

INTEGRUOTA NUSIKALSTAMUMO PREVENCIJOS INFORMACINĖ SISTEMA

Donatas Čiukšys, Antanas Mitašiūnas, Saulius Ragaišis

Vilniaus universitetas

Alfredas Kiškis

Lietuvos teisės akademija

1. Įvadas. Nusikalstamumo prevencijos veikla remiasi plataus spektro informacija, kaupiama įvairiose institucijose. Norint šią informaciją analizuoti ir jos pagrindu priimti pagrįstus sprendimus, ji turi būti integruota. Tokio tipo informacinės sistemos kūrimo iškyla dvi pagrindinės problemos.

Tradiciškai, kuriant informacines sistemas, laikomasi vienos iš dviejų schemų. Tuo atveju, kai taikomoji sritis gerai apibrėžta, yra sukuriama automatizuojamos taikomosios srities modelis, kuris išreiškiamas struktūrizuotų duomenų modeliu. Struktūrizuotų duomenų modelyje yra sąvokos, atitinkančios taikomosios srities sąvokas. Savo ruožtu taikomosios srities informacinė sistema, paremta struktūrizuotų duomenų modeliu, irgi operuoja taikomosios srities sąvokomis [1].

Tuo atveju, kai taikomoji sritis nėra aiškiai apibrėžta, nėra galimas ir specifinis taikomosios srities modelis, sudarytas iš specifinių taikomosios srities sąvokų. Blieka taikomosios srities žodinio aprašymo tekstai (dokumentai) kaip tam tikro modelio išraiška. Tokiame modelyje operuojama tikrai žodžiais bei jų rinkiniais, o taip pat bendrais struktūrizuotais dokumento atributais, tokiais kaip autorius, data ir pan. Faktiškai, toks modelis yra sudarytas iš dokumentų, turinčių tam tikrus struktūrizuotus atributus, ir tokie dokumentai priklauso nuo taikomosios srities tik tiek, kiek skirtingos taikomosios srities apibūdinamos skirtingais žodžiais. Tokį modelį galėtume vadinti nestruktūrizuotų duomenų modeliu. Informacinė sistema, besiremianti nestruktūrizuotų duomenų modeliu, operuoja žodžiais arba žodžių rinkiniais ir neturi specifinių taikomosios srities sąvokų.

Tokiame kontekste iškyla klausimas: ar, esant "blogai" apibrėžtai taikomajai sričiai, nėra galimas modelis, kitoks nei paremtas dokumentais bei žodžiais juose, labiau orientuotas į taikomąją sritį; ar nėra galima informacinė sistema, pasižyminti labiau specifinėmis funkcijomis, negu operacijos su žodžiais ar jų rinkiniais. Prieš atsakydami į šį klausimą, pateiksime antrą problemą, su kuria susiduriama šiame darbe.

Paprastai, kuriant informacines sistemas, yra fiksuojami pirminiai faktai ar įvykiai ir jų pagrindu yra pateikiama išvestinė informacija. Toks pirminių faktų ar įvykių fiksavimas yra gana imlus darbo sąnaudoms, todėl dažnai yra apsiribojama tik tam tikro laikotarpio veiklos rezultatų fiksavimu. Duomenys už laikotarpį yra pateikiami lentelinių ataskaitų, vadinamų statistinėmis ataskaitomis, forma. Tokių statistinių ataskaitų pagrindu yra parengiamos suvestinės ataskaitos, jos analizuojamos, lyginamos, formuluojamos tam tikros išvados.

Statistinių ataskaitų naudojimas yra gana paplitęs įvairaus pobūdžio institucijų veikloje, suteikiant statistinėms ataskaitoms pirminių duomenų statusą. Iš pirmo žvilgsnio, informacinių technologijų požiūriu toks statistinių ataskaitų traktavimas yra "neteisingas" ir nepriimtinas - reikia turėti pirminius duomenis apie veiklos faktus bei įvykius, o reikiama statistinė ataskaita turėtų būti gaunama kaip įprasta ataskaita turimų pradinio duomenų pagrindu. Toks "teisingas" požiūris į informacines technologijas verčia įstaigos ar organizacijos veikloje įdiegti procesus, kurių rezultate būtų fiksuojami ir kaupiami pradiniai duomenys apie veiklos faktus bei įvykius. Tačiau tai reikalauja gana rimtų plataus masto pokyčių, o tai visada yra problematiška.

Todėl iškyla antras klausimas: ar negalima sukurti tam tikro informacinio serviso, tam tikros informacinės sistemos, nereikalaujant bent kiek žymesnių pokyčių įstaigos ar organizacijos veikloje, neįdiegiant pirminių duomenų fiksavimo bei kaupimo procesų.

Pasirodo, tiek į pirmąjį klausimą dėl informacinės sistemos sukūrimo neaiškiai apibrėžtoje taikomojoje srityje, tiek ir į antrąjį klausimą dėl informacinės sistemos sukūrimo nerestruktūrizuojant įstaigos ar organizacijos veiklos, galima tam tikru mastu atsakyti teigiamai, įvedant tokį duomenų primityvą kaip ataskaitas bei apibrėžiant informacinės sistemos funkcijas su jomis.

Straipsnyje yra pristatomas informacinės sistemos, besiremiančios statistinėmis ataskaitomis, modelis.

2. Sąvokų sistema. Panašiai, kaip nestruktūrizuotų duomenų informacinė sistema remiasi dokumentu, taip čia pateikiamas informacinės sistemos, sprendžiančios minėtas dvi problemas, modelis remiasi ataskaitos samprata.

2.1. Ataskaitos. Pateiksime ataskaitas apibūdinančią sąvokų sistemą.

• **Ataskaita:** tam tikras lentelės pavidalo šablonas, turintis eilučių ir stulpelių rinkinį, bei laikotarpio ir teritorinio vieneto (institucijos padalinio, teikiančio ataskaitą) atributus.

• **Ataskaitą teikiančių padalinių rinkinys:** institucijos struktūrinių padalinių, teikiančių duotą ataskaitą, rinkinys. Šis rinkinys apibrėžia visas galimas ataskaitos teritorinio vieneto atributo reikšmes.

• **Ataskaitos egzempliorius** (arba tiesiog **egzempliorius**): ataskaita, užpildyta duomenimis už tam tikrą laikotarpį (laikotarpio atributas) institucijos padalinio, priklausančio tą ataskaitą teikiančių padalinių rinkiniui (teritorinio vieneto atributas).

• **Egzempliorių paketas:** ataskaitą teikiančių padalinių rinkiniu apibrėžtas egzempliorių, turinčių tą pačią laikotarpio atributo reikšmę, rinkinys.

- **Periodiškumas:** ataskaitos atributas, parodantis, koku periodišku yra gaunami šios ataskaitos egzempliorių paketai. Periodiškumas gali būti: mėnesinis, ketvirtinis, pusmetinis, metinis.
- **Teritorinė eilutė:** eilutė, kurios duomenys yra pateikti už tam tikrą teritorinį vienetą (institucijos padalinį, teikiantį ataskaitą). Konkrečios ataskaitos visos eilutės yra teritorinės arba ne. *Pastaba:* jei ataskaitos *stulpeliai teritoriniai*, tai ji sistemoje saugoma transponuota (žr. *Transponuota ataskaita*).
- **3-dimensinės ataskaitos:** ataskaitos, kurių egzempliorius teikia daugiau nei vienas institucijos struktūrinis padalinys. Šios ataskaitos pavadintos 3-dimensinėmis dėl to, kad, neskaitant laikotarpio, jos turi tris matavimus: eilutes, stulpelius ir teritorinius vienetus. *Pastaba:* 3-dimensinių ataskaitų eilutės ir stulpeliai negali būti teritoriniai.
- **2-dimensinės teritorinės ataskaitos:** ataskaitos, kurių egzempliorius teikia tik vienas institucijos struktūrinis padalinys, turinčios tik teritorines eilutes. Šios ataskaitos pavadintos 2-dimensinėmis dėl to, kad, neskaitant laikotarpio, turi tik du matavimus: eilutes ir stulpelius.
- **2-dimensinės neteritorinės ataskaitos:** ataskaitos, kurių egzempliorius teikia tik vienas institucijos struktūrinis padalinys, neturinčios nei teritorinių eilučių, nei teritorinių stulpelių.
- **Transponuota ataskaita:** ataskaita, kurios eilutės vaidina stulpelių vaidmenį, o stulpeliai – eilučių vaidmenį. Duomenų bazėje tokios ataskaitos stulpeliai saugomi eilučių vietoje, o eilutės – stulpelių vietoje.
- **Auganti ataskaita:** ataskaita, kurios duomenys už tam tikrą periodą yra duomenys, surinkti už laiko intervalą nuo metų pradžios iki periodo pabaigos. Pvz., ataskaitoje už II-ą ketvirtį bus duomenys už pirmus šešis metų mėnesius.
- **Neauganti ataskaita:** ataskaita, kurios duomenys už tam tikrą periodą yra duomenys už laiko intervalą nuo šio periodo pradžios iki pabaigos. Pvz., ataskaitoje už II-ą ketvirtį bus duomenys už trečią, ketvirtą ir penktą mėnesius.
- **Skirtuminis stulpelis:** stulpelis, kurio duomenis galima atimti iš ankstesnio (laiko atžvilgiu) tos pačios ataskaitos to paties institucijos struktūrinio padalinio pateikto egzemplioriaus stulpelio duomenų, norint gauti duomenis už tam tikrą periodą. *Pastaba:* skirtuminiai stulpeliai gali būti tik augančiose ataskaitose (bet nebūtinai visi augančios ataskaitos stulpeliai yra skirtuminiai).
- **Sumuojamas laike stulpelis:** stulpelis, kurio duomenis galima sumuoti su tos pačios ataskaitos to paties institucijos struktūrinio padalinio pateikto egzemplioriaus, tik kito laikotarpio duomenimis, norint gauti suminius duomenis už tam tikrą laiko intervalą.
- **Sumuojamas teritoriškai stulpelis:** stulpelis, kurio duomenis galima sumuoti teritoriškai. Pvz.: remiantis duomenimis už žemesnio lygio teritorinius vienetus, galima gauti duomenis už visą respubliką. *Pastaba:* ataskaitos stulpelis gali būti sumuojamas teritoriškai, bet nebūti sumuojamas laike.
- **Teritorija:** teritorinis vienetas, apribotas tam tikro rajono riba, į save įtraukiantis rajoną kartu su to rajono ribose esančiu miestu(-ais).
- **Taškinis laikotarpis:** laikotarpis, apibrėžtas konkrečia data ataskaitos periodiškumo tikslumu. Pvz.: ataskaitai, kurios periodiškumas ketvirtinis, taškinis laikotarpis galėtų būti “1995 metų II-as ketvirtis”.
- **Intervalinis laikotarpis:** taškinių laikotarpių seka, apibrėžta trejetu (*pradžia, pabaiga, žingsnis*). Čia *pradžia* ir *pabaiga* yra taškiniai laikotarpiai, apibrėžiantys atitinkamai intervalinio laikotarpio pradžią ir pabaigą, *žingsnis* yra skaičius, lygus arba kartotinis ataskaitos periodiškumui, apibrėžiantis, koku žingsniu yra išsidėstę taškiniai laikotarpiai intervalinio laikotarpio viduje.
- **Suminis laikotarpis:** intervalinis laikotarpis, pagal kurį pateikiami suminiai tam tikrų stulpelių, sumuojamų laiko atžvilgiu, duomenys.

2.2. Teritoriniai lygmenys. Lietuvos administracinis-teritorinis suskirstymas yra pakankamai stabilus, todėl tikslinga pateikti jo modelį, kuris gali būti panaudotas įvairiuose taikymuose.

Kiekviena teritorinė ataskaita turi savo teritorinį suskirstymą, atitinkantį tą ataskaitą teikiančių institucijos struktūrinių padalinių teritorinę struktūrą. Norint visas ataskaitas saugoti vienodai, reikia sudaryti bendrą visoms ataskaitoms teritorinį klasifikatorių, kuris taip pat svarbus ataskaitų sumavimui teritoriškai.

Teritorinis klasifikatorius gali susidėti iš tokių lygmenų:

1. **Šalių lygmuo:** susideda iš vienos šalies – Lietuvos.
2. **Apygardų lygmuo:** susideda iš 5 apygardų, naudojamų teisėsaugos institucijų struktūrinių padalinių skirstymui. Apygarda susideda iš teritorijų.
3. **Apskričių lygmuo:** susideda iš 10 apskričių, atitinkančių LR administracinį skirstymą į apskritis. Apskritis susideda iš teritorijų.
4. **Teritorijų lygmuo:** susideda iš 44 teritorijų (žr. terminą *teritorija*).
5. **Miestų/rajonų lygmuo:** susideda iš 12 miestų ir 44 rajonų, atitinkančių LR administracinį skirstymą į respublikinio pavaldumo miestus ir rajonus.
6. **Padalinių lygmuo:** susideda iš visų konkrečios institucijos struktūrinių padalinių.

Pastabos:

- terminas *rajonas* vartojamas turint omenyje rajoną be to rajono ribose esančio miesto(-ų);
- skirstymas į apygardas ir apskritis yra nepriklausomas;
- institucijos padalinys gali priklausyti bet kuriam iš teritorinių lygmenų, pvz., Vilniaus mieste yra 4 apylinkių teismai, Panevėžio teismas apima Panevėžio teritoriją, Vilniaus apygardos teismas - Vilniaus apygardą.

Vienos teritorijos susidės tik iš rajono (pvz., Anykščių teritorija), kitos - iš rajono ir miesto (pvz., Vilniaus teritorija), dar kitos (tokių šiuo metu yra tik viena) - iš rajono ir dviejų miestų (Klaipėdos teritorija susidės iš Klaipėdos rajono, Klaipėdos miesto ir Neringos miesto).

Tokiu būdu duomenys turi pavidalą lentelės, kuri pateikia informaciją pagal tam tikrus rodiklius eilutėse bei stulpeliuose, už tam tikrą laikotarpį ir tam tikrai teritorijai. Iš vienos pusės, lentelinių duomenų primumas yra pakankamai universalus statistinių ataskaitinių duomenų modelis, iš kitos pusės, jis įgalina operuoti lenteliniais duomenimis, juos integruoti bei vizualizuoti įvairiais grafiniais pavidalais.

3. Sistemos funkcijos. Apibrėžiant sistemos funkcijas, laikotarpius traktuosime kaip dar vieną ataskaitos dimensiją. Ko pasekoje 3-dimensinės ataskaitos yra esybės, turinčios keturis matavimus: eilutes, stulpelius, teritorinius vienetus, laikotarpius - o 2-dimensinės ataskaitos - tris matavimus: eilutes (teritorinius vienetus, jei ataskaita teritorinė), stulpelius, laikotarpius.

Kitai sakant, tam, kad vienareikšmiškai nurodyti tam tikrą rodiklį, 3-dimensinei ataskaitai būtina nurodyti eilutę, stulpelį, teritorinį vienetą ir laikotarpį, o 2-dimensinei ataskaitai - eilutę (arba teritorinį vienetą 2-dimensinei teritorinei ataskaitai), stulpelį ir laikotarpį. Iš šito seka, kad vartotojo interfeisas turi numatyti eilučių, stulpelių, teritorinių vienetų ir laikotarpių pasirinkimo galimybę.

Laikysime, kad atitinkama ataskaitos dimensija yra vartotojo nurodyta taškiniu būdu, jei

- eilutėms: yra pasirinkta viena eilutė,
- stulpeliams: yra pasirinktas vienas stulpelis,
- teritoriniams vienetams: yra pasirinktas vienas teritorinis vienetas arba keli teritoriniai vienetai, pagal kuriuos duomenys yra susumuojami,

- laikotarpiams: yra pasirinktas taškinis arba suminis laikotarpis.

Kitais atvejais laikysime, kad atitinkama ataskaitos dimensija yra nurodyta rinkiniu.

Pastaba: 2-dimensinei teritorinei ataskaitai eilučių nurodymas apibrėžia ir teritorinę dimensiją, ir atvirkščiai.

Vartotojo užklausų formavimo interfeisas susideda iš tokių dalių:

- institucijos, teikiančios ataskaitas, pasirinkimas,
- ataskaitos pasirinkimas,
- rezultatų vaizdavimo būdo pasirinkimas: lentelė, žemėlapis, linijinė diagrama arba histograma,
- rodiklių (eilučių ir stulpelių) pasirinkimas,
- laikotarpio arba laikotarpio intervalo bei periodiškumo pasirinkimas,
- teritorinio lygmens pasirinkimas: šalis, apygardos, apskritys, teritorijos, miestai ir rajonai, padaliniai (padalinių lygmens nėra, jei duomenys bus vaizduojami žemėlapyje),
- teritorinių vienetų pasirinkimas (nėra, jei duomenys bus vaizduojami žemėlapyje).

Detaliau panagrinėsime, kaip vartotojo pasirinktas rezultatų vaizdavimo būdas įtakoja galimus dimensijų reikšmių pasirinkimus ir gaunamus užklausos rezultatus.

Vaizduojant rezultatus lentele, nėra apribojimų dimensijų reikšmių pasirinkimui. Vartotojo užklausos rezultate gaunama:

- **0-dimensinė lentelė** (susidedanti iš vienos eilutės) - jei taškiniu būdu nurodomos visos ataskaitos dimensijos.
- **1-dimensinė lentelė** (susidedanti iš vienos eilutės arba vieno stulpelio, priklausomai nuo to, kaip patogiau pavaizduoti) - jei viena ataskaitos dimensija nurodoma rinkiniu, o visos likusios dimensijos taškiniu būdu.
- **2-dimensinė lentelė** (susidedanti iš kelių eilučių ir kelių stulpelių) - jei dvi ataskaitos dimensijos nurodomos rinkiniu, o visos likusios dimensijos taškiniu būdu.
- **2-dimensinių lentelių seka** - jei bent trys ataskaitos dimensijos nurodomos rinkiniu.

Pastaba: sakydami "lentelė" čia turime omenyje ne pradinės statistinės ataskaitos lentelę, o lentelinių rezultatų pavaizdavimo būdą, t.y., tokios lentelės eilutėse ir stulpeliuose gali būti išsidėstę ir laikotarpiai, ir teritorijos, priklausomai nuo vartotojo užklausos.

Kad užklausos rezultatai galėtų būti pavaizduoti **žemėlapyje**, vartotojas turi stulpelių ir laikotarpio dimensijas nurodyti taškiniu būdu, pasirinkti vieną iš teritorinių lygmenų, o 3-dimensinės ataskaitos atveju dar ir taškiniu būdu nurodyti eilučių dimensiją.

Linijinėje diagramoje vaizduojamas rodiklio(-ių) kitimas laike, todėl laikotarpio dimensija turi būti nurodyta rinkiniu (intervalu su reikiamu periodiškumu). Kitų dimensijų reikšmių pasirinkimui principinių apribojimų nėra, nuo jų pasirinkimo priklauso laužčių linijinėje diagramoje skaičius, todėl iš praktinių sumetimų (kad gautą diagramą dar būtų galima suprasti) tikslinga apriboti, kad ne daugiau kaip dvi kitos dimensijos būtų nurodomos rinkiniu ir rezultate gaunamas laužčių skaičius neviršytų sistemos konfigūracijoje nurodyto skaičiaus.

Vaizduojant rezultatus **histograma**, irgi nėra principinių apribojimų, bet praktiškai tikslinga vartotojui užtikrinti galimybę gauti:

- vieno rodiklio reikšmės kitimą laiko atžvilgiu - visos dimensijos, išskyrus laiko, turi būti nurodytos taškiniu būdu;
- vieno rodiklio reikšmės pasiskirstymą teritoriniuose vienetuose - visos dimensijos, išskyrus teritorinę, turi būti nurodytos taškiniu būdu;
- kelių rodiklių reikšmės nurodytą laikotarpį teritoriniame vienetu - laikotarpio ir teritorinės dimensijos turi būti nurodytos taškiniu būdu.

Kadangi dauguma statistinių ataskaitų yra augančios (jose pateikiama informacija nuo metų pradžios), todėl sistemoje reikalingos funkcijos, įgalinančios vartotoją gauti tam tikro laikotarpio informaciją, kuri apskaičiuojama kaip dominančio ir atitinkamo ankstesnio laikotarpių rodiklių reikšmių skirtumas. Šis skirtumų skaičiavimas įveda papildomus apribojimus dimensijų reikšmių pasirinkimui: tam, kad būtų galima skaičiuoti skirtumus, turi būti pasirinkti tik skirtuminiai stulpeliai ir laikotarpiai bei teritoriniai vienetai, kurių egzemplioriai turi atitinkamo ankstesnio laikotarpio egzempliorius.

Nusikalstamumo prevencijos informacinė sistema skirta plačiam, geografiškai pasiskirsčiusių vartotojų ratui, todėl ji pasiekama per Internetą ir su ja dirbama, naudojant standartinę WWW naršyklę.

4. Išvados. Darbe pasiūlytas informacinės sistemos modelis, paremtas pradiniais duomenimis, turinčiais fiksuotos struktūros lentelinės statistinės ataskaitos formą ir galimus teritorijos bei laikotarpio atributus. Apibrėžtos informacinės sistemos funkcijos reikiamos informacijos formulavimui statistinių ataskaitų terminais bei tos informacijos pavaizdavimui lentelėmis, žemėlapiu, linijinėmis diagramomis bei histogramomis.

Pasiūlytas informacinės sistemos modelis susišaukia su duomenų sandėliavimo modeliais [2], tačiau operuoja specifiniais objektais - ataskaitomis - bei specifiniais veiksmais su jais.

Literatūra

[1] **T. J. Teorey.** Database Modeling & Design. The Fundamental Principles. *Addison-Wesley*, 1994, 277 pp.

[2] **B. Devlin.** Data Warehouse : From Architecture to Implementation. *Addison-Wesley*, 1996, 450 pp.

INTEGRATED CRIME PREVENTION INFORMATION SYSTEM

Donatas Čiukšys, Antanas Mitašiūnas, Saulius Ragaišis

Vilnius University

Alfredas Kiškis

Lithuanian Law Academy

Summary

This paper presents the integrated crime prevention information system which data model is based on statistical reports. Information system of such kind can be applied when initial data about the activity of the institution is described as set of reports in table mode.

The model consists of enhanced definition of the reports and the functions of reports based information system.

Keywords: information systems, domain independent data model, statistical reports.