Problem određivanja izbornih jedinica

Dukić, Sven

Sveučilište Josipa Jurja Strossmayera u Osijeku

Fakultet primijenjene matematike i informatike

Sveučilišni prijediplomski studij Matematika i računarstvo

Sven Dukić

Problem određivanja izbornih jedinica

Završni rad

Osijek, 2023.

Sveučilište Josipa Jurja Strossmayera u Osijeku

Fakultet primijenjene matematike i informatike

Sveučilišni prijediplomski studij Matematika i računarstvo

Sven Dukić

Problem određivanja izbornih jedinica

Završni rad

Mentor: prof. dr. sc. Kristian Sabo

Osijek, 2023.

Problem određivanja izbornih jedinica

Sažetak: U ovom je završnom radu prezentiran model za određivanja izbornih jedinica u Republici Hrvatskoj. Spektralno klasteriranje korišteno je za raspodjelu županija po izbornim jedinicama. Županije su predstavljene težinskim centrima izračunatim korištenjem Gauss-Krügerovih koordinata pripadajućih jedinica lokalne samouprave. Uvjetovano je da se izborne jedinice sastoje od županija koje međusobno graniče. D'Hondtova metoda primijenjena je za raspodjelu broja zastupničkih mjesta po izbornim jedinicama. Model raspoređuje županije u izborne jedinice na način da glasovi svih birača imaju što je više moguće jednaku težinu. S tom su svrhom izračunati indeksi koji mjere neujednačenost u težini biračkog glasa. Prezentirani model je implementiran u programskom jeziku Python. Budući da postoje brojni čimbenici koji mogu utjecati na konfiguraciju izbornih jedinica, potrebna su daljnja istraživanja kako bi se poboljšao model i njegova upotrebljivost.

Ključne riječi: izborne jedinice, težina biračkog glasa, matematički model, spektralno klasteriranje, Python

The problem of determining constituencies

Abstract: A mathematical model for the determination of constituencies in the Republic of Croatia was presented in this thesis. Spectral clustering was used to distribute counties into constituencies. Counties are represented by centroids calculated using the Gauss-Krüger coordinates of the local government units. It was required that the constituencies should consist of counties that border each other. The D'Hondt method was applied to allocate the number of seats to each constituency. The model distributes counties into constituencies in such a way that all voters' votes are as equal as possible. For this purpose, disproportionality indices that measure inequality in the weight of votes were calculated. The presented model is implemented in the Python programming language. Since there are many factors that can affect the configuration of constituencies, further research is needed to improve the model and its usability.

Keywords: constituencies, weight of votes, mathematical model, spectral clustering, Python

Sadržaj

1. Uvod

U Republici Hrvatskoj se prema važećoj zakonskoj regulativi zastupnici biraju u jednodomni parlament na neposrednim izborima primjenom razmjernog sustava, uz mogućnost preferencijskog glasanja. Republika Hrvatska podijeljena je na 10 izbornih jedinica, a u svakoj od njih bira se po 14 zastupnika (pripadnici dijaspore i nacionalnih manjina glasaju u posebnim izbornim jedinicama). Granice i broj zastupnika u izbornim jedinicama se tijekom sedam parlamentarnih izbora, održanih od 2000. do 2020. godine, nije mijenjao, iako je uslijed migracijskih, demografskih i drugih promjena broj birača u njima sve više odstupao od propisanog raspona. S obzirom da je tako narušena ustavom zajamčena jednakost biračkog prava, a time i zakonitost samih izbora, to je pitanje postalo predmetom stalnih rasprava. Ustavni sud Republike Hrvatske još se 2010. godine očitovao o tom problemu. Budući da u međuvremenu nisu ispravljene uočene nepravilnosti koje mogu utjecati na rezultate izbora, Ustavni sud u veljači 2023. godine donosi odluku prema kojoj je Hrvatski sabor do 1. listopada 2023. godine dužan osigurati jednakost biračkog prava. U skladu s tim je podneseno više prijedloga promjena izbornih jedinica, a cilj je ovog završnog rad bio doprinijeti aktualnoj raspravi i ukazati na njihove moguće konfiguracije.

U ovom radu prezentirani model raspodjele županija Republike Hrvatske u izborne jedinice temelji se na metodi spektralnog klasteriranja. Matematički model je implementiran u programskom jeziku Python. Županije su u modelu predstavljene težinskim centrima, dobivenim na temelju Gauss-Krügerovih koordinata svih pripadajućih jedinica lokalne samouprave. Svaki težinski centar je svrstan u klaster koji označava izbornu jedinicu. U izbornim jedinicama, neovisno o njihovom broju, trebaju biti razvrstane županije koje međusobno graniče. Željeni je broj klastera jedna od ulaznih varijabli programa pomoću kojeg se pronalaze raspodjele županija u izborne jedinice s najujednačenijim težinama biračkog glasa. Na temelju toga, može se donijeti zaključak o najprimjerenijoj konfiguraciji izbornih jedinica.

Završni rad podijeljen je u sedam glavnih poglavlja. U poglavlju koje slijedi nakon uvoda objašnjen je pojam izbornog sustava i kako se ostvaruje jednakost biračkog prava, s posebnim osvrtom na modele izbora zastupnika u Hrvatski sabor i problem nejednake težine biračkog glasa u izbornim jedinicama. Sljedeće se poglavlje, koje je podijeljeno na nekoliko potpoglavlja, bavi metodom spektralnog klasteriranja s teorijskog aspekta. U tom su poglavlju pojašnjeni najvažniji pojmovi i navedeni osnovni algoritmi spektralnog klasteriranja, uključujući onaj koji je korišten u ovom radu. Pri tome su objašnjene razlike između nenormaliziranog i normaliziranog spektralnog klasteriranja, a također su navedene prednosti i nedostaci te metode. Raspodjela broja zastupničkih mjesta i mjere ujednačenosti težina biračkih glasova teme su četvrtog poglavlja. U njemu su uvedene korištene oznake te su objašnjeni pojmovi prezastupljenosti i podzastupljenosti izbornih jedinica. Osim toga, navedeni su indeksi kojima se mjeri ujednačenost težine biračkog glasa. Primjena modela za konfiguraciju izbornih jedinica opisana je u petom poglavlju. Detaljno objašnjenje svakog od koraka popraćeno je vizualnim prikazima i dijelovima programskog koda. U šestom poglavlju prezentiran je postupak određivanja primjerenog broja izbornih jedinica na temelju indeksa koji mjere kvalitetu raspodjele broja zastupničkih mjesta po izbornim jedinicama. Rad završava zaključkom u kojem su rezimirani njegovi glavni rezultati i dane smjernice za buduća istraživanja. Na kraju rada je navedena korištena literatura.

2. Izborni sustav i jednakost biračkog prava

Izborni sustav jednostavno se može definirati kao skup pravila koja propisuju kako se glasa na izborima za predstavnička tijela i kako se glasovi birača raspodjeljuju u zastupnička mjesta [7]. Tako se također naziva pravnim propisima uređen skup društvenih odnosa pri izboru kandidata u predstavnička tijela, pri čemu odluka o odabiru određene vrste izbornog sustava predstavlja jednu od najvažnijih ustavnopravnih i političkih odluka [23, 30]. Izborni sustav rezultanta je društvene stvarnosti, političkih odnosa i interesa one stranke koja ima mogućnost, samostalno ili uz podršku drugih stranaka, donositi izborne zakone [31]. Biračko pravo temeljni je element izbornog sustava koji ukazuje na stupanj političkih sloboda [20]. Ono je u modernim demokratskim državama zajamčeno ustavom, a podrazumijeva pravo svakog građana koji ispunjava određene uvjete da bira svoje predstavnike i da bude biran. Jednakost biračkog prava danas je univerzalno prihvaćeno načelo, prema kojem svakom biraču na izborima pripada isti broj glasova [17].

2.1. Modeli izbora zastupnika u Hrvatski sabor

Od stjecanja samostalnosti i neovisnosti, u Republici Hrvatskoj održano je devet parlamentarnih izbora. U tom je razdoblju izborni sustav doživio nekoliko promjena. Herceg Zeba [9] navodi da su prvi parlamentarni izbori nakon osamostaljenja održani u kolovozu 1992. godine prema kombiniranom izbornom modelu u kojem je isti broj mandata određen većinskim i razmjernim načinom. Naime, polovina zastupnika birala se s državne liste D'Hondtovom metodom, uz izborni prag od 3%, a druga polovina u jednomandatnim izbornim jedinicama, primjenom načela relativne većine. Izborna jedinica prosječno je imala oko 60 tisuća birača, ali su postojala značajna odstupanja od tog broja. Sljedeći parlamentarni izbori održani su 1995. godine, pri čemu je zadržan kombinirani izborni sustav, no za razliku od prethodnih izbora, otprilike 75% zastupnika biralo se razmjernim, a ostali principom relativne većine. Također je uveden varijabilan izborni prag, ovisno o tome izlazi li stranka samostalno na izbore ili u koaliciji s drugim strankama. Promijeni izbornog zakonodavstva ponovo se pristupilo prije parlamentarnih izbora 2000. godine. Tada je kao rezultat konsenzusa relevantnih političkih aktera usvojen razmjeran izborni sustav s 10 izbornih jedinica u kojima se bira po 14 zastupnika. Zadržana je raspodjela glasova u mandate prema D'Hondtovoj metodi, dok je izborni prag na razini izbornih jedinica postavljen na 5%. Osim toga, određeno je da broj birača u izbornim jedinicama ne smije odstupati za više od $\pm$ 5%. Zakonom je propisano i da pripadnici dijaspore i nacionalnih manjina zastupnike biraju u posebnim izbornim jedinicama, s tim da se zastupnici nacionalnih manjina biraju većinom glasova. Izvorni Zakon o izborima zastupnika u Hrvatski sabor iz 1999. godine doživio je do sada više izmjena, no one, kako ističe Herceg Zeba [9], nisu bile supstancijalne. Svi parlamentarni izbori od 2000. do danas provedeni su u skladu s tada prihvaćenim modelom.

Počevši od 2011. godine i konstituiranja sedmog saziva (broji li se od prvih višestranačkih izbora održanih 1990. godine), Hrvatski sabor ima 151 zastupnika. U 10 izbornih jedinica na području Republike Hrvatske bira se ukupno 140 zastupnika, tri zastupnika u XI. izbornoj jedinici bira hrvatska dijaspora (njihov broj se do parlamentarnih izbora 2011. godine određivao prema tzv. nefiksnoj kvoti), a osam zastupnika u XII. izbornoj jedinici biraju pripadnici nacionalnih manjina [25]. Iako se uz određene izmjene održao tijekom više izbornih ciklusa, pokazalo se da takav model nije bez nedostatka. Zbog toga, kao i važnost koju ima za cjelokupno društvo, godinama privlači pozornost kako stručne, tako i šire javnosti.

2.2. Nejednaka težina biračkog glasa u izbornim jedinicama u Republici Hrvatskoj

Prema Kasapović [15], kod nas se ni o jednoj političkoj temi nije toliko intenzivno raspravljalo kao o izbornim sustavima. Spomenuta autorica smatra da mnogo toga izrečenog i napisanog nema racionalne osnove, pri čemu izdvaja nekoliko pitanja o kojima se najviše polemizira. Prvo od njih je razmjerni izborni sustav. Prema jednima, takav sustav pogoduje dvjema najvećim političkim strankama, dok drugi smatraju da omogućava ulazak u Hrvatski sabor malim beznačajnim strankama. Na udaru kritika je i D'Hondtova metoda, kojoj se često pripisuje prevelik utjecaj na rezultate izbora. Za mnoge je također problematičan utvrđeni izborni prag, pri čemu se izdvajaju dva oprečna stajališta. Prema prvom, izborni prag treba povisiti, a prema drugom sniziti. Uz kritiku zatvorenih listi, koja je rezultirala ozakonjenjem opcijskog preferencijskog glasanja, stalni je predmet rasprava i podjela Republike Hrvatske na izborne jedinice. Kasapović se slaže da je kritika koja se odnosi na podjelu na deset izbornih jedinica jednake veličine opravdana. Naime, dosadašnja podjela ne uvažava u dovoljnoj mjeri granice povijesnih regija, koje su i danas ekonomski i kulturno međusobno povezane. Drugi je problem vezan uz promjenu broja stanovnika i birača u izbornim jedinicama, zbog čega je ozbiljno narušena jednakost kao jedno od temeljnih načela biračkog prava u demokratskim zemljama. Na to je još 2010. godine upozorio i Ustavni sud Republike Hrvatske, no tadašnja su nastojanja da se prekroje granice izbornih jedinica doživljena kao svojevrsni izborni inženjering te nisu dobila potrebnu podršku [26].

Postojanje značajnog odstupanja u broju birača različitih izbornih jedinica u kojima se bira identičan broj zastupnika jedan je od sofisticiranijih oblika izborne manipulacije kojom se narušava jednakost biračkog prava [24]. U takvim izbornim sustavima glasovi birača u pojedinim izbornim jedinicama imaju veću važnost od glasova ostalih birača. S obzirom da može utjecati na rezultate izbora, prekomjerno odstupanje u broju birača po izbornim jedinicama nije u suglasju s Ustavom. Ujednačavanjem težine biračkog glasa unapređuje se izborni postupak i doprinosi legitimnosti izbora [12]. No, usprkos upozorenju Ustavnog suda i zahtjevima političara, stručnjaka, organizacija civilnog društva i ostale zainteresirane javnosti, izborne jedinice godinama se nisu usklađivale s promjenama u broju birača uzrokovanim prirodnim kretanjima i migracijama. U tom smislu, Žugaj i Šterc [35] navode da su se najmanja i najveća izborna jedinica 2016. godine razlikovale za više od 93 tisuće birača te predlažu novi ustroj sa šest izbornih jedinica i nejednakim brojem zastupnika. Na nejednaku težinu biračkog glasa i njezin utjecaj na rezultate izbora ukazao je i Čular [4]. Prema podacima koje je iznio u svom radu, čak osam od deset izbornih jedinica izlazilo je 2016. godine izvan zakonskog okvira od $\pm$ 5%. Zbog toga je, kako pokazuje njegova simulacija, na parlamentarnim izborima, koji su održani te godine, 12 mjesta u Hrvatskom saboru pripalo drugim zastupnicima i strankama od onih kojima bi pripali da su birači po izbornim jedinicama bili ravnomjerno raspodijeljeni.

Budući da se nije pristupilo izmjeni izbornih jedinica, Ustavni sud Republike Hrvatske pokrenuo je postupak za ocjenu suglasnosti s Ustavom Zakona o izbornim jedinicama za izbor zastupnika u Zastupnički dom Hrvatskoga državnog sabora, koji je izglasan 1999. godine i od tada nije bio mijenjan, i 7. veljače 2023. godine donio odluku kojom se taj Zakon ukida. [32] Prema odluci Ustavnog suda, ukinuti Zakon prestaje važiti 1. listopada 2023. godine, a Hrvatski sabor dužan je do tog datuma osigurati provedbu ustavnog jamstva jednakog biračkog prava. U skladu s odlukom Ustavnog suda, 28. rujna 2023. godine izglasan je novi Zakon o izbornim jedinicama za izbor zastupnika u Hrvatski sabor, koji će zasigurno nastaviti izazivati polemike.

3. Metoda spektralnog klasteriranja

Strojno učenje, kao skup metoda za automatsko otkrivanje obrazaca u podacima i njihovo korištenje za predviđanje budućih ishoda, uobičajeno se dijeli na nadzirano i nenadzirano učenje [21]. Glavna je razlika između nadziranog i nenadziranog učenja u tome što nadzirano koristi labelirane (označene) podatke, a nenadzirano nelabelirane (neoznačene) podatke [8]. U nenadziranom učenju algoritam pokušava otkriti skrivene strukture ili uzorke u podacima bez predefiniranih kategorija ili oznaka. Klasteriranje pripada algoritmima nenadziranog učenja [10]. Cilj je klasteriranja grupirati podatke sa sličnim svojstvima u isti klaster [19]. Za rješavanje problema kojim se bavi ovaj rad korištena je metoda spektralnog klasteriranja. Metoda spektralnog klasteriranja povezana je s teorijom grafova jer se s podacima postupa kao da su vrhovi grafa. Stoga se na spektralno klasteriranje može gledati kao na problem particioniranja grafa [11].

3.1. Algoritmi spektralnog klasteriranja

U okviru spektralnog klasteriranja razvijeno je više algoritama. Oni se prvenstveno razlikuju po tome koriste li nenormaliziranu ili normaliziranu Laplaceovu matricu. Liu i Han [18] navode da se spektralno klasteriranje sastoji od tri glavna koraka:

- Konstruira se graf sličnosti i matrica sličnosti za sve podatke.
- Podaci se transformiraju iz višedimenzionalnog u nižedimenzionalni prostor koristeći svojstvene vektore Laplaceove matrice.
- Provede se neka od metoda klasteriranja (npr. k-means algoritam) kako bi se podaci grupirali.

U nastavku je, prema [18] i [33], objašnjeno nenormalizirano i normalizirano spektralno klasteriranje.

3.1.1. Graf sličnosti i matrica sličnosti

Označimo sa $S = \{x_1, ..., x_n\}$ u $\mathbb{R}^m$ skup podataka koje želimo grupirati u k klastera. Prvo se podaci prezentiraju u obliku neusmjerenog "grafa sličnosti" $G = (V, E)$. Skup V je skup vrhova, pri čemu vrh v_i predstavlja podatak x_i, a E predstavlja skup bridova. Primijetimo da je x_i vektor koji označava podatak, a v_i vrh koji nema svojstva. Nadalje, definiramo nenegativnu matricu $W \in \mathbb{R}^{n\times n}$ kojom opisujemo graf G. Matricu W nazivamo matricom sličnosti. Pri tome vrijedi $W = (W_{ij})_{i,j=1,...,n}$ te $W_{ij} = 0$, ako vrhovi v_i i v_j nisu povezani. Bitno je izabrati učinkovitu metodu za računanje matrice W zato što spektralnim klasteriranjem želimo svrstati vrhove u isti klaster, ako su međusobno slični, a u različite klastere, ako nisu slični. U nastavku su navedena tri načina za konstrukciju grafa G, a samim time i matrice sličnosti W:

1. **Graf k-najbližih susjeda**
   Vrhovi v_i i v_j su povezani kada je v_j među k-najbližih susjeda v_i, ili obratno, kada je v_i među k-najbližih susjeda v_j. Na taj način dobivamo neusmjeren graf.

2. **Graf ϵ-okoline**
   Vrhovi v_i i v_j su povezani ako je mjera udaljenosti $||x_i - x_j||^2 < \epsilon$. Ovaj pristup može rezultirati grafovima s nepovezanim komponentama ako ϵ nije ispravno odabran.

3. **Potpuno povezan graf**
   U ovom slučaju povezuju se svi vrhovi za koje je mjera sličnosti pozitivna. S obzirom da bi vrijednost W_{ij} trebala biti veća ako su vrhovi v_i i v_j sličniji, odnosno manja ako im sličnost nije velika, funkciju sličnosti nije uvijek jednostavno odabrati. Kao primjer se može navesti sljedeća funkcija sličnosti između vrhova v_i i v_j:

$$W_{ij} = e^{-\frac{||x_i - x_j||^2}{\sigma^2}}.$$

   O parametru σ ovisi veličina okoline svakog vrha.

3.1.2. Nenormalizirano spektralno klasteriranje

Pretpostavimo da je konstruiran graf sličnosti G te pripadajuća matrica sličnosti W. Kako je G neusmjeren graf, vrijedi $W_{ij} = W_{ji}$. Nadalje se za svaki vrh v_i izračuna njegov stupanj:

$$d_i = \sum_{j=1}^{n} W_{ij}.$$

Matrica stupnjeva D definirana je kao dijagonalna matrica za koju vrijedi $D_{ii} = d_i$. Nenormalizirana Laplaceova matrica tada je definirana izrazom:

$$L = D - W.$$

Propozicija 3.1. *Nenormalizirana Laplaceova matrica L zadovoljava sljedeća četiri svojstva:*

1. *Za proizvoljni vektor $f \in \mathbb{R}^n$ vrijedi:*

$$f^T L f = \frac{1}{2} \sum_{i,j=1}^{n} W_{ij}(f_i - f_j)^2.$$

2. *L je simetrična i pozitivno semidefinitna.*

3. *Najmanja svojstvena vrijednost od L je 0, s pripadajućim svojstvenim vektorom $\mathbb{1}$.*

4. *L ima n realnih, nenegativnih svojstvenih vrijednosti $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$.*

Dokaz tvrdnje može se pronaći u [18].

Nakon toga se određuju svojstveni vektori $f_1, f_2, \ldots, f_k$, koji su pridruženi najmanjim svojstvenim vrijednostima matrice L. Ti svojstveni vektori su reprezentanti originalnih podataka iz skupa S u nižoj dimenziji. Stoga se nad njima provodi neka od metoda klasteriranja, kao što je k-means algoritam.

Algoritam nenormaliziranog spektralnog klasteriranja može se sažeti u sljedeće korake:

1. Konstruira se graf sličnosti G, pripadajuća matrica sličnosti W i matrica stupnjeva D.
2. Izračuna se nenormalizirana Laplaceova matrica L, gdje je $L = D - W$.
3. Odredi se k svojstvenih vektora, $f_1, f_2, \ldots, f_k$, koji su pridruženi najmanjim svojstvenim vrijednostima matrice L.
4. Konstruira se matrica $F \in \mathbb{R}^{n \times k}$ čiji su stupci navedeni svojstveni vektori.
5. Svaki red matrice F čini vrh u $\mathbb{R}^k$. Vrhovi se grupiraju u k klastera primjenom neke od klasičnih metoda klasteriranja.

3.1.3. Normalizirano spektralno klasteriranje

Dvije varijante normalizirane Laplaceove matrice definirane su kako slijedi:

$$L_{sym} = D^{\frac{-1}{2}} L D^{\frac{-1}{2}} = I - D^{\frac{-1}{2}} W D^{\frac{-1}{2}},$$

$$L_{rw} = D^{-1} L = I - D^{-1} W.$$

Propozicija 3.2. *Normalizirana Laplaceova matrica L zadovoljava sljedećih šest svojstava:*

1. *Za proizvoljni vektor $f \in \mathbb{R}^n$ vrijedi:*
$$f^T L_{sym} f = \frac{1}{2} \sum_{i,j=1}^{n} W_{ij} \left(\frac{f_i}{\sqrt{d_i}} - \frac{f_j}{\sqrt{d_j}}\right)^2.$$
2. *λ je svojstvena vrijednost matrice L_{sym} sa svojstvenim vektorom u ako i samo ako je λ svojstvena vrijednost matrice L_{rw} sa svojstvenim vektorom w, tako da vrijedi $u = D^{\frac{1}{2}} w$.*
3. *λ je svojstvena vrijednost matrice L_{rw} sa svojstvenim vektorom w ako i samo ako λ i w rješavaju generalizirani svojstveni problem $Lw = \lambda D w$.*
4. *L_{sym} i L_{rw} su pozitivno semidefinitne.*
5. *Najmanja svojstvena vrijednost i za L_{sym} i za L_{rw} je 0, sa svojstvenim vektorom $D^{\frac{1}{2}}\mathbb{1}$ za matricu L_{sym} te svojstvenim vektorom $\mathbb{1}$ za matricu L_{rw}.*
6. *L_{sym} i L_{rw} imaju n realnih, nenegativnih svojstvenih vrijednosti $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$.*

Dokaz tvrdnje može se pronaći u [18].

Algoritam za normalizirano spektralno klasteriranje, pri čemu se računa Laplaceova matrica L_{sym}, provodi se u sljedećim koracima:

1. Konstruira se graf sličnosti G, pripadajuća matrica sličnosti W i matrica stupnjeva D.
2. Izračuna se normalizirana Laplaceova matrica L_{sym}, gdje je $L_{sym} = D^{\frac{-1}{2}} L D^{\frac{-1}{2}}$.

3. Odredi se k svojstvenih vektora, $f_1, f_2, \ldots, f_k$, koji su pridruženi najmanjim svojstvenim vrijednostima matrice L_{sym}.

4. Konstruira se matrica $F \in \mathbb{R}^{n\times k}$, čiji su stupci navedeni svojstveni vektori.

5. Normiraju se retci matrice F tako da je $\forall i \leq n, \sum_j F_{ij}^2 = 1$.

6. Svaki red matrice F čini točku u $\mathbb{R}^k$. Točke se grupiraju u k klastera nekom od klasičnih metoda klasteriranja.

Algoritam normaliziranog spektralnog klasteriranja za Laplaceovu matricu L_{rw} sličan je algoritmu za nenormalizirano spektralno klasteriranje:

1. Konstruira se graf sličnosti G, pripadajuća matrica sličnosti W i matrica stupnjeva D.

2. Izračuna se nenormalizirana Laplaceova matrica L, gdje je $L = D - W$.

3. Odredi se k svojstvenih vektora, $f_1, f_2, \ldots, f_k$, koji su pridruženi najmanjim svojstvenim vrijednostima generaliziranog svojstvenog problema $Lf = \lambda Df$.

4. Konstruira se matrica $F \in \mathbb{R}^{n\times k}$, čiji su stupci navedeni svojstveni vektori.

5. Svaki red matrice F čini točku u $\mathbb{R}^k$. Točke se grupiraju u k klastera nekom od klasičnih metoda klasteriranja.

3.2. Algoritam spektralnog klasteriranja za rješavanje problema određivanja izbornih jedinica

Prethodno opisani algoritmi pružaju uvid i omogućavaju razumijevanje procedure provođenja spektralnog klasteriranja. Algoritam koji je u ovom radu korišten za određivanje izbornih jedinica u Republici Hrvatskoj razvili su Ng, Jordan i Weiss [22] i donekle se razlikuje od prethodno opisanih. Za skup točaka $S = \{s_1, s_2, ..., s_n\}$ u R^m, algoritam se provodi u sljedećim koracima:

1. Konstruira se matrica sličnosti $A \in \mathbb{R}^{n\times n}$, gdje je
$$A_{ij} = e^{-\frac{||s_i - s_j||^2}{2\sigma^2}}.$$

2. Izračuna se matrica stupnjeva D i Laplaceova matrica L, gdje je $L = D^{\frac{-1}{2}} A D^{\frac{-1}{2}}$.

3. Odredi se k svojstvenih vektora, $x_1, x_2, \ldots, x_k$, koji su pridruženi najvećim svojstvenim vrijednostima matrice L (te koji su međusobno ortogonalni zbog slučaja ponavljanja svojstvenih vrijednosti).

4. Formira se matrica $X \in \mathbb{R}^{n\times k}$, čiji su stupci navedeni svojstveni vektori.

5. Formira se matrica Y normiranjem redaka matrice X, odnosno vrijedi:

$$Y_{ij} = \frac{X_{ij}}{\sqrt{\sum_j X_{ij}^2}}.$$

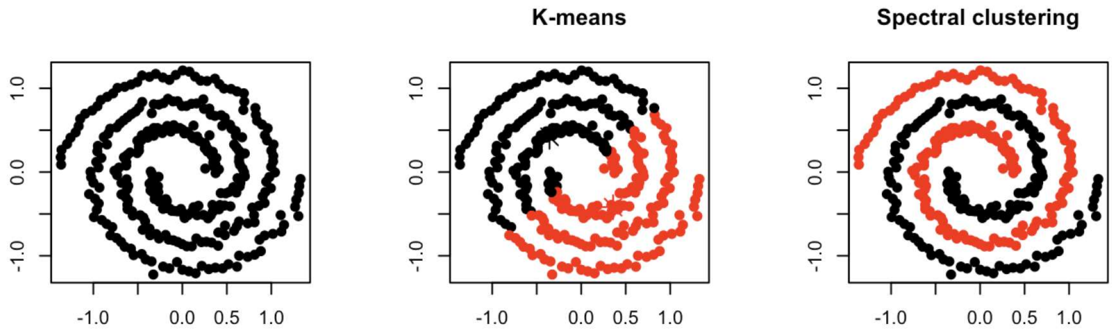
6. Svaki red matrice Y čini točku u $\mathbb{R}^k$. Točke se grupiraju u k klastera nekom od klasičnih metoda klasteriranja.

7. Pridruži se originalni podatak s_i klasteru j ako i samo ako je red i matrice Y pridružen klasteru j.

3.3. Prednosti i nedostaci spektralnog klasteriranja

Kao i svaka metoda klasteriranja, spektralno klasteriranje ima svojih prednosti i nedostataka. Neke su prednosti spektralnog klasteriranja [5, 18, 33]:

- Spektralno klasteriranje je fleksibilno u smislu uključivanja različitih vrsta mjera sličnosti.
- Spektralno klasteriranje ima dobro postavljene teorijske osnove.
- Spektralo klasteriranje je vrlo jednostavno za implementaciju i može se učinkovito riješiti standardnim metodama linearne algebre.
- S obzirom na uobičajene metode klasteriranja (poput k-means algoritma), spektralno klasteriranje se ne oslanja na pretpostavku o konveksnim klasterima (tj. oblicima klastera) ili euklidskoj udaljenosti među podacima, nego na njihovu povezanost. Razlika u klasteriranju podataka između k-meansa algoritma i spektralnog klasteriranja može se vidjeti na slici 1.



Slika 1: Razlika između k-means algoritma i spektralnog klasteriranja [16]

S druge strane, spektralno klasteriranje ima i svoje nedostatke [13, 18, 34]:

- Spektralno klasteriranje može biti računalno zahtjevno, posebno za velike skupove podataka (npr. računanje svojstvenih vrijednosti i svojstvenih vektora može postati praktički neprovedivo).

- Odabir metode za konstrukciju grafa i matrice sličnosti može biti vrlo kompleksan problem. Trenutno nema teorijske studije koja može ukazati na najbolji odabir.
- Određivanje optimalnog broja klastera može biti zahtijevan zadatak, a spektralno klasteriranje često zahtijeva specificiranje ovog parametra unaprijed ili korištenje dodatnih tehnika za njegovu procjenu.

4. Raspodjela broja zastupničkih mjesta i mjere ujednačenosti težina biračkih glasova

Različite metode koriste se za raspodjelu zastupničkih mjesta na temelju glasova birača [2, 3, 7]. U prezentiranom modelu određivanje broja zastupničkih mjesta po izbornim jedinicama provodi se na temelju D'Hondtove, odnosno Jeffersonove metode [1, 6].

Pretpostavimo da smo područje Republike Hrvatske podijelili u k izbornih jedinica. Uvodimo sljedeće oznake [27]:

- Q - ukupan broj birača u zemlji;
- S - ukupan broj zastupnika biranih iz k izbornih jedinica;
- Q_i - broj birača u i-toj izbornoj jedinici ($i = 1, ..., k$);
- S_i - broj zastupnika u i-toj izbornoj jedinici ($i = 1, ..., k$);
- $q_i = \frac{Q_i}{Q}$ - udio broja birača i-te izborne jedinice u ukupnom broju birača;
- $s_i = \frac{S_i}{S}$ - udio broja zastupnika i-te izborne jedinice u ukupnom broju zastupnika;
- $w_i = \frac{S_i}{Q_i}$ - težina biračkog glasa u izbornoj jedinici i;
- $w = \frac{S}{Q}$ - opća težina biračkog glasa.

Budući da se u izbornim jedinicama u Republici Hrvatskoj s općih lista bira 140 zastupnika, u prezentiranom modelu će S biti konstantan broj i iznositi 140. Tada vrijedi:

$$\sum_{i=1}^{k} S_i = S = 140, \quad \sum_{i=1}^{k} Q_i = Q.$$

Svaki birač ima jednaku težinu biračkog glasa ako vrijedi:

$$w_i = w, \quad \forall i = 1, ..., k. \tag{1}$$

Prema Sabo i Scitovski [27], uvjet (1) može se postići ako je cijela zemlja jedna izborna jedinica, odnosno ako je $k = 1$. No, za $k > 1$, uvjet (1) općenito neće biti ispunjen. To znači da glasovi birača iz različitih izbornih jedinica neće imati jednaku težinu, odnosno neke izborne jedinice će biti prezastupljene, a neke podzastupljene. Birač u prezastupljenoj izbornoj jedinici ima veću težinu biračkog glasa od birača iz podzastupljene izborne jedinice. Izborna jedinica i je prezastupljena ako vrijedi $w_i > w$, a podzastupljena ako je $w_i < w$. Ujednačene mjere zastupljenosti izbornih jedinica ukazuju na ujednačenost težina biračkih glasova. Dakle, problem ujednačavanja mjere zastupljenosti izbornih jedinica zapravo je ekvivalentan problemu ujednačavanja težina biračkih glasova po izbornim jedinicama.

U skladu s navedenim, cilj je razviti model koji će područje Republike Hrvatske podijeliti u k izbornih jedinica i u okviru kojeg će se odrediti broj zastupničkih mjesta za svaku izbornu jedinicu tako da što je više moguće bude ujednačena težina biračkih glasova. S tom su namjerom u nastavku navedeni indeksi kojima se mjeri kvaliteta raspodjele broja zastupničkih mjesta po izbornim jedinicama i, shodno tome, ujednačenost težine biračkog glasa. Mjere su preuzete od Karpova [14] i Sabe i Scitovskog [27], koji su indekse iz tehničkih razloga pomnožili sa 100. Potrebno je napomenuti da manja vrijednost indeksa korespondira s ujednačenijom težinom biračkih glasova.

- Indeks maksimalne devijacije:
$$MDI = 100 \max_{i=1,\ldots,k} |s_i - q_i|.$$
- Raeov indeks:
$$RAE = \frac{100}{k} \sum_{i=1}^{k} |s_i - q_i|.$$
- Loosemore-Hanbyjev indeks:
$$LHI = 50 \sum_{i=1}^{k} |s_i - q_i|.$$
- Grofmanov indeks:
$$GRO = \frac{100}{\sum_{i=1}^{k} q_i^2} \sum_{i=1}^{k} |s_i - q_i|.$$
- Gallagherov indeks:
$$GAI = 100 \sqrt{\frac{1}{2} \sum_{i=1}^{k} (s_i - q_i)^2}.$$
- Monroev indeks:
$$MI = 100 \sqrt{\frac{\sum_{i=1}^{k} (s_i - q_i)^2}{1 + \sum_{i=1}^{k} q_i^2}}.$$
- Gatev indeks:
$$GATI = 100 \sqrt{\frac{\sum_{i=1}^{k} (s_i - q_i)^2}{\sum_{i=1}^{k} (s_i^2 + q_i^2)}}.$$
- Ryabtsev indeks:
$$RYAI = 100 \sqrt{\frac{\sum_{i=1}^{k} (s_i - q_i)^2}{\sum_{i=1}^{k} (s_i + q_i)^2}}.$$
- Szalaiev indeks:
$$SZI = 100 \sqrt{\frac{\sum_{i=1}^{k} (\frac{s_i - q_i}{s_i + q_i})^2}{k}}.$$

- Aleskerov-Platonov indeks:

$$API = \frac{100}{k} \sum_{i \in I_0} \frac{s_i}{q_i}, \quad I_0 = \{i : w_i > w\}.$$

U modelu se za definiranje izbornih jedinica, odnosno analizu dobivenog rješenja koristi i indeks maksimalnog postotnog odstupanja težina biračkih glasova:

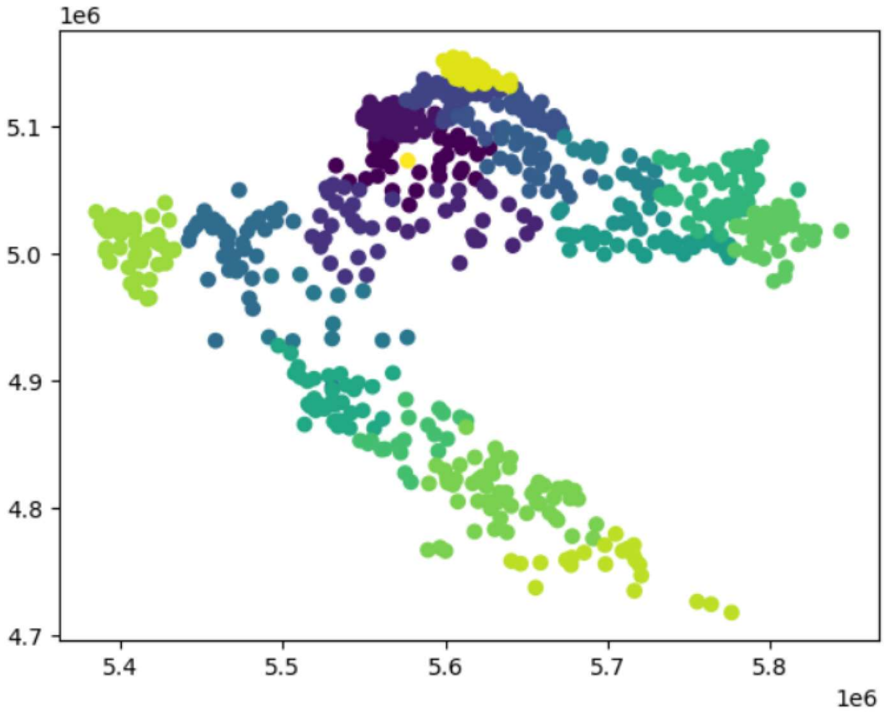
$$MPO = 100 \frac{\max\limits_{i=1,\ldots,k} w_i - \min\limits_{i=1,\ldots,k} w_i}{\min\limits_{i=1,\ldots,k} w_i}.$$

MPO predstavlja relativnu razliku između najveće i najmanje težine biračkog glasa. Ako je $p\%$ unaprijed zadana tolerancija, odnosno zakonski propisani raspon, i ako je MPO manji od $p\%$, u svake će se dvije izborne jedinice težine biračkih glasova međusobno razlikovati za najviše $p\%$ [27].

5. Primjena modela za konfiguraciju izbornih jedinica

Prezentirano rješenje temelji se na prethodnim istraživanja [27, 28]. Matematički model razvijen je u interaktivnom računalnom okruženju Jupyter Notebook, u programskom jeziku Python. Podaci koji se učitavaju u programski kod (podaci o koordinatama svih jedinica lokalne samouprave Republike Hrvatske te broju punoljetnih građana i ukupnom broju stanovnika po županijama) pohranjeni su u Excel datotekama. Metodom spektralnog klasteriranja želimo podijeliti područje Republike Hrvatske u k izbornih jedinica.

Raspolažemo Gauss-Krügerovim koordinatama [29] svih jedinica lokalne samouprave Republike Hrvatske, kojih ima 556. Na slici 2 možemo vidjeti njihov grafički prikaz.



Slika 2: Grafički prikaz koordinata svih jedinica lokalne samouprave Republike Hrvatske

Za početak odredimo skup $\mathcal{A} = \{a_i, i = 1, ..., 21\} \subset \mathbb{R}^2$, čiji su elementi težinski centri županija. Oni predstavljaju težinske centroide svih jedinica lokalne samouprave unutar županije. Pri tome se za težine uzima broj stanovnika sukladno popisu stanovništva iz 2021. godine. Funkciju u Pythonu kojom određujemo težinski centroid županije i definiramo na sljedeći način:

```
def tezcent(i):
    x = 0;
    y = 0
    w = 0;
    for j in range(len(rad)):
        if(rad["Broj zupanije"][j] == i):
            x += rad["Population"][j]* rad["X"][j]
```

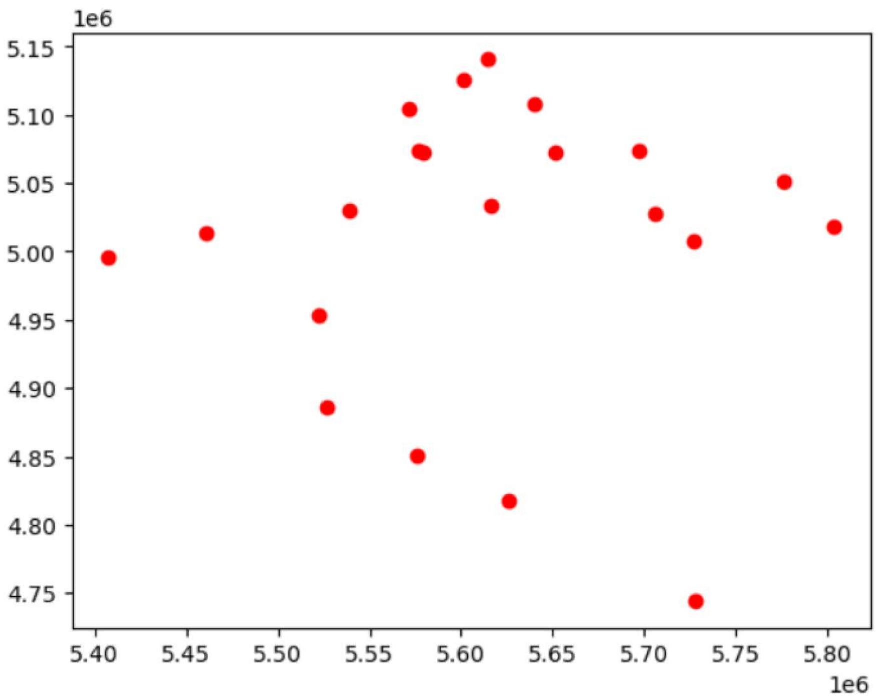
```
            y += rad["Population"][j]* rad["Y"][j]
            w += rad["Population"][j]
    x = 1/w * x
    y = 1/w * y
    return [x,y]
```

Svaki težinski centroid predstavlja jednu županiju. Cilj je primijeniti metodu spektralnog klasteriranja i smjestiti svaki težinski centroid, odnosno svaku županiju, u jednu izbornu jedinicu. Programski je kod za funkciju koja služi za spremanje svih skaliranih težinskih centroida županija:

```
def svitezcent():
    lista = []
    for i in range(1,21+1):
        centar = [round((tezcent(i)[0]/1000000),5),
                  round((tezcent(i)[1]/1000000),5)]
        lista.append(centar)
    return lista

tez_centr = svitezcent()
```

Slika 3 predstavlja grafički prikaz koordinata težinskih centroida svih županija Republike Hrvatske. Težinske centroide svih županija spremamo u listu te njihove koordinate dijelimo s 10^6 radi jednostavnijeg računanja u koracima koji slijede.



Slika 3: Grafički prikaz koordinata svih jedinica lokalne samouprave Republike Hrvatske

Nadalje, formiramo matricu sličnosti između županija $M = (m_{ij}), i, j = 1, ..., 21$, tako da je $m_{ij} = 100$, ako županija i graniči s županijom j, a $m_{ij} = 0.05$ inače.

Definiramo matricu $B(\gamma) = (b_{ij}(\gamma))$, $i, j = 1, ..., 21$, koja će služiti za izračun matrice stupnjeva i Laplaceove matrice. Za funkciju sličnosti koristi se sljedeća funkcija:

$$b_{ij}(\gamma) = \begin{cases} e^{-\gamma \frac{||x_i - x_j||^2}{m_{ij}}}, & i \neq j \\ 0, & \text{inače.} \end{cases}$$

Matrica stupnjeva $D(\gamma)$ na dijagonali i ima elemente:

$$d_i(\gamma) = \sum_{j=1}^{21} b_{ij}(\gamma).$$

U modelu je korištena metoda normaliziranog spektralnog klasteriranja i simetrična Laplaceova matrica:

$$L(\gamma) = D^{\frac{-1}{2}}(\gamma) B(\gamma) D^{\frac{-1}{2}}(\gamma)$$

Funkcije za računanje kvadratne udaljenosti između a_i i a_j te konstrukciju matrica $B(\gamma)$, $D^{\frac{-1}{2}}(\gamma)$ i $L(\gamma)$ definirane su u nastavku:

```
def distcent(x, y):
    return (x[0]-y[0])**2 + (x[1]-y[1])**2

def matrixB(matr_slicnosti , gama):
    B = np.zeros((21,21))
    for i in range(21):
        for j in range(21):
            if (i==j):
                B[i][j] = 0
            else:
                B[i][j] = np.exp(-gama * distcent(tez_centr[i],
                tez_centr[j])/matr_slicnosti[i][j])
    return B

def matrixDsqrtinv(B):
    D = B.sum(axis=1)
    Dsqrtinv = np.sqrt(1/D)
    return Dsqrtinv

def matrixL(B, Dsqrtinv):
    L = np.multiply(Dsqrtinv[np.newaxis,:],
        np.multiply(B, Dsqrtinv[:, np.newaxis]))
    return L
```

Za matricu $L(\gamma)$ odredimo k svojstvenih vektora, koji su pridruženi najvećim svojstvenim vrijednostima. Programski je kod funkcije koja vraća svojstvene vrijednosti i pripadajuće svojstvene vektore (u padajućem sortiranom poretku):

```
def svek(L):
    eigenvalues, eigenvectors = np.linalg.eig(L)
    eigenvectors = np.transpose(eigenvectors)
    eiva = list(eigenvalues)
    eiva.sort(reverse=True)
    eiei = dict(zip(eigenvalues, eigenvectors))
    eiei = {i: [eiva[i],eiei[eiva[i]]] for i in range(len(eiva))}
    return eiei
```

Potom formiramo matricu $X(\gamma) \in \mathbb{R}^{21\times k}$ tako da njene stupce čini k svojstvenih vektora. Naposlijetku, definiramo matricu $Y(\gamma) = y_{ij}(\gamma) \in \mathbb{R}^{21\times k}$, gdje je:

$$y_{ij}(\gamma) = \frac{x_{ij}(\gamma)}{||X_i(\gamma)||}.$$

Dakle, matrica $Y(\gamma)$ se dobije normiranjem redaka matrice $X(\gamma)$. Programski kod funkcija za konstrukciju matrica $X(\gamma)$ i $Y(\gamma)$ glasi:

```
def matrixX(k, eiei):
    X = []
    for i in range(k):
        X.append(eiei[i][1])
    X = np.transpose(X)
    X = [[float(val.real) for val in inn_list] for inn_list in X]
    return X

def matrixY(X):
    Y = normalize(X)
    return Y
```

Retci matrice shvaćaju se kao nova reprezentacija županija u $\mathbb{R}^k$ na koje se primjenjuje k-means algoritam. Za početne centre k-means algoritma uzima se k približno najokomitijih redaka matrice $Y(\gamma)$. Funkcije su za aproksimaciju početnih centara i metode spektralnog klasteriranja:

```
def pocetni_centri(Y, k, sth):
    z0 = [0] * k
    z0[0] = sth
    for i in range(1, k):
        Y1 = [y for y in Y if not any(np.all(y == z)
              for z in z0[:i])]
        z0[i] = Y1[0]
        sp0 = sum(np.abs(np.dot(z0[i], z0[j])) for j in range(i))
        for s in range(1, len(Y1)):
```

```
                x = Y1[s]
                if sum(np.abs(np.dot(x,z0[j])) for j in range(i))<sp0:
                    z0[i] = x
                    sp0 = sum(np.abs(np.dot(z0[i], z0[j]))
                                  for j in range(i))
    return z0


def spectralClustering(matr_slicnosti, gama, k):
    B = matrixB(matr_slicnosti, gama)
    Dsqrtinv = matrixDsqrtinv(B)
    L = matrixL(B, Dsqrtinv)
    eiei = svek(L)
    X = matrixX(k, eiei)
    Y = matrixY(X)
    kmeans_instances = []
    for i in range(21):
        z0 = Y[i]
        k_pred_sc = KMeans(n_clusters=k, n_init=1,
            init=np.array(pocetni_centri(Y,k,Y[i])),
            tol=0.0000001).fit_predict(Y)
        kmeans_instances.append(k_pred_sc)
    return kmeans_instances
```

Važno je napomenuti da se parametar $Y(\gamma) > 0$ određuje tako da županije koje se nalaze u istoj izbornoj jedinici imaju zajedničku granicu te da indeks maksimalnog postotnog odstupanja (MPO) bude što je moguće manji.

Return vrijednost funkcije *spectralClustering* je array s 21 elementom (indeks svakog elementa označava redni broj županije umanjen za 1) te je svaki element jednak broju između 0 i $k-1$, koji označava kojoj izbornoj jedinici županija pripada. Potrebna je funkcija koja na temelju oznaka (brojeva) dodijeljenim županijama stvara listu listi (u svakoj se takvoj listi nalaze županije u istoj izbornoj jedinici). Njezin je programski kod zapisan u nastavku:

```
def raspodjelaZupanija(raspodjela, k):
    rasZup = []
    for _ in range(k):
        rasZup.append([])
    for i in range(21):
        for j in range(k):
            if (raspodjela[i]==j):
                rasZup[j].append(i)
    return rasZup
```

Metodu spektralnog klasteriranja želimo provesti za više različitih vrijednosti parametara γ i odrediti optimalno rješenje. Za tu svrhu definirana je funkcija *listMSK* pomoću koje se pohranjuju rješenja metode spektralnog klasteriranja za sve vrijednosti γ iz zadanog intervala s obzirom na postavljeni korak iteracije:

```
def listMSK(matr_slicnosti,gama,stop, step, k):
    n = math.floor((stop-gama)/step)
    listaMSK = []
    for i in range(n):
        gama = gama + i*step
        listaMSK.append(spectralClustering(matr_slicnosti,gama,k))
    return listaMSK
```

Prvo, među rješenjima treba odbaciti ona koja sadrže barem jednu izbornu jedinicu u kojoj se nalaze županije koje ne graniče, odnosno koje nisu povezane. Stoga, svaku izbornu jedinicu smatramo grafom za koji treba provjeriti je li povezan. U tu se svrhu računa Laplaceova matrica grafa i Fiedlerova vrijednost. Fiedlerova vrijednost je druga najmanja svojstvena vrijednost Laplaceove matrice. Ta vrijednost je veća od 0 ako i samo ako je graf povezan. U trivijalnom slučaju, ako graf sadrži samo jedan vrh, smatramo ga povezanim. Navedena funkcija definirana je na sljedeći način:

```
def povezano(klaster, matr_slicnosti):
    D = degreeMatrix(klaster, matr_slicnosti)
    A = adjacencyMatrix(klaster, matr_slicnosti)
    Lap = D - A
    eigenvalues, eigenvectors = np.linalg.eig(Lap)
    eiva = list(eigenvalues)
    eiva.sort()
    if len(eiva) == 1:
        return True
    else:
        if eiva[1] > 1e-10:
            if np.any(np.diag(D) == 0):
                return False
            return True
        else:
            return False
```

Pomoću funkcije *povezano* definiramo funkciju *povezanostRjesenja* koja vraća sve raspodjele županija u izborne jedinice, a u kojima su županije u istoj izbornoj jedinici povezane:

```
def povezanostRjesenja(listaMSK, matr_slicnosti, k):
    dobra_raspodjela = []
    for instanca in listaMSK:
        for raspodjela in instanca:
            br = 0
            rasZup = raspodjelaZupanija(raspodjela, k)
            k = len(rasZup)
            for klaster in rasZup:
                if (povezano(klaster, matr_slicnosti)):
```

```
                br+=1
            if (br == k):
                dobra_raspodjela.append(rasZup)
    return dobra_raspodjela
```

Kako bismo odredili rješenja za koja je $MPO < \epsilon$, trebamo odrediti ukupan broj birača po izbornim jedinicama. Funkcija *izborne_jedinice_biraci* računa broj birača u svakoj izbornoj jedinici te vraća rješenje u obliku liste:

```
def izborne_jedinice_biraci(dobra_raspodjela):
    k = len(dobra_raspodjela)
    izb_jed_bi = list(np.zeros(k))
    for i, klaster in enumerate (dobra_raspodjela):
        for j in klaster:
            for z in range(21):
                if(j+1 == rad2["Broj zupanije"][z]):
                    izb_jed_bi[i]+=rad2["Punoljetni drzavljani"][z]
    return izb_jed_bi
```

Gore navedenu funkciju koristimo za funkciju *optimalnaRaspodjela*, koja od svih korektnih raspodjela (to su sve konfiguracije u kojima su u svakoj od izbornih jedinica županije povezane) vraća raspodjelu čiji je $MPO < \epsilon$, gdje ϵ predstavlja zadanu toleranciju. Funkcija *optimalnaRaspodjela*, čiji je programski kod naveden u nastavku, koristi D'Hondtovu metodu, preuzetu iz Pythonovog modula *apportionment*, za raspodjelu broja zastupnika:

```
def optimalnaRaspodjela(lista_dobri, epsilon):
    maliMPO = []
    for i in range(len(lista_dobri)):
        Q_list = np.array(izborne_jedinice_biraci(lista_dobri[i]))
        S_list = np.array(app.compute("dhondt",
            izborne_jedinice_biraci(lista_dobri[i]),S,
            list(range(1,len(lista_dobri[i])))))
        mpo = MPO(Q_list, S_list)
        if (mpo < epsilon):
            maliMPO.append([mpo, lista_dobri[i]])
    return maliMPO
```

Ako želimo izdvojiti samo najbolju konfiguraciju izbornih jedinica, a to je ona s najmanjim MPO, koristimo funkciju *najboljaRaspodjela*:

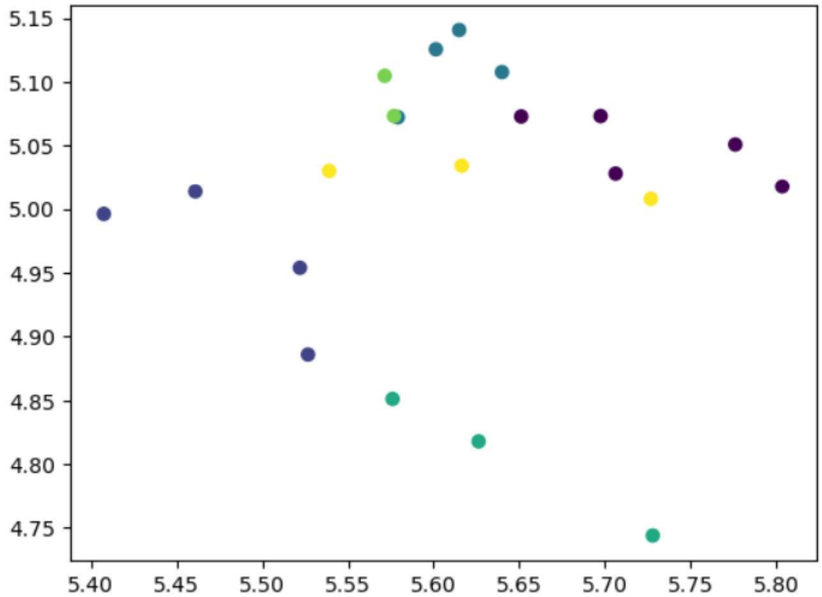
```
def najboljaRaspodjela(matr_slicnosti, gama, stop, step, k, epsilon):
    listaMSK = listMSK(matr_slicnosti, gama, stop, step, k)
    lista_dobri = povezanostRjesenja(listaMSK, matr_slicnosti, k)
    maliMPO = optimalnaRaspodjela(lista_dobri, epsilon)
    najmanjiMPO = min(maliMPO)
    return najmanjiMPO
```

Return vrijednost funkcije *najboljaRaspodjela* je lista čija je prva vrijednost MPO, a druga vrijednost lista listi koja sadrži raspodjelu županija po izbornim jedinicama. Budući da se vrijednosti kreću od 0, želimo li vidjeti kojoj izbornoj jedinici pripada npr. Osječko-baranjska županija, koja je 14. po redu županija, trebamo tražiti listu u kojoj se nalazi broj 13, jer je $13+1=14$. Sljedeći odsječak programskog koda prikazuje najbolju konfiguraciju (onu za koju je MPO najmanji) za $k=6$ s obzirom na sve γ iz intervala $[0,4]$ s korakom 0.05:

```
najbolja = najboljaRaspodjela(matr_slicnosti,0,4,0.05,6,5)
print(najbolja)

OUTPUT:
[2.1018197231468116,
    [[6, 9, 10, 13, 15], [7, 8, 12, 17], [0, 4, 5, 19],
    [14, 16, 18], [1, 20], [2, 3, 11]]]
```

Optimalna raspodjela za $k=6$ prikazana je slikom 4.



Slika 4: Grafički prikaz optimalne raspodjele za $k=6$

Dakle, ako se definira da postoji 6 izbornih jedinica, tada će se u prvoj izbornoj jedinici naći 7., 10., 11., 14. i 16. županija, u drugoj izbornoj jedinici 8,. 9., 13. i 18. županija, u trećoj izbornoj jedinici 1., 5., 6. i 20. županija, u četvrtoj izbornoj jedinici 15., 17. i 19. županija, u petoj izbornoj jedinici 2. i 21. županija, a u šestoj izbornoj jedinici 3., 4. i 12. županija. Unutar funkcije je izračunat i broj birača u svakoj izbornoj jedinici, kao i broj zastupničkih mjesta. Tako je u slučaju sa šest izbornih jedinica dobiveno da prva izborna jedinica ima 528176 birača (u njoj se bira 23 zastupnika), u drugoj izbornoj jedinici ima 557591 birač (u njoj se bira 24 zastupnika), u trećoj izbornoj jedinici ima 546602 birača (u njoj se bira 24 zastupnika), u četvrtoj izbornoj jedinici ima 523358 birača (u njoj se bira 23 zastupnika), u petoj izbornoj jedinici ima 729935 birača (u njoj se bira 32 zastupnika), a u šestoj izbornoj jedinici ima 319295 birača (u njoj se bira 14 zastupnika).

Također, pomoću funkcije *rjecnikIndeksa* možemo izračunati sve indekse navedene u prethodnom poglavlju (indeks maksimalne devijacije, Raeov indeks, Loosemoreo-Hanbyjev indeks, Grofmanov indeks, Gallagherov indeks, Monroev indeks, Gatev indeks, Ryabtsev indeks, Szalaiev indeks i Alesker-Platonov indeks) za najbolju raspodjelu u k izbornih jedinica.

```
indeksi = [MDI, RAE, LHI, GRO, GAI, MI, GATI, RYAI, SZI, API]

def rjecnikIndeksa(matr_slicnosti, k, indeksi):
    naj_k = najboljaRaspodjela(matr_slicnosti, 0, 10, 0.05, k, 5)
    Qlista_r = np.array(izborne_jedinice_biraci(naj_k[1]))
    Slista_r = np.array(app.compute("dhondt",
               izborne_jedinice_biraci(naj_k[1]),
               S, list(range(1,len(naj_k[1])))))
    rjecnik = {}
    for indeks in indeksi:
        rjecnik[indeks] = indeks(Qlista_r, Slista_r)
    return rjecnik
```

S obzirom da je Republika Hrvatska administrativno podijeljena u 21 županiju, u svim prethodno navedenim funkcijama koje sadrže k kao parametar, njegova vrijednost može biti bilo koji cijeli broj između 1 i 21, odnosno $k \in \{1, 2, ..., 21\}$.

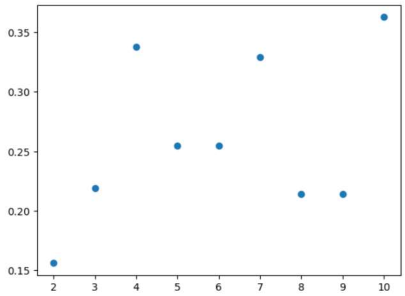
6. Određivanje primjerenog broja izbornih jedinica

Prethodno je objašnjeno na koji se način uz primjenu metode spektralnog klasteriranja županije raspodjeljuju u željeni broj izbornih jedinica. No, koliki je uopće primjereni broj izbornih jedinica? Kako bi se odgovorilo na to pitanje, prvo je potrebno odrediti optimalnu raspodjelu za k izbornih jedinica, a zatim za svaku izbornu jedinicu determinirati pripadajući broj zastupnika tako da je njihov ukupni broj jednak 140. Nakon toga se za svaki k izračunaju indeksi koji ukazuju na ujednačenost težine biračkog glasa. Funkcija *svi_indeksi* izračunava i sprema vrijednost svakog indeksa s obzirom na broj izbornih jedinica k, pri čemu *Qlista_naj* i *Slista_naj* označavaju liste u kojima se nalaze brojevi birača, odnosno brojevi zastupnika za svaku izbornu jedinicu s obizrom na k. Ispitana je primjerenost raspodjela županija u 2, 3, 4, 5, 6, 7, 8, 9 i 10 izbornih jedinica.

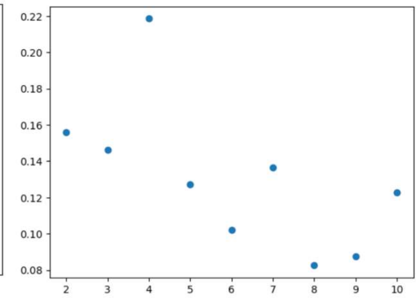
```
Qlista_naj = []
Slista_naj = []
for naj in lista_najboljih:
    Qlista_naj.append(np.array(izborne_jedinice_biraci(naj)))
    Slista_naj.append(np.array(app.compute("dhondt",
    izborne_jedinice_biraci(najbolja),S,list(range(1,len(naj))))))

def svi_indeksi(Qlista_naj, Slista_naj):
    ind_MDI, ind_RAE, ind_LHI = [], [], []
    ind_GRO, ind_GAI, ind_MI = [], [], []
    ind_GATI, ind_RYAI, ind_SZI = [], [], []
    ind_API = []
    for i in range(9):
        ind_MDI.append(MDI(Qlista_naj[i],Slista_naj[i]))
        ind_RAE.append(RAE(Qlista_naj[i],Slista_naj[i]))
        ind_LHI.append(LHI(Qlista_naj[i],Slista_naj[i]))
        ind_GRO.append(GRO(Qlista_naj[i],Slista_naj[i]))
        ind_GAI.append(GAI(Qlista_naj[i],Slista_naj[i]))
        ind_MI.append(MI(Qlista_naj[i],Slista_naj[i]))
        ind_GATI.append(GATI(Qlista_naj[i],Slista_naj[i]))
        ind_RYAI.append(RYAI(Qlista_naj[i],Slista_naj[i]))
        ind_SZI.append(SZI(Qlista_naj[i],Slista_naj[i]))
        ind_API.append(API(Qlista_naj[i],Slista_naj[i]))
    return [ind_MDI, ind_RAE, ind_LHI, ind_GRO, ind_GAI, ind_MI,
            ind_GATI, ind_RYAI, ind_SZI, ind_API]
```

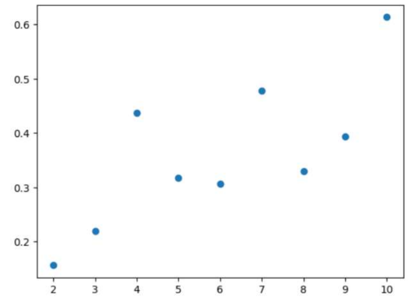
Za primjereni broj izbornih jedinica ima smisla uzeti k kod kojeg dolazi do naglog skoka vrijednosti indeksa. Na temelju sljedećih slika, koje prikazuju vrijednosti indeksa (za svaki $k = 2, ..., 10$), zaključuje se da je primjereni broj izbornih jedinica jednak 3, 6 i 9.



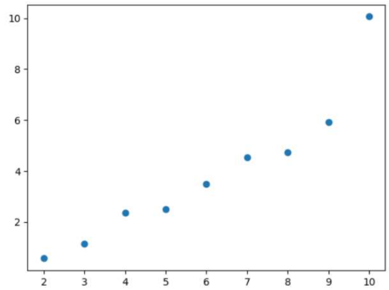
(a) MDI



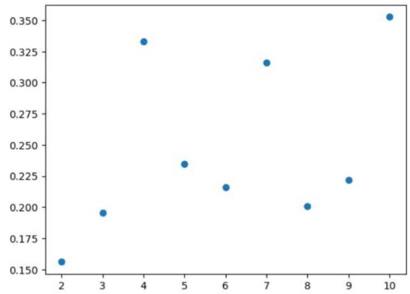
(b) RAE



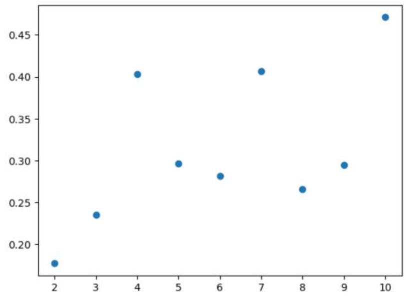
(c) LHI



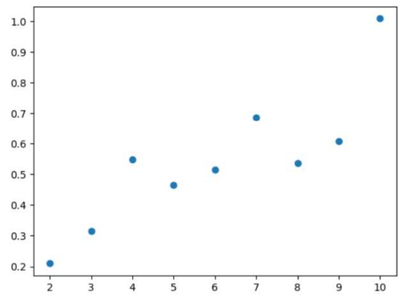
(d) GRO



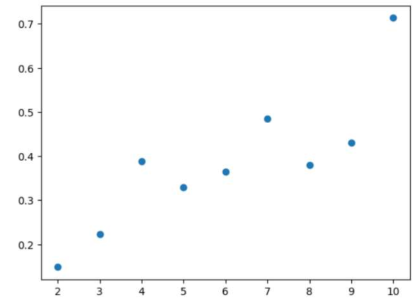
(e) GAI



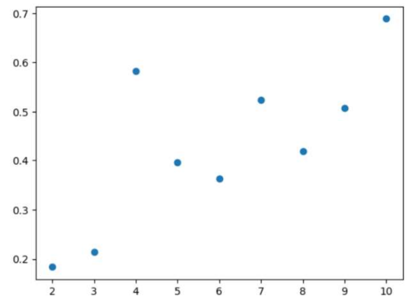
(f) MI



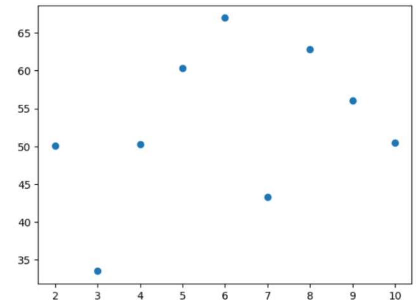
(g) GATI



(h) RYAI



(i) SZI



(j) API

Slika 5: Grafički prikazi vrijednosti indeksa za svaku optimalnu raspodjelu s obzirom na broj izbornih jedinica k

7. Zaključak

U ovom je završnom radu prikazan matematički model određivanja izbornih jedinica u Republici Hrvatskoj koji se temelji na metodi spektralnog klasteriranja. Pokazano je da model može biti primijenjen sa svrhom ujednačavanja težine biračkog glasa pronalaženjem raspodjele koja čuva granice županija, pri čemu one koje međusobno graniče smješta u istu izbornu jedinicu. Zbog toga je broj zastupničkih mjesta po izbornim jedinicama varijabilan te ovisi o pripadajućem broju birača. Definiranjem željene tolerancije, odnosno maksimalnog postotnog odstupanja osigurava se pronalaženje najprimjerenijeg rješenja s obzirom na relativnu razliku u težinama biračkih glasova. D'Hondtova metoda, koja se inače koristi za preračunavanje glasova birača u mandate, primijenjena je u modelu za određivanje broja zastupničkih mjesta po izbornim jedinicama. Implementacijom većeg broja indeksa za mjerenje ujednačenosti težine biračkog glasa i njihovom usporedbom dobiva se bolji uvid u rezultate modela. Njihove vrijednosti mogu nam pomoći pri određivanju konfiguracije izbornih jedinica koja najbolje zadovoljava naše zahtjeve. Program napisan u programskom jeziku Python omogućava generiranje rješenja do kojih bi bez računalne podrške bilo praktički nemoguće doći.

Prezentirani model uzima u obzir broj stanovnika i birača u županijama, te vodi računa o tome graniče li županije ili ne. Budući da njihovo utvrđivanje izlazi izvan okvira ovog završnog rada, u modelu su zbog praktičnih razloga zanemareni drugi čimbenici koji mogu utjecati na konfiguraciju izbornih jedinica. No, sa svrhom unaprijeđenja modela, u njega bi u budućim istraživanjima trebalo uključiti čimbenike kao što su prometna, gospodarska, povijesna i kulturna povezanost županija. U tom je smislu poželjno ispitati ne samo stavove stručnjaka, nego i stanovnika iz različitih krajeva Republike Hrvatske. Na taj način zasigurno bi bilo moguće napraviti korektniju raspodjelu birača po izbornim jedinicama. Neovisno o nedostacima, model ukazuje na moguća rješanja problema nejednakosti težine biračkog glasa.

Literatura

[1] Balinski, M. L., & Young, H. P. (1975). The quota method of apportionment. *The American Mathematical Monthly, 82*(7), 701–730. https://doi.org/10.1080/00029890.1975.11993911

[2] Balinski, M. L., & Young, H. P. (2001). *Fair representation: Meeting the ideal of one man, one vote.* (2nd ed.). Washington, DC: Brookings Institution Press.

[3] Cortona, P.G., Manzi, C., Pennisi, A., Rocca, F., & Simeone, B. (1999). *Evaluation and optimization of electoral systems.* Philadelphia, PA: SIAM.

[4] Čular, G. (2018). Metodološki izazovi ustavnog sudovanja: učinci podjele na izborne jedinice na rezultate izbora u Hrvatskoj 2000-2016. *Anali Hrvatskog politološkog društva, 15*(1), 7-28. https://doi.org/10.20901/an.15.01

[5] Fouedjio, F. (2017). A spectral clustering approach for multivariate geostatistical data. *International Journal of Data Science and Analytics, 4*(4), 301-312. https://doi.org/10.1007/s41060-017-0069-7

[6] Gallagher, M. (1992). Comparing proportional representation electoral systems: Quotas, thresholds, paradoxes and majorities. *British Journal of Political Science, 22*(4), 469–496. https://doi.org/10.1017/S0007123400006499

[7] Gallagher, M., & Mitchell, P. (2005). Introduction to electoral systems. In M. Gallagher, & P. Mitchell (Eds.), *The politics of electoral systems* (pp. 3-23). Oxford: Oxford University Press.

[8] Greene, C. S., Tan, J., Ung, M., Moore, J. H., & Cheng, C. (2014). Big data bioinformatics. *Journal of Cellular Physiology, 229*(12), 1896-1900. https://doi.org/10.1002/jcp.24662

[9] Herceg Zeba, J. (2016). Izmjene hrvatskoga izbornog zakonodavstva iz 2014. godine – put prema većoj kvaliteti izbornog sustava?. *Poslovna izvrsnost, 10*(1), 121-150.

[10] Huo, Z., Mei, G., Casolla, G., & Giampaolo, F. (2020). Designing an efficient parallel spectral clustering algorithm on multi-core processors in Julia. *Journal of Parallel and Distributed Computing, 138*, 211-221. https://doi.org/10.1016/j.jpdc.2020.01.003

[11] Jia, H., Ding, S., Xu, X., & Nie, R. (2014). The latest research progress on spectral clustering. *Neural Computing and Applications, 24*(7-8), 1477-1486. https://doi.org/10.1007/s00521-013-1439-2

[12] Jozić-Ileković, A. (2012). Kako unaprijediti izbornu administraciju u Hrvatskoj?. *Sveske za javno pravo, 3*(8), 27-37.

[13] Kanavos, A., Karamitsos, I., & Mohasseb, A. (2023). Exploring clustering techniques for analyzing user engagement patterns in Twitter data. *Computers, 12*(6), Article 124. https://doi.org/10.3390/computers12060124

[14] Karpov, A. (2008). Measurement of disproportionality in proportional representation systems. *Mathematical and Computer Modelling, 48*(9-10), 1421–1438. https://doi.org/10.1016/j.mcm.2008.05.027

[15] Kasapović, M. (2017). Jesu li izborni sustavi sredstva dramatična utjecaja na sudbine zemalja?. *Političke analize: tromjesečnik za hrvatsku i međunarodnu politiku, 8*(32), 17-21.

[16] Keerthana, V. (2022). What, why and how of Spectral Clustering! Preuzeto 25.9.2023. s https://www.analyticsvidhya.com/blog/2021/05/what-why-and-how-of-spectral-clustering/

[17] Lauc, Z. (2011). Znanstveno mišljenje o ustavnosti Ustavnog zakona o pravima nacionalnih manjina i izbornog zakona. *Političke analize: tromjesečnik za hrvatsku i međunarodnu politiku, 2*(8), 59-66.

[18] Liu, J., & Han, J. (2014). Spectral clustering. In C. C. Aggarwal, & C. K. Reddy (Eds.), *Data clustering: Algorithms and applications* (pp. 177-199). Boca Raton, FL: CRC Press.

[19] Ma, E. W. M., & Chow, T. W. S. (2004). A new shifting grid clustering algorithm. *Pattern Recognition, 37*(3), 503-514. https://doi.org/10.1016/j.patcog.2003.08.014

[20] Milušić, A. (2000). Temeljna obilježja izbornog sustava za Hrvatski sabor u njegovome izbornom redu iz 1918. godine. *Pravni vjesnik, 16*(1-2), 123-150.

[21] Murphy, K. P. (2012). *Machine learning: A probabilistic perspective.* Cambridge, MA: MIT Press.

[22] Ng, A. Y., Jordan, M. I. & Weiss, Y. (2002). On spectral clustering: Analysis and an algorithm. In T. G. Dietterich, S. Becker, & Z. Ghahramani (Eds.), *Advances in neural information processing systems 14: Proceedings of the 2001 Conference* (Vol. 2, pp. 849-856). Cambridge, MA: MIT Press.

[23] Palić, M. (2012). Učinci primjene razmjernog izbornog sustava u Republici Hrvatskoj. *Zbornik radova Pravnog fakulteta u Splitu, 49*(1), 49-58.

[24] Palić, M., & Rac, G. (2011). Nejednako biračko pravo u Republici Hrvatskoj. *Pravni vjesnik, 27*(3-4), 333-342.

[25] Petković, K. (2011). Prilog: Rezultati izbora za Hrvatski sabor od 4. prosinca 2011. *Političke analize: tromjesečnik za hrvatsku i međunarodnu politiku, 2*(8), 26-32.

[26] Podolnjak, R. (2013). Suvremeni hrvatski izborni inženjering kao sofisticirani oblik izborne manipulacije. *Zbornik Pravnog fakulteta u Zagrebu, 63*(1), 155-187.

[27] Sabo, K., & Scitovski, R. (2023). *Nova metoda za definiranje izbornih jedinica u Hrvatskoj.* Neobjavljen manuskript.

[28] Sabo, K., Scitovski, R., & Taler, P. (2012). Ravnomjerna raspodjela broja birača po izbornim jedinicama temeljem matematičkog modela. *Hrvatska i komparativna javna uprava, 12*(1), 229-249.

[29] Scitovski, R., & Sabo, K. (2020). *Klaster analiza i prepoznavanje geometrijskih objekata.* Osijek: Sveučilište Josipa Jurja Strossmayera u Osijeku, Odjel za matematiku.

[30] Smerdel, B., & Sokol, S. (2006). *Ustavno pravo.* Zagreb: Pravni fakultet Sveučilišta u Zagrebu.

[31] Šiber, I. (2000). Politički marketing i politički sustav. *Politička misao, 37*(2), 149-167.

[32] Ustavni sud Republike Hrvatske (2023). *Odluka Ustavnog suda Republike Hrvatske broj: U-I-4089/2020 i dr. od 7. veljače 2023. i šest izdvojenih mišljenja sudaca.* Narodne novine 24/2023.

[33] Von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and Computing, 17*(4), 395-416. https://doi.org/10.1007/s11222-007-9033-z

[34] Wang, X., Zheng, X., Qin, F., & Zhao, B. (2013). A fast spectral clustering method based on growing vector quantization for large data sets. In H. Motoda, Z. Wu, L. Cao, O. Zaiane, M. Yao. & W. Wang (Eds.), *Proceedings of the 9th International Conference on Advanced Data Mining and Applications – ADMA 2013* (Vol. 2, pp. 25-33). Heidelberg: Springer. https://doi.org/10.1007/978-3-642-53917-6_3

[35] Žugaj, M., & Šterc, S. (2016). Hrvatske izborne jedinice – postojeći nesklad i buduće promjene. *Pilar: časopis za društvene i humanističke studije, 11*(22(2)), 9-33.