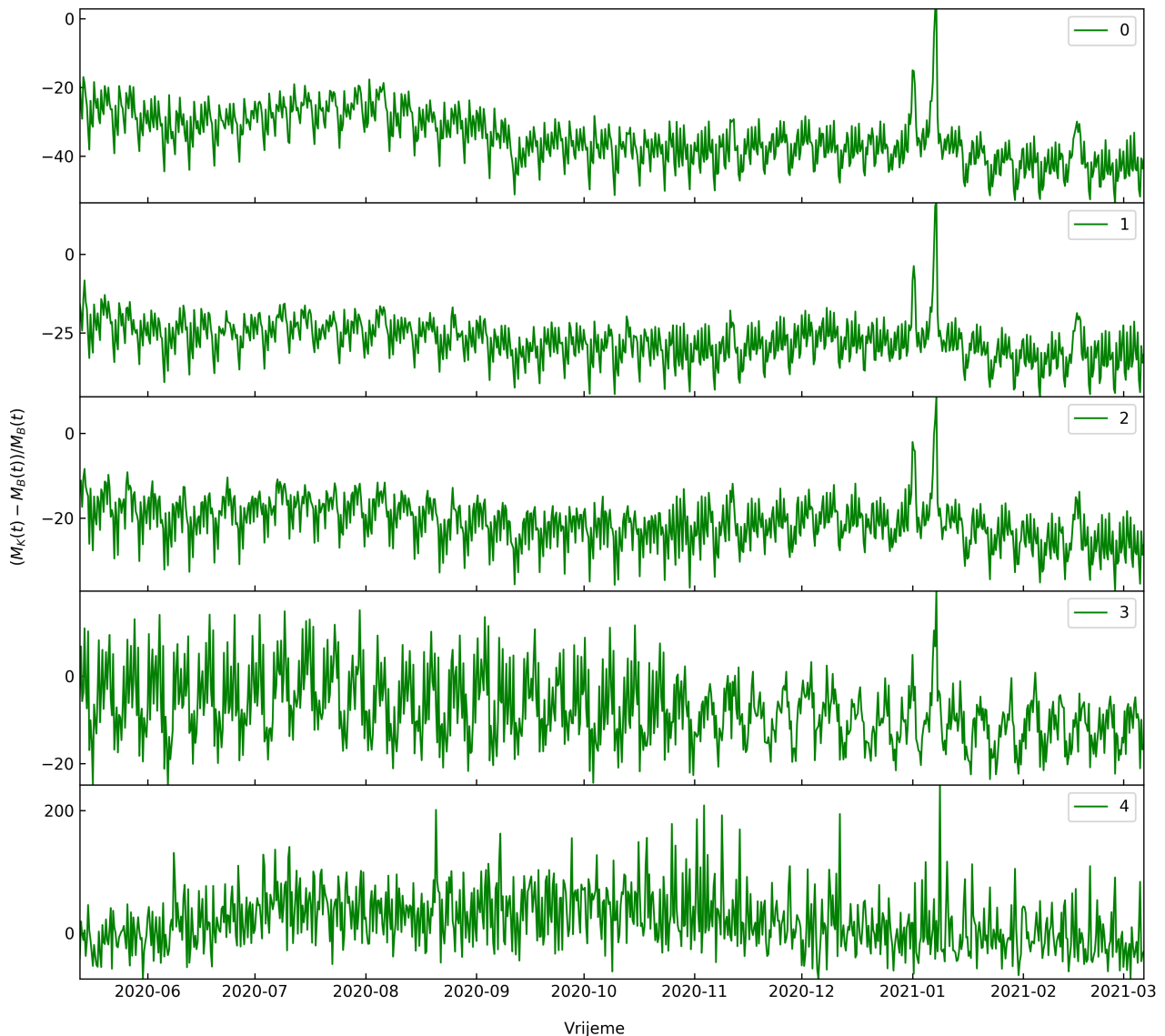
rezultata. Čelije su podijeljene u 5 klastera. Prag vrijednosti iznad kojeg su odbačeni unosi u matrici sličnosti je 0.7. Raspodjela ćelija u klasterne na mapi je prikazana na slici 8.



Slika 8: Za razne vrijednosti početnih uslova pokretan je K-means algoritam klasterovanja na vremenskim serijama relativne promjene broja kretanja unutar ćelija. Čelije čije su vremenske serije najčešće pridruživane istim klasterima su obojene istom bojom. Ovo predstavlja rješenje ansamblskog K-means algoritma. Crna linija predstavlja granice gradskih opština Grada Beograda.

Na Slici 8 su prikazane ćelije podijeljene u klasterne prema njihovoj funkciji u Gradu Beogradu dobijene pomoću ansamblskog K-means algoritma. Na slici 9 su prikazani centriodi ovih klastera. Oni predstavljaju srednju vrijednost relativne promjene broja kretanja unutar svih ćelija koje pripadaju datom klasteru. Na slici 8 se vidi da je obuhvaćena većina ranije opisanih pravilnosti u dodjeljivanju ćelija u klasterne koje su se dobijale pomoću algoritma K-means klasterovanja za razne početne uslove. Čelija koja predstavlja oblast na kojoj se nalazi Aerodrom Beograd je sama u klasteru 4. Centralni dio Grada Beograda, kao i centralni dijelovi opština Obrenovac, Lazarevac i Mladenovac su klaster 3. Oblasti u okolini centralnog dijela Beograda i centralnih dijelova opština Obrenovac, Lazarevac i Mladenovac (klaster 3) su takođe grupisane u klaster, klaster 2. Klasteri 1 i 2 predstavljaju slabije naseljene oblasti Grada Beograda. Centriodi klastera ćelija predstavljenih

na slici 8 su prikazani na slici 9.



Slika 9: Vremenske serije na slici predstavljaju centre klastera dobijenih pomoću ansamblskog K-means algoritma. Veličina koju predstavljaju vremenske serije je srednja vrijednost relativne promjene broja kretanja unutar ćelija (izražena u procentima) za svaki od klastera određenih na ovaj način.

Prvo što se može primjetiti iz posmatranja grafika centroida je da se mobilnost unutar ćelija smanjila u centroidima klastera 0, 1 i 2. Centroid klastera 4 predstavlja ćeliju koja predstavlja oblast na kojoj je Aerodrom Beograd. Centroid klastera 4 je zapravo jednak jedinoj vremenskoj seriji dodijeljenoj u ovaj klaster. Jedino se na centroidu ovog klastera primjećuje povećana unutrašnja mobilnost unutar čitavog perioda posmatranja mobilnosti. Centroid klastera 3 je u početku perioda posmatranja obrazaca mobilnosti imao mobilnost nešto veću nego u periodu prije krize. I u ovom klasteru se unutrašnja mobilnost smanjila nakon novembra 2020. Ovi podaci ukazuju na to da su se ljudi manje posjećivali i družili, i uopšte, manje kretali tokom posmatranog perioda. Na graficima centroida klastera 0, 1, 2 i 3 se primjećuje približna periodičnost u relativnoj promjeni broja kretanja unutar ćelija. Minimum unutrašnje mobilnosti kod ovih klastera odgovara periodu petak-nedjelja. Radnim danima se relativna promjena broja kretanja umjereno povećava od ponedjeljka do četvrtka. Karakteristične maksimalne vrijednosti na graficima centroida klastera 0, 1, 2 i 3 odgovaraju 31. decembru 2020. godine i 7. januaru 2021. godine. Približna periodičnost koja se primjećuje kod centroida klastera 0, 1, 2 i 3 se

ne može primjetiti na centroidu klastera 4. Očigledna razlika grafika ove vremenske serije od drugih centroida može biti objašnjenje nemogućnosti klasterovanja te vremenske serije sa drugim vremenskim serijama prilikom pokretanja algoritma K-means klasterovanja.

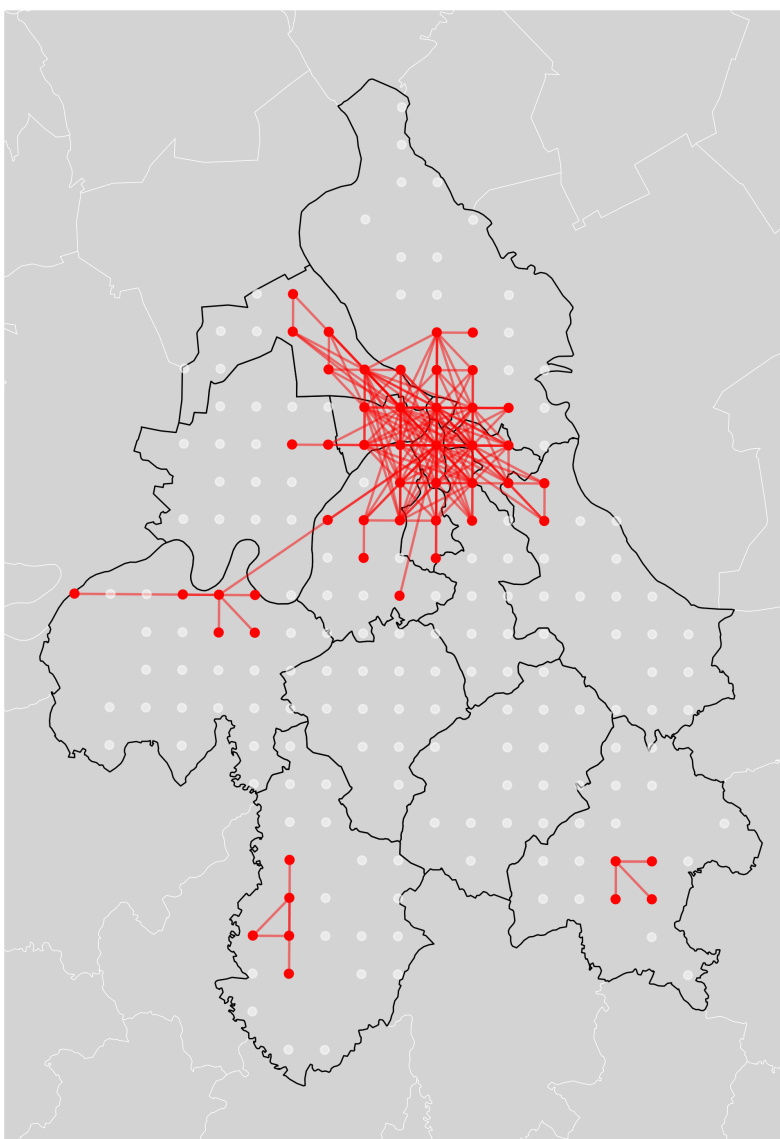
Posmatranjem položaja ćelija podijeljenim u klaster na teritoriji Grada Beograda (slika 8) zajedno sa centroidima tih klastera, može se primjetiti povećana mobilnost za vrijeme krize unutar ćelije na kojoj se nalazi Aerodrom Beograd, klaster 4. Mobilnost se nije drastično promijenila u centralnom dijelu Beograda, Obrenovcu, Mladenovcu i Lazarevcu. Ove oblasti su predstavljene klasterom 3. Unutar ćelija koje predstavljaju stambene oblasti i manje naseljene oblasti Beograda primjećuje se značajno smanjenje mobilnosti.

4.4 Mobilnost između ćelija

Mobilnost između ćelija je druga komponenta posmatranih obrazaca mobilnosti na teritoriji Grada Beograda, pored mobilnosti unutar ćelija. Iz podataka o mobilnosti je određen broj kretanja između ćelija i i j za vrijeme krize $W_{ij}^K(t)$ i broj kretanja u slučaju da nema krize $W_{ij}^B(t)$. Njihova relativna razlika $W_{ij}^R(t)$ se mijenja sa vremenom i predstavljena je kao vremenska serija koja se računa pomoću:

$$W_{ij}^R(t) = \frac{W_{ij}^K(t) - W_{ij}^B(t)}{W_{ij}^B(t)} \quad (4.2)$$

Ćelije na teritoriji Grada Beograda se posmatraju kao čvorovi u mreži. Između $N = 201$ čvora može biti maksimalno $(201 \cdot (201 - 1))/2 = 20100$ veza. Iz podataka o mobilnosti je određeno da postoji 180 aktivnih veza u navedenom vremenskom periodu. Ove veze su prikazane na slici 10.



Slika 10: Kružići na slici predstavljaju čvorove u mreži (centre ćelija). Kružići koji su obojeni bijelom bojom predstavljaju čvorove (ćelije) kod kojih u podacima postoji samo unutrašnje kretanje. Kružići obojeni crvenom bojom predstavljaju čvorove između kojih postoji kretanje. Ako je između dva čvora bar u jednom trenutku u posmatranom periodu postojalo kretanje, veza je formirana između ta dva čvora na slici. Za svaku od veza na slici (njih 180) postoji jedna vremenska serija relativne promjene broja kretanja korisnika između dva čvora.

Aktivne veze su veze koje su bar u jednom trenutku povezivale dva čvora (ćelije). Za svaku od 180 veza između ćelija i i j na teritoriji Grada Beograda odgovara jedna vremenska serija relativne promjene broja kretanja $W_{ij}^R(t)$. Vremenske serije relativne promjene broja kretanja između dva čvora predstavljaju aktivnost svake od 180 veza za vrijeme krize u navedenom periodu od 298 dana. U svakom danu postoje tri vremenska intervala od 8 sati, pa je dužina vremenskih serija relativne promjene broja kretanja $298 \cdot 3 = 894$. U slučaju da se prilikom računanja vremenskih serija relativne promjene broja kretanja između ćelija jave nedostajuće vrijednosti, one su izračunate interpolacijom u odnosu na susjedne tačke unutar vremenskih serija. Na ovim vremenskim serijama je pokrenut algoritam K-means klasterovanja da bi se veze podijelile u klustere. Prilikom pokretanja algoritma K-means klasterovanja na ovim vremenskim serijama, dobijana su različita rješenja za različite početne uslove. Da bi se riješio problem nestabilnih rješenja koji se javlja kod primjene algoritma K-means klasterovanja, vremenske serije relativne promjene broja kretanja između ćelija su takođe podijeljene u klustere primjenom ansamblskog K-means algoritma da bi se dobilo stabilno rješenje. Rješenje ovog problema podrazumijeva vremenske serije relativne promjene broja kretanja između ćelija grupisane u klustere. Svako od ovih vremenskih serija relativne promjene broja kretanja između ćelija odgovara veza između tih ćelija prikazana na mapi.

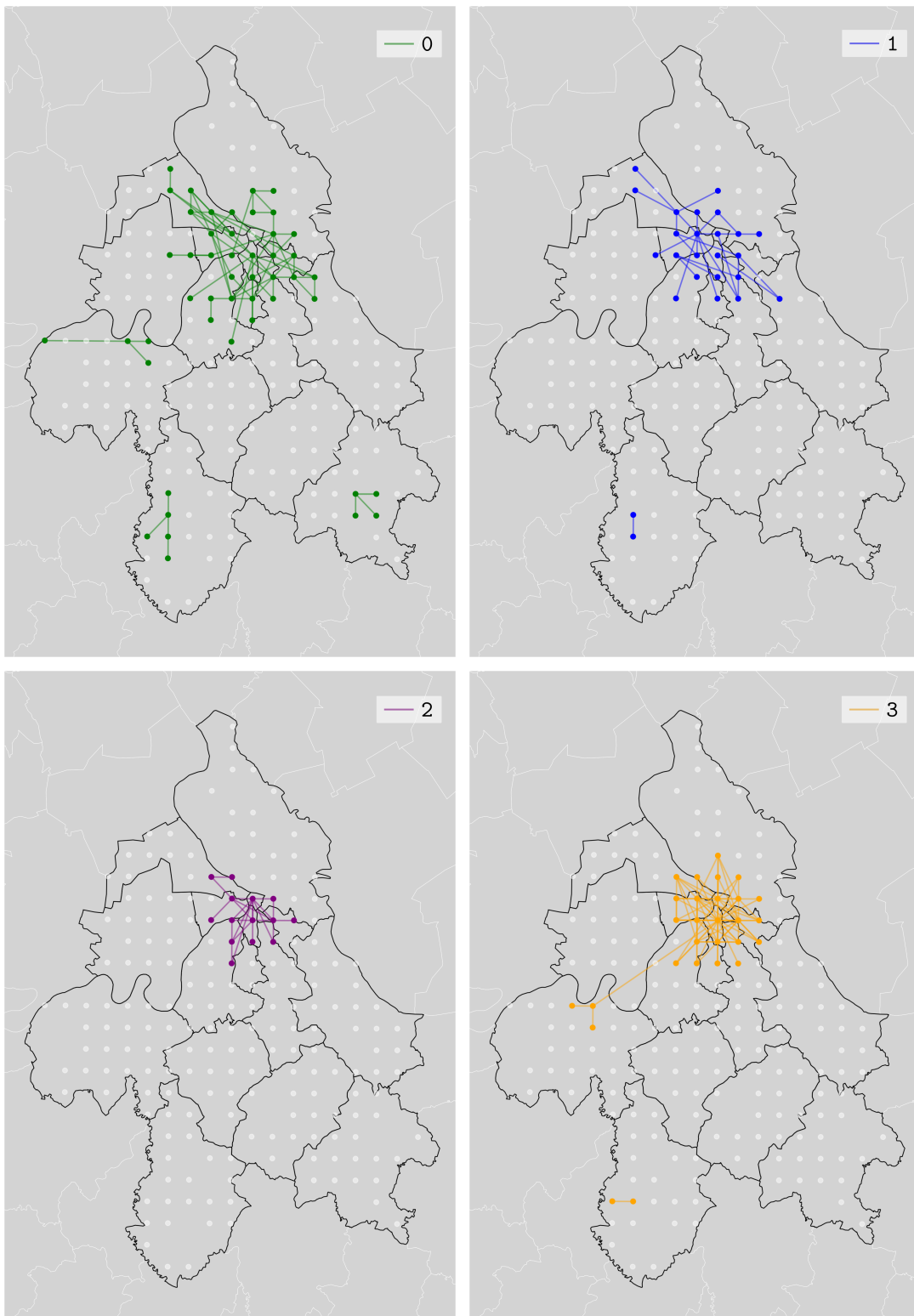
4.4.1 Klasterovanje veza pomoću ansamblskog K-means algoritma

Primjenom algoritma K-means klasterovanja na vremenske serije relativne promjene broja kretanja između ćelija, veze su raspoređivane na različite načine u 4 klastera. Prilikom pokretanja algoritma, primjećene su neke pravilnosti prilikom podjele veza u klustere. Često se pojavljivao klaster koji sadrži veze između ćelije koje se nalaze u stambenim dijelovima Beograda van centralnog dijela Beograda, te oblasti oko najnaseljenih dijelova gradskih opština Obrenovac, Lazarevac i Mladenovac.

Da bi se riješio problem nestabilnosti rješenja, primjenjen je ansamblski K-means algoritam na vremenskim serijama relativne promjene broja kretanja između ćelija. Prilikom primjene algoritma K-means klasterovanja, očekuje se da će velik broj inicijalizacija algoritma davati velik broj parova veza između ćelija koje se jako često javljaju zajedno u klasteru za razne vrijednosti početnih uslova. Takođe se i očekuje da se javi veliki broj parova veza između ćelija koje se jako rijetko pojavljuju zajedno u klasteru za razne početne uslove. Na osnovu učestalosti zajedničkog pojavljivanja parova veza između ćelija u klasterima, konstruisana je matrica sličnosti. Ovim se eliminiše proizvoljnost odabira početnih uslova i samim tim zavisnost rješenja algoritma od proizvoljno izabranih početnih uslova. U ovom slučaju, broj inicijalizacija algoritma K-means klasterovanja je $n_i = 1200$. Izabrana vrijednost broja inicijalizacija je optimalna vrijednost između računarskog vremena izvršenja algoritma i kvaliteta dobijenih rezultata. Veze između ćelija su podijeljene u 4 klastera na osnovu podjele njihovih vremenskih serija relativne promjene broja kretanja u klustere. Prag vrijednosti iznad kojeg su odbačeni unosi u matrici sličnosti je 0.9. Raspodjela veza između ćelija u klustere na mapi je prikazana na slici 8.

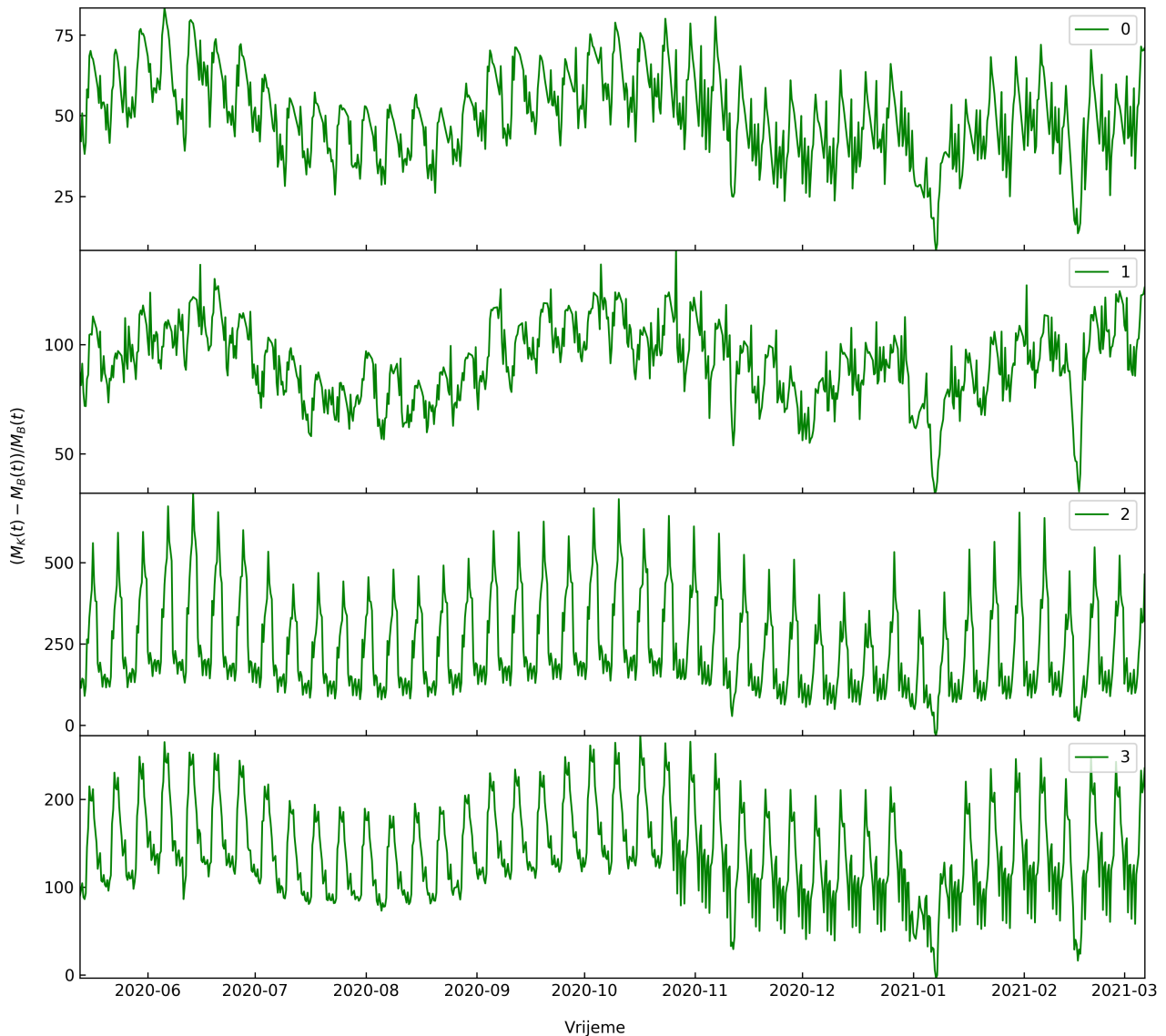
Sa slike 8 se vidi raspodjela klastera veza na mapi. Klasteru 0 pripadaju veze koje povezuju oblasti u okolini najnaseljenijih oblasti (pogledati sliku 3). Veze koje povezuju stambene oblasti van centralnog dijela Beograda su pridružene klasteru 1. Ovom klasteru je takođe pridružena jedna veza sa oblasti gradske opštine Lazarevac. Najmanje veza se nalazi u klasteru 2, koje povezuju najuži centar Beograda. Ove veze povezuju ćelije na teritoriji centralnog dijela Beograda i gradskih opština Zvezdara, Novi Beograd i Zemun. Klaster 3 obuhvata

veze koje povezuju centralni dio Beograda sa stambenim oblastima najbližim centru Beograda. Ovom klasteru je pridružena i veza između centralnog dijela Beograda i gradske opštine Obrenovac.



Slika 11: Za razne vrijednosti početnih uslova pokretan je K-means algoritam klasterovanja na vremenskim serijama relativne promjene broja kretanja između ćelija. Da bi se riješio problem nestabilnosti rješenja, primjenjen je ansamblski K-means algoritam. Veze između čvorova (ćelija) su podijeljene u 4 klastera. Svaka od vremenskih serija relativne promjene broja kretanja korisnika između ćelija je predstavljena vezom na slici. Veze čije vremenske serije pripadaju istom klasteru su obojene istom bojom. Kružići na slici predstavljaju čvorove u mreži (centre ćelija). Kružići obojeni bijelom bojom predstavljaju čvorove (ćelije) kod kojih u podacima postoji samo unutrašnje kretanje.

Za svaki od klastera prikazanih na slici 11, iscrtani su grafici centroida na slici 12.



Slika 12: Vremenske serije na slici predstavljaju centroide klastera dobijenih pomoću ansamblskog K-means algoritma na vremenskim serijama relativne promjene broja kretanja između ćelija. Veličina koju predstavljaju vremenske serije je srednja vrijednost relativne promjene broja kretanja između ćelija (izražena u procentima) za svaki od klastera određenih na ovaj način.

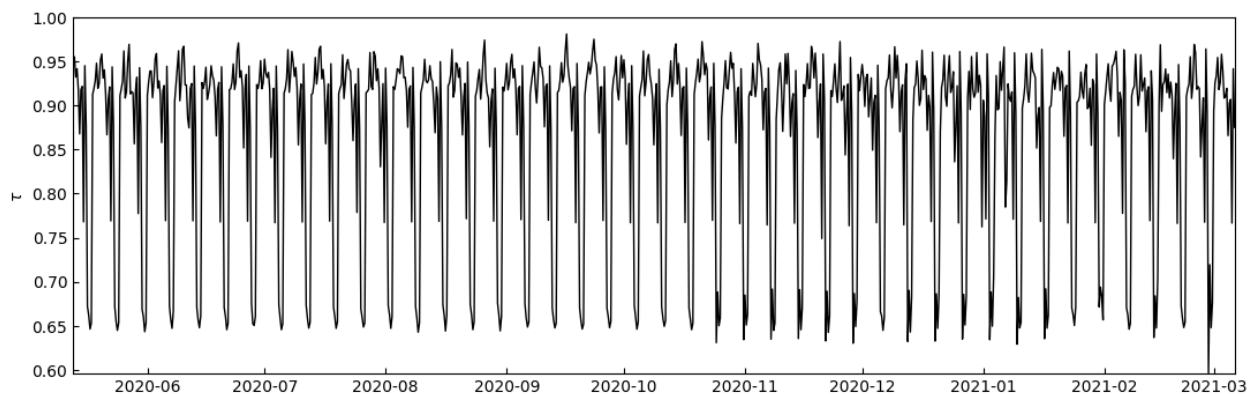
Prvo što se primjećuje kod grafika centroida klastera je da se mobilnost između ćelija drastično povećala za vrijeme krize u odnosu na period prije krize. Najveće povećanje mobilnosti između ćelija je u klasteru 2, klasteru koji sadrži veze koje povezuju centar Beograda sa gradskim opštinama Zemun, Novi Beograd i Zvezdara. Najmanja promjena mobilnosti između ćelija pripada klasteru 0, klasteru koji povezuje stambene oblasti u blizini najnaseljenijih oblasti na teritoriji Grada Beograda. Kao i kod centroida klastera vremenskih serija relativne promjene broja kretanja unutar ćelija, i ovdje se može primjetiti približna periodičnost sa periodom od 7 dana. Karakteristične maksimalne vrijednosti na grafiku odgovaraju maksimalnoj relativnoj promjeni broja kretanja između ćelija koja se dešava danom subota. Relativna promjena broja kretanja je približno konstantna na sedmičnom nivou, dok se petkom primjećuje minimalna vrijednost u poređenju sa ostalim danima u sedmici. Posmatranjem trenda grafika vremenskih serija primjećuje se da se mobilnost između ćelija u prosjeku povećavala u početku perioda prikupljanja podataka za vrijeme krize. Ovo odgovara periodu popuštanja epidemioloških

mjera. Nakon ovog perioda mobilnost između ćelija se u prosjeku smanjivala pa opet povećavala prema kraju 2020. godine. Poslije ovog je nastupio stabilniji period mobilnosti između ćelija. Minimum mobilnosti koji se primjećuje na graficima svih centroida odgovara 7. januaru 2021. godine.

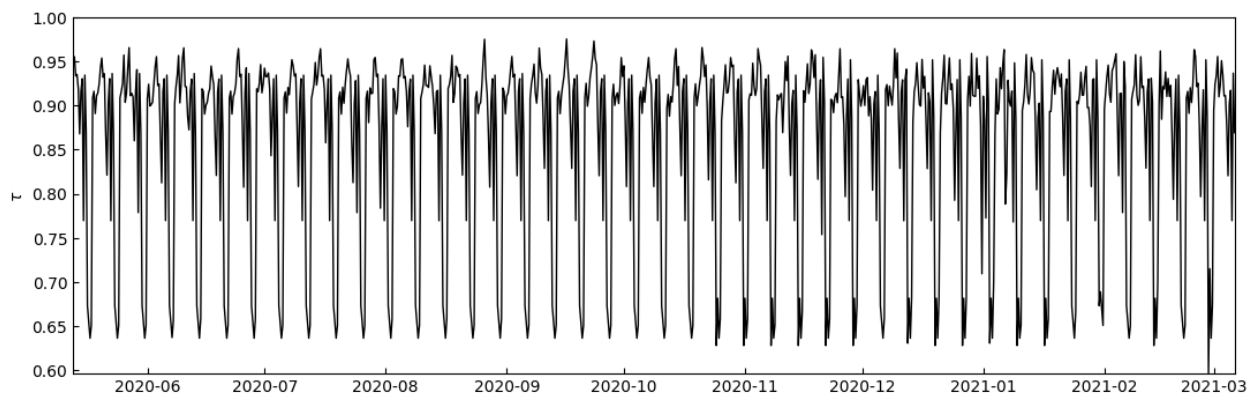
Posmatranjem raspodjele klastera veza između ćelija na mapi i centroida ovih klastera primjećuje se da je najveće povećanje mobilnosti u navedenom periodu u klasteru 3 koji obuhvata veze koje povezuju centralni dio Beograda sa stambenim oblastima najbližim centru Beograda. Takođe se primjećuje značajnija promjena u mobilnosti između ćelija u klasteru 3 koji obuhvata veze koje povezuju centralni dio Beograda sa stambenim oblastima najbližim centru Beograda. U klasteru 0 koji povezuju oblasti u najnaseljenijim oblastima (centralnom dijelu Beograda i centrima opština Obrenovac, Lazarevac i Mladenovac) i klasteru 1 koji predstavlja veze koje povezuju stambene oblasti van centralnog dijela Beograda nije primjećena drastična promjena u mobilnosti između ćelija.

4.5 Poređenje slojeva višeslojne mreže prema rangiranju čvorova i veza

U cilju određivanja korelacije između rangiranja čvorova i veza u različitim trenucima, generisane su višeslojne mreže iz podataka o kretanju [5]. Iz podataka o mobilnosti su generisane dvije višeslojne mreže, jedna koja predstavlja mobilnost za vrijeme krize $W_{ij}^K(t)$ i druga koja predstavlja mobilnost za navedeni period za slučaj kad ne bi bilo krize $W_{ij}^B(t)$. U podacima za navedeni period postoji 298 intervala od 8 sati pa tako i svaka od ove dvije mreže ima 894 sloja koji predstavljaju jedan vremenski trenutak. Topološki su slojevi ovih mreža identični u istom vremenskom trenutku, razlikuju se samo po težinama veza $w_{ij}(t)$. Čvorovi u ovim mrežama su ćelije između kojih se vrši kretanje, a veze postoje u datom sloju ako je u tom trenku postojalo kretanje između čvorova. Ako je veza u nekom trenutku aktivna, njoj se pripisuje težina veze jednaka broju kretanja korisnika u datom trenutku. Takođe se svakom čvoru pripisuje veličina jačina čvora s_i^X gdje je X indeks višeslojne mreže: B za višeslojnu mrežu koja predstavlja mobilnost u slučaju da nema krize i K za višeslojnu mrežu kojom je opisana mobilnost za vrijeme krize. Jačina čvora je jednaka sumi svih težina veza sa kojim je taj čvor povezan sa drugim čvorovima: $s_i^X = \sum_j W_{ij}^X$. Ovo nam dozvoljava da poredimo veličinu koja opisuje čvor ili vezu od jednog sloja mreže do drugog u jednoj višeslojnoj mreži ili u obje mreže u istom vremenskom trenutku [5].

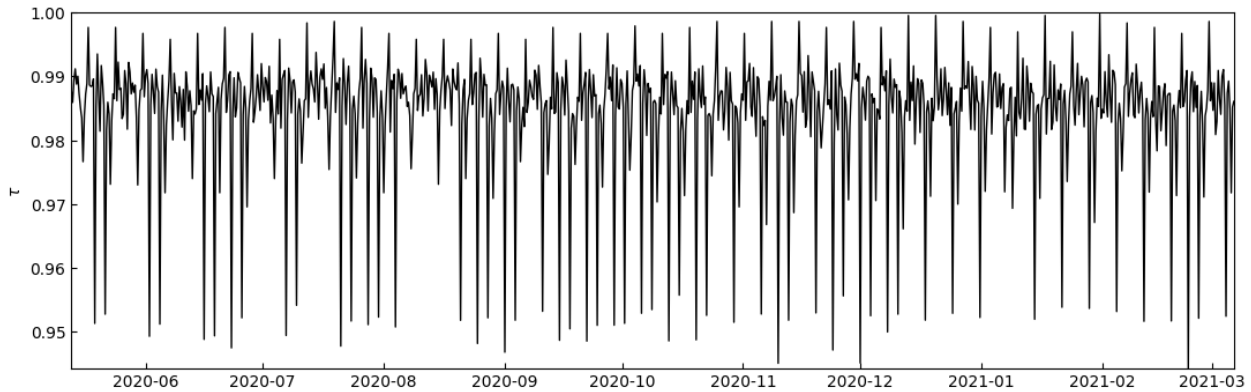


Slika 13: Kendallov koeficijent ranga je određen za rangirane nizove vrijednosti jačine čvorova s_i u višeslojnoj mreži u slučaju krize $W_{ij}^K(t)$. Kendallov koeficijent se određuje za rangove niza vrijednosti jačine čvorova s_i dva susjedna sloja višeslojne mreže. Kako jedan sloj mreže predstavlja jedan vremenski trenutak, na grafiku je prikazano kako se τ mijenja za rangove nizova vrijednosti jačine čvorova s_i od jednog vremenskog trenutka do narednog vremenskog trenutka.

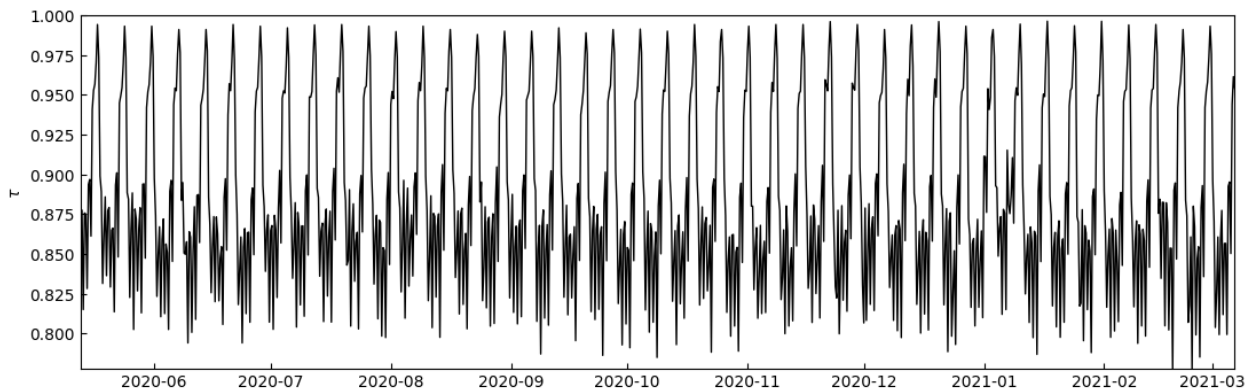


Slika 14: Kendallov koeficijent ranga je određen za rangirane nizove vrijednosti jačine čvorova s_i u višeslojnoj mreži za slučaj da nema krize $W_{ij}^B(t)$. Kendallov koeficijent se određuje za rangove niza vrijednosti jačine čvorova s_i dva susjedna sloja višeslojne mreže. Kako jedan sloj mreže predstavlja jedan vremenski trenutak, na grafiku je prikazano kako se τ mijenja za rangove nizova vrijednosti jačine čvorova s_i od jednog vremenskog trenutka do narednog vremenskog trenutka.

Poređenje se vrši Kendalovim koeficijentom korelacije ranga [36]. Posmatranjem susjednih slojeva mreže (dva susjedna vremenska trenutka) unutar jedne od dvije višeslojne mreže ($W_{ij}^K(t)$ ili $W_{ij}^B(t)$) vidi se da se važnost čvorova (ćelija) ne mijenja bitno tokom radnih dana, dok se najveća promjena primjećuje tokom subota i nedjelja (slika 13 i slika 14).



Slika 15: Posmatrane su dvije višeslojne mreže - prva koja predstavlja slojeve mreže za vrijeme krize (vremenska mreža u različitim vremenskim trenucima) $W_{ij}^K(t)$ i druga koja predstavlja slojeve mreže u slučaju da nema krize $W_{ij}^B(t)$. Kendalov koeficijent ranga je određivan za rangirane nizove vrijednosti jačine čvorova s_i u slojevima ove dvije mreže koji predstavljaju isti vremenski trenutak. Na grafiku je prikazano kako se τ mijenja sa vremenom.



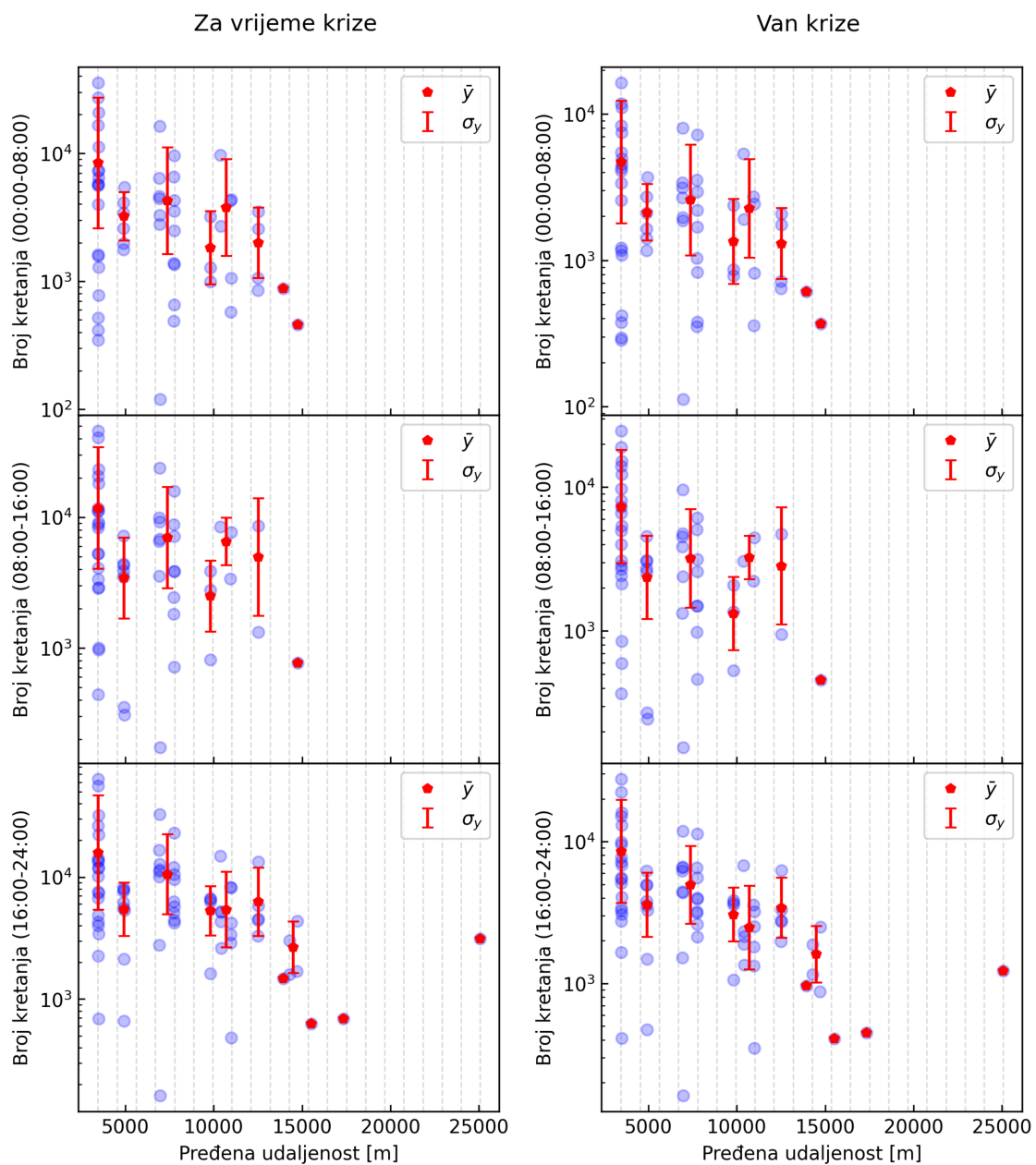
Slika 16: Posmatrane su dvije višeslojne mreže - prva koja predstavlja slojeve mreže za vrijeme krize (vremenska mreža u različitim vremenskim trenucima) $W_{ij}^K(t)$ i druga koja predstavlja slojeve mreže u slučaju da nema krize $W_{ij}^B(t)$. Kendalov koeficijent ranga je određivan za rangirane nizove vrijednosti težine veza w_{ij} u slojevima ove dvije mreže koji predstavljaju isti vremenski trenutak. Na grafiku je prikazano kako se τ mijenja sa vremenom.

Poređenjem rangiranja čvorova u obje višeslojne mreže u slojevima koji predstavljaju isti vremenski trenutak vidi se da se važnost čvorova (ćelija) ne mijenja bitno, iako postoji značajna promjena u mobilnosti (slika 15). Rangirane su težine veza u obje višeslojne mreže u slojevima koji predstavljaju isti vremenski trenutak. Primjećuje se da se rangiranje veza promijenilo do određene mjere zbog epidemije. Ova razlika je značajnija tokom radnih dana gdje $\tau < 0.9$, dok je promjena manja tokom subota i nedjelja (slika 16).

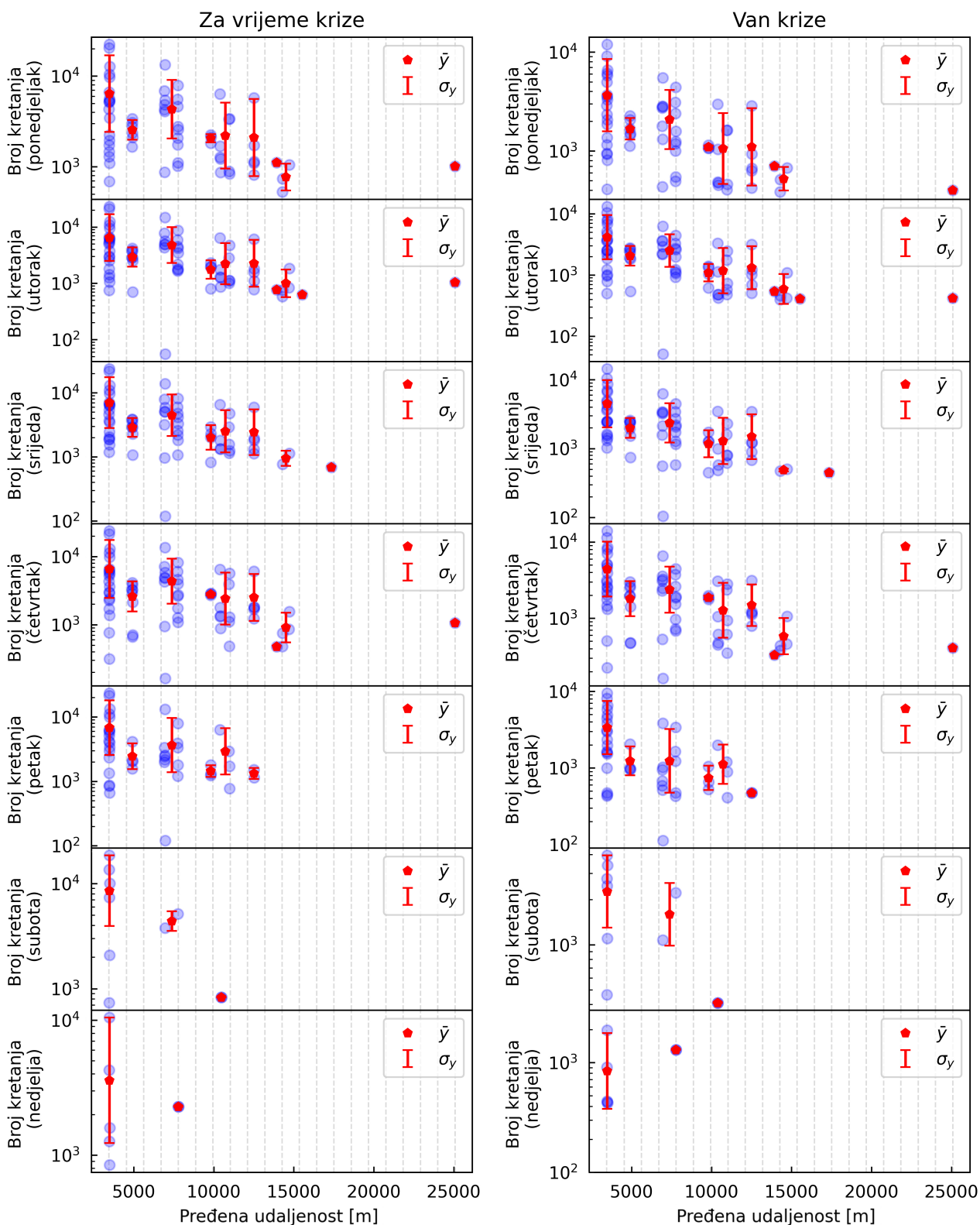
Vidimo i da fluktuacije koje postoje u vremenskim serijama relativne promjene broja kretanja unutar ćelija i relativne promjene broja kretanja između ćelija (koje su posljedica povećanog broja slučajeva COVID-19) se ne primjećuju u samim mrežama.

4.6 Broj kretanja u funkciji udaljenosti

Sa ciljem da se ispita da li i kako se mijenjaju obrasci kretanja u Gradu Beogradu usljed epidemije COVID-19, analizirana je zavisnost broja putovanja od rastojanja usrednjena po vremenskim trenucima u toku dana (slika 17) i danima u sedmici (slika 18). Ova zavisnost je izračunata posebno za obrasce kretanja prije krize i u toku krize. Broj putovanja je normiran tako što je za svaki od 3 vremenska intervala u danu određena maksimalna vrijednost putovanja M_{max}^t , $t \in \{00-08, 08-16, 16-00\}$, i minimalna vrijednost M_{min}^t u posmatranom periodu od 298 dana. Zatim svaku tačku u vremenskoj seriji koja odgovara jednom od ova tri vremenska intervala od njene originalne vrijednosti $M(t)$ oduzmemo minimalnu vrijednost M_{min} nađenu za interval dana kome ta tačka pripada i normiramo ovu razliku sa $M_{max} - M_{min}$. Normiranje je napravljeno zbog različitog broja ukupnih putovanja prije i tokom krize. Podaci su podijeljeni u 20 binova na horizontalnoj osi i određene su srednje vrijednosti broja kretanja.



Slika 17: Broj kretanja u funkciji udaljenosti za tri intervala u danu.



Slika 18: Broj kretanja u funkciji udaljenosti za dane u sedmici.

Kao što su i prethodne analize obrazaca mobilnosti pokazale, broj putovanja između dvije ćelije opada sa njihovim međusobnim rastojanjem [34], odnosno, najveći broj putovanja je unutar same i između susjednih ćelija, dok je ovaj broj između ćelija koje su udaljene 15 i više kilometara za cijeli red veličina manji. Na osnovu slike 17 može se zaključiti da se ovi obrasci razlikuju za vrijeme različitih vremenskih intervala, posebno za duža putovanja. Konkretno, dok se putovanja od 25 km ne pojavljuju tokom noći i u toku perioda do 16h, ova

putovanja se događaju tokom popodnevnih časova, period od 16h do 24h. Takođe, na osnovu slike 17 može se zaključiti da se zavisnost broja putovanja ne mijenja tokom krize u odnosu na period prije krize, kada se ova zavisnost usrednji u odnosu na vremenske periode u toku dana.

Ista zavisnost je usrednjena po pojedinim danima. Na osnovu slike 18 može se zaključiti da se obrasci mobilnosti razlikuju po danima, odnosno da postoji jasna razlika između perioda ponedjeljak-petak i subota-nedjelja. Sa slike 18 slijedi da su obrasci mobilnosti stabilni radnim danima, sa malim razlikama koje se primjećuju u slučaju dugih putovanja i obrasca tokom petka. Konkretno, duga putovanja se dešavaju ponedjeljkom, utorkom i četvrtkom, dok se ona nisu prisutna srijedom i petkom. Petkom je broj dužih putovanja generalno manji. Vikend dane karakterišu vrlo kratka putovanja, što ukazuje na to da su duga putovanja direktna posljedica poslovnih aktivnosti građana Grada Beograda, a da se njihova putovanja tokom vikenda dešavaju u nekom radiusu od 8 km. Sa slike 18 slijedi i da se ovi obrasci ne mijenjaju usljed COVID-19 krize. Iako COVID-19 kriza ne utiče mnogo na zavisnost broja putovanja u odnosu na rastojanje, mora se primjetiti jedna očigledna razlika. Naime, na slici 18 vidi se da je broj putovanja dužine oko 8 km u toku dana nedjelja bio veći prije nastanka krize nego u toku krize. Ovo ukazuje na činjenicu da su se ljudi ograničili na mjesta koja su im bliža usljed njihovih slobodnih aktivnosti. Dalje, vidimo da je fluktuacija broja putovanja za srednje udaljenosti, od 7 km do 8 km, veća prije krize nego tokom krize, što dodatno podržava zaključak da su ljudi smanjili broj dužih putovanja, i svoje slobodne aktivnosti ograničili na mjesta koja su bliža njihovom mjestu boravka.

5 Zaključak

Iako kretanje pojedinaca može djelovati nasumično usljed raznih faktora koji djeluju na njihovo kretanje, kada se izvrši prostorno-vremenska analiza obrazaca mobilnosti primjećuju se neke pravilnosti. Najveći dio ranijih istraživanja je zasnovan na podacima dobijenih iz normalnih perioda. Dinamika sistema opisanih ovim podacima je lakša za razumijevanje od sistema u kojima su nastupile krize kao što su požari, zemljotresi i epidemije. Ovakve krize značajno utiču na prirodu interakcija u takvom sistemu.

Posmatrani su podaci dobijeni sa Facebook-ovog portala Data for good. Ovi podaci su prostorno predstavljeni preko ćelija kvadratnog oblika određenih dimenzija, a vremenski unutar intervala od 8 sati. Uz ove podatke o mobilnosti korisnika su takođe obezbjeđeni podaci o prostornoj raspodjeli tih korisnika, pa je prije analize obrazaca mobilnosti posmatrana prostorna raspodjela tih korisnika. Kako podaci obezbjeđuju obrasce mobilnosti za period krize i za period prije krize, najzgodnije je bilo posmatrati relativnu promjenu broja kretanja za vrijeme krize u odnosu na period prije krize. Da bi se ovi podaci bolje razumjeli, izvršena je podjela podataka u zavisnosti od prirode kretanja pojedinaca čije je kretanje posmatrano.

Prva komponenta obrazaca mobilnosti su bila kretanja unutar pomenutih ćelija na teritoriji Grada Beograda. Ova kretanja su uglavnom kretanja korisnika u blizini tačke u kojoj provode najviše vremena tokom krize. Primjećeno je broj kretanja unutar ćelija manji za vrijeme krize u odnosu na period prije krize. Prvo se primjećuje vremenska pravilnost u obrascima kretanja ljudi na nivou sedmice. Ovo je posljedica načina na koji ljudi obavljaju svoje aktivnosti tokom sedmice. Unutar većine ćelija je došlo do relativnog smanjenja broja putovanja za vrijeme krize. Takođe se može zaključiti da se prilikom analize kretanja unutar ćelija mogu pronaći ćelije sa netipičnim obrascima kretanja (npr. aerodrom sa velikim brojem kretanja).

Druga komponenta obrazaca mobilnosti sa Facebook-a su kretanja između ćelija na teritoriji Grada Beograda. Za razliku od relativne promjene broja kretanja unutar ćelija, ovdje se primjećuje znatno povećanje mobilnosti korisnika između ćelija za period krize. Takođe se primjećuje približna periodičnost na sedmičnom nivou. Najmanja mobilnost je za dan petak, a najveća za dan subota na sedmičnom nivou. U periodu nedjelja-četvrtak, mobilnost korisnika između ćelija je približno konstantna.

Primjenom metoda kompleksnih mreža, izvršena je analiza mreža generisanih iz dobijenih podataka o mobilnosti. Poređenjem veličine koja opisuje važnost čvorova u datom trenutku u mreži, vidi se da se važnost čvorova ne mijenja bitno sa vremenom, što je interesatno jer postoji značajna promjena u mobilnosti između čvorova u posmatranom periodu. Posmatranjem veličine koja opisuje važnost veze u mreži u datom trenutku, primjećuje se da se važnost veza promijenila do određene mjere usred epidemije.

Da bi se napravila još jedna procjena u razlici ili sličnosti obrazaca kretanja za vrijeme krize i van krize, posmatran je broj kretanja korisnika u funkciji pređene udaljenosti. Kretanja su podijeljena na dane u sedmici u kojim se odvijaju i na intervale u danu u kojim se odvijaju. U oba slučaja je primjećeno da broj putovanja između dvije ćelije opada sa njihovim međusobnim rastojanjem i da kriza ne utiče na ovu zavisnost.

Prostorno-vremenska analiza obrazaca mobilnosti u Gradu Beogradu predstavljena u ovom radu je pokazala da univerzalnost ovih obrazaca u smislu sedmičnih ciklusa se ne mijenja usred pojave epidemije. Analiza je pokazala i da broj putovanja u toku krize zavisi od broja zaraženih, odnosno, povećanje u broju zaraženih neminovno vodi ka smanjenju mobilnosti u Gradu Beogradu. Smanjenju broja putovanja najviše doprinose putovanja unutar iste ćelije, dok je došlo do relativnog povećanja broj putovanja između ćelija. Iako se značaj putanja između ćelija u pogledu njihovih težina ne mijenja drastično, vidimo da ovakve promene ipak postoje.

One su najizraženije za radne dane, dok su značajni putanja između ćelija nedeljom isti tokom i prije krize. Rezultati predstavljeni u ovom radu ukazuju na to da su obrasci mobilnosti zadržali određeni stepen stabilnosti i predvidljivosti tokom krize. Ipak, postoje izvjesne razlike u mobilnosti tokom krize u odnosu na period prije krize koje se moraju uzeti u obzir prilikom modeliranja. Pored toga, u slučaju modeliranja mobilnosti u toku epidemija neophodno je uzeti u obzir broj zaraženih ljudi u populaciji.

Još ne postoji kompletno razumijevanje obrazaca kretanja usljed kriza. Moguće je da će se razumijevanje ovih sistema unaprijediti sa prilivom veće količine podataka i sa povećanjem rezolucije tih podataka. Sa povećanjem količine podataka o mobilnosti pojedinaca i njihove rezolucije, nailazi se na jasan problem narušenja privatnosti pojedinaca čiji su podaci prikupljeni. Iako su podaci sa Facebook Data for good anonimizovani u cilju zaštite privatnosti korisnika, njihovom analizom se mogu primjetiti neke pravilnosti koje se mogu iskoristiti u cilju ublažavanja posljedica nekih kriza.

6 Reference

- [1] F. Asgari, V. Gauthier, M. Becker *A survey on Human Mobility and its applications* arXiv:1307.0814 (2013)
- [2] C. L. Archer, G. Cervone, M. Golbazi, N. A. Fahel, C. Hultquist *Changes in air quality and human mobility in the USA during the COVID-19 pandemic* Bulletin of Atmospheric Science and Technology volume 1, 491–514 (2020)
- [3] F. Battiston, V. Nicosia, V. Latora *Structural measures for multiplex networks* Phys. Rev. E 89, 032804 (2014)
- [4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang *Complex networks: Structure and dynamics* Physics Reports, 424, 4–5, 175-308 (2006)
- [5] S. Boccaletti, G. Bianconi, R. Criado, C.I. del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, M. Zanin *The structure and dynamics of multilayer networks* Physics Reports, 544, 1, 1-122 (2014)
- [6] V. D. Blondel et al. *Fast unfolding of communities in large networks* J. Stat. Mech. P10008 (2008)
- [7] D. Brockmann, L. Hufnagel, T. Geisel *The scaling laws of human travel* Nature 439 (2006)
- [8] A.-L. Barabási *Network Science* Cambridge University Press (2016)
- [9] C. Castellano, S. Fortunato, V. Loreto *Statistical physics of social dynamics*, Reviews of modern physics 81, 591 (2009)
- [10] A. Chatterjee, M. Mitrović, S. Fortunato *Universality in voting behavior: an empirical analysis*, Sci. Repos. 3, 1049 (2013)
- [11] R. Cohen, K. Erez, D. ben-Avraham, S. Havlin *Breakdown of the Internet under Intentional Attack* Physical Review Letters. 16 (86): 3682–5 (2001)
- [12] J. R. Cloug, J. Gollings, T. V. Loach, T. S. Evans *Transitive reduction of citation networks* Journal of Complex Networks. 3 (2): 189–203 (2015)
- [13] T. Cristiano, S. L. Zhao *Machine Learning in Complex Networks*, Springer (2016)
- [14] E. G. Dada, J. S. Bassi, H. Chiroma, S. M. Abdulhamid, A. O. Adetunmbi, O. E. Ajibuwa *Machine learning for email spam filtering: review, approaches and open research problems* Heliyon, Volume 5, Issue 6 (2019)
- [15] S. Eubank et al. *Modelling disease outbreaks in realistic urban social networks* Nature 429, 180–184 (2004)
- [16] B. Evkoski, I. Mozetič, N. Ljubešić, N. P. Kralj *Community evolution in retweet networks* PLoS ONE 16(9): e0256175 (2021)

- [17] A. Esteva, B. Kuprel, R. Novoa et al. *Dermatologist-level classification of skin cancer with deep neural networks* Nature 542, 115–118 (2017)
- [18] A. M. Fahim, A. M. Salem, F. A. Torkey, M. A. Ramadan *An efficient enhanced k-means clustering algorithm* Journal of Zhejiang University-SCIENCE A 7, 1626–1633 (2006)
- [19] M. Gosak, R. Markovič, J. Dolenšek, M. S. Rupnik, M. Marhl, A. Stožer, M. Perc *Network science of biological systems at different scales: A review*, Physics of life reviews 24, 118–135 (2018)
- [20] M. C. Gonzalez, C. A. Hidalgo, A.-L. Barabási *Understanding individual human mobility patterns* Nature 453, 779–782 (2008)
- [21] R.E. González, R.P. Muñoz, C.A. Hernández *Galaxy detection and identification using deep learning and data augmentation* Astronomy and Computing 25, 103-109 (2018)
- [22] L. Gabrielli, E. Deutschmann, F. Natale et al. *Dissecting global air traffic data to discern different types and trends of transnational human mobility* EPJ Data Sci. 8, 26 (2019)
- [23] P. Holme, J. Saramäki *Temporal networks* Physics Reports, 519, 3, 97-125 (2012)
- [24] G. Krings, F. Calabrese, C. Ratti, V. D. Blondel *Urban gravity: A model for inter-city telecommunication flows* J. Stat. Mech.-Theor. Exp. L07003 (2009)
- [25] X. Lu, L. Bengtsson, P. Holme *Predictability of population displacement after the 2010 Haiti earthquake*, Proceedings of the National Academy of Sciences 109, 11576–11581 (2012)
- [26] Franck Legendre, Thrasyvoulos Spyropoulos, Theus Hossmann *A complex network analysis of human mobility* 3rd IEEE International Workshop on Network (2011)
- [27] J. B. MacQueen *Some Methods for classification and Analysis of Multivariate Observations*, Berkeley, University of California Press, 281-297 (1967)
- [28] Makse, H. A., Havlin, S. and Stanley, H. E. *Modelling urban growth patterns* Nature 377, 608–612 (1995)
- [29] P. Maas et al *Facebook Disaster Maps: Aggregate Insights for Crisis Response and Recovery* WiPe Paper - Social Media in Crisis and Conflicts Proceedings of the 16th ISCRAM Conference (2019)
- [30] M. Mitrović, G. Paltoglou, and B. Tadić, *Networks and emotion-driven user communities at popular blogs*, The European Physical Journal B 77, 597–609 (2010)
- [31] Andreas C. Müller and Sarah Guido *Introduction to Machine Learning with Python* (2016)
- [32] M. E. Newman, S. H. Strogatz, and D. J. Watts *Random graphs with arbitrary degree distributions and their applications* Physical review E, vol. 64, no. 2 (2001)
- [33] Newman, M. E. J. *Coauthorship networks and patterns of scientific collaboration* Proceedings of the National Academy of Sciences 101(suppl 1), 5200-5205 (2004)
- [34] V. Palchykov, M. Mitrović, H.-H. Jo, J. Saramäki, R. K. Pan *Inferring human mobility using communication patterns*, Sci. Reps. 4, 6174 (2014)

- [35] A. Radovic, M. Williams, D. Rousseau et al. *Machine learning at the energy and intensity frontiers of particle physics* Nature 560, 41–48 (2018)
- [36] N. J. Salkind *Encyclopedia of Measurement and Statistics* SAGE Publications, Inc. (2007)
- [37] C. Song, T. Koren, P. Wang, A.-L. Barabási *Modelling the scaling properties of human mobility* Nature Physics volume 6, pages 818–823 (2010)
- [38] D. Helbing *Traffic and related self-driven many-particle systems*, Reviews of modern physics 73, 1067 (2001)
- [39] P. Sen, B. K. Chakrabarti *Sociophysics: an introduction*, Oxford University Press (2014)
- [40] H. Tamura, K. Nakano, M. Sengoku, S. Shinoda *On applications of graph/network theory to problems in communication systems* ECTI TRANSACTIONS ON COMPUTER AND INFORMATION TECHNOLOGY VOL.5, NO.1 (2011)
- [41] R. Tavenard, J. Faouzi, G. Vandewiele, F. Divo, G. Androz, C. Holtz, M. Payne, R. Yurchak, M. Rußwurm, K. Kolar, E. Woods *Tslearn, A Machine Learning Toolkit for Time Series Data* Journal of Machine Learning Research, 21(118), 1–6 (2020)
- [42] D. J. Watts, S. H. Strogatz *Collective dynamics of ‘small-world’ networks*, Nature 393, 440–442 (1998)
- [43] K. Wong *A Short Survey on Data Clustering Algorithms* 2015 Second International Conference on Soft Computing and Machine Intelligence (ISCMI), 64-68 (2015)
- [44] Z. Wu, L.A. Braunstein, S. Havlin and H.E. Stanley *Transport in weighted networks: Partition into super-highways and roads* Phys Rev Lett 96:148702 (2006)
- [45] L. Zhang, G. Zeng, D. Li, H. Huang, H. E. Stanley, S. Havlin *Scale-free resilience of real traffic jams* Proceedings of the National Academy of Sciences 116(18), 8673-8678 (2019)
- [46] Y. Zhou, B. Pik Lik Lau, C. Yuen, B. Tuncer, E. Wilhelm *Understanding Urban Human Mobility through Crowdsensed Data* IEEE Communications Magazine 56.11, 52-59 (2018)
- [47] Bing Maps Tile System - Bing Maps (<https://docs.microsoft.com/en-us/bingmaps/articles/bing-maps-tile-system>)
- [48] Disease Prevention Maps, Facebook Data for Good (<https://dataforgood.fb.com/tools/disease-prevention-maps/>)