



Kaip balsuoja Seimas?

2008-2014 m. LR Seimo balsavimų vizualizavimas ir analizė

Projekto „Mokslo pieva“ ataskaita

Komandos vadovai:

Tomas Krilavičius

Vaidas Morkevičius

Tyrimą atliko:

Vytautas Mickevičius

Greta Ciganaitė

Izabelė Nebilevičiūtė

Vilnius, Kaunas, 2014

Kas mes esame?

Vytautas Mickevičius – Vytauto Didžiojo universiteto taikomosios matematikos 1 kurso magistrantas.

Greta Ciganaitė – Vytauto Didžiojo universiteto matematikos ir jos taikymo programos (specializacija: finansų ir draudos matematika) bakalauro 3 kurso ir informatikos gretutinių studijų studentė.

Izabelė Nebilevičiūtė – Tarptautinių santykių ir politikos mokslų instituto bakalauro politikos mokslų programos 3 kurso studentė.

Tomas Krilavičius – Vytauto Didžiojo universiteto profesorius, Baltijos pažangių technologijų instituto vyresnysis mokslo darbuotojas.

Vaidas Morkevičius – Kauno technologijos universiteto Viešosios politikos ir administravimo instituto vyresnysis mokslo darbuotojas.

Ką mes darome?

Mūsų projekto tikslas – sukurti interneto puslapį, kuriame būtų galima aiškiai ir suprantamai matyti Lietuvos Respublikos Seimo narių balsavimų tendencijas.

Kuo mes skiriamės nuo kitų?

Jau egzistuoja keletas interneto puslapių, skirtų LR Seimo stebėsenai (pvz., www.manoseimas.lt, www.seime.lt). Tad natūraliai gali kilti klausimas – o kuo gi mūsų išskirtinis?

Pagrindinis mūsų puslapio skiriamasis bruožas – bendrųjų viso Seimo balsavimo tendencijų analizė. Panorėję sužinoti, kurie Seimo nariai stropiausiai vaikšto į posėdžius, ar daugiausia kartų balsuoja „Prieš“ („Už“), galite tai greitai sužinoti interneto svetainėje www.seime.lt, o jei norite išsiaiškinti kuriuo konkrečiu klausimu ir kaip balsuoja konkretūs Seimo nariai, galite pasinaudoti www.manoseimas.lt puslapiu internete. Mes akcentuojame bendrųjų balsavimų tendencijų analizę, kitaip tariant – nagrinėjame Seimo narių panašumus ir skirtumus pagal jų balsavimo daugeliu klausimų rezultatus. Nedarome subjektyvių išvadų, kad vienas ar kitas parlamentaras yra „geras“ ar „blogas“, balsuoja „teisingai“ ar „neteisingai“ ir pan. Tačiau galime pasakyti, kad Seimo nario A balsavimo rezultatai daug panašesni į Seimo nario B, negu į Seimo nario C, nors A ir C priklauso vienai frakcijai, o B – kitai. Galime pasakyti, kurie parlamentarai balsuoja skirtingiausiai nuo savos frakcijos, o kurie jai lojaliausi. Taip pat galime pasakyti, kurie Seimo nariai balsuoja labai panašiai į pozicijos atstovus, nors priklauso opozicijai.

Bendrųjų parlamentarų balsavimo tendencijų analizė yra gerokai sudėtingesnė, nei paprastų suvestinių, kiek kartų balsuota „Už“, kiek – „Prieš“ ir t.t., sudarymas, nes jai atlikti būtina taikyti matematinis-statistinius metodus.

Kodėl neužtenka paprastų balsavimų suvestinių?

Iš pirmo žvilgsnio gali pasirodyti, jog du parlamentarai, 5 kartus balsavę už, 3 kartus susilaikę ir 2 kartus balsavę prieš, balsuoja identiškai. Bet taip yra tik iš pirmo žvilgsnio. Šie parlamentarai išties balsuotų identiškai, jeigu balsuotų lygiai taip pat kiekviename konkrečiame balsavime. Analogiškai šiuo atveju galima ir tokia situacija, kuomet nesutaptų nei vienas dviejų parlamentarų balsavimo rezultatas. Kitaip tariant –

svarbu ne tik kiek kartų yra balsuota vienaip ar kitaip, bet ir kuriuose konkrečiuose balsavimuose. Į šiuos du veiksnius būtina atsižvelgti, siekiant parodyti, kaip atrodo tam tikro vieno Seimo nario balsavimai, lyginant su visų kitų Seimo narių balsavimais.

Kaip mes nagrinėjame bendrąsias LR Seimo balsavimo tendencijas?

Kadangi LR Seime balsuojama labai dažnai (daugiau nei tūkstantį kartų per metus), tokiam kiekiui informacijos apdoroti turime taikyti matematinius-statistinius metodus. Šie metodai kiek sudėtingi suprasti jų neišmanančiam žmogui, tačiau vienas šio teksto uždavinių yra įmanomai paprastai paaiškinti taikomų metodų principus ir rezultatų interpretavimo galimybes.

Kokius metodus naudojame?

Pagrindinis mūsų naudojamas metodas yra klasikinė daugiamačių skalių analizė (angl. multidimensional scaling, sutrumpintai MDS). Tai duomenų struktūrizavimo metodikų grupė, skirta atskleisti latentinę („nematomą“) duomenų struktūrą (plačiau žr. www.lidata.eu/mokymai/metodologiniai/SA). Savo esme ji panaši į daugelį kitų matematinių duomenų struktūrizavimo metodikų grupių (pavyzdžiui, faktorinę analizę). Pagrindinis MDS uždavinys – atvaizduoti objektus kuo mažesnio rango mažamateje erdvėje taip, kad artumai tarp objektų, suskaičiuoti pagal kokius nors objektų požymius, kuo labiau atitiktų atstumus tarp jų toje erdvėje. Mūsų atveju objektai yra LR Seimo nariai, jų požymiai pagal kuriuos skaičiuojame jų artumus – balsavimai dėl teisės aktų (taigi suskaičiuojame šiuos artumus gauname LR Seimo narių balsavimo panašumų matricą), o mūsų analizės tikslas – kelių matavimų (2 ar 3) euklidinėje erdvėje atvaizduoti tuos balsavimo panašumus taip, kad atstumai gautame grafike kuo labiau atitiktų suskaičiuotus panašumus.

LR Seimo balsavimams analizuoti taip pat taikome klasterizavimą k-means metodu. Tai kitas duomenų struktūrizavimo metodas, kurio tikslas – pagal objektų artumus suklasifikuoti juos į tam tikrą skaičių grupių, vadinamų klasteriais (plačiau žr. www.lidata.eu/mokymai/metodologiniai/SA). Trumpai galima aprašyti k-means algoritmą keliais nesudėtingais etapais:

1. Nustatyti klasterių skaičių. Šis veiksmas atliekamas dar prieš klasterizavimo procedūrą. Šiuo atveju, numatytasis klasterių skaičius yra frakcijų skaičius laikotarpyje.
2. Nustatyti centrus. Centrai nustatomi atsitiktinai ir vėliau tikslinami algoritmo vykdymo metu.
3. Nustatyti atstumus tarp centrų ir kiekvieno objekto. Atstumo matavimo būdą galima pasirinkti (mūsų atveju: Euklido arba Manheteno). Pasirinktas būdas bus taikomas visuose atstumo skaičiavimuose.
4. Priskirti likusius objektus prie artimiausio nustatyto centro.
5. Nustato atstumus tarp centro ir klasteryje esančių objektų
6. Nustatomi nauji centrai. Naujas centras, tai toks klasterio objektas, nuo kurio atstumų su kitais objektais suma yra mažiausia.
7. Perskaičiuojami atstumai tarp objektų ir naujų klasterių centrų.
8. Perrenkami objektai prie tų centrų, kur atstumas yra trumpiausias.

9. Algoritmas nustoja veikti, jei centrai po perrinkimo lieka tie patys ir visi objektai yra priskirti prie artimiausio centro, kitu atveju pradedama nuo penkto punkto.

Vėlgi, mūsų atveju objektai yra LR Seimo nariai, o požymiai, pagal kuriuos skaičiuojame jų artumus – balsavimai dėl teisės aktų. Taigi, k-means klasterizavimo pagalba LR Seimo narius galima suskirstyti į tam tikrą kiekį grupių pagal balsavimo panašumą. Mūsų analizėje šis metodas taikomas siekiant atskleisti, ar pagal balsavimo panašumus statistiniais metodais sudarytos parlamentarų grupės atitinka LR Seimo narių politines institucines pozicijas: priklausymą frakcijoms ir pozicijai, arba opozicijai.

Kaip sugebame sutalpinti 300-700 tūkst. balsavimų rezultatų į vieną grafiką?

Būtent tokiems atvejams ir yra skirtas daugiamačių skalių metodas. Žiūrint techniškai, LR Seimo nariai yra objektai, kuriuos aprašo n kintamųjų, kur n – balsavimų skaičius. Priklausomai nuo laikotarpio, šis skaičius gali siekti 5000 ir daugiau balsavimų. Taigi, mūsų analizėje parlamentaras yra taškas n -matėje erdvėje. Nors apie erdves, turinčias daugiau nei 3 matavimus, labiausiai mėgsta kalbėti teorinės fizikos (arba mokslinės fantastikos) atstovai, tačiau mūsų atveju tai tiesiog reiškia, jog taško padėtį erdvėje apibūdina n koordinatė. Ko gero visi prisimena dar vidurinėje mokykloje ant lentos pieštus taškus plokštumoje, kurių koordinatės (-3; 1) ar (14; 17). Tas pats ir trimatėje plokštumoje, tik tašką apibūdina 3 koordinatės, tarkim (-2; 58; 0). Mūsų atveju vieną tašką apibūdina daugybė koordinatė: (-1; 0; -2; 2; ...; 0; 1; 0). Šios koordinatės yra vieno LR Seimo nario balsavimų rezultatai – kiekviena skirtinga skaitinė reikšmė žymi skirtingą balsavimo baigmę, tarkime, 2 – „Už“, 1 – „Neatvyko į posėdį“, -1 – „Susilaikė“, -2 – „Prieš“, ir t.t.

Naudodami daugiamačių skalių analizę mes siekiame šį didžiulį koordinatė kiekį sumažinti iki 2 ar 3, t. y. gauti paprastą dvimatį ar trimatį grafiką, kurio erdvėje LR Seimo nariai būtų išsidėstę pagal balsavimo panašumus, kuo tiksliau išlaikant pradinius artumus tarp taškų (LR Seimo narių).

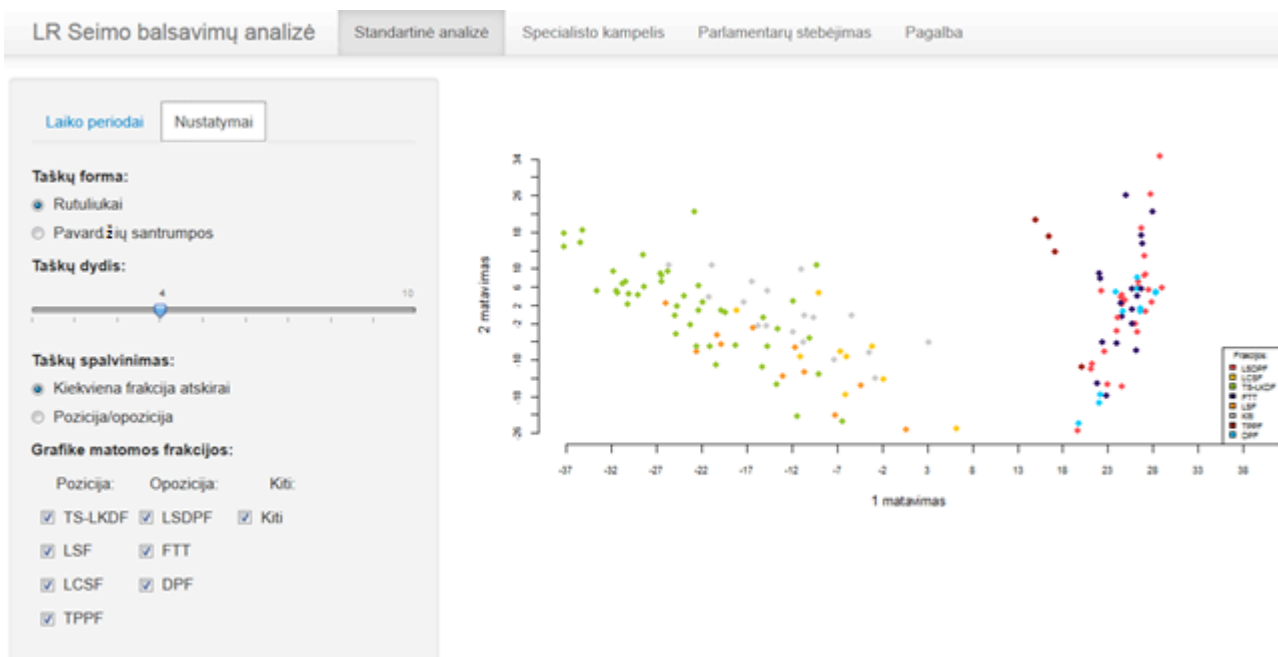
Ar yra skirtumas, kokiomis reikšmėmis koduojame balsavimus?

Taip. Skirtingai koduodami balsavimų baigmes (t. y. priskirdami joms kitokias skaitines reikšmes), gauname skirtingus daugiamačių skalių analizės rezultatus. Kita vertus, kodavimo schemos pasirinkimas yra svarbus ir turinio prasme, nes tam tikros skirtingos baigmės gali būti užkoduotos vienodomis skaitinėmis reikšmėmis – tada tariame, kad pagal prasmę šios baigmės yra tapačios. Pavyzdžiui, dažnai balsuojant LR Seime teisės aktui priimti reikalinga posėdyje dalyvaujančių parlamentarų balsų dauguma. Tai reiškia, kad balsų „Už“ dalis turi būti didesnė, nei balsų „Prieš“ ir „Susilaikė“ kartu sudėjus. Kitaip tariant, visiškai nesvarbu, ar LR Seimo narys balsuoja „Prieš“, ar „Susilaiko“ – abiem atvejais jis išreiškia nepritarimą teisės akto projektui. (galioja principas: ne „Už“ = „Prieš“). Standartinis balsavimų kodavimas, taikomas daugelyje tokio pobūdžio tyrimų, yra toks, kuomet baigmei „Už“ suteikiama reikšmė 1, o visoms kitoms – 0. Jei vartotojas ar tyrėjas mano, kad toks balsavimų kodavimo būdas nėra tinkamas, statistiko skiltyje galima sudaryti savo kodavimo schemą, tik reikia nepamiršti, kad tokia schema turi atitikti kiekybinės, tvarkos ar dvireikšmės skalės sudarymo principus (kitais sudarytomis skalėmis užkoduotiems duomenims nebus galima taikyti daugiamačių skalių analizės).

Kodėl nagrinėjame 5 atskirus laikotarpius, o ne vieną bendrą?

Visų pirma, LR Seimo sudėtis nėra pastovi. Pavyzdžiui, mandato atsisakiusius parlamentarų pakeičia kiti (kai kurie išeina ir lieka nepakeisti). Kita vertus, yra „perbėgėlių“ Seimo narių, kurie iš vienos frakcijos

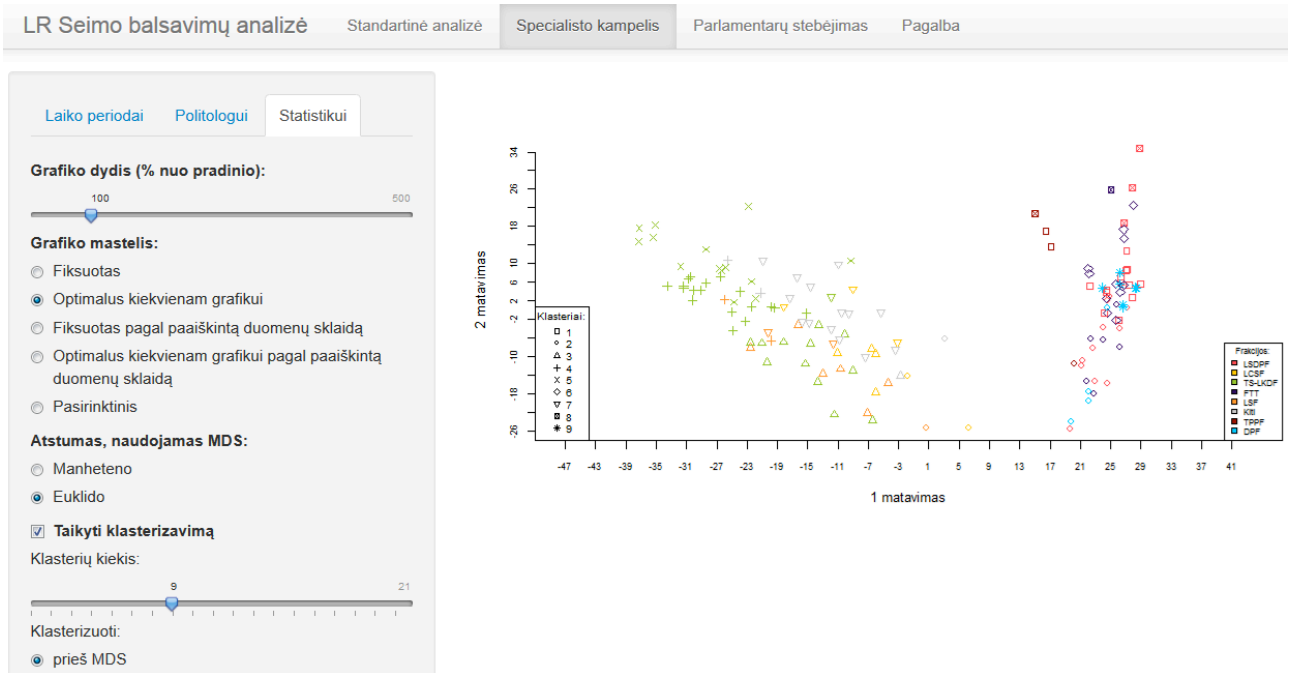
„perbėga“ į kitas ir atitinkamai pasikeičia jų balsavimo pasirinkimai. Tačiau svarbiausi balsavimo pokyčiai Seime vyksta tada, kai skyla (arba jungiasi) frakcijos arba ištisos frakcijos pereina iš pozicijos į opoziciją. Šie pokyčiai yra susiję su radikaliais ir plataus masto parlamentarų balsavimo pokyčiais, kurie statistinės analizės metu apimant visos kadencijos balsavimus „susiniveliuotų“ ir nebūtų matomi rezultatų grafikuose. Žinoma, tokia situacija nėra pageidautina. Todėl ties reikšmingais frakcijų skilimo ir jungimosi epizodais, bei momentais, kai ištisos frakcijos pereina iš pozicijos į opoziciją, balsavimai yra padalinti į atskirus laikotarpius, taip suteikiant galimybę rezultatų grafikuose stebėti parlamentarų balsavimo pokyčius, susijusius su šiais pasikeitimais.



1 pav. Puslapio vaizdo pavyzdys (screenshot). Standartinė analizė.

Kokių žinių reikia, norint naudotis mūsų puslapiu?

Puslapį sudaro trys pagrindinės dalys, kuriose galima pasirinkti skirtingus nustatymus ir gauti kiek kitokius rezultatus. Pirmoji dalis (1 pav.) skirta standartiniams rezultatams ir jų vaizdavimui (čia galima pasirinkti objektų atvaizdavimo ir spalvinimo būdus). Antroji dalis labiau skirta politologams, t. y. vartotojams, kurie turi daugiau politologinių žinių ir nori gauti atsakymus į tam tikrus specifinius klausimus (čia, be standartinėje analizėje naudojamų parinkčių, vartotojai gali pasirinkti atvaizduojamų balsavimų aibę, bei „perbėgėlių“ klasifikavimą). Trečioji dalis (2 pav.) skirta statistikams ar matematikams ir apima šios srities ekspertams galimai svarbių nustatymų spektrą (čia vartotojų laukia platus spektras *k*-means klasterizavimo ir MDS nustatymų, balsavimų kodavimo parinkty, bei grafiko dydžio ir mastelio keitimo parinkty). Nustatymų pasirinkimas standartinėje bei politologinėje dalyse nereikalauja iš esmės jokių matematinių žinių, tačiau trečiojoje (statistikams skirtoje) dalyje dirbantis vartotojas turėtų būti gerai susipažinęs su naudojamais metodais (daugiamačių skalių analize ir *k*-means klasterizavimu).



2 pav. Puslapio vaizdo pavyzdys (screenshot). Matematiko-statistiko skiltis.

Kokie mūsų ateities planai?

Nors išsikeltus tikslus įvykdėme ir sukūrėme pilnai funkcionuojantį interneto puslapį, skirtą LR Seimo balsavimų analizei, yra daug darbų, kuriuos atlikti esame numatę netolimoje ateityje. Svarbiausi iš jų:

- Grafikų tobulinimas – planuojame sukurti naujus, informatyvesnius, o taip pat interaktyvius grafikus;
- Puslapio dizaino tobulinimas – planuojame sukurti patrauklesnį ir patogesnį vartotojams puslapio dizainą;
- Puslapio veikimo efektyvumo didinimas – planuojame optimizuoti puslapį ir maksimaliai paspartinti jo veikimą;
- Puslapio reklamavimas – įvykdę aukščiau aprašytus tikslus, planuojame puslapį populiarinti akademinės bendruomenės (numatytas pristatymas KTU birželio 19 d.) ir politika besidominčios visuomenės (domisi „Transparency International“ skyrius Lietuvoje) tarpe.

Projektą „Mokslo pieva“ organizuoja mokslininkų ir dėstytojų komanda iš Baltijos pažangių technologijų instituto, Kauno technologijos universiteto, Socialinių inovacijų instituto, Vilniaus universiteto ir Vytauto Didžiojo universiteto.

Projekto metu atliekami įvairūs tyrimai iš fizikos, IT, socialinių mokslų bei kitų disciplinų.

2014 m. projekto „Mokslo pieva“ idėją rėmė UAB „Philip Morris Baltic“. Tyrimų temos, išvados bei rekomendacijos išreiškia autorių asmeninę nuomonę.

Daugiau informacijos: www.mokslopieva.lt