ISSN 2029-249X
ISSN 2029-4824 (elektroninis leidinys)

VYTAUTO DIDŽIOJO UNIVERSITETAS
KAUNO TECHNOLOGIJOS UNIVERSITETAS
VILNIAUS UNIVERSITETO KAUNO HUMANITARINIS FAKULTETAS

INFORMACINĖS TECHNOLOGIJOS

XVIII TARPUNIVERSITETINĖS TARPTAUTINĖS MAGISTRANTŲ IR DOKTORANTŲ KONFERENCIJOS „INFORMACINĖ VISUOMENĖ IR UNIVERSITETINĖS STUDIJOS" (IVUS 2013) MEDŽIAGA

2013 M. BALANDŽIO 25 D., KAUNAS, LIETUVA

2013
VYTAUTO DIDŽIOJO UNIVERSITETAS

Redaktorių kolegija

Vytautas Barzdaitis, Vytauto Didžiojo universitetas, Lietuva
Vidas Daudaravičius, Vytauto Didžiojo universitetas, Lietuva
Andrius Davidsonas, Vytauto Didžiojo universitetas, Lietuva
Algirdas Deveikis, Vytauto Didžiojo universitetas, Lietuva
Vladislav V. Fomin, Vytauto Didžiojo universitetas, Lietuva
Kęstutis Kapočius, Kauno technologijos universitetas, Lietuva
Tomas Krilavičius, Vytauto Didžiojo universitetas, Lietuva
Jieming Ma, Xi'an Jiaotong-Liverpool universitetas, Kinija
Aušra Mackutė-Varoneckienė, Vytauto Didžiojo universitetas, Lietuva
Ka Lok Man, Xi'an Jiaotong-Liverpool universitetas, Kinija
Vaidas Morkevičius, Kauno technologijos universitetas, Lietuva
Lucio Tommaso De Paolis, Salento universitetas, Italija
Sigita Pečiulytė, Vytauto Didžiojo universitetas, Lietuva
Gailius Raškinis, Vytauto Didžiojo universitetas, Lietuva
Rasa Ruseckaitė, Monaš universitetas, Australija
Tomas Skersys, Kauno technologijos universitetas, Lietuva
Roma Uzdanavičiūtė, Vilniaus Gedimino technikos universitetas, Lietuva
Virginija Uzdanavičiūtė, Kauno technologijos universitetas, Lietuva
Audrius Varoneckas, Vytauto Didžiojo universitetas, Lietuva
Aušra Vidugirienė, Vytauto Didžiojo universitetas, Lietuva
Vincentas Vobolevičius, ISM Vadybos ir ekonomikos universitetas, Lietuva
Ivan S. Zapreev, ASML, Olandija
Indrė Žliobaitė, Aalto universitetas, Suomija

Konferencijos organizacinis komitetas

Tomas Krilavičius, Vytauto Didžiojo universitetas, Lietuva
Audrius Varoneckas, Vytauto Didžiojo universitetas, Lietuva
Aušra Mackutė-Varoneckienė, Vytauto Didžiojo universitetas, Lietuva
Aušra Vidugirienė, Vytauto Didžiojo universitetas, Lietuva
Andrius Davidsonas, Vytauto Didžiojo universitetas, Lietuva

Leidinio redaktoriai

Audrius Varoneckas, Vytauto Didžiojo universitetas, Lietuva
Aušra Mackutė-Varoneckienė, Vytauto Didžiojo universitetas, Lietuva

Rėmėjai:

Kauno aukštųjų ir informacinių technologijų parkas
Lietuvos energetikos institutas

ISSN 2029-249X
ISSN 2029-4824 (electronic issue)

VYTAUTAS MAGNUS UNIVERSITY
KAUNAS UNIVERSITY OF TECHNOLOGY
VILNIUS UNIVERSITY KAUNAS FACULTY OF HUMANITIES

INFORMATION TECHNOLOGY

Proceedings of the XVIII international master and phd students conference "Information Society and University Studies" (IVUS 2013)

25 April 2013, Kaunas, Lithuania

2013
VYTAUTAS MAGNUS UNIVERSITY

Editorial board

Vytautas Barzdaitis, Vytautas Magnus University, Lithuania
Vidas Daudaravičius, Vytautas Magnus University, Lithuania
Andrius Davidsonas, Vytautas Magnus University, Lithuania
Algirdas Deveikis, Vytautas Magnus University, Lithuania
Vladislav V. Fomin, Vytautas Magnus University, Lithuania
Kęstutis Kapočius, Kaunas University of Technology, Lithuania
Tomas Krilavičius, Vytautas Magnus University, Lithuania
Jieming Ma, Xi'an Jiaotong-Liverpool University, China
Aušra Mackutė-Varoneckienė, Vytautas Magnus University, Lithuania
Ka Lok Man, Xi'an Jiaotong-Liverpool University, China
Vaidas Morkevičius, Kaunas University of Technology, Lithuania
Lucio Tommaso De Paolis, University of Salento, Italy
Sigita Pečiulytė, Vytautas Magnus University, Lithuania
Gailius Raškinis, Vytautas Magnus University, Lithuania
Rasa Ruseckaitė, Monash University, Australia
Tomas Skersys, Kaunas University of Technology, Lithuania
Roma Uzdanavičiūtė, Vilnius Gediminas Technical University, Lithuania
Virginija Uzdanavičiūtė, Kaunas University of Technology, Lithuania
Audrius Varoneckas, Vytautas Magnus University, Lithuania
Aušra Vidugirienė, Vytautas Magnus University, Lithuania
Vincentas Vobolevičius, ISM University of Management, Lithuania
Ivan Zapreev, ASML, The Netherlands
Indrė Žliobaitė, Aalto University, Finland

Steering committee

Tomas Krilavičius, Vytautas Magnus University, Lithuania
Audrius Varoneckas, Vytautas Magnus University, Lithuania
Aušra Mackutė-Varoneckienė, Vytautas Magnus University, Lithuania
Aušra Vidugirienė, Vytautas Magnus University, Lithuania
Andrius Davidsonas, Vytautas Magnus University, Lithuania

Proceedings editors

Audrius Varoneckas, Vytautas Magnus University, Lithuania
Aušra Mackutė-Varoneckienė, Vytautas Magnus University, Lithuania

Sponsors:

Kaunas High-Tech & IT Park
Lithuanian Energy Institute

TURINYS / CONTENT

B. Aleksandravičius, D. Matulis
SNMP ŽURNALŲ ANONIMIZAVIMO TYRIMAS 8

D. Ašeriškis
PROJEKTO VALDYMO SISTEMOS ŽAIDIMIZAVIMAS 13

J. Balčas, T. Krilavičius, Ž. Medelis
LIETUVIŠKŲ DATŲ ATPAŽINIMAS 17

D. Baronas, A. Žilinskas, F. Ivanauskas
COMPUTATIONAL MODELLING AND VALIDATION OF A MULTILAYER AMPEROMETRIC BIOSENSOR 22

K. Bytautas, T. Krilavičius
LRS NARIŲ GRUPAVIMAS PAGAL BALSAVIMĄ NAUDOJANT KLASTERIZAVIMĄ 27

V. Barzdaitis, A. Degutis, J. Sabulytė
NAUJŲ RIEDĖJIMO GUOLIŲ DEFEKTŲ DIAGNOSTINIAI TYRIMAI 33

L. Dovydaitis
UNIVERSALIOJO VIENETINIO PRISIJUNGIMO MODELIS 39

E. Dzidolikaitė
MAŽŲ IT PROJEKTŲ APIMTIES VALDYMO METODAS 43

D. Jankūnas, R. Chavez-Santiago, I. Balasingham, V.V. Fomin
SIMULATION OF A COGNITIVE RADIO SYSTEM FOR INTERFERENCE REDUCTION ON THE 2.4 GHZ ISM BAND IN HOSPITAL ENVIRONMENTS 49

S. Kuzminskis, V.V. Fomin
ELEKTRONINIO DIENYNO ĮVEDIMO MOKYKLOSE ĮTAKOS ORGANIZACINIŲ PROCESŲ EFEKTYVUMUI TYRIMAS 53

L. Litvinas
APPLICATION OF ARTIFICIAL NEURAL NETWORKS TO DETERMINE CONCENTRATIONS OF MIXTURE 58

J. Ma, P.W.H. Wong, K.L. Man, N. Zhang, T.O. Ting, Sheng-Uei Guan
INTRODUCTION OF PHOTOVOLATIC BEHAVIORAL MODELS 63

J. Mandravickaite
STYLOMETRY FOR VISUALIZING RHETORICAL DIFFERENCES IN PARLIAMENTARY SPEECH 68

A. Mialik
SKOLOS SPĄSTŲ LOGISTINIS MODELIAVIMAS IT PRIEMONĖMIS 73

D. Mockus, J. Čeponis
TINKLALAPIŲ APSAUGOS PRIEMONIŲ ĮTAKA ATSAKO LAIKUI 77

P. Nanevičius, A. Venčkauskas
PIRŠTO KRAUJAGYSLIŲ TINKLO TAIKYMO SLAPTŲ RAKTŲ GENERAVIMUI GALIMYBIŲ TYRIMAS 81

A. Ogorodov
RINKOS NUOTAIKŲ TECHNINIAIS INDIKATORIAIS GRINDŽIAMŲ PREKYBOS STRATEGIJŲ PELNINGUMO TYRIMAS 85

A. Orlova
FORMALIZUOTAS JUDRIŲJŲ METODIKŲ APRAŠYMAS 91

N. Pažareckas
VIETOVĖS NUSTATYMO (GPS) PAKLAIDŲ ATSIRADIMO PRIEŽASTYS, JŲ MAŽINIMO METODAI IR TYRIMAI 97

G. Petraitis, N. Morkevičius
DISKO ANALIZEI ATSPARUS DUOMENŲ SLĖPIMO METODAS 102

I. Plauska
DEVELOPMENT OF ROBOT PROGRAMMING ONTOLOGY 108

V. Pnevski
ATSTUMŲ SKAIČIAVIMAS IKI IEŠKOMŲ ŽODŽIŲ IŠTISINĖJE ŠNEKOJE 114

T. Rasymas
BALSO KOMANDŲ APTIKIMAS KALBOS SIGNALE NAUDOJANT NAIVŲ BAJESO KLASIFIKATORIŲ 118

F. Ricciardi, C.L. Manes, L.T. De Paolis
AUGMENTED REALITY IN RADIOFREQUENCY ABLATION OF THE LIVER TUMOURS 123

N. Saikauskas
NUSIKALSTAMOS VEIKLOS „SKAIČIAVIMŲ DEBESIES“ SAUGYKLOSE ATKŪRIMO METODAS 129

Z. Shen, D.O. Afolabi, K.L. Man, Hai-Ning Liang, N. Zhang, Sheng-Uei Guan, P.W.H. Wong
WIRELESS SENSOR NETWORKS FOR ENVIRONMENT MONITORING 135

G. Sukarevičienė, V.V. Fomin
ANALYSIS AND MODELLING OF DYNAMIC SPECTRUM ACCESS INFORMATION INFRASTRUCTURE (DSA II) 138

T. Šalnaitė
STATISTINIŲ KALBOS MODELIŲ TAIKYMAS NATŪRALIOS KALBOS DIALOGUOSE 143

E. Šarapnickytė, K. Kapočius
SAITYNO INFORMACINIŲ SISTEMŲ VARTOTOJO SĄSAJOS TESTAVIMO METODIKA 147

B. Šilanskaitė, D. Dilijonas
MAŽMENINĖS BANKININKYSTĖS SAVITARNOS TERMINALŲ PASLAUGŲ VERTINIMO MODELIS 153

R. Širvinskas, T. Krilavičius, Ž. Medelis
AUTOMATINIS CITATŲ ATPAŽINIMAS LIETUVIŠKUOSE NAUJIENŲ PORTALUOSE 159

T. Trainys, V. Uzdanavičiūtė, R. Butleris
RELIACINIŲ DUOMENŲ BAZIŲ DUOMENŲ INTEGRAVIMAS TAIKANT ONTOLOGIJAS 164

R. Užupytė, V. Morkevičius
LIETUVOS SEIMO NARIŲ BALSAVIMŲ TYRIMAS PASITELKIANT SOCIALINIŲ TINKLŲ ANALIZĘ: TINKLO KONSTRAVIMO METODOLOGINIAI ASPEKTAI 170

V. Vaičiūnas
NUOLAIDŲ IR LOJALUMO KORTELIŲ PERKĖLIMO Į IŠMANIUOSIUS ĮRENGINIUS TYRIMAS 176

M. Vasiljevas, K. Šidlauskas
ROBOTO VALDYMO SISTEMOS NEURONINĖS KOMPIUTERIO SĄSAJOS TYRIMAS 181

Ž. Veiverys, R. Kubiliūnas
TARPUSAVIO BENDRADARBIAVIMU PAGRĮSTAS NUOTOLINIO MOKYMOSI PARAMOS SISTEMOS MODELIS 185

T. Vileiniškis, R. Butkienė
SEMANTINIŲ INFORMACINIŲ SISTEMŲ KŪRIMAS VIRTUALIZUOTŲ RELIACINIŲ DUOMENŲ BAZIŲ PAGRINDU 190

A. Zocco, D. Cannone, A. Sindico, L.T. De Paolis
STEREOSCOPIC HUMAN-COMPUTER INTERFACE FOR NETWORK CENTRIC OPERATIONS 196

A. Zonys, J. Čeponis
STATINĖS KODO ANALIZĖS ĮRANKIO „GENDARME“ TOBULINIMAS 201

V. Žitkus
KARKASO „ZEND FRAMEWORK“ PAGRINDU SUKURTŲ INTERNETINIŲ SISTEMŲ PRITAIKYMO IŠMANIESIEMS ĮRENGINIAMS TECHNOLOGINIO SPRENDIMO SUDARYMAS IR TYRIMAS 205

SNMP žurnalų anonimizavimo tyrimas

Bartas Aleksandravičius
KTU Informatikos fakultetas, Studentų g. 50, Kaunas, Lietuva
bartas.aleksandravicius@stud.ktu.lt

Darius Matulis
KTU Informatikos fakultetas, Studentų g. 50, Kaunas, Lietuva
darius.matulis@ktu.lt

Santrauka. **Kompiuterinių tinklų valdytojai plačiai naudoja paprastąjį tinklų valdymo protokolą, kuris padeda aptikti tinklų anomalijas ir įvairius gedimus. Norint tobulinti patį protokolą ir/ar stebėti specifinės tinklinės įrangos elgseną, būtina disponuoti šio protokolo registracijos žurnalais. Siūlomas modelis leidžia anonimizuoti šiuos žurnalus, išanalizuoti jų anonimiškumo laipsnį ir perduoti trečioms šalis be grėsmės pakenkti disponuojamo tinklo saugumui.**

Reikšminiai žodžiai: **SNMP įvykių žurnalai; anonimiškumas; entropija.**

I. ĮVADAS

Realiame pasaulyje interneto paslaugų tiekėjai ir tinklinių sistemų administratoriai disponuoja dideliais informacijos k iekiais, susijusiais s u ti nklo in ventoriaus veikimo būsenomis, incidentais, įvairiausio pobūdžio informaciniais pranešimais, kurie parodo ir padeda sistemų administratoriams, kai sistema nukrypsta n uo natūralios būsenos.

Problema iškyla, kai norima tą informaciją perduoti kitiems, ar kitų sistemų administratoriai, analitikai, inžinieriai ar mokslininkai norėdami geriau suprasti, išanalizuoti ir patobulinti informavimo protokolą ar metodą, tam ne turi visos informacijos di sponavimo laisvės. Natūraliai kyla klausimas, kodėl taip yra?

Sistemų administratoriai, interneto paslaugų tiekėjai nelinkę skleisti informaciją viešam naudojimui apie savo valdomų tinklų įrenginių būsenų, adresų, siuntėjų, gavėjų, gedimų, klaidų ir kitos informacijos iš esmės tik todėl, kad nebūtų galima analizuoti jų tinklų informacijos ir nebūtų kuriamos priemonės, nukreiptos prieš jų pačių tinklus bei jų silpnąsias vietas. Kitaip sakant, pagrindinis argumentas yra saugumo grėsmės.

Jeigu sistemų administratoriai ar interneto tiekėjai turėtų galimybę anonimizuoti savo disponuojamų tinklų informaciją iki tokio lygio, kad nebūtų galima atsekti realių įrenginių adresų, datų ir kitos informacijos, kuri galėtų kaip nors pakenkti kompiuterių tinklų valdytojams, tikrai da ug da ugiau r eikiamos i nformacijos moksliniams tyrimams atsirastų viešuose terpėse ir būtų laisvai pasiekiama.

Problemos s prendimas. Sukurtas modelis, leidžiantis pagal kelis a tributus a nonimizuoti S NMP žurnalus, pasirenkant naudingų duomenų išliekamumo žurnale lygmenį. Taip pat jis leidžia įvertinti SNMP žurnalų anonimizuotų atributų įtaką bendram anonimizavimo l ygmeniui i r na udingos i nformacijos išliekamumui.

Esminiai reikalavimai modeliui

- Anonimizavimo l ygmuo t uri a psaugoti organizacijų teikiamą informaciją iki tokio lygio, kad nebūtų galimybės jai pačiai pakenkti piktavališkais tikslais.
- Anonimizavimo l ygmuo t uri būti optimizuotas iki to kio l ygio, k ad nepažeidžiant pirmo reikalavimo, anonimizuota i nformacija vi s d ar turėtų tiriamąją vertę moksliniu ir analitiniu aspektu.
- Modelis turi turėti įrankį, le idžiantį įvertinti anonimizuotų duomenų anonimiškumo lygį.

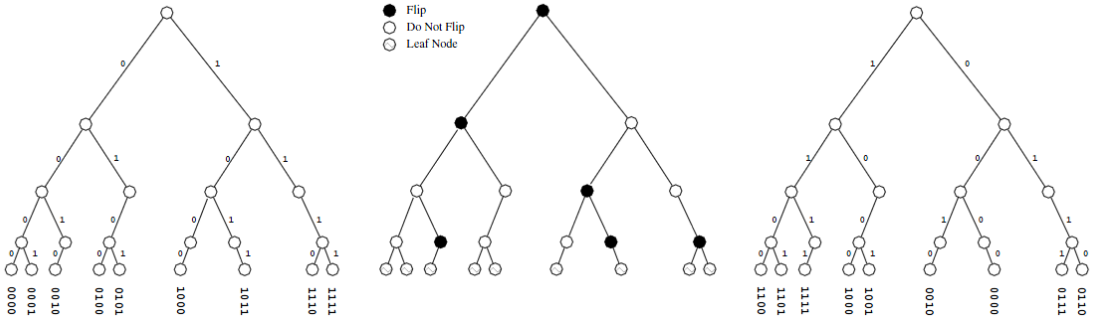
II. SNMP APŽVALGA IR ANALIZĖ

Paprastas tin klo v aldymo p rotokolas (angl. Simple N etwork M anagement Protocol - SNMP) ve ikia pagal serverio/agento modelį [11]. Agentai siunčia signalus apie įvykius (angl. Trap) serveriui, kai atsitinka tam t ikra s ituacija, p vz., kintamojo reikšmė viršija nustatytąją ar pasiekia kritinę vertę.

Kadangi pastarieji pranešimai (Trap) turi didžiausią kiekį su tinklų sauga susijusios jautrios informacijos, mes ir pasirinkome analizuoti būtent šiuos pranešimus, n es i š jų galima s usidaryti o rganizacijos tinklo topologiją, anomalijų ir gedimų bendrą vaizdą, galimus atakos taškus bandant įsiskverbti į organizacijos tinklą. Būtent dėl šių priežasčių Trap pranešimai yra griežtai saugomi organizacijų vidaus saugumo politikų, nes p asinaudodamas š iais d uomenimis p iktavalis ga li bandyti pakenkti organizacijai per kompiuterinį tinklą pasinaudodamas šia jautria informacija.

III. ANONIMIZAVIMO METODŲ APŽVALGA

Kadangi mūsų darbas remiasi anonimizavimo principais, tai iš pradžių būtina apžvelgti ir išanalizuoti duomenų anonimizavimo specifiką. Privatumas y ra atskleistas, kai specifinis kenkėjas sugeba teisingai sujungti asmenis su jų jautriomis vertėmis. Metrinis anonimiškumas nu stato lygį, k aip gerai anonimizavimo technika s lepia at stovo t apatybes ar s antykius p rieš specifinį kenkėją. Norint, kad a nonimizuoti SNMP žurnalai būtų naudingi ir tolimesniems ty rimams bei nebūtų atskleista jautri interneto tiekėjų informacija, būtina anonimizuoti ne tik siuntėjo ir gavėjo IP adresus, bet kitus jautrius žurnalų laukus, tokius kaip siuntimo data ir laikas, įrenginio ar objekto identifikatorius bei jų būsenas ar įvykius. Toliau m es j uos va dinsime kvazi-identifikatoriais (netiesiogiai objektą identifikuojančių atributų aibė, toliau – QI). Apžvelkime keletą anonimizavimo algoritmų, kurie panaudoti panašiems informacijos anonimizavimo tikslams.



(a) Originalus adresų medis (b) Anonimizavimo funkciją f_i (c) Anonimizuotas adresų medis

1 pav. Prefikso išsaugojimo anonimizavimo funkcijos geometrinė interpretacija [14]

A. *RYŠIŲ NUTRAUKIMO ALGORITMAS*

Weijia Y ang i r S anzheng Q iao [8] pasiūlė anonimizavimo algoritmą, kurio tikslas y ra n utraukti ryšius t arp k -identifikatorių ir su jais susijusia jautria informacija. Tai nereikalauja anonimizuoti visų QI reikšmių. Pagrindinė idėja yra ta, kad atsitiktiniu būdu (angl. random) pakeisti dalį QI reikšmių kiekvienam įrašui, vartojant originalių verčių požymių pasiskirstymą. Tokiu būdu jokia nauja informacija nėra pridėta prie anonimizuotų duomenų, ryšiai tarp QI ir jautrių duomenų yra nutraukti, o originalus duomenų pasiskirstymas yra išsaugotas.

B. *IP ADRESO ANONIMIZAVIMAS IŠSAUGANT PREFIKSĄ IR LEKSIKOGRAFINĮ EILIŠKUMĄ*

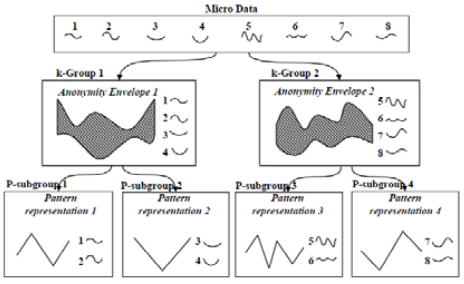
Jun Xu, J inliang Fan ir Mostafa H. Ammar [14] pasiūlė naują IP adreso anonimizavimo algoritmą, kuris išsaugo tokį patį IP adreso prefiksą, j ei yra anonimizuojami d u t o paties p otinklio a dresai. Šio anonimizavimo tipo nauda yra ta, kad tinklų ir potinklių struktūra yra visiškai išsaugota, kai tuo pačiu metu ir nežinoma. Kenkėjas bandydamas skenuoti tokį anonimizuotą IP adresą jokios apčiuopiamos naudos negaus, ne s nei r ealus p otinklis, nei t ikras I P a dreso savininkas nebus žinomas.

Apibrėžimas. Du IP adresai $a = a_1a_2 \ldots a_n$ ir $b = b_1b_2 \ldots b_n$ turi *k* bitų ilgio prefiksą ($0 \le k \le n$), jei $a_1a_2 \ldots a_k = b_1b_2 \ldots b_k$ ir $a_{k+1} \neq b_{k+1}$, tada $k < n$. Anonimizavimo funkcija F yra apibrėžta kaip vienas vienam funkcija nuo $\{0, 1\}^n$ į $\{0, 1\}^n$. Anonimizavimo funkcija F yra priešdėlio išsaugojimas: jei duotiems dviems IP adresams a ir b, kurie turi k – bitų ilgio priešdėlį, tada F (a) ir F (b) turi k – bitų ilgio priešdėlį taip pat.

Įsivaizduokime anonimizavimo metodo i šsaugant prefiksą geometrinę interpretaciją. Atkreipkime dėmesį, kad visai I P ad reso er dvei g ali at stovauti užbaigtas dvejetainis medis. Kadangi IPv4 adresai, šis medis turėtų 32 mazgus tuo metu, kai IPv6 turės 128 mazgus. Kiekvienam IP adresui at stovauja mazgas. B e to, kiekvienas mazgas atitinka bito poziciją (išreikštas pagal mazgo aukštį) ir vertes (išreikštas pagal jos tėvinio mazgo šakos kryptį).

C. *(K,P)- ANONIMIŠKUMO MODELIS: LAIKO DUOMENŲ STRUKTŪRĄ SAUGANTIS ANONIMIZAVIMAS*

Xuan S hang [10] pasiūlė algoritmą, skirtą laiko duomenims a nonimizuoti. J o metode yra numatyta, ka d kiekviena laiko eilutė yra publikuota tr ijuose komponentuose: apibendrinti QI atributai, QI struktūros pavaizdavimas i r j autri i nformacija. P atikslinant (k, P) - anonimiškumo modelis gali būti apibūdintas kaip konceptualus tr adicinio k -anonimiškumo išplėtimas. V is dėlto (k, P) - anonimiškumas nepriklauso tradiciniam k-anonimiškumo algoritmui.



2 pav. (k,P)- anonimiškumo modelis [10]

Kaip iliustruoja 2 pav., jų modelis ga rantuoja anonimiškumą dviejuose lygmenyse. Pirmame lygmenyje QI atributai yra apibendrinti, kad įvykdytų tradicinį k-anonimiškumą, nepriklausomai nuo QI struktūros pavaizdavimo. Apibendrinimo r ezultatai tu ri daugiau padalijimų, ž inomų kaip k -grupės. Pažymėtina, ka d apibendrinti Q I atributai panašėja į tuos tradiciniame k-anonimiškume. Antro l ygmens įrašų anonimiškumas laikomas kiekvienoje k -grupėje. Bet kokiam įrašui *r* k-grupėje, jei ten egzistuoja bent jau $P - 1$ kiti įrašai, kurie turi tą patį struktūros pavaizdavimą kaip *r*, teigiama, kad P-anonimiškumas yra priverstinai sukuriamas visai šiai k-grupei. Todėl galima padalinti k-grupę toliau į pogrupius, iš kurių kiekvienas talpina mažiausiai P įrašų, turinčių identišką struktūros pavaizdavimą.

Apibendrinant a pžvelgtus a nonimizavimo algoritmus ir metodus galima juos pritaikyti skirtingiems SNMP žurnalų laukams; priklausomai nuo kintamųjų reikšmių formato, taikyti vieną ar kitą anonimizavimo būdą:

- **Registracijos d atos** (angl. time stamp) laukų anonimizavimui tik slinga ta ikyti 3 .3 p unkte apžvelgtą metodą „Laiko duomenų struktūrą saugantis a nonimizavimas“, ne s j uo ga lima anonimizuoti l aiko d uomenis i ntervalais, kuriuos taip pat galima dar papildomai skaidyti į mažesnius pogrupius, jeigu reikia padidinti anonimiškumą.
- **Objekto I P a dresui** (angl. agent ad dress) tikslinga būtų panaudoti 3.2 punkte apžvelgtą IP adreso anonimizavimą, išsaugant prefiksą ir leksikografinį jo eiliškumą.
- **Objekto i dentifikatoriams** (toliau – OID) **ir jų reikšmėms** tikslinga būtų panaudoti 3.1 punkte apžvelgtą algoritmą, kuris leidžia nutraukti ryšius tarp realių identifikatorių ir jautrios informacijos.

IV. ANONIMIZAVIMO LYGIO ANALIZĖ

Prieš pateikiant trečioms šalims anonimizuotus SNMP žurnalus, juos būtina ištirti, kiek naudingos a rba vis dar atitinkančios realybę informacijos yra išlikę šiuose anonimizuotose duomenyse bei kokia tikimybė yra susieti realius duomenis su jų būsenomis ar kitais realiais atributais (pvz., kada įvyko įvykis, iš kurio tinklo įrenginio su kuriuo jo objektu?). Analizei buvo pasitelktas anonimiškumo matavimo pa rametras – entropija. N orint visapusiškai įvertinti ar anonimizuoti, duomenys ne turi galimybės atskleisti realios informacijos ten, kur nenorima; reikia įvertinti visus informacijos laukus, patikrinti jų individualias, sandūros (angl. joint) ir sąlygines (angl. conditional) entropijas, o t aip pa t duomenų stulpelių priklausomybes vienam nuo kito (pvz., žinant ar pašalinant vieną vieno stulpelio reikšmę, kokia įtaka daroma kitam stulpeliui individualiai ir sandūros entropijai).

Naudojamos formulės vertėms gauti:

$$H(X) = -\sum_{i=n} p(x_i) * \log\big(p(x_i)\big) \quad (1)$$

$$H(Y) = -\sum_{i=n} p(y_i) * \log\big(p(y_i)\big) \quad (2)$$

$$H(X,Y) = -\sum_{i=n}\ \sum_{j=n} p\big(x_i, y_j\big) * \log\Big(p\big(x_i, y_j\big)\Big), \quad (3)$$

kur

$p(x_i)$ yra kiekvieno s kirtingo X aibės argumento pasikartojimo tikimybė x_i,

$p(y_i)$ yra kiekvieno skirtingo Y aibės argumento pasikartojimo tikimybė y_i,

$p\big(x_j, y_i\big)$ yra X ir Y skirtingų pasikartojimų tikimybė x_j, y_i,

n yra suminis įvykių kiekis.

Entropijos apskaičiuojamos kiekvienam anonimizuotam stulpeliui (žiūrėti 2 lentelę), o ten, kur egzistuoja tamprus ryšys tarp dviejų stulpelių (kaip šiuo atveju IP-OID), vertinama ir sandūros (angl. joint) entropija [15].

Remiantis na ujuoju a nonimiškumo matavimu vadovaujantis d aline e ntropija [16], i švadose p ateikiama saugumo lygio skalė nurodanti rizikos grėsmę, priklausomai nuo e samo anonimiškumo lygio, ku r yra vertinama e ntropijos la ipsnis (H), skirtingų QI kiekis ir pasikartojimo dažnio pasiskirstymas visoje duomenų aibėje:

- NESAUGU. Labai mažas neapibrėžtumas (H< 1.5) le idžia le ngvai a tstatyti r yšius ta rp tin klo įrenginių ir jo objektų, nes pasikartojimo dažnis pasiskirstęs labai nevienodai ir jų yra labai mažai.
- NEPATIKIMA. Neapibrėžtumas minimalus (H=1.5-2.5) le idžia le ngvai a tstatyti r yšius ta rp tinklo įrenginių ir jo objektų, nes pasikartojimo dažnis pasiskirstęs daugiau nei keliuose objektuose vienodai, bet jų yra mažai.
- SAUGU. Neapibrėžtumas pakankamas (H=2.5-5), kur is neleidžia le ngvai a tstatyti r yšius ta rp tinklo įrenginių ir jo objektų, nes pasikartojimo dažnis pasiskirstęs daugiau nei keliuose objektuose vi enodai, o dominuojančių pasikartojimų daugiau nei 2.
- LABAI SAUGU. Neapibrėžtumas labai didelis (H>5) neleidžia atstatyti ryšius tarp tinklo įrenginių ir jo objektų, nes pasikartojimo dažnis pasiskirstęs labai tolygiai ir jų yra labai daug (artima ar lygu n).

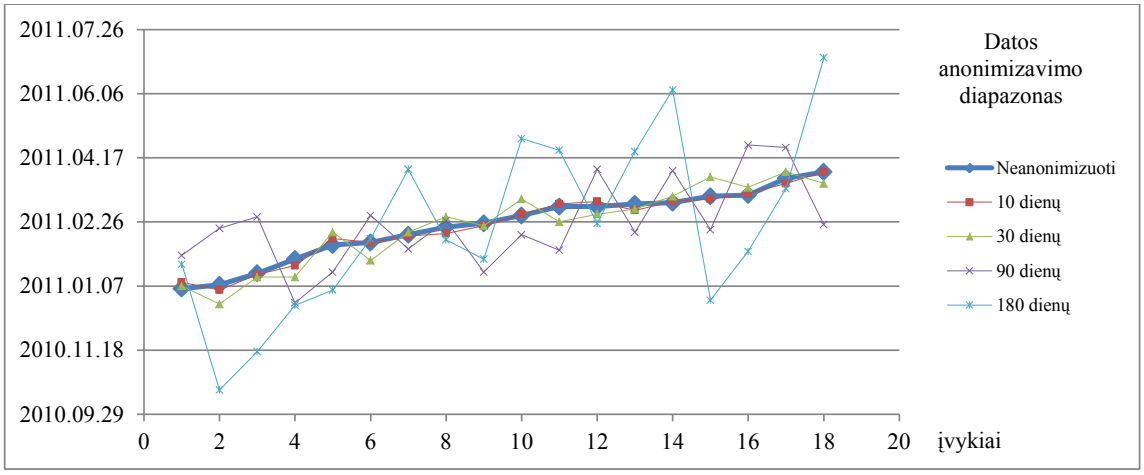
V. EKSPERIMENTAS IR JO REZULTATAI

Eksperimento t ikslas – įvertinti modelio anonimizavimo metodus i r išanalizuoti a nominiškumo lygį.

Mūsų sukurtas programinės įrangos prototipas (toliau – prototipas), leidžiantis sparčiai atlikti visus pristatomo modelio veiksmus nuo duomenų nuskaitymo iki išanalizuotų duomenų perdavimo trečioms šalims, įskaitant duomenų atrinkimą, anonimizavimą ir jų analizę, mums leidžia efektyviai įvertinti modelio pagrįstumą. Mūsų prototipas turi funkcijas, leidžiančias įtakoti anonimizavimo procesą, priklausomai k iek ir k okios realios in formacijos norima p alikti a nonimizuotose duomenyse.

- Anonimizuojant datą ir laiką (naudojamas 3. 3 punkte a prašymas a lgoritmas) galima nu statyti, kokiame diapazone generuoti datos nukrypimą nuo realios (pvz, paslėpti tikslias įvykusio įvykio datas ir /ar la ikus, arba keičiant minimaliai, kad būtų galima įvertinti įvykių tendencijas laiko atžvilgiu).
- Anonimizuojant IP adresą (naudojamas 3. 2 punkte ap rašymas al goritmas) galima n ustatyti prefikso ilgį (jei r eikia, išsaugoma o rganizacijos tinklo struktūra).

Anonimizuojant OID ir jų reikšmes (naudojamas 3.1 punkte ap rašymas al goritmas) yra galimybė sumaišyti esamas reikšmes, nutraukiant realius gedimų ar įvykių ryšius su realiais, juos generuojančiais, objektais. B e to , esant poreikiui, galima įterpti papildomų neegzistuojančių objektų ar atsitiktinai sugeneruotų ne tikrų įvykių (naudinga, kai norima paslėpti tikruosius įvykius).



3 pav. Priklausomai, koks pasirinktas datos anonimizavimo diapazonas, gautų reikšmių lyginimas su realiais duomenimis

Testiniams žurnalų duomenims mes naudosime 1000 realių SNMP Trap pranešimų, kuriame yra du potinkliai po 50 tinklinių įrenginių, turinčių savo unikalius IPv4 adresus. Ištyrėme, kaip tankiai „lūžinėja“ organizacijos vietinis tinklas ir kurie įrenginiai ir/ar moduliai tai įtakoja. Būtina pateikti visus įrašus, kur yra pranešta apie SNMP standartinį aliarmą (angl. generic trap) su jo būsena LinkDown (GT=3). Išfiltravę informaciją iš visų duomenų, gavome 86 įrašus, kurie atitiko mūsų užsibrėžtą tikslą.

Lentelė I. Pirminiai SNMP pranešimo duomenys

Data laikas	IP adresas	OID	Įvykis
2011-01-12 20:41:11	172.30.10.34	a3Com	linkDown
2011-01-12 04:08:18	172.30.20.39	a3Com	linkDown
2011-01-11 07:15:37	172.30.10.33	microsoft.1.1.3.1.1	linkDown
2011-01-09 06:37:11	172.30.10.12	a3Com	linkDown
2011-01-01 03:02:10	172.30.20.18	a3Com	linkDown
2011-01-04 07:46:54	172.30.10.19	snmpTraps	linkDown

Duomenų anonimizavimui buvo pasirinkta anonimizuoti datą/laiką, IP adresą ir OID reikšmes, o įvykis paliktas nekeistas. Dabar tas pats duomenų masyvas atrodo kitaip (rodomas tik fragmentas):

Lentelė II. Anonimizuoti SNMP pranešimo duomenys

Data laikas	IP adresas	OID	Įvykis
2011-01-01 03:04:10	44.83.61.18	snmpTraps	linkDown
2011-01-05 07:48:54	44.83.10.204	a3Com	linkDown
2011-01-10 06:39:11	44.83.10.222	microsoft.1.1.3.1.1	linkDown
2011-01-11 07:13:37	44.83.10.139	a3Com	linkDown
2011-01-12 20:39:11	44.83.10.196	snmpTraps	linkDown
2011-01-12 04:06:18	44.83.61.238	microsoft.1.1.3.1.1	linkDown

Apskaičiuojant entropijas būtina patikrinti, ar nėra galimybės susieti realių tinklo įrenginių (IP adresų) su jų objektais (OID), todėl buvo skaičiuojamos entropijos H(IP), H(OID) ir sandūros entropija H(IP, OID). Naudojamų algoritmų anonimizavimo sparta nebuvo vertinama, nes sparta l abai pr iklauso nuo pr ograminio išpildymo ir techninės įrangos galimybių. Norime pabrėžti, kad mūsų metodas skirtas saugumo lygiui įvertinimui, o ne anonimizavimo spartai tirti, nors atliekant tyrimus su dideliais (daugiau nei 100 000 SNMP žurnalo įrašų) duomenų kiekiais, sparta tenkino poreikius. Sustačius reikšmes į formules, gauti rezultatai.

X aibės argumento pasikartojimo dažnis x_i apskaičiuojamas suskaičiuojant, kiek tokių pačių elementų yra aibėje ir surandama jo dalis visoje aibėje X. Pvz.: IP adresas 23.239.43.192 pasikartojo 7 kartus iš 86 įrašų, taigi jo pasikartojimo dažnis aibėje X yra 7/86 = 0.0814. Tokiu būdu surandami visų egzistuojančių skirtingų IP adresų pasikartojimo dažniai ir sustatomi į 1 formulę:

H(IP) = -1 * [(0.0814* log(0.0814)) + (0.1279* log(0.1279)) + (0.1395* log(0.1395)) + (0.0814 * log(0.0814)) + (0.1163* log(0.1163)) + (0.093* log(0.093)) + (0.1744* log(0.1744)) + (0.186* log(0186))]
H(IP) = 2.93558

Tokiu pačiu metodu apskaičiuojami visų OID pasikartojimo dažniai (naudojama 2 formulė):

H(OID) = -1 * [(0.3256 * log(0.3256)) + (0.3256 * log(0.3256)) + (0.3488 * log(0.3488]
H(OID) = 1.58419

Sandūros entropijos (H(IP,OID)) įvykių porų pasikartojimų dažnių x_j, y_i apskaičiavimas atliekamas suskaičiuojant, kiek tokių pačių elementų porų yra aibėje ir surandama jų dalis sandūros aibėje X,Y.

Pvz.: IP adresas 23.239.43.192 su OID microsoft.1.1.3.1.1 pasikartojo 3 kartus iš 86 įrašų, ta igi j o pasikartojimo dažnis sandūros aibėse X ,Y yra 0.0349. Tokiu būdu surandamos visos porų kombinacijos, apskaičiuojami jų pasikartojimų dažniai.

$p(x_i,y_i)$ =[0.0814, 0.0698, 0.0698, 0.0698, 0.0698, 0.0581, 0.0581, 0.0581, 0.0465, 0.0465, 0.0349, 0.0349, 0.0349, 0.0349, 0.0349, 0.0349, 0.0233, 0.0233, 0.0233, 0.0233, 0.0233, 0.0233, 0.0116, 0.0116]

Gautos reikšmės $p(x_i,y_i)$ sustatomos į 3 formulę ir gaunamas rezultatas yra H(IP, OID) = 4.41407.

VII. IŠVADOS IR TOLIMESNI DARBAI

Atlikus mūsų pasirinktų anonimizuotų duomenų analizę, galime lengvai nustatyti, kad pagal mūsų sudarytą vertinimo skalę IP adresų neapibrėžtumo laipsnis atitinka kriterijui S AUGU, o OID skirtingų objektų kiekis tik 3, tad ir neapibrėžtumo laipsnis tik NEPATIKIMA, b et sandūros entropija atitinka kriterijų SAUGU, n es r yšiai tarp tikrų tinklų įrenginių ir jų objektų pasiskirstę pakankamai tolygiai ir jų dominuojančių daugiau nei 3 (0.0814, 0.0698, 0.0698, 0.0698, 0.0698).

Išliekamoji vertė anonimizuotame duomenų masyve iš dalies išsaugota, nes IP adresų potinklių struktūra išsaugota su jų leksikografiniu eiliškumu, OI D visi realūs, tik nutraukti ryšiai su realiais tinklo įrenginiais.

Trap pranešimų data ir laikas gali būti anonimizuojamas minimaliai (pvz., 2 dienų tikslumu). Tokiu būdu reali įvykio data nežinoma, bet tendencijas ir/ar įvykių pasikartojimus galima identifikuoti. Esant poreikiui, naudojantis prototipo galimybėmis, galima anonimizuoti n et 180 dienų tikslumu (neapibrėžtumo laipsnis d augiau n ei 6 , t. y. LABAI S AUGU), bet tada visiškai prarandamas ryšys su įvykio realia data (3 pav.).

Atlikta modelio realizacija ir j os p atikrinimas pirminiais duomenimis užsibrėžtus tikslus pateisino. Kitame modelio vy stymo etape paaiškės, ar visais atvejais yra patikima taikyti šį modelį, ar gali reikėti atlikti pataisymus, ar įtraukti papildomų išimčių. Jau dabar galima pastebėti, kad būtina atlikti visapusišką anonimizuotos i nformacijos analizę. Būtina įvertinti ir kitus, kai kuriais at vejais n ereikšmingus, parametrus, tokius kaip sąlyginė entropija ar abipusė informacija, kuri atrodo nereikšminga, bet pasikeitus situacijai ar atsiradus galimybei atskleisti vieną iš anonimizuotų duomenų aibės, galima padaryti nepataisomą žalą organizacijai ar jos įvaizdžiui su visomis blogomis pasekmėmis.

LITERATŪROS SĄRAŠAS

[1] J. S chonwalder, A. P ras, M . H arvan, J . S chippers, and R . v an d e M eent, "SNMP T raffic Analysis: Approaches, T ools, an d F irst R esults," Proc. 10t h IFIP/IEEE International Symposium o n In tegrated Network Management, May 2007.

[2] M. H arvan i r J . S chonwalder, „Prefix- and Lexicographical-order-preserving I P A ddress Anonymization," 10th I EEE/IFIP Network Operations a nd M anagement S ymposium p.519–526, 2006.

[3] Latanya S weeney, "K-anonymity: a M odel f or Protecting Privacy," 2002.

[4] M. F . C hi-Yin C how a nd X . L iu, "A P eer-to-Peer Spatial C loaking Algorithm f or Anonymous Location-Based Services.", Proceedings of the 14th annual ACM international symposium on Advances in g eographic i nformation s ystems. A rlington, Virginia, USA, 2006.

[5] Gedik B and Ling Liu, "Protecting Location Privacy with P ersonalized k -Anonymity: Architecture a nd Algorithms", IEEE Transactions o n Mo bile Computing. U niversity o f I llinois a t U rbana-Champaign. P.1-18, 2008.

[6] Raymond, J.-F., "Traffic analysis: Protocols, attacks, design issues, and open problems. In Proceedings of Designing Privacy E nhancing T echnologies: Workshop o n D esign I ssues i n Anonymity a nd Unobservability," 2000.

[7] Serjantov, A . a nd N ewman, "On t he a nonymity o f timed pool mixes. In Proceedings of the Workshop on Privacy and Anonymity Issues in Networked and Distributed S ystems. K luwer, A thens, G reece," p. 427–434, 2003.

[8] Weijia Y ang i r Sanzheng Q iao, "A no vel anonymization a lgorithm: P rivacy p rotection a nd knowledge preservation," 2009.

[9] Aggarwal, C. C., & Yu, P. S., "A general survey of privacy-preserving d ata mining models a nd algorithms. Privacy-preserving data mining," p. 11–52, 2008.

[10] Xuan S hang, K e C hen, Lidan Shou, G ang Chen, Tianlei H u, "(k,P)-Anonymity: T owards P attern-Preserving Anonymity o f T ime-Series D ata," CIKM'10, O ctober 26 –30, T oronto, O ntario, Canada, 2010.

[11] RFC 3411 – 3418 SNMP pr otokolo standartų aprašymas [prieiga per internetą] http://www.rfc-editor.org/rfc/std/std62.txt, žiūrėta 2012-01-20

[12] M. E dman, B . Yener, "On Anonymity i n a n Electronic S ociety: A S urvey of Anonymous Communication S ystems," ACM C omputing Surveys, Vol. 42, No. 1, Article 5, 2009.

[13] M. Bezzi, "An Entropy based method for Measuring Anonymity," Proceedings o f the th ird in ternational conference o n s ecurity an d p rivacy i n communication networks. N ice, F rance. p . 2 8 – 32, 2007.

[14] J. Xu, J. Fan, and M. H. Ammar, "Prefix-Preserving IP Address Anonymization: Measurement-based Security Evaluation and a New Cryptography-based Scheme," Proceedings of the 10th IEEE International Conference on Network Protocols (ICNP'02) , 2002.

[15] Paul Meagher, "Calculating Entropy for Data Miners," [prieiga per internetą] http://onlamp.com/pub/a/php/2005/03/24/joint_entropy.html?page=1, žiūrėta 2013-01-15

[16] Guihua Duan, Weiping Wang+, Jianxin Wang, Luming Yang, "A new anonymity measure based on partial entropy," IEEE Communications Society publication in the ICC, 2008.

INFORMACINĖS TECHNOLOGIJOS ● IVUS 2013 ● ISSN 2029-249X ● eISSN 2029-4824

Projekto valdymo sistemos žaidimizavimas

Darius Ašeriškis
Programų inžinerijos katedra
Kauno technologijos universitetas
Studentų g. 50-406
Kaunas, Lietuva
daraser@gmail.com

Santrauka. Žaisti pradedame visi nuo vaikystės ir žaidimas yra natūrali žmogaus vystymosi dalis. Kartu su Web 2.0 tendencijomis keičiasi žmonės ir visuomenė tampa atvira. Būtent tokie pokyčiai keičia žmonių poreikius, todėl keičiasi ir programinė įranga. Straipsnyje aprašomas praktinis projektų valdymo sistemos žaidimizavimas.

Reikšminiai žodžiai: žaidimizacija, projektų valdymo sistemos, Web 2.0.

I. ĮVADAS

Žaisti mes pradedame nuo ankstyvos vaikystės ir žaidimas yra natūralus reiškinys, priešingai nei darbas. Žaidimizacija – tai procesas, kai sistemai, neturinčiai nieko bendro su žaidimais, yra suteikiama žaidimo mechanika ir elementai. Žaidimizacijos terminas pasiūlytas 2008 m. ir ypač išpopuliarėjo 2010-2011 m. Žaidimizacija kyla iš šeimos įvairių žaidimų metodikų [1], tarp kurių yra rimtų žaidimų (ang. „serious games“), kurie orientuoti į mokymą, žaismingas dizainas (ang. „playfull design“) ir kiti. Bendru atveju žaidimizacija apibrėžiama kaip nežaidybinės veiklos transformavimas į žaidimą, t.y. nežaidybinei veiklai sukuriamas žaidimo sluoksnis.

Web 2.0 atnešė mūsų visuomenei didelių permainų. Iš uždaros visuomenės mes tampame atvira visuomene. Web 2.0 permainos padeda populiarinti žaidimizavimą mūsų kasdieninėje veikloje. Atvirumas, permatomumas tampa mūsų kasdienybe, žmonės nori būti pastebimi, matomi ir įvertinami darbe [2]. Žaidimizacija ir Web 2.0 verčia natūraliai uždaras organizacijas keistis bei keisti savo veikimo procesus [3]. Šioms kitokioms organizacijoms reikia ir kitokios projektų valdymo įrangos.

Žadimizacija yra svarbus elementas norint nesudėtingai padidinti atvirumą ir permatomumą įmonėje. Šiame straipsnyje pateiksiu praktinį tradicinės projektų valdymo sistemos žaidimizavimo pavyzdį. Šiandien sutinkame vis daugiau sistemų, kurioms yra suteiktos žaidimų savybės. Labiausiai įkvepiantys žaidimizacijos taikymo atvejai:

- „Taskville“ [2] – tarptautinės ir pasiskirsčiusios darbuotojų komandos sujungimas ir matomumo organizacijoje padidinimas statant bendrą komandos miestą.
- „Eco-driving“ [4] – projektas, kuris pristato sistemą, skatinančią vairuotojus tausoti gamtą.
- „Power House“ [5] – projektas, kuris skatina tausoti gamtą taupant elektros energiją.

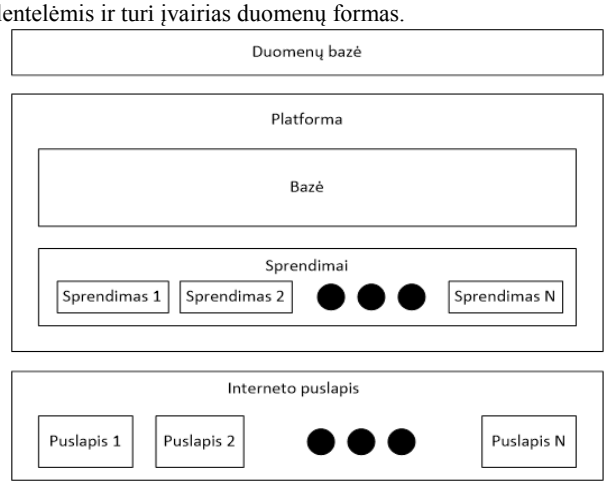
Žaidimizuojant sistemą būtina įvertinti technologinius ir socialinius aspektus [1]:

- Sąsajos šablonus, pavyzdžiui, ženkliukus, lygius ar reitingus.
- Žaidimų dizainą ir (arba) žaidimų mechaniką.
- Žaidėjų tipus.
- Žaidimų sąsajos šablonus projektuose, kuriuose jie sėkmingai taikomi, kaip „Wikipedia“ ir „StackOverflow“.
- Naujos kartos darbuotojus, kurie užaugo su kompiuteriniais žaidimais ir turi įgūdžių, kuriuos norėtų panaudoti įmonės.

Pastaruoju metu keičiasi darbo samprata, bandoma pašalinti darbas – rimtumas sąsają. Taip bandoma pakeisti darbo suvokimą iš prievolės į malonumą.

II. PROJEKTŲ VALDYMO SISTEMA

Projektų valdymo sistema – programinė įranga, skirta valdyti projektus, registruoti darbus ir valdyti darbo procesą. Tokio tipo sistemos dažniausiai manipuliuoja įvairiomis lentelėmis ir turi įvairias duomenų formas.



1 pav. „Trogon“ aukšto sistemos lygio architektūra

Šiame straipsnyje pateikiamas praktinis „Trogon“ projektų valdymo sistemos žaidimizavimas. 1 paveikslėlyje pateikta „Trogon“ projektų valdymo sistemos aukšto lygio architektūra. Pastaroji susideda iš trijų pagrindinių sluoksnių. Interneto svetainės puslapių sluoksnis – tai sistemos sluoksnis, atsakingas už vartotojo sąsają. Platformos sluoksnis – tai sistemos sluoksnis, sudarytas iš bazinių sistemos elementų ir atskirų verslo logikos sprendimų. Duomenų bazės lygis atsakingas už visos sistemos informaciją. Mūsų atveju žaidimizacija yra vienas iš sprendimų, sudarančių galutinę sistemą.

„Trogon“ buvo realizuotas taikant Microsoft ASP.NET technologiją ir C# programavimo kalbą, duomenų bazei buvo panaudota MSSQL duomenų bazė.

III. PROJEKTŲ VALDYMO SISTEMOS ŽAIDIMIZAVIMAS

A. *Taisyklės*

Neįmanoma sistemos žaidimizuoti, neįvedus žaidimo taisyklių. Taigi sistemos žaidimizavimas pradedamas nuo žaidimo taisyklių sukūrimo.

Pagrindinės taisyklės:

- Kiekvienas darbas gali turėti parinktą ženklelį ir nurodytą planuojamą trukmę.
- Kiekvienas darbuotojas nurodo, kiek užtruko atlikdamas darbą.
- Kiekvienam darbui nurodomas sudėtingumas.
- Ženklelis pelnomas, jei jis išlaikomas iki darbo užbaigimo. Jį galima prarasti projektų vadovo sprendimu, pavyzdžiui, jei prasta atlikto darbo kokybė ar per ilgas atlikimo laikas. Kokybės užtikrinimo metu ženklelis gali būti pašalintas, jei randama daug trūkumų.

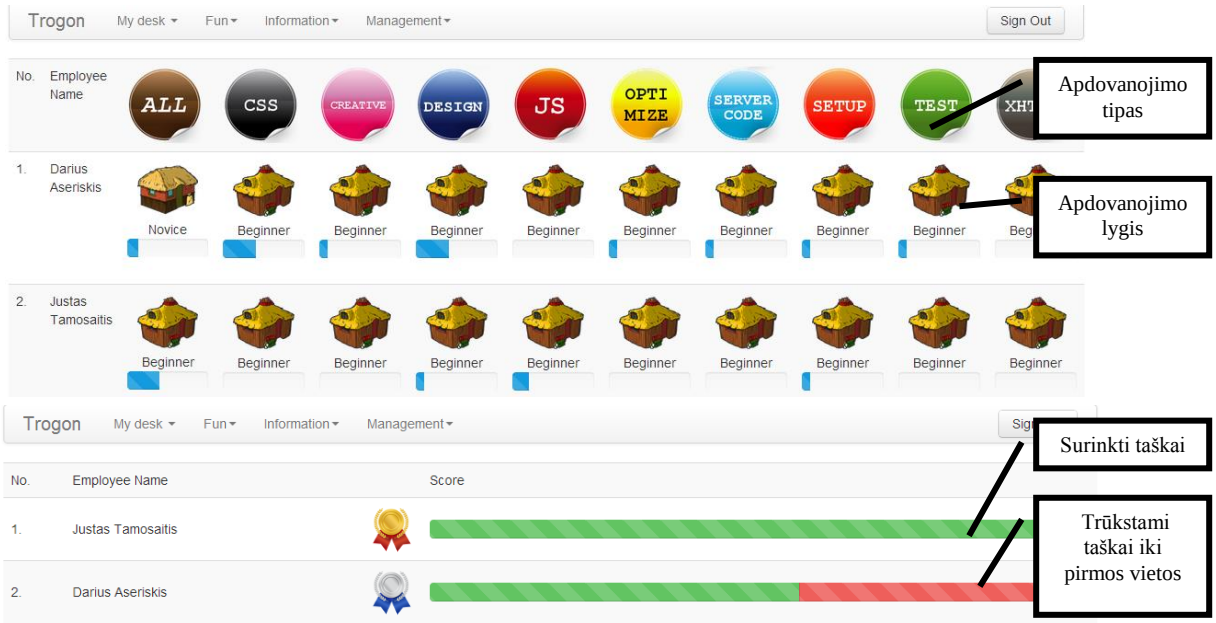
Taškų skaičiavimas:

- Sumuojamos visos tą savaitę baigtų darbų planuojamos trukmės.
- Rezultatas dauginamas iš darbo sudėtingumo koeficiento.
- Pridedami papildomi taškai už ankščiau baigtus darbus.
- Atimami taškai už vėliau baigtus darbus.
- Papildomi taškai pridedami už gautus ženkliukus:
  - Gautas ženkliukas prideda papildomus 20% darbo vertės.
  - Per savaitę surinkus 5 ir daugiau vienos rūšies ženklelių, pridedami papildomi 20% už visus darbus, kuriais buvo pelnyti tos rūšies ženkleliai.

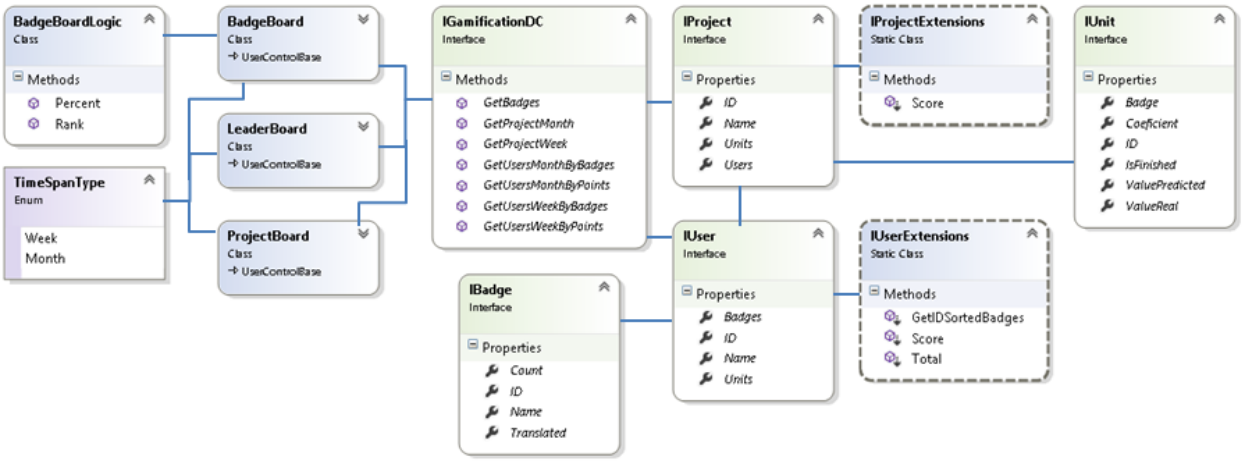
Žaidimo pabaiga ir apdovanojimai:

- Pasibaigus savaitei, pagal surinktus taškus renkamas geriausias savaitės darbuotojas, kuriam suteikiama teisė rinktis iš projektų vadovo siūlomų darbų.
- Mėnesio geriausiam darbuotojui suteikiama viena papildoma atostogų dieną.

Jokiu būdu negalima laikyti šių taisyklių vienintelėmis ir nekeičiamomis. Tai yra tik siūlomos taisyklės, o galutinė jų versija priklauso nuo įmonės tikslų, politikos ir galimybių.



2 pav. Žaidimizacijos elementai – ženklelių statistika ir žaidėjų reitingai



3 pav. Įgyvendinimo klasių diagrama

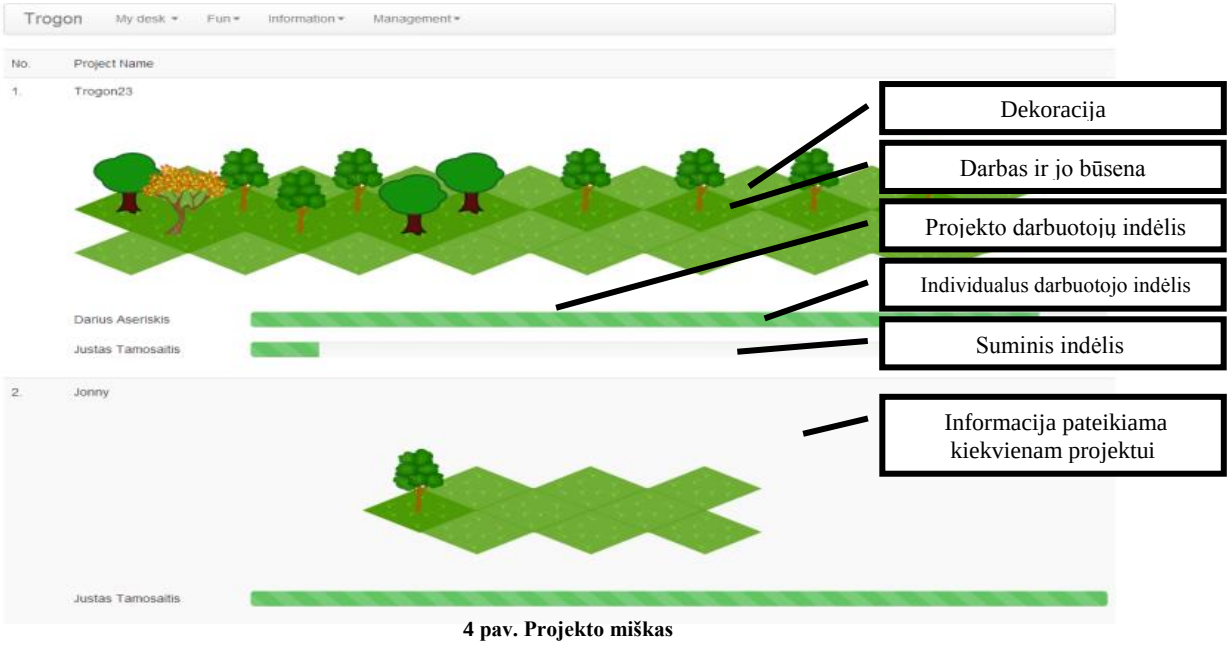
B. Elementai

Žaidimas susideda iš savaitės ir mėnesio rezultatų lentelių (2 pav.). Taip pat pateikiami ir surinkti ženkleliai per savaitę ar mėnesį. Žaidimizacija įdiegta kaip papildomas visos sistemos elementas, sukuriant naują meniu punktą „Fun“. Ženkleliai yra įvairių tipų ir gali kisti priklausomai nuo įmonės veiklos. Apdovanojimo lygis atvaizduojamas tiek progreso laukeliu, tiek piktograma su tekstiniu aprašymu. Žaidėjų reitingų puslapyje rodomi visų darbuotojų pasiekimai per pasirinktą laikotarpį. Žaliai žymimi vartotojo surinkti taškai, raudonai – taškai, kuriuos reikia surinkti norint pasivyti žaidėją, esantį pirmoje vietoje. Pirmos trys vietos atskiriamos medaliais.

C. Techninis sprendimas

Klasių diagramoje (3 pav.) matome sistemą, padalintą į dvi dalis: atvaizdavimo ir duomenų. Duomenis sudaro bendras duomenų kontraktas „IGamificationDC“. Šiuo duomenų kontraktu buvo apibrėžti žaidimizavimo elementai. Pagrindinių duomenų sąsajos dedamosios:

- „IProject“ – visi projekto duomenys: pavadinimas, darbuotojai ir darbų informacija.
- „IUser“ – darbuotojo duomenys, kurie sudaryti iš ženklelių, darbuotojo informacijos ir darbų informacijos.
- „IUnit“ – darbų informacija, kuri sudaryta iš ženklelio,



4 pav. Projekto miškas

sudėtingumo koeficiento, darbo pabaigos žymės, prognozuotos ir realios darbo trukmės.

- „IBadge“ – ženklelio informacija, sudaryta iš ženklelių kiekio ir pavadinimo.

„IProjectExtensions“ ir „IUserExtensions“ suteikia atitinkamai „IProject“ ir „IUser“ reikalingą skaičiavimo logiką, atitinkančią ankščiau apibrėžtas žaidimo taisykles. „BadgeBoard“, „LeaderBoard“ ir „ProjectBoard“ grafiniai komponentai atsakingi už žaidmizacijos elementų atvaizdavimą.

D. Galimybės ir rezultatas

Vienas pagrindinių žaidimizuotos sistemos privalumų – tai galimybė pabėgti nuo klasikinio sistemos suvokimo. Projekto miškas – tai „Trogon“ projektų valdymo sistemoje pristatytas naujas projektų vizualizavimo metodas.

Panagrinėkime 4 paveikslėlį. Žolės spalva rodo darbo būseną: šviesi žolė – dekoracija, tamsi žolė – pradėtas darbas, tamsi žolė su medžiu – baigtas darbas. Medžiai kinta, priklausomai nuo darbo vertės taškais, kurie gaunami pagal anksčiau minėtas žaidimizacijos taisykles.

Pagrindiniai tokio atvaizdavimo privalumai:

- Didesnis informacijos matomumas lyginant su tradiciniu atvaizdavimu lentele.
- Matomas individualaus darbuotojo indėlis.
- Padeda suvokti projekto sudėtingumą.

Šitoks informacijos atvaizdavimo transformavimas yra neatsiejama žaidimizacijos dalis, tačiau reikia nepamiršti, kad būtina išlaikyti projektų valdymo sistemos konservatyvumą.

Taigi gavome žaidimizuotą sistemą, kurios tikslas pakeisti darbuotojų nusiteikimą darbui. Būtent bandymas pakeisti darbuotojų nusiteikimą vienu ar kitu klausimu yra ryškiausia žaidimizacijos savybių [4].

IV. IŠVADOS

Straipsnyje pristatytas praktinis projektų valdymo sistemos žaidimizavimo pavyzdys. Šiam pavyzdžiui pristačiau žaidimo taisykles ir išskyriau žaidimo elementus. Taip pat buvo pateiktas sistemos praktinis įgyvendinimas ir parodyta, kad sistemos žaidimizavimas leidžia pateikti informaciją kiek kitaip negu klasikiniais būdais.

Sistemos privalumai: sistema papildo įprastą projektų valdymo procesą žaidimo elementais, padidina organizacijos skaidrumą ir atvirumą, praktinė realizacija leidžia integraciją į kitas projektų valdymo sistemas.

Sistemos trūkumai: nėra tyrimų, įrodančių projektų valdymo sistemos žaidimizavimo naudą, trūksta sistemos išvystymo, kuris pasiekiamas ilgai vystant tokio tipo sistemą. Tačiau žaidimizacija kitose srityse rodo gerus rezultatus, kurie ir skatina žaidimizacijos taikymą vis naujose srityse.

Būsimi darbai: žaidimizuotos projektų valdymo sistemos tyrimas, siekiant įrodyti tokios sistemos naudą programinę įrangą kuriančiai įmonei.

REFERENCES

[1] Yongwen X. Literature Review on Web Application Gamification and Analytics. CSDL Technical Report 11-05.

[2] Nikkila S., Linn S., Sundaram H., Kelliher A. Playing in Taskville: Designing a Social Game for the Workplace. Gamification Workshop, CHI 2011.

[3] Ašeriškis D., Tamošaitis J. Kitokia PĮ kompanija, Mag&Doc IT2012: tarpuniversitetinės konferencijos pranešimų medžiaga [Kaunas, 2012 m. balandžio 20 d.]. Kaunas, 2012, p. 49-52.

[4] Inbar O., Tractinsky N., Tsimhoni O., Seder T., Driving the scoreboard: Motivating eco-driving through in-car gaming. Gamification Workshop, CHI 2011.

[5] Reeves B., Cummings J. J., Anderson D. Leveraging the engagement of games to changes energy behavior Gamification Workshop, CHI 2011.

INFORMACINĖS TECHNOLOGIJOS ● IVUS 2013 ● ISSN 2029-249X ● eISSN 2029-4824

Lietuviškų datų atpažinimas

Justas Balčas
Vytauto Didžiojo universitetas
Kaunas, Lietuva
balcas.justas@gmail.com

Žygimantas Medelis
UAB "Tokenmill"
Vilnius, Lietuva
zygimantas.medelis@tokenmill.lt

Tomas Krilavičius
Vytauto Didžiojo universitetas
Kaunas, Lietuva
Baltijos pažangių technologijų institutas
Vilnius, Lietuva
t.krilavicius@gmail.com

***Santrauka* — Lietuviškų datų atpažinimas reikalingas media monitoringe, policijos ataskaitų, ligonio istorijos analizei. Anotuojant tekstyną, buvo atsižvelgta į lietuvių kalbos standartus ir neformalios kalbos datos užrašymo formatus. Apibrėžti angliškos datos užrašymo formatai, jų naudojimas internetiniuose dienraščiuose. Aptarti anotavimo įrankiai, pasirinkus GATE programą, parašytas taisyklių rinkinys lietuviškų datų atpažinimui. Atliktas rankinis ir automatinis anotavimas, kurio metu buvo anotuota 500 straipsnių iš įvairių internetinių šaltinių ir pateikiami sukurto automatinio datų žymėjimo rezultatai ir aprašyti klaidos vertinimo metodai.**

Reikšminiai žodžiai* — *kalbos technologijos, kompiuterinė lingvistika, esybių atpažinimas.

I. Įvadas

Informacinių technologijų plėtra ir sklaida per paskutinius pora dešimtmečių labai pakeitė mūsų gyvenimą. Atsirado nauji įrankiai ir įpročiai, gyvenimas be mobilaus telefono, žiniatinklio, elektroninio pašto, kai kam ir be socialinių tinklų (pvz., Facebook), pokalbių programų (pvz., Skype), paieškos sistemų (pvz., Yahoo, Google, Bing) yra sunkiai įsivaizduojamas. Sukuriami beprecedenčiai skaitmeninių duomenų kiekiai. Visi šie pokyčiai ne tik suteikia naujas galimybes, bet ir naujus reikalavimus informacinių ir telekomunikacinių sistemų kūrėjams. Atsiranda poreikis naujiems metodams bei technologijoms, nes klasikinių nepakanka.

Iš teksto galima atpažinti ir suprasti daug informacijos, net ir ne mokant tos kalbos, bet tam prireikia identifikavimo programos. Objektai, kuriuos galima išskirti:

1. Miestų pavadinimai.
2. Asmenvardžiai, t.y. vardai, pavardės.
3. Datos (mėnesiai, dienos, valandos).
4. Elektroninio pašto, internetinio puslapio adresai.
5. Vietovardžiai.
6. Valiutos.

Esybių nustatymo algoritmai [1] (angl. *Named Entity Recognition – NER*) leidžia pasirinkto tipo tekstuose identifikuoti nurodytus faktus ar teksto elementus.

Visi turimi dokumentai gali būti analizuojami ir anotuojami rankiniu būdu prie kiekvieno tokio dokumento prijungiant atitinkamus metaduomenis pateikiančius reikiamus faktus. Tačiau rankinis dokumentų anotavimas reikalauja ženklių išteklių, o rankiniu būdu anotuoti didelius bei nuolat besikeičiančius tekstinės informacijos dokumentus praktiškai neįmanoma. Todėl esybių nustatymas turi būti automatizuojamas, naudojant atitinkamus algoritmus.

Natūralios kalbos apdorojimo algoritmai, atliekantys esybių identifikavimą (angl. *named entitiy recognition, NER*), pasižymi dideliu sudėtingumu, reikalaujančiu gebėjimų ne tik greitai ir tiksliai apdoroti didelius informacijos kiekius, bet ir mokėti atskirti žodžių naudojimo kontekstą bei dokumento tematiką.

Lietuviškos datos atpažinimo metodai aprašyti [2], kuriame identifikuojama dauguma galimų datos užrašymo formatų. Vertinimui naudojamas 400 tūkstančių žodžių turintis tekstynas.

Angliškų datų atpažinimui galima naudotis jau sukurtomis taisyklėmis, parašytomis GATE Developer [3] programai. Užrašoma data ne visada reiškia, kad ji yra data. Data, laikas ir metai GATE Developer programoje yra pažymimi atskirai (TempDate, TempTime, TempYear) ir formuojama data iš išsaugotų duomenų. Taip atliekama, nes, pvz., 1234 gali reikšti metus ar tik skaičių, 9 – 10 gali reikšti tarpą tarp laiko ar tarpą tarp dienų, arba dviejų skaičių skirtumą.

II skyriuje aprašomi datos užrašymo formatai lietuvių ir anglų kalboje. III skyriuje aprašyti esybių išskyrimo metodai ir įrankiai (Regexp, JAPE, GATE, LingPipe); IV skyriuje pateikiamos JAPE taisyklės lietuviškos datos atpažinimui. V skyriuje pateikiami atlikti eksperimentiniai tyrimai; VI skyriuje pateikiamo išvados, pasiektos atliekant šį tyrimą.

II. Datos užrašymo formatai

A. *Datos formatas lietuvių kalboje*

Kalendorinė data yra skirta pažymėti svarbias dienas, pvz., Tėvo dieną, straipsnio publikavimo datą, redagavimo datą, gimimo, mirties datas.

1. LR standartuose numatytas datos rašymo formatas [4].
   1.1. Rašoma skaitmenimis (2011-11-05 arba 2011 11 05).
   1.2. Ilgesniu būdu, pvz., 2011 m. gegužės 7 d., 2012 m. gruodžio mėn.
   1.3. Data užrašoma, pvz., 2011 liepos 20, 2011, sausio 14, 2011 liepos mėn. 20 d.
   1.4. Skiriant tašku „.“ pvz., 2011.11.01.
   1.5. Sutrumpinant mėnesių ir dienų simbolius, nerašant 0.
2. LR standartuose numatytas datos rašymo formatas [5].

2.1. 3 va l. 15 m in. arba 3.15 va l. (arba t ik 3.15). Pvz., Pradžia 19.30 val. Įėjimas nuo 11.30 iki 18 val. (arba nuo 11.30 iki 18.00), 18.00 – Žinios.
2.2. Nurodant laikotarpį vartojamas brūkšnys (ne brūkšnelis) ir nepaliekama tarpų, pvz., 1996–2005 m. programa, spalio 6–7 d., 14.00–17.00.

3. LR standartuose neapibrėžti datos rašymo formatai.
3.1. Metai.
3.1.1. Nusakant amžių: XXa., XXIa.
3.1.2. Žodiniu f ormatu: du tūkstančiai vienuoliktų metų, tūkstantis devyniasdešimt šeštais.
3.1.3. Skaitmeniniu f ormatu, bet p irmiau r ašoma diena – mėnuo – metai, pvz., 15-02-2012, 11 03 2011.
3.2. Mėnesiai.
3.2.1. Žodiniu f ormatu nur odant vasaros, ž iemos, pavasario, rudens laikotarpį, pvz., pirmąjį vasaros mėnesį.
3.3. Dienos.
3.3.1. Nurodant tipą: darbo dienomis, savaitgaliais.
3.3.2. Nurodant šventę: pirmąjį gegužės sekmadienį (Motinos diena).
3.4. Laikas.
3.4.1. Žodiniu formatu: šešioliktą valandą.

B. *Datos formatas anglų kalboje*

Anglų kalba turi žymiai daugiau datos formatų [6]. Taip yra dėl to, kad anglų kalba naudojama ne vien Jungtinėje karalystėje, b et ir J AV, Australijoje, N aujojoje Z elandijoje, Indijoje ir kaip tarptautinė kalba. Skirtumas nuo lietuvių kalbos – anglų kalboje dažnai yra trumpinami mėnesių pavadinimai arba mėnesio pavadinimas pradedamas rašyti iš mažosios raidės, kuomet rašomas tik mėnuo ir metai. Lietuviškų datų užrašyme yra svarbu apibrėžti angliškų datų užrašymą, nes lietuviškuose por taluose ne turėtų būti naudojamas angliškos datos formatas, bet dažnai pasitaiko.

Datos formatai naudojami anglų kalboje:

1. Data pradedama dienos numeriu arba dienos pavadinimu.
1.1. Neapibrėžti tarp tarptautinių ISO 8601 standartų [7] („9 November 2003" arba "9. November 2003", 9/11/2003 arba 9.11.2003, 09-Nov-2003, 9-11-2003, [The] 9th [of] November 2003, Sunday, 9 November 2003, 9/xi/03, 9.xi.03).
2. Pradedama metais.
2.1. Apibrėžti ISO 8601 standartuose (2003-11-09, 2003 November 9 , 2003Nov9, 2003Nov09, 2003-Nov-9, 2003-Nov-09).
3. Pradedama mėnesiu.
3.1. Atitinka Kanados, JAV ir dar keletos šalių standartus (Sunday, N ovember 9, 2003 , November 9, 2003 , Nov. 9, 2003).
4. Laiko užrašymas.
4.1. Skiriant va landas, m inutes, s ekundes d vitaškiu a rba nededant jokio skirtuko (hh:mm:ss arba hhmmss).
4.2. Naudojant A M a rba PM. Neatitinka I SO 8601 standartų, bet atitinka JAV datų rašymo standartus (8:45am, 3:17 PM).
4.3. JAV naudojamas l aiko f ormatas [8]. Atitinkantis JAV standartus (11:32 am EST).

III. ESYBIŲ IŠSKYRIMO METODAI IR ĮRANKIAI

Esybių atpažinimas naudingas daugeliui kalbos apdorojimo technologijų, tokių kaip informacijos gavyba [9], klausimų atsakymas – apdorojimas [10], m ašininis v ertimas [11], semantinių tinklų metaduomenų pažymėjimo pagrindas [12]. Įprastai esybių atpažinimo s istemos yra na udojamos s ubjekto identifikavimui molekulinėje biologijoje, b ioinformatikoje ir medicininėse natūralios kalbos apdorojimo bendruomenėse [13]. Esybių atpažinimo sistemos pirmiausia buvo pradėtos naudoti esybių žymėjimui straipsniuose, vėliau automatinė turinio gavyba buvo pradėta naudoti ir neformalaus teksto stiliuose, tokiuose kaip tinklaraščiai, forumai, tekstų nuorašai iš telefono pokalbių.

A. *NER metodai*

Taisyklėmis grįsti metodai – vienas iš populiariausių esybių atpažinimo metodų. Šiuo atveju užrašomos taisyklės, skirtos atpažinti tekste tam tikroms esybėms, p vz., datoms, įmonių pavadinimams, pavardėms ir pan. Dažnai taisyklėmis grįsti metodai derinami su žodžių sąrašais.

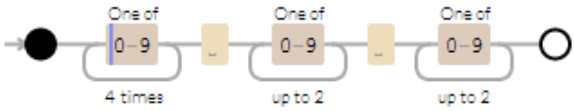
1) Regexp

Dažnai t am na udojamos r eguliarios i šraiškos (angl. Regexp) [14]. Skaitmeninės datos (2013 02 2 8) a tpažinimo Regexp išraiška:

LENTELĖ I. Regex užklausa skaitmeninei datai

```
[0-9]{4} [0-9]{1,2} [0-9]{1,2}
```

Skaitmeninės datos (2013 02 28) reguliariosios išraiškos užklausa



Paveikslas 1. Regexp užklausa grafiškai [15]

LENTELĖ II. Datos, užrašytos ilgesniu būdų Regexp atpažinimo išraiška

```
[0-9]{4} ?(m\.?|met[ųu])?
?(sausio|vasario|kovo|baland[zž]io|gegu[zž][eė]s|bir[zž]e
lio|liepos|rugpj[uū][cč]io|rugs[ėe]jo|spalio|lapkri[cč]io
|gruod[zž]io) (m[ėe]n(\.|esio)?)? ?([0-3]?){1}[0-9]{1}
(d(\.|ien(os|a|ą)))?
```

Taisyklės, užrašytos reguliariosiomis išraiškomis, yra sunkiai skaitomos, tad lietuviškų datų atpažinimui p asirinkta JAPE taisyklių rašymo metodika.

2) JAPE

JAPE [16] gramatikos taisyklės susideda iš modelių, veiksmų, taisyklių. Gramatika v isada turi d vi puses – kairę ir dešinę. Taisyklėmis yra nustatomas anotacijos modelis, kurių kairioji pusė gali būti reguliarios išraiškos o peratoriai (pvz., *,?, +), o dešinioji pusė apibrėžia, kokių veiksmų reikia imtis, kad anotacija atitiktų peržiūrint aptiktos a notacijos manipuliavimo modelį.

B. NER įrankiai

1) GATE

GATE [3] (General Architecture for Text Engineering) yra teksto analizės sistema, sukurta J ava ka lba i r pr ieinama ka ip atviro kodo programinė įranga. Tai teksto apdorojimo karkasas, kuris suteikia galimybę, naudojant interaktyvią grafinę vartotojo sąsają, kurti programinės įrangos komponentus, skirtus i nformacijos g avybai i r jos apdorojimui. G ATE Developer y ra l aisvai platinimas produktas, naudojantis specializuotas duomenų struktūras bei algoritmus, tokius kaip anotacijų grafus, baigtinių automatų ir atraminių vektorių metodus. Šiame darbe pateikiamos taisyklės yra naudojamos kaip GATE Developer programos įskiepis.

2) LingPipe

LingPipe [17] įrankio pagrindinė paskirtis – identifikuoti ir išskirti žmonių, organizacijų ar vietovių pavadinimus iš naujienų srautų. LingPipe pagalba galima klasifikuoti Twitter trumpųjų pranešimų sistemos paieškos rezultatus ir suskirstyti į kategorijas bei panaudoti įrankį rašybos klaidų tikrinimui. Šis įrankis naudoja Java platformą. Kaip ir dauguma kitų modulių, LingPipe e sybių atpažinimui naudoja žodynus, reguliariąsias išraiškas, taip pat mašininio mokymo metodus.

IV. TAISYKLĖS DATOS ATPAŽINIMUI

Datos atpažinimui pasirinkta JAPE taisyklių rašymo metodika. Jos k odas y ra l engviau s kaitomas, g alima u žrašyti sudėtingas užklausas, išvengiama klaidų aprašant vieną ilgą eilutę (kaip R egexp d atos a tpažinimo), išskaidant esybės atpažinimą į atskiras dalis.

Iš v iso buvo išskirstytos 24 datų rašymo formos, kurios aprašytos 25 JAPE taisyklėmis:

1. Skaitmeniniai metai.
2. Skaitmeninė data.
3. Data su trumpiniais ir mėnesio pavadinimu.
4. Data be trumpinių ir su mėnesio pavadinimu.
5. Žodiniai metai.
6. Skaitmeninių metų laikotarpis.
7. Skaitmeniniai mėnesiai.
8. Žodiniai mėnesiai.
9. Skaitmeniniai metai su mėnesiu.
10. Žodiniai metai su mėnesiu.
11. Žodinis mėnuo su diena.
12. Skaitmeninių mėnesių laikotarpis.
13. Žodinių mėnesių laikotarpis.
14. Skaitmeninė diena.
15. Žodinė diena.
16. Laikas skaitmeniniu formatu.
17. Skaitmeninis laikas su trumpiniais.
18. Laikas užrašytas žodiniu formatu.
19. Skaitmeninis laiko laikotarpis.
20. Neapibrėžti datos standartai.
21. Amžius (romėniški skaitmenys).
22. Metai, užrašyti žodiniu formatu.
23. Žodiniu f ormatu nur odant va saros, žiemos, pavasario, rudens laikotarpį.
24. Šventės diena.
25. Žodiniu f ormatu nus akomas l aikas: vi durdienis, vidurnaktis, po pietų.

A. JAPE taisyklės datos atpažinimui

Datos a tpažinimui, pa sinaudojus M acro (naudojamas ka ip greitos komandos, žinomų struktūrų atpažinimui), apsibrėžiame tikslias vietas, žinodami kur t ekste ko kia bus naudojama reikšmė. Macro leidžia sukurti apibrėžimą simbolio, kuris gali būti daug kartų naudojamas JAPE taisyklėse.

LENTELĖ III. Macro funkcijos skaitmenų atpažinimui

```
Macro: Metai
({Token.kind=="number", Token.length==4})
Macro: Menuo
(({Token.kind=="number", Token.length==1}) |
({Token.kind=="number", Token.length==2}))
Macro: Diena
(({Token.kind=="number", Token.length==1}) |
({Token.kind=="number", Token.length==2}))
```

LENTELĖ IV. Macro funkcija skirtukų atpažinimui

```
Macro: Skirtukai
((({Token.string =="."}) |
({Token.string ==","}) |
({Token.string=="-"}) |
({Token.string=="/"})))
```

LENTELĖ V. JAPE taisyklė Skaitmeninių metų atpažinimui

```
Rule: SkaitMetai
((Metai)):GautaData -->
:GautaData.Data = {rule = "Skaitmeniniai metai"}
```

LENTELĖ VI. Skaitmeninės datos atpažinimo taisyklė

```
Rule: DataSkaitmenine
(((Metai)(Menuo)(Diena)) |
((Metai)(Skirtukai)(Menuo)(Skirtukai)(Diena)))
:GautaData -->
:GautaData.Data = {rule = "Skaitmeninė data"}
```

LENTELĖ VII. Datos, užrašytos su mėnesio pavadinimu ir su arba be trumpinių

```
Rule: VardMetSuMen
(({Lookup.majorType == VardMetai})
({Lookup.majorType == Datos}) |
({Lookup.majorType == VardMetai})
({Lookup.majorType == Datos})
```

```
(Trumpiniai)|

({Lookup.majorType == VardMetai}) (Menuo) (Trumpiniai))
:GautaData -->

:GautaData.Data = {rule = "Žodiniai metai su skaitmeniniu
mėnesiu ir vardiniu mėnesiu, ir su trumpiniu ar be"}
```

V. EKSPERIMENTINIAI TYRIMAI

A. *Kokybės įvertinimo metodai*

Esybių atpažinimo sistema paprastai yra vertinama, remiantis jos siūlymais su žmogaus kalba. Tai gali generuoti penkias skirtingas klaidas:

1. Sistema identifikuoja esybę, kurios nebuvo.
2. Esybė buvo sistemos visai klaidingai atpažinta.
3. Sistema atpažino esybę, bet suteikė jai neteisingą etiketę.
4. Sistema atpažino esybę, bet klaidingai atpažino jos ribas.
5. Sistema suteikė esybei neteisingą etiketę, bet klaidingai atpažino jos ribas.

Sistemos rezultatas priklauso nuo daromų klaidų kiekio.

1 ir 2 tipo klaidos užrašomos lentelėje taip [18]:

1. Teisingai priskirta klasė (tp, true positive).
2. Neteisinga priskirta klasė (fp, false positive).
3. Neteisinga nepriskirta klasė (fn, false negative).
4. Teisingai nepriskirta klasė (tn, true negative).

Precision [19] – tikslumas, t.y. kuri dalis iš atpažintų datų tikrai yra datos (teisingai atpažinti, padalinti iš teisingai atpažintų ir neteisingai atpažintų sumos):

$$precision = \frac{tp}{tp+fp} \quad (1)$$

Recall [19] – išsamumas, t.y. kuri dalis iš datų atpažinta (teisingai atpažinti, padalinti iš teisingai atpažintų ir neteisingai neatpažintų sumos, t.y. visų datų):

$$recall = \frac{tp}{tp+fn} \quad (2)$$

F įvertis [20] – matas, kuris kombinuoja tikslumą ir išsamumą:

$$F_\beta = (1+\beta^2) * \frac{precision*recall}{(\beta^2 * precision)+recall} \quad (3)$$

Eksperimentų rezultatuose naudojamas F_1 matas.

Tekstynas

Tekstynui pasirinkti 500 atsitiktinių straipsnių iš internetinių dienraščių:

1. 200 Straipsnių iš www.delfi.lt.
2. 200 straipsnių iš http://www.culture.lt/satenai/.
3. 100 straipsnių iš įvairių blogų.

Iš viso 500 anotuotų dokumentų, 516336 termų. Ilgiausio straipsnio žodžių kiekis – 4305 žodžiai, trumpiausio – 210 žodžių.

B. *Eksperimentų rezultatai*

Anotuoto tekstyno ir parašytų JAPE taisyklių rinkinio rezultatai.

Rasta datų, rankiniu būdu anotavus tekstyną: 1944.

Parašytų JAPE taisyklių rinkinio rezultatai:

1. Sistema spėjo esybę, kurios nebuvo – 115.
2. Esybė buvo visai klaidingai atpažinta – 184.
3. Atpažino esybę, bet priskyrė neteisingą etiketę – 0.
4. Atpažino esybę, bet klaidingai atpažino ribas – 210.
5. Suteikė esybei neteisingą etiketę, bei klaidingai atpažino ribas – 0.

LENTELĖ VIII. Rezultatai

Teisingai teigiami (TP)	Klaidingai teigiami (FP)
1734	184
Atpažinta data	Nesama data
Klaidingai neigiami (FN)	**Teisingai neigiami (TN)**
210	115
Neatpažinta data	Nepriskirta teisinga klasė datai

Dažniausiai pasitaikančios klaidos:

1. Nurodytus skaičius sistema pažymi kaip datą, pvz., 1997, 2000, 3500, 19 – 20, 2000 - 2010. Reikia įdėti, kaip ir angliškų datų atpažinime, laikinuosius kintamuosius, kurie būtų išsaugomi tik tuo atveju, jei tekste yra mėnesio ir dienos užrašymas.
2. Nurodyti mėnesiai ne visada reiškia mėnesį (tekstyne mėnuo liepa reiškė vardą, bet ne mėnesį).
3. Taisyklių rinkinyje nenustatytas taisyklių prioritetas, todėl datos (2011 – 2012 metų, kovo – balandžio mėn.) yra atpažįstamos, bet radus pirmą atitikmenį, kitos taisyklės nėra pritaikomos.

Pagal pateiktas kokybės įvertinimo metodų formules, gauti tokie parašytų JAPE taisyklių rinkinio rezultatai:

1. Tikslumas – 0.903.
2. Išsamumas – 0.892.
3. F_1 įvertis – 0.895.

VI. REZULTATAI IR IŠVADOS

Šio tyrimo metu buvo atliktos užduotys:

1. Sukauptas ir anotuotas tekstynas.
2. Išanalizuoti Lietuvos Respublikos datos užrašymo standartai.
3. Išanalizuoti angliškos datos užrašymo standartai.
4. Pasiūlyti ir realizuoti datų atpažinimo metodai.
5. Atlikti eksperimentai su metodais.

Pagal gautus rezultatus ir anotuotą tekstyną pastebėtos tokios išvados, kurių laikantis būtų pasiekiami geresni rezultatai:

1. Pridėti likusius galimus datos užrašymo metodus: prieš 20 dienų, per 2 valandas, už 2 min. Šiame tekstyne tokių datos formatų nebuvo, dėl to jų atsisakyta. Tam r eikalingas d idesnis t ekstynas, kurio metu būtų atrasti ir ištirti visi g alimi d atos užrašymo būdai.
2. Skaitmenizuoti datos formatą ir patikrinti, ar ta data yra eg zistuojanti, p vz., 2013-02-31, t okios d atos niekada nebus.
3. Kur neužtenka paprastų taisyklių aprašymo, panaudoti mašininio mokymo metodus.
4. Atpažinti ir skaitmenizuoti datą, kai yra rašoma prieš 10 dienų.
5. Tiksliau a tpažinti d atos r ibas, p apildyti taisyklių rinkinį taisyklėmis.
6. Sudėti taisyklių prioritetus.
7. Atpažintį datą, kuri yra tiksli straipsnio publikavimo data, jei tokia yra.

LITERATŪROS SĄRAŠAS

[1] V. C. H. Y. Diana Maynard, „Named Entity Recognition from Diverse Text Types, Recent Advances in Natural Language,“ 2001.

[2] G. R. Jurgita Kapočiūtė, „Rule-based annotation of lithuanian text corpora,“ *INFORMATION TECHNOLOGY AND CONTROL*, pp. 290-296, 2005.

[3] „gate.ac.uk,“ [Tinkle]. Prieiga per internetą: http://gate.ac.uk/ [Žiūrėta 2013 03 01].

[4] „Valstybine Lietuviu kalbos komisija. Data, datos rašymas,“ [Tinkle]. Prieiga per internetą: http://www.vlkk.lt/lit/892 [Žiūrėta 2013 03 01].

[5] „Valstybine Lietuviu kalbos komisija. Rašyba: data ir laikas,“ [Tinkle]. Prieiga per internetą: http://www.vlkk.lt/lit/nutarimai/rasyba/data-laikas.html [Žiūrėta 2013 03 20].

[6] „Wikipedia. Calendar date.,“ [Tinkle]. Prieiga per internetą: http://en.wikipedia.org/wiki/Calendar_date [Žiūrėta 2013 03 14].

[7] „Wikipedia. Calendar dates ISO 8601,“ [Tinkle]. Prieiga per internetą: http://en.wikipedia.org/wiki/ISO_8601#Calendar_dates [Žiūrėta 2013 03 12].

[8] „Time and Date AS. EST – Eastern Standart Time,“ [Tinkle]. Prieiga per internetą: http://www.timeanddate.com/library/abbreviations/timezones/na/est.html [Žiūrėta 2013 03 11].

[9] V. J. a. M. d. R. Mahboob Alam Khalid, „The Impact of Named Entity Normalization on,“ ISLA, University of Amsterdam, Amsterdam, 2008.

[10] M. v. Z. D. S. Diego Molla, „Named Entity Recognition for Question Answering,“ Macquarie University, Sydney, 2007.

[11] A. S. Neeraj Agrawal, *Using Named Entity Recognition to improve Machine Translation.*

[12] H. S. Nadine Ludwig, „Named Entity Recognition for User-Generated Tags,“ Ludwig TIR, 2011.

[13] L. Thierry Poibeau, „Proper Name Extraction from Non-Journalistic Texts, Computational Linguistics in the Netherlands,“ 2001.

[14] J. Goyvaerts, „Regular-Expressions.info,“ [Tinkle]. Prieiga per internetą: http://www.regular-expressions.info/ [Žiūrėta 2013 02 25].

[15] S. Toarca, „A visual regular expression debugger, tester, and helper.,“ [Tinkle]. Prieiga per internetą: http://www.debuggex.com/?re=%5B0-9%5D%7B4%7D+%5B0-9%5D%7B1%2C2%7D+%5B0-9%5D%7B1%2C2%7D&str= [Žiūrėta 2013 02 27].

[16] „GATE JAPE manual,“ [Tinkle]. Prieiga per internetą: http://gate.ac.uk/sale/thakker-jape-tutorial/GATE%20JAPE%20manual.pdf

[17] „LingPipe: Named Entity Tutorial,“ [Tinkle]. Prieiga per internetą: http://alias-i.com/lingpipe/demos/tutorial/ne/read-me.html [Žiūrėta 2013 03 01].

[18] „Sensitivity and specificity,“ Wikipedia, [Tinkle]. Prieiga per internetą: http://en.wikipedia.org/wiki/Sensitivity_and_specificity [Žiūrėta 2013 03 04].

[19] „Precision and recall,“ Wikipedia, [Tinkle]. Prieiga per internetą: http://en.wikipedia.org/wiki/Precision_and_recall [Žiūrėta 2013 03 10].

[20] „Wikipedia F1 score,“ [Tinkle]. Prieiga per internetą: http://en.wikipedia.org/wiki/F1_score [Žiūrėta 2013 02 27].

[21] S. Toarca, „A visual regular expression debugger, tester, and helper.2,“ Toarca, [Tinkle]. Prieiga per internetą: http://www.debuggex.com/?re=%5B0-9%5D%7B4%7D+%3F%28m%5C.%3F%7Cmet%5B%C5%B3u%5D%29%3F+%3F%28sausio%7Cvasario%7Ckovo%7Cbaland%5Bz%C5%BE%5Dio%7Cgegu%5Bz%C5%BE%5D%5Be%C4%97%5Ds%7Cbir%5Bz%C5%BE%5Delio%7Cliepos%7Crugpj%5Bu%C5%AB%5D%5Bc%C4%8D%5Dio%7Crugs%5B%C4%9 [Žiūrėta 2013 02 27]

INFORMACINĖS TECHNOLOGIJOS • IVUS 2013 • ISSN 2029-249X • eISSN 2029-4824

Computational modelling and validation of a multilayer amperometric biosensor

Darius Baronas, Antanas Žilinskas
Institute of Mathematics and Informatics
Vilnius University
Vilnius, Lithuania

Feliksas Ivanauskas
Faculty of Mathematics and Informatics
Vilnius University
Vilnius, Lithuania

***Abstract*—This paper presents a mathematical model of an amperometric mediated biosensor based on an enzyme layer and two supporting porous membranes. The analytical system is modelled by non-stationary reaction-diffusion equations containing a non-linear term related to an enzymatic reaction. The mathematical model involves four layers: the enzyme layer where the enzymatic reaction as well as the mass transport by diffusion takes place and three diffusion limiting regions where only the mass transport by diffusion takes place. A computational model as well as the dimensionless model were developed for multiple simulation of the biosensor response. The numerical simulation was performed using the finite difference technique. The mathematical model and the numerical solution were validated by experimental data.**

I. Introduction

Biosensors are analytical devices combined mainly of biologically active substance, usually an enzyme, and a physiochemical transducer converting biochemical reaction result to a measurable quantity [1], [2], [3]. Biosensors are classified by their transducer extraction. Amperometric biosensors measure the current change on the working electrode due to the direct oxidation or reduction of chemical reaction products. The measured current is usually proportional to the concentration of the analyte (substrate). The amperometric biosensors are relatively cheap, sensitive and reliable devices for clinical diagnostics, drug detection, food analysis and environment monitoring [4], [5], [6], [7].

When designing biosensors the understanding their kinetic peculiarities is of crucial importance. To improve the efficiency of the development of a novel biosensor and to optimize its configuration the biosensor should be modelled [8], [9], [10]. Starting from the seventies various mathematical models have been successfully used as important tools to study and optimize analytical characteristics of actual biosensors [11], [12], [13]. A comprehensive review on the modelling of the amperometric biosensors has been presented by Schulmeister [14] and more recently by Baronas et al. [15].

Recently, an innovative approach in design of biosensors based on carbon nanotubes (CNT) layer deposited on the polycarbonate perforated membrane has been proposed [17] and computationally modelled [18]. Then the approach was enhanced by replacing the CNT-based electrode with a more simple and cheaper carbon paste electrode combining with an additional lavsan membrane. The aim of this work was to develop a mathematical model of an amperometric mediated biosensor based on an enzyme layer and two supporting porous membranes. The proposed model is based on nonlinear non-stationary reaction-diffusion equations. The model was described in a one-dimensional-in-space domain and comprises four layers (compartments): a layer of enzyme solution entrapped on an electrode, a polyvinyl alcohol (PVA) membrane, a lavsan membrane and an outer diffusion layer. The numerical simulation at transient conditions was carried out using the finite difference technique [14], [15], [16]. The mathematical model and the numerical solution were validated by experimental data. The obtained agreement between the simulation results and the experimental data was admissible at different concentrations of the substrate to be analyzed.

II. Mathematical model

Fig. 1 shows a schematic representation of a biosensor to be modelled. The considered biosensor consists of several layers of different materials and sizes. An electrode and a relatively thin layer of an enzyme (an enzyme membrane, region L_1) applied onto the electrode surface are the basic parts of the biosensor. The enzyme layer is covered by two permeable membranes: a polyvinyl alcohol (PVA) membrane (region L_2) and a lavsan membrane (L_3). In the enzyme layer, the enzymatic reaction as well as the mass transport by diffusion of the substrate, mediator and reaction product take place. Due to the immobilization no mass transport of the enzyme occurs. In two other membranes as well as in the diffusion limiting region only the mass transport by diffusion takes place. The model also involves a convective region (L_4) where the concentrations of all the species are maintained constant.

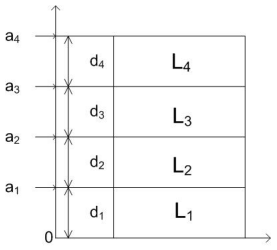


Fig. 1. The principal structure of the modelled biosensor.

In the enzyme layer we consider a two-stage enzyme-

catalyzed reaction

$$\mathrm{E_{ox}} + \mathrm{S} \xrightarrow{k_1} \mathrm{E_{red}} + \mathrm{P}, \tag{1a}$$

$$\mathrm{E_{red}} + \mathrm{M_{ox}} \xrightarrow{k_2} \mathrm{E_{ox}} + \mathrm{M_{red}}, \tag{1b}$$

where the substrate (S) combines with an enzyme (E) to form a product (P) in a presence of the mediator (M) [18]. There $\mathrm{E_{ox}}$ stands for the oxidised while $\mathrm{E_{red}}$ - for the reduced forms of the enzyme. Respectively, $\mathrm{M_{ox}}$ and $\mathrm{M_{red}}$ represent the oxidised and reduced forms of the mediator.

At the electrode surface the mediator is electrochemically re-oxidised ant electrons are released creating the current as an output result,

$$\mathrm{M_{red}} - \mathrm{e} \longrightarrow \mathrm{M_{ox}}. \tag{2}$$

As it is common for the amperometric biosensors, the electrochemical reaction (2) is assumed to be very fast [2], [3].

The substrate S to be analysed gets into the biosensor area from the buffer solution through the porous lavsan and PVA membrane. The external diffusion layer can be treated as the Nernst diffusion layer [16]. According to the Nernst approach the layer of the thickness d_4 remains unchanged with time. It is also assumed that away from it the solution is a uniform in the concentration.

A. Governing equations

Assuming the symmetrical geometry of the electrode and a homogeneous distribution of the immobilized enzyme in the enzyme layer of a uniform thickness, the mathematical model of the biosensor action can be defined in a one-dimensional-in-space domain [14], [15], [19]. Coupling the enzyme catalyzed reaction (1) in the enzyme layer with the mass transport by diffusion, described by Fick's law, leads to the following system of nonlinear reaction-diffusion equations:

$$\frac{\partial E_{ox}}{\partial t} = k_2 M_{ox,1} E_{red} - k_1 E_{ox} S_1, \tag{3a}$$

$$\frac{\partial E_{red}}{\partial t} = -k_2 M_{ox,1} E_{red} + k_1 E_{ox} S_1, \tag{3b}$$

$$\frac{\partial M_{ox,1}}{\partial t} = D_{M_{ox,1}} \Delta M_{ox,1} - k_2 M_{ox,1} E_{red,1}, \tag{3c}$$

$$\frac{\partial M_{red,1}}{\partial t} = D_{M_{red,1}} \Delta M_{red,1} + k_2 M_{ox,1} E_{red,1}, \tag{3d}$$

$$\frac{\partial S_1}{\partial t} = D_{S_1} \Delta S_1 - k_1 E_{ox,1} S_1, \quad x \in (0, a_1),\ t > 0, \tag{3e}$$

where x and t stand for space and time, S_1 is the concentration of the substrate (S) in the enzyme layer, E_{ox}, E_{red}, $M_{ox,1}$ and $M_{red,1}$ stand for the molar concentrations of the oxidised and reduced forms of the enzyme and mediator, respectively, D_{S_1}, $D_{M_{ox,1}}$, $D_{M_{red,1}}$ are the constant diffusion coefficients, and $a_1 = d_1$ is the thickness of the enzyme layer [12], [13], [14]. The coefficients k_1 and k_2 are the rates of the reactions (1a) and (1b), respectively. Since the product P has no influence to the biosensor response, the dynamics of the P concentration is not considered.

In the outer layer only the mass transport by diffusion of both species takes place ($t > 0$),

$$\frac{\partial M_{ox,i}}{\partial t} = D_{M_{ox,i}} \Delta M_{ox,i}, \tag{4a}$$

$$\frac{\partial M_{red,i}}{\partial t} = D_{M_{red,i}} \Delta M_{red,i}, \tag{4b}$$

$$\frac{\partial S_i}{\partial t} = D_{S_i} \Delta S_i, \quad x \in (a_{i-1}, a_i), \quad i = 2, 3, 4, \tag{4c}$$

where S_i is the substrate concentrations in the layer (a_{i-1}, a_i), $M_{ox,i}$ and $M_{red,i}$ are the concentrations of the oxidised and reduced forms of the mediator, respectively, D_{S_1}, $D_{M_{ox,1}}$, $D_{M_{red,1}}$ are the constant diffusion coefficients, $i = 2, 3, 4$.

B. Initial and boundary conditions

Let $x = 0$, $x = a_1$, $x = a_2$, $x = a_3$ and $x = a_4$ respectively represent the following boundaries: between the electrode and the enzyme layer, the enzyme layer and the PVA membrane, the PVA membrane and the lavsan membrane, the lavsan membrane and the external diffusion layer, the external diffusion layer and the bulk solution. The biosensor operation starts when the substrate and the mediator appear in the bulk solution ($t = 0$),

$$\begin{aligned} &M_{ox,i}(x,0) = 0, \quad M_{red,i}(x,0) = 0, \\ &S_i(x,0) = 0, \quad x \in [a_{i-1}, a_i], \quad i = 2, 3, \end{aligned} \tag{5}$$

$$\begin{aligned} &M_{ox,4}(x,0) = 0, \quad M_{red,4}(x,0) = 0, \\ &S_4(x,0) = 0, \quad x \in [a_3, a_4), \end{aligned} \tag{6}$$

$$\begin{aligned} &M_{ox,4}(a_4,0) = M_0, \quad M_{red,4}(a_4,0) = 0, \\ &S_4(a_4,0) = S_0, \end{aligned} \tag{7}$$

$$E_{ox,0}(x,0) = E_0, \quad E_{red,0}(x,0) = 0, \quad x \in (0, a_1), \tag{8}$$

where S_0 and M_0 are the concentrations of the substrate and the mediator in the buffer solution, and E_0 stands for the initial concentration of the enzyme.

Due to the electrode polarization the concentration of the reduced mediator $\mathrm{M_{red}}$ at the electrode surface ($x = 0$) is permanently reduced to zero ($t > 0$) [14], [15],

$$M_{red,1}(0,t) = 0. \tag{9}$$

The mediator re-oxidation reaction (2) is assumed to be so fast, that whole diffusive mediator $\mathrm{M_{red}}$ touching the electrode surface ($x = 0$) is immediately re-oxidised, i.e. $\mathrm{M_{red}}$ is converted to $\mathrm{M_{ox}}$ ($t > 0$),

$$D_{M_{red,1}} \left.\frac{\partial M_{red,1}}{\partial x}\right|_{x=0} = - D_{M_{ox,1}} \left.\frac{\partial M_{ox,1}}{\partial x}\right|_{x=0}. \tag{10}$$

The substrate concentration flux on the electrode surface equals zero because of the substrate electro-inactivity ($t > 0$),

$$\left.\frac{\partial S_1}{\partial x}\right|_{x=0} = 0. \tag{11}$$

The outer diffusion layer ($a_3 < x < a_4$) is treated as the Nernst diffusion layer [16]. According to the Nernst approach the layer of the thickness $d_4 = a_4 - a_3$ remains unchanged

with time, and away from it the solution is uniform in the concentration ($t > 0$),

$$\begin{aligned} &M_{ox,4}(a_4,t) = M_0, \quad M_{red,4}(a_4,t) = 0, \\ &S_4(a_4,t) = S_0. \end{aligned} \tag{12}$$

The fluxes of the substrate and the mediator (in oxidised and reduced forms) through the stagnant external layer are assumed to be equal to the corresponding fluxes entering the surface of the enzyme membrane. Because of this, on the boundary between adjacent regions having different diffusivities, we define the following matching conditions ($t > 0,\ i = 1, 2, 3$):

$$\begin{aligned} D_{S_i}\frac{\partial S_i}{\partial x}\bigg|_{x=a_i} &= D_{S_{i+1}}\frac{\partial S_{i+1}}{\partial x}\bigg|_{x=a_i}, \\ S_i(a_i,t) &= S_{i+1}(a_i,t), \end{aligned} \tag{13}$$

$$\begin{aligned} D_{M_{ox,i}}\frac{\partial M_{ox,i}}{\partial x}\bigg|_{x=a_i} &= D_{M_{ox,i+1}}\frac{\partial M_{ox,i+1}}{\partial x}\bigg|_{x=a_i}, \\ M_{ox,i}(a_i,t) &= M_{ox,i+1}(a_i,t), \end{aligned} \tag{14}$$

$$\begin{aligned} D_{M_{red,i}}\frac{\partial M_{red,i}}{\partial x}\bigg|_{x=a_i} &= D_{M_{red,i+1}}\frac{\partial M_{red,i+1}}{\partial x}\bigg|_{x=a_i}, \\ M_{red,i}(a_i,t) &= M_{red,i+1}(a_i,t). \end{aligned} \tag{15}$$

C. Biosensor response

The anodic or cathodic current is measured as a result of a physical experiment. The current is proportional to the gradient of the mediator concentration at the electrode surface, i.e. on the border $x = 0$. The biosensor current $I(t)$ at time t can be obtained explicitly from Faraday's and Fick's laws [14],

$$I(t) = n_e F A D_{M_{red,1}}\frac{\partial M_{red,1}}{\partial x}\bigg|_{x=0} \tag{16}$$

$$= -n_e F D_{M_{ox,1}}\frac{\partial M_{ox,1}}{\partial x}\bigg|_{x=0}, \tag{17}$$

where n_e is a number of electrons involved in a charge transfer, A is the area of the electrode surface and F is the Faraday constant, $F = 96486\,\mathrm{C/mol}$.

We assume that the system (3)–(15) reaches equilibrium when $t \to \infty$,

$$I_\infty = \lim_{t\to\infty} I(t). \tag{18}$$

where I_∞ is the the steady state biosensor current. The steady-state current is the main characteristic in commercial amperometric biosensors acting in the batch mode [1], [2], [3].

D. Dimensionless model

In order to extract the main governing parameters of the mathematical model, thus reducing a number of model parameters in general, a dimensionless model is often derived [8], [14], [20]. The replacement of the parameters is based on parameter mappings eliminating model dimensional parameters. Table I presents all the dimensionless parameters of the model.

TABLE I. MODEL PARAMETERS

Dimensional	Dimensionless
x, cm	$\hat{x} = x/d_1$
t, s	$\hat{t} = tD_{M_{ox,1}}/d_1^2$
E_{ox}, M	$\hat{E}_{ox} = E_{ox}/E_0$
E_{red}, M	$\hat{E}_{red} = E_{red}/E_0$
S_i, M	$\hat{S}_i = S_i/S_0$
$M_{ox,i}$, M	$\hat{M}_{ox,i} = M_{ox,i}/M_0$
$M_{red,i}$, M	$\hat{M}_{red,i} = M_{red,i}/M_0$
$D_{M_{ox,i}}$, cm^2/s	$\hat{D}_{M_{ox,i}} = D_{M_{ox,i}} \,/\, D_{M_{ox,1}}$
$D_{M_{red,i}}$, cm^2/s	$\hat{D}_{M_{red,i}} = D_{M_{red,i}} \,/\, D_{M_{ox,1}}$
D_{S_i}, cm^2/s	$\hat{D}_{S_i} = D_{S_i} \,/\, D_{M_{ox,1}}$
I, A	$\hat{I} = Id_1/(n_e F A D_{M_{red,1}} M_0)$

For the enzyme layer, the reaction-diffusion equations (3) can be rewritten in the following dimensionless form ($\hat{t} > 0$, $\hat{x} \in (0, 1)$):

$$\frac{\partial \hat{E}_{ox}}{\partial \hat{t}} = \alpha_2 \hat{M}_{ox,1}\hat{E}_{red} - \alpha_1 \hat{E}_{ox}\hat{S}_1, \tag{19a}$$

$$\frac{\partial \hat{E}_{red}}{\partial \hat{t}} = -\alpha_2 \hat{M}_{ox,1}\hat{E}_{red} + \alpha_1 \hat{E}_{ox}\hat{S}_1, \tag{19b}$$

$$\frac{\partial \hat{M}_{ox,1}}{\partial \hat{t}} = \hat{D}_{M_{ox,1}}\frac{\partial^2 \hat{M}_{ox,1}}{\partial \hat{x}^2} - \alpha_2 \hat{M}_{ox,1}\hat{E}_{red,1}, \tag{19c}$$

$$\frac{\partial \hat{M}_{red,1}}{\partial \hat{t}} = \hat{D}_{M_{red,1}}\frac{\partial^2 \hat{M}_{red,1}}{\partial \hat{x}^2} + \alpha_2 \hat{M}_{ox,1}\hat{E}_{red,1}, \tag{19d}$$

where all the unknowns and model parameters are dimensionless. α_1 and α_2 are the diffusion moduled, also known as Damköhler numbers [14],

$$\alpha_1 = \frac{k_1 E_0 d_1^2}{D_{M_{ox,1}}}, \tag{20a}$$

$$\alpha_2 = \frac{k_2 E_0 d_1^2}{D_{M_{ox,1}}}. \tag{20b}$$

The diffusion module α_j compares the rate of the enzyme reaction $k_j E_0$ with the diffusion rate $D_{M_{ox,1}}/d_1^2$, $j = 1, 2$. It is rather well known that an ordinary enzyme electrode acts under diffusion limitation when the diffusion modulus is much greater than unity [12], [14]. If the diffusion modulus is significantly less than unity then the enzyme kinetics predominates in the biosensor response.

The governing equations (4) are transformed into the following dimensionless form ($\hat{t} > 0$, $\hat{x} \in (a_{i-1}/d_1, a_i/d_1)$):

$$\frac{\partial \hat{M}_{ox,i}}{\partial \hat{t}} = \hat{D}_{M_{ox,i}}\frac{\partial^2 \hat{M}_{ox,i}}{\partial \hat{x}^2}, \tag{21a}$$

$$\frac{\partial \hat{M}_{red,i}}{\partial \hat{t}} = \hat{D}_{M_{red,i}}\frac{\partial^2 \hat{M}_{red,i}}{\partial \hat{x}^2}, \tag{21b}$$

$$\frac{\partial \hat{S}_i}{\partial \hat{t}} = \hat{D}_{S_i}\frac{\partial^2 \hat{S}_i}{\partial \hat{x}^2}, \quad i = 2, 3, 4. \tag{21c}$$

The initial and boundary conditions can be transformed to dimensionless form easily.

III. Numerical simulation

Some practical experiments were performed to produce biosensor data. From physical experiments the following parameters were given:

$$S_0^{(1)} = 1.99\,\mathrm{mM}, \quad S_0^{(2)} = 19.6\,\mathrm{mM}, \tag{22a}$$
$$S_0^{(3)} = 0.49\,\mathrm{mM}, \quad S_0^{(4)} = 0.99\,\mathrm{mM}, \tag{22b}$$
$$S_0^{(5)} = 4.98\,\mathrm{mM}, \quad S_0^{(6)} = 9.9\,\mathrm{mM}, \tag{22c}$$
$$M_0 = 50\,\mu\mathrm{M}. \tag{22d}$$

The initial boundary value problem (3)–(15) was solved numerically because of nonlinearity of governing equations (3) [16], [14], [15]. Solving the biosensor model, an explicit finite difference scheme was developed on a uniform discrete grid. The model was implemented in C programming language [21]. Computations were performed using supercomputer. Assuming biosensor geometry and after performing number of experiments. Some of the parameters were fixed and kept constant [17], [18]:

$$\begin{aligned}
&D_{M_{ox,1}} = D_{M_{ox,1}} = D_{S_1} = 3\,\mu\mathrm{m}^2/\mathrm{s},\\
&D_{M_{ox,2}} = D_{M_{ox,2}} = D_{S_2} = 4.2\,\mu\mathrm{m}^2/\mathrm{s},\\
&D_{M_{ox,3}} = D_{M_{ox,3}} = D_{S_3} = 3.75\,\mu\mathrm{m}^2/\mathrm{s},\\
&D_{M_{ox,4}} = D_{M_{ox,4}} = D_{S_4} = 6\,\mu\mathrm{m}^2/\mathrm{s},\\
&k_1 = 6.9\times10^8 s^{-1}, \quad k_2 = 4\times10^{10} s^{-1},\\
&d_1 = d_2 = 0.2\,\mathrm{mm}, \quad d_3 = 1\,\mathrm{mm}, \quad d_4 = 15\,\mathrm{mm},\\
&n_e = 2, \quad A = 0.03\,\mathrm{cm}^2, \quad E_0 = 12\,\mu\mathrm{M}.
\end{aligned} \tag{23}$$

The precise determination of values of some model parameters was impossible [17]. Because of this values of those parameters were determined by multiple simulation of the biosensor response by fitting the simulation results to experimental data.

IV. Results and discussion

Fig. 2 shows the evolution of the biosensor current I obtained experimentally (curves 2, 4, 6, 7, 9 and 11) and numerically (curves 1, 3, 5, 8,10 and 12) at values of the model parameters defined in (23), the mediator concentration $M_0 = 50\,\mu$M and six concentrations of the substrate S_0. The physical experiments were carried out more than 100 seconds, while the computer simulation was performed till a desired accuracy of the steady state current was achieved.

As one can see that the obtained agreement between the simulation results and the experimental data is admissible at relatively low concentrations of the substrate, i.e. the computational model matches the physical experiments. The simulated current tends to be bigger compared to physical results when the concentration S_0 is greater than $4.98\,$mM.

At S_0 = 1.99 mM (curves 7 and 8) the relative error decreases from about 0.3 at $t = 10\,$s to approximately 0 at $t = 42\,$s, while at the greater values of t the error slightly increases. At S_0 = 0.99 mM (curves 9 and 10) the relative error monotonous decreases from 1 at the initial stage of the biosensor operation to 0.4, while at S_0 = 0.49 mM (curves 11 and 12) the relative error slightly increases from 0.25 to 0.4 both at the final one ($t = 55\,$s). Rather great values of the relative error of the simulated responses indicate that the physical experiments are really more complex than as defined by the mathematical model. Despite this inadequacy with the experiments, the model seems to be suitable for investigating the kinetic peculiarities and optimizing the configuration of the multilayer amperometric biosensor.

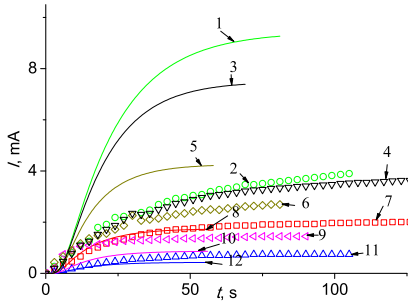


Fig. 2. The dynamics of the biosensor current I obtained experimentally (2, 4, 6, 7, 9 and 11) and numerically (1, 3, 5, 8,10 and 12) at $M_0 = 50\,\mu$M and six concentrations of the substrate S_0: 0.49 mM (11, 12), 0.99 mM (9, 10), 1.99 mM (7, 8), 4.98 mM (5, 6), 9.9 mM (3, 4), 19.6 mM (1,2)

V. Conclusions

Modelling of such biosensor is time consuming as we have to fit many parameters and to obtain an admissible agreement between the simulation results and the experimental data. To solve this task one must have many physical experiments to validate the simulation results. In this work only six validation curves were used, but this number of physical experiments was enough to conclude that model was not quite to match experimental data in a wide range of substrate concentrations. On the other hand, the model was quite accurate at relatively low concentrations of the substrate. It must be taken into consideration that experimental data can be also inaccurate because of external environment impact. The selected accuracy of the steady state current did not have perceptible effect to the simulation results.

Acknowledgment

This work has been supported by the project "Theoretical and engineering aspects of e-service technology development and application in high-performance computing platforms" (No. VP1-3.1-ŠMM-08-K-01-010) funded by the European Social Fund.

The authors are very grateful to Dr. J. Razumienė for the experimental data.

References

[1] H. Gutfreund, Kinetics for the Life Sciences, Cambridge: Cambridge University Press, 1995.

[2] A.P.F. Turner, I. Karube, and G.S. Wilson, Biosensors: Fundamentals and Applications, Oxford: Oxford University Press, 1987.

[3] F.W. Scheller, and F. Schubert, Biosensors, Amsterdam: Elsevier Science, 1992.

[4] U. Wollenberger, F. Lisdat, and F.W. Scheller, Frontiers in Biosensorics 2, Practical Applications, Basel: Birkhauser Verlag, 1997.

[5] B.E. Rapp, F.J. Gruhl, and K. Länge, "Biosensors with label-free detection for diagnostic applications," Anal. Bioanal. Chem., vol. 398, pp. 2403–2412, 2010.

[6] S. Viswanathan, H. Radecka, and J. Radecki, "Electrochemical biosensors for food analysis," Monatsh. Chem., vol. 140, pp. 891–899, 2009.

[7] A.K. Wanekaya, W. Chen, and A. Mulchandani, "Recent Biosensing Developments in Environmental Security," J. Environ. Monit., vol. 10, pp. 703-710, 2008.

[8] C. Amatore, A. Oleinick, I. Svir, N. da Mota, and L. Thouin, "Theoretical modeling and optimization of the detection performance: a new concept for electrochemical detection of proteins in microfluidic channels," Nonlinear Anal. Model. Contr., vol. 11, pp. 345–365, 2006.

[9] L. Liu, "Optimal inputs and sensitivities for parameter estimation in bioreactors," J. Math. Chem., vol. 47, pp. 1154–1176, 2010.

[10] M.E.G. Lyons, "Modelling the transport and kinetics of electroenzymes at the electrode/solution interface," Sensors, vol. 6, pp. 1765–1790, 2006.

[11] L.D. Mell, and T. Maloy, "A model for the amperometric enzyme electrode obtained through digital simulation and applied to the immobilized glucose oxidase system," Anal. Chem., vol. 47, pp. 299–307, 1975.

[12] J. Kulys, "The development of new analytical systems based on biocatalysts," Anal. Lett., vol. 14, pp. 377–397, 1981.

[13] P.N. Bartlett, and R.G. Whitaker, "Electrochemical immobilisation of enzymes. Part I. Theory," J. Electroanal. Chem., vol. 224, pp. 27–35, 1987.

[14] T. Schulmeister, "Mathematical modeling of the dynamic behaviour of amperometric enzyme electrodes," Sel. Electrode Rev., vol. 12, pp. 203–260, 1990.

[15] R. Baronas, F. Ivanauskas, and J. Kulys, Mathematical Modeling of Biosensors, Springer Series on Chemical Sensors and Biosensors, vol. 9, Dordrecht: Springer, 2010.

[16] D. Britz, Digital Simulation in Electrochemistry, Lecture Notes in Physics, vol. 666, Berlin: Springer, 2005.

[17] J. Razumienė, J. Gurevičienė, J. Barkauskas, V. Bukauskas, and A. Šetkus, "Novel combined template for amperometric biosensors with changeable selectivity," in Biodevices 2009: Proceedings of the International Conference on Biomedical Electronics and Devices, Porto, 2009, pp. 448–452.

[18] R. Baronas, J. Kulys, K. Petrauskas, and J. Razumiene, "Modelling carbon nanotube based biosensor," J. Math. Chem., vol. 49, pp. 995–1010, 2011.

[19] J.P. Kernevez, Enzyme Mathematics. Studies in Mathematics and its Applications, Amsterdam: Elsevier Science, 1980.

[20] M.E.G. Lyons, T. Bannon, G. Hinds, and S. Rebouillat, "Reaction/diffusion at conducting polymer ultramicroelectrodes," Analyst, vol. 123, pp. 1961–1966, 1998.

[21] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, Numerical Recipes: The Art of Scientific Computing, 3rd ed., New York: Cambridge University Press, 2007.

LRS narių grupavimas pagal balsavimą naudojant klasterizavimą

Kęstutis Bytautas
Baltijos pažangių technologijų institutas (BPTI)
Vilnius, Lietuva
Kestutis.Bytautas@bpti.lt

Tomas Krilavičius
Informatikos fakultetas
Vytauto Didžiojo universitetas ir BPTI
Kaunas ir Vilnius, Lietuva
t.krilavicius@bpti.lt

Santrauka. **Vienas iš pagrindinių būdų kontroliuoti politikus - jų stebėsena. Balsavimas viena iš esminių charakteristikų, reprezentuojanti politiko poziciją. Grupuojant juos pagal balsavimą galima pastebėti tam tikras tendencijas, ir klasterizavimas yra vienas iš metodų leidžiančių grupuoti objektus pagal panašumą, ir taip aptikti tendencijas, jų pokyčius. Šiame straipsnyje klasterizavimas taikomas politikų grupavimui pagal jų balsavimą Lietuvos Respublikos Seime (LRS). Siekiama atsakyti į klausimus, ar parlamentarų priklausomybė frakcijai ar pozicijai/opozicijai lemia jų balsavimą, ar galima aptikti politikų elgesio pokyčius? Analizuojama 2008-2012 metų Seimo kadencijos veikla. Tyrime lyginami skirtingi balsų kodavimo, panašumo skaičiavimo ir klasterizavimo metodai, taip pat naudojami įvairūs temporaliniai duomenų grupavimo metodai.**

Reikšminiai žodžiai: politinių duomenų analizė, duomenų gavyba, klasterizavimas

I. ĮVADAS

Politikai tiek priešrinkiminiuose debatuose, tiek būdami Seimo nariais įvairiai deklaruoja savo elgesį, todėl vienas iš pagrindinių būdų juos kontroliuoti – stebėjimas. Stengiamasi pažvelgti į Seimo darbo virtuvę, konkrečiai į balsavimus: ar tikrai Seimo nariai balsuoja ir elgiasi taip kaip deklaruoja ir skelbiasi patys, ar parlamentarų priklausomybė partijai (frakcijai) ir pozicijai (opozicijai) lemia jų balsavimą Seime, kokie dėsningumai iškyla, gal egzistuoja „slaptų“ grupių. Taip stengiamasi atsakyti į klausimą: ar informacinių technologijų įrankiai gali leisti nustatyti/atsakyti šiuos iškilusius klausimus?

LR Seimą [1] sudaro 141 narys. Rinkimuose į Seimą taikoma mišri rinkimų sistema: vienmandatė ir daugiamandatė. Įstatymai laikomi priimtais, jeigu už juos balsavo dauguma Seimo narių, dalyvaujančių posėdyje. Savo politiniams tikslams įgyvendinti Seimo nariai gali jungtis į frakcijas, kurios dažniausiai sudaromos išrinktų partijų pagrindu. Frakcijos gali jungtis į koalicijas, kurios gali veikti kaip viena frakcija.

Darbas turi kelis tikslus: 1. Seimo narių grupavimas – darbe siekiama pasiūlyti metodiką LR Seimo narių grupavimui pagal balsavimą. Norima pažvelgti kaip atrodys klasterizavimo pagalba sudarytos Seimo narių grupės, kaip jos kinta skirtingais laikotarpiais. 2. Seimo narių balsavimo pokyčių aptikimas – pasiūlyti metodiką balsavimo pokyčiams aptikti. Konkrečių Seimo narių balsavimų analizė: kokiai frakcijai priklausė balsuodamas ir kokiai grupei priklauso, atlikus klasterinę analizę, kaip kinta balsavimas įvairiais laikotarpiais.

Parlamento darbo analizė nėra naujas dalykas. Pasaulyje jau gana seniai atliekami įvairūs tyrimai [2]. Lietuvoje taip pat yra atlikta Seimo darbo tyrimų, pvz., [3] bandoma pažvelgti į praktinę partijų programų įgyvendinimo pusę, kaip jos įgyvendina savo programas, patekusios į Seimą. Pagrindinis rodiklis, šiuo atveju (parlamentarų balsavimas dėl įvairių teisės aktų projektų) - ar partinę ideologiją atspindinčius teisės aktus balsuodami paremia tai ideologijai atstovaujantys Seimo nariai. Seimo veikla analizuojama [4]: lyginami du duomenų analizės metodai – klasterizavimas ir daugiamačių skalių metodas. Pagrindiniai iškylantys klausimai - kaip pasirinkti tiriamus laikotarpius, duomenų kodavimą, panašumų matų palyginimą, hierarchinių klasterizavimo metodų palyginimas. Daugiamačių skalių metodas ir korespondentinė analizė plačiau išnagrinėti [5,6], tiriant 2008 m. LR Seimo narių balsavimus, kuriame atskleidžiami MDS trūkumai - jie priklauso nuo keleto subjektyviai parenkamų parametrų: duomenų kodavimo būdų, panašumo matų ir modeliavimo tipų

II. DUOMENYS

A. LRS 2008-2012

Duomenys tiriamajam darbui buvo paimti iš projekto *atviras-seimas.info* [7]. Analizuoti pavasario, rudens ir neeilinės sesijos balsavimai (žiūr. 1 lentelę) nuo 2011.03.10 iki 2012.01.19.

Pasirinktame laikotarpyje iš viso balsavimuose dalyvavo 8 frakcijos: Darbo partijos frakcija (DPF), Frakcija „Tvarka ir teisingumas“ (FTT), Liberalų sąjūdžio frakcija (LSF), Liberalų ir centro sąjungos frakcija (LCSF), Lietuvos socialdemokratų partijos frakcija (LSDPF), Mišri Seimo narių grupė (MG), Tėvynės sąjungos Lietuvos krikščionių demokratų frakcija (TSLKDF), Krikščionių partijos frakcija (KPF).

LENTELĖ 1. 2008-2012 LRS STATISTIKA

Pavadinimas	Laikotarpis	Balsavimų skaičius	Seimo narių skaičius
Visi balsavimai	2011.03.10-2012.01.19	3543	144
6 eilinė sesija	2011.03.10-2011.06.30	1724	143
7 eilinė sesija	2011.09.10-2011.12.23	1721	141
7 neeilinė sesija	2012.01.17-2012.01.19	102	140

Pasirinktame laikotarpyje balsavimuose dalyvavo 144 Seimo nariai. Frakcijas paliko ir kitas papildė 7 skirtingi Seimo nariai. Vienas Seimo narys sugebėjo pabuvoti trijose frakcijose.

LENTELĖ 2. Frakcijų pasiskirstymas

Pasiskirstymas	Frakcijos
Pozicija	TSLKDF, LCSF, LSF
Opozicija	DPF, TTF, LSDPF, MG, KPF

Balsavimus galima aprašyti matrica *V(K,I)*, kur *K* yra parlamentarų skaičius, *I* yra balsavimų skaičius, ir *V(k,i)* yra *k*-ojo parlamentaro balsavimas dėl *i*-ojo klausimo (t.y. *i*-asis balsavimas). Kitame skyriuje aprašyta, kaip balsavimus galima koduoti kiekybinėmis reikšmėmis.

B. Kodavimas

Balsavimas Seime nėra vien tik paprastas balsavimas už ar prieš. Galimas rezultatas, kai Seimo narys balsuoja, yra:

- už – balsavo už įstatymo projektą ar pasiūlymą;
- prieš – balsavo prieš įstatymo projektą ar pasiūlymą;
- susilaikymas – balsavimo metu susilaikė;
- nebalsavimas – užsiregistravo balsavime, bet nebalsavo;
- nedalyvavimas – plenariniame posėdyje nedalyvavo, kai vyko balsavimas;
- nėra informacijos - senesniuose šaltiniuose gali išvis nebūti pateikta informacija apie parlamentaro dalyvavimą/nedalyvavimą ir pan.

Paprasčiausias būdas yra užkoduoti duomenis taip: 1 – už, 0 – visi kiti [8]. Šiame tyrime buvo naudojamos šios balsavimo kodacijos (žr. 3 lentelę):

LENTELĖ 3. Balsavimo kodavimas

Koduotė	Aprašymas
100000	Už 1, visi kiti 0
101000	Už 1, prieš -1, kiti 0

C. Laikotarpiai

Duomenys buvo analizuojami ir atliekamas tyrimas šiais laikotarpių etapais (žr. 4 lentelę):

LENTELĖ 4. Pasirinkti laikotarpiai

Laikotarpis	Aprašymas
Visi balsavimai	Laikotarpis nuo 2011.03.10 iki 2012.01.19.
Sesijos	Laikotarpiai 2011.03.10-2011.16.10, 2011.09.10-2011.12.23, 2012.01.17-2012.01.19
Mėnesiai	2011 kovo, balandžio, gegužės, birželio, rugsėjo, spalio, lapkričio, gruodžio, 2012 sausio mėnesiai
Persidengiantys mėnesiai	2011 kovo-balandžio, balandžio-gegužės, gegužės-birželio, birželio-rugsėjo, rugsėjo-spalio, spalio-lapkričio, lapkričio-gruodžio, gruodžio-2012 sausio laikotarpiai.

III. KLASTERIZAVIMAS

Klasteris – tai panašių objektų grupė [9], [10]. Klasterinė analizė - tai matematinių metodų visuma, kurios pagalba galima objektų arba reikšmių aibes pagal jų panašumus suskirstyti į tam tikras grupes (klases, klasterius; šiuo atveju būtų frakcijos) taip, kad tos pačios grupės objektai būtų „artimi“ vienas kitam, kitaip tariant skirtumai būtų kuo mažesni, o objektai iš skirtingų grupių - „tolimi“; kitaip tariant skirtumai būtų kuo didesni. Klasterizavimo algoritmų pagrindinis privalumas – gebėjimas atpažinti grupavimo struktūrą be jokios išankstinės informacijos.

A. Panašumai

Objektų homogeniškumui nustatyti yra naudojami metriniai atstumo matai. Dažniausiai naudojami metodai:

- Minkowski [11]

$$d_{Minkowski}{}^{p}(x,y)=\left(\sum_{i=1}^{n}(x_i-y_i)^p\right)^{\frac{1}{p}} \quad (1)$$

- Manheteno (Minkowski, kur *p=1*) [12]

$$d_{Manheteno}(x,y)=\sum_{i=1}^{n}|x_i-y_i| \quad (2)$$

- Euklido (Minkowski, kur *p=2*) [13]

$$d_{Euklido}(x,y)=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2} \quad (3)$$

- Čebyševo [14]

$$d_{Čebyševo}(x,y)=\max_i|x_i-y_i| \quad (4)$$

B. Klasterizavimo metodai

Plokščiasis klasterizavimas. K-means yra vienas iš populiariausių klasterizavimo metodų [15], dar vadinamas kvadratinės paklaidos algoritmu (angl. *squared error clustering algorithm)*. Objektai skirstomi į *k* pradinių klasterių, apskaičiuojami klasterių centrai, apskaičiuojamas kiekvieno objekto atstumas iki klasterių centrų ir objektas skiriamas artimiausiam klasteriui; klasterių centrai perskaičiuojami, skaičiuojama kvadratinė paklaida, 2-5 etapai kartojami tol, kol gauname mažiausią kvadratinę paklaidą, kuri yra mažesnė už

pasirinktą reikšmę. Arba objektai nepersiskirsto po kitus klasterius.

Hierarchinis klasterizavimas. Tai yra taip pat vienas iš svarbių klasterizavimo metodų [16]. Kitaip negu K-Means, kur objektai paskirstomi į klasterius, čia sukuriama tam tikra objektų hierarchija. Šio algoritmo rezultatas yra klasterių hierarchija ir dendrograma. Pagal vykstančius procesus algoritmai skirstomi į jungimo (*angl. agglomerative*) arba skaidymo (*angl. divisive*). Vienas taip pat svarbesnių veiksnių hierarchiškai klasterizuojant yra objektų jungimas į klasterius. Populiariausi dabar egzistuojantys objektų jungimo metodai [17] yra:

- *Artimiausio kaimyno* jungimo metodas (AK) (*angl.furthest neighbor)* (5), žinomas kaip *Single-Link*.

$$p_{\min}(S_l, S_m) = \min_{\substack{X(i)\in S_l \\ X(j)\notin S_m}} d(X(i), X(j)) \quad (5)$$

- *Tolimiausio kaimyno* jungimo metodas (TK) (*angl. nearest neighbor)* (6), žinomas kaip *Complete-Link.*

$$p_{\max}(S_l, S_m) = \max_{\substack{X(i)\in S_l \\ X(j)\notin S_m}} d(X(i), X(j)) \quad (6)$$

- *Vidutinių atstumų* (VID) jungimo metodas (7), žinomas kaip *Average.*

$$p_{vid}(S_l, S_m) = \frac{1}{n_l n_m} \sum_{X(i)\in S_l} \sum_{X(j)\in S_m} d(X(i), X(j)) \quad (7)$$

- *Ward* jungimo metodas (8).

$$p_w(S_l, S_m) = \frac{d(\bar{S}_l, \bar{S}_m)}{1/n_i + 1/n_m} \quad (8)$$

C. Klasterizavimo kokybės įvertinimas

Rand indeksas (RI) [18] – lyginimas objektų panašumas, apskaičiuojamas formule (9) (žiūr. 5 lentelę):

$$RI = \frac{TP + TN}{TP + FP + FN + TN} \quad (9)$$

LENTELĖ 5. TP, FP, FN, TN išsidėstymas

	Priklauso klasei	Nepriklauso klasei
Priskirtas klasteriui	Teisingai priskirtas (true positive, TP)	Neteisingai priskirtas (false positive, FP)
Nepriskirtas klasteriui	Neteisingai nepriskirtas (false negative, FN)	Teisingai nepriskirtas (true negative, TN)

PURITY metodas [19] – paprasta ir nesudėtinga vertinimo priemonė. Apskaičiuojamas klasteryje daugiausia objektų turinčios klasės objektų skaičius, susumuojama ir padalinama iš objektų skaičiaus (10) :

$$PURITY(\Omega, \mathrm{B}) = \frac{1}{N} \sum_k \max_j (w_k \cap c_j) \quad (10)$$

kur $\Omega = (w_1, w_2, .., w_k)$ - klasterių aibė, $B = (c_1, c_2, .., c_j)$ - klasių aibė, N – objektų skaičius.

Neteisingai sugrupuoti - tiesiog skaičius, kuris rodo, kiek objektų neteisingai sugrupuota, objektas yra vienos klasės, o klasteris – kitos.

NMI [9]**,** [20] (*Normalized Mutual Information)* - normalizuota abipusė informacija. Skirtingoms aibėms X ir Y, NMI nustatomas (11) formule:

$$NMI(\Omega, \mathrm{X}) = \frac{I(\Omega; X)}{(H(\Omega) + H(X))/2} \quad (11)$$

kur - $\Omega = (w_1, w_2, .., w_k)$ klasterių aibė, $X = (c_1, c_2, .., c_j)$ - klasių aibė.

$$I(\Omega, X) = \sum_k \sum_j P(w_k \cap c_j) \log \frac{P(w_k \cap c_j)}{P(w_k) P(c_j)} = \sum_k \sum_j \frac{|w_k \cap c_j|}{N} \log \frac{N |w_k \cap c_j|}{|w_k||c_j|} \quad , \quad (12)$$

kur $P(w_k)$, $P(c_j)$, $P(w_k \cap c_j)$ yra tikimybės objektams atsidurti atitinkamai klasteryje w_1, klasėje c_j ir sankirtoje $w_k \cap c_j$. $H(\Omega)$ ir $H(X)$ yra Ω ir X entropijos.

$$H(\Omega) = -\sum_k P(w_k) \log P(w_k) = -\sum_k \frac{|w_k|}{N} \log \frac{|w_k|}{N} \quad (13)$$

Atitinkamai skaičiuojama ir $H(X)$. Kuo didesnis NMI, tuo galima tvirčiau teigti, jog klasterizavimas yra geresnis.

D. Įrankiai

Klasterinei analizei atlikti buvo naudojami šie įrankiai: ***Weka*** (*Waikato Enviromemt for Knowledge Analysis)* [21] yra populiarus mašininio mokymo programinės įrangos paketas, sukurtas JAVA kalbos pagrindu. Tai yra nemokama programinė įranga. Weka paketą sudaro algoritmų ir vizualizacijos įrankių rinkiniai, skirti duomenų analizei ir prognozavimo modeliavimui; ***RapidMiner*** [22] yra mašininio mokymo programinės įrangos paketas, skirtas duomenų išrinkimui, tekstų, prognozavimo analizėms.

IV. EKSPERIMENTŲ EIGA IR REZULTATAI

A. Duomenų paruošimas

Duomenys buvo paruošti specialiuose ARFF [23] tipo failuose, kuriuose duomenų struktūrą sudaro matrica, joje stulpeliuose nurodomi balsavimai, o eilutėse konkretūs Seimo nariai. Sukurti dviejų tipų duomenų failai, skirtingai užkoduojant balsavimo duomenis:

a) už – 1, kiti - 0. Kodavimas žymimas **10000**.

b) už – 1, prieš -1, kiti – 0.Kodavimas žymimas **10100**.

Suformuoti tyrimams reikalingi ARFF failai:

a) viso laikotarpio;

b) sesijos (pavasario, rudens, neeilinė);

c) mėnesiai (kovas, balandis, gegužė, birželis, rugsėjis, spalis, lapkritis, gruodis, sausis);

d) slenkantys mėnesiai (03-04, 04-05, 05-06, 06-09, 09-10, 10-11, 11-12, 12-01).

B. Seimo narių grupavimas į frakcijas

Įrankis - WEKA. Pasirinktas laikotarpis - visi balsavimai.

LENTELĖ 6. Pasirinkti klasterizavimo metodai

Klasterizavimo metodas	Kodavimas	Atstumų skaičiavimas	Klasterių skaičius
K-Means	10000	Euklido	8
K-Means	10000	Manheteno	8
Hierarchinis AK	10000	Euklido	8
Hierarchinis VID	10000	Euklido	8
Hierarchinis TK	10000	Euklido	8
K-Means	10100	Euklido	8

Klasterizavimo kokybės įvertinimas atliktas naudojant PURITY metodą.

LENTELĖ 7. Klasterizavimo įvertinimas

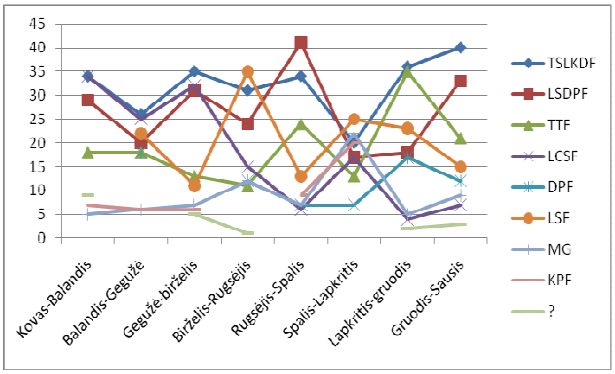
Klasterizavimo metodas	Purity (max)	Purity (class)	Neteisingai sugrupuoti
K-Means. Euklido atstumas. 10000	0.5763	0.4861	74
K-Means. Manheteno atstumas. 10000	0.5625	0.4305	77
Hierarchinis AK. 10000	0.368	0.375	92
Hierarchinis VID. 10000	0.604	0.55555	64
Hierarchinis TK. 10000	0.5833	0.5486	65
K-Means. Euklidas. 10100	0.5	0.4444	80
Hierararchinis TK. 10100	0.5902	0.5277	68

Pagal įvertinimo rezultatus (žr. 7 lentelė) matyti, jog rezultatai yra panašūs, išskyrus *artimiausio kaimyno* (AK) metodą. Vis dėlto geriausiai klasterizuoja hierarchiniai *tolimiausio kaimyno* (TK) ir *vidutinių atstumų* (VID) metodai.

LENTELĖ 8. Klasterizavimo rezultatai

Klasteris	TTF	TSLKDF	LSDPF	LSF	DPF	MG	KPF	LCSF	Priskirta
0	12				2	3	1		TTF
1	3		11		8	3	7		DPF
2	1	14		**12**	1	1		10	LSF
3		32		1				2	TSLKDF
4			11						LSDPF
5						4			MG
6	2		2						?
7								1	LCSF

Pasirinkus balsavimo periodus, galima stebėti, kaip per laiką kinta frakcijų sudėtis



Pav. 1 Frakcijų kaita persislenkančiais mėnesiais

Galima pastebėti (žr. 1 pav.), jog praktiškai nėra stabilių frakcijų (balsų skaičius nėra tolygus). Analizuojant frakcijų kaitą persislenkančiais mėnesiais (žr. 1 pav.), matome, jog galbūt TSLKDF kiek pastovesnė, išskyrus spalio-lapkričio mėn. Pradžioje pradingsta DPF, pabaigoje KPF. Matomas ryškus LSF svyravimas tai aukštyn, tai žemyn.

C. Seimo narių kitimo aptikimas

Priešingai negu kituose tyrimuose, čia didelis dėmesys kreipiamas į kiekvieną Seimo narį, lyginami kaip jie balsuoja būdami vienos frakcijos atstovais ir kokiai frakcijai yra priskirti po klasterizavimo. Lygindami kelis periodus, galime aptikti Seimo narių „klajones“ po Seimo frakcijas. Klasterizavimo kokybės įvertinimui naudojome NMI ir RAND (žr. 9 lentelė) (kuo didesnė indekso reikšmė, tuo klasterizavimas teisingesnis).

LENTELĖ 9. Klasterizavimo įvertinimas

RAND			NMI		
Periodas	Metodas	Rodiklis	Periodas	Metodas	Rodiklis
Slenkantys	**KM Eu**	**0.83**	**Visi**	**HR VID Eu**	**0.54**
Sesijos	HR TK Eu	0.82	**Visi**	**HR VID Ma**	**0.54**
Mėnesiai	HR TK Eu	0.82	Visi	HR CO Eu	0.53
Slenkantys	HR TK Eu	0.82	Sesijos	HR VID Ma	0.51
Slenkantys	KM Ma	0.82	Slenkantys	HR VID Ma	0.51
Slenkantys	HR TK Ma	0.82	Slenkantys	HR VID Eu	0.5

Viename pavyzdyje pateiktas balsavimo kitimas, aptiktas sesijose (žr. 10 lentelę). Naudotas K-Means metodas, Euklido atstumas, kodavimas 10000. Pavyzdyje matosi Seimo nario priklausomybė frakcijai, kaip jis balsavo ir kokiai frakcijai (nuspalvota) jis yra priskirtas klasterizavimo metu. Spalvų pagalba galima matyti kaip kinta Seimo narių priklausomybė frakcijose pagal balsavimą. Atlikus eksperimentus, buvo gauta, jog 44 nariai (8 – LSF, 1 – LSDPF, 35 – TSLKDF) per visą laikotarpį priskirti tai pačiai frakcijai, kurioje ir buvo balsuodami.

LENTELĖ 10. Seimo narių balsavimo kitimo aptikimas

Seimo narys	Pavasario sesija		Rudens sesija		Neeilinė sesija	
Abra.Vil	TSLKDF	TSLKDF	TSLKDF	TSLKDF	TSLKDF	TSLKDF
Ačas.Rem	TTF	LSF				
Adom.Man	TSLKDF	TSLKDF	TSLKDF	TSLKDF	TSLKDF	TSLKDF
Andr.Vyt	LSDPF	KPF	LSDPF	LSDPF	LSDPF	TTF
Anuš.Arv	TSLKDF	TSLKDF	TSLKDF	TSLKDF	TSLKDF	?
Aušt.Pet	LSF	LSF	LSF	LSF	LSF	LSF
Ažub.Aud	TSLKDF	LSF	TSLKDF	LSF	TSLKDF	LCSF
Babi.Vin	LCSF	LSF	LCSF	DPF	LCSF	LSF

Kitame pavyzdyje (žr. 11 lentelę) pateikiamas balsavimo kitimo aptikimas mėnesiais. Naudotas Hierarchinis tolimiausio kaimyno metodas, Euklido atstumų skaičiavimas, kodavimas – 10000. Paryškintos eilutės žymi stabilius Seimo narius, kurie priklauso tai pačiai frakcijai po klasterizavimo ir prieš klasterizavimą. Žalsvai pažymėtos pozicinės frakcijos, rausvai – opozicinės. Pagal rezultatus galima matyti, jog vėl stabiliausiai atrodo TSLKDF.

LENTELĖ 11. Seimo narių balsavimo kitimo aptikimas

Seimo narys	**Frakcija**	**03**	**04**	**05**	**06**
Abra.Vil	**TSLKDF**	**TSLKDF**	**LSF**	**TSLKDF**	**TSLKDF**
Bace.Vai	TSLKDF	TSLKDF	LSF	TSLKDF	TSLKDF
Balt.Vir	DPF	DPF	MG	DPF	DPF
Bara.Dai	TTF	TTF	TTF	LSDPF	TTF
Bast.Min	LSDPF	LSDPF	DPF	TTF	LSDPF
Bašk.Rim	MG	DPF	MG	MG	MG
Bauk.Ast	MG	?	TSLKDF	?	TSLKDF
Baur.Ant	MG	MG	MG	MG	MG
Beki.Dan	**TSLKDF**	**TSLKDF**	**TSLKDF**	**TSLKDF**	**TSLKDF**
Bilo.Agn	**TSLKDF**	**TSLKDF**	**TSLKDF**	**TSLKDF**	**TSLKDF**
Seimo narys	**09**	**10**	**11**	**12**	**01**
Abra.Vil	**TSLKDF**	**TSLKDF**	**TSLKDF**	**TSLKDF**	**TSLKDF**
Bace.Vai	TSLKDF	TSLKDF	TSLKDF	TSLKDF	?
Balt.Vir	KPF	LSDPF	LSDPF	DPF	LSDPF
Bara.Dai	LSF	TTF	TTF	TTF	LSDPF
Bast.Min	LCSF	LSDPF	TTF	LSDPF	TTF
Bašk.Rim	MG	MG	MG	MG	TTF
Bauk.Ast	TTF	MG	LSF	LSDPF	LSF
Baur.Ant	MG	MG	DPF	MG	TTF
Beki.Dan	**TSLKDF**	**TSLKDF**	**TSLKDF**	**TSLKDF**	**TSLKDF**
Bilo.Agn	**TSLKDF**	**?**	**TSLKDF**	**TSLKDF**	**TSLKDF**

V. REZULTATAI IR IŠVADOS

(1) Klasterizavimas yra tinkama metodika Seimo balsavimų analizei atlikti. Grupuojant narius matomi dėsningumai kaip Seimo nariai pasiskirsto po suklasterizuotas frakcijas. Ypač išsiskiria pagrindinės opozicijai ir pozicijai priklausančios frakcijos (TSLKDF ir LSDPF); **(2) Matome drausmę.** Pasirinktos priemonės leidžia įvertinti parlamentarų drausmę, frakcijų vieningumą. Skirtumai tarp Seimo nariams priklausančių frakcijos balsavimu metu ir frakcijų, priskirtų klasterizavimo metu, parodo, kiek Seimo nariai yra lojalūs savo frakcijai ir kiek frakcija yra nevieninga; **(3) Dinamikos aptikimas.** Grupuodami Seimo narius į grupes ir pasirinkdami įvairius laikotarpius galima pastebėti priklausomybės frakcijoms dinamiką. Pagal tai nesunku pamatyti, kurie nariai yra lojalūs frakcijoms, kurie linkę balsuoti kitaip ir galbūt ateityje atsidurti kitose frakcijose. (4) **Metodų tinkamumas**. Pagal išnagrinėtus metodus ir atliktus eksperimentus galima daryti išvadas: Hierarchinis *artimiausio kaimyno* klasterizavimo metodas visiškai yra netinkamas. Klasterizavimo rezultatai ir įvertinimo kriterijų rodikliai tai įrodo. Balsavimo kodavimas buvo atliktas dviem variantais. Pagal rezultatus labiau tinkamesnis kodavimas 10000. Atstumus skaičiuoti geriau Euklido atstumų skaičiavimo metodu. Vis dėlto apibendrinant visus rezultatus, geriausius klasterizavimo rezultatus parodė hierarchinis *tolimiausio kaimyno* ir *vidutinių atstumų* metodai. (5) **Metodų tobulinimas.** Išanalizavus gautus rezultatus, negalima pasakyti, kuris klasterizavimo metodas naudojant įvairius parametrus yra geriausias. Vis dėlto rezultatai priklauso ir nuo duomenų kodavimo, ir nuo atstumų skaičiavimo metodų, paties klasterizavimo metodo parinkimo, duomenų laikotarpių parinkimo, todėl siekiant išgauti geresnius ir teisingesnius rezultatus, reikia metodus tobulinti.

Planuojamos tyrimai. (1) Eksperimentai su duomenų sekų (*angl. data stream*) analizės metodais ir įrankiais, pvz. MOA [24] bei vizualaus balsavimo kitimo aptikimas; (2) balsavimų laikotarpių optimalių lango dydžių balsavimo kitimams aptikimas; (3) užmiršimo įtraukimas, pvz., senesniems balsavimams suteikiant mažesnį svorį; (3) socialinių tinklų naudojimas analize, pvz., kur viršūnės - parlamentarai, o lankai aprašo balsavimų panašumą [25]; (4) bendradarbiavimas su politologais tikrinant prasminius atitikimus, darbas su konkrečiais atrinktais įstatymais ir pan.

PADĖKOS

Dėkojame V. Morkevičiui už komentarus.

LITERATŪRA

[1] Lietuvos Respublikos Seimas. *LRS*. 2012 04 16. http://www.lrs.lt.

[2] Hix, S., Noury, A. and Roland, G. Dimensions of Politics in the European Parliament. *American Journal of Political Science*. 2006 m., 50(2): 494-520.

[3] Morkevičius, Vaidas. Neideologinis Seimas? Statistinė svarbių 2004-2008 m. kadencijos Lietuvos Seimo balsavimų analizė. Partinės demokratijos pabaiga? Politinis atstovavimas ir ideologijos. Vilnius : Versus Aureus, 53-87, 2009.

[4] Krilavičius, T. and Žilinskas A. *On Structual Analysis of Parlamentarian Voting Data. Informatica 19, 3 (August 2008), 377-390.* URL: http://www.mii.lt/informatica/pdf/INFO727.pdf.

[5] Tomas Krilavičius, Vaidas Morkevičius. *Mining Social Science Data: A Study of Voting of Members of The Seimas of Lithuania Using Multidimensional Scaling and Homegeneity Analysis*. 2009.

[6] V., Krilavičius T. and Morkevičius. *Lietuvos parlamentarų ideologinės pozicijos ir jų raida 2008-2012 m. kadencijos Seime: statistinė Seimo narių balsavimų analizė [Ideological positions of Lithuanian MPs and their dynamics in the term 2008-2012: Statistical Analysis of MPs Voting]*. Paper presented at the annual conference of the Lithuanian Political Science Association, 2010 11 20, Vilnius.

[7] Krilavičius, T., Cimmperman, P. and Žalandauskas, T. *Duomenų užteks visiems [Plentyof data for everyone]*. Vilnius : s.n., 2010.

[8] Hix, S. and Noury, A. Government-Opposition or Left-Right? The Institutional Determinants of Voting in Fourteen Parliaments. 2008 m.

[9] Christopher D.Manning, Prabhakar Raghavan, Hinrich Schutze. *Introduction to Information Retrieval*. 2009.

[10] Tan, P-N, Steinbach, M and Kumar, V. *Introduction to datamining*. 2006.

[11] Vikipedijos dalyviai. Minkowski distance. *Vikipedija, Laisvoji enciklopedija*. 2012 m. 03 12 d. http://en.wikipedia.org/wiki/Minkowski_distance.

[12] 16. Mark Young.,Tom Radcliffe Ph.D. Manhattan. *Improved Outcomes Software.* 2012 m. 01 08 d. http://www.improvedoutcomes.com/docs/WebSiteDocs/Clustering/Clustering_Parameters/Manhattan_Distance_Metric.htm.

[13] 17. Nathan Egge, Aaron Krowne. The PlanetMath Story. *Euclidean distance.* 2012 m. 14 23 d. http://planetmath.org/EuclideanDistance.html.

[14] 18. Teknomo, Kardi. Chebyshev distance. *Kardi Teknomo's Homepage.* 2012 m. vasaris 10 d. http://people.revoledu.com/kardi/tutorial/Similarity/ChebyshevDistance.html.

[15] Bernatavičienė, Jolita. *Vizualios žinių gavybos metodologija ir jos tyrimas.* Daktaro disertacija, 2008. http://www.mii.lt/files/mii_dis_08_bernataviciene.pdf.

[16] **Dipartimento di elettronica e informazione.** Hierarchical Clustering Algorithms. *Politecnico di Milano.* 2012 m. 05 13 d. http://home.dei.polimi.it/matteucc/Clustering/tutorial_html/hierarchical.html.

[17] Cambridge University Press. Single-link and complete-link clustering. *The Stanford Natural Language Processing Group* . 2012 m. 03 30 d. http://nlp.stanford.edu/IR-book/html/htmledition/single-link-and-complete-link-clustering-1.html.

[18] Vikipedijos dalyviai. Rand Index. *Vikipedija, Laisvoji enciklopedija.* 2012 m. 03 04 d. http://en.wikipedia.org/wiki/Rand_index.

[19] Delip Rao. A note of cluster purity. *Department of Computer Science and Engineering, Indian Institute of Technology Madras.* 2012 m. 03 24 d. http://www.cse.iitm.ac.in/~cs672/purity.pdf

[20] Yong Wang, Bao-Gang Hu. *Derivations of Normalized Mutual Information.* http://arxiv.org/ftp/arxiv/papers/0711/0711.3675.pdf.

[21] Vikipedijos dalyviai. Weka. *Vikipedija, Laisvoji enciklopedija. 2012 03 04.* http://en.wikipedia.org/wiki/Weka_%28machine_learning%29.

[22] Vikipedijos dalyviai. RapidMiner. *Vikipedija, Laisvoji enciklopedija.* 2012 m. 02 03 d. http://en.wikipedia.org/wiki/RapidMiner.

[23] Waikato, Department of Computer Science. The University of Waikato. Attribute-Relation File Format (ARFF). *Machine Learning Group at University of Waikato.* 2008 m. Lapkritis 01 d. http://www.cs.waikato.ac.nz/ml/weka/arff.html.

[24] Team, MOA**.** Stream Clustering. *MOA Massive Online Analysis.* 2012 m. balandis 12 d. http://moa.cs.waikato.ac.nz/details/stream-clustering/.

[25] Užupytė R., Morkevičius V., Krilavičius T., Lietuvos Respublikos Seimo narių balsavimų tyrimas pasitelkiant socialinių tinklų analizę: tinklo konstravimo parametrų įtaka. Studentų mokslinė praktika 2012, Konferencijos pranešimų santraukos, I dalis, Vilnius

Naujų riedėjimo guolių defektų diagnostiniai tyrimai

Vytautas Barzdaitis[1], Aurimas Degutis[2], Justina Sabulytė[3]
Informatikos fakultetas,
Vytauto Didžiojo universitetas
Kaunas, Lietuva
v.barzdaitis@if.vdu.lt [1], a.degutis@if.vdu.lt [2], j.sabulyte@if.vdu.lt [3]

Santrauka. **Straipsnyje pateikiamas metodas leidžia atlikti naujo riedėjimo guolio ir guolio elementų diagnostiką ir įvertinti surinkimo kokybę, analizuojant aukštų dažnių virpesius.**

Reikšminiai žodžiai: riedėjimo guolis, mikrodefektai, identifikavimas, aukšto dažnio virpesiai.

I. ĮVADAS

Pramonėje svarbu prognozuoti naujo riedėjimo guolio surinkimo kokybę bei jo atskirų mechaninių elementų kokybę. Naudojama diagnostinė įranga „Adash 4101", „Adash 4300", „Pruftechnik AG Viprotip", „OROSS" atlieka riedėjimo guolių, žadinamų žemo dažnio virpesių, analizę, kas svarbu vykdant naudojamų guolių diagnostiką [6].

Šiandien pramonėje būtina prognozuoti naujo riedėjimo guolio surinkimo bei jo atskirų mechaninių elementų kokybę tam, kad mechanikai išvengtų guolių netikėtų gedimų, pateiktų informaciją apie riedėjimo guolio atskirų mechaninių elementų defektus ar jų mikrodefektus. Šiems tyrimams svarbu atlikti naujo guolio diagnostiką nagrinėjant aukšto dažnio virpesius [7].

Dėl pramonėje naudojamos diagnostinės įrangos ribotų galimybių kuriama programinė įranga leis atlikti neardomą naujo riedėjimo guolio diagnostiką ir įvertinti jo būklę nagrinėjant aukšto dažnio virpesius.

Norint nustatyti naujo riedėjimo guolio elementų ir jų surinkimo kokybę, atliekami mechaninių virpesių, naudojant aukšto dažnio pjezoakcelerometrus, matavimai ir matavimo rezultatų analizė [6]. Mechaninių virpesių matavimo rezultatai diagnostikoje pateikiami virpesių duomenų vaizdavimo formatais, kurie vaizduojami virpesių pagreičių grafikais, virpesių pagreičių spektruose (spektrogramose), virpesių pagreičių gaubiančiaisiais spektrais, nusakomi virpesių pagreičių reikšmių kvadratiniais vidurkiais (a_{rms}, g_{rms}), virpesių pagreičių maksimaliomis vertėmis (a_{max}, g_{max}), ekscesо koeficientais ir kitais matavimais [2]. Mechaninių virpesių duomenų vaizdavimo formatuose matosi riedėjimo guolių elementų mikrodefektai, kurie leidžia įvertinti (1) bendrą riedėjimo guolio kokybę ir (2) guolio atskirų elementų kokybę. Guolio elementų kokybė nusakoma guolio vidinio ir išorinio žiedų riedėjimo takelių mikronuokrypiais, taip pat riedėjimo elementų (ritinukuose, rutuliukuose) ovalumu ir briaunuotumu [7].

II. TYRIMO OBJEKTAS

Statistiškai nustatyta [1], kad 90% - 95% riedėjimo guolių gedimų sudaro riedėjimo guolio išorinio ir vidinio riedėjimo takelių gedimai, metalo mikroištrupėjimai. Labai svarbu šiuos defektus nustatyti nagrinėjant naujų guolių žadinamus aukšto dažnio virpesius. Dabartiniai metodai [2], [3] taikomi guolių eksploatacijai nagrinėjant guolių žemųjų dažnių virpesius. Žemųjų dažnių virpesiai (iki 1kHz) susiję su makrodefektais riedėjimo guolio žiedų riedėjimo takeliuose ir riedėjimo elementuose.

Nagrinėdami naujo riedėjimo guolio mechaninių virpesių duomenų vaizdavimo formatus aukštųjų dažnių intervale (1kHz – 10kHz ir aukščiau), įvertiname riedėjimo guolio kiekvieno riedėjimo elemento atskirai žadinamus mikrodefektus. Nagrinėdami žemųjų dažnių mechaninius virpesius ir atsižvelgdami į ISO 492-94 ir ISO 199-97 standartų rekomendacijas, negalime įvertinti guolio riedėjimo takelių mikrodefektų ir prognozuoti guolių netikėtų gedimų.

Praktikoje naudojama diagnostinė įranga „Adash 4101", „Adash 4300", „Pruftechnik AG Viprotip", „OROSS", kuria gali būti atliekami tyrimai, yra standartinė ir taikoma pramonėje. Minėta įranga nepateikia pakankamai parametrų, tinkamų naujų riedėjimo guolių diagnostikai, pavyzdžiui, guolio riedėjimo rutuliukų tikslumo, guolio vidinio ir išorinio žiedų riedėjimo takelių paviršių šiurkštumo, ovalumo ir banguotumo. Visa tai leidžia identifikuoti naujo guolio mechaninių elementų mikrodefektus. Toks apribojimas neįvertina naujo riedėjimo guolio atskirų mechaninių elementų kokybės, todėl kuriamas naujas diagnostinis algoritmas ir programinė įranga.

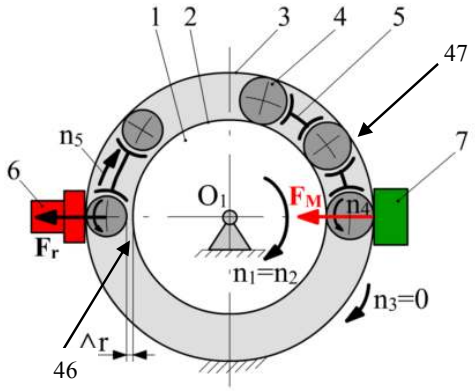
Pramonėje svarbu prognozuoti naujo riedėjimo guolio surinkimo kokybę bei jo atskirų mechaninių elementų kokybę. Analizuojant aukšto dažnio virpesių pagreičių spektrus, gautus nagrinėjant naujo riedėjimo guolio elementų mikronelygumų žadinamus virpesius, galime nustatyti:

- Naujo guolio išorinio žiedo rezonanso virpesių pagreičio amplitudę ir dažnį, jos priklausomybę nuo mikrodefektų.
- Naujo guolio išorinio ir vidinio žiedų mikrodefektus.
- Naujo guolio riedėjimo elementų nuokrypius.

Naudojama įranga nors ir pateikia minėtus parametrus, tačiau nepateikia informacijos apie mikronelygumus, esančius

kiekviename guolio riedėjimo elemente atskirai. Įvertinus turimos įrangos apribojimus, kuriame algoritmą ir programinę įrangą, skirtą naujo riedėjimo guolio atskirų riedėjimo elementų mikrodefektų diagnostikai.

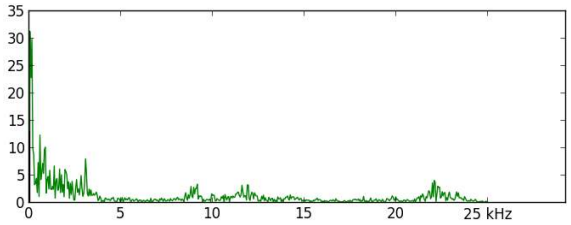
Šiame straipsnyje nagrinėjamas guolių elementų kokybės įvertinimas, metodai ir priemonės, leidžiantys įvertinti guolio riedėjimo elementų kokybę. Guolis bandomas virpesių tyrimo stende (3 pav. ir 5 pav.), kurio rotorius sukasi slydimo trinties guoliuose. Tai leidžia tiksliai įvertinti aukšto dažnio virpesius, žadinamus riedėjimo guolių (pavyzdžiui, 6205 guolio) elementų mikrodefektų [3]. Tyrimo stende tas pats guolis matuojamas esant dviem skirtingoms apkrovoms, kurias nustato (1) nuolatinis magnetas, tvirtinamas prie pjezoakcelerometro (3 pav.), ir (2) pjezoakcelerometras, tvirtinamas be magneto, o magnetas tvirtinamas prie guolio išorinio žiedo 180^O kampu nuo pjezoakcelerometro (5 pav.). Tiriamojo guolio vidinio žiedo sukimosi greitis yra pastovus 3000 aps./min. Šis greitis parinktas taip, kad guolio riedėjimo elementai išvystytų pakankamą išcentrinę jėgą, prispaudžiančią juos pastoviu kontaktu prie išorinio žiedo [4]. Tai leidžia įvertinti atskirai išorinio ir vidinio žiedų mikronelygumus.



1 pav. Rutulinio riedėjimo guolio virpesių matavimas, veikiant nuolatinio magneto M magnetinei jėgai F_M ir rutuliuko masės m_r išcentrinei jėgai F_r: 1-velenas, 2-guolio besisukančio vidinio žiedo riedėjimo takelis, 3-nesisukančio išorinio žiedo riedėjimo takelis, 4, 46, 47-rutuliukai, 5–separatorius, 6-pjezoakcelerometras, 7-nuolatinis magnetas, Δr-radialusis tarpas, n_1= n_2=3000 r/min veleno ir guolio vidinio žiedo sukimosi greitis

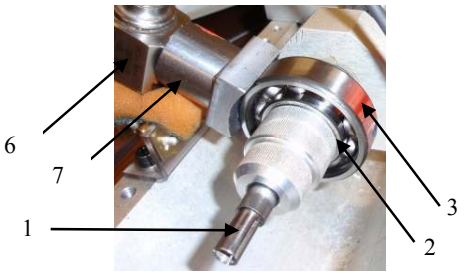
Pagal mechaninių virpesių pagreičių matavimus, realizuota programinė įranga pateikia 2 pav. ir 4 pav. grafikuose atvaizduotus virpesių pagreičių spektrus – virpesių pagreičių vaizdavimo duomenų formate. 2 pav. parodytas virpesių pagreičių spektras, kuriame vidinis ir išorinis žiedų takeliai liečiasi su riedėjimo elementu.

Naujo riedėjimo guolio išorinio ir vidinio žiedų virpesių pagreičius matuojame virpesių tyrimo stende 3 pav., kai riedėjimo guolio virpesių pagreičiai matuojami pjezoakcelerometru. Riedėjimo guolio išorinis žiedas yra veikiamas nuolatinio magneto apkrovos. Pjezoakcelerometras yra pritvirtinamas prie nuolatinio magneto, ir vidinis, ir išorinis žiedų takeliai liečiasi su riedėjimo elementu [4].

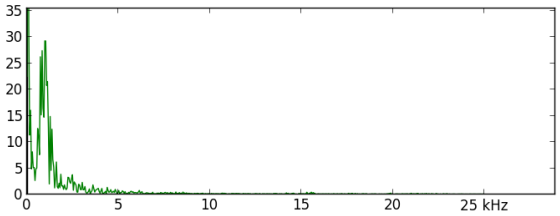


2 pav. Naujo guolio išorinio ir vidinio žiedų riedėjimo takelių žadinami aukšto dažnio virpesių pagreičių spektras. Matavimo stendas pateiktas 3 pav., riedėjimo guolis be tepalo. X ašimi – virpesių dažniai (kHz), Y ašimi – virpesių pagreičiai (μm/s^2).

Naujo riedėjimo guolio išorinio ir vidinio žiedų virpesių pagreičius matuojame virpesių tyrimo stende (5 pav., kai riedėjimo guolio virpesių pagreičiai matuojami pjezoakcelerometru. Antrasis eksperimentas atliekamas tyrimo metu matuojami naujo guolio tik išorinio žiedo ir riedėjimo elementų žadinami aukšto dažnio virpesių pagreičiai. Riedėjimo guolio išorinis žiedas yra veikiamas nuolatinio magneto apkrovos. Pjezoakcelerometras tvirtinamas be magneto, o magnetas tvirtinamas prie guolio išorinio žiedo 180^O kampu nuo pjezoakcelerometro nuolatinio magneto.



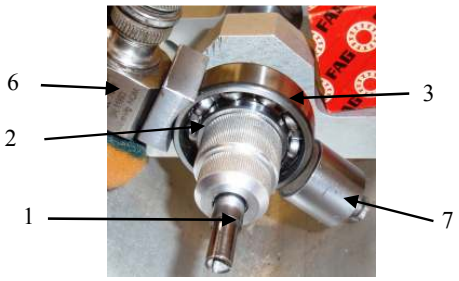
3 pav. Virpesių tyrimo stendas (PK+M+0, riedėjimo guolis be tepalo): matuojami guolio išorinio ir vidinio žiedų riedėjimo takelių žadinami virpesiai. Čia: 6-pjezoakcelerometras (PK), 7-nuolatinis magnetas (M), 1-velenas, 2- vidinis žiedas, 3- išorinis žiedas.



4 pav. Naujo guolio išorinio žiedo ir riedėjimo elemento žadinami aukšto dažnio virpesių pagreičių spektras. Matavimo stendas pavaizduotas 5 pav. guolis be tepalo. X ašimi – virpesių dažniai (kHz), Y ašimi – virpesių pagreičiai (μm/s^2).

Guolio 6205 vidinio žiedo sukimosi greitis yra parinktas 3000 aps./min. Tai garantuoja riedėjimo elementų pastovų kontaktą tik su išorinio žiedo riedėjimo takeliu ir jų nekontaktavimą su vidinio žiedo riedėjimo takeliu. Riedėjimo elementas nutolsta nuo vidinio žiedo riedėjimo takelio per radialųjį tarpelį Δr (iki 30μm) [4, 5]. Todėl 4 pav. grafike

pavaizduotas naujo guolio tik išorinio žiedo ir riedėjimo elementų žadinami aukšto dažnio virpesių pagreičiai.



5 pav. Virpesių tyrimo stendas (PK+0+M guolis be tepalo): matuojami guolio išorinio žiedo ir riedėjimo elementų žadinami virpesiai. Čia: 6-pjezoakcelerometras (PK), 7-nuolatinis magnetas (M), 1-velenas, 2- vidinis žiedas, 3- išorinis žiedas.

2 pav. ir 4 pav. atvaizduoti virpesių pagreičių spektrai neleidžia kiekybiškai įvertinti vidinio žiedo riedėjimo takelio ir riedėjimo elementų žadinamų virpesių [4]. Analizuojant virpesių pagreičių spektrus matome didelį virpesių intensyvumo skirtumą, kuris sąlygoja ir vidinio žiedo mikronelygumų žadinamus virpesius.

III. Tyrimo metodas

Straipsnyje pateiktu tyrimu aprašomas metodas, leidžiantis įvertinantis riedėjimo elementų mikronelygumų žadinamus aukšto dažnio virpesius. Sukurtas algoritmas realizuojamas programine įranga, o nauju metodu gaunamų virpesių analizės rezultatai palyginami su standartine, pramonėje naudojama, diagnostine įranga gautais rezultatais.

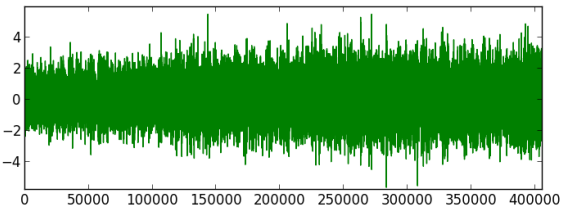
Žemiau pateikiami darbe naudojami terminai, jų matematinės išraiškos ir tyrimo metu gauti skaičiavimai atvaizduoti grafiškai.

3.1. Virpesių pagreitis

Virpesių pagreitį *x(t)* sudaro naujo riedėjimo guolio tyrimo metu gauti signalai visame virpesių matavimo intervale.

x(*t*) – virpesių pagreitis matavimų intervale.

Virpesių matavimo stendai, kuriuos naudojant atliekami matavimai, pavaizduoti 3 pav. ir 5 pav. nuotraukose. Visas rutulinio guolio virpesių pagreičio matavimų intervalas pateiktas 6 pav.



6 pav. Riedėjimo guolio virpesių pagreitis. X ašimi – matavimo taškų kiekis, Y ašimi – virpesių pagreičiai (μm/s^2).

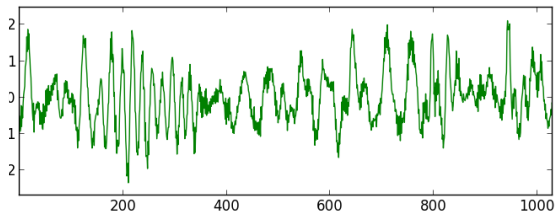
3.2. Virpesių pagreičio skenavimo intervalas

Virpesių pagreičio skenavimo intervalas nusako vieno rutulinio guolio apsisukimas, kurį sudaro 1024 matavimo taškai.

Virpesių pagreičio skenavimo intervalas – x(*n*, *l*).

x(n, *l*) - virpesių pagreitis *l*-tame skenavimo intervale, čia *n* – virpesių pagreičio skenavimo intervalo abscisė (*n* = 1, 2, .. , 1024); *l* – skenavimo intervalo eilės numeris (*l* = 1, 2, ..., L), L – virpesių pagreičio matavimų intervale esantis rutulinio guolio apsisukimų kiekis.

Šis virpesių pagreitis yra sudalinamas į atskirus apsisukimus, vadinamus virpesių pagreičio skenavimo intervalu, 7 pav.



7 pav. Riedėjimo guolio virpesių pagreitis skenavimo intervale (vienas rotoriaus veleno apsisukimas). X ašimi skenavimo intervalo taškai, Y ašimi – virpesių pagreičiai (μm/s^2).

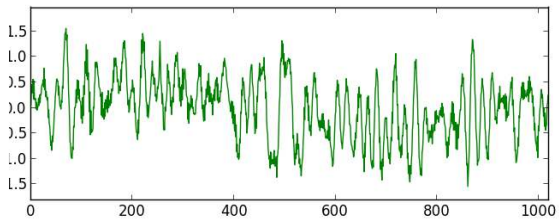
3.3. Virpesių pagreičio skenavimo intervalų vidurkis

Skenavimo intervalo vidutinė vertė nusako virpesių pagreičio svyravimus, kurie yra nusistovėję kiekviename apsisukime.

Skaičiavimai vykdomi sumuojant atitinkamus skenavimo intervalų taškus tarpusavyje ir dalinant iš jų kiekio.

$$y(n) = \frac{1}{L}\sum_{l=1}^{L} x(n,l) \qquad (1)$$

x(n, *l*) - virpesių pagreitis *l*-tame skenavimo intervale, čia *n* – virpesių pagreičio skenavimo intervalo abscisė (*n* = 1, 2, .. , 1024); *l* – skenavimo intervalo eilės numeris (*l* = 1, 2, ..., L), L – virpesių pagreičio matavimų intervale esantis rutulinio guolio apsisukimų kiekis.



8 pav. Virpesių pagreičio skenavimo intervalų vidurkis. X ašimi skenavimo intervalo vidurkio taškai, Y ašimi – virpesių pagreičiai (μm/s^2).

Tokiu būdu gaunamas normuotas taškų intervalas y(*n*), kurį atitinka skenavimo intervalo imtis. Virpesių pagreičio skenavimo intervale vidutinė vertė pavaizduota 8 pav.

Atlikus formulėje 1 pateiktus skaičiavimus, gaunama virpesių pagreičio skenavimo intervale vidutinė vertė, pavaizduota 8 pav. Gautas rezultatas atitinka skenavimo intervalo imties dydį, kurį sudaro 1024 taškai.

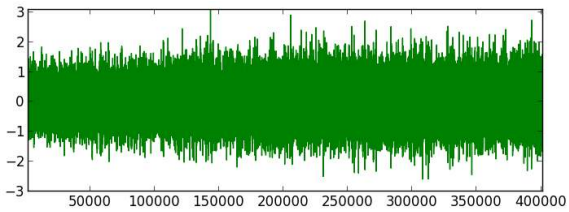
3.4. Centruotas virpesių pagreitis matavimo intervale

Norint turėti tikslesnes kiekvieno apsisukimo reikšmes, atliekamas virpesių pagreičio signalo centravimas. Šio etapo metu panaikiname nusistovėjusį svyravimą vykdant formulėje 2 pateiktus skaičiavimus. Skaičiavimai atliekami iš kiekvieno skenavimo intervalo (x(*n,l*) atimant gautą skenavimo intervalo vidutinę vertę y(*n*).

$$z(n,l) = y(n) - x(n,l) \qquad (2)$$

Čia y(*n)* - virpesių pagreičio skenavimo intervalų vidurkis, x(*n,l*) - virpesių pagreitis *l*-tame skenavimo intervale; *n* – virpesių pagreičio skenavimo intervalo abscisė (*n* = 1, 2, .. , 1024); *l* – skenavimo intervalo eilės numeris (*l* = 1, 2, ..., L), L – virpesių pagreičio matavimų intervale esantis rutulinio guolio apsisukimų kiekis.

Tai leistų pamatyti tikslesnes virpesių pagreičio reikšmes, nebepriklausančias nuo nusistovėjusių svyravimų, tuo pačiu leidžiant lengviau analizuoti kiekvieną apsisukimą. Centruotas rutulinio guolio virpesių pagreitis matavimo intervale pateiktas 9 paveiksle.

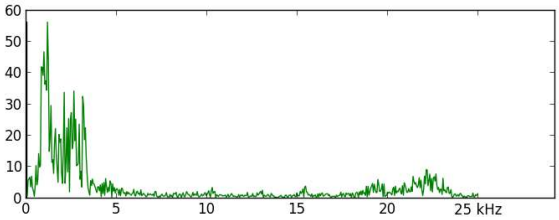


9 pav. Centruotas virpesių pagreitis matavimo intervale. X ašimi – centruoto virpesių pagreičio taškų kiekis, Y ašimi – virpesių pagreičiai (μm/s^2).

Vaizdingesnis analogas šio skaičiavimo būtų, kai norima pamatuoti medžių aukštį, kurie sustatyti vienoje eilėje, bet jie stovi ant banguotos kalvos, ir norint pamatyti aiškesnį skirtumą tarp medžių aukščių, reikia išlyginti kalvą. Centravimo operacija būtų analogiška kalvos išlyginimui. Valymas vykdomas, kai iš kiekvieno apsisukimo, arba dar vadinamų, virpesių pagreičio, skenavimo intervalų yra atimama skenavimo intervalo vidutinė vertė, tai yra visų apsisukimų vidurkis. Atimtis vykdoma, kai kiekviena skenavimo intervalo vidutinės vertės atitinkamos eilės reikšmė yra atimama iš atskiro skenavimo intervalo atitinkamos eilės reikšmės. Valymo proceso metu gautą rezultatą vadiname centruotu virpesių pagreičiu. Kadangi paprastai dirbant su dideliais duomenų kiekiais tarp centruoto ir necentruoto virpesių pagreičio skirtumo nematyti laiko erdvėje, iš jų yra paskaičiuojami dažnių spektrai, kurie parodo, kiek apvalymas turėjo įtakos virpesiams.

3.5. Centruotas virpesių pagreičio dažnių spektras

Naudojantis Furjė transformacija centruotam virpesių pagreičiui, gaunamas centruotas virpesių pagreičio dažnių spektras $F(z(n,l),f)$, kuris parodo, kokie esantys dažniai sudaro vidutinę vertę, 10 pav. Čia, z(*n, l*) - centruotas virpesių pagreitis matavimo intervale; *f* – virpesio pagreičio dažnis, *n* – virpesių pagreičio skenavimo intervalo abscisė (*n* = 1, 2, .. , 1024); *l* – skenavimo intervalo eilės numeris (*l* = 1, 2, ..., L), L – virpesių pagreičio matavimų intervale esantis rutulinio guolio apsisukimų kiekis.



10 pav. Centruotas virpesių pagreičio dažnių spektras. X ašimi – virpesių dažniai (kHz), Y ašimi – virpesių pagreičiai (μm/s^2).

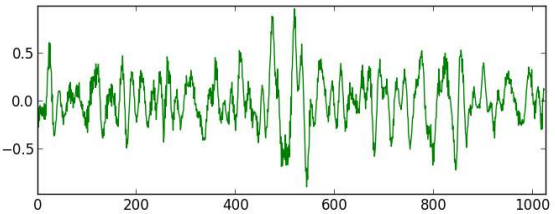
3.6. Centruotas virpesių pagreičio vidurkis

Kadangi norima gauti virpesį, kuris yra sužadinamas esant kontaktui tarp riedėjimo kūno ir mikronelygumo, ir žinant, jog jis yra labai mažos amplitudės ir tuo pačiu paskendęs triukšme, atliekamas centruoto virpesių pagreičio skenavimo intervalo vidutinės vertės skaičiavimas, formulė 3.

$$v(n) = \frac{1}{L}\sum_{l=1}^{L} z(n,l) \qquad (3)$$

čia z(*n, l*) - centruotas virpesių pagreitis matavimo intervale; *n* – virpesių pagreičio skenavimo intervalo abscisė (*n* = 1, 2, .. , 1024); *l* – skenavimo intervalo eilės numeris (*l* = 1, 2, ..., L), L – virpesių pagreičio matavimų intervale esantis rutulinio guolio apsisukimų kiekis.

Centruoto virpesių pagreičio skenavimo intervale vidurkio v(*n*) rezultatas yra vidutinė vertė, gaunama sumuojant atitinkamus centruoto virpesių pagreičio skenavimo intervalų taškus tarpusavyje ir padalinus iš centruoto skenavimo intervalų kiekio. Taip gaunamas normuotas taškų intervalas v(*n*), kurį atitinka centruoto skenavimo intervalo imtis, 11 pav.



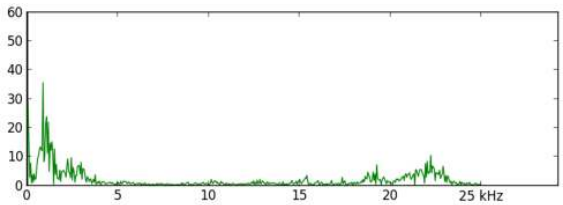
11 pav. Centruotas virpesių pagreičio vidurkis. X ašimi – centruoto virpesių pagreičio skenavimo intervale taškų kiekis, Y ašimi – virpesių pagreičiai (μm/s^2).

3.7. Centruotas virpesių pagreičio vidurkio spektras, vidurkintas laiko erdvėje

Gautas 3 formulės skaičiavimuose aprašytas centruoto virpesių pagreičio skenavimo intervale vidurkis v(*n*), naudojant greitą Furjė transformaciją [8], transformuojamas į dažnių spektrą $F(v(n,l),f)$. Šį signalą vadiname centruotu virpesių

pagreičio vidurkio spektru, vidurkintu laiko erdvėje. Jis atvaizduotas 12 pav.

Čia v(*n,*) - centruotas virpesių pagreičio vidurkis; *n* – virpesių pagreičio skenavimo intervalo abscisė (*n* = 1, 2, .. , 1024); *l* – skenavimo intervalo eilės numeris (*l* = 1, 2, ..., L), L – virpesių pagreičio matavimų intervale esantis rutulinio guolio apsisukimų kiekis.



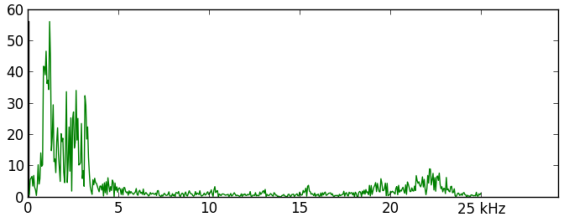
12 pav. Centruotas virpesių pagreičio vidurkio spektras. X ašimi – virpesių dažniai (kHz), Y ašimi – virpesių pagreičiai (µm/s^2).

3.8. Centruotas virpesių pagreičio vidurkio spektras vidurkinant dažnių erdvėje

Centruoto virpesių pagreičio vidurkio skenavimo intervale rezultatas *w(n)* yra vidutinė vertė, gaunama sumuojant atitinkamus centruoto virpesių pagreičio skenavimo intervalų spektro taškus tarpusavyje ir padalinus iš centruoto skenavimo intervalų kiekio. Gaunamas normuotas centruoto virpesių pagreičio vidurkio dažnių spektras *w(n),* matematinė išraiška pateikta 4 formulėje, grafiškai atvaizduotas 13 pav.

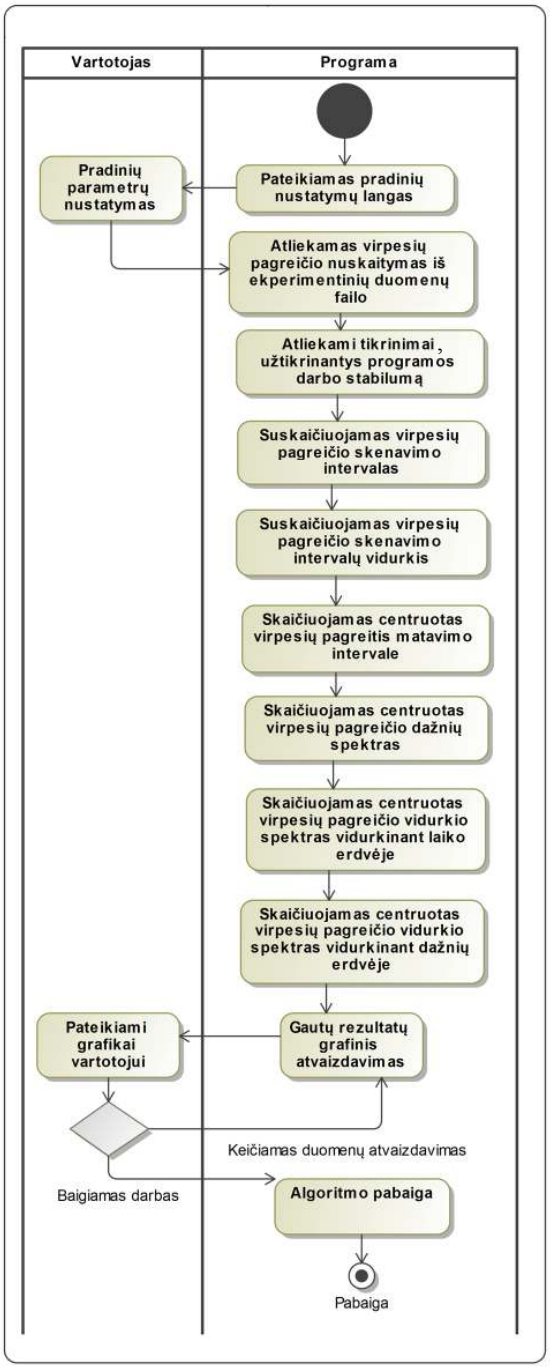
$$w(n) = \frac{1}{L}\sum_{l=1}^{L}(F(z(n,l)), f) \qquad (4)$$

Čia $F(z(n,l),f)$ - centruotas virpesių pagreičio z(*n, l*) dažnių spektras matavimo intervale, *n* – virpesių pagreičio skenavimo intervalo abscisė (*n* = 1, 2, .. , 1024); *l* – skenavimo intervalo eilės numeris (*l* = 1, 2, ..., L), L – virpesių pagreičio matavimų intervale esantis rutulinio guolio apsisukimų kiekis.



13 pav. Centruotas virpesių pagreičio vidurkio dažnių spektras. X ašimi – virpesių dažniai (kHz), Y ašimi – virpesių pagreičiai (µm/s^2).

Toks vidutinės vertės skaičiavimas reikalingas tam, kad sumažintų matavimo signaluose esantį triukšmo lygį. Tai atliekame sumuodami visus centruoto virpesių pagreičio skenavimo intervalų spektrus. Kadangi triukšmo reikšmės yra atsitiktinės, nepastovios ir nekoreliuotos, todėl atliekant skenavimo intervalo spektrų tarpusavio sumavimą didesnėje matavimų imtyje triukšmas sumažėja, t.y. sumažėja triukšmo amplitudės.



14 pav. Programinės įrangos veiklos modelis.

Triukšmui sumažėjus, išryškėja procesai, arba virpesiai, kurie yra nusistovėję laiko erdvėje. Vadinasi, jų amplitudė turėjo nors ir nedidelę vertę, taip pat jos reikšmės buvo pastovios tame pačiame dažnių ir/ar laiko intervale. Atliekant pavienių virpesių pagreičio skenavimo intervalų dažnių ir/ar

laiko erdvėje stebėjimą, pastovaus virpesių pagreičio reikšmės yra apgaubtos triukšmo, todėl yra atliekamas kiekvieno centruoto skenavimo intervalų vidurkinimas. Triukšmo sumažinimui ir/ar pasikartojančio proceso išryškinimui vykdomi 3 ir 4 formulėse aprašyti skaičiavimai.

IV. REALIZACIJA

Straipsnyje aprašytas tyrimo metodas, realizuotas programine įranga, kurios veiklos modelis pavaizduotas 14 pav.

IŠVADOS

Šiuolaikinė įranga ribotai leidžia įvertinti kiekvieno riedėjimo elemento mikronelygumus atskirai ir neidentifikuoja naujo guolio mikrodefektų.

Sukurtas naujas diagnostinis algoritmas ir programinė įranga leidžia įvertinti riedėjimo guolio vidinio ir išorinio žiedų riedėjimo takelių kokybę.

Tolimesniuose tyrimuose tobulinsime algoritmą ir programinę įrangą, kuri leistų atskirai įvertinti guolio išorinio ir vidinio riedėjimo takelių žiedų mikrodefektus.

LITERATŪRA

[1] R.Bosmans. Rebam – getting past the limitations of seismic transducers to a more thorough root cause analysis of a rolling element failure. Orbit, USA, 1999. No.1. p.7-10.

[2] P.Mažeika. Diagnostics and Faillures Prevention Researches of Rotors with Rolling element Bearing. Summary of Doctoral Dissertation. KTU, Technologija. Kaunas, 2008. p.p. 36.

[3] R.B.Randal, J.Antoni, Rolling element bearing diagnostics.A tutorial, Mechanical Systems and Signal Processing. Volume 25, Issue 2, February 2011. p.485-520. ISSN 0888-3270.

[4] J.E.Shigley, Ch.R.Mischke. Standard Handbook of Machine Design. McGraw-Hill Book Company, N.Y. 1986. ISBN 0-07-056892-8.

[5] S KF General Catalog 4000 US, Second Edition. Reg.47, 1997-01. Printed in USA, by Command Web Inc. pp. 800.

[6] N. Tandon, A. Choudhury A review of vibration and acustic measurement methods for the detection of defects in rolling element bearings. Tribology International 32 (1999) 469-480.

[7] A. Rafsanjani, S. Abbasion, A. Farshidianfar, H. Moeenfard, Nonlinear dynamic modeling of surface defects in rolling element bearings systems. Journal of Sound and Vibration 319 (2009) 1150-1174.

[8] Vibration analysis.
Prieiga internete: http://freevibrationanalysis.blogspot.com .
Paskutinės peržiūros data 2013 03 10.

INFORMACINĖS TECHNOLOGIJOS ● IVUS 2013 ● ISSN 2029-249X ● eISSN 2029-4824

Universaliojo vienetinio prisijungimo modelis

Laurynas Dovydaitis
Informatikos katedra, Kauno humanitarinis fakultetas
Vilniaus Universitetas
Kaunas, Lietuva

Santrumpa. **Straipsnyje analizuojamas universaliojo vienetinio prisijungimo (Single sign-on) (SSO) modelio prototipas, skirtas perduoti vartotojo tapatybę iš organizacijos valdomos centralizuotos vartotojų tarnybos išoriniam debesies paslaugų tiekėjui. Straipsnyje taip pat apžvelgiami SSO įgyvendinimo scenarijai. Aprašomi pateikto universaliojo SSO prisijungimo modelio prototipo patikrinimo eksperimento rezultatai. Pateikiamos prototipo vertinimo išvados ir galimos grėsmės. Pateikiami siūlymai, į ką reikėtų atsižvelgti vertinant galimas realizacijas.**

Reikšminiai žodžiai: universalusis SSO, Single Sign On, Universaliojo SSO prototipas.

I. ĮŽANGA

Šiuo metu informacinės technologijos išgyvena naują, debesų kompiuterijos, etapą. Debesies technologija yra elastingų, lengvai kontroliuojamų resursų infrastruktūra, kurią galima talpinti organizacijos vidinėse IT sistemose (privatus debesis) arba nuomotis kaip paslaugą iš išorinio tiekėjo (viešas debesis). Kai šias dvi technologijas apjungiame ir naudojame vienu metu - tai hibridinis debesis.

Organizacijos jau senai siekia apsaugoti vartotojų tapatybę. Nuolatos keičiantis ir didėjant technologijų naudojimui, šis uždavinys tampa dar sudėtingesnis. Organizacijose naudojami įprastiniai autentifikacijos, autorizacijos ir apskaitos sprendimai neturi naudojamose sistemose integruotų ir universalių (skirtingiems tiekėjams pritaikomų) įrankių, skirtų vartotojo tapatybę perduoti į debesies tiekėjų sistemas. Tokių produktų pavyzdžiai Microsoft Active Directory, IBM Tivoli ir kiti vidinei įmonės autentifikacijai skirti produktai.

Organizacijos intraneto ribose Microsoft Active directory, IBM Tivoli turi realizuotas galimybes vartotojui prisijungti centralizuotu būdu prie vidinių organizacijos resursų, tapatybę pateikiant vieną kartą ir pasinaudojant vienetiniu prisijungimu (angl. Single-sign-on – SSO). Tobulinant šių sistemų galimybes ir atliekant modifikacijas, jos gali būti pritaikytos perduoti autentifikavimo informaciją ir viešojo debesies paslaugų tiekėjams.

Temos aktualumą patvirtina 2012 metų pabaigoje rinkoje pasirodęs naujas komercinis produktas - IBM Security Access Manager. Šis produktas išplečia vidinės sistemas, įdiegiant papildomus programinės įrangos paketus ir susiejant juos su turimais autentifikacijos produktais. Naudojantis IBM Security Access Manager, papildomai kainuoja programinės įrangos licencijos, reikia skirti IT sistemų administratoriaus laiką produkto priežiūrai. Šie autentifikavimo sistemų papildymai priverčia diegti naujus produktus, taip patiriant papildomas sąnaudas tiek programinės įrangos licencijų įsigijimui, tiek sistemų administravimui.

Paminėtas IBM produktas neatitinka straipsnio apimtyje suformuotų apribojimų – panaudoti turimus autentifikacijos produktus, nepatiriant papildomų išlaidų jų įsigijimui ir priežiūrai.

II. SSO ANALIZĖ

Ankstesniuose darbuose [1] kalbant apie SSO buvo išskirti du pagrindiniai tipai: pirmasis – SSO naudojamas vidiniame įmonės tinkle, prisijungiant prie vidinių (privačių debesų) įmonės resursų, kai vartotojų tapatybės autentifikaciją atlieka vidinių resursų serveriai; antrasis – išoriniai (viešųjų debesų) resursai, dažniausiai sutinkami kaip Web SSO sistemos, kai tapatybės autentifikaciją atlieka organizacijai nepriklausantys serveriai, pasiekiami per viešąsias prieigas. Pirmuoju atveju SSO suteikia prielaidas organizacijoms administruoti vartotoją, naudojant vieną, centralizuotą, autentifikacijos mechanizmą. Antruoju atveju tapatybės valdymą užtikrina viešojo debesies tiekėjai.

Pagal populiarumą šiuo metu rinkoje su vieno prisijungimo principu suderinami produktai: Microsoft Active directory, OpenLDAP, IBM Tivoli, Novel eDirectory ir kt. [2]. Antrojo tipo sistemų pavyzdžiai: OpenID, Facebook Connect ir kt. [2].

Dažnai sutinkame sistemas, siūlančias hibridinį variantą - įmonės verslo procesų infrastruktūra, teikiama privačiuose duomenų centruose, o papildoma infrastruktūra saugoma išoriniuose šaltiniuose. Hibridinio debesies infrastruktūra tampa vis populiaresnė ir kaip rodo įvairių analitinių agentūrų tyrimų metu atlikti skaičiavimai, planuojamas debesų kompiuterijos augimas per ateinančius keletą metų viršys 50% [1].

Galime teigti, kad SSO yra paprasčiausias ir efektyviausias sprendimas tapatybės perdavimui. Organizacijos viduje naudojamas paprastasis SSO. Tuo tarpu, viešųjų debesų infrastruktūra autentifikacijai dažniausiai naudoja Web-SSO, kurį suteikia debesies paslaugų tiekėjai.

Apžvelgiant šiuo metu prieinamus hibridinių SSO prisijungimo scenarijus, užtikrinančius produktus PasswordBank Esso; Ping identity, kt., turime išskirti jų teikiamus privalumus: organizacija pati valdo tapatybės ciklus (angl. identity provisioning); iš dalies yra valdomi autentifikacijos mechanizmai; sąlyginai pritaikoma viena

autentifikacijos politika. Paminėtų produktų trūkumai: reikalingas pasitikėjimas šių sistemų kūrėjais; išauga kaštai dėl papildomai kainuojančios programinės įrangos; įmonėje autentifikacijos autonomija užtikrinama tik iki tarpininko sistemų.

Norint apjungti viešųjų debesų tiekėjų ir vidinės organizacijos naudojamas sistemas, galime analizuoti ir Universaliojo SSO koncepcija [1].

Įgyvendinus universalųjį SSO ir užtikrinus jo suderinamumą, atsiranda papildomas saugumo lygis valdant vartotojų elektroninę tapatybę. Šios sistemos pagalba yra sumažinamas vartotojo rekvizitų kiekis ir sukuriama prielaida vartotojui naudotis vienu, sudėtingu, sunkiai atspėjamu rekvizitu: slaptažodžiu ar kitu autentifikatoriumi [1].

Išskiriant SSO tipus, galime trumpai apžvelgti organizacijos SSO ir Web SSO naudojamus standartus.

1. Organizacijos SSO (privataus debesies) taikyme dažniausiai naudojamas Kerberos protokolas, kuris veikia „bilietų“ suteikimo principu [3].

2. Web SSO (viešojo debesies) taikyme naudojama bent keletas atvirai prieinamų tapatybės autentifikavimo sprendimų, tarp jų - OpenID, kuris yra plačiai paplitęs (pagal OpenID organizacijos teikiamą informaciją OpenID jau turi apie 1 milijardą vartotojų [4]).

Vienas iš universalių sprendimo variantų - Universalusis SSO (skirtas hibridinio debesies scenarijui) - siūlomas autentifikacijos duomenų perdavimo prototipas, kurio pagalba galime panaudoti organizacijos turimus, su SSO suderinamus produktus ir juose saugomą prisijungimo informaciją (vartotojo rekvizitus). Tokiu atveju rekvizitai būtų perduoti tiesiai iš vartotojo darbo vietos. Šie rekvizitai būtų saugomi centralizuotai turimoje tapatybės duomenų bazėje.

Naudojantis pateiktu prototipu, nereikia iš naujo kurti jau naudojamų saugumo politikų, siūlomas sprendimas gali būti paleistas su naudojamais produktais, ir autentifikacijos valdymas priklauso nuo organizacijos.

III. Universaliojo SSO panaudojimas ir prototipo kūrimas

Norint išnaudoti jau įdiegtas vartotojų autentifikacijos, autorizacijos ir apskaitos sistemas prisijungimui prie debesies paslaugų, tapatybės informaciją reikia saugoti vartotojų tapatybės tarnybų servisuose. Universalusis SSO turėtų veikti nepriklausomai nuo naudojamo produkto rūšies: ar tai būtų Active Directory, Tivoli ar kitas produktas. Siūlomas sprendimas – organizacijoje naudojama vartotojų autentifikavimo sistema išplečiama papildomais atributais. Pagal naudojamos sistemos reikalavimus aprašomas skriptas, kuris veikdamas vartotojo darbo vietoje gali pasiimti prisijungimo prie debesies tiekėjo duomenis iš centralizuotos duomenų bazės ir juos pateikti per interneto naršyklę.

Universaliojo SSO prototipo panaudojimas galimas realizuojant organizacijos SSO tapatybės perdavimo principus ir paleidžiant vartotojo darbo vietoje algoritmą, kuris galėtų „paimti“ centrinėje autentifikacijos bazėje turimą informaciją ir tiesiai iš vartotojo darbo vietos ją perduoti išoriniam debesų kompiuterijos tiekėjui.

Šio scenarijaus panaudojimas yra naudingas, nes nepadaugėja sistemų, kurias reikia administruoti, ir nepatiriamos papildomos sistemų administravimo sąnaudos.

Siūlomas prototipo modelis išsprendžia suderinamumo klausimus su įvairiomis autentifikacijos sistemomis ir skirtingais debesų kompiuterijos teikėjais. Siekiant tai įgyvendinti, prisijungimo informacija turi būti perduota iš vartotojo darbo vietos. Šiuo atveju vartotojo darbo vieta jau turi ryšį su centrine autentifikavimo sistema, o vartotojas įsijungdamas į savo paskyrą yra autentifikuojamas automatiškai. Daroma prielaida, kad vartotojo darbo vieta yra suderinama su debesų kompiuterijos tiekėjo sistemomis.

Norint apjungti šias sistemas į vieną sprendimą, vartotojo darbo vietoje galime paleisti skriptą, kuris vykdytų užklausą tapatybės direktorijos tarnybai. Užklausos pagalba, jungiantis prie debesies tiekėjo, būtų gauta informacija apie vartotojui suteiktas teises. Skriptas, savo ruožtu gavęs autorizacijos informaciją, galėtų ją pateikti viešojo debesies servisui, vartotojui neįvedant papildomų rekvizitų ir duomenų. Toliau įvyktų prisijungimas prie debesies paslaugos tiekėjo sukurtos ir organizacijos paskyros.

Pirmoje lentelėje pateikiamas bazinis informacijos, kurią turėtume saugoti centrinėje vartotojų duomenų bazėje, rinkinys. Šis rinkinys pagal organizacijos poreikį gali būti plečiamas ar kitaip modifikuojamas, įtraukiant papildomas informacijos reikšmes. Šiuo atveju pasirinktas bazinis informacijos rinkinys, kurio reikia siekiant patikrinti siūlomo prototipo veiksnumą.

LENTELĖ I. SAUGOMOS REIKŠMĖS

Pavadinimas	**Reikšmė**
Adresas	Paslaugų tiekėjo puslapio, prie kurio yra jungiamasi, www adresas.
Vardas	Vartotojo prisijungimo vardas, kuris turi būti panaudotas bandant pasiekti tiekėjo paslaugą.
Slaptažodis	Vartotojo prisijungimo slaptažodis, kuriuo galima pasiekti tiekėjo paslaugą.

Pasirinktos reikšmės pagal pirmosios lentelės informaciją yra tekstinio lauko tipo (angl. String). Tiriant konkretaus organizacijos SSO produkto specifiką šis parametras gali būti daugiareikšmis (angl. Multi-valued), t.y. aprašytas vienu duomenų bazės įrašu arba išskaidytas į tris atskirus atributus.

IV. Siūlomas sprendimas

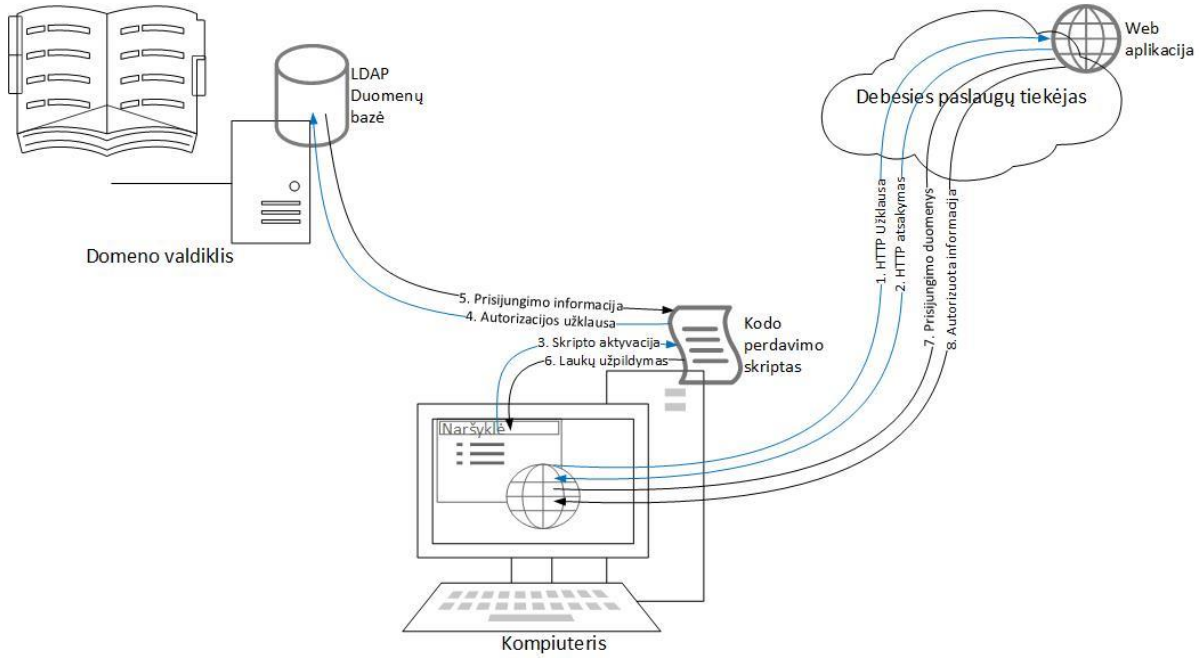
Kadangi siūlomas Universaliojo SSO sprendimas skirtas saugoti informaciją centralizuotoje tapatybės duomenų bazėje, pirmoje lentelėje aprašytomis reikšmėmis reikia papildyti vartotojų duomenų bazės schemą.

Atlikus schemos keitimus, galime aprašyti algoritmą, kuriame šios reikšmės būtų išgaunamos ir panaudojamos

autentifikacijai atlikti. Siūlomos autentifikacijos eiga pavaizduota pirmame paveikslėlyje.

Diagramoje atvaizduojami informacijos perdavimo mainai nuo 1 iki 4 žingsnio atliekami dar neautorizavus vartotojo į debesies tiekėjo paskyrą. Nuo 5 iki 8 žingsnio veiksmai vykdomi, kai vartotojas yra autorizuotas pasiekti paslaugos tiekėją. Diagramoje atskirai nėra detalizuojamas pagalbinės servisų informacijos perdavimas (tokie servisai kaip DNS užklausos ir kt.).

Schemoje pažymėtas kodo perdavimo skriptas – tai vykdomųjų komandinių eilučių seka, kuri gali naudoti HTTP GET/POST ir LDP užklausas.



Pav. 1. Siūloma autentifikacijos perdavimo schema Universaliajam SSO.

Siūlomas informacijos perdavimo veiksmų eiliškumas

1. Vartotojas surenka naršyklėje debesų tiekėjo serviso adresą. Pateikiama HTTP arba HTTPS užklausa. Užklausa vykdo kreipimąsi į paslaugos tiekėjo serverį ir prašo puslapio informacijos.

Reikia pažymėti, jog šiame etape taip pat yra vykdomos DNS užklausos ir kiti susiję procesai. Šios užklausos nėra detalizuojamos, kadangi tai tiesiogiai neįtakoja Universaliojo SSO prototipo realizacijos.

2. Tiekėjas pagal užklausą išsiunčia HTTP arba HTTPS atsakymą su puslapio turiniu. Po šio veiksmo vartotojo naudojamoje naršyklėje yra atvaizduojamas puslapio pradinis langas.

Šiame žingsnyje daroma prielaida, jog pradiniame puslapio lange tiekėjas suteikia galimybę įvesti vartotojo vardo ir slaptažodžio informaciją. Jei prisijungimo informacijos laukai yra pateikiami paskesniuose naršyklės puslapiuose, vartotojo kompiuteryje skriptas neaktyvuojamas, kol naršyklėje neatvaizduojami reikiami laukai.

3. Atsiradus prisijungimo laukų informacijai, kai puslapyje atvaizduojami vartotojo vardo ir slaptažodžio laukai, yra aktyvuojamas skriptas. Atsiradus šiems laukams, aktyvuojama prisijungimo duomenų užklausos procedūra. Užklausos tikslas yra pasitikrinti, ar vartotojas turi autorizaciją pasiekti debesų tiekėjo prieigą.

4. Duomenų užklausos procedūra vykdo kreipimąsi į centralizuotą vartotojų duomenų bazę. Iš centralizuotos vartotojų tarnybos duomenų bazės užklausa nuskaito aktyvaus vartotojo prisijungimo prie debesies paslaugų informaciją. Jei šis laukelis turi įrašą apie bandomą pasiekti debesų tiekėjo puslapį, vartotojui yra suteikta autorizacija.

Ši duomenų užklausa kompiuteryje yra vykdoma prisijungusio vartotojo kontekste. Visos užklausoje vykdomos komandos yra paleidžiamos prie kompiuterio prisijungusio vartotojo teisėmis. Jeigu vartotojas neturi teisės kreiptis į domeno valdiklį ir iš ten gauti informacijos apie jam priskirtus debesies atributus, papildoma informacija į kompiuterį negrąžinama.

5. Darant prielaidą, jog vartotojas autorizuotas, skirptui yra perduodama prisijungimo prie puslapio informacija, pagal pirmoje lentelėje suformuotas reikšmes: puslapio adresas, vartotojo vardas, bei vartotojo slaptažodis. Užklausa gali būti pateikta per LDAP protokolą arba kitomis priemonėmis, kurios yra prieinamos iš vartotojo darbo vietos ir gali pasiekti vartotojų duomenų bazę.

6. Gautą atsakymo informaciją skriptas įkelia į vartotojo vardo ir slaptažodžio laukus, esančius atidarytame naršyklės puslapyje. Kadangi prisijungimo skriptas veikia vartotojo darbo vietos aplinkoje, prisijungimo laukai yra užpildomi nepriklausomai nuo to, ar pasiekiamas puslapis yra šifruotas HTTPS protokolu, ar tiesiog pateikiamas atviruoju tekstu (HTTP protokolu).

7. Interneto naršyklės pagalba prisijungimo duomenys yra persiunčiami debesies paslaugų tiekėjo serveriui.

Tiekėjo serveryje, priklausomai nuo saugumo procedūrų, yra patikrinama, ar pateikta prisijungimo informacija yra teisinga.

Šios patikrinimo procedūros yra priklausomos nuo debesų tiekėjo algoritmų. Jei po atlikto patikrinimo pateikta informacija yra teisinga, įvykdoma vartotojo autorizacija.

8. Darant prielaidą, jog paslaugų tiekėjo puslapyje autorizacija įvyko sėkmingai, vartotojas pradeda gauti apsaugotą (reikalaujančią autorizacijos) informaciją iš paslaugų tiekėjo.

Pasiūlytas prototipas suteikia šias naudas:

- Atlikus visus žingsnius, vartotojas nesužino prisijungimo prie tiekėjo informacijos rekvizitų.
- Organizacijos administratorius gali centralizuotai suteikti prisijungimo teises prie vieno ar prie kito paslaugos tiekėjo, teises priskirdamas centralizuotos vartotojų tarnybos duomenų bazėje.

Kadangi vartotojų direktorijų tarnybos įgyvendinimai gali skirtis, pasiūlyto prototipo realizacija reikia adaptuoti konkrečiam centralizuotos vartotojų tarnybos produktui.

Siekiant išsaugoti organizacijos prisijungimo informaciją jau turimose vartotojų direktorijose, reikia įvertinti esamus laisvus atributus ir pasirinkti geriausią papildomos informacijos saugojimo strategiją.

V. Prototipo patikrinimas

Siekiant įsitikinti pasiūlyto prototipo veiksnumu, buvo atliktas dalinis pateikto prototipo eksperimentas, pasitelkiant Microsoft Active directory autentifikavimo tarnybą.

Kadangi pagal nutylėjimą atributai, esantys Active directory duomenų bazėje, gali būti rezervuoti kitoms organizacijų naudojamoms aplikacijoms, buvo nuspręsta atlikti naujų atributų sukūrimo procedūrą. Prototipo eksperimentinio patikrinimo metu buvo praplėsta duomenų bazės schema pasinaudojant pirmoje lentelėje pateiktu naujų atributų sąrašu. Šio proceso metu buvo modifikuojamas vartotojo (angl. User) klasės atributų rinkinys, juos papildant daugiareikšmiu tekstiniu įrašu (angl. Multi-valued string).

Atlikus eksperimentą ir įvykdžius pagal veiksmų eigą siūlomą autorizacijos užklausą (žingsnis nr. 4), informacija į skriptą buvo perduota sėkmingai.

Tęsiant papildomą analizę pastebėta, jog pagal nutylėjimą naujai sukurto atributų lauko informacija pateikiama visiems tarnybos vartotojams, nepriklausomai nuo jiems suteiktų teisių. Todėl pasinaudojant Active Directory schemos keitimo įrankiu, reikėjo atlikti sukurto atributo paieškos lauko korekciją, kuris nurodo parametro paieškos reikšmę.

Sukurtam atributui buvo priskirta searchFlags parametro reikšmė, suteikiant jai 128 žymą - parametro konfidencialumą. Šis pakeitimas sukurtą atributo informacijos lauką pažymi kaip konfidencialų ir apsaugo sukurtą lauką nuo informacijos pateikimo kitiems vartotojams.

Kadangi autorizavimo informacijos gavimo į vartotojo kompiuterį procedūra buvo įvykdyta sėkmingai, galima teigti, jog pasiūlytas prototipas gali pasiimti informaciją iš centralizuotos vartotojų duomenų bazės.

Visa siūlomo prototipo veiksmų seka šio eksperimento apimtyje nebuvo atlikta, eksperimento metu nevykdytos HTTP GET ir POST užklausos.

VI. Išvados

Šiame straipsnyje buvo pateiktas Universaliojo SSO informacijos pateikimo prototipas. Pateiktas prototipas įgyvendina pagrindinius keliamus reikalavimus – universalumą, paprastumą ir galimybę centralizuotai valdyti vartotojus. Informacija, pateikiama debesų kompiuterijos tiekėjams, yra perduodama iš kliento (vartotojo) darbo vietos.

Pasiūlytas prototipas galėtų būti pritaikomas skirtingoms sistemoms be papildomo programinės įrangos įsigijimo, atliekant modifikacijas jau turimiems organizacijos vartotojų valdymo produktams. Įgyvendinus šį prototipą, organizacijos gali panaudoti jau turimus produktus ir papildomai užtikrinti aukštesnius vartotojų rekvizitų apsaugos lygius.

Prieš pasiūlyto prototipo realizaciją reikia atlikti planuojamų pakeitimų įvertinimą. Eksperimentinio prototipo patikrinimo metu buvo rasti viešai prieinami sukurtų bandomųjų vartotojų rekvizitai. Rekomenduotina atlikti analizę, kurios metu turėtų būti patikrinti saugumo reikalavimai. Prisijungimo informacija, kuri yra saugoma centralizuotai, neatlikus papildomų parametrų keitimo, gali būti atiduota neautorizuotiems vartotojams. Neatlikus papildomos analizės, yra galimybė prarasti konfidencialią informaciją, todėl gali būti sukompromituota visa organizacijos viduje saugoma debesies paslaugų tiekėjo vartotojų bazė.

Literatūros sąrašas

[1] L. Dovydaitis (2012), "Vienetinio prisijungimo galimybės hibridinių debesų kompiuterijoje". Informacinės technologijos 2012, VU, ISSN 2029-249X p. 177-180.

[2] M. Vizard (2010), "Authentication in the Cloud" [Interaktyvus], prieiga per internetą http://www.itbusinessedge.com/blogs/it-unmasked/authentication-cloud.

[3] J.Y. Migeon "The MIT Kerberos Administrator's How-to Guide Protocol, Installation and Single Sign On" [Interaktyvus], prieiga per internetą http://www.kerberos.org/software/adminkerberos.pdf

[4] "OpenID Authentication 2.0" (2007), [Interaktyvus], prieiga per internetą http://openid.net/specs/openid-authentication-2_0.html

Mažų IT projektų apimties valdymo metodas

Eglė Dzidolikaitė
Vilniaus universitetas
Kauno humanitarinis fakultetas
Muitinės g. 8, LT-44280 Kaunas
El. paštas: e.dzidolikaite@gmail.com

***Santrauka*. Straipsnyje siekiama supažindinti su IT projektų valdymo samprata, identifikuojama jo reikšmė įmonėms. Taip pat išanalizuoti IT projektų apimties valdymo metodai. Pirmojoje straipsnio dalyje atskleidžiama IT projektų valdymo sąvoka, samprata ir reikšmė, analizuojamos įvairių autorių nuomonės apie tai. Toliau pateikiamas projektų valdymo metodikų palyginimas bei tyrimo, kuris parodė, nuo ko priklauso projekto sėkmė, rezultatai bei pasiūloma nauja metodika mažų projektų valdymui.**

Reikšminiai žodžiai: informacinės technologijos, projektų valdymas, IT projektų apimtis, projektų sėkmė.

I. ĮŽANGA

Šiame technologijų amžiuje, kai informacinės sistemos laikomos neatsiejama sėkmingo verslo dalimi, svarbu neatsilikti nuo siūlomų naujovių. Lietuvos statistikos departamento duomenimis [12], mokslo tiriamųjų darbų apimtys IT sektoriaus įmonėse 2010 m., palyginti su 2009 m., sudarė 24,6 procento visų mokslo tiriamųjų darbų išlaidų verslo sektoriuje, o 2011 m. pradžioje kompiuteriais ir internetu naudojosi 98,1 procento gamybos ir paslaugų įmonių, kuriose dirbo 10 ir daugiau darbuotojų. Tai reiškia, kad įmonės, suprantančios IT svarbą, diegia naujus IT produktus ir tokiu būdu didina konkurencinį pranašumą. Diegiant IT naujoves įmonėse susiduriama su projektų apimties valdymo problema. Pagal Chaos Report [5], projektų apimties neapibrėžimas ar apimties reikalavimų pakeitimas – viena iš dažniausių IT projektų žlugimo priežasčių. Nepaisant to, kad informacinės technologijos vystosi labai greitai, darbuotojams, atsakingiems už projektų valdymą, trūksta kompetencijos ir žinių, kad efektyviai išnaudotų turimus resursus, tikslingai paskirstytų darbus.

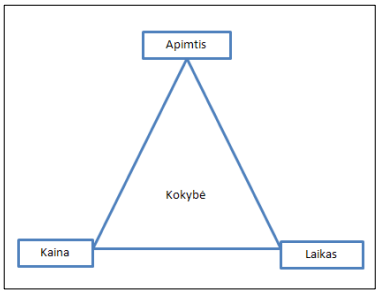
Yra sukurtų IT projektų valdymo metodikų - PMBOK, PRINCE2, RUP ir kt., tačiau paprastos ir suprantamos, skirtos mažiems projektams, daug nėra. Tradiciniai projektų valdymo metodai geriausiai pritaikomi didelės apimties projektams, todėl maži IT projektai susiduria su problemomis valdant apimtį.

Tikslas - patikrinti apimties valdymo metodikų efektyvumą mažų įmonių IT projektų valdyme. Darbo objektas - IT projektų apimčių valdymas. Antrame ir trečiame skyriuje aprašomi ištirti IT projektų apimties valdymo metodai, sąvokos, samprata ir reikšmė bei nustatyti IT projektų valdymo įgyvendinimo specifikos mažų projektų apimčių valdyme. Toliau pateikti atlikto eksperimentinio tyrimo rezultatai, kurie atskleidė, kokie faktoriai įtakoja sėkmingo projekto įvykdymą. Galiausiai pasiūlomas mažų projektų apimties valdymo metodas, atsižvelgiant į eksperimento rezultatus.

Analizuojant IT projektų apimties valdymo metodus panaudota palyginamoji analizė, kuri leido atskleisti metodikų privalumus ir trūkumus. Analizės rezultatai apibendrinti ir pateiktos išvados. Sukurtas apimties nustatymo metodas ir patikrintas apklausos būdu. Apklausos rezultatai panaudoti eksperimentiniam tyrimui, kurio metu atlikta koreliacinė analizė. Rezultatai apibendrinti ir pateikti paskutinėje darbo dalyje.

II. IT PROJEKTŲ VALDYMO SAMPRATA IR REIKŠMĖ

Projektų valdymas kaip atskira disciplina buvo oficialiai pripažinta 1950 m., atskiriant ją nuo vadybos disciplinos. Dabar projektų valdymas laikomas kaip nauja aukšto prioriteto planuojamų pakeitimų įgyvendinimo priemonė visose kompanijose ir organizacijose, neatsižvelgiant į jų dydį. Projektų valdymas pagal Atkinson [1] yra priemonių ir metodų taikymas, kurie nukreipti į sudėtingą unikalų tikslą, išpildant laiko, apimties ir kaštų apribojimus. Atkinson įvardino projekto sėkmės kriterijus (1 pav.), kurie vadinami „Geležiniu trikampiu“.



Šaltinis [1]:ATKINSON, R. (1999) Project Management: cost, time and quality, two best guesses and a phenomenon, its time to accept other success criteria, p 338

1 pav. Geležinis trikampis

Geležinis trikampis, vaizduojamas 1 paveiksle, nurodo, kad norint pasiekti kokybiškumo, turi būti palaikomas kainos, apimties ir laiko santykis. Geležinis trikampis vis dar naudojamas apibūdinti projektų valdymą šiomis dienomis.

Britų standartų instituto išleistame Projektų vadybos vadove [4] projektu vadinamas unikalus koordinuojamų veiklų, pasižyminčių apibrėžta pradžia ir pabaiga, rinkinys. Veikla vykdoma individo ar organizacijos, siekiant tam tikrų tikslų ir vadovaujantis apibrėžtais laiko, kaštų ir rezultatų parametrais.

Kerzner H. [10] projektu vadina veiklą, kuri turi specifinius tikslus, iš anksto nustatytą pradžią ir pabaigą, finansavimo limitą. Projektui atlikti naudojami žmogiškieji ir nežmogiškieji ištekliai, o pats projektas yra daugiafunkcinis produktas. Projektas pagal ISO standartą (ISO 10006:2003) [9] yra unikalus procesas, apimantis koordinuojamų ir valdomų veiklų visumą, turintis pradžios ir pabaigos datas ir skirtas tikslui, atitinkančiam specifinius reikalavimus, pasiekti su laiko, kainos ir išteklių apribojimais.

Violeta Bozhikova [3] teigia, kad pagrindinis projektų valdymo tikslas yra "gauti darbą”, pasiekti išsikeltus tikslus

laiku, vykdyti projektą biudžeto ribose, siekti efektyvumo. Šie trys kintamieji yra kritiniai projekto apribojimai, kurie reikalauja nuolatinio projekto vadovo dėmesio.

Turner, R. [14] išleistame Gower projektų vadybos vadove (angl. *„The Gower Handbook of Project Management“)* projektu vadinamas veiklų ciklas, kuriuo per numatytą laikotarpį siekiama sukurti tam tikrą kokybę ir kaštus atitinkantį produktą, paslaugą ar informaciją. Projektų valdymo instituto (angl. „*Projekt management institute – PMI“*) išleistame Projektų vadybos standarte (angl. *„Project Management Body of Knowledge, PMBOK®“*) [11] projektas apibrėžiamas lakoniškai – laikinos pastangos sukurti unikalų produktą ar paslaugą.

PRINCE2 projektų valdymo vadove [2] rašoma, kad projektas yra valdymo aplinka, kuri yra sukurta tiekti vieną ar daugiau verslo produktus, atsižvelgiant į konkretų verslo atvejį. PRINCE2 projektai turi šiuos skiriamuosius bruožus:

1. Baigtinė ir apibrėžta gyvenimo trukmė.
2. Apibrėžti ir įvertinami galutiniai produktai.
3. Atitinkamos veiklos pasiekti galutiniams produktus.
4. Apibrėžti resursai.
5. Organizacinė struktūra, turinti aiškiai apibrėžtą atsakomybę valdyti projektą.

Visame pasaulyje įmonėms pradėjus masiškai diegti projektus, projektų valdymo metodai tapo esminiai siekiant ne tik sėkmingo projektų įgyvendinimo bei jų vertės didinimo, bet ir visų kompanijų verslo sėkmės. Kerzner [10] projektų valdymą apibūdina kaip santykinai trumpalaikių tikslų, kurie buvo suformuoti siekiant įgyvendinti specifinius tikslus, planavimą, organizavimą, režisavimą, ir kontrolę.

Nors visi šie projekto apibrėžimai šiek tiek skiriasi, esmė lieka ta pati: projektas apibrėžiamas kaip "unikali" ir laikina veikla, vykdoma siekiant tam tikro apibrėžto tikslo. Esminė projekto savybė, retai randama apibrėžimuose, yra projekto gyvavimo ciklas. Visi projektai, kad ir kokie sudėtingi ar paprasti būtų, nuo eilinės "ne projektinės" veiklos skiriasi tuo, kad pereina kelis vienodus gyvavimo ciklo etapus.

Analizuojant mokslinę literatūrą [8] pastebima, kad išskiriami labai mažos, mažos, vidutinės, didelės ir stambios apimties projektų tipai. Šie tipai atskiriami pagal darbų apimtį, trukmę ir biudžetą.

III. IT PROJEKTŲ VALDYMO METODOLOGIJOS IR STANDARTAI, JŲ VERTINIMAS

Projekto apimties valdymas užtikrina, kad būtų vykdomi būtini darbai, nevykdomi nebūtini darbai ir kad darbai, kurie yra vykdomi, neštų reikiamą naudą. Apimtimi apibrėžiamas galutinis laukiamas rezultatas bei tarpiniai darbo produktai. Projekto vadovas, tik prasidėjus projektui, užima pagrindinį vaidmenį, nes tik pradžioje turi būti apibrėžiami reikalavimai, apimtis. Tokiu principu vadovaujasi tradiciniai projektų valdymo metodai, kurių pagrindinis tikslas - fiksuoti apimtį. Vikrieji (angl. *Agile*) ir RUP metodikos projekto apimtį apibrėžia sprinto ar iteracijos pradžioje, nes manoma, kad tokiu būdu reikalavimai iš vartotojo gaunami detalesni, kai jis suvokia, kokios sistemos jam reikia.

A. Tradiciniai projektų apimties valdymo būdai

Apibrėžiant apimtį būtina dokumentuoti reikalavimus tam, kad projekto eigoje nebūtų iškreipiami rezultatai. Stellman A. ir Greene J. [13] naudoja vizijos ir apimties dokumentą, kuriame atsispindi visi reikalavimai sistemai. Gera vizija ir apimties dokumentas padės išvengti brangiai kainuojančių problemų. Stellman A., ir Greene J. savo veikale *Applied Software Project Management* (2006) siūlo dokumentuoti projekto apimties reikalavimus tam tikra tvarka.

1) Problemos identifikavimas
   a) *Projekto pagrindimas.* Apibendrinama problema, kuri turi būti išspręsta. Turi būti trumpa problematikos istorija ir paaiškinimas kaip organizacija pagrindžia savo apsisprendimą kurti sistemą. Ši skiltis turi atskleisti priežastis, kodėl problema egzistuoja, nurodyti ankščiau vykdytus projektus, kurie buvo susiję su problemos sprendimu.
   b) *Suinteresuotosios šalys*. Tai sąrašas žmonių. Nurodomas kiekvieno poreikis trumpais sakiniais.
   c) *Vartotojai*. Nurodomas kiekvieno poreikis trumpais sakiniais.
   d) *Rizika*. Įvardinamos potencialios rizikos. Rizikos surenkamos komandinio susirinkimo metu. Apima tiek vidinius, tiek išorinius faktorius.
   e) *Prielaidos*. Sudaro vartotojai, suinteresuotos šalys ir projekto komanda.
2) Sprendimo vizija
   a) *Vizijos formuluotė*. Tikslas – apibūdinti, ko tikimasi iš projekto. Turi paaiškint, koks yra projekto tikslas. Tikslas turi būti pagrįstas.
   b) *Funkcijų sąrašas*. Patartina, kad sistema turėtų ne daugiau kaip dešimt funkcijų, kitaip bus susiduriama su dideliu sudėtingumu.
   c) *Laipsniškas paleidimas*. Kartais patogu, kad kuriant IT projektus, pakeitimai būtų keliami laipsniškai. Pirmiausia keliamas poaibis funkcijų – pirminė produkto versija, o vėlesnės versijos keliamos vėliau.
   d) *Funkcijų, kurios negali būti plėtojamo, sąrašas*. Funkcijos dažnai nepatenka į projekto rėmus. Kai taip atsitinka, visos nepageidaujamos funkcijos turi būti surašytos šioje skiltyje, kad užsakovas matytų, jog jis to atsisakė.

Reikalavimai surenkami iš projekto suinteresuotųjų šalių, kurie yra pagrindiniai proceso dalyviai, žinantys kaip turi veikti sistema. Jie sprendžia, koks funkcionalumas jiems reikalingas.

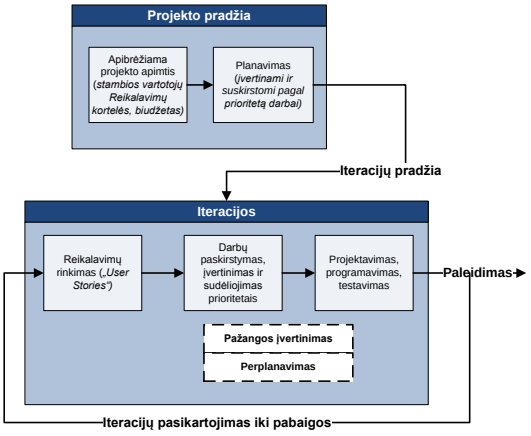
Pagal PMBOK (projektų valdymo gerosios praktikos standartas)[6], apimties valdymas atliekamas tokiais žingsniais:

1. Reikalavimų surinkimas. Apibrėžiami ir dokumentuojami suinteresuotų asmenų poreikiai kuriamai sistemai. Projekto sėkmė yra tiesiogiai susijusi su tuo, kaip teisingai ir tiksliai yra nustatomi reikalavimai produktui ir kaip suprantamas vartotojo poreikis.
2. Apimties apibrėžimas. Kuriamas dokumentas, apibrėžiantis visus darbus, kurie turės būti vykdomi. Šis dokumentas atsako į pagrindinius klausimus: kokia kuriamo produkto paskirtis, numatomi atlikti darbai, numatomi kriterijai, apribojimai ir prielaidos.
3. WBS dokumento sudarymas. Tai procesas, kuris suskirsto rezultatus ir projekto darbus į smulkesnius, labiau suvaldomus komponentus. Tai hierarchinis projekto apimties išskaidymas arba grupavimas.
4. Apimties patikrinimas. Tai procesas, kurio metu formalizuojama sukomplektuoti projekto rezultatai.

5. Apimties kontrolė. Procesas, kurio metu prižiūrimas projekto statusas.

B. Vikrieji metodai

Vienas iš Vikriųjų grupės projektų valdymo metodikų yra XP [7], kuris labiau pritaikomas mažiems projektams [8] ir mažoms įmonėms, kur vidutinis komandos dydis 3-12 žmonių. XP valdymas pagrįstas iteracijomis. Iteracinis XP procesas vaizduojamas 2 paveiksle.



Šaltinis [7]: EXTREME PROGRAMMING (2001). Construx software builders, p. 9

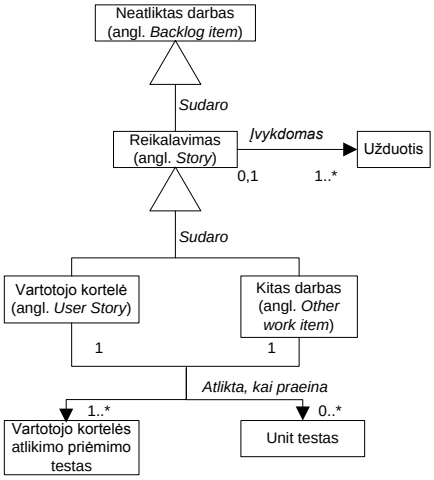
2 pav. XP darbų sekų modelis

2 paveiksle vaizduojama XP proceso darbų seka, kurios pradinis žingsnis yra apimties aprašymas. XP aprašydamas apimtį remiasi Ganto grafiko sudarymu. Ganto schema yra seniai naudojama planavimo priemonė. Toks būdas yra santykinai paprastas, todėl jį racionalu naudoti rengiant nedidelius, paprastesnius projektus, kai kiti planavimo būdai tiesiog reikalauja per didelių laiko sąnaudų. Svarbiausia, kad turima informacija leistų įvertinti, kiek laiko reikia kiekvienai užduočiai atlikti.

XP apimties valdyme naudoja WBS dokumentą. WBS dokumento sudarymas - procesas, kuris suskirsto rezultatus ir projekto darbus į smulkesnius, labiau suvaldomus komponentus. Tai hierarchinis projekto apimties išskaidymas arba grupavimas. XP nereikalauja WBS sudaryti labai tiksliai, bet čia yra naudojami vartotojų reikalavimai (angl. *User story*), kurie gana tiksliai nusako, kas turi būti padaryta. Tai yra vartotojo reikalavimų surinkimo technika, kuomet iš kiekvieno potencialaus kuriamos sistemos vartotojo yra surenkamas poreikis. Šio proceso kuratorius – užsakovas. Treneris naudoja klausimų seriją apie konkretų funkcionalumą, kuris palengvina išsiaiškinti tikslius vartotojų poreikius. Tačiau klausimai turi būti suformuluoti taip, kad neįtikinėtų vartotojo, o treneris nedominuotų procese. Vartotojų *stories* surašomos ant kortelių. Jei randamas toks apibūdinimas, kuris nėra išsamus arba per daug sudėtingas, jis yra perrašomas iki bus tenkinami trenerio ir užsakovo reikalavimai. *User stories* aprašymui yra naudojamas INVEST metodas.

SCRUM [15] reikalavimams surinkti yra naudojami tie patys metodai kaip ir XP. Pasitelkiama *User stories* kortelės, kurios audituojamos INVEST būdu. SCRUM naudojamos sąvokos ir terminai:

Product Backlog - tai prioritetų tvarka sudarytas norimų funkcijų sistemai sąrašas. Naudojant Scrum nėra būtina iš anksto dokumentuoti visus numatytus reikalavimus. Paprastai pradedama tuo, kad projekto komanda ir produkto užsakovas užrašo tai, kas yra lengviausiai įvardinama. Dažniausiai to visiškai užtenka, kad būtų pradėtas pirmasis *Sprint* ciklas. Tuomet *Product Backlog* paliekamas „augti". Scrum reikalavimų surinkimo procesas pateiktas 3 paveiksle.



Šaltinis [15]: WESTLEY, A (2011). Agile Software requirements, p. 61

3 pav. Reikalavimų surinkimo valdymas

3 paveiksle matyti, kad reikavimas (angl. *Story*) yra neatlikto darbų sąrašo (angl. *Product Backlog)* dalis ir yra sudarytas iš vartotojo reikalavimų aprašymo (angl. *User story)* ir kito darbo (angl. *other work item*). Kitas darbas gali būti, pvz., palaikymas, aptarnavimas, infrastruktūros darbai. Pats reikalavimas yra skaidomas į užduotis, kitaip tariant reikalavimas yra įgyvendinamas, atlikus užduotis. Vartotojų kortelės priėmimo testas (angl. *Story acceptance test*) naudojamas kaip patvirtinimas, kad reikalavimas yra atliktas korektiškai. Ryšys (1 ; 1*) rodo, kad kiekvienas reikalavimas turi vieną arba daugiau priėmimo testų. *Unit* testas skirtas tam, kad patvirtintų žemiausio lygio aplikacijos (procedūros, klasės, metodo ar funkcijos) tinkamą veikimą. Testas visada turi būti sukurtas programuotojo, kuris įgyvendina reikalavimą. Išskaidant reikalavimų korteles į užduotis, gaunamas galutinis *Backlog* sąrašas.

IV. IT PROJEKTŲ APIMTIES VALDYMO METODŲ PALYGINIMAS

1 lentelėje pateikiami projektų valdymo (informacinių sistemų kūrimo) metodai, kurie palyginami pagal kelis kriterijus.

1. Taikymo sritis. Nurodoma, kokioms įmonėms ar projektams gali būti pritaikomas modelis, ar tinkamas sprendžiant IT problemas ir procesus.
2. Apimties įvedimas į procesą. Nurodoma, kokio pobūdžio yra apimtis – ar apibrėžiama projekto pradžioje, ar yra numatyti pokyčiai.
3. Apimties aprašymui naudojamos technikos.
4. Privalumai.
5. Trūkumai

1 LENTELĖ. IT PROJEKTŲ APIMTIES VALDYMO METODŲ PALYGINIMAS

Kriterijus	PMBOK	Stellman A., ir Greene J	Scrum	XP
Taikymo sritis	Daugiau didelėms įmonėms, nebūtinai IT projektams	Skirta programinės įrangos kūrimo projektams	Pagal konfigūraciją galima pritaikyti ir didelėms įmonėms ir projektams	IT projektams, mažos apimties įmonės
Apimties įvedimas į procesą	Apibrėžiama planavimo stadijoje, griežtai dokumentuota, nekinta visą projekto laikotarpį	Projekto startavimo metu, tačiau esant problemoms, gali būti keičiama su laiko ir resursų sąnaudomis	Apibrėžiama pradžioje ir kiekvieno sprint metu. Vykstant sprint, apimtis „užšaldoma“.	Apibrėžiama pradžioje per user stories ir kinta iteracijų metu
Privalumai	Griežtai dokumentuota, paremta logiškais sprendimo priėmimo metodais.	Skirta programinės įrangos kūrimui, išvystytas visas valdymo ciklas.	Gali būti pritaikomas jau vykstančiam projektui	Skirta programinės įrangos kūrimui, prisitaiko prie kintančių reikalavimų, gerai išvystas testavimo valdymas.
Trūkumai	Tai bendrų dėsnių standartas, reikalinga atskira modifikacija norint pritaikyti konkrečiam projektui.	Apibrėžia reikalavimus programinei įrangai.	Silpnai atliekamas analizės etapas, todėl sunku prognozuoti projekto trukmę. Gali tęstis ilgai.	Jei programinės įrangos kūrimo technologija nesuderinama su besikeičiančiais reikalavimais, ji negali būti panauda XP

Buvo išanalizuoti vikrieji ir tradiciniu procesų valdymu pagrįstos metodikos. Tradicinės - panašios tarpusavyje – jų taikymo sritis - dažniausiai didelių įmonių IT projektai, o valdymo procesas pagrįstas etapų seka. Apimtis apibrėžiama projekto pradžioje – inicijavimo, startavimo ar planavimo stadijose. Numatoma, kad ji nekis visą laikotarpį arba kis su laiko ir resursų sąnaudomis. Šių valdymo metodikų privalumas, kad stiprioji jų vieta - analizės etapas. Ši vieta svarbi apibrėžiant apimtį, nes aukšto lygio procesų analizė yra būdas sustabdyti apimties nuokrypį, fiksuoti resursus ir pabaigos datą.

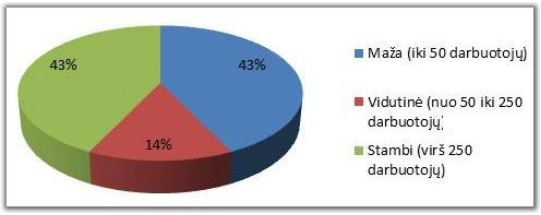
Vikriuosiuose metoduose peršokama iškart į sprintą ar iteraciją, kurių metu atliekama ir analizė, ir apimties apibrėžimas. Pasibaigus sprintui, lieka nedaug laiko, kad būtų tinkamai išanalizuoti kiti reikalavimai, todėl išlieka tikimybė į sprintą įvesti klaidingus reikalavimus, kurie įtakos prailgėjusius terminus ir padidėjusius kaštus. Kita vertus, Scrum ir XP siūlo patraukliausią sprendimą analizuojant mažos apimties IT projektus. Jie geba prisitaikyti prie kintančių reikalavimų, o galutinis rezultatas nėra apibrėžiamas iš anksto, todėl užsakovas gauna visiškai realizuotą sistemą, o ne funkcijas, kurias apibrėžė projekto pradžioje (iš praktinės pusės yra žinoma, kad vykstant projektui išaiškėja įvairūs funkciniai reikalavimai, kurie tikrai reikalingi sistemai, tačiau dėl griežtai apibrėžtos apimties jų realizuoti negalima).

Taigi, išanalizavus visus metodus galima teigti, kad nėra vieno „geriausio“ varianto. Kiekviena metodologija gali būti pritaikoma pagal įmonės modifikaciją. Tačiau analizuojant mažos apimties IT projektų kūrimo ir valdymo metodus išskiriami SCRUM ir XP, nes jie leidžia jau pradiniuose projekto etapuose gauti greitus rezultatus. Taip vartotojai gali greičiau reaguoti į netinkamus, pakeisti savo reikalavimus sistemai ar pakreipti programavimo darbus kita, jiems tinkama linkme. Šie metodai taip pat naudingi, nes yra linkę keisti reikalavimus. Kitos metodologijos griežtai pasisako prieš, o tai gali įtakoti projekto nesėkmę.

V. IT PROJEKTŲ SĖKMĖS FAKTORIAI

Tyrimas vyko 2012 m. gruodžio – 2013 sausio mėn. El. paštu išsiųsta 80 anketų, grįžo 21 (26.2%). Tyrimo imčiai sudaryti taikyta atsitiktinė tikslinė atranka.

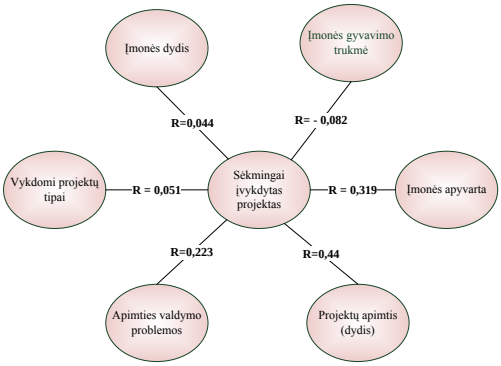
Pagrindiniai tyrimo dalyvių atrankos kriterijai – mažos įmonės, potencialiai vykdančios projektinę veiklą. Tyrime dalyvavo 21 įmonė, iš jų 9 mažos. Tyrime dalyvavusių įmonių pasiskirstymas pagal dydį pateikiamas 4 pav.



4 pav. Apklaustų įmonių pasiskirstymas pagal dydį (N=21), %

Mažiausiai – 14 % apklausoje dalyvavusių vidutinio dydžio įmonių. Vienodas pasiskirstymas (43%) pastebėtas tarp vidutinių ir mažų įmonių.

Atlikta koreliacinė analizė parodė ryšių stiprumą tarp sėkmingai įvykdyto projekto ir (1) įmonės dydžio, (2) gyvavimo trukmės, (3) apyvartos, (4) projektų apimties (projekto dydžio), (5) apimties valdymo problemų bei (6) vykdomų projektų tipų. Galima daryti išvadą, kad sėkmingai įvykdytam projektui didžiausią įtaką daro įmonės apyvarta, projekto dydis ir apimties valdymo problemos, kadangi koreliaciniai ryšiai tarp šių rodiklių patys didžiausi. 5 paveiksle demonstruojami koreliacinės analizės rezultatai, apibendrinus visus rodiklius.



5 pav. Koreliacinės analizės rodikliai

Išanalizavus anketos duomenis, paaiškėjo su kokiomis problemomis susiduria įmonės, vykdydamos mažos apimties projektus. Rezultatai pateikiami 6 pav.

6 pav. Mažų IT projektų apimties valdymo problemos

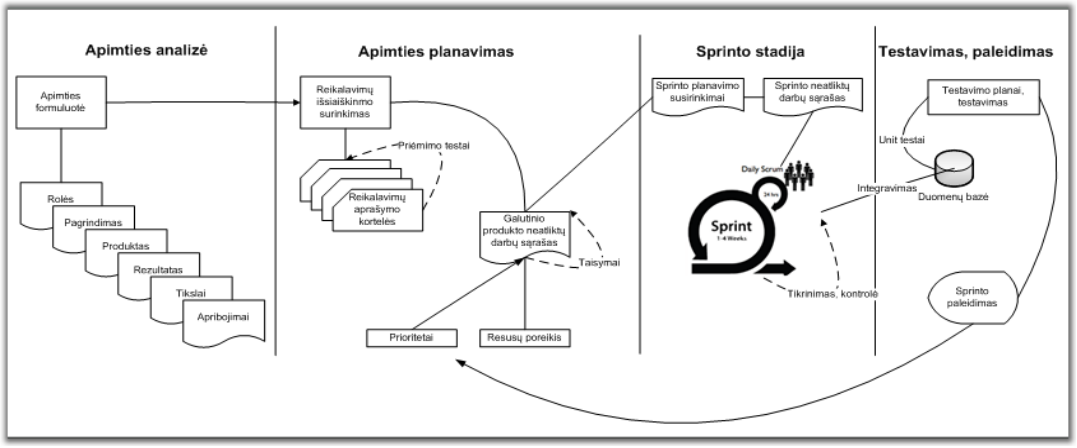
Iš 6 paveikslo matyti, kad maži projektai dažniausiai kenčia nuo projekto apimties augimo. Nekontroliuojami vartotojo reikalavimų pokyčiai taip pat opi problema. Su ja susiduria 60 % respondentų. Netikslus apimties planavimas taip pat daro įtaką projekto sėkmei.

Iš 7 paveikslo matyti, kad mažiems projektams visai neaktualios vartotojų reikalavimų surinkimo kortelės ir jų patikrinimo testai. Visos įmonės, kurios vykdo mažos apimties projektus, teigia, kad dokumentavimas vienas iš svarbiausių būdų suvaldyti apimties pokyčius. Jos taip pat naudoja WBS dokumentą ir einamojo etapo apimties sudarymą. Apibendrinant galima teigti, kad norint mažos apimties projektus vykdyti efektyviau, reikėtų didesnį dėmesį kreipti į apimties analizės valdymą.

VI. SIŪLOMAS MAŽOS APIMTIES IT PROJEKTŲ VALDYMO METODAS

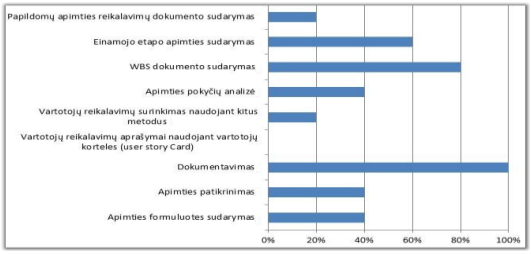
Atsižvelgus į analizės rezultatus ir į anksčiau išanalizuotus projektų apimties valdymo metodus, siūlomas modifikuotas modelis, leidžiantis efektyviau organizuoti mažų IT sistemų kūrimo projektų apimties valdymą (8 pav.).

Sudarant modelį remtasi PMBOK bei vikriaisiais XP ir SCRUM metodais. Šie metodai pasirinkti, nes labiausiai atitinka nagrinėjamą temą. XP ir Scrum puikiai pritaikomi mažos apimties IT projektų valdymui, o



Pav. 8 Mažos apimties projektų valdymo siūlomas metodas

Analizės metu taip pat paaiškėjo, kad mažiems projektams suvaldyti dažniausiai naudojami 7 pav. rodomi metodai ir būdai. Kaip patogiausią projektų valdymo problemų sprendimo būdą pasirinko dokumentavimą. Tai patogus ir naudingas, bet išsprendžia ne visas problemas. Nei vienas respondentas nepasirinko vartotojų poreikių kortelės aprašymo testų, kaip alternatyvos problemoms spręsti, bet šias korteles naudoja, norėdami pagerinti reikalavimų surinkimo procesą. Apimties pokyčių analizė – dar vienas labai dažnai naudojamas problemų sprendimo būdas.



7 pav. Mažų IT projektų apimties valdymo problemos

PMBOK [11] sukuria analizės valdymo struktūrą, kuri prastai apibrėžta XP ir Scrum. Modelį sudaro 4 etapai:

1. Apimties analizė
2. Apimties planavimas
3. Sprint stadija
4. Testavimo, paleidimo stadija

Atlikus projektų apimties valymo analizę ir koreliacinę analizę pastabėta, kad XP ir Scrum metoduose analizė atliekama labai paviršutiniškai. Šie du modeliai koncentruojasi į sprinto ir iteracijos valdymą, produkto kūrimą. Tačiau analizės nebuvimas įtakoja kitas problemas - neprognozuojamą apimties projekto terminų ir išlaidų kaitą. Vykdant projektus pagal Vikriuosius metodus rekomenduojama atlikti išsamią apimties analizę pagal PMBOK. Po analizės sudaromas apimties formuluotės dokumentas, kuris apibrėžia:

- Projekto pagrindimas – suformuluojama problema, dėl ko reikalinga kurti sistemą.
- Apibrėžiamas galutinis produktas. Bendrais bruožais pasakoma, kas turi būti padaryta. Neturi būti labai detalu.

- Planuojami rezultatai. Pasakoma, kokia nauda bus organizacijai, jei bus išpildyti apimtyje numatyti reikalavimai.
- Apimties tikslai. Apibrėžia sistemos funkcionalumą. Neturi būti aprašomas labai detaliai.
- Apribojimai – galimos problemos, kuriant sistemą.

Šis dokumentas apibrėžia planuojamus atlikti pakeitimus ir tikslus, kad vykdant kitus etapus, nebūtų galimybės išaugti apimčiai. Šioje vietoje galioja SCRUM ir XP principas, kad apimtis gali kisti, tačiau norima kontroliuoti kaitos ribas.

VII. IŠVADOS

Projekto apimties valdymas užtikrina, kad būtų vykdomi būtini darbai, nevykdomi nebūtini darbai ir kad darbai, kurie yra vykdomi būtų naudingi.

Atlikus projektų valdymo metodologijų analizę paaiškėjo, kad labiausiai tinkami modeliai mažos apimties IT projektams yra SCRUM ar XP metodai, nes jie leidžia per greitą laiką pateikti vartotojui veikiančios sistemos karkasą. Trūkumas – silpnai atliekama analizė.

Atlikus koreliacinę analizę matoma, kad IT projektų sėkmė labiausiai priklauso nuo apimties valdymo. Norint suvaldyti šio proceso kaitą, pasiūlytas modifikuotas metodas, leidžiantis daugiau dėmesio skirti apimties valdymui. Siūlomos modifikacijos esmė – detaliai aprašyti laukiamus viso projekto rezultatus, kuo daugiau sumažinant riziką padidinti projekto kaštus ar terminus.

LITERATŪRA

[1] Atkinson, R. (1999). Project Management: cost, time and quality, two best guesses and a phenomenon, its time to accept other success criteria. Computer science. International Journal of Project Management. March, p. 337 – 342.

[2] Bentley, C. (2005) *Practical PRINCE2* [interaktyvus]. London: TSO. 100 p.ISBN: 0-11-703544-0 [žiūrėta 2012m. sausio 2 d.]. Prieiga per internetą <http://www.google.lt/books?hl=lt&lr=&id=a46v0GieDvAC&oi=fnd&pg=PP9&dq=Prince2&ots=TgPX8Pf2a8&sig=X8cGm65x8fl58KxBRTOlZ5CFMPY&redir_esc=y#v=onepage&q=Prince2&f=false>

[3] Bozhikova V., Stoeva M., Tsonev K. (2009). *A practical approach for software project management.*

[4] BRITISH STANDARDS INSTITUTE (2006). *Project management – Part 4: Guide to project management in the construction industry*. London: BSI. 82 p. ISBN 0 580 49045 9

[5] CHAOS REPORT (2009). Chaos summary 2009 [Interaktyvus]. JAV: The Standish Group International. standishgroup.com [žiūrėta 2011, Gruodžio 8 d.] Prieiga per internetą:<http://docs.google.com/viewer?a=v&q=cache:a1zv4XyswLAJ:www.portal.state.pa.us/portal/server.pt/document/690692/standish_group_chaos_summary_2009_pdf%3Fqid%3D44137104%26rank%3D1+chaos+report+2009&hl=lt&gl=lt&pid=bl&srcid=ADGEESjdQaMcZ4aeTHh8NSLc30XKCvWF0qXoPzBDxCUYA3CsA1Q3-b6elGQOydxMhzls67656L0J-Ouv1NRnfXO3lOu5D10dq6eeXWKiT86yUKOHrANTMLghiZpCUdRPRgxuQwQWYWSJ&sig=AHIEtbQ_7-oRXEDBUIxO4Ei3tlUcJc8Itw

[6] Dugan W.(2009). *A Guide to the Project Management Body of Knowledge*. JAV: Project Management Institute

[7] EXTREME PROGRAMMING (2001). Construx software builders, p. 9

[8] Hastie Shane. (2008) Project Classification [interaktyvus]. Australia: Software Educatiopn Group. softed.com. [žiūrėta 2011 gruodžio 5 d.]. Prieiga per internetą <http://www.softed.com/Resources/Docs/ProjectClassification.pdf >

[9] INTERNATIONL ORGNIZTION OF STNDARTIZATION (2003). Quality management systems – Guidelines for quality management in projects. [Interaktyvus, žiūrėta 2013 Vasario 5 d.] Prieiga per internetą: <http://marcelrzmut.comxa.com/2APII/ISO10006 2003.pdf>

[10] Kerzner H. (2006). *Project management: a systems approach to planning, scheduling, and controlling.* New York: John Wiley and Son.

[11] PROJECT MANGEMENT INSTITUTE (2004). *Guide to Project Management Body of Knowledge.* Pensilvnija:405 p.ISBN 1-930699-50-6.

[12] STATISTIKOS DEPARTAMENTAS (2011). I*nformacinės technologijos Lietuvoje 2011*. Vilnius: Statistikos departamentas prie Lietuvos Respublikos Vyriausybės. 149 p. ISSN 2029-3615

[13] STELLMAN, A.; GREENE, J. (2006) *Applied Software Project ManagemenT.* Tokyo: O'Reilly Media, Inc. 318 p. ISBN: 978-0-596-00948-9.

[14] Turner, R.. (2007) *Gower handbook of project management*. [interaktyvus]. United Kingdom: Gower Publishing, 912 p. ISBN: 978-0-566-08806-3 [žiūrėta 2011m. gruodžio 20d.] http://www.google.lt/books?hl=lt&lr=&id=hAPsnKp-OhEC&oi=fnd&pg=PR17&dq=Gower+Handbook+of+Project+Management&ots=OW3IrbwDa4&sig=WZBdxERBcs4v4b1NqQwYBYeHouY&redir_esc=y#v=onepage&q=Gower%20Handbook%20of%20Project%20Management&f=false

[15] Westley, A (2011). *Agile Software requirements.* 558 p. ISBN-13: 978-0-321-63584-6

INFORMACINĖS TECHNOLOGIJOS ● IVUS 2013 ● ISSN 2029-249X ● eISSN 2029-4824

Simulation of a cognitive radio system for interference reduction on the 2.4 GHz ISM band in hospital environments

Dainius Jankūnas
Department of Applied Informatics
Vytautas Magnus University
Kaunas, Lithuania
dainius.jankunas@live.vdu.lt

Raúl Chávez-Santiago
The Intervention Centre
Oslo University Hospital
Oslo, Norway
raul.chavez-santiago@rr-research.no

Ilangko Balasingham
The Intervention Centre
Oslo University Hospital
Oslo, Norway
ilangko.balasingham@medisin.uio.no

Vladislav V. Fomin
Department of Applied Informatics
Vytautas Magnus University
Kaunas, Lithuania
v.fomin@if.vdu.lt

***Abstract*—Wireless communications technologies are used to support many applications which require mobility and flexibility. A number of wireless devices operate in the 2.4 GHz ISM band and the interference is inevitable. ZigBee and Wi-Fi are two technologies used in hospital environments for which the co-existence problem has to be solved. We find that cognitive radio (CR) can provide a solution by intelligently scanning the environment and assigning proper channels to different communication applications. In this paper, at first, the 2.4 GHz ISM band and wireless technologies operating in these bands are presented. Then, a cognitive radio (CR) algorithm for medical environment is explained. The CSMA/CA with RTS/CTS protocol is simulated. Finally, multi-channel scenario is introduced and a possible database structure defined for the CR system is discussed.**

Keywords—2.4 GHz ISM band; ZigBee; Wi-Fi; Cognitive Radio; Database

I. INTRODUCTION

Wireless communications is a key technology to improve mobility and flexibility, instead of having to be tethered to a limited wired connection. Since a good number of wireless devices operate in the 2.4 GHz ISM band the co-existence problem of multiple wireless technologies arises [1, 2]. The possible technologies being used in hospital environment are ZigBee [3], which is used for sensor networks and Wi-Fi networks, which are everywhere. However, most medical devices used in hospital environment are sensitive to electromagnetic interference (EMI) caused by wireless transmissions. Cognitive Radio (CR), which is software-defined radio, has emerged as key technology to increase the radio spectrum utilization and reduce interference. A cognitive radio transceiver can intelligently detect which channels are in use and which are not, and instantly move to vacant channels while avoiding occupied ones. Furthermore, CR transceiver can adjust wireless transmission parameters by sensing neighboring wireless devices in the operating environment. The objective of this article is to explain a proposal to evaluate CR in 2.4 GHz ISM band.

The rest of the paper is structured as follows. Section 2 presents an overview of the 2.4 GHz ISM band and describes potential co-existence problems of using multiple medical applications. In Section 3 a cognitive radio system [4] for medical environments is explained. Section 4 introduces a simulation of medical communication in the 2.4 GHz ISM band and proposes a multi-channel scenario. Finally, Section 5 concludes the paper.

II. THE 2.4 GHZ ISM BAND

In this section a short overview of the various radio technologies operating in the 2.4 GHz unlicensed ISM band is presented. The main focus is on ZigBee (802.15.4) and Wi-Fi (802.11) co-existence problems.

The license-free industrial, scientific and medical (ISM) radio band is reserved for the use of radio frequency (RF) spectrum for industrial, scientific and medical purposes. With so many devices built in to operate in 2.4 GHz ISM band, there is no doubt that crowding is a problem. Different medical applications use this band. For example, ZigBee is used for sensor networks due to its long battery life and low cost. Wi-Fi is used for medical telemetry for transmission of data related to patient's health. Both, ZigBee and Wi-Fi, belong to the class of spread-spectrum technologies, which occupy much larger bandwidth than required for data to be sent. Because the signal is spread over a large bandwidth, it can co-exist with other signals, but performance decreases.

The RF spectrum and available channels for ZigBee and Wi-Fi overlap as shown in Fig. 1. ZigBee divides the 2.4 GH band into 16 non-overlapping channels, which are 2 MHz wide and 5 MHz apart. Wi-Fi channel is 22 MHz wide and this allows up to three channels to be used simultaneously without overlapping each other. Careful channel selection and assignment can be the solution [5] to avoid interference. This can be achieved by selecting ZigBee channels that use the free space between two neighboring Wi-Fi channels, plus channels 25 and 26. Therefore, Wi-Fi channels 1, 6 and 11 can be used with ZigBee channels 15, 20, 25 and 26 simultaneously without a risk of interference.

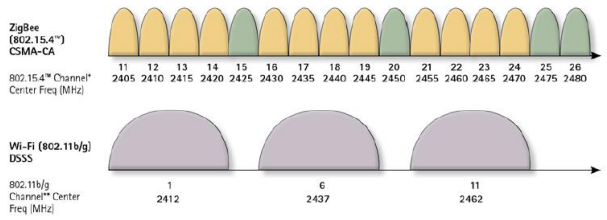


Fig. 1. ZigBee and Wi-Fi channels [3]

III. COGNITIVE RADIO FOR MEDICAL ENVIRONMENTS

This section explains an algorithm of a cognitive radio system for medical environments [4], which will be used as a core for our simulation with some modifications.

The proposed cognitive radio system consists of the following components: a database, the cognitive radio controller and cognitive radio clients. The database will be used to store all the information about medical devices. The cognitive radio controller will be used to control the transmission parameters of the cognitive radio clients for wireless access. There will be two types of users: ZigBee and Wi-Fi clients. In [4] every cognitive radio client transmits data through the cognitive radio controller, which has two radio interfaces, one for the control channel and the other for the data channel. We will extend this to use more than one channel, which will be explained later. The controller gets information about medical devices from the database, which is required to compute the transmission parameters for the users. The controller can control the transmission power of the cognitive radio clients using EMI-aware request-to-send/clear to send (RTS/CTS) protocol [4]. The detailed RTS/CTS protocol is described in Section 4.

At first, the cognitive radio controller calculates and broadcasts the maximum transmit power for transmitting an RTS message without any harmful interference to any medical device. Since the users change their position and medical devices can be turned on or off, the transmit power is calculated and broadcasted periodically. Next part of the algorithm is RTS/CTS protocol extended with possibility to drop the user if the controller cannot find a feasible transmit power that satisfies QoS requirements (i.e., minimum data rate) or the EMI constraint on the medical devices. This is calculated after each successful RTS message. If two or more users send an RTS message at the same time then the collision occurs. After collision each user waits different back-off time and then resends an RTS message. After number of retries the user is dropped. Also, the controller can drop users randomly to avoid congestion. If the CTS message is successfully received by the user, the user is placed in the queue before it can really transmit its data. This happens because all the users use the same data channel to transmit data. This could be avoided if we apply multichannel scenario in which there are more possible channels to use.

IV. SIMULATION OF MEDICAL COMMUNICATION IN THE 2.4 GHZ ISM BAND

A cognitive radio system proposed in [4] considers one-channel only. It means that only one user can access the channel at a time. If channel is occupied the user has to wait while channel becomes free. As explained in [7, 8] using more channels can satisfy the cognitive radio wireless sensor nodes' bandwidth requirements and interference problem. In our case we consider a multi-channel scenario. As mentioned before, Wi-Fi and ZigBee can co-exist by using proper channel selection and assignment (see Fig.1). There are the following co-existence scenarios: when no Wi-Fi channels are assigned in the area, when one or two, or free channels are assigned. In our simulation we will try the worst case scenario, which is when all three Wi-Fi channels are assigned and ZigBee has only four possible channels to use without interference.

First, the CSMA/CA with RTS/CTS protocol will be implemented for comparison with the cognitive radio algorithm which will be implemented after that. The performance of the RTS/CTS protocol is evaluated by simulations using MATLAB. The simulation parameters are placed in Table 1. One of the characteristics [6] to evaluate if our algorithm is working properly is to measure the throughput. Throughput is the number of packets that are successfully transmitted during time specified. Fig. 2 shows the throughput of our simulation with the different number of users. Increasing number of users decreases the throughput because collisions are more likely. To calculate the throughput we use the following formula:

$$Throughput = \frac{time\ used\ for\ successful\ transmission\ in\ an\ interval}{length\ between\ two\ consecutive\ transmissions} \quad (1)$$

Next step will be to implement the cognitive radio algorithm as described in Section 3 and compare the results with the traditional CSMA/CA RTS/CTS protocol. We will measure the interference and outage probabilities for both protocols. The interference probability indicates the chance that the user transmission causes interference to the medical devices. The interference occurs when the required transmit power is higher than the acceptable level. The outage probability is the probability that the user cannot transmit. If we get the similar results to [4] we will assume that our simulation is good and continue with the multi-channel approach. As mentioned before we will modify the cognitive radio algorithm to work under multi-channel scenario. We will compare the interference and outage probabilities again once we implement multi-channel scenario with the previous CSMA/CA RTS/CTS and EMI-aware RTS/CTS protocols. We expect that using the multi-channel scenario will decrease the outage probability. And finally, the database will be implemented under the structure we define in this article.

A. CSMA/CA with RTS/CTS Protocol

CSMA/CA (Carrier Sense Multiple Access/Collision Avoidance) protocol is well known in IEEE 802.11 networks. Unlike wired networks where collisions can be detected (e.g. CSMA/CD), in a wireless environment collisions cannot be detected and CSMA/CA protocol is beneficial in such networks. The CSMA/CA protocol works as follows. When a user wants to transmit data, the system has to wait until the channel is idle for a distributed inter frame space (DIFS) period. If the channel is clear after DIFS period of time, the

user can transmit its data. If the channel is not clear, the user uses a random number generator to set a back-off time. In the following time slot back-off time decreases by one if channel is found idle. In case channel is busy, the user stops counting and waits till channel is idle for DIFS again. Once the back-off time reaches zero and channel is found busy, the user has to set the back-off time again. Note, that after each collision to this data packet the contention window (CW) is doubled. CW is interval in which random number is generated. After successful data transmission the controller waits for a short inter frame space (SIFS) and sends an ACK message after receiving without error. At this point the process is repeated.

RTS/CTS is an additional method to implement virtual carrier sensing in CSMA/CA. The main difference is that RTS and CTS messages are sent to reserve channel. RTS/CTS mechanism works as follows. Before transmitting data each user has to wait until the channel is sensed idle for a DIFS period. After DIFS time user sends an RTS message to the controller. Once the controller successfully receives an RTS message, it waits for a SIFS and then sends a CTS message to the user. Any other users hear an RTS or a CTS message refrain from sending data for a time specified in RTS/CTS message. This solves hidden node problem. After CTS message user starts transmitting data immediately after SIFS time. When data transmission is completed controller notifies the user by sending an ACK message after SIFS time. As we see the data packets are protected from collisions, but RTS messages are not. They are still vulnerable to collisions. However, RTS messages are short compared to data packets, collisions are less likely and can be resolved much quicker.

TABLE I. SIMULATION PARAMETERS

Max Payload Size	DIFS	SIFS	RTS/CTS/ACK
127	3	1	2

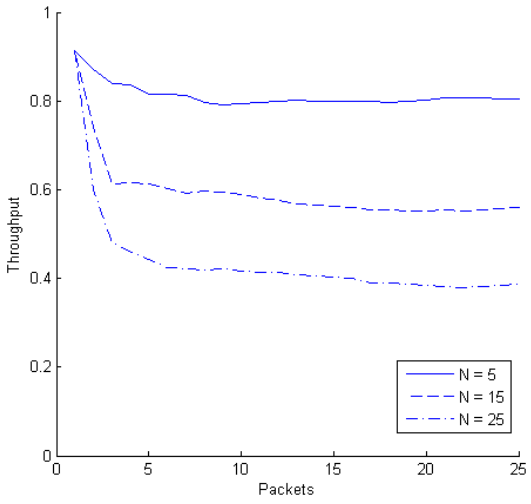


Fig. 2. The ThroughPut of CSMA/CA with RTS/CTS

B. Database Structure for Cognitive Radio System

A database in a cognitive radio system will be used to store information about all medical devices and users in the hospital (e.g., location, activity status, and EMI immunity level). This information will be used by cognitive radio controller to calculate transmit power and other transmission parameters. Fig. 3 shows the possible database structure. To calculate the transmit power the following parameters are required: the distance from the user to the non-life-supporting and life-supporting medical devices, the EMI immunity levels for non-life-supporting and life-supporting devices in terms of the electric field (measured in volts per meter) in which the medical devices can operate normally [4]. All medical devices will be stored in table Device. Every device has the following attributes:

- Location, which describes the position of the device in hospital environment. The location can be any room or hall way in any floor. It can be updated by using any tracking system.
- Type, which is non-life-supporting device or life-supporting device.
- Status, which is on or off. Since medical devices are turned on or off it is important to keep such information in a database. Furthermore, it is crucial to update and synchronize this attribute with the controller on time because incorrectly reported status of the medical device can cause interference with the wireless device.
- EMI immunity level here is defined in terms of the electric field (measured in volts per meter) in which the medical devices can operate normally [4].
- Desc here is just human readable description of the medical device or additional information about it.

Second table is for storing all the users with the following attributes:

- Location, which describes the position of the user in hospital environment. The location can be any room or hall way in any floor. It can be updated by using any tracking system..
- Type, which is ZigBee or Wi-Fi user.
- Power, which is the maximum transmit power calculated by the cognitive radio controller.
- Channel, which is any vacant channel assigned for the user.
- Active, which represents if user is currently active or shut down.
- Desc, which is human readable description

The third table will keep all the possible places in hospital environment.

The cognitive radio controller takes all the active users from the database and calculates the maximum transmit power for each user. Depending on the location of the user controller can easily get the other medical devices in the same location as these objects are connected by location foreign keys. We propose the minimal database structure for the cognitive radio system operate properly.

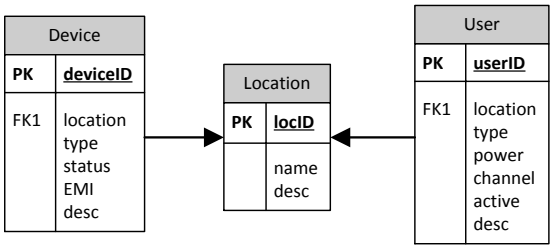


Fig. 3. Database structure for Cognitive Radio System

V. CONCLUSIONS

In this article the main problem of interference between multiple wireless technologies in 2.4 GHz ISM band was presented. The CSMA/CA RTS/CTS and EMI-aware RTS/CTS protocols were described but only the first one was implemented in our simulation for now. The possible database structure was proposed to work under the cognitive radio system. This work is still ongoing. In our future work we will implement CR algorithm and apply multi-channel scenario described in this article. The expectations are that CR allows for better use of the electromagnetic spectrum.

ACKNOWLEDGMENT

This work was supported in part by COST Action IC0905 “Techno-Economic Regulatory framework for Radio spectrum Access for Cognitive Radio/Software Defined Radio” (TERRA) (www.cost-terra.org). Additional funding was provided by ERASMUS.

REFERENCES

[1] ZigBee and Wireless Frequency Coexistances, ZigBee White Paper, June 2007, [online] Available: http://docs.zigbee.org/zigbee-docs/dcn/07-5219.PDF

[2] G. Thonet, P. Allard-Jacquin, and P. Colle. ZigBee - WiFi Coexistence: White Paper and Test Report. Technical report, Schneider Electrics, 2008.

[3] Daintree Networks, "Getting Started with ZigBee and IEEE 802.15.4", [online] Available: http://www.daintree.net/downloads/whitepapers/zigbee_primer.pdf

[4] P. Phunchongharn, E. Hosssain, D. Niyato, and S. Camorlinga, "A cognitive radio system for e-health applications in a hospital environment," IEEE Wireless Communications Magazine, pp. 20-28, Feburary 2010.

[5] MICAz-Based ZigBee and WiFi Coexistence, Avoiding RF Interference Between WiFi and Zigbee, [online] Available:

http://www.willow.co.uk/ZigBeeandWiFiInterference.pdf

[6] T.S. Ho and K.C. Chen, Performance analysis of IEEE 802.11 CSMA/CA medium access control protocol, in: *Proc. PIMRC*, Taipei, Taiwan (October 1996) pp. 407–411.

[7] Rehmani, Mubashir Husain, Stéphane Lohier, and Abderrezak Rachedi. "Channel bonding in cognitive radio wireless sensor networks." *Mobile and Wireless Networking (iCOST), 2012 International Conference on Selected Topics in*. IEEE, 2012.

[8] Raman, Vijay, and Nitin H. Vaidya. "Adjacent channel interference reduction in multichannel wireless networks using intelligent channel allocation." *Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Tech. Rep* (2009).

INFORMACINĖS TECHNOLOGIJOS ● IVUS 2013 ● ISSN 2029-249X ● eISSN 2029-4824

Elektroninio dienyno įvedimo mokyklose įtakos organizacinių procesų efektyvumui tyrimas

Saulius Kuzminskis
Taikomosios informatikos katedra
Vytauto Didžiojo universitetas
Kaunas, Lietuva
saulius.kuzminskis@live.vdu.lt

Vladislav V. Fomin
Taikomosios informatikos katedra
Vytauto Didžiojo universitetas
Kaunas, Lietuva
v.fomin@if.vdu.lt

Santrauka. ***Vis daugiau procesų perkeliama į elektroninę erdvę − tokia tendencija neaplenkė ir Lietuvos mokyklų. Šiame straipsnyje pristatoma problema apžvelgiant situaciją vienoje Kauno gimnazijų, kur remiantis kitų mokyklų praktika, Europos Sąjungos nuorodomis ir Lietuvos Respublikos įstatymais, buvo įdiegtas elektroninis dienynas. Skirtingai nei elektroninės komercijos atveju, skaitmenizavus dienyną, tiesioginę įtaką organizacinių procesų efektyvumui išmatuoti yra sunku tiek dėl ankstesnių tyrimų trūkumo, tiek ir trūkstant visuotinai priimtinos praktikos. Praeitais metais pasiūlytas efektyvumo tyrimo būdas šiame darbe analizuojamas, remiantis Harold Leavitt bei James Gaskin modeliais, aiškinančiais technologijos įtaką kitiems organizacijos komponentams, šią įtaką matuojant našumo, produktyvumo, kokybės ir nuoseklumo dimensijose. Vėliau pristatomi moksliniai tyrimo metodai, tarnausiantys duomenų rinkimui, analizei ir rezultatų apibendrinimui.***

Reikšminiai žodžiai—elektroninis dienynas, švietimas, mokykla, efektyvumas, informacinė visuomenė.

I. ĮŽANGA

Šiais laikais vis daugiau sistemų perkeliamos į elektroninę erdvę – internetu balsuojama, perkami bilietai į renginius, pildomos deklaracijos ir kiti dokumentai. E-komercija, e-sveikata, e-valdžia ir daugelis kitų pavyzdžių. E-komercijos atveju nauda aiški – pardavinėjant internete sutaupomi kaštai, nereikia nuomotis patalpų, sutrumpėja logistinė grandinė, todėl elektroninėje erdvėje pateikiamos mažesnės kainos nei realiose parduotuvėse. Gautas rezultatas - pritraukti klientai ir padidėjęs pelnas. Tačiau kaip pamatuoti elektroninio dienyno naudą mokyklai? Šiame darbe siekiama atsakyti į šį klausimą, detaliai aprašant probleminę sritį, bei atliekant jau padarytų tyrimų apžvalgą, kuri padės surasti tinkamą modelį (-ius) efektyvumo matavimui.

Lietuvoje e-dienynų platinimo pradžioje buvo teigiama, jog jie padės mažinti biurokratizmą, vystyti kokybiškesnį bendradarbiavimą su tėvais, mokytojams sumažės darbo krūvis, pagerės mokinių pamokų lankomumas [1]. Įvedus e-dienyną, kardinalius pokyčius pastebėti sunku - mokiniai gauna tuos pačius pažymius, mokytojai tuos pačius atlyginimus, todėl svarbu išsiaiškinti ar anksčiau deklaruotos tiesos iš tiesų yra teisybė, bei kokią realią efektyvumo įtaką mokyklos bendruomenėje ir jose vykstantiems procesams daro ši technologija. Atlikus paiešką Lietuvos Elektroninių tezių ir disertacijų duomenų bazėje[1], randamas tik vienas darbas, susijęs su e-dienyno efektyvumu organizacijoje vykstantiems procesams – Giedrės Česonytės praeitais metais atliktas tyrimas vienoje Vilniaus mokyklų. Vis dar nėra oficialiai priimtos metodikos, kaip reiktų atlikti tokį tyrimą, todėl bus siekiama pakartoti praeitais metais atliktą mokslinį darbą, panaudojant pateiktą modelį kaip tyrimo pagrindą Kauno gimnazijoje, ir išsiaiškinti, ar jis yra tinkamas praktikoje, bei esant reikalui atlikti patobulinimus.

II. PROBLEMOS APRAŠYMAS

Mokyklose dažniausiai sutinkami organizaciniai procesai, kuriems didžiausią įtaką daro elektroninio dienyno įvedimas.

1. *Mokinių lankomumas*: suvedamas mokinių lankomumas - tėvai, auklėtojai, mokyklos vyresnybė jį patikrina.
2. *Pažymių surašymas*: surašomi mokinių pažymiai, vėliau išvedami pusmečiai.
3. *Ugdymo planų rengimas*: klasės vadovas parengia ugdymo planus, pagal kuriuos organizuojamas mokinių užimtumas.
4. *Tvarkaraščių sudarymas*: mokytojai sudaro mokinių tvarkaraščius, mokiniai juos pasitikrina.

Europos Sąjungos Komisijos prioritetas – Europe 2020[2] strategija ateinančiam dešimtmečiui, kurios vienas pagrindinių tikslų yra tęsti anksčiau įvestų priemonių, skirtų švietimo sektoriaus augimui, išnaudojant IKT (informacinės ir komunikacinės technologijos) galimybes bei naudojimą, taip skatinant piliečių pažangumą ir integraciją informacinės visuomenės amžiuje[3]. Informacinės visuomenės kūrimo strategijose[4] ne kartą minimas e-pasiekiamumo poreikis

[1] Lietuvos elektroninės tezės ir disertacijos (ETD duomenų bazė) http://goo.gl/yjMsn
[2] "Europe 2020 - Europe's growth strategy" - European Commission. http://ec.europa.eu/europe2020/index_en.htm
[3] Implementation of the European strategic objectives in Education and Training (ET 2020) http://eacea.ec.europa.eu/llp/funding/2012/call_et_2012_en.php
[4] Digital Strategy, i2010 Strategy, eEurope Action Plan, Digital Strategy Programmes http://europa.eu/legislation_summaries/information_society/strategies/index_en.htm

įvairiose srityse, tarp jų ir mokyklose, kaip pavyzdį minint e-dienyno diegimą[5]. Teigiama, jog e-pasiekiamumo įrankiai skatina platesnį interneto ir e-paslaugų naudojimą, taigi ir žmonių įgūdžių naudojant e-paslaugas gerinimą. Šie įgūdžiai tikimasi, prisidėtų prie technologinių procesų spartesnės kaitos, modernizacijos, verslų konkurencingumo, struktūrinių ekonomikos pokyčių bei jos augimo, kartu padidinant ir atskirų organizacijų efektyvumą.

Lietuvoje elektroninio dienyno tvarką reguliuojantis įstatymas yra Lietuvos Respublikos Švietimo ir mokslo ministro įsakymu patvirtintas *pradinio ugdymo, pagrindinio ugdymo, vidurinio ugdymo ir vidurinio ugdymo mokytojų dienynų sudarymo elektroninio dienyno duomenų pagrindu tvarkos aprašas*, kurio nuostatomis turi vadovautis visos sprendimą naudotis elektroniniu dienynu priėmusios mokyklos [2].

Įvardintas įsakymas leidžia daryti kelias išvadas:

1. Elektroninio dienyno naudojimas mokyklose yra neprivalomas, pasirinktinis. Taigi, mokyklos nėra priverstos diegti šią technologiją ir mokyklos tarybos nutarimu galima toliau naudoti popierinį dienyną.
2. Elektroninio dienyno struktūra ir sudarymo principai turi būti identiški popieriniam dienynui, todėl mokykla gali neabejoti, jog galės atlikti visas užduotis, kaip ir naudodama popierinį dienyną.
3. Įdiegus elektroninį dienyną, galima atsisakyti popierinio dienyno, taip teoriškai sumažinant papildomo darbo krūvį, popieriaus sunaudojimą bei dienynų skaičių mokykloje. Visgi net atsisakius popierinio dienyno, visiškai panaikinti popieriaus sunaudojimo nepavyks, kadangi įstatyme reikalaujama išspausdinti tam tikras dienyno dalis, o mokyklai pasirinkus išspausdinti jas visas (kadangi tas dalis, kurių nereikalaujama spausdinti galima saugoti ir išorinėse laikmenose), popieriaus sunaudojimas ne sumažinamas, o greičiau padidinamas, taip išaugant piniginiams kaštams (abonentinis mokestis už elektroninio dienyno naudojimą, popieriaus ir rašalo kaina, prieštaravimas Europos Sąjungos „žalios“ ekonomikos siekiams[6]). Tokiu atveju mokyklai veikiausiai pigiau yra išlaikyti popierinį dienyną.

Apibendrinant pateiktus Europos Sąjungos bei Lietuvos Respublikos teisės aktus, iš jų išplaukiančias išvadas, galima teigti, jog mokykla apsispręsti diegti e-dienyną gali dėl kelių priežasčių:

politinių – kuomet remiantis įstatymais, valstybėje gera praktika yra laikomas elektroninio diegimas ir naudojimas, taip skatinant mokyklas diegti e-dienynus (tokia situacija yra Lietuvoje, Švietimo ir mokslo ministerijos duomenimis[7] 80% mokyklų naudoja e-dienyną),

praktinių, kuomet mokykla įžvelgia galimą kaštų mažinimą – greitesnį darbų atlikimą, sumažėjusį popieriaus sunaudojimą, kas, remiantis įstatymu ir dažna praktika, yra praktiškai neįmanoma – popierinio varianto vis tiek reikia.

Kitas mokykloje vykstančių procesų efektyvumo padidėjimo scenarijus – kuomet mokykla pasiekia tuos pačius ar geresnius rezultatus kaip ir naudodama popierinį dienyną, tačiau daro tai greičiau ir efektyviau atsisakydama nereikalingų procesų ar pagerindama jų atlikimą (mažiau pastangų ir laiko), bei jausdama teigiamą įtaką mokyklos bendruomenei.

III. Problemos Apibrėžimas

Kauno „Varpo“ gimnazijos vadovybė savo noru priėmė sprendimą įdiegti elektroninį dienyną 2010 metais, kuris funkcionalumu visiškai atstoja popierinį dienyną. Tai padariusi, turi laikytis LR įstatymų, kuriais remdamasi nutarė, jog duomenų saugojimo tikslais ir toliau spausdins didžiąją dalį dienyno, nepasinaudodama galimybe sumažinti sunaudojamo popieriaus apimtis, taip neišnaudodama privalumų, kuriuos suteikia informacinių technologijų naudojimas.

Toliau pateikiami pastebėjimai apie elektroninį dienyną, naudotą gimnazijoje 2010-2012 metais. *Pasviruoju šriftu* apibendrinamos naudojimosi metu kilusios problemos.

Stebint mokytojų naudojimąsi elektroniniu dienynu išsiaiškinta, jog didžiausią nepasitenkinimą jiems kėlė ilgas krovimosi laikas, neaiški vartotojo sąsaja, negalėjimas prisijungti prie sistemos.

Mokiniai buvo nepatenkinti naudojamu dienynu, daugelio nuomone tai yra „visiška nesąmonė“. *Administracija* ir *mokytojai* kaip vieną svarbiausių neigiamų savybių taip pat paminėjo nuolatinį strigimą ir ilgą krovimosi laiką, ypač pusmečio pabaigoje, kuomet visos mokyklos *veda* pusmečius.

Galima suformuluoti tarpinę problemą nr. 1 – Kauno „Varpo“ gimnazijoje įdiegtas dienynas nepateisino vadovybės ir visos organizacijos lūkesčių.

Vietoje popierinio dienyno pildymo, pažymių suvedimą ir kitas užduotis reikia atlikti virtualioje erdvėje, o tam reikia turėti darbo kompiuteriu įgūdžių (keičiasi žmonių įpročiai), išmokti naudotis nauja sistema (atsiranda papildomas krūvis). Atsirado naujų pareigybių, susijusių su dienyno priežiūra ir duomenų saugojimu, pasikeitė bendravimas su tėvais (struktūros pakitimai). Vienas pagrindinių dienyno pasirinkimo motyvų buvo jo kaina, kadangi naudota Nacionalinė švietimo procesų valdymo sistema „Žinios“ [4] buvo nemokama.

Galima daryti išvadą, jog tinkamai neįvertinta sudėtinga organizacinė struktūra, skirtingas žmonių pasirengimas naudoti IT ir kompiuterinis raštingumas (tarpinė problema nr.2), nes kaip teigiama H. Leavitt "Deimanto modelyje" [3], organizacijoje dirbantys žmonės, jos struktūra, vykstantys procesai ir nauja technologija yra glaudžiai tarpusavyje susiję.

[5] Key Data on Learning and Innovation through ICT at School in Europe 2011 http://eacea.ec.europa.eu/education/eurydice/documents/key_data_series/129EN.pdf

[6] European Commision. Europe 2020 http://goo.gl/PQW9u

[7] Švietimo ir mokslo ministerija http://www.smm.lt/naujienos/pranesimai.htm?id=4061

Išvardintas problemas galima suglausti į vieną - Kauno „Varpo“ gimnazija, atsižvelgdama į kainą, o ne į funkcionalumą ir vartotojų poreikių atitikimą, tinkamai neįvertinusi sudėtingos organizacinės struktūros, skirtingo žmonių pasirengimo naudoti IT ir kompiuterinio raštingumo, įdiegė nemokamą dienyną, kuris nepateisino vadovybės ir visos organizacijos lūkesčių.

Taip pat nepasitenkinimas visose mokyklos bendruomenės grupėse (mokiniai, tėvai, mokytojai, administracija) lėmė įvykių pabaigą – mokykla atsisakė lūkesčių nepateisinančio dienyno ir perėjo prie mokamos alternatyvos – „Tamo“ dienyno [5] tikėdamasi, jog tai padės išspręsti problemas. Iš tiesų buvusį dienyną kamavusios techninės problemos išnyko – sistema nestringa, nereikia ilgai laukti, kol užsikraus norimas langas ir bus galima įvykdyti užduotis – tai žymiai sumažino bendrą nepasitenkinimą. Tačiau kai kuriais atžvilgiais kritikos strėlė toliau sminga ir į naują sistemą. Techninės problemos išspręstos, tad kyla klausimas − jeigu tai ne jos, tuomet kas gi įtakoja šį nepasitenkinimą?

Jeigu e-dienyno diegimas, kaip parodyta problemos aprašyme, nepadeda sumažinti piniginių kaštų ar juos net padidina, svarbu susikoncentruoti į mokykloje vykstančių organizacinių procesų efektyvumą. Nepatenkinti naudotojai yra vienas iš patikimų ženklų, rodančių sistemos neefektyvumą. Todėl svarbu išsiaiškinti, kokios to priežastys, atlikus e-dienyno efektyvumo mokyklos organizaciniams procesams tyrimą. Norint jį atlikti, reikia dar labiau įsigilinti į buvusio dienyno situaciją ir surasti tikrąsias priežastis. Reikia gydyti ligos priežastis, o ne stengtis sumažinti jos simptomus. Todėl apibendrinus situacijos aprašymą ir išdėstytas problemas, galima teigti, jog *pagrindinė problema* yra trūkumas tyrimų, kuriuose būtų pakankamai išanalizuota, kaip elektroninio dienyno įdiegimas įtakoja organizacijoje dirbančių žmonių darbo ir joje vykstančių procesų efektyvumą. Trūksta vieningo modelio, galinčio padėti mokykloms priimti tinkamą sprendimą ir išvengti panašių padarinių, iškilusių Kauno “Varpo“ gimnazijoje, bei apskritai pateisinti arba paneigti visą Lietuvos mokyklų e-modernizacijos procesą elektroniniais dienynais.

IV. Literatūros Apžvalga

Harold Leavitt [3] pasiūlė būdą, dabar vadinamą „Leavitt‘s diamond“, organizacijų efektyvumo vertinimui, kuriame teigiama, jog kiekvienoje organizacijoje yra keturios tarpusavyje susijusios dimensijos:

1. žmonės, 2. procesai, 3. organizacinė struktūra, 4. technologija.

Organizacijos sėkmingą vystymąsi nulemia šių keturių komponentų tarpusavio ryšys. Autorius teigia, jog pakitimai vienoje dimensijoje įtakoja pakitimus kitose.

Toliau pateikiami dimensijų paaiškinimai kartu su pavyzdžiais, rodančiais kaip yra įtakojama aprašoma dimensija, keičiantis kitiems komponentams.

Dimensija *Žmonės*. Kalbant apie žmones vertinamos ne pareigybės (buhalteris ar vadybininkas), o sugebėjimai, atliekamo darbo efektyvumas, žinios ir produktyvumas visos organizacijos kontekste. *Pavyzdys*: kinta komponentas Procesai - atsiranda nauja užduotis. Reikia išmokyti žmones ją tinkamai įvykdyti.

Dimensija *Procesai*. Dėmesys kreipiamas į tai, kaip atliekamos užduotys ir ko norima pasiekti. Čia kaip ir prieš tai buvusioje dimensijoje svarbu ne užduoties ar siekio apibrėžimas, bet kokybiniai jų aspektai – kaip juos atliekant teigiamai įtakojama organizacijos veikla. *Pavyzdys*. Keičiantis darbuotojams reikia modifikuoti užduotis, kad optimaliai būtų išnaudojamos naujų darbuotojų žinios ir galimybės.

Dimensija *Struktūra*. Tai ne tik organizacinė struktūra, bet ir santykiai tarp darbuotojų, tarpusavio komunikacija ir skirtingų organizacijos lygių darni veikla. *Pavyzdys*: į darbą priimant puikiai kvalifikuotus darbuotojus, organizacinėje struktūroje galima atsisakyti šiems žmonėms vadovaujančių, prižiūrinčių pareigybių, kurios buvo reikalingos dirbant mažiau kvalifikuotiems darbuotojams.

Dimensija *Technologija*. Komponentas, kuris padeda organizacijos darbuotojams atlikti užduotis. *Pavyzdys*: įvedus naują technologiją, reikalingi apmokymai darbuotojams (keičiasi žmonės), išnaudojant technologijos privalumus galima atsisakyti kai kurių darbo vietų ar sukurti naujas (keičiasi struktūra), o natūralus noras išnaudoti naujas technines galimybes ir pasiekti geresnių rezultatų reikš tikslų ir būdų (keičiasi procesai) jiems pasiekti pasikeitimą.

Surinkus duomenis apie Kauno „Varpo“ gimnaziją ir kitas įstaigas, patvirtintas Leavitt modelio teisingumas: jose įvedus naują technologiją, atsirado naujos pareigybės (struktūra) bei užduotys (procesai) susijusios su technologijos naudojimu, žmonėms buvo rengiami mokymai.

Interpretuojant tyrimų metu surinktą medžiagą ir panaudojant Leavitt teorinius teiginius galima daryti išvadą, jog tinkamai neįvertinus sudėtingos organizacinės struktūros, kiekviename iš skirtingų organizacinių lygių vykstančių užduočių ir jų tarpusavio ryšių, žmonių techninio pasirengimo ir žinių lygio, galimas dalinis arba visiškas technologijos nepasiteisinimas, patiriami psichologiniai bei materialiniai praradimai, suardant visą organizacijos veiklą.

Vis dėlto šios išvados nepadeda surasti atsakymo į klausimą − kaip e-dienynas veikia mokyklos organizacinių procesų efektyvumą. Veikiausiai tik padeda suprasti, kodėl Kauno „Varpo“ gimnazija atsisakė buvusio dienyno.

Norint surasti tinkamą teorinį modelį šiam tyrimui, buvo atlikta paieška elektroninių tezių duomenų bazėje. Rastas vienas darbas[8], nagrinėjantis panašią problemą. Būdas efektyvumui tirti autorės išbandytas vienoje sostinės mokyklų, paremtas James Gaskin teoriniu modeliu[9], kuriame teigiama, jog technologijos efektyvumą organizacijoje galima matuoti keturiomis dimensijomis – produktyvumo, našumo, kokybės ir nuoseklumo. Jos buvo matuojamos respondentams pateikiant klausimynus.

[8] G. Česonytė, „Elektroninio dienyno vaidmuo mokykloje. "Technologija-įstaiga" sandaros efektyvumo tyrimas“, Vytauto Didžiojo universitetas, Kaunas, 2012

[9] J. Gaskin, Opening the black box of system use: the impact of the individual usage behaviours on performance. Working paper. The Weatherhead School of Management, Case Western Reserve University, Cleveland, OH, 2011.

Giedrės Česonytės darbas yra tvirtas pagrindas situacijos Kauno „Varpo“ gimnazijoje vertinimui, tačiau jame pastebėti tam tikri netikslumai, galintys nulemti iškreiptus ir netikslius atsakymus, vedančius į galimai nepilnai teisingus rezultatus. Šiuo darbu bus siekiama patobulinti pateiktą klausimyną, bei patvirtinti ar paneigti teorijos teisingumą, remiantis James Gaskin modeliu, kuriame kaip jau minėta minimos keturios efektyvumo dimensijos [6]:

Produktyvumas – matuojama, kiek užduočių per tą patį laiko tarpą atlieka darbuotojai naudojantys naują technologiją.

Našumas – atsako į klausimą kiek laiko, žmonių pastangų ir kitų išteklių yra sutaupoma naudojant naują technologiją.

Kokybė – leidžia apibūdinti padaromų klaidų, atliekant užduotis, sumažėjimą (ar padidėjimą), gautų rezultatų patikimumą.

Nuoseklumas – parodo, jog darbuotojai, naudojantys technologiją ir pastoviai atliekantys tas pačias užduotis, turėtų būti patikimai vienodai rezultatyvūs savo darbe

Remiantis anksčiau atliktu Giedrės Česonytės darbu [7], bei Leavitto [3] ir Gaskino [6] modeliais galima daryti išvadą, jog įmanoma apibūdinti (aprašyti) organizacijoje esantį (nesantį) efektyvumo padidėjimą (sumažėjimą) įvedus naują technologiją, atsižvelgiant į tai, jog technologija yra tampriai susijusi su (ir juos veikia) ten dirbančiais žmonėmis, vykstančiais procesais ir esama struktūra, bei pati yra šių komponentų veikiama. Šį poveikį efektyvumo prasme galima išmatuoti keturiomis dimensijomis: produktyvumo, našumo, kokybės ir nuoseklumo.

V. Moksliniai Tyrimo Metodai

Moksliniai tyrimo metodai, pasirinkti duomenų rinkimui, bei analizei, nesiskirs nuo ankstesniame darbe parinktų [7], norint įsitikinti ir patikrinti tų metodų tinkamumą naudoti kitoje aplinkoje, taip įrodant ar paneigiant teorijos teisingumą ne tik vienai, bet daugeliui organizacijų.

Pasirinkti kokybiniai empiriniai tyrimo metodai, Kadangi tyrimo objektas bus stebimas ir analizuojamas jo natūralioje aplinkoje [8].

Vienas iš kokybinių tyrimo metodų, tarnausiantis kaip „gaubiamasis“ ir apibendrinantis darbo įrankis kitiems metodams ir gautiems rezultatams yra atvejų analizė.

Atvejų analizės rezultatai gali turėti labai didelę įtaką tyrimui. Nesuvaržytas griežtų klausimynų ar mokslinių metodų apribojimų, toks tyrimas gali vesti link naujų kūrybiškų įžvalgų, naujos teorijos sukūrimo ir specialistų (bene svarbiausias tokio tyrimo naudotojas) noro pritaikyti rezultatus praktikoje. Šiuo būdu naujos teorijos yra sukuriamos ne toli nuo tiriamo objektų esančių akademikų, bet šalia dirbančių, turinčių artimą kontaktą su objektu žmonių [9].

Atvejų analizė yra tinkamas metodas šiam tyrimui, kadangi pagal apibrėžimą jos metu analizuojama viena ar kelios esybės (žmogus, grupė ar organizacija), pasirinkto vieneto sudėtingumas yra analizuojamas intensyviai, dėmesys kreipiamas į šiuolaikinius įvykius, fenomenas tiriamas natūralioje aplinkoje, nėra eksperimentinio valdymo ar manipuliacijos, o esant reikalui, kuomet tyrėjas sukuria naujas hipotezes, galimi tyrimo vietos ar duomenų rinkimo metodų pakeitimai, tyrėjas klausia „kaip“ ir „kodėl“ [10].

A. Kokybiniais metodais paremtas duomenų rinkimas

Duomenys renkami šiais tyrimo metodais [8]:

Stebėjimas – vienas iš pasirinktų kokybinių metodų duomenims surinkti. Kadangi straipsnio autorius pats dirba tiriamo objekto aplinkoje, todėl gali „iš vidaus“ stebėti kitų asmenų darbą su dienynu natūralioje aplinkoje, be jokios papildomos įtakos jų naudojimuisi sistema.

Apklausa (anketinė apklausa) - apklausos dažniausiai atliekamos standartizuotomis metodikomis, vykdoma vienpusė komunikacija, kurią valdo interviu gavėjas.

Apklausa (Interviu) viena iš apklausos rūšių, apibrėžiama kaip dviejų asmenų pokalbis, kurios tikslas – gauti būtiną tyrimo uždaviniams informaciją. Pasirinktas nestruktūrizuotas interviu tipas, kuomet klausinėjama laisva forma, situacija atvira, apklausiamasis jaučiasi patogiai.

VI. Išvados ir Perspektyvos

Straipsnyje pristatyta įvairių sričių skaitmenizavimo problema: ten kur tiesioginė nauda perkeliant veiklas į elektroninę erdvę yra tiesiogiai pastebima e-komercijoje, kitokia situacija yra švietimo įstaigose. Beveik visose Lietuvos mokyklose įvedus e-dienyną vis dar nėra daug tyrimų, kurie paaiškintų šios technologijos įtaką mokyklose vykstančių procesų ir jos bendruomenės efektyvumui, nors tokie teiginiai buvo plačiai skelbiami.

Tyrimų svarbą ir modelio reikalingumą gali iliustruoti vienos Kauno gimnazijos pavyzdys, kur pirminis įdiegtas dienynas nepateisino bendruomenės lūkesčių ir buvo pakeistas nauju, patiriant materialinius ir nematerialinius nuostolius.

Galimos šio reiškinio priežastys, aiškinamos straipsnyje pristatytu Leavitt modeliu, teigiančiu, jog keturi komponentai yra tiesiogiai susiję ir veikia vienas kitą kiekvienoje organizacijoje: technologija, struktūra, žmonės, procesai. Remiantis atliktu preliminariu tyrimu, priimtos išvados, jog organizacijai tinkamai neįvertinus sudėtingų šių komponentų tarpusavio ryšių, neįmanomas efektyvus technologijos įdiegimas, o tai gali grėsti destruktyviais organizacijai padariniais.

Vienintelis e-dienyno efektyvumo tyrimas Lietuvoje, kuris buvo atliktas praeitais metais rėmėsi James Gaskin modeliu, kuriame yra įrodyta, jog žmonių ir aplamai visos organizacijos efektyvumas naudojant informacines technologijas gali būti matuojamos keturiomis dimensijomis: produktyvumo, našumo, kokybės ir nuoseklumo.

A. Perspektyvos

Išanalizavus ankstesniame tyrime pateiktą klausimyną, sudarytą remiantis James Gaskin dimensijomis, ir jame pastebėjus tam tikrų netikslumų, iškilo poreikis jį pakoreguoti. Tikimasi gauti tikslesnius, bei patikimesnius rezultatus.

Modifikuotas modelis (klausimynas) pritaikytas minėtoje Kauno gimnazijoje, kuri po nesėkmingo bandymo įsidiegė naują dienyną. Surinkti rezultatai bus analizuojami, o pateikus išvadas, tikimąsi patvirtinti arba paneigti modelio tinkamumą mokykloms. Kadangi tai jau būtų antra mokykla, kurioje pritaikoma anksčiau iškelta teorija (atliktas teorijos testavimas), galima būtų daryti išvadas apie modelio teisingumą, o pasiūlytą klausimyną naudoti plačiu mastu, padedant išvengti nepavykusių mėginimų įdiegti technologijas, suprasti jų naudą ir svarbą ar atvirkščiai nenaudą.

LITERATŪROS SĄRAŠAS

[1] Pasitarimas dėl elektroninio dienyno tvarkymo mokykloje Švietimo ir mokslo ministerija. Prieiga internetu: http://www.smm.lt/ugdymas/docs/formos/ed/Pasitarimas%202010%2001%2021.pdf

[2] Dienynų sudarymą elektroninio dienyno duomenų pagrindu reglamentuojantys teisės aktai. Prieiga internetu: http://www.smm.lt/ugdymas/docs/formos/ed/Dienynu%20sudarymas.pdf

[3] H.J. Leavitt, "Diamond Model." In Handbook of Organizations, edited by J. G. March. Chicago, IL: Rand McNally & Co, 1965.

[4] Nacionalinė švietimo procesų valdymo sistema „Žinios". Prieiga internetu: http://www.dienynas.lt

[5] Informacinė sistema švietimo įstaigoms „Tamo". Prieiga internetu: http://www.tamo.lt

[6] J. Gaskin, Opening the black box of system use: the impact of the individual usage behaviours on performance. Working paper. The Weatherhead School of Management, Case Western Reserve University, Cleveland, OH, 2011.

[7] G. Česonytė, „Elektroninio dienyno vaidmuo mokykloje. "Technologija-įstaiga" sandaros efektyvumo tyrimas", Vytauto Didžiojo universitetas, Kaunas, 2012

[8] K. Kardelis, Mokslinių tyrimų metodologija ir metodai. 2 leidimas. Šiauliai, 2002.

[9] R. Dul, T. Hak. Case Study Methodology in Business Research. Elsevier, Ltd., 2008.

[10] I. Benbasat, D. K. Goldstein, M. Mead. „The case research strategy in studies of information systems". MIS Quarterly, 11 Nr. 3: 369—386, 1987

Application of Artificial Neural Networks to Determine Concentrations of Mixture

Linas Litvinas
Faculty of Mathematics and Informatics
Vilnius University
Vilnius, Lithuania
linas.litvinas@mif.stud.vu.lt

***Abstract*—A biosensor is capable to measure the concentration of a substrate. In this case, the linear dependency between biosensors signal at steady state and substrate concentration is used. In case of multiple substrates, such simple method does not work. In this case, the nonlinear dependency between biosensor response and substrate concentrations exists. As this dependency cannot be determined analytically, it can be approximated by artificial neural networks. The goal of this paper is to investigate the possibility to use a single amperometric biosensor and artificial neural networks to determine multiple substrates concentrations. For that purpose mathematical model of amperometric biosensor was used to generate experimental data. The principal component analysis was applied to compress the data. Finally artificial neural networks were applied and the results were compared to an optimization based method. As artificial neural networks gave sufficiently accurate results, the direction for the future research was given.**

***Keywords*—modeling; biosensor; concentration measurement; principal components analysis; artificial neural networks**

I. Introduction

Biosensor is a device composed of a biologically active component with signal transducer and an amplifier [1]. In biologically active component an enzymatic reaction occurs and the measured substrate becomes to product. The product of this reaction is converted to a signal by a transducer. The strength of the signal is proportional to concentration of a measured substrate (or a group of them). In case of an amperometric biosensor the current on an indicator electrode is measured (due to direct oxidation or reduction of the products of the biochemical reaction), while keeping the potential on an electrode constant. The amperometric biosensors are relatively cheap, sensitive and reliable devices for environment, clinical and industrial purposes [9-12].

Mathematical models of biosensor action are widely used in the design phase - to get a biosensor of needed characteristics and parameters [13-18]. Detailed review of modeling an amperometric biosensor is presented in [6, 7].

Number of papers are devoted for investigating possibility to determine multiple substrate concentration using a single amperometric biosensor [3, 4, 8]. In [3, 4] use an artificial neural network to classify multiple substrate concentrations to fixed predefined categories. In practice, such classification is irrelevant, because substrate concentrations in mixture vary in continuous space. In [8] use optimization of error function, to determine substrate concentrations in continuous space. Error function in such case is Euclidean norm between an actual biosensor response and the response combined of separate responses having only a single substrate. In the minimum of Error function, the determined values are concentrations of single substrate. In this paper [8] shown that concentrations of multiple substrate can be determined by using a single biosensor. On the other hand, the method proposed by [8] requires computing resources for optimization, so embedded system would be complex and expensive. Meanwhile, artificial neural networks are simple and cheap as it requires less computing resources. Consequently the possibility to use neural networks to determine multiple substrates concentrations was investigated. This possibility is not trivial as for some substrates the contribution to final biosensor response is negligible [8].

This paper as [3, 4, 8] use biosensor simulation data. Simulations are based on mathematical model of amperometric biosensor. The simulation data was compressed by principal component analysis as this method gave good results in case of classification [3, 4]. Finally artificial neural networks were applied to determine multiple substrates concentrations and accuracy was analyzed. The results were compared to optimization based method [8]. In compare it gave slightly worse, but sufficiently accurate results.

II. Mathematical Model

Consider an enzyme-catalysed reaction:

$$S_k \xrightarrow{E_k} P_k, \quad k = 1, \dots, K, \tag{1}$$

where substrate (S_k) becomes to product (P_k) using enzymes (E_k), K is number of substrates in mixture. Reaction rate in this case is the rate of a product appearance, which is known to depend on the concentration of substrate. An amperometric biosensor can be treated as an enzyme electrode. Enzyme layer is put into the surface of probe. Lets assume these conditions are satisfied: no interaction between analysed substrates of the mixture, symmetrical geometry of the electrode, homogeneous distribution of enzyme in the enzyme membrane, mixture is well stirred, relative thin enzyme layer. Considering one-dimensional diffusion described by Fick's law and coupling with enzyme reaction leads to the following equations [2, 3]:

This work was partly supported by project "Promotion of Student Scientific Activities" (VP1-3.1-ŠMM-01-V-02-003) from the Research Council of Lithuania

$$\frac{\partial S^{(k)}}{\partial t} = D_{\mathrm{S}}^{(k)} \frac{\partial^2 S^{(k)}}{\partial x^2} - \frac{V_{max}^{(k)} S^{(k)}}{K_M^{(k)} + S^{(k)}}, \quad (2)$$

$$\frac{\partial P^{(k)}}{\partial t} = D_{\mathrm{P}}^{(k)} \frac{\partial^2 P^{(k)}}{\partial x^2} + \frac{V_{max}^{(k)} S^{(k)}}{K_M^{(k)} + S^{(k)}}, \quad (3)$$

$$0 < x < d, \qquad 0 < t \le T, \qquad k = 1, \dots, K,$$

where K - number of considered substrates. $V_{max}^{(k)}$ - maximal enzymatic rate reachable with that amount of enzyme, when the enzyme is fully saturated with substrate S_k. $K_M^{(k)}$ - the Michaelis constant. $S^{(k)}$ – concentration of substrate S_k. $P^{(k)}$ – concentration of product P_k. d - thickness of enzyme layer. t - time, T - duration of biosensor operation. $D_{\mathrm{S}}^{(k)}$ and $D_{\mathrm{P}}^{(k)}$ are diffusion coefficients of the substrate S_k and product P_k respectively. $k = 1, \dots, K$ [1, 5, 7].

The biosensor start to operate when some substrate appears over the surface of the enzyme layer. It can be expressed in the following initial conditions ($t = 0$):

$$S^{(k)}(x, 0) = \begin{cases} 0, & 0 \le x < d, \\ S_0^{(k)}, & x = d, \end{cases} \quad (4)$$

$$P^{(k)}(x, 0) = 0, \quad 0 \le x \le d, \quad (5)$$

$$k = 1, \dots, K,$$

where $S_0^{(k)}$ – substrate S_k concentration over the surface of biosensor.

The biosensor operates in injection mode (as this mode more suited for determination of concentrations of multiple substrates [3, 4, 8]), so substrate appears over the surface of the enzyme layer for short period of time (T_M). This fact can be expressed by these boundary conditions ($0 < t \le T$):

$$\left.\frac{\partial S^{(k)}}{\partial x}\right|_{x=0} = 0, \quad (6)$$

$$S^{(k)}(d, t) = \begin{cases} 0, & t > T_M, \\ S_0^{(k)}, & t \le T_M, \end{cases} \quad (7)$$

$$P^{(k)}(0, t) = P^{(k)}(d, t) = 0, \quad (8)$$

$$k = 1, \dots, K.$$

The result of a physical experiment is a biosensor response – the current measured at predefined time values. The current is proportional to the amount of product appeared on the surface of electrode ($x = 0$). Consequently, current density $I^{(k)}(t)$ can be computed using Faraday's law:

$$I^{(k)}(t) = n_e F D_{\mathrm{P}}^{(k)} \left.\frac{\partial P^{(k)}}{\partial x}\right|_{x=0}, \quad k = 1, \dots, K, \quad (9)$$

where n_e - number of electrons involved in the charge transfer, F - Faraday constant. Total biosensor response to a mixture represents a sum of individual responses $I^{(k)}(t)$ to each constituent substrate:

$$I^*(t) = \sum_{k=1}^{K} I^{(k)}(t) \quad (10)$$

III. Generation of Data Set

As seen in Part II, the model of amperometric biosensor is based on a system of nonlinear partial differential equations. Because of nonlinearity (1)-(8) system was solved by numerical finite difference method [19] using the explicit scheme. More details on solving (1)-(8) system can be found in [2, 3].

Further defined (1)-(8) system parameters values was the same as in [3, 4, 8] to get comparable results. The only difference is in predefined substrates concentration values set.

In this paper the mixture of four substrates were considered i.e. $K = 4$. The values of these parameters were constant during numeric experiments:

$$D_{\mathrm{S}}^{(k)} = D_{\mathrm{P}}^{(k)} = 3 \times 10^{-6} \mathrm{cm}^2/\mathrm{s}, \quad k = 1, \dots, K,$$

$$K_M^{(k)} = 10^{-7} \mathrm{mol/cm}^3, \quad k = 1, \dots, K, \quad n_e = 2. \quad (11)$$

Individual maximal enzymatic rate characterizes each component of the mixture:

$$V_{max}^{(k)} = 10^{-6-k}\ \mathrm{mol/cm}^3\mathrm{s}, \qquad k = 1, \dots, K \quad (12)$$

Previous works [3, 4, 8] for each of K substrates $\mathrm{S}_1, \dots, \mathrm{S}_K$ used in geometric progression manner generated values of the concentrations $S_0^{(k)}$:

$$S_0^{(k)} \in \{S_{0,m} : S_{0,m} = \alpha_m \times S_0, m = 1, \dots, M\}, \ k = 1, \dots, K,$$

$$S_0 = 10^{-8} \mathrm{mol/cm}^3, \quad K = 4, \ M = 8,$$

$$\alpha_1 = 1, \alpha_2 = 2, \alpha_3 = 4, \alpha_4 = 8,$$

$$\alpha_5 = 12, \alpha_6 = 16, \alpha_7 = 32, \alpha_8 = 64. \quad (13)$$

Combination of all possible values of concentrations $S_0^{(k)}$ was used to generate data set (totaling $M^K = 8^4 = 4096$ records). This kind of data set is not suited as learning set for artificial neural network, because it contains gaps in the domain of interest - i.e. there are no records with concentrations $S_0^{(k)}$ from the interval $(\alpha_7 * S_0,\ \alpha_8 * S_0) = (32 * S_0,\ 64 * S_0)$. Artificial neural network results in this interval is unpredictable (this problem irrelevant for [3, 4], because they classify only

predefined values: $\alpha_1, \dots, \alpha_8$). To avoid this problem, data set, uniformly covering all domain space, was generated:

$$S_0^{(k)} \in \{S_{0,m} : S_{0,m} = \alpha_m \times S_0, m = 1, \dots, M\},\ k = 1, \dots, K,$$

$$S_0 = 10^{-8}\text{mol/cm}^3, \quad K = 4,\ M = 64,$$

$$\alpha_1 = 1, \alpha_2 = 2, \dots, \alpha_{64} = 64. \tag{14}$$

Combination of all possible values of concentrations $S_0^{(k)}$ was used to generate the data set (totaling $M^K = 64^4 = 16777216$ records).

To verify accuracy of determination of substrate concentrations in mixture, some test data set was used. This data set was generated by randomly selecting substrate concentrations from the domain of interest space:

$$S_0^{(k)} \in [\alpha_1 \times 10^{-8}, \alpha_{64} \times 10^{-8}]\ (\text{mol/cm}^3)$$

$$\in [1 \times 10^{-8}, 64 \times 10^{-8}]\ (\text{mol/cm}^3),$$

$$k = 1, \dots, K. \tag{15}$$

This way 1000 records were generated.

Previous works [3, 4, 8] gives reasoning what the thickness of enzyme layer d should be used. As good results with $d = 0.02$ were achieved, this value was used in the current work.

Biosensor work was conducted for 300 s ($T = 300$ s), first 10 s being inserted in mixture and then pulled out ($T_M = 10$ s). Values of the biosensor current density were stored in a file every second of simulated work, thus one record has 300 values (see Fig. 1).

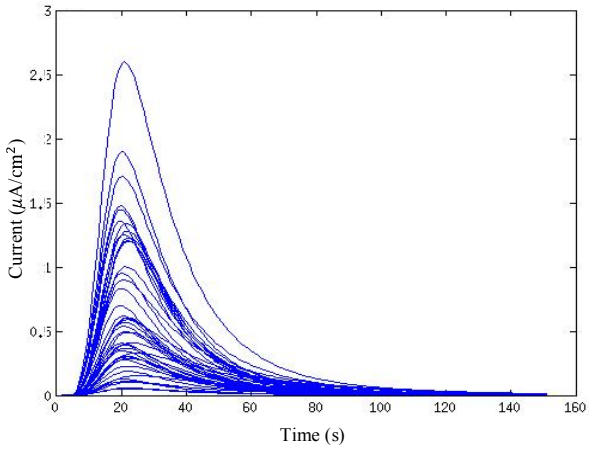


Fig. 1. Randomly selected biosensor responses. Curves are plotted until 160 s as then it becomes negligible.

IV. Application of Artificial Neural Networks

In this section application of artificial neural networks to determine concentrations of mixture is presented. First subsection describes data compression approach. Next subsection presents application of artificial neural networks. Finally, results are presented and analyzed.

A. Data Compression

Biosensor response, as mentioned previously, is 300 floating point numbers. This data is used to determine concentrations of mixture. This number or attributes is hardly manageable by artificial neural network [23]. On the other hand, this data is strongly correlated, so statistical data compression methods could be used [24]. One of these methods is principal components analysis (PCA) [24, 25]. This method was used in previous works and gave good results in case of classification [3, 4]. This method extracts, so-called principal components, which are uncorrelated data in orthogonal space. First principal component has the biggest dispersion, whereas, the last one has the lowest [24, 25].

Let's say we have the data matrix: $X = (X_1, \dots, X_N)^T$, where X_i is one record - biosensor response in our case. S - data covariance matrix. Then, using eigendecomposition of a covariance matrix [24], we get:

$$S = G\Delta G^T, \tag{16}$$

where $G = (G_1, \dots, G_P)$ – orthogonal matrix of normalized eigenvectors G_i (note: $G^T = G^{-1}$ as G is orthogonal). $\Delta = \text{diag}(\delta_1, \dots, \delta_P)$ is the diagonal matrix whose diagonal elements are the corresponding eigenvalues organized in decreasing order. Note: in our case $P = 300$, $N = M^K = 64^4 = 16777216$.

G is used in orthogonal linear transformation of domain space [24], i.e.:

$$X_{new} = XG, \tag{17}$$

where $X_{new} = (X_{new_1}, \dots, X_{new_N})^T$ is transformed data set of records: $X_{new_i} = (x_{new_{i_1}}, \dots, x_{new_{i_P}})$. New attribute $x_{new_{i_j}}$ is j-th principal component, having δ_j dispersion.

For the further analysis few first (having biggest dispersions) principal components will be used. There is no direct rule how many principal components ought to be used. In this paper we used seven principal components, as then variance rapidly decreases.

B. Artificial Neural Networks

In this paper feedforward neural network of a single hidden layer was used (see Fig. 2). The input layer consist of $J = 7$ nodes $X = (x_1, \dots, x_J)$. Each node represents principal component of principal component analysis. Hidden layer consisting of $p = 30$ nodes using sigmoid functions as their

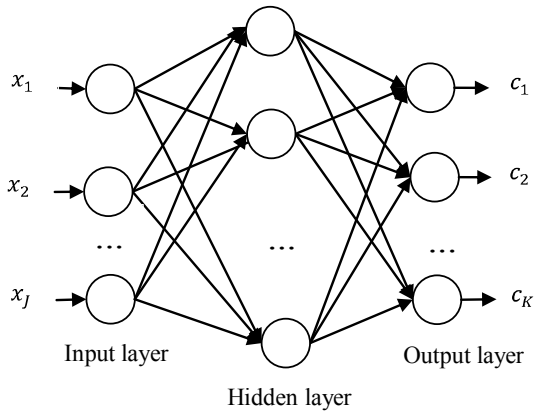


Fig. 2. Principle sheme of feedforward neural network: x_1, x_2 ... x_J - input values, c_1, c_2 ... c_K - output values (determined concentrations of substrates).

activation (transfer) functions. Output layer has $K = 4$ nodes $C = (c_1, \dots, c_K)$ with linear activation functions. Each output neuron c_i represents a concentration of substrate S_i. Such neural network scheme is universal approximator [20].

Artificial neural network mapping from X to C can be expressed by following function:

$$c_k = \sum_{s=1}^{p} \alpha_{s,k} \varphi(B_s X^T + \gamma_s) + \varepsilon_k, \quad k = 1, \dots, K, \qquad (18)$$

where c_i is value of i-th output neuron representing a concentration of substrate S_i, p is number of nodes in hidden layer, $\alpha_{s,k}$, $B_s = (b_1, \dots, b_j)$, γ_s, ε_k are the weights and $\varphi(u) = 1/(1 + \exp(-u))$ - sigmoid activation function.

Artificial neural network was trained by supervised batch learning procedure [23], using the advanced variant of back-propagation optimization algorithm - Levenberg–Marquardt [21, 22]. Previous section showed that learning data set consists of huge amount of records ($N = M^K = 64^4 = 16777216$). This number of records is hardly covered by any of unparalleled optimization algorithms. For this reason some randomized artificial neural network training algorithm was developed. In essence it randomly takes a small subset (4000 records) of learning data set and uses it to train artificial neural network. Algorithm continues until subsequent five times performance of artificial neural network fails to improve. This randomized algorithm uses less than $64^3 = 262144$ records, significantly reducing artificial neural network learning time. As gradient descent (Levenberg–Marquardt [21, 22] is the case) is local optimization algorithm, it was run multiple times to get global optimum [23].

C. Results

To validate artificial neural network accuracy, the absolute value of difference between artificial neural network output and real value was measured: $\Delta c_k = |c_k - S_{rltv_0}^{(k)}|$, $k = 1, \dots, K$. Note that in this work as in [8] forecasted not real concentration values $S_0^{(k)} \in [1 \times 10^{-8}, 64 \times 10^{-8}]$ (mol/cm^3), but instead some relative values: $S_{rltv_0}^{(k)} \in [1, 64]$, $S_0^{(k)} = S_{rltv_0}^{(k)} \times 10^{-8}$ mol/cm^3. This is done to get more intuitive error values. Artificial neural network was validated on a test data set and results are presented in Table 1. Optimization based method proposed by [8] gives results presented in Table 2. From these

TABLE I. ACCURACY OF ARTIFICIAL NEURAL NETWORK DETERMINED CONCENTRATION VALUES

	Δc_1	Δc_2	Δc_3	Δc_4
Mean	0.1131	0.0990	0.0927	0.1222
Std.	0.0944	0.0813	0.0754	0.0989

TABLE II. ACCURACY OF OPTIMIZATION BASED METHOD DETERMINED CONCENTRATION VALUES [8]

	Δc_1	Δc_2	Δc_3	Δc_4
Mean	0.0041	0.0108	0.0408	0.0464
Std.	0.0388	0.1663	0.8404	0.9755

tables it is seen that accuracy of [8] method is about $0.1222/0.0464 \approx 2.63$ times better in compare to artificial neural network results (as Δx_4 gives worst result for both methods, so characterizes accurateness). On the other hand more complex artificial neural network, using whole data set ($N = M^K = 64^4 = 16777216$) for training and more principal componenets ($J > 7$) as input values could give more accurate results. Also, as mentioned previously, the method proposed by [8] requires computing resources for optimization, so embedded system would be complex and expensive. Meanwhile, artificial neural networks need less computing resources as it only need evaluate (18) function.

V. CONCLUSIONS

In this paper the use of artificial neural networks to determine concentrations of mixture was proven. In compare with [8] proposed optimization based method, artificial neural network gives slightly worse results, on the other hand, artificial neural network need less computing resources, so embedded system would be simple and cheap.

The topic of this paper demands some further investigation. The experiments with real word data are needed as it contains some additional distortions so results of multiple substrate measurement may be less accurate.

ACKNOWLEDGMENT

Author acknowledges support by project "Promotion of Student Scientific Activities" (VP1-3.1-ŠMM-01-V-02-003) from the Research Council of Lithuania. This project is funded by the Republic of Lithuania and European Social Fund under the 2007-2013 Human Resources Development Operational Programme's priority 3.

REFERENCES

[1] F. Scheller, F. Schubert, "Biosensors". Amsterdam, Elsevier, 2004.

[2] R. Baronas, J. Christensen, F. Ivanauskas, J. Kulys, "Computer simulation of amperometric biosensor response to mixtures of compounds", Nonlinear Analysis: Modelling and Control, vol. 7(2), pp. 3–14, 2002.

[3] R. Baronas, F. Ivanauskas, R. Maslovskis, P. Vaitkus, “An analysis of mixtures using amperometric biosensors and artificial neural networks”, Journal of Mathematical Chemistry, vol. 36(3), pp. 281–297, 2004.

[4] R. Baronas, F. Ivanauskas, R. Maslovskis, M. Radavičius, P. Vaitkus, “Locally weighted neural networks for an analysis of the biosensor response”, Kybernetika, vol. 43(1), pp. 21–30, 2007.

[5] R. Baronas, F. Ivanauskas, J. Kulys, “Modelling dynamics of amperometric biosensors in batch and flow injection analysis”, Journal of Mathematical Chemistry, vol. 32(2), pp. 225–237, 2002.

[6] R. Baronas, F. Ivanauskas, J. Kulys, “Mathematical Modeling of Biosensors. Springer Series on Chemical Sensors and Biosensors”. Dordrecht, Springer, 2010.

[7] T. Schulmeister, “Mathematical modelling of the dynamics of amperometric enzyme electrodes”, Selective Electrode Rev., vol. 12, pp. 260–303, 1990.

[8] A. Ţilinskas, D. Baronas, “Optimization-based evaluation of concentrations in modeling the biosensor-aided measurement”, Informatica, vol. 22(4), pp. 589–600, 2011.

[9] U. Wollenberger, F. Lisdat, F.W. Scheller, “Frontiers in Biosensorics. 2. Practical Applications”. Basel, Birkhäuser, 1997.

[10] B.E. Rapp, F.J. Gruhl, K. Länge, “Biosensors with label-free detection designed for diagnostic applications”, Analytical and Bioanalytical Chemistry, vol. 398(6), pp. 2403–2412, 2010.

[11] S. Viswanathan, H. Radecka, J. Radecki, “Electrochemical biosensors for food analysis”, Monatshefte für Chemie, vol. 140(8), pp. 891–899, 2009.

[12] A.K. Wanekaya, W. Chen, A. Mulchandani, “Recent biosensing developments in environmental security”, Journal of Environmental Monitoring, vol. 10(6), pp. 703–712, 2008.

[13] C. Amatore, A. Oleinick, I. Svir, N. da Mota, L. Thouin, “Theoretical modeling and optimization of the detection performance: a new concept for electrochemical detection of proteins in microfluidic channels”, Nonlinear Analysis: Modelling and Control, vol. 11(4), pp. 345–365, 2006.

[14] L. Liu, “Optimal inputs and sensitivities for parameter estimation in bioreactors”, Journal of Mathematical Chemistry, vol. 47(3), pp. 1154–1176, 2010.

[15] M.E.G. Lyons, “Modelling the transport and kinetics of electroenzymes at the electrode/solution interface”, Sensors, vol. 6(12), pp. 1765–1790, 2006.

[16] L.D. Mell, T. Maloy, “A model for the amperometric enzyme electrode obtained through digital simulation and applied to the immobilized glucose oxidase system”, Analytical Chemistry, vol. 47(2), pp. 299–307, 1975.

[17] J. Kulys, “The development of new analytical systems based on biocatalysts”, Analytical Letters, vol. 14(B6), pp. 377–397, 1981.

[18] P.N. Bartlett, R.G. Whitaker, “Electrochemical immobilisation of enzymes. Part I. Theory.”, Journal of Electroanalytical Chemistry, vol. 224, pp. 27–35, 1987.

[19] A. Samarskii, “The Theory of Difference Schemes”. New York-Basel, Marcel Dekker, 2002.

[20] G. Cybenko, “Approximations by superpositions of sigmoidal functions”, Mathematics of Control, Signals, and Systems, vol. 2(4), pp. 303–314, 1989.

[21] K. Levenberg, “A method for the solution of certain non-linear problems in least squares”, Quarterly of Applied Mathematics, vol. 2, pp. 164–168, 1944.

[22] D. Marquardt, “An algorithm for least-squares estimation of nonlinear parameters”, SIAM Journal on Applied Mathematics, vol. 11(2), pp. 431–441, 1963.

[23] Š. Raudys, “Ţinių išgavimas iš duomenų”. Klaipėda, Klaipėdos universiteto leidykla, 2008.

[24] G.Dzemyda, O. Kurasova, J.Ţilinskas, “Daugiamačių duomenų vizualizavimo metodai”. Vilnius, Mokslo aidai, 2008.

[25] K. Pearson, “On lines and planes of closest fit to systems of points in space”, Philosophical Magazine, vol. 2(11), pp. 559–572, 1901.

INFORMACINĖS TECHNOLOGIJOS ● IVUS 2013 ● ISSN 2029-249X ● eISSN 2029-4824

Introduction of Photovolatic Behavioral Models

Jieming Ma, Prudence W.H. Wong
Department of Computer Science,
University of Liverpool,
Liverpool, UK
{jieming, pwong}@liverpool. ac.uk

K.L. Man, Nan Zhang, T.O. Ting, Sheng-Uei Guan
Xi'an Jiaotong-Liverpool University,
Suzhou, P.R. China
{ka.man, nan.zhang, toting, steven.guan}@xjtlu.edu.cn

***Abstract*—Recent developments in Photovoltaic (PV) generation system have heightened the need for reliable behavioral models that is capable of predicting the electrical characteristics of PV modules. Over the years, various types of PV models have been developed with different factors being considered. This paper introduces the most commonly used PV models together with their inherent properties. The advantages and disadvantages of the corresponding parameter estimation methods are extensively discussed, which provide a reference for choosing the preferable PV models.**

***Keywords*—Photovolatic cells, behavioral models, equivalent circuit, simulation.**

I. INTRODUCTION

Solar energy is clean, abundant and renewable - making it a good energy source for the future. Photovoltaic (PV) cells are semiconductors that are capable of producing electrons when photons strike the surface. They are normally assembled into modules or arrays on some kind of mounting systems. In the last decade, the decrease of the PV module prices and the increasing needs for green energy sources have given rise to the continuing and dramatic growth of the world PV market. The statistics released by European Photovoltaic Industry Association reported that the global cumulative capacity of PV systems grew 49-fold between 2000 and 2011 [1].

Although PV module prices fell 74% from 1995 to 2011 [2], the initial cost of a PV system is still relatively high. An accurate assessment for the electrical characteristic is therefore essential in the system design [3]. PV manufactures normally provide Maximum Power Point (MPP) properties of their PV modules, such as the current at MPP (Imp), the voltage at MPP (V_{mp}), and the power at MPP (P_{max}). These values are generally measured at the Standard Test Conditions (STCs) which correspond to a module temperature of 25 ºC and an irradiance of 1000 W/m^2 at 1.5 air mass spectral distributions. The Current and Voltage (I-V) characteristic curves under several different test conditions may also be presented by manufactures. Despite this, the data available in manufactures' data sheet are still limited and usually cannot fulfill the engineering requirement because PV modules always operate under environments far from these test conditions.

PV model, with the ability to predict I-V characteristics of a PV module under the working environment other than STCs, is a predictive performance tool that allows consumers maximize the cost effectiveness of the system before installing it [4]. It is widely applied to estimate whether a PV system will be economically feasible. Recently, many Maximum Power Point Tracking (MPPT) techniques have been proposed to overcome the problems caused by partial shadowing conditions [5]. For instance, Chen et al. [6] utilized model-based Particle Swam Optimization (PSO) to search the global MPP. In [7], the solar array was adaptively reconfigured by a control algorithm integrated with emulated PV module models. These methods have high-lighted the need for a reliable model with high accuracy but low computation complexity.

Significant research efforts have been put into the development of behavioral models of PV systems [8]. These models are generally analytical equations based on a physical description formulating photovoltaic generated current (I) with the most crucial technical characteristics and environment variables, such as operating voltage (V), the ambient temperature (T), and the irradiance (G). Even though the other environment factors like dust and wind velocity may vary the electrical characteristics of PV modules, it is quite impossible to obtain a model that accounts for every single effect on the performance of a PV model [9]. Among numerous modelling approaches, the single-diode model is usually considered to offer a good compromise between simplicity and accuracy [10]. In consideration of the recombination loss in the depletion region, Sah [11] introduced a more accurate model known as double-diode model. The three-diode PV model can be found in [12]. Although it takes into account the influence of grain boundaries and leakage current through the peripheries, the extra diode increases the number of parameters. Accordingly, more computational effort is required to predict the electrical characteristics via a three-diode PV model.

In the next section we step through a wide variety of PV models. Their parameter estimation methods, including the analytical, numerical, and evolutionary algorithm extraction techniques are introduced in section 3. The last section concludes this paper.

II. PV MODULE MODELS

There exists several single-diode PV models in literature describing the behaviors of a PV module for the arbitrary time profiles of V, T and G. Depending on the factors taken into account in the modelling, PV models can be classified into three types: ideal model, Rs-model and Rp-model, shown in Fig. 1(a), Fig. 1(b), and Fig. 1 (c), respectively. Fig. 1(d)

illustrates an accurate double-diode model. For simplicity, the more complex models, such as three-diode models are not studied in this paper.

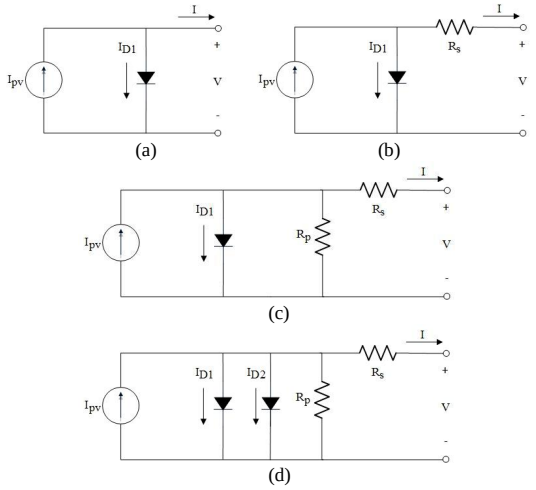


Fig. 1. Different models for a PV module: (a) ideal model, (b) R_s-model, (c) R_p-model and (d) two-diode model.

An ideal PV model is generally a linear independent current source in parallel to a diode [13], [14]. A reverse current, which depends linearly on *G* and is also slightly influenced by the *T*, is generated across the junction on the condition that the p-n junction of a PV module is exposed to incident light [15]. This current is known as photocurrent (I_{pv}). By eliminating the effect of photocurrent, a PV module behaves like a conventional diode and the I-V characteristics of an ideal PV module can be derived by an equivalent Thevenin circuit equation:

$$I = I_{pv} - I_o\left(e^{\frac{V}{nN_sV_T}} - 1\right) \tag{1}$$

In this equation, *n* is the diode ideality constant and its value varies between 1 and 2 depending on the material and the physical structure of the diode. The number of series connecting cells in the module is denoted by N_s. V_T is a constant called thermal voltage, whose value is a function of *T* that can be written as:

$$V_T = \frac{kT}{q} \tag{2}$$

where *k* is the Boltzmann constant (1.380650×10^{-23} J/K), *q* is the electron charge (1.602176×10^{-19} C), I_o is the saturation current. The thermal voltage at nominal temperature can be denoted as V_{tn}.

Rs-model improves the ideal model by recognizing the series resistance (R_s) in solar modules, which basically depends on the contact resistance of the metal base with the *p* semiconductor layer, the resistances of the *p* and *n* bodies, the contact resistance of the n layer with the top metal grid, and the resistance of the grid [16]. The values of PV current and voltage forming I-V curve are given in Equation (7):

$$I = I_{pv} - I_o\left(e^{\frac{V+IR_s}{nN_sV_T}} - 1\right) \tag{3}$$

The shunt resistance (R_p) exists mainly due to the leakage current of the p–n junction and depends on the fabrication method of the PV cell [3]. By identifying the effect of R_s, the R_p-model demonstrates a more accurate I-V relationship:

$$I = I_{pv} - I_o\left(e^{\frac{V+IR_s}{nN_sV_T}} - 1\right) - \frac{V+IR_s}{R_p} \tag{4}$$

In practice, the current source is also shunted by another diode that models the space charge recombination current [17]. Double-diode model improves the accuracy by including the additional diode, however, the corresponding parameter quantity increases to seven. The terminal current of a double-diode model therefore can be rewritten as:

$$I = I_{pv} - I_{o1}\left(e^{\frac{V+IR_s}{n_1N_sV_T}} - 1\right) - I_{o2}\left(e^{\frac{V+IR_s}{n_2N_sV_T}} - 1\right) - \frac{V+IR_s}{R_p} \tag{5}$$

TABLE I. PROPERTIES OF DIFFERENT PV MODELS

Model Name	Diode Quantity	Parameter Quantity	Parameters
Ideal model	1	3	I_{pv}, I_o, n
Rs-model	1	4	I_{pv}, I_o, n, R_s
Rp-model	1	5	I_{pv}, I_o, n, R_s, R_p
Two-diode model	2	7	I_{pv}, I_{o1}, I_{o2}, n_1, n_2, R_s, R_p

As seen in Equation (2)-(5), nonlinear and implicit relationships exist between the terminal voltage and current of these PV models. Saloux [3] reported that these mathematical equations normally can only be solved using tedious iterative numerical methods.

Table I shows several types of PV models with their different properties. They differ according to the parameters involved in the model. Thus, in literature, models are also named after the parameter quantity (e.g. Rs-model is called four-parameter model, Rp-model is called five-parameter model, etc.)

III. PARAMETER ESTIMATION

A PV model consists of several parameters, namely, I_{pv}, I_o, n, R_s, and R_p. After determining which PV models will be used, it is essential to estimate the optimal values of these unknown parameters. In general, the traditional parameter estimation

methods can be classified into two groups: analytical procedures ([15], [18], [19]) and numerical extraction ([20]-[22]) techniques. More recently, there has been an increasing interest in the Evolutionary Algorithm (EA) technique ([24]-[26]) due to its ability to handle nonlinear functions without requiring derivatives information.

A. Analytical technique

The Analytical technique utilizes mathematical equations to describe the parameters of PV models. In literature, there are many analytical expressions describing model parameters in terms of the physical parameters, such as the coefficient of diffusion of electrons in the semiconductor, lifetime of minority carriers, the intrinsic carrier density, etc. [12]. However, the values of these physical parameters are normally not provided by manufactures, which impels researchers to explore an alternative way of formulating parameters by using the information available in datasheet (e.g. short circuit current coefficient K_i and open circuit voltage coefficient K_v, short circuit current I_{sc}, open circuit voltage V_{oc}, the voltage at the MPP V_{mp}, the current at the MPP I_{mp}, etc.).

In view of the fact that linear relationship between I_{pv} and G, I_{pv} can be expressed as [3]:

$$I_{pv} = \left(I_{pvn} + K_i \Delta T\right)\frac{G}{G_n} \tag{6}$$

where I_{pvn}, G_n, T_n are the photocurrent, solar irradiance, cell temperature measured at STCs, respectively. ΔT is the difference between T and T_n. I_{pvn} is usually assumed equal to I_{scn}.

The value of the saturation current is conventionally described as a nonlinear function [16]:

$$I_o = I_{on}\left(\frac{T_n}{T}\right)e^{\frac{qE_g}{nk}\left(\frac{1}{T_n}-\frac{1}{T}\right)} \tag{7}$$

where

$$I_{on} = \frac{I_{scn}}{e^{(V_{ocn}/nV_{tn})}-1}$$

E_g is the band gap energy of the semiconductor (E_g=1.12 eV for the polycrystalline Si at 25 ºC). Aiming at matching the V_{oc} with the experimental data for a very large range of temperature, Villalva [3] further simplified Equation (3) to the following form:

$$I_o = \frac{I_{scn} + K_i \Delta T}{e^{(V_{ocn}+K_i\Delta T)/(nN_sV_T)}-1} \tag{8}$$

In Equation (8), I_{scn} and V_{ocn} denote the short circuit current and the open voltage at STC, respectively.

A large number of analytical methods have been applied to determine the values of R_s and R_p over the years. In [9], mathematical formulas are derived to predict R_s and R_p. However, they must estimate the slopes at the open-circuit point and short-circuit that are not given in datasheet. Iterative process was proposed in [3] and [4] based on several analytical conditions. This approach may decrease the absolute error, yet the computation complexity increases. Considering the fact that R_s and R_p vary in almost inverse linear mode with the solar irradiance, Brano [4] demonstrated an improved expression for the series and shunt resistances:

$$R_s(\alpha_G) = \frac{R_{sn}\cdot G_n}{G}\ ,\ R_p(\alpha_G) = \frac{R_{pn}\cdot G_n}{G} \tag{9}$$

The ideality factor (n) is an important parameter that can be used to decide whether the *p-n* junction device behaves closely or apart from the ideal case. Its value lies on the kind of the semiconductor. With the measured slopes at the open-circuit point and short-circuit, mathematical expressions are available to predict n. Based on a Lambert W-function, a simple method for extracting the n was proposed in [28]. For simplicity, n can be set to constant arbitrarily between 1and 1.5 [3]. It may be later modified in order to improve the model fitting, if necessary.

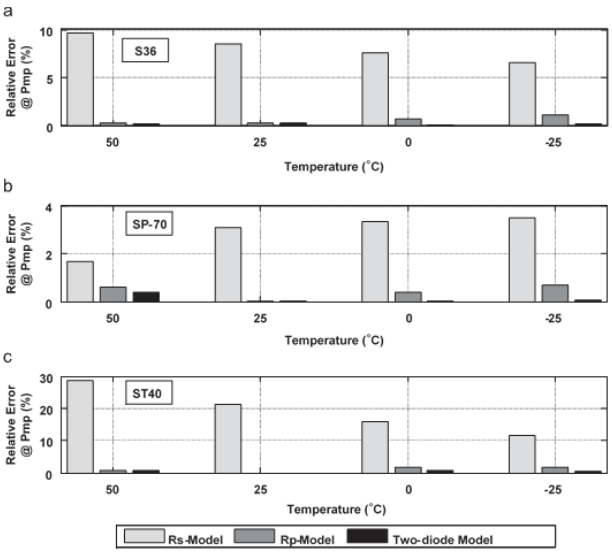


Fig. 2. Relative error for Pmp of Rs-, Rp- and the two-diode model for temperature variation (a) S36 (multi-crystalline) (b) SP-70 (mono-crystalline) and (c) ST40 (thin-film) (obtained from [29]).

TABLE II. RELATIVE ERRORS ON THE V_{MP} OF RS, RP AND TWO-DIODE MODEL AT DIFFERENT TEMPERATURE FOR S36 (MULTI-CRYSTALLINE) MODULE.

Temp (ºC)	Error		
	Rs-model	***Rp-model***	***Two-diode Model***
50	10.82	2.055	1.369

Temp (°C)	Error		
	Rs-model	*Rp-model*	*Two-diode Model*
25	9.393	1.818	1.212
0	8.315	2.174	1.630
-25	7.537	1.478	0.246

Obtained from [29]

By using the analytical technique, the accuracy of Rs-, Rp-, and double-diode models were compared in [29]. The modeling methods were validated by the measured data of selected modules: multi-crystalline (S36), mono-crystalline (SP-70) and thin-film (ST40). Fig. 2 shows the Relative error for P_{max} of Rs-, Rp- and the two-diode model for temperature variation. The relative errors of V_{mp} of the three models at different for S36 module can be seen in TABLE II. In the three model, the simulation and experimental results proved that the model with more parameters obtained a higher accuracy.

In short, analytical techniques are basically mathematical equations estimating the parameters in PV models and are normally simple to compute. The accuracy of these analytical techniques, however, hinges on the known coefficients and key operating points. The errors can be relatively significant if the values of these known data are incorrectly specified,

B. Numerical extraction technique

Aided by a statistical method, the numerical extraction technique fits a great many operating points on the I-V curves to obtain a more accurate solution. These curve fitting methods minimize the current error ε given in [9] as:

$$\varepsilon = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(I(V_i;p)-I_i)^2} \qquad (10)$$

where N is the number of measured I-V pairs represented by (V_i, I_i). p denotes the estimated parameters in the PV models. $I(V_i, p)$ is the calculated current using the predicted model parameters. By minimizing the root mean square error ε, the numerical extraction techniques can obtain accurate values for the relevant parameters. On the other hand, this approach may decrease the computational speed. Furthermore, its consistency, which means the accuracy of the final solution is significantly affected by the selection of the initial values [26].

C. EA technique

The EA technique, a stochastic optimization method, is very efficient in optimizing real-valued multi-modal objective functions [24]. Over the last decade, Genetic Algorithm (GA) [25], Particle Swarm Optimization (PSO) [26] and differential evolution (DE) [27] have been employed for estimating parameters of various PV module models. Before proceeding with the parameter extraction operations, an objective function, as well as some performance criterion, should be first determined. It profoundly influences the calculation of the solar cell parameters [25]. In most cases, Equation (10) is used as the objective function. It is worthwhile to remark that the EA technique may obtain the most accurate solution compared to any other strategies. However, this technique is usually computationally expensive. Moreover, the performance of EA may affected by a number of factors like initial points, and thus it is necessary to evaluate the set of algorithms via some criteria (e.g. accuracy of solution, consistency of solution, speed of convergence, the required number of control parameters, the choice of initial values, etc.).

IV. CONCLUSION

This paper has stepped through the ideal, Rs-, Rp- and two diode-PV behavioral models taken from the literature. In addition, three most commonly used parameter estimation methods have been discussed by elaborating their pros and cons. The comparative study provides some hints to choose the most suitable model, as well as the corresponding parameter estimation technique, according to the PV users' requirements.

REFERENCES

[1] European Photovoltaic Industry Association, Global Market Outlook for Photovoltaics until 2016, May 2012.

[2] U.S. DOE, Annual Energy Review 2011, Energy Information Administration (EIA) , 2012.

[3] M.G. Villalva, J.R. Gazoli, E.R. Filho, "Comprehensive approach to modeling and simulation of photovoltaic arrays", IEEE Trans. Power Electronics, vol. 24, no. 5, pp. 1198 – 1208, 2009.

[4] V. L. Brano, A. Orioli, G. Ciulla, A.D. Gangi, "An improved five-parameter model for photovoltaic modules," Solar Energy Materials and Solar Cells, vol. 94, no. 8, pp. 1358-1370, August 2010.

[5] S.A. Spanoche, J.D. Stewart, S.L. Hawley, I.E. Opris, "Model-Based Method for Partially Shaded PV Module Hot-Spot Suppression," IEEE Journal of Photovoltaics, vol.PP, no.99, pp.1-6, 2012.

[6] L.R. Chen, C.H. Tsai, Y.L. Lin, Y.S. Lai, "A biological swarm chasing algorithm for tracking the PV maximum power point," IEEE Trans. Energy Conversion, vol.25, no.2, pp. 484-493, June 2010

[7] D. Nguyen, B. Lehman, "An adaptive solar photovoltaic array using model-based reconfiguration algorithm," IEEE Trans. Industrial Electronics, vol.55, no.7, pp. 2644-2654, July 2008.

[8] J.H. Jung, S. Ahmed, "Real-time simulation model development of single crystalline photovoltaic panels using fast computation methods," Solar Energy, vol. 86, no. 6, pp. 1826-1837, June 2012.

[9] A.N. Celik, N. Acikgoz, "Modelling and experimental verification of the operating current of mono-crystalline photovoltaic modules using four- and five-parameter models," Applied Energy, vol. 84, no. 1, pp. 1-15, January 2007.

[10] C. Carrero, J.Amador, and S. Arnaltes, "A single procedure for helping PV designers to select silicon PV module and evaluate the loss resistances," Renewable Energy, vol. 32, no. 15, pp. 2579–2589, Dec. 2007.

[11] C.T. Sah, R.N. Noyce, W. Shockley, "Carrier Generation and Recombination in P-N Junctions and P-N Junction Characteristics", Proceedings of the IRE , vol.45, no.9, pp.1228-1243, September 1957.

[12] K. Nishioka, N. Sakitani, Y. Uraoka, T. Fuyuki, "Analysis of multicrystalline silicon solar cells by modified 3-diode equivalent circuit model taking leakage current through periphery into consideration," Solar Energy Materials and Solar Cells, vol. 91, no. 13, pp. 1222–1227, 2007.

[13] L. Castaner, S. Silvestre, Modelling Photovoltaic System using PSpice, John Wiley & Sons Inc, 2002.

[14] Y.T. Tan, D.S. Kirschen, N. Jenkins, "A model of PV generation suitable for stability analysis," IEEE Trans. Energy Conversion. vol. 19, no. 4, pp. 748-755, 2004.

[15] W.D. Soto and S.A. Klein and W.A. Beckman "Improvement and validation of a model for photovoltaic array performance," Solar Energy, vol. 80, no. 1, pp. 78 – 88, 2006.

[16] F. Lasnier and T. G. Ang, Photovoltaic Engineering Handbook, New York: Adam Hilger, 1990.

[17] M.F. AlHajri, K.M. El-Naggar, M.R. AlRashidi, A.K. Al-Othman, "Optimal extraction of solar cell parameters using pattern search," Renewable Energy, vol. 44, pp. 238-245, August 2012.

[18] J.P. Charles, G. Bordure, A. Khoury, P. Mialhe, "Consistency of the double exponential model with physical mechanisms of conduction for a solar cell under illumination," Journal of Physics D: Applied Physics, vol. 18, no.11, pp. 2261-2268, 1985.

[19] A. Ortiz-Conde, F.J. García Sánchez, J. Muci, "New method to extract the model parameters of solar cells from the explicit analytic solutions of their illuminated I–V characteristics," Solar Energy Materials and Solar Cells , vol. 90, pp. 352–61, 2006.

[20] M. Chegaar, Z. Ouennoughi, F. Guechi, "Extracting dc parameters of solar cells under illumination," Vacuum, vol. 75, no. 4, pp. 367-372, August 2004.

[21] M. Haouari-Merbah, M. Belhamel, I. Tobías, J.M. Ruiz, "Extraction and analysis of solar cell parameters from the illuminated current–voltage curve," Solar Energy Materials and Solar Cells, vol. 87, no. 1–4, pp. 225-233, May 2005.

[22] M. Chegaar, G. Azzouzi, P. Mialhe, "Simple parameter extraction method for illuminated solar cells," Solid-State Electronics, vol. 50, no. 7–8, pp. 1234-1237, July-August 2006.

[23] K. Ishaque, Z. Salam, H. Taheri, A. Shamsudin, "A critical evaluation of EA computational methods for Photovoltaic cell parameter extraction based on two diode model," Solar Energy, vol. 85, no. 9, pp. 1768-1779, September 2011.

[24] J. A Jervase, H Bourdoucen, A. Al-Lawati, "Solar cell parameter extraction using genetic algorithms," Measurement Science and Technology, vol. 12, pp. 1922–1925, 2001.

[25] M. Ye, X. Wang, Y. Xu, "Parameter extraction of solar cells using particle swarm optimization," Journal of Applied Physics, vol. 105, no. 9, pp. 094502-094502-8, May 2009.

[26] R. Gottschalg, M. Rommel, D.G. Infield, M.J. Kearney, "The influence of the measurement environment on the accuracy of the extraction of the physical parameters of solar cells," Measurement Science and Technology, vol.10, no. 9, pp. 796–804, 1999.

[27] W.T. da Costa, J.F. Fardin, D.S Simonetti, L.de V.B.M. Neto, "Identification of photovoltaic model parameters by Differential Evolution," IEEE International Conference on Industrial Technology (ICIT), pp. 931-936, March 2010.

[28] H. Bayhan, M. Bayhan, "A simple approach to determine the solar cell diode ideality factor under illumination," Solar Energy, vol. 85, no. 5, pp. 769-775, May 2011.

[29] K. Ishaque, Z. Salam, H. Taheri, Simple, "fast and accurate two-diode model for photovoltaic modules," Solar Energy Materials and Solar Cells, vol. 95, no. 2, pp. 586-594, February 2011.

INFORMACINĖS TECHNOLOGIJOS ● IVUS 2013 ● ISSN 2029-249X ● eISSN 2029-4824

Stylometry for Visualizing Rhetorical Differences in Parliamentary Speech

Justina Mandravickaite
Baltic Institute of Advanced Technology
Vilnius, Lithuania
justina@bpti.lt

***Abstract*—This paper presents an on-going experimental work on visualization of rhetorical differences in parliamentary speech using stylometric methods. As material for analysis, speeches of parliamentary factions from three transcripts of Lithuanian Parliament plenary sittings were chosen. Furthermore, stylometric analysis was performed with R - free software environment for statistical computing and graphics. Objective of the experiments was to explore if stylometry can produce significant results in above mentioned domain. Results obtained showed that stylometry with R has vast capabilities for visualization and that the list of at least 1000 words is necessary for analysis. Finally, the method allowed visualizing rhetorical differences of different factions in Lithuanian Parliament fairly well, however, for stronger assertions and implications further research must be realised.**

Keywords—stylometry; computational stylistics; parliamentary speech; R; statistical analysis

I. INTRODUCTION

Stylometry has been used for authorship attribution via analysis of measurable features of literary style. The basics of stylometry were introduced by Polish philosopher Wincenty Lutosławsky at the end of 19th century (Principes de stylométrie, 1890). Stylometry as a whole is based on an assumption that there are evidences of "personal style" occurring in texts. As such, the method is used often for literary texts, especially seeking an author identification of anonymous texts. There has been variety of authors performing stylometric experiments on literary texts, seeking to see stylistic differences (or similarities) in chosen set of texts – [1], [2], [3], [4], [5] and others.

This paper presents a little bit different application of stylometry – an on-going experimental work on visualization of rhetorical differences in parliamentary speech using stylometric methods. The objective of experiments is to explore visualization capabilities using stylometric analysis in R[1] (free software environment for statistical computing and graphics) to see rhetorical differences in parliamentary speech with focus on political groups (factions) in the Lithuanian Parliament.

This experiment is work in progress as we are exploring if this method can produce significant results in the above mentioned domain. So, for the analysis three transcripts of Lithuanian parliamentary plenary sittings were chosen. While performing stylometric analysis, many parameters play significant role, among them: language, number of texts, length of word list, choice of distance measure, etc. [6]. As there are no fixed parameters on performing stylometry [5], all the results presented in this paper were obtained via experiments by using different parameters as research on how stylometry operates on Lithuanian non-literary texts has not been performed as far as we know.

II. DATA AND METHODS

For this small experiment three transcripts of 2008-2012 Lithuanian Parliament plenary sittings were chosen: transcript of plenary sitting nº 43 (10 March 2009, morning session) [2], transcript of plenary sitting nº 244 (16 September 2010, morning session)[3] and transcript of plenary sitting nº 413 (20 March 2012, morning session)[4]. Speeches of Parliament members were extracted from the transcripts and put into separate text files according to the faction and the date of the sitting. Parliamentary speech on faction level was chosen because all the legislative decisions in the Parliament are adopted at this level. So, members of the Parliament join factions in order to seek their political objectives [7].

In order to explore stylistic differences of parliamentary speech on faction level, speeches that came in discussions regarding changes in legislation or adoption of new legislation were analyzed as there stylistic differences were the most clearly pronounced. Stylometric analysis was performed regarding usage of the most frequent words, assuming that the most frequent the word is the more importance it carries [8]. Therefore, the analysis shows stylistic correspondence regarding the most frequent words in the parliamentary speech on the level of factions.

Regarding analysis, most of the results [4], [5] show that the certain lower bound exists for the length of word list for the analysis. Preliminary experiments revealed that the list of at least 50 most frequent words (MFW) are necessary. So, if the whole text of certain faction consists of less than 50 words during the plenary sitting, it was ignored.

[1] http://www.r-project.org/

[2] http://www3.lrs.lt/pls/inter2/dokpaieska.showdoc_l?p_id=338732
[3] http://www3.lrs.lt/pls/inter3/dokpaieska.showdoc_l?p_id=381077
[4] http://www3.lrs.lt/pls/inter2/dokpaieska.showdoc_l?p_id=420348

TABLE I. PARAMETERS FOR STYLOMETRIC ANALYSIS

Parameter	Value
Corpus format	plain
Corpus language	English ALL
Analyzed features	words
Ngram size	1
Most Frequent Words (MFW)	Experiments performed with 100, 200, 300, 400, 500, 700, 1000, 2000, 3000, 5000, 7000 MFW
Start at	1
Culling min	0
Culling max	0
Analysis types	Cluster Analysis, Multidimensional Scaling, Bootstrap Consensus Tree
Consensus strength if Bootstrap Consensus Tree is selected	0.8
Distance measures	Classic Delta, Eder's Delta
Sampling	no sampling

Also, only political groups that existed during all three plenary sittings were investigated. So, speeches of the following factions in the selected transcripts of Lithuanian Parliament plenary sittings were used for experiment: Labor Party Political Group (Darbo partijos frakcija, DPF), Lithuanian Social Democratic Party Political Group (Lietuvos socialdemokratų partijos frakcija, LSDPF), Liberals Movement Political Group (Liberalų sąjūdžio frakcija, LSF), Homeland Union - Lithuanian Christian Democrat Political Group (Tėvynės sąjungos-Lietuvos krikščionių demokratų frakcija, TS-LKDF) and Order and Justice Political Group (frakcija "Tvarka ir teisingumas", TTF).

For the stylometric analysis with the purpose of visualization, R - free software environment for statistical computing and graphics – was used. R was chosen because it satisfactorily processes texts, provides all necessary tools for statistical analysis and visualization, and its performance is sufficient as well. Finally, R scripts execute all the analysis and computations from the initial text processing to visualization of results without the need of any additional tools.

Stylometric experiments on the selected transcripts of parliamentary plenary sittings were performed using "stylo" script for R [6]. It computes Principal Components Analysis, Cluster Analysis, Multidimensional Scaling, and Bootstrap Consensus Trees [6]. Cluster Analysis and Multidimensional Scaling is performed on Delta distance [1] table while Bootstrap Consensus Tree a compromise between quantity of virtual Cluster Analysis for distinct lists of most frequent words and culling, that refers to rate of words that were not in all the texts to be removed [6]. Also, it allows selection out of many of parameters for analysis (language, pronoun deletion, different distance measures, length of word list, etc.).

For the experiment on transcripts of parliamentary plenary sittings with focus of speeches of political groups, Cluster Analysis, Multidimensional Scaling, and Bootstrap Consensus Trees [6] were used with different length of word lists. As for the distance measures, variations of Burrows Delta were selected as it is the most basic distance measure in stylometry [9].

III. EXPERIMENTAL RESULTS

The objective of the experiments was to identify differences in rhetoric of parliamentary speech via visualization provided by stylometry. Parameters for the analysis were chosen based on recommendations in literature [5], [6] and experimental evaluation.

In order to explore differences in rhetoric of parliamentary groups, a number of plots have been generated using parameters displayed in Table 1. Depending on the chosen parameters, the results varied. Two examples (though there were many generated) can be seen in Fig. 1-2. In Fig.1 Eder's Delta was used while in Fig. 2 – Classic Delta. Additionally, as culling (ignoring the most characteristic words in the selected texts for analysis) did make no change for results – possibly because of small set of texts selected for the analysis, further experiments went with culling 0%.

TABLE II. PARAMETERS FOR FIG. 1-2

Parameters for Fig. 1	Parameters for Fig. 2
corpus.format = "plain" corpus.lang = "English.all" analyzed.features = "w" ngram.size = 1 mfw.min = 300 mfw.max = 300 mfw.incr = 100 start.at = 1 culling.min = 0 culling.max = 0 analysis.type = "CA" distance.measure = "ED" sampling = "no.sampling"	corpus.format = "plain" corpus.lang = "English.all" analyzed.features = "w" ngram.size = 1 mfw.min = 300 mfw.max = 300 mfw.incr = 100 start.at = 1 culling.min = 0 culling.max = 0 analysis.type = "CA" distance.measure = "CD" sampling = "no.sampling"

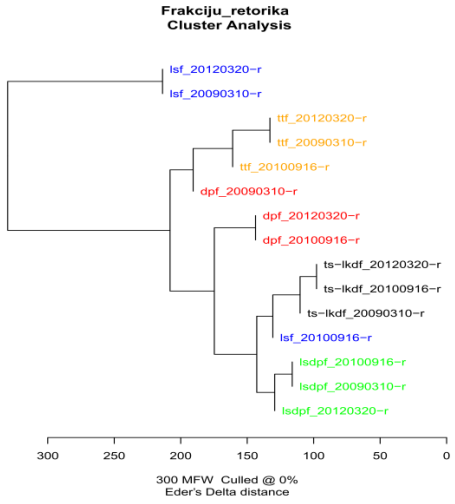


Fig. 1

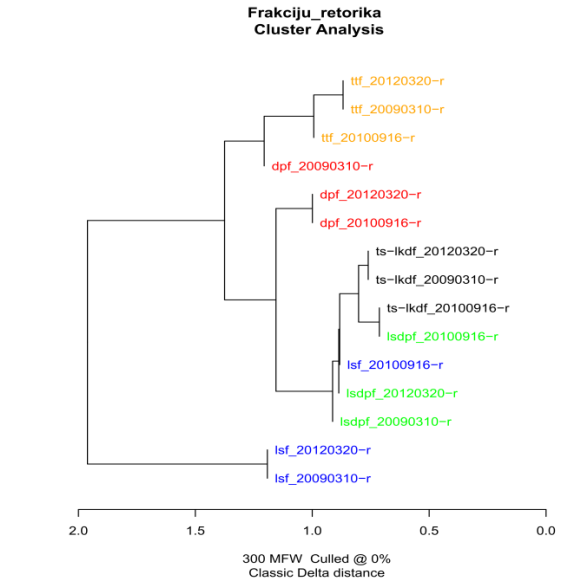


Fig. 2

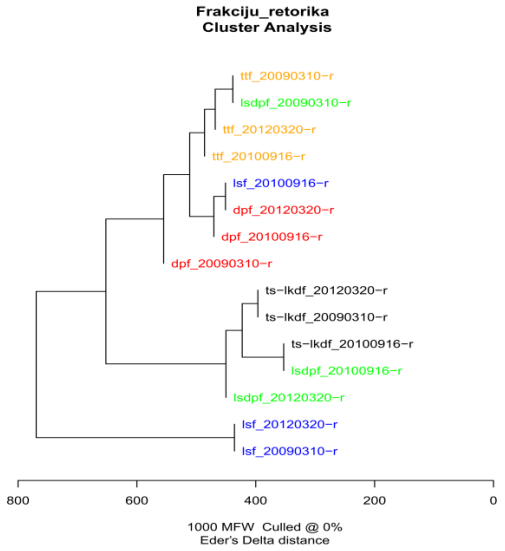


Fig. 3

As it can be seen in Fig. 1-2, different Delta distances with word list of 300 MFW, results generated were different (see parameters of the analysis in Table 2). Parliamentary speeches of the same political group come close together with exception of LSF (plenary sitting n^{o} 244, 16 September 2010) in both plots. Also, speeches made by LSDPF (plenary sitting n^{o} 244, 16 September 2010) come closer to TS-LKDF in Fig. 3. Thus, generating more plots with different parameters let us see if this tendency continues or disappears.

Exploring results presented in Fig. 3-4, one can observe differences in grouping. However, speeches of TS-LKDF and LSDPF appeared to have the least changes in position switch. On the other hand, speeches of DPF were jumping while using Classic Delta. Furthermore, LSDPF (plenary sitting n^{o} 43, 10 March 2009) found its place by the TTF and LSF (plenary sitting n^{o} 244, 16 September 2010) sat by the DPF.

As Fig. 5-6 displays, notwithstanding the choice of distance measure, two big groups are noticed in both graphs. One group consists of speeches of TS-LKDF and LSDPF, another – of speeches of TTF, LSF and DPF. Though, those two groups can be distinguished much clearer with Classic Delta than with Eder's Delta. Regarding the second group, speeches of DPF and TTF tend to go close to each other.

So, taking different lengths of word lists for analysis, trend occurred that speeches of TS-LKDF and LSDPF stick closer though ideologically being on the opposite sides of political spectrum. Meanwhile, TTF and DPF are grouping together and LSF, though changing positions quite a lot during the experiments, remains somewhere in the middle.

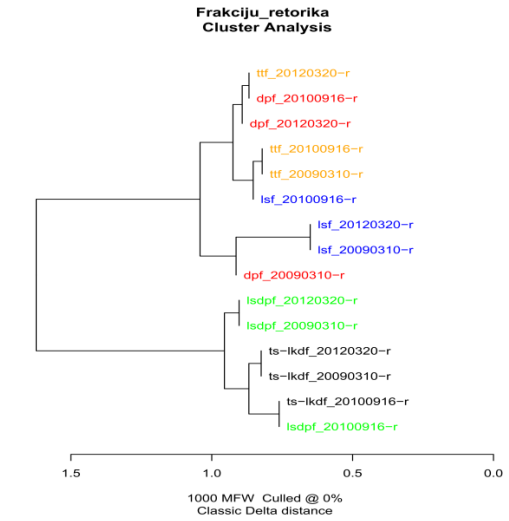


Fig. 4

Taking our small experiment further, Multidimensional Scaling was used to see if similar results can be observed. Two examples can be seen in Fig. 7-8. Again, different distance measures (Eder's Delta – in Fig. 8 and Classic Delta – in Fig. 9) were used. And similar results as in Fig. 6-7 were attained while performing analysis with word list of 3000 MFW. It appeared that results of the various experiments differed depending on the parameters selected. Only starting with 1000 MFW the results became steadier. Author [4] also recommended using more than 1000 MFW for stylometric analysis. However, according to [5], there is no universal solution for choosing length of word list for stylometry.

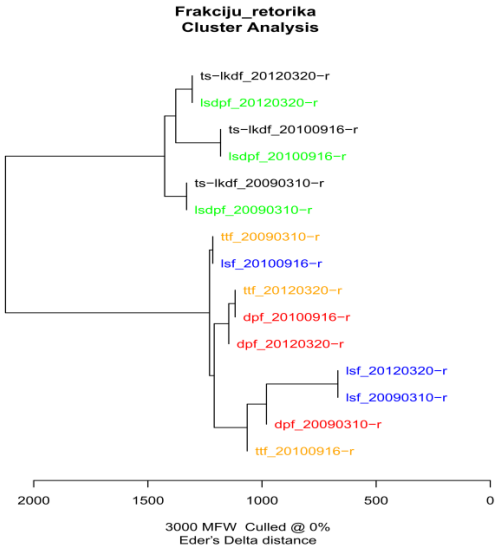


Fig. 5

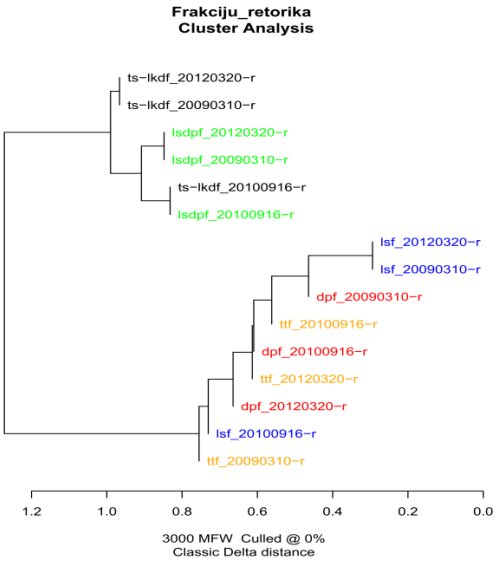


Fig. 6

Finally, taking one more stylometric method – Bootstrap Consensus Tree, that is a combined method of Cluster Analysis and Decision Tree, – plots depicted in Fig. 9-10 were generated. Parameter of Consensus 0.8 was taken meaning that there is a link between two texts that appears in 80% of invisible Cluster Analyses. Additionally, Bootstrap Consensus Tree computes compromise results of the latter.

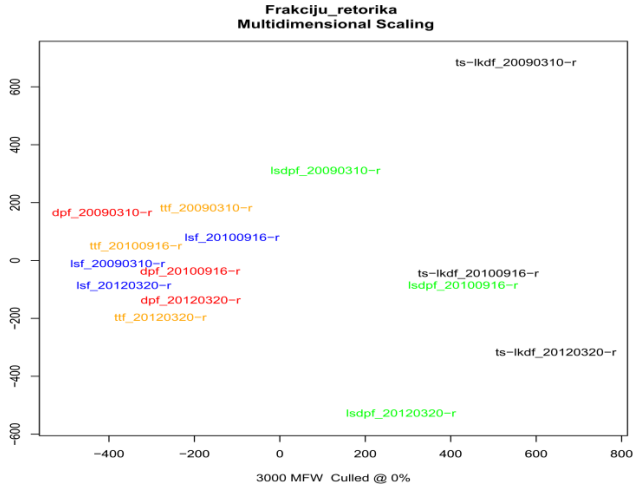


Fig. 7

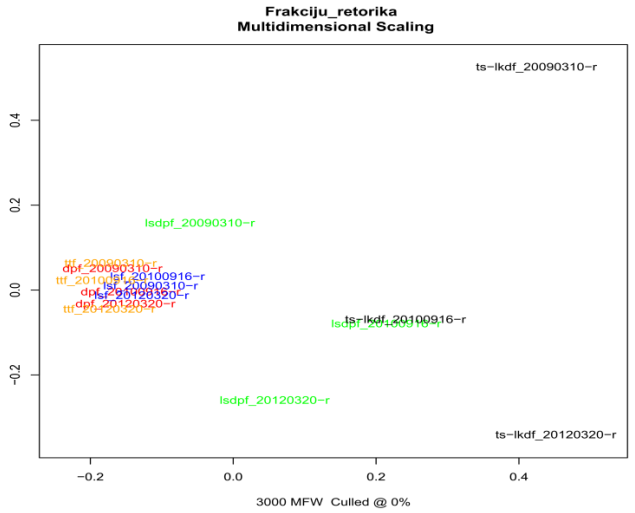


Fig. 8

As the graphs show, the results of the latter experiments were almost identical with exception of DPF (plenary sitting n^{o} 244, 16 September 2010 and plenary sitting n^{o} 413, 20 March 2012) and LSF (plenary sitting n^{o} 244, 16 September 2010) that change places among themselves. As with Cluster Analysis and Multidimensional Scaling, TS-LKDF and LSDPF remains a close group. Meanwhile, DPF and TTF, with some variations, tend to stand by each other. Finally, LSF positions itself in the corner of the latter group, but not really close with exception of LSF (plenary sitting n^{o} 244, 16 September 2010).

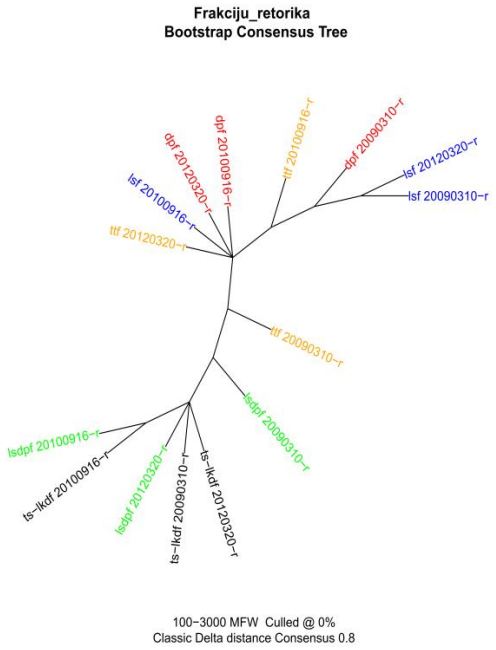


Fig. 9

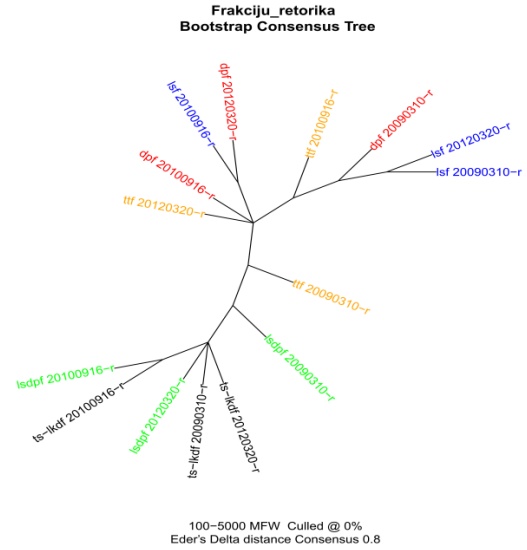


Fig. 10

To summarize, stylometric methods as means to visualize rhetoric differences in a case of speeches of Lithuanian Parliament political groups provided interesting results and proved to have enough capabilities of graphic presentation. But, taking limitations into consideration, this experiment could take 5000 MFW at max because of the chosen set of texts that is small. For further analysis, using longer than 5000 MFW word list, bigger set of texts and/or longer texts are needed.

IV. CONCLUSIONS

The idea to use stylometry to visualize rhetoric differences of non-literary texts in Lithuanian proved fulfilling as well as challenging as stylometric methods are mainly used for literary texts. Even though this experiment was very limited, some findings can be reported. Firstly, stylometry in R gave opportunity to obtain interesting results and proved to have vast capabilities of visualization. Secondly, to attain as reliable results as possible, as it can be assumed after performing several tens of experiments, at least 1000 MFW for analysis are needed. Thirdly, regarding the results of generated plots, TS-LKDF and LSDPF tend to stay close, meaning that their usage of words was similar. Following that, TTS and DPF though results varied for some degree, tended to position themselves near each other. And LSF in the performed experiments took place in some cases closer to DPF and TTF and in other ones – between the two mentioned groups – TS-LKDF & LSDPF and DPF & TTF. However, for stronger assertions and implications further research must be realised.

REFERENCES

[1] J. F. Burrows, "Delta': A Measure of Stylistic Difference and a Guide to Likely Authorship," Literary and Linguistic Computing, 17 vol. no. 3, pp. 267-87, 2002.

[2] P. Juola, J. Sofko and P. Brennan, "A Prototype for Authorship Attribution Studies," Literary and Linguistic Computing, vol. 21 no. 2, pp. 169-78, 2006.

[3] M. L. Jockers, D. M. Witten and C. S. Criddle, "Reassessing Authorship in the 'Book of Mormon' Using Delta and Nearest Shrunken Centroid Classification," Literary and Linguistic Computing, vol. 23 no. 4, pp. 465-91, 2008.

[4] D. L. Hoover, "Testing Burrows's Delta," Literary and Linguistic Computing, vol. 19 no. 4, pp. 453-475, 2004.

[5] J. Rybicki and M. Eder, "Deeper Delta across genres and languages: do we really need the most frequent words?" Literary and Linguistic Computing vol. 26 no. 3, pp. 315-21, 2011.

[6] M. Eder and J. Rybicki, "Stylometry with R," Digital Humanities 2011: Conference Abstracts, Stanford University, Stanford, CA, pp. 308-11, 2011.

[7] A. Gelbukh and G. Sidorov, "Zipf and Heaps Laws' coefficients depend on language," Computational Linguistics and Intelligent Text Processing, pp. 332-335, 2001.

[8] Statute of Parliament of the Republic of Lithuania. http://www3.lrs.lt/pls/inter2/dokpaieska.showdoc_e?p_id=284135, accessed 2013 04 08

[9] S. Argamon, "Interpreting Burrows's Delta: Geometric and Probabilistic Foundations," Literary and Linguistic Computing, vol. 23 no. 2, pp. 131-47, 2008.

Skolos spąstų logistinis modeliavimas IT priemonėmis

Andželа Mialik
Vilniaus universitetas
Kauno humanitarinis fakultetas
Muitinės g. 8, LT-44280, Kaunas, Lietuva
andzelina.m@gmail.com

Santrauka. Pastaruoju metu atliekami ekonominio augimo tyrimai vis labiau įtikina, kad nagrinėjant kapitalo (produkto) augimą būtina vertinti prisotinimą. Tam tenka apibrėžti augimo ribą. Straipsnyje pateikiamas investuoto kapitalo augimo modelis, kurio viena iš grandžių modeliuoja augimo ribos dydį. Apžvelgus mokslinę literatūrą ir išanalizavus žinomas augimo funkcijas, skirtas populiacijos dydžio, gimstamumo ar mirtingumo prognozavimui, nuspręsta naudoti klasikinį logistinį biologinių populiacijų raidos modelį, kurį 2002 metais S.Girdzijauskas esmingai modifikavo ir pritaikė ekonominių reiškinių tyrimui [1].

Pastarojo modelio panaudojimas leido išaiškinti nestabilių ekonominių situacijų susiformavimo mechanizmą. Naujai aptiktos šio mechanizmo pagrindinės grandys – didėjančio pelningumo ir skolos spąstų paradoksai. Naujai aptiktieji paradoksai vaidina labai svarbų (dažnai lemiantį) vaidmenį, formuojantis nestabilioms ekonominėms situacijoms. Viena iš tokių situacijų pasireiškimo sričių – skolų spąstų susiformavimas investuojant su svertu.

Straipsnyje nagrinėjama rinkų prisotinimo (dalinai ir perprodukcijos) įtaka investicinių skolos spąstų susidarymui.

Reikšminiai žodžiai: logistinė funkcija, augimas, investavimas, modelis, analizė, prisotinimas, skolos spąstai, svertas, augimo riba.

I. ĮVADAS

Ekonominių procesų modeliavimas yra gana plačiai paplitęs šiuolaikinis krizinių situacijų analizės būdas. Modeliuojant kapitalo augimą dažniausia neatsižvelgiama į įvairius augimo trukdžius ir apribojimus. Nagrinėjant praktinius aspektus kapitalas daugumoje modelių yra laikomos neriboto augimo. Tačiau toks kapitalo vertinimas nėra tikslus.

Apžvelgus mokslinę literatūrą ir išanalizavus žinomas augimo funkcijas [2,3], buvo nuspręsta naudoti P.Verhulsto logistinį biologinių populiacijų raidos modelį, kuris 2002 metais S.Girdzijausko buvo modifikuotas ir pritaikytas ekonominių reiškinių tyrimui [1].

Darbe nagrinėjamas investuotas kapitalas yra siejamas su rinkos, į kurią yra investuojama, baigtine talpa. Dėl to ir investuojamo kapitalo kiekis turi būti ribotas. Šiame darbe pagrindinis dėmesys sutelktas į ribotos augimo, arba logistines funkcijas, aprašančias kapitalo kaupimo (augimo) procesą. Reikia turėti galvoje ir tai, kad logistinis augimas yra būdingas ne tik kapitalui, bet ir bet kokioms populiacijoms, kurių augimo greitis yra proporcingas jų pačių dydžiui [2,14]. Tai ne tik įvairios biologinės populiacijos, bet ir tokie ekonominiai rodikliai, kaip BNP, BVP ir kt. Mes tokias populiacijas dažnai vadinsime tiesiog produktu [2,3,4].

Darbe kalbama apie investavimą su svertu. Finansinis svertas – tai skolinto kapitalo naudojimas investicijos pelnui padidinti. Bankinėje veikloje finansinis svertas suprantamas kaip santykis tarp skolinto ir nuosavo kapitalo. Kita vertus, finansinis svertas – tai banko paskola investuotojui [11,13].

Investicijų pajamingumo nustatymui taikomas svertinis logistinis augimo modelis, kadangi atlikti tyrimai parodė, jog taikant logistinį kapitalo kaupimo modelį, mes pradedame vertinti rinkos prisotinimą, bei atlikus modeliavimo veiksmus, galime pamatyti investicijos augimo elgesį egzistuojant skirtingiems rinkų prisotinimo lygiams [13]. Šiuolaikinė investavimo praktika ir literatūra apie investavimą su svertu nurodo, kad toks investavimas yra labai rizikingas, tačiau nepateikia ir nepaaiškina tos rizikos šaltinių [3,4].

Straipsnio tikslas – pristatyti logistinį investavimo su svertu modelį, leidžiantį tirti skolos spąstų efektą.

Straipsnio 2-jame, pagrindiniame, skyriuje skaitytojas yra supažindinamas su rinkų prisotinimu, jo vertinimo svarba, pateikiama originalaus modelio struktūra, kurioje nagrinėjama investavimo su svertu rizika.

II. LOGISTINIS MODELIAVIMAS SU SVERTU

Duomenų eilučių (pvz., žaliavų kaina, BVP ir kt.) analizei taikomi įvairūs augimo modeliai, kurių pagalba yra analizuojamas ir prognozuojamas įvairių sistemų elgesys. Nuo seno yra žinomi T.Malthus'o (eksponentinis), B.Gompertz'o, P.F.Verhiulst'o (logistiniai) augimo modeliai [2,3,4]. Šiuos modelius įvairūs mokslininkai: J.D.Sterman, T.Modis, S.M.Perrin, W. Yung Jason ir H.A.Jesse taikė savo tyrimuose populiacijos augimo, mirtingumo, gimstamumo – vaisingumo analizei. Dažniausiai buvo modeliuojami biologiniai, fiziniai, medicininiai ir panašūs procesai [2,3,4,6].

Kadangi S.Girdzijausko modernizuota logistinė funkcija susieta su palūkanų norma [1], o ekonominiuose procesuose tai yra labai svarbu, sukurtame algoritme ji ir yra naudojama kapitalo kaupimui modeliuoti.

Logistinės analizės taikymas atliktuose tyrimuose parodė, kad prisotintos sistemos augantis produktas (mūsų atveju - kapitalas) elgiasi skirtingai nei neprisotintos. Kai kapitalas investuojamas uždaroje ar pusiau uždaroje rinkoje, o didėjant prisotinimui (kartu ir paslėptajai perprodukcijai), investicijos pelningumas auga, pasireikšdamas nevaržomu vidinės grąžos

didėjimu, tuomet pasireiškia ***didėjančio pelningumo paradoksas*** [11,13].

Logistinės analizės tyrimuose vartojama nemažai specifinių terminų, betarpiškai susietų su naudojamais modeliais. Nagrinėjant ekonominius investicinius procesus svarbu tinkamai suprasti, kas yra rinka. Mūsų nagrinėjamu aspektu rinka tai mainų sfera, kurioje, veikiant pasiūlos ir paklausos dėsniams, vyksta pirkimo ir pardavimo procesas. Tai sfera, kurioje (neskaitant įvairių tarpininkų) veikia du dalyviai: pardavėjas ir pirkėjas. Investuotojo požiūriu kiekviena produktų grupė, ar net pavienis produktas, turi savo rinką. Kita vertus, rinka gali būti kelių (daugelio) pavienių rinkų visuma [7,11,13].

Investicine rinkos talpa, arba potencialiuoju kapitalu, galime laikyti kapitalo kiekį, kuris gali būti efektyviai įsisavintas konkrečioje rinkoje (konkrečioje investicinėje aplinkoje) [13].

Rinkos prisotinimas – tai ribotos talpos (uždarosios ar dalinai uždarosios) rinkos dalinis ar visiškas užpildymas kapitalu, pasireiškiančiu atitinkamų rinkos produktų pertekliumi (perprodukcija). Kaip teigia S.Girdzijauskas, rinkos prisotinimas prasideda tada, kai rinkos talpa iš begalinės tampa išmatuojama arba, kai investuojama yra sparčiau nei plečiasi pati rinka (kai pasiūla auga sparčiau už paklausą) [11].

Rinkos prisotinimas stipriai veikia finansinio sverto investicinį mechanizmą. Neigiamas skolinto kapitalo poveikis bendram pelno lygiui dėl investavimo į prisotintas rinkas yra finansinio sverto rizika. Pats finansinis svertas bankinėje veikloje, kaip jau buvo minėta, suprantamas kaip skolinto ir nuosavo kapitalo poveikis bendram investicijos pelno lygiui, finansuojant tą investiciją nuosavo ir skolinto kapitalo sąskaita [11].

Finansinio sverto modelyje panaudota anksčiau minėta 2002 m. Girdzijausko sudaryta logistinė kapitalo augimo formulė:

$$K = \frac{K_p \cdot K_0 \cdot (1+i)^n}{(K_p - K_0) + K_0 (1+i)^n}, \qquad (1)$$

kurioje: K_0 – nuosavos investuotojo lėšos, i – svertą teikiančio banko palūkanų norma, K_p – rinkos talpa, n – investavimo trukmė [9,10].

Panaudojant formulę (1) sudarome investavimo su svertu prisotintoje rinkoje skaičiavimo logistinį kaupimo modelį [13]:

$$K = (1+m) \cdot \frac{K_p \cdot K_0 \cdot (1+j)^n}{(K_p - K_0) + K_0 (1+j)^n} - m \cdot \frac{K_p \cdot K_0 \cdot (1+i)^n}{(K_p - K_0) + K_0 (1+i)^n}. \qquad (2)$$

Čia kintamieji aprašyti (1) modelyje bei papildomai naudojami j – numatomas planuojamos investicijos pajamingumas (pajamingumo norma) ir m – svertas (sverto koeficientas).

Atliekant prisotintų sistemų tyrimus kartu su didėjančio pelningumo paradoksu, buvo aptiktas kitas paradoksalus reiškinys – **skolos,** arba **kredito spąstai.** Jį galėtume apibrėžti taip: jei investuojama uždaroje (riboto talpumo) rinkoje, tai skolinto kapitalo (skolos) augimo greitis pralenkia nuosavo kapitalo augimo greitį; be to, šis augimo greičių skirtumas, pradžioje buvęs nežymus, po tam tikro periodų skaičiaus esmingai išauga. Vadinas, jei rinka yra neprisotinta, tai tiek nuosavas, tiek skolintas kapitalas auga identiškai. Tuo tarpu prisotintoje rinkoje skola auga išskirtinai greitai [11,13].

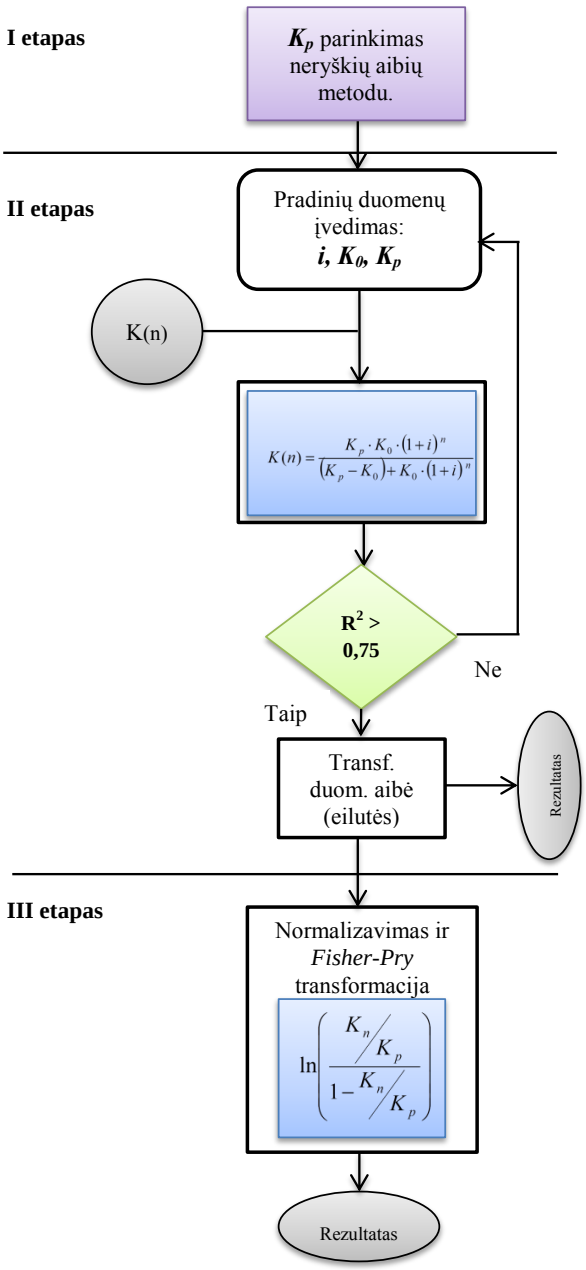
Reikia pabrėžti, kad rinkų prisotinimas glaudžiai susijęs su tose pačiose rinkose vykstančia perprodukcija (dažniausiai ji būna paslėptoji).

Kapitalo (produkto) perprodukcija yra lydima pelningumo mažėjimo tik atvirose rinkose, kur nėra prisotinimo. Tuo tarpu dalinai ar visai uždarose (ir prisotintose) rinkose perprodukciją lydi pelningumo augimas, o tai savo ruožtu suaktyvina vertybinių popierių rinką. Pastaroji perprodukciją dar labiau didina [13]. Taip išpučiamas kainų burbulas. Burbului subliuškus ar sprogus, sulaukiame nuosmukio ar krizės. [9,12]. Eurozonos krizė taip pat didele dalimi yra suformuota tų pačių sąlygų – prisotinimo ir paslėptosios perprodukcijos. Minėtos sąlygos laiku nebuvo reikiamai įvertintos, dėl to rinkos pradžioje ėmė kaisti (pasireiškė didėjančio pelningumo paradoksas), vėliau susiformavo skolos spąstai.

Modeliuojant rinkos kaitimą ir burbulo susiformavimą pastebime, kad didelę (gal net esminę) įtaką turi rinkos talpa arba potencialiojo kapitalo parametras K_p. Šio parametro nustatymas yra vienas problemiškiausių modeliavimo uždavinių. Šios problemos sprendimui buvo nutarta naudoti neraiškių aibių metodą. 1965 metai L.A.Zadeh pirmasis pristatė fuzzy (neraiškių aibių) logiką – taisyklių rinkinių apibrėžimus bei taikymo galimybes. Kiekvienas rinkinys apima apibrėžtus tarpusavio modelio narių (parametrų) ryšius, aprašytus paties vartotojo. Neraiškių aibių taikymas įgijo pripažinimą ir yra sėkmingai taikomas visame pasaulyje optimizuojant žmonių elgesio taisyklių rinkinius tam tikrų procesų valdymui [5,11,15,16]. Šis metodas buvo panaudotas ir investavimo su svertu modelyje.

Pav. 1 pavaizduota svertinio modelio struktūrinė schema. Modeliavimas pradedamas nuo potencialiosios augimo reikšmės (augimo ribos) K_p nustatymo. Šiame etape yra naudojamas fuzzy logikos metodai [5]. Apibrėžus parametrą K_p, parenkami ir kiti, mažiau problemiški ir lengviau nustatomi parametrai. Jie visi, kartu ir tiriamų statistinių duomenų eilutė, yra įvedami į modelį. Pati modeliavimo eiga suskirstyta į tris etapus.

I etapas: šiuo metu atliekamas tyrimas, kuriame K_p nustatymui naudojamas fuzzy logikos metodas. Sudarytos taisyklės (JEI sąlyga, TAI išvada), pagal kurias vertinant statistinių duomenų vidurkį bei statistinių reikšmių pasiskirstymą aplink vidurkį mes nustatome kapitalo kaupimo ribą.



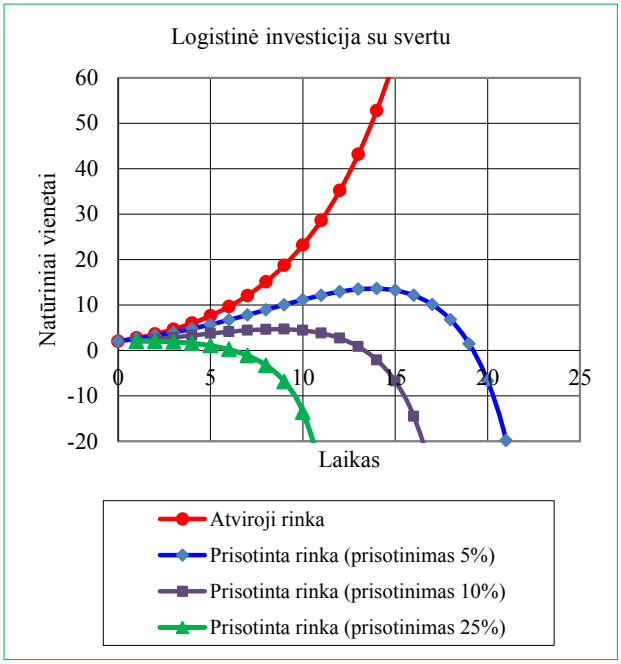
Pav.1 Investavimo su svertu struktūrinis modelis

II etapas: apskaičiuojamos logistinės regresijos lygties koeficientų reikšmės įvertinant išorinius kintamuosius. Gautų duomenų patikimumą patikriname apskaičiuodami logistinės funkcijos determinacijos koeficientą. Pagal apskaičiuotą determinacijos koeficientą atliekama pradinių duomenų korekcija. Laikome, kad R^2 (determinacijos koeficientas) mūsų modelyje turėtų būti didesnis negu 0,75. Jei ši sąlyga netenkinama, modelis reikalauja tikslinti pradinius duomenis. Jei sąlyga tenkinama – sudaroma augimo regresijos lygtis, kuri toliau pertvarkoma (trečiasis etapas).

III etapas: gautos reikšmės normalizuojamos ir atliekama jų *Fisher – Pry* transformacija. Logaritmuojant gautoji lygtis pertvarkoma į tiesę [6].

Rezultatas: gauname grafiškai pavaizduotą kapitalo kaupimo (augimo) regresijos modelį. Kontroliuodami statistinį parametrą – determinacijos koeficientą – optimizuojame pagrindinius logistinio modelio parametrus. Radus regresijos lygtį su optimaliais parametrais, įvertiname rinkos prisotinimo laipsnį ir sprendžiame apie investavimo į duotąją rinką patrauklumą ir sverto panaudojimo tikslingumą.

Prisotinimo svarbą finansų rinkose mums parodo sumodeliuotas pavyzdys (pav. 2). Asmuo, turintis 2 natūrinius vnt. (K_0) nuosavo kapitalo bei tiek pat nat. vnt. ir skolinto kapitalo, investuoja su svertu į atvirą rinką bei į rinkas, turinčias skirtingus prisotinimo lygius. Šiame „imitaciniame“ modelyje kintamieji parinkti taip: m=1,5, K_0=2, K_p=4, 10, 20 natūriniai vnt. (prisotinimas: 25%, 10%, 5%), j=0,2, i=0,1. Matome, kad investicijos augimo dinamika, keičiantis prisotinimo laipsniui, būna skirtinga.



Pav.2 Investavimas su svertu neprisotintoje ir prisotintoje rinkoje

Iš grafiko pastebimas skirtingas investicijos elgesys, priklausantis nuo rinkos tipo, atspindintis ekonomikos teorijos fundamentalų dėsnį, kuris nėra plačiai paplytęs, todėl kartojamas įvairiomis formomis. Pačioje pradžioje augimas būna gana vienodas, bet po kelių laiko periodų augimai išsiskiria ir prasideda nuosmukis. Matome, kad investicijos pelningumas bei rezultatas aiškiai priklauso nuo rinkos tipo (prisotinimo laipsnio), gaunami skirtingi rezultatai. Investavimo rizika užauga iškart, kai tik prisotinimo lygis padidėja. Kreivių skirtumas, grafike pavaizduotame auksčiau, parodo mums investavimo riziką.

Atliekant investavimo su svertu logistinį modeliavimą galima įvertinti investicijos pelningumą atviros (begalinės) rinkos sąlygomis ir esant prisotinimui. Didėjant prisotinimo laipsniui atsiranda ir auga investavimo rizika, kurią šiuo atveju patogu vertinti natūriniais vienetais.

Šio modelio pagalba mes galime apskaičiuoti investicijos dydį, praėjus *n* periodų. Modelio pagrindinė esmė – skaičiuoti investicijos sumą vertinant rinkos prisotinimą, bei nuspėti skolinto ir nuosavo kapitalo pelningumą. Logistiniai finansinio sverto modeliai parodė, kaip ta pati investicija, priklausomai nuo rinkos sąlygų, gali duoti visiškai priešingus rezultatus: neprisotintoje (begalinėje) rinkoje gali būti gaunamas didelis pelnas, tuo tarpu prisotintoje rinkoje – dar didesnis nuostolis.

Dmitry Kucharavy ir *Roland De Guio* rašė: „galiausiai, svarbu pabrėžti esminį labai svarbų prognozuotojų vaidmenį. Vien tik patikimo prognozavimo metodo neužtenka sudaryti tikslų ilgalaikį spėjimą. Tai niekada nebus vien mygtuko paspaudimo veiksmas. Kaip ir kiekvienas efektyvus įrankis, protingai naudojamas metodas gali duoti puikių rezultatų, tačiau nepatyrusiose rankose gali pateikti klaidingas išvadas“ [8].

III. IŠVADOS

- Logistinė analizė parodė, kad pagrindinis rizikos šaltinis investuojant su svertu yra rinkų prisotinimas ir iš to išplaukianti perprodukcija
- Atliekant investavimo su svertu logistinį modeliavimą galima įvertinti investicijos pelningumą atviros (begalinės) rinkos sąlygomis ir esant prisotinimui. Didėjant prisotinimo laipsniui atsiranda ir auga investavimo rizika, kurią šiuo atveju patogu vertinti natūriniais vienetais.
- Grafike pavaizduotas rezultatas parodo, kaip jautriai reaguojama į pelningumo pokyčius investuojant su finansiniu svertu, kai investicijos rinka yra bent dalinai prisotinta. Pasikeičia tokio investavimo sąlygos, kur skola nusveria turimą investuotą kapitalą, kas stumteli prie didelių nuostolių.
- Logistiniai finansinio sverto modeliai parodė, kad ta pati investicija, priklausomai nuo rinkos sąlygų, gali duoti visiškai priešingus rezultatus: neprisotintoje rinkoje galimas didelis pelnas, tuo tarpu prisotintoje rinkoje – dar didesnis nuostolis.

LITERATŪROS SĄRAŠAS

[1] Girdzijauskas S. Draudimas : kiekybinė finansinė analizė. Kaunas: Naujasis lankas, 104p. ISBN 9955-03-114-X. 2002.

[2] Sterman J. D. Business Dynamics, Systems Thinking and Modeling for a Complex World, 9 – S-Shape Growth: Epidemics, Innovation Diffusion, and the Growth of New Products. 950 p. ISBN 0-07-231135-5. 2000.

[3] Sornette D. Why Stock Markets Crash: Critical Events in Complex Financial Systems. Princeton University Press. 2003.

[4] Modis T. The End of the Internet Rush, Technological Forecasting & Social Change, Vol. 72, 940-941 p.p. 2005.

[5] Zadeh L. A. Fuzzy Sets. Information and Control 8, 338-353 p.p. 1965.

[6] Perrin S. M., Jason W. Yung and Jesse H. Ausubel. A Primer on Logistic Growth and Substitution: the Mathematics of the Loglet Lab Software [interaktyvus] [žiūrėta 2008 m. birželio 20 d.]. Prieiga per internetą: <http://phe.rockefeller.edu/LogletLab/logletlab.pdf>. 1999.

[7] Girdzijauskas S., Štreimikienė D. Application of Logistic Models for Stock Market Bubbles Analysis // *Journal of Business Economics and Management.* ISSN 1611-1699. Vilnius: Technika, Vol. 10, No. 1, p. 45-51, 2009.

[8] Kucharavy D. and R. De Guio. Logistic Substitution Model And Technological Forecasting. Procedia Engineering 9, 402-416 p.p, 2011. Available online at TRIZ Future Conference 2008 – Synthesis in Innovation.

[9] Girdzijauskas S., Štreimikienė D. Logistic Analysis of Business Cycles, Economic Bubbles and Crises. Chapter 3. Business Intelligence in Economic Forecasting: Technologies and Techniques. Publisher in the USA by Information Science Reference (an imprint sof IGI Global). 45-64 p.p. 2010.

[10] Girdzijauskas, S., Mialik, A., Mackevičius, R. Modeling the Fragile Economic Situations. *2nd* International Conference *on Computer Supported Education 2010*. Valencia, Spain, 7 – 10 April. ISBN 9789896740245. 2010.

[11] Girdzijauskas S. Sovereign Debt Crisis: Logistic Analysis. Vilnius: Vilniaus universiteto leidykla. ISBN 978-9955-634-52-2. 40p.p. 2011.

[12] Girdzijauskas S., Mialik A., Jociūtė E., Zujytė N. Modeling the Economic Saturation. 3rd Workshop AKTB 2011, 14th International Conference on Business Information Systems – BIS 2011 Poznan. Poland, 15 – 17 June, 2011.

[13] Girdzijauskas S., Mialik A., Dubnikovas M. Financial Leverage Trap. International Journal of Business and Social Science (IJBSS). ISSN 2219-1933 (Print), 2219-6021 (Online). 3(11) , 29-37 p.p. June 2012.

[14] Girdzijauskas S. The Logistic Theory of Capital Management: Deterministic Methods; Monograph No 1, published a peer-reviewed Supplement A to Transformations in Business & Economics, Vol. 7, No 2(14), 163 p., 2008.

[15] Strigunaite S., Kriksciuniene D. Fuzzy Expert System for Virtual Team Collaboration and Work Evaluation. BIS (Workshops) 2012: 37-43 p.p.

[16] Simutis R., Masteika S. Intelligent stock trading systems using fuzzy-neural networks and evolutionary programming methods. Solid State Phenomena, Vols. 97-98, 59-64 p.p., 2004.

Tinklalapių apsaugos priemonių įtaka atsako laikui

Dainius Mockus
Kompiuterių katedra
Kauno technologijos universitetas
Kaunas, Lietuva
dainius.mockus1987@gmail.com

Jonas Čeponis
Kompiuterių katedra
Kauno technologijos universitetas
Kaunas, Lietuva
jonas.ceponis@ktu.lt

***Santrauka*—Išaugus internetinių sistemų naudojimo populiarumui, vis daugiau slaptų, asmeninių duomenų talpinama internete. Toks reiškinys tapo viena iš esminių priežasčių, lėmusių išaugusį nusikaltimų skaičių internetinėje erdvėje. Dažnai informacijos perėmimas ar kitos kenkėjiškos operacijos atliekamos įterpiant programinį kodą į vartotojo peržiūrimą tinklapį. Siekiant apsaugoti nuo šio tipo atakų, realizuojami įvairūs apsaugos mechanizmai, tačiau jie sulėtina tinklalapio atsako laiką. Straipsnyje pristatomas tinklalapių atsako laiko tyrimas taikant skirtingus apsaugos būdus.**

Reikšminiai žodžiai—XSS, tinklalapių apsauga, atsako laikas

I. ĮVADAS

Dabar, kai informacijos ir ryšių technologijos sudaro šiuolaikinės visuomenės ir ekonomikos pagrindą, tinklų ir informacijos saugumas tampa vis aktualesnis. Prieiga prie informacinių sistemų dabar yra įmanoma iš bet kur ir bet kuriuo metu. Informacijai laisvai kertant valstybių sienas, plinta ir tinklų bei informacijos saugumo problemos.

Pagrindinis saugojimo objektas yra vertinga informacija. Pašaliniam asmeniui priėjus prie šių vertybių, kyla didelis pavojus, kad su šia informacija bus atlikti nereglamentuoti veiksmai. Todėl pirmaeiliu uždaviniu visada išliks informacijos sistemų apsauga nuo nesankcionuoto jų panaudojimo, o žemesnį prioritetą turės pačios informacijos saugumo užtikrinimas šiose sistemose.

Šiandien vyraujančios ir aktualiausios grėsmės, su kuriomis susiduria informacinės sistemos: programinio kodo įterpimas į vartotojo peržiūrimą tinklapį (angl. Cross-Site Scripting XSS) [1, 4], užgrobti kompiuteriniai tinklai [1], piktnaudžiavimo įrankiai (angl. exploit tools) [1], loginės bombos [1], paketiniai šniukštinėtojai (angl. sniffer) [1], SQL injekcijos [1, 4], Trojos arkliai [1] ir kitos kenkėjiškos priemonės. Jau daug metų iš eilės nekomercinės organizacijos OWASP (angl. Open Web Application Project) kiekvienais metais sudaromame dažniausiai internete aptinkamų atakų sąraše kodo įterpimo atakos yra aukščiausiuose pozicijose. 2013 metų plačiausiai paplitusių pažeidžiamumų dešimtuke pasaulyje XSS atakos yra trečioje vietoje [3].

Internete pradėjus saugoti daug informacijos ir teikti įvairias paslaugas, pradėjo gausėti įvairiausių kenkėjiškų atakų, naudojant įvairius būdus ir priemones. Kadangi XSS tipo atakų yra nuolat sukuriama vis naujų ir nutaikytų į vis skirtingus tinklalapių pažeidžiamumus, nuo šių atakų apsisaugoti yra ypač sunku [14]. Apsauga gali ir turi būti realizuojama ne tik tinklalapyje, bet ir vartotojo kompiuteryje [6].

II. PROGRAMINIO KODO ĮTERPIMAS Į VARTOTOJO PERŽIŪRIMĄ TINKLALAPĮ

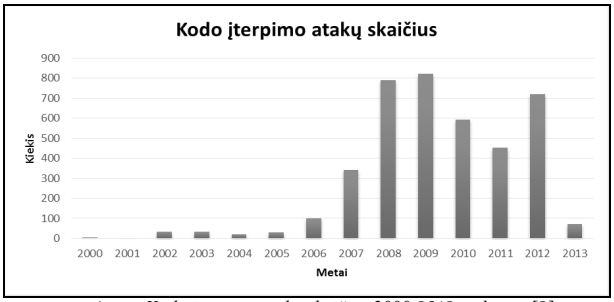
Internete veikiančios taikomosios programos tampa vis populiaresnės ir į jas keliama vis daugiau svarbios informacijos. Dažniausios atakos nukreiptos prieš internete veikiančias programas yra programinio kodo įterpimo, arba XSS tipo, atakos. Šių atakų metu bandoma įterpti į svetainę kenkėjišką kodą [1], kurį naršyklė įspraus vartotojui atsidarius svetainę. Kodo įterpimo atakos yra sunkiai sustabdomos, nes internete veikiančios programos tampa vis dinamiškesnės ir jos suteikia vartotojui vis daugiau laisvės įvairiausiems veiksmams. Kodo įterpimo atakos į tinklalapį tampa įmanomomis tuomet, kai:

1. Duomenys į sistemą yra įvedami nepatikimų šaltinių (neautorizuotų vartotojų), dažniausiai tai daroma sistemos reikalavimu.

2. Duomenys yra dinamiškai įvedinėjami vartotojų ir vėliau puslapio turinys rodomas kitiems vartotojams, neatliekant tinkamo įvedamų duomenų skenavimo. Tai gali būti komentarai, forumai, diskusijos ir pan.

XSS atakoms dažniausiai yra naudojami JavaScript programavimo kalba parašyti kodai, tačiau gali būti panaudoti ir HTML, Flash ar bet kokio kito tipo kodai, kuriuos gali įvykdyti vartotojo naršyklė [3]. Šio tipo atakos nukreiptos prieš vartotojus, o ne serverius. Užpuolikai dažniausiai renkasi patikimo turinio svetaines, kuriomis vartotojai pasitiki ir taip tampa lengviau pažeidžiami [1].

Pirmą kartą šio tipo atakos buvo pastebėtos 2000 metais ir nuo to laiko jos smarkiai išpopuliarėjo (1 pav.).



1 pav. Kodo įterpimo atakų skaičius 2000-2013 m. kovas [2]

Vis populiarėjant internetui ir vis daugiau informacijos perkeliant į sistemas internete, išpopuliarėjo ir XSS atakos [1, 3]. Pagrindinė to priežastis buvo ta, kad pirmosios šio tipo atakos nebuvo labai žalingos ir niekas nemanė, kad į šiuos pažeidžiamumus reikia atkreipti dėmesį. Atakoms vis įžūlėjant ir pridarant vis daugiau nuostolių, buvo pradėtos kurti apsaugos priemonės.

XSS atakos dažniausiai realizuojamos naudojant 3 pagrindinius būdus – nuolatinis, vienkartinis ir paremtas dokumento objektų modeliu (DOM) [1]. Nuolatinis kodo įterpimas - tai yra atakos tipas, kai kodas yra įterpiamas į puslapį, kuriame jis yra išsaugomas ir nuolat vykdomas, puslapio vartotojams peržiūrint tinklalapį. Šio tipo atakos įmanomos tik tuomet, jeigu puslapyje, prieš saugant vartotojo įvestus pranešimus ar informaciją, nėra tinkamai patikrinami įvedami duomenys [1, 5, 10]. Vienkartinės XSS atakos būdingos dinamiškai generuojamiems tinklalapių puslapiams. Ši ataka dažniausiai atliekama vartotojui pateikiant įvedimo formą, į kurią suvesti duomenys panaudojami visai kitais tikslais nei numatyta vartotojo [10]. DOM XSS pažeidžiamumas paremtas dokumento objektų modeliu. Šio tipo ataka paprastai naudojasi JavaScript [6] ar kita programavimo kalba ir keičia tinklalapio puslapio atvaizdą naršyklėje. Šis pažeidžiamumas įvyksta vartotojo kompiuteryje, o ne serveryje. Kadangi interneto naršyklės operacinėje sistemoje turi nemažai teisių, o su tokios rūšies ataka įmanoma paleisti visus procesus su tomis pačiomis teisėmis kaip ir naršyklės proceso teisės, įvykdytos atakos gali padaryti daug žalos [11].

III. Apsauga nuo kodo įterpimo (XSS) atakų

Beveik kiekvienas dabartinis tinklalapis turi dinamiško turinio formavimo dalis [6], kuriuose rašomi komentarai, bendraujama su kitais vartotojais arba tiesiog galima atlikti paiešką tame puslapyje ir po paieškos yra parodoma paieškos frazė. Visos šios funkcijos piktavaliams vartotojams suteikia galimybę bandyti įvykdyti kodo įterpimo atakas. Tai reiškia, kad kiekviename puslapyje programuotojai turi imtis apsaugos priemonių, kuriomis tinklalapio vartotojai būtų apsaugomi nuo XSS atakų. Apsaugos būdų nuo kodo įterpimo atakų yra trijų rūšių ir jie skirstomi pagal tai, kurioje vietoje yra realizuojami: serveryje, vartotojo kompiuteryje (naršyklėje) ir hibridinėje apsaugoje [3].

Vykdant apsaugą nuo kodo įterpimo atakų, serverio pusėje dažniausiai yra diegiama programinio lygio ugniasienė, kuri atliktų visus saugumo patikrinimus, perduodant informaciją tarp vartotojo kompiuterio ir serverio. Programinio lygio ugniasienės gali tikrinti užklausas, einančias iš vartotojo kompiuterio į serverį, tikrinti serverio atsakus į vartotojo užklausas arba gali atlikti šiuos abu veiksmus vienu metu. Efektyviausią apsaugą užtikrina visos, tiek įeinančios, tiek išeinančios informacijos srauto tikrinimas, tačiau tai gali labai apkrauti serverį ir sulėtinti tinklalapio atsako laiką.

Apsauga nuo kodo įterpimo atakų gali būti vykdoma ir vartotojo kompiuteryje. Tokiai apsaugai realizuoti reikalingi vartotojo kompiuteryje įdiegti tam tikri įrankiai. Tai gali būti vartotojo naršyklėje įdiegtas JavaScript kodo tikrinimo komponentas ir įsibrovimo aptikimo programinė įranga. Taip pat kaip ir serveryje, vartotojo kompiuteryje gali būti įdiegiama ir programų lygio ugniasienė, tačiau tokių priemonių įdiegimas į kompiuterį reikalauja, kad vartotojas atliktų apsaugos priemonių įdiegimą ir konfigūravimą sistemoje.

Hibridinė XSS apsauga – tai tokia apsauga, kai serveryje yra įdiegiama tam tikra programinė įranga ir taip pat yra papildoma ar modifikuojama vartotojo kompiuteryje veikianti naršyklės apsauga. Tokiu atveju dažniausiai serveryje veikianti programinė įranga atsakinga už turinio pradinį patikrinimą ir saugumo taisyklių nustatymą, o naršyklė privalo reaguoti į nusakytas taisykles ir atlikti galutinį turinio patikrinimą.

XSS pažeidžiamumas dažniausiai atsiranda dėl nepakankamo įvedamos informacijos filtravimo [15]. Radęs būdą įterpti kenksmingą kodą, piktavalis gali gauti prieigą prie tinklapio turinio, sesijos slapukų ar kitos informacijos. Beveik visas XSS galimybes sukuria tos programos, kurios nesugeba saugiai valdyti HTML įvesties ir išvesties. Prieš bet kur ir bet kada naudojant vartotojo įvestus duomenis reikia juos patikrinti, ar jie atitinka nustatytus duomenų tipus ir struktūras. Ypač dažnai pažeidimai yra susiję su HTML žymėmis laužtiniuose skliaustuose (< ir >) ir tarp kelių kitų simbolių, kurie pateikiami pirmojoje lentelėje [7].

Lentelė 1. Simboliai, kurie dažniausiai naudojami XSS atakose

Simbolis	Unikodas
"	U+0022 (34)
&	U+0026 (38)
’	U+0027 (39)
<	U+003C (60)
>	U+003E (62)

Tinklalapių turinio saugumo užtikrinimui yra sukurta įvairių apsaugos priemonių. Norint visiškai apsisaugoti nuo kodo įterpimo atakų, vartotojo kompiuteryje galima tiesiog išjungti programinio kodo vykdymą, tačiau tai labai apriboja puslapių funkcionalumą, kas yra nepatogu vartotojui.

Vien vartotojo pastangų apsisaugoti nuo kodo įterpimo atakų nepakanka. Nemaža dalis saugumo problemų kyla serverio pusėje, todėl kiekvieno tinklalapio saugumą serverio pusėje privalo užtikrinti puslapio kūrėjai [7]. Kiekviename dinaminio turinio tinklalapyje privaloma naudoti išsamų vartotojo įvedamų duomenų filtravimą, taip pat tikrinti rodomą informaciją, kartais net būtina naudoti turinio šifravimą [1]. Serverio pusėje tinklalapio apsaugai programuotojai, įvertinę visas rizikas ir pavojus, privalo pasirinkti tinkamus apsaugos lygius ir būdus. Tai nėra lengvas uždavinys, nes apsaugos būdų yra siūloma labai daug. Nuo paprasčiausių programavimo kalbos komandų (pavyzdžiui Java programavimo kalboje naudojamos funkcijos: replaceAll(), escapeXml(), regex()) iki serverio pusėje API (angl. Application programming interface) principu veikiančių tinklalapio tekstų šifravimo arba kodavimo bibliotekų. Tokių bibliotekų internete gausu. Jos kuriamos įvairioms programavimo kalboms tiek mokamos, tiek ir nemokamos. Pagrindinis šio tyrimo tikslas ir bus ištirti, kurios iš populiariausių nemokamų turinio filtravimo bibliotekų pakankamai apsaugo tinklalapį ir per daug nesulėtina tinklalapio atsako laiko.

IV. TINKLALAPIO ATSAKO LAIKO TYRIMAS

Norint atlikti testavimus pirmiausia sukurtas tinklalapis testavimui - Java programavimo kalba parašytas tinklalapis (naudojamas Spring Framework 3.0.2 karkasas). Sistema veikia vietiniame kompiuteryje (localhost).

Tinklalapis turi pavyzdinę vartotojams skirtą komentavimo formą ir jau įvestų komentarų sąrašą. Komentarų forma pasirinkta todėl, kad jie aptinkami beveik kiekviename dinaminio turinio puslapyje ir yra visiems viešai prieinami, todėl dažnai tampa piktavalių taikiniu. Komentarai saugomi duomenų bazėse. Tinklalapių testavimui naudojama Google Chrome naršyklės (versija 24.0.1312.57) programuotojams skirtas įrankis. Testuojant buvo matuojamas tinklalapio atsako laikas be jokios apsaugos ir rezultatai fiksuojami. Po to tinklalapis buvo bandomas su integruotomis saugos priemonėmis: Java programavimo kalbos komandomis, kurios aptinka ir pakeičia neleistinus atvaizduojamojo HTML kodo simbolius (escapeXML, replaceAll, regex); ESAPI ir OWASP Java HTML Sanitizer kodo tikrinimo bibliotekomis.

Eksperimentų metu buvo fiksuojamas tinklalapių atsako laikas. Atsako laikas matuojamas tiek duomenis nuskaitant iš duomenų bazės ir apdorojant juos prieš parodant vartotojui ekrane, tiek ir saugant naują vartotojo įvestą komentarą, prieš tai iš jo išfiltravus duomenis su parinktomis saugumo priemonėmis. Abu eksperimentai atliekami atskirai, kiekvieną kartą suvienodinant duomenų bazėje esančios informacijos kiekį (ištrinamas po kiekvieno eksperimento išsaugotas įrašas iš duomenų bazės, jeigu tai yra įrašo išsaugojimo laiko matavimas). Informacijos kiekis duomenų bazėje turi būti vienodas, nes apdorojamos informacijos kiekis turi įtakos atsako laikui [13]. Kuo daugiau informacijos bus bandoma atvaizduoti ekrane, prieš tai ją patikrinus, tuo daugiau laiko užims informacijos tikrinimas. Informacijos atvaizdavimo bandymams buvo naudojama OWASP projekto tinklalapyje pateikiami visų dažniausiai aptinkamų tipų XSS atakų pavyzdžiai [12], kurie, iš anksto nenaudojant jokios apsaugos, surašomi į duomenų bazę ir tuomet vykdomi atidarant testavimo tinklalapį. Su kiekvienu iš apsaugos variantų testavimai atliekami po 10 kartų ir apskaičiuojamas aritmetinis laiko verčių vidurkis, kas ir bus laikomas to apsaugos būdo atsako laikas. Gautus rezultatus palyginame, taip nustatant, kurios bibliotekos ir kaip keičia tinklalapio atsako laiką.

Matuoti tik atsako laiką, kad galėtume spręsti, jog tai yra geriausia apsauga, būtų klaidinga. Bibliotekos, kurios atlieka tik neleistinų simbolių paiešką ir jų pašalinimą arba konvertavimą į leistinus HTML simbolius, visada veiks greičiau už sudėtingus tekstų filtravimus atliekančias bibliotekas, tačiau jos negali tinkamai užtikrinti saugumo tinklalapiuose. Galima lengvai nuspėti, kad programa, kurioje nenaudojama jokia papildoma biblioteka ir apsaugai panaudotos tik programavimo kalbos komandos, veiks greičiausiai. Būtent dėl visų šių priežasčių būtina įvertinti kiekvienos bibliotekos patikimumą saugumo prasme. Pats paprasčiausias būdas patikrinti bibliotekos patikimumą yra apžvelgti jos funkcionalumą ir taip bus galima įvertinti, nuo kurio tipo atakų padeda apsisaugoti analizuojama biblioteka.

Java programavimo kalba parašytas tinklalapis buvo testuojamas su šiomis tinklalapio kodo filtravimo priemonėmis:

- Nenaudojant jokio filtravimo.
- Filtruojant Java programavimo kalbos komandomis (escapeXml, replaceAll, regex) [9].
- ESAPI [5].
- OWASP Java HTML Sanitizer – OWASP projekto sukurtas HTML kodo analizatorius, parašytas Java programavimo kalba [3].

Testavimui buvo pasirinktos OWASP projekto kuriamos ir plėtojamos bibliotekos. Šios bibliotekos kuriamos ir patariamos naudoti OWASP projekto dalyvių ir kūrėjų, kurie aktyviai dalyvauja kovojant su kodo įterpimo atakomis [3, 5]. Apsauga realizuota programavimo kalbos komandomis escapeXml, replaceAll ir regex į eksperimentą yra įtraukta siekiant pademonstruoti, kad nors vien paprastais metodais realizuota minimali apsauga mažai keičia atsako laiką, tačiau ji nėra pakankama.

Testavimo metu matavimai atliekami nuosekliai, vienas po kito. Atliekant testavimus, kompiuteris papildomų uždavinių neatliko. Testavimui naudojamo kompiuterio aparatūrinės ir programinės įrangos sąrašai pateikti žemiau esančiose lentelėse.

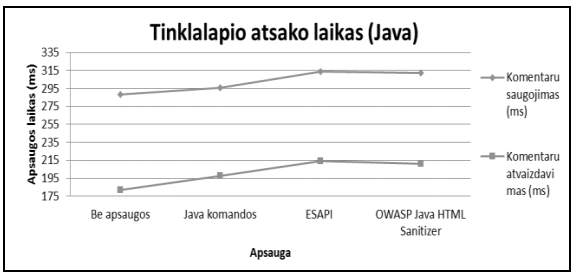
LENTELĖ 2. TESTAVIMUI NAUDOJAMO KOMPIUTERIO APARATINĖ ĮRANGA

Procesorius:	Intel(R) Core(TM) i7-2600 3,4 GHz
Spartinančioji atmintis:	DDR3 8,00 GB
Vaizdo plokštė:	NVIDIA GeForce GTX 560
Kietasis diskas:	SAMSUNG SATA3, Talpa: 1TB

LENTELĖ 3. TESTAVIMUI NAUDOJAMO KOMPIUTERIO PROGRAMINĖ ĮRANGA

Operacinė sistema:	Windows 7 Professional 64-bit
Interneto naršyklė:	Google Chrome 24.0.1312.57 versija
Antivirusinė programa:	(nėra)
Serverio programinė įranga:	Apache Tomcat v7.0 (paleidžiamas ir veikiantis kartu su Eclipse Java EE IDE for Web Developers. Versija: Indigo Service Release 2) WampServer Version 2.2 pakete naudojama duomenų bazė: MySQL 5.5.24 projekte naudojamas Spring Framework 3.0.2 karkasas

Atliekant testavimus atsako laikas buvo matuojamas, naudojantis Google Chrome naršyklėje integruotomis tinklalapių programuotojų priemonėmis. Matavimai buvo atliekami nuosekliai, vienas po kito. Testavimo metu kompiuteris papildomų uždavinių neatliko, kad testavimo rezultatai nebūtų vėlinami pašalinių veiksnių. Atlikus atsako laiko testavimus su visomis anksčiau išvardintomis priemonėmis, buvo gauti rezultatai, kurie pateikiami 2 paveiksle.



2 pav. Tinklalapio atsako laiko matavimų rezultatai Java tinklalapyje

Iš gautų rezultatų galima pastebėti, kad tinklalapio atsako laiką labiausia didino ESAPI biblioteka, o mažiausiai – programavimo kalbos komandos. Iš nagrinėtų saugumo bibliotekų greičiausiai veikė OWASP Java HTML Sanitizer.

Gavus atsako laikų testavimo rezultatus, būtina įvertinti bibliotekų saugumo užtikrinimo lygius. Pagal saugumo funkcionalumą sudaryta ketvirta lentelė, iš kurios bus galima įvertinti priemonių užtikrinamą saugumo lygmenį.

LENTELĖ 4. JAVA PROGRAMAVIMO KALBOS HTML KODO TIKRINIMO BIBLIOTEKŲ FUNKCIONALUMO ĮVERTINIMAS

Galimi pažeidimai / Biblioteka	Koduotės simbolių filtravimas	HTML kodo formatavimas (užtikrinimas, kad žymos uždarytos ir pan.)	Atributų filtravimas	Kodo įterpimo skenavimas
Java komandos	Dalinai	Ne	Ne	Dalinai
ESAPI	Taip	Ne	Taip	Taip
OWASP Java HTML Sanitizer	Taip	Taip	Taip	Taip

Pateiktoje Java HTML kodo filtravimo priemonių saugumo įvertinimo lentelėje matome, kad geriausią saugumo lygmenį užtikrina OWASP Java HTML Sanitizer. Ši biblioteka taip pat yra geriausia ir atsako laiko prasme, atmetant Java programavimo kalbos komandas, kurios neužtikrina patikimo saugumo lygio. Pagal gautus rezultatus galima teigti, kad Java programavimo kalba kuriamuose tinklalapiuose geriausia naudoti OWASP projekto sukurtą priemonę OWASP Java HTML Sanitizer.

V. IŠVADOS

Tyrimo metu buvo sukurtas Java programavimo technologijas naudojantis tinklalapis. Naudojantis šiuo tinklalapiu, ištestuotos pasirinktos apsaugos priemonės ir išmatuotas tinklalapio atsako laikas. Buvo stebima, kaip keičiasi atsako laikas taikant skirtingas saugumo priemones sukurtame tinklalapyje.

Testavimo rezultatai parodė, kad mažiausiai atsako laikus didina programavimo kalbos komandos, tačiau jos negali užtikrinti pakankamo tinklalapių saugumo. Atsirenkant geriausius rezultatus demonstravusias apsaugos priemones buvo nustatyta, kad mažiausiai atsako laikus įtakoja: OWASP Java HTML Sanitizer.

Ištestavus tinklalapių atsako laikus su skirtingomis priemonėmis, buvo būtina išsiaiškinti, ar tos apsaugos priemonės, kurios veikia greičiausiai, tikrai gali užtikrinti tinkamą apsaugos lygį. Išanalizavus saugumo priemones, išsiaiškinus jų funkcijas, buvo nustatyta, kad geriausią saugumo lygį užtikrina OWASP Java HTML Sanitizer.

Pagal turimus tyrimo rezultatus galima teigti, kad Java tinklalapiams pakankamą saugumo lygį užtikrinanti priemonė, kuri, palyginus su kitomis nagrinėtomis saugumo priemonėmis, mažiausiai padidina atsako laiką - OWASP Java HTML Sanitizer.

Tolimesniuose tyrimuose numatoma atlikti XSS atakų apsaugos priemonių tyrimą tinklalapiuose, sukurtuose naudojant PHP ir ASP.NET programavimo technologijas.

LITERATŪRA

[1] J. Grossman, R. Hansen, P. D. Petkov, A. Rager, S. Fogie, XSS Attacks. Cross site scripting exploits and defense, United State of America, 2007.

[2] Nacionalinis standartų ir technologijų instituto (angl. National Institute of Standarts and Technology - NIST) statistiniai duomenys apie aptiktas XSS atakas. Paskutinis apsilankymas: 2013.02.26. http://nvd.nist.gov/.

[3] OWASP atvirojo kodo internetinių aplikacijų saugumo tyrimo projektas. Apsaugos būdai ir priemonės nuo XSS tipo atakų. Dažniausiai pasitaikančių atakų reitingavimas ir aprašymas. Paskutinis apsilankymas: 2013.02.26. https://www.owasp.org/index.php/XSS_(Cross_Site_Scripting)_Prevention_Cheat_Sheet

[4] Adam Kiezun, Philip J. Guo, Karthick Jayaraman, Michael D. Ernst, Automatic Creation of SQL Injection and Cross-Site Scripting Attacks, 2009.

[5] Yi Wang, Zhoujun Li, Tao Guo, Program Slicing Stored XSS Bugs in Web Application, Beijing, China, 2011.

[6] A. E. Nunan, E. Souto, E. Feitosa. Automatic classification of cross-site scripting in web pages using document-based and URL-based features. ISCC (p. 702-707). IEEE. ISBN: 978-1-4673-2712-1, 2012.

[7] Lwin Khin Shar, Hee Beng Kuan Tan, Automated removal of cross site scripting vulnerabilities in web applications, Nanyang Avenue, Singapore, 2011.

[8] OWSAP ESAPI bibliotekos, skirtos apsaugai nuo XSS atakų, internetinis tinklalapis. Paskutinis apsilankymas: 2013.02.26. https://www.owasp.org/index.php/ESAPI

[9] Java programavimo kalbos bibliotekų aprašymas ir dokumentacija. Paskutinis apsilankymas: 2012.02.26. http://docs.oracle.com/javaee/5/jstl/1.1/docs/tlddocs/fn/escapeXml.fn.html

[10] R. Putthacharoen, P. Bunyatnoparat, Protecting Cookies from Cross Site Script Attacks Using Dynamic Cookies Rewriting Technique, Technology Ladkrabang, Bankok Thailand, 2011.

[11] A. Klein, DOM Based Cross Site Scripting or XSS of the Third Kind, 2005. Straipsnis internete: http://www.webappsec.org/projects/articles/071105.shtml.

[12] OWASP XSS atakų pavyzdžiai kuriamų tinklalapių testavimui. Paskutinis apsilankymas: 2013.04.07. https://www.owasp.org/index.php/XSS_Filter_Evasion_Cheat_Sheet#Tests

[13] Lwin Khin Shar, Hee Beng Kuan Tan, Defending against Cross-Site Scripting Attacks, Nanyang Avenue, Singapore, 2012.

[14] Mario Heiderich, Marcus Niemietz, Felix Schuster, Thorsten Holz, Jörg Schwenk. Scriptless attacks: stealing the pie without touching the sill. In Proceedings of the ACM conference on Computer and communications security (CCS '12). 2012.

[15] Theodoor Scholte, William Robertson, Davide Balzarotti, Engin Kirda. Preventing Input Validation Vulnerabilities in Web Applications through Automated Type Analysis. In Proceedings of the IEEE 36th Annual Computer Software and Applications Conference (COMPSAC '12). 2012.

Piršto kraujagyslių tinklo taikymo slaptų raktų generavimui galimybių tyrimas

Povilas Nanevičius
Informatikos Fakultetas
Kauno Technologijos universitetas
Kaunas, Lietuva
povilas.nanevicius@stud.ktu.lt

doc. dr. Algimantas Venčkauskas
Informatikos Fakultetas
Kauno Technologijos universitetas
Kaunas, Lietuva
algimantas.venckauskas@ktu.lt

Santrauka. **Biometrinės sistemos, skirtos slaptų raktų generavimui iš piršto kraujagyslių tinklo, nenaudojant palyginimo su iš anksto išsaugotu šablonu, galimybių tyrimas. Tokia sistema leistų naudoti vartotojo piršto kraujagyslių atvaizdą nesudėtingiems raktams generuoti be originalaus šablono išsaugojimo ir palyginimo. Darbe naudojamos tradicinėse (palyginimo) sistemose taikomos piršto kraujagyslių tinklo išskyrimo funkcijos ir papildomos raktų generavimo funkcijos. Bandymo metu iš pavyzdinio kraujagyslių tinklo atvaizdo, apdoroto Miura „Pasikartojančių linijų sekimo“ algoritmu ir morfologinėmis funkcijomis, pritaikius darbe siūlomą „Kontūro sekimo iteracijų skaičiaus“ metodą, gautas 12 simbolių ilgio slaptas raktas.**

Reikšminiai žodžiai — piršto kraujagyslių tinklas; raktų generavimas; biometrija

I. Įvadas

Piršto kraujagyslių atvaizdų skenavimas yra vienas iš naujausių, bet jau plačiai taikomų biometrinių vartotojų autentifikavimo metodų. Lyginant su viena dažniausiai biometrijoje naudojamų savybių – piršto antspaudais – kraujagyslių tinklas yra saugesnis, nes jis yra nematomas, nuskaitomas tik tol, kol pirštu teka kraujas [1]. Taip pat kraujagyslių išskyrimo algoritmų bendroji klaidos tikimybė EER=|FAR-FRR| (angl. Equal Error Rate), kur FAR yra klaidingo priėmimo tikimybė (angl. False Acceptance Rate), o FRR – klaidingo atmetimo tikimybė (angl. False Rejection Rate) dažnai yra mažesnė, lyginant su piršto antspaudų autentifikavimo sistemų EER. Miura „Pasikartojančių linijų sekimo“ algoritmo EER = 0,145%, o skirtingų piršto antspaudus naudojančių sistemų EER rodikliai svyruoja tarp 0,2 ir 4% [2]. Piršto kraujagyslių taikymas autentifikacijai patogus ir dėl to, kad, lyginant su kai kuriais kitais biometriniais autentifikavimo metodais, atvaizdo gavimui mažiau įtakos turi tokie aplinkos veiksniai kaip purvas, piršto ar skaitytuvo paviršiaus drėgnumas ir kt. [1].

Daugelio šiuolaikinių piršto kraujagyslių tinklą autentifikacijai naudojančių sistemų veikimas paremtas vartotojo pateikto kraujagyslių tinklo žemėlapio palyginimu su iš anksto išsaugotu pavyzdžiu arba tam tikru būdu suformuotu šablonu. Plintant sistemoms, naudojančioms piršto kraujagyslių tinklą autentifikacijai, iškyla keletas naujų problemų: to paties kraujagyslių tinklo naudojimas skirtingos sistemose kelia abejonių dėl tokių sistemų saugumo – kraujagyslių tinklas tampa saugotina savybe. Norint lyginti atvaizdus su šablonais reikia brangesnės nuskaitymo įrangos su daugiau vidinės atminties arba tenka siųsti nuskaitytus atvaizdus į centrinį serverį – tai gali sukurti papildomą tinklo srautą ir padidinti galimų sistemos pažeidžiamumų kiekį.

Šiame darbe aptariamos slapto pastovaus rakto generavimo galimybės iš piršto kraujagyslių tinklo, nuskaitant laikinai saugomą iš dalies kintantį piršto kraujagyslių tinklo žemėlapį. Tokios sistemos sugeneruoti raktai priklausytų nuo naudojamo generavimo algoritmo ir skirtingose autentifikavimo sistemose būtų unikalūs. Kadangi nuskaitytas piršto kraujagyslių tinklo žemėlapis, sugeneravus slaptą raktą, būtų iš karto pašalinamas, iš sistemoje saugomų raktų nebūtų galima atkurti pradinio kraujagyslių tinklo savybių.

Viena iš pagrindinių problemų, susijusių su tiesioginiu slaptų raktų generavimu, yra ribotas galimų raktų ilgis ir unikalumas. Taip pat, lyginant su palyginimo sistemomis, tiesioginio generavimo sistemų veikimą galėtų labiau trikdyti nežymios kraujagyslių tinklo nuskaitymo ir pirminio apdorojimo paklaidos.

Šiame darbe aptariamos kelios skirtingos raktų generavimo funkcijos ir jų savybės bei išbandomas rakto generavimo, skaičiuojant kontūro sekimo iteracijas, metodas, kuris iš pateikto kraujagyslių tinklo generuoja 12 skaitmenų ilgio raktą.

II. Atvaizdo gavimas

A. Piršto kraujagyslių tinklo nuskaitymas

Piršto kraujagyslių tinklas yra nuskaitomas padedant pirštą tarp infraraudonųjų spindulių šaltinio (skirtingose sistemose naudojamas bangos ilgis svyruoja tarp 750 iki 950 nm) ir kameros, fiksuojančios vaizdą. Kraujyje esantis hemoglobinas sugeria infraraudonąją spinduliuotę, todėl atvaizde kraujagyslės atrodo kaip tamsesnės linijos [3]. Dėl kraujagyslių geometrijos, kraujagyslių centrai yra tamsesni, o juos supantys kraštai tolygiai šviesesni. Kraujagyslės matomos kaip tamsūs slėniai.

Tarp komercinių sistemų, skirtų pirštų kraujagyslių tinklo atvaizdams gauti ir apdoroti, dominuoja Hitachi, Fujitsu, M2SYS, Sony ir kitų gamintojų produktai.

Šiame darbe aptariamoms raktų generavimo funkcijoms analizuoti naudojamas pavyzdinis piršto kraujagyslių atvaizdas [4], pateiktas 1 paveiksle (a), gautas naudojant 830 nm bangos ilgį. Vėlesniuose darbuose numatoma atlikti bandymus su didesnėmis pirštų kraujagyslių duomenų bazėmis.

B. *Pirminis kraujagyslių išskyrimas*

Kraujagyslių atvaizdas bandomoje sistemoje yra apdorojamas Miura „Pasikartojančių linijų sekimo“ funkcija [2]. Ši funkcija, pradėdama nuo atsitiktinių taškų atvaizde, ieško tamsesnių plotų, apsuptų tolygiai šviesesnio fono. Nuo surastų taškų atliekamas linijų sekimas tamsiais kontūrais (slėniais). Nuolatinis aplinkinių plotų šviesumo tikrinimas leidžia sumažinti atsitiktinai sekamų triukšmų įtaką galutiniam rezultatui ir efektyviai pašalinti dėl kaulų ir sąnarių struktūros atvaizde atsirandančius šešėlius. Priklausomai nuo pasirinkto sekimo iteracijų skaičiaus (sekimo kartų), funkcija gali praeiti per aptiktas tamsios spalvos linijas vieną ar keletą kartų. Tikėtina, kad taškai, per kuriuos linijų sekimo algoritmas praeina keletą kartų, yra kraujagyslės [3].

Taip pat darbe buvo išbandyti Miura „Maksimalaus linkio“ bei Huang „Plačių linijų sekimo“ [5] algoritmai, kurių veikimas daugeliu atveju buvo greitesnis, leido pasiekti geresnių rezultatų apdorojant blogesnės kokybės atvaizdus, bet gautas tinklas buvo labiau tinkamas palyginimo operacijoms nei numatomam linijų sekimui atlikti. Priklausomai nuo sekamų taškų skaičiaus, „Pasikartojančių linijų sekimo“ algoritmas gali būti vienas iš lėčiausių, tačiau jo rezultatas pastoviausias.

Darbe naudotų Matlab ir Lee „Piršto ploto nustatymo“ funkcijų realizacijos buvo atliktos Bram Ton [4]. Lee „Piršto ploto nustatymo“ funkcija naudojama riboms, kuriose atliekamas apdorojimo funkcijų sekimas, išskirti.

C. *Morfologinis kraujagyslių tinklo apdorojimas*

Siekiant supaprastinti ir papildomai pataisyti gautus kraujagyslių tinklų atvaizdus, daugelyje kraujagyslių autentifikaciją naudojančių biometrinių sistemų taikomos įvairios morfologinių funkcijų aibės. Šiame darbe naudotos morfologinės funkcijos yra standartinės Matlab paketo funkcijos, tačiau jų aibės pasirinkimas ir taikymo seka yra svarbūs žingsniai siekiant suformuoti norimą kraujagyslių tinklo atvaizdą.

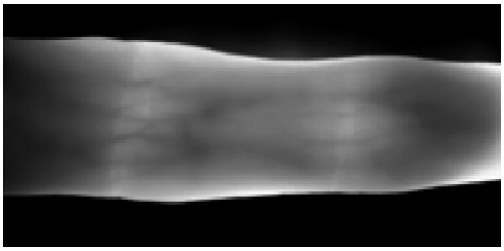
Miura „Pasikartojančių linijų sekimo“ funkcijos išvestyje (2 pav. atvaizdas dešinėje) yra daug pavienių taškų, nedidelių neužpildytų plotų, su kitomis kraujagyslėmis nesusijungiančių atšakų ir kitokių triukšmų. Taip pat ryškiausiai matomos arba tam tikrose vietose esančios kraujagyslės gali būti sekamos daugiau kartų, todėl atrodo platesnės.

Siūlomiems raktų generavimo metodams reikalingas vieno pikselio pločio kraujagyslių tinklas, todėl pasirinkta tokia morfologinių funkcijų aibė: pavienių taškų pašalinimas (*clear*[inf]), vieno pikselio dydžio plotų, visiškai apsuptų priešingos vertės elementais, invertavimas (*fill*[1]), vieno pikselio pločio linijų trūkių sujungimas (*bridge*[inf]), pakartotinė morfologinė „uždarymo“ funkcija (*close*[1]), pikselių invertavimas, jei penki ar daugiau supantys pikseliai yra priešingos vertės 3x3 elemente (*majority*[3]), „genėjimo“ funkcija, pašalinanti vienoje pusėje neprisijungusius elementus (*spur*[inf]), objektų susiaurinimas iki vieno pikselio pločio linijos, neleidžiant objektams išsiskirti (*skel*[inf]).

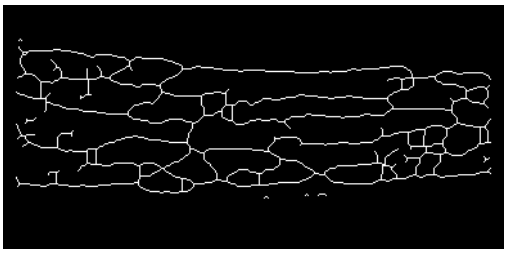
Kai kurie tarpiniai taškai (susijungimai ar žiedai tinkle) tokio apdorojimo metu gali būti prarandami. Taikant kitokius raktų generavimo metodus morfologinių funkcijų aibė gali būti koreguojama [6][7].

D. *Galutinis atvaizdas*

Galutinis atvaizdas po pirminio ir morfologinių funkcijų apdorojimo yra [0,1] verčių matrica, kurioje linijomis išskirtos piršto kraujagyslės. Galutinio atvaizdo pavyzdys pateikiamas 1 paveiksle (b).

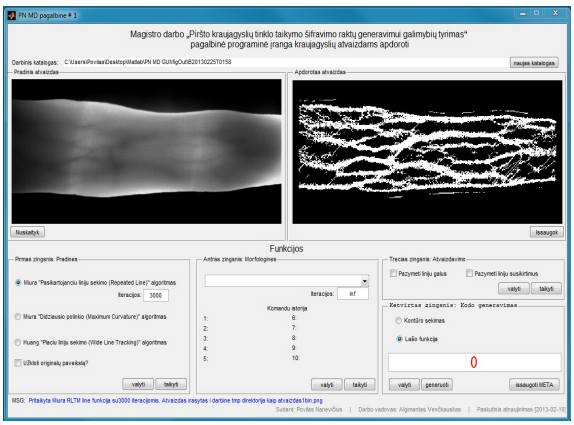


(a) Neapdorotas piršto kraujagyslių tinklo atvaizdas



(b) Galutinis apdorotas kraujagyslių tinklo atvaizdas

1 paveikslas. Pradinis (a) ir galutinis (b) piršto kraujagyslių atvaizdai.



2 paveikslas. Pagalbinė Matlab grafinė sąsaja pirminių, morfologinių ir rakto generavimo funkcijų taikymui ir atvaizdavimui.

III. RAKTO GENERAVIMO ALGORITMAI

Tradicinėse autentifikavimo sistemose atvaizdas, apdorotas pirminėmis arba pirminėmis ir morfologinėmis funkcijomis,

galėtų būti lyginamas su iš anksto išsaugotu šablonu. Tačiau šiame darbe, užuot atlikus palyginimą, atvaizdas toliau apdorojamas, siekiant išskirti reikšminius požymius ir sugeneruoti pastovų šifravimo raktą [8]. Šiame skyriuje aptariamos siūlomos funkcijos ir metodai raktui generuoti.

A. Reikšminių koordinačių nustatymo funkcija

Tolesnis skaitinių verčių gavimas iš kraujagyslių tinklo gali būti pagrįstas kontūrų sekimu. Tokiam metodui svarbu „žinoti“ reikšmines tinklo koordinates: linijų pradžių ir pabaigų koordinates, linijų susikirtimo taškus, pradinį sekimo tašką ir kt. Šioms reikšminėms koordinatėms nustatyti darbe naudojamos matematinės morfologijos sankryžų ir galų nustatymo funkcijos. Funkcijų rezultatas yra sankryžų, pradžios ir pabaigos taškų koordinačių matricos.

B. Kontūro sekimo funkcija

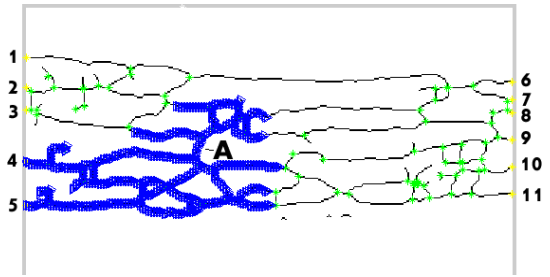
Antras žingsnis pasirinktame tinklo apdorojimo metode yra kontūro sekimas. Nuo sekimo funkcijos realizacijos priklauso ir galutinis kontūro sekimo rezultatas. Darbui realizuota kontūro sekimo funkcija, sekanti tašką iki susikirtimo. Po susikirtimo aptikimo visos linijos toliau sekamos vienu metu (3 pav.). Alternatyvi sekimo funkcija galėtų sekti tik vieną iš šakų iki galutinio taško, atsitiktinai pasirinkdama kelią per sankryžas arba taikydama tam tikras ėjimo per sankryžas taisykles. Šie funkcijos veikimo aspektai turėtų įtakos galutiniam rezultatui ir jo maksimaliam kardinalumui. Tolesnėse darbo stadijose numatoma ištirti kitokių kontūro sekimo funkcijų realizacijų rezultatus.

3 paveikslas. kraujagyslių tinklo atvaizdas, uždengtas sluoksniu, paryškinančiu sekamą kontūrą. Paryškinta dalis yra lygi 100 sekamo kontūro taškų, vaizdas padidintas.

Realizuota kontūro sekimo funkcija „slenka“ per kraujagyslių tinklą ir priklausomai nuo jo formos, sudėtingumo ir kitų savybių, pasiekia skirtingus kraujagyslių tinklo taškus (galus ar susikirtimus) skirtingais santykiniais laiko momentais. Išvestiniai kontūro sekimo funkcijos veikimo parametrai darbe naudojami nesudėtingam raktui generuoti.

C. Kontūro sekimo iteracijų skaičiaus metodas

Šio metodo idėja yra tam tikrų, iš anksto numatytų, verčių priskyrimas kiekvienam kontūro pradžios ir pabaigos taškui ir vertinimas, kada šiuos taškus pasiekia pasirinkta kontūro sekimo funkcija. Pavyzdyje (4 pav.) kiekvienam iš kraujagyslių tinklo galų priskiriamos skaitinės vertės nuo 1 iki 11. Pasirenkamas pradinis sekimo taškas (vertė 4) ir atliekamas sekimas. Pavaizduotas sekimo funkcijos rezultatas, kai pasiekiamas pirmasis „išėjimas“, atlikus 826 taškų kontūre sekimą.



4 paveikslas. Kraujagyslių tinklo kontūro sekimas pradedant nuo taško nr. 4.

- Kiekvienam pradžios ir pabaigos taškui priskirtos skaitinės vertės.

- Paryškinta sekama kontūro atkarpa po 826 kontūro sekimo algoritmo iteracijų.

- Kritinis kontūro taškas pažymėtas raide A.

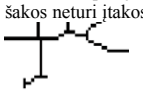
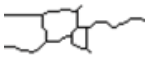
Atlikus visą sekimo ciklą, metodo rezultatas – slaptas raktas, sudarytas iš skaičių 532111910678. Skaitmenys šiame skaičiuje yra lygūs po tam tikro sekimo iteracijų skaičiaus pasiekto tinklo įėjimo/išėjimo skaitinei vertei. Algoritmo generuojamas raktas priklausytų nuo įėjimams/išėjimams priskirtų pradinių verčių ir kontūro sekimo funkcijos realizacijos. Šiame pavyzdyje naudojama III. B skyriuje aprašyta kontūro sekimo funkcija. Dalis verčių, dėl to, kad jos pasiekiamos labai artimais laiko momentais ir, atsiradus netikslumams kontūro sudaryme (pavyzdžiui netiksliai atlikus piršto lokalizaciją), galėtų įtakoti galutinį rezultatą, gali būti atmetamos. Numatomas patikimas galutinis rakto ilgis galėtų būti apie 4-8 skaitmenys.

Šio metodo privalumas – santykinai maža paklaida, jei kraujagyslių tinkle yra nedidelių neatitikimų. Metodui reikšminės įtakos neturėtų linijų krypties ar vietos pasikeitimai, atsiradę pašaliniai prie pagrindinio sekamo tinklo neprisijungę elementai, žiedai, sudėtingos geometrinės formos kontūro viduje, atsišakojimai ir netiksliai nustatytos kraujagyslių sankryžų vietos.

Daugiausiai neigiamos įtakos šiam metodui galėtų turėti netinkamai nustatyti (arba nenustatyti) kritiniai linijų susijungimai. Metodo generuojamo rakto vertė galėtų tapti netiksli, jeigu kontūre atsirastų trūkių, dėl kurių kai kurie pradžios arba pabaigos taškai taptų nepasiekiami arba pasiekiami po reikšmingai daugiau arba mažiau kontūro sekimo iteracijų. Pavyzdinio kritinio kontūro taško pavyzdys 3 paveiksle pažymėtas simboliu A. Jei šios jungiančiosios

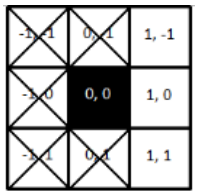
kontūro dalies nebūtų, viršutinė kraujagyslių tinklo dalis būtų pasiekiama tik per kur kas tolesnį jungiamąjį tašką tarp išėjimų 8 ir 9 ir reikšmingai pakeistų galutinį generavimo rezultatą. Taip pat metodui neigiamos įtakos turėtų sudėtingas kraujagyslių tinklas kontūro galuose, ypač jei piršto lokalizacija būtų netiksli.

1 lentelė. „Kontūro sekimo iteracijų skaičiau metodo“ privalumai ir trūkumai.

Privalumai	Trūkumai
Atskiros neprisijungusios tinklo šakos neturi įtakos rezultatui. 	Susijungimai arba trūkiai reikšminėse kraujagyslių atvaizdo tinklo vietose (pvz., tarp arti viena kitos esančių linijų) gali visiškai keisti rezultatą.
Žiedai ir sudėtingi išsišakojimai kraujagyslių tinkle turi mažai įtakos rezultatui. 	

D. „Vandens lašo“ metodas

Kontūro sekimo iteracijų skaičiaus ir kitų kontūro sekimo funkcija paremtų metodų veikimui naudojamas visas piršto kraujagyslių kontūras. Tokiems algoritmams įtakos galėtų turėti bet kokie tinkle atsiradę netikslumai. Priešingai nei daugiakrypčiai algoritmai, siūlomas „vandens lašo“ metodas yra vienakryptis. Šiame metode naudojamas kraujagyslių kontūro sekimas vyksta tik išėjimo kryptimi ir algoritmui leidžiama sukti tik +/- 45 laipsnių kampu išėjimo link (5 pav.).

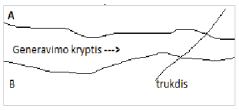


5 paveikslas. Galimos „Vandens lašo“ metode naudojamos kontūro sekimo funkcijos kryptys, kai kontūras sekamas iš kairės į dešinę

Toks algoritmas padėtų išvengti neigiamos tam tikra kryptimi atsiradusių netikslumų ar susijungimų/nesusijungimų reikšminėse kraujagyslių tinklo vietose įtakos galutiniam rezultatui, pavyzdžiui, sekant kontūrą, pavaizduotą 4 pav., nuo 1, 2, 3 arba 5 kairėje pusėje esančių pradžios taškų, reikšminis sujungimas A įtakos galutiniam rezultatui neturėtų.

Tačiau priklausomai nuo išpildymo būdo, galutinis vienakrypčio metodo generuojamo rakto ilgis galėtų būti trumpesnis nei naudojant daugiakrypčius generavimo metodus. Taip pat atliekant sekimą priešinga nei numatyta kryptimi, sekimo netikslumai galimai būtų didesni nei naudojant daugiakryptį kontūro sekimą.

2 lentelė. „Vandens lašo“ metodo privalumai ir trūkumai.

Privalumai	Trūkumai
Algoritmui įtakos turėtų tik tam tikra kryptimi atsiradę netikslumai. 	Algoritmas negalėtų įveikti T formos arba tinklo vietų, kur linijos sukasi atgal ir negeneruotų dalies rezultato.

E. Papildomų reikšminių koordinačių vertinimas rakto generavimo algoritmuose

Siekiant generuoti ilgesnius slaptus raktus, anksčiau aptarti generavimo algoritmai gali būti papildyti reikšminių taškų fiksavimo funkcijomis. Pavyzdžiui, kontūro sekimo algoritmui „slenkant“ per kraujagyslių kontūrą, galima fiksuoti pakeliui pereitų sankryžų skaičių arba matuoti santykinius atkarpų, per kurias atliekamas kontūro sekimas, ilgius. Toks papildomų reikšminių koordinačių fiksavimas leistų padidinti generuojamų raktų ilgį kelis kartus, tačiau raktų generavimo metodai taptų labiau priklausomi nuo tikslių pirminių funkcijų rezultatų.

IV. Išvados

Šiame darbe aptariamos pirštų kraujagyslių tinklo tinkamumo slaptų raktų generavimui, nenaudojant tradicinių palyginimo funkcijų, galimybės, pristatomos kelios pagrindinės pirminio kraujagyslių tinklo apdorojimo, morfologinio išskyrimo funkcijos ir aptariami galimi slapto rakto generavimo metodai. Kituose darbuose numatoma realizuoti papildomus rakto generavimo metodus, atlikti algoritmų bandymus su skirtingomis pirminių ir morfologinių funkcijų aibėmis ir pateikti išvadas apie siūlomų metodų tinkamumą slaptų raktų generavimui.

Šaltiniai

[1] Alginamtas Venčkauskas, Egidijus Kazanavičius. „Informacinių technologijų saugos metodai“, 2011, Kaunas.

[2] N. Miura, A. Nagasaka ir T. Miyatake, „Automatic Feature Extraction from non-uniform Finger Vein Image and its Application to Personal Identification“ IAPR Workshop on Machine Vision Applications, Dec. 11 - 13.2002, Nara- ken New Public Hall, Nara, Japan

[3] N. Miura, A. Nagasaka ir T. Miyatake „Feature extraction of finger vein patterns based on repeated line tracking and its application to personal identification“, Machine Vision and Applications, Volume 15, Number 4 (2004), pp. 194—203

[4] Bram Ton, “Miura et al. vein extraction methods”, http://www.mathworks.com/matlabcentral/fileexchange/35716-miura-et-al-vein-extraction-methods, žiūrėta 2013-02-20

[5] Beining Huang, Yanggang Dai, Rongfeng Li, Darun Tang and Wenxin Li, „Finger-vein Authentication Based on Wide Line Detector and Pattern Normalization“, 2010 International Conference on Pattern Recognition

[6] Petros Maragos,Ronald W. Schafer,Muhammad Akmal Butt „Mathematical Morphology and its applications to image and sygnam processing“, Kluwer Academic Publishers, 1996

[7] Henk J. A. M. Heijmans „Mathematical morphology: a modern approach in image processing based on algebra and geometry“ Society for Indostrial and Applied Mathematics, Vol.37, No.1,pp 1-36, March 1995

[8] Wende Zhung, Yao-Jen Chang, and Tsuhun Chen “Biometric-Based cryptographic Key Generation” 2004 IEEE International Conference on Multimedia and Expo (ICME)

Rinkos nuotaikų techniniais indikatoriais grindžiamų prekybos strategijų pelningumo tyrimas

Anton Ogorodov
Vilniaus Universitetas
Kauno Humanitarinis fakultetas, Informatikos katedra
Kaunas, Lietuva
anton.ogorodov@gmail.com

***Santrauka*. Straipsnyje siekiama supažindinti su faktoriais turinčiais įtaką investuotojų nuotaikai. Išskirti techniniai indikatoriai galintys nurodyti optimistinį arba pesimistinį investuotojų nusiteikimą. Atliktas indikatorių pelningumo tyrimas lyginant indikatorius su S&P 500 indeksu.**

Reikšmingi žodžiai: indeksai, investuotojų nuotaika, strategija.

I. ĮVADAS

Klasikinėje investavimo teorijoje nėra vietos investuotojų nuotaikoms - racionalūs investuotojai priima sprendimus remdamiesi tik fundamentalia analize. Vis dėlto jausmai, nuojauta ir nuotaika visada lydi investuotojų sprendimus. Investuotojų nuotaika daro stiprią įtaką akcijų kainoms ir rinkų svyravimams [1]. Sprendimai priimti, remiantis tik fundamentalia analize, laikomi labiau klaidingais negu sprendimai, priimti remiantis techniniais indikatoriais ar investuotojo nuosava investavimo strategija. Nebekyla klausimas, ar investuotojų nuotaika turi įtaką akcijų kainoms, o formuluojama problema - kaip apskaičiuoti nuotaikos įtaką prekybai ir numatyti jos pokyčius [13].

Dažniausiai sprendimo priėmimui naudojamos techninė arba fundamentali analizė, kuriose nėra išgryninami investuotojų nuotaikos faktoriai. Investuotojų nuotaiką bandytą įvertinti sudarant paieškos užklausomis paremtą FEARS indeksą [13] arba sudarant naujus techninius indikatorius [10], tačiau pirmu atveju naudoti tik abstraktūs duomenys, antru – panaudota tik dalis prieinamos prekybos informacijos.

Techniniai indikatoriai yra apdorotų prekybos duomenų skaitinė išraiška. Pasitelkus techninius indikatorius, kurie atspindi investuotojų nuotaiką, galima sudaryti tik nuotaikomis pagrįstą investavimo strategiją. Iki šiol tik nuotaikų techniniais indikatoriais paremtos investavimo strategijos nebuvo nagrinėjamos. Tokio tipo strategija gali būti pagrindu priimant investavimo sprendimus arba kaip pagalbinė priemonė, patvirtinanti arba paneigianti jau priimtą sprendimą. Techniniais indikatoriais paremta strategija turės tikslią skaitinę išraišką, gali būti lengvai keičiama pridedant arba pašalinant indikatorius, taip išnaudojant reikiamą maksimalią prekybos duomenų apimtį.

Straipsnyje pateikti faktoriai, atspindintys investuotojų nuotaiką, išskirti techniniai indikatoriai, kurie taip pat atspindi investuotojų optimistinį arba pesimistinį nusiteikimą, bei atliktas indikatorių pelningumo tyrimas.

II. ANKSTESNI TYRIMAI

Pirmieji investuotojų nuotaikos ir prekybos strategijų pelningumo tyrimai buvo atliekami 1980-aisiais. Tyrimai buvo labai abstraktūs, nebuvo vykdomi kryptingai ir apėmė visą akcijų rinką, o ne konkrečias sritis. Autoriai ieškojo ryšių tarp akcijų grąžos ir dividendų išmokų, pelningumo tendencijų. Ritter ieškojo rinkos svyravimo dėsningumų ir suformulavo mintį, kad pardavimų apimtys ne visada atitinka teorinius, paremtus skaičiavimais, pardavimus [8]. Tai buvo pirmieji tyrimai, kuriuose prisiliesta prie investuotojų nuotaikos analizavimo, tačiau terminas „rinkų nuotaikos indikatoriai“ atsirado vėliau.

Naujesni tyrimai išryškino investuotojų nuotaikos įtaką akcijų kainoms. Baker ir Wurgler atliko tyrimą, kuriame naudojo žodžių junginį „nuotaikų finansai“ ir pasiūlė daug testų, kurių pagalba galima nustatyti, ar rinka yra veikiama jos dalyvių nuotaikos [1]. Remiantis DeLong, Shleifer, Summers ir Waldmann tyrimu, investuotojai yra išskiriami į dvi grupes: racionalūs investuotojai ir nuotaikoms pasiduodantys investuotojai, arba iracionalūs investuotojai. Racionalūs investuotojai rinkoje varžosi nustatinėdami „protingas“ kainas, patvirtintas skaičiavimais ir pagrįstas įmonių finansinėmis analizėmis. Tuo pat metu racionalūs investuotojai turi savo taisykles. Dažniausiai tai yra rizikos prisiėmimo taisyklės, dėl kurių racionalių investuotojų suskaičiuojama akcijos kaina skiriasi nuo akcijos rinkos kainos. Šis akcijų kainos skirtumas yra įrodymas, jog tuo metu rinkoje veikia investuotojų nuotaikos [2].

Jeigu atsiranda investuotojų, manančių, kad iš tam tikros įmonės akcijų galima pasipelnyti daugiau negu rodo įmonės finansų analizė, tuomet akcijos kaina pradeda nepagrįstai kilti. Kylant akcijos kainai, į prekybą įsitraukia vis daugiau optimistiškai nusiteikusių investuotojų. Dažniausiai tai būna iracionalūs, nuotaikos ir jausmų vedami, nepatyrę investuotojai. Tam tikru laiko momentu šie investuotojai netgi gali fiksuoti didelius pelnus, tačiau jų elgesys yra lyginamas su dalyvavimu loterijoje. Yra manoma, kad kuo vėliau investuotojas įžengia į kylančią poziciją, tuo mažiau jis yra patyręs ir tuo mažesnė tikimybė, kad jis „laimės“ šią loteriją.

Tačiau tokie nepagrįsti akcijų kainų kilimai ar kritimai yra mėgstami spekuliantų.

Baker ir Wurgler atsako į klausimą, kodėl vienos akcijos naudojamos spekuliavimu labiau nei kitos. Yra tikima, kad sunkiausia nustatant akcijų kainas yra objektyviai ir teisingai įvertinti įmonės finansinę būklę. Kaip pavyzdys pateikiama nesenai įkurta įmonė, šiuo metu nepelninga, bet galimai ypač pelninga ateityje. Pelno nebuvimas praeityje ir nepatikima įmonės ateitis sudaro sąlygas, kai įmonės akcijų kaina nustatoma, vadovaujantis investuotojų prielaidomis ir jausmais [1]. Palyginimui - ilgai ir stabiliai dirbanti įmonė, turinti ilgą pelningumo istoriją ir stabilius rodiklius, yra mažiau įtakojama investuotojų jausmų. Miller pastebi, kad netgi tą pačią informaciją turinčių investuotojų nuomonė gali labai skirtis ir šie skirtumai formuoja polinkį akcijų spekuliavimui [4].

G.D'Avolio teigia, kad investuoti į spekuliantų pamėgtas akcijas yra rizikinga. Tokių akcijų kainos kitimas gali būti chaotiškas, staigų kilimą gali keisti toks pat staigus kritimas. Tas pats taikoma ir akcijų prekybos apimtims, kai aktyvi prekyba akcijomis staiga aprimsta arba priešingai. Esminė mintis yra ta, kad kuo mažiau informacijos turima vertinant įmonę, tuo labiau jos akcijomis spekuliuojama [3].

III. RINKOS NUOTAIKŲ NUSTATYMO BŪDAI IR JŲ PALYGINIMAS

Investuotojų nuotaiką gali paveikti bet koks įvykis, kuris gali sukelti kitus įvykius, kurie įtakos investuotojo veiksmus. Kaip pavyzdį galima pateikti 1990 metai įvykusį interneto burbulo sprogimą [9]. Visus šiuos įvykius galima stebėti bei fiksuoti jų įtaką investuotojo veiksmams. Pirmiausia investuotojų nuotaikos pasikeitimus galima užfiksuoti atliekant apklausas. Šie nuotaikos pasikeitimai gali priversti investuotoją keisti savo prekybos strategiją. Investuotojų elgesyje įmanoma įžvelgti dėsningumą esant tam tikroms aplinkybėms, todėl prekybos strategijos keitimui galima pritaikyti tam tikrus šablonus ir formuotes. Tuomet prekyboje gali atsirasti skirtumai tarp racionalių investuotojų analizės ir spekuliantų siūlomos akcijų kainos. Šiuos skirtumus įmanoma pastebėti naudojant įvairius fundamentinius rodiklius. Žmonės, susiję su analizuojamomis įmonėmis, pavyzdžiui, įmonės vadovai, turėdami svarbią informaciją, gali išnaudoti susidariusius kainų skirtumus savo naudai. Kiekvienas šioje įvykių grandinėje esantis įvykis turi savo pamatuojamą išraišką, kurią vadiname rodikliu arba indikatoriumi.

Toliau aprašomi investuotojų nuotaiką atspindintys faktoriai.

A. Prekybos apimtis

Prekybos apimtis (angl. trading volume) yra investuotojų nuotaikos faktorius, kurį galima stebėti realiu laiku.

$$\text{Prekybos apimtis} = \text{pirkimo sandoriai} + \text{pardavimo sandoriai} \quad (1)$$

Kaip taisyklė yra žinoma, kad didelės apimties prekyba reiškia, jog rinkos investuotojų nuotaika dar išliks tokia, kokia yra. Tuomet, kai akcijos kaina kyla, tačiau mažėja prekybos apimtis, galima numatyti nuotaikos pasikeitimą ir kainos kilimo lėtėjimą arba kainos kritimą. Ta pati taisyklė taikoma ir akcijos kainai krentant.

Privalumai: informaciją apie prekybos apimtį galima gauti iš daugybės vietų, realiu laiku, internetu. Nereikia sudėtingo informacijos apdorojimo. Visos akcijų prekybos sistemos pateikia prekybos apimties informaciją.

Trūkumai: nėra labai patikimas investuotojų nuotaikos indikatorius, nes labiau padeda susidaryti bendrą vaizdą negu priimti tikslų sprendimą dėl pirkimo ar pardavimo.

B. Akcijos kainos kintamumas

Akcijos kainos kintamumas (angl. price volatility) - tai akcijos kainos pokyčio santykis. Akcijos kintamumas - kiek daug akcijos kaina kinta nuo laikotarpio vidutinės kainos reikšmės. Fundamentalioje akcijų analizėje akcijos kainos kintamumą žymi β koeficientas. Jeigu per pasirinkta laikotarpį akcijos kaina keičiasi nedaug ($0 < \beta < 1$), tai šios akcijos kintamumas yra mažas ir sakoma, kad akcija yra rami. Didelis akcijos kainos kintamumas ($\beta > 1$) yra naudingas investuotojams, perkantiems opcionus, nes tokiu atveju didesnė tikimybė, kad pasieks norimą kainą.

Privalumai: investavimas spekuliavimo tikslais trumpuoju laikotarpiu, padeda nustatyti ypač aktyvias akcijas, tinkamai įvertinti riziką investuojant ilguoju laikotarpiu.

C. Parduotų skolintų akcijų santykis

Skolintų akcijų pardavimas yra priešingas veiksmas akcijų pirkimui. Parduodamos asmeniui nepriklausančios ir pasiskolintos akcijos su tikslu uždirbti iš akcijos kainos kritimo. Po kurio laiko akcijos atperkamos ir grąžinamos savininkui. Visada galima žinoti, kiek rinkoje yra parduotų skolintų ir dar neatpirktų akcijų. Šis santykis skaičiuojamas, visą kompanijos skolintų parduotų ir neatpirktų akcijų kiekį padalinus iš akcijos prekybos apimties.

Santykiui padidėjus, galima teigti, kad atsirado daugiau žmonių, kurie tiki, kad ateityje akcijos kaina kris. Norint investuot į tokią akciją, reikėtų ypač tiksliai pasverti galimą riziką. Esant mažam santykiui, vyrauja optimistinės nuotaikos ir tikima, kad kaina kils.

D. Investuotojų apklausos

Klausiant investuotojų kaip optimistiškai ar pesimistiškai jie nusiteikę, galima išskirti vidutinį iracionalų investuotoją. Robert Shiller atlieka investuotojų apklausas nuo 1989 metų. „UBS/Gallup" apklausia atsitiktinai parinktus investuotojus, „Investors Intelligence" jau ilgą laiką daro specialistų, rašančių apie finansus, apklausas. Qiu ir Welch atkreipia dėmesį į tai, kad nors sudarant Mičigano universiteto vartotojų pasitikėjimo indeksą tiesioginių klausimų apie investavimą jame nėra, indekso pokyčiai koreliuoja su pokyčiais „UBS/Gallup" pateikiamuose apklausų rezultatuose. Taip pat autoriai pastebi, kad vartotojų pasitikėjimo indeksas stipriai koreliuoja su smulkių investuotojų portfeliuose laikomų akcijų grąža. Tai reiškia, kad smulkieji investuotojai yra labai įtakojami rinkos nuotaikų ir tuo pat metu patys gali sukelti nepagrįstas optimizmo ar pesimizmo bangas [5].

E. Investuotojų emocijos.

Buvo atliekami tyrimai, kurių pagalba buvo bandoma nustatyti ryšį tarp akcijų kainų ir žmogaus emocijų pokyčių. Kamstra, Kramer ir Levi (2003) teigia, kad rinkų grąža vidutiniškai yra žemesnė rudenį ir žiemą. Šis sezoniškas sumažėjimas grindžiamas depresinėmis investuotojų nuotaikomis dėl trumpėjančių saulėtų dienų. Buvo pastebėtas ir švenčių dėsningumas – prieš didžiausias metų šventes rinkose pastebimas nedidelis kilimas. Autoriai taip pat pastebėjo dėsningumą ir skirtingose laiko juostuose ar žemės pusrutuliuose. Kitas dėsningas emocijų pokytis buvo pastebėtas po tarptautinių futbolo rungtynių. Pralaimėjusios komandos šalies rinkoje pastebimas sumažėjęs aktyvumas, dažniausiai vienos dienos [6].

F. Smulkių investuotojų sandoriai

Yra manoma, kad smulkūs ir nepatyrę investuotojai prekiaujant yra labiau linkę į emocijas negu profesionalūs [1]. Interneto burbulo metu buvo pastebėta, kad jauni investuotojai daugiausia akcijų nupirko prieš pat burbulo „sprogimą“. Tai reiškia, kad esant optimistinėms nuotaikoms, netgi mažiausiai patyrę investuotojai pradeda aktyviai prekiauti akcijomis. Taip pat buvo pastebėta, kad smulkūs investuotojai sprendimus pirkti ar parduoti priima lyg susitarę ir veiktų kartu. Toks jų elgesys sutampa su visos rinkos nuotaikomis, todėl dažnai siūloma smulkių investuotojų sandorius išreikšti santykiu ir naudoti kaip efektyvų indikatorių.

G. Investicinių fondų kryptys

Informacija apie investicinių fondų prekybos srautus ir apimtis yra lengvai prieinama. Yra manoma, kad investuotojų nuotaiką gali įtakoti investicinių fondų sprendimai. Jeigu fondas investuoja į saugiomis laikomas valstybių obligacijas, investuotojui tai gali būti signalas susilaikyti nuo rizikingų investicijų ir pasirinkti konservatyvesnę ir tuo pačiu saugesnę strategiją. Kitu atveju, jeigu fondas investuoja į įmonių akcijas, vadinasi, fondas tiki optimistinėmis ateities perspektyvomis ir investuotojas gali prisiimti didesnę riziką taip pat. Frazzini ir Lamont stebėdami investicinių fondų veiklą ir įmonių akcijų apyvartas, pastebėjo, kad fondui investavus į konkrečios įmonės akcijas, jų tolimesnis pelningumas sumažėja [7].

H. Pirminis akcijų siūlymas

Kartais pirminio akcijų siūlymo metu per pirmą dieną įmonės surenka neįprastai dideles pajamas. Tokias aukštas pajamas galima pagrįsti tik dideliu investuotojų optimizmu. Netscape 1995 metais per pirmą dieną surinko 108 procentais daugiau negu buvo tikėtasi. Pirminio akcijų siūlymo metu vyksta dideli apyvartų svyravimai. Investuotojai šiuos svyravimus vadina galimybių langais, kurių metu parduodami didžiuliai akcijų kiekiai.

I. Susijusių asmenų prekyba

Įmonių vadovai ir su jais susiję asmenys turi daugiau informacijos apie įmonės būklę negu kiti investuotojai. Šie asmenys dažniausiai į prekybą įsitraukia rinkoje esant vienai iš dviejų situacijų. Įmonės akcijų kaina yra neįvertinta ir tikroji akcijų vertė yra didesnė, todėl laukdami kainų kilimo, vadovai perka akcijas. Kita situacija: kai rinkoje vyrauja ypač pesimistinės nuotaikos, vadovai perka įmonės akcijas norėdami parodyti investuotojams, kad įmonė veikia stabiliai.

IV. RINKŲ NUOTAIKŲ INDIKATORIŲ LYGINAMOJI ANALIZĖ

Rinkų nuotaikų indikatorius galima skirstyti pagal:

- Trukmę – dažniausiai sutinkami indikatoriai gali būti apskaičiuojami vienos dienos, savaitės ar mėnesio tikslumu. Retais atvejais rinkų nuotaikų indikatoriai skaičiuojami smulkesniu tikslumu (valandos ar minučių).
- Investuotojų patirtį – manoma, kad didžiausiomis apyvartomis rinkoje pasižymi labiau patyrę investuotojai ir investuotojai, turintys neviešos informacijos, ir jų veiksmai yra pagrįsti ir paremti žiniomis. Šių investuotojų prekyba vadinama „protingų pinigų prekyba“ (angl. smart money flow). Priešingai yra su nepatyrusiais investuotojais, kurių apyvartos yra ypatingai mažos. Pagal prieštaringą investavimo teoriją, tokių investuotojų visuma niekada nepriims teisingo sprendimo, todėl ir indikatoriai, kurie yra paremti neprofesionalių investuotojų prekyba, vadinami „neprotingų pinigų prekyba“ (angl. dumb money flow).

Indikatoriai, atspindintys investuotojų pesimistinį arba optimistinį nusiteikimą.

A. Program Trading Buy/Sell Spread

Šis indikatorius apima mažiausiai 15-os vertybinių popierių prekybos informaciją. Bendra skaičiuojamų vertybinių popierių sandorių vertė turi viršyti 1 milijoną dolerių. Tikima, kad tokiais kiekiais prekiauja investicinės kompanijos, kurių prekyba yra kompiuterizuota. Investicinės kompanijos laikomos prieštaringais investuotojais, t.y. taip pat stebinčiais rinkos nuotaikas.

Indikatorius parodo skirtumą tarp 15-os didelės vertės vertybinių popierių pasiūlos ir paklausos (2). Indikatorius priskiriamas „protingų“ pinigų indikatoriams.

Buy/Sell spread = pirkimo sandoriai / pardavimo sandoriai (2)

Indikatorius atnaujinamas kartą per savaitę. Indikatoriaus autoriai - kompanija „Wall street courier“ [10].

Signalai: visos reikšmės mažesnės nei -10 sudaro pirkimo signalą, reikšmės didesnės nei 35 – pirkimo skolon signalą.

B. Put/Call ratio

Indikatorius, parodantis santykį tarp pirkimo ir pardavimo opcionų (3).

Put/Call ratio = pirkimo opcionų sk. / pardavimo opcionų sk. (3)

Šis indikatorius parodo dabar vyraujančią investuotojų nuotaiką dėl ateities prekybos. Indikatorius yra priskiriamas prieštaringos nuomonės indikatoriams. Yra manoma, kad esant ypač aukštam optimizmui opcionų rinkoje, greitu metu keisis prekybos kryptis į pesimistišką ir atvirkščiai. Indikatorius yra sudaromas pagal CBOE (Čikagos opcionų biržos) opcionų prekybos duomenis. Gali būti keli indikatoriaus tipai – tik

akcijų, tik indeksų ir bendras. Indeksų opcionais prekiauja ne tik smulkieji investuotojai, bet ir didžiosios investicinės kompanijos, todėl akcijų opcionų santykis atspindi daugumos investuotojų nuomonę labiau negu indeksų opcionų santykis.

C. Smart Money Flow

Indikatoriaus pavadinimas reškia „protingų“ pinigų judėjimą. Šis indikatorius sukurtas 1997 metais kompanijos „WallStreetCourier“ [10]. Indikatorius yra apskaičiuojamas, remiantis dviem laikotarpiais. Nagrinėjama pasirinkto indekso dienos prekybos pradžia – pirmos 30 minučių ir prekybos dienos paskutinė valanda. Per pirmas 30 minučių dominuoja baimės ir gobšumo vedami investuotojai, kurie bando fiksuoti savo pelnus ar nuostolius dėl per naktį atsiradusios informacijos, pvz., kotiruojamos kompanijos ataskaitos pateikimo. Taip pat rytinės prekybos metu uždaromos skolon pirktų akcijų pozicijos.

Paskutinė prekybos valanda yra įdomiausia, nes tuo metu daugiausiai „manevrų“ atlieka profesionalūs, daug žinių ir patirties turintys investuotojai, kurie disponuoja didelėmis pinigų sumomis. Paskutinės valandos pradžioje jie tikrina rinkos reakciją atlikdami mažesnius pirkimus ar pardavimus, o valandos pabaigoje vykdo didžiausius sandorius. Šie didieji investuotojai gali turėti daugiau informacijos negu kiti rinkos dalyviai, pvz., organizacijos vidinę neviešą informaciją.

Jeigu analizuojamas indeksas artėja prie didžiausios ar mažiausios reikšmės, o smart money flow indikatorius to neparodo, tai reiškia, kad indekso reikšmės nepagrįstos ir didieji investuotojai netiki tolimesniu kainos judėjimu ta pačia linkme.

Indikatorius gali būti skaičiuojamas kiekvieną dieną arba kartą per savaitę.

Signalai: jeigu indeksas artėja prie didžiausios ar mažiausios reikšmės, o indikatorius nekoreliuoja, tuomet tikimasi kainos judėjimo krypties pokyčio.

D. Globalus ateities sandorių dugno indikatorius

Šis indikatorius sudarytas iš kassavaitinių ateities sandorių pirkimo ir pardavimo apimties santykių. Indikatorius nepasižymi dažnu signalo suformavimu, tačiau tokiam atsiradus, ypač tikėtinas krypties pasikeitimas. Indikatorius vertinamas ilgo laikotarpio investuotojų, nes puikiai apsaugo nuo pirkimų netinkamu laiku. Ypač geri atsiliepimai investuotojų į ateities sandorius.

Deja, šis indikatorius neparodo pardavimo ar atpirkimo signalo. Indikatorius atnaujinamas kas savaitę ir yra priskiriamas „protingų“ pinigų indikatorių grupei. Nulinės ir visos reikšmės žemiau nulio parodo galimą dugną. Nuo indikatoriaus sukūrimo, lyginant su DOW JONES indeksu, dugnas neteisingai buvo nurodytas tik du kartus 1990 metų rugpjūčio mėnesį dėl karinių veiksmų Artimuosiuose rytuose.

Signalai: reikšmės tarp 1 ir 5 patikimai nurodo rinkos dugną ir signalizuoja būsimą kainos augimą.

E. Kassavaitinė investuotojų apklausa

Amerikos individualių investuotojų asociacijos rengiama kassavaitinė investuotojų apklausa, kurioje nurodoma investuotojų nuotaika ateinantiems 6 mėnesiams [11]. Investuotojai gali atsakyti nurodydami pesimistišką, neutralią arba optimistišką savo nusiteikimą. Investuotojas gali pasirinkti tik vieną variantą.

Ši apklausa laikoma prieštaringos nuomonės indikatoriumi, nes yra tikima, kad investuotojas, kuris investavęs didžiąją daugumą savo lėšų, nebeturi perkamos galios ir tikisi augimo, todėl rinka yra pasiekusi savo aukštumą. Kita vertus, kai investuotojai pranašauja rinkos kritimą, jų portfeliai jau būna tušti, todėl tikima, kad rinka tik augs.

Kiekvieną savaitę yra apklausiami 170 tūkstančių asociacijos narių. Optimistiškai ir pesimistiškai nusiteikusių investuotojų santykis ir yra naudojamas kaip indikatorius. Didelės indikatoriaus reikšmės nurodo rinkos pakilimą, mažos – apačią.

V. TIRIAMŲ INDEKSŲ PALYGINIMAS

A. S&P 500 JAV įmonių akcijų kainų indeksas

Indeksas S&P 500 yra laikomas vienu geriausių JAV akcijų biržos rodiklių. Jį apskaičiuojant yra įtrauktos 500 pirmaujančių įmonių svarbiausiuose JAV sektoriuose, suskirstytuose pagal GICS[1] klasifikaciją, akcijų kaina. S&P 500 rodo apie 75%. Jungtinių Amerikos valstijų akcijų biržos kapitalizacijos. Pagrindiniu kriterijumi atrenkant įmones yra laikoma rinkos kapitalizacija, tačiau atsižvelgiama ir į akcijos likvidumą, augimo tempus, įmonės veiklos rūšį. Beveik pusę indekso reikšmės sudaro finansų, informacinių technologijų ir energetikos sektorių akcijos. Šis indeksas buvo pradėtas apskaičiuoti 1923 metais. Tuo metu apskaičiuojant indekso reikšmę atsižvelgta tik į 233 įmonių akcijų kainas, o 1957 metais jis buvo pradėtas apskaičiuoti atsižvelgiant į 500 įmonių akcijų kainas. Indekso sudėtis yra peržiūrima tada, kai reikia, t. y. nėra nustatyta konkrečių datų, kada turi susirinkti komitetas.

B. NASDAQ COMPOSITE indeksas

Indeksas vienija daugiau nei 3000 akcijų, kuriomis prekiaujama per NASDAQ vertybinių popierių biržą. Indeksas daugiausia sudarytas iš Jungtinių Amerikos Valstijų kompanijų, bet įtrauktos ir kitų šalių kompanijos. Panašiai kaip ir S&P 500 indeksas, kuris atspindi bendrą rinkos būseną ir nusiteikimą, NASDAQ COMPOSITE indeksas atspindi siauresnės šakos būklę. Į indeksą daugiausia įtrauktos technologijų kompanijų akcijos, todėl indeksas geriausiai atspindi vyraujančią būseną technologijų srityje

C. DOW JONES INDUSTRIAL indeksas

Vertybinių popierių rinkos indeksas, dar kitaip vadinamas “Industrial average”, “Dow Jones” arba “Dow 30”. Šis indeksas, sukurtas XIX – ame amžiuje populiaraus leidinio

[1] GICS (angl. Global Industry Classification Standard) pasaulio pramonės klasifikavimo standartas, kurio struktūrą sudaro 10 sektorių, 24 ekonominės veiklos grupės, 68 veiklos šakos ir 154 veiklos porūšiai.

„The Wall Street Journal“ redaktoriaus Charles Dow, atspindi Jungtinių Amerikos Valstijų industrinio sektoriaus akcijų kursus. Tai yra vienas seniausių indeksų Jungtinėse Valstijose, einantis po to paties verslininko sukurto „Dow Jones Transportation Average“ indekso. Indeksas apskaičiuojamas naudojant 30-ties didžiausių ir paklausiausių JAV kompanijų. Žodis industrinis yra istorinis reliktas, nes šiuo metu daug didžiausių firmų neturi nieko bendro su senąja sunkiąja industrija

VI. Siūlomo sprendimo metodika

Siūloma išskirti naudojamus techninius rinkų nuotaikų indikatorius ir ištirti jų efektyvumą naudojant istorinę informaciją. Tyrimą pateikti įrankio pavidalu, pasiekiamu per interneto naršyklę. Sudaryti galimybę atlikti išskirtų indikatorių efektyvumo tyrimą naudojant pasirinktus indeksus.

Tyrimo metu bus atliekamas rinkų nuotaikų indikatorių efektyvumo tyrimas. Šiai užduočiai atlikti bus naudojami istoriniai duomenys. Istorinių duomenų apdorojimas, indikatorių skaičiavimas, prekybos strategijų pelningumo skaičiavimai, testavimo rezultatų apdorojimas ir kitos loginės užduotys bus atliekamos naudojant Microsoft Excel.

Tiriant indikatorių efektyvumą yra svarbu atlikti testavimus, pasitelkiant skirtingų laikotarpių praeities duomenis. Istoriniai akcijų kainų bei apyvartos kitimo duomenys bus parsiunčiami iš http://finance.yahoo.com ir http://www.nasdaq.com.

Tyrime naudojamų rinkos nuotaikų indikatorių lyginamosios reikšmės ir indeksų reikšmės gaunamos iš http://www.wallstreetcourier.com.

VII. Eksperimentinis tyrimas

Eksperimentinio tyrimo metu siekiama nustatyti efektyviausius investuotojų nuotaikos indikatorius. Kiekvieno indikatoriaus efektyvumas vertinamas nepriklausomai nuo kitų. Tyrimo metu bus keičiami indikatorių parametrai ir fiksuojami pirkimo ir pardavimo ar pirkimo skolon ir atpirkimo signalai. Indikatoriai bus lyginami su S&P 500 indekso reikšmėmis.

Kadangi indikatoriai lyginami su S&P 500 indeksu, o ne su konkrečios akcijos prekybos duomenimis, tai indekso reikšmė yra ir lyginamoji pelningumo reikšmė.

Rinkos nuotaikos pokyčius tiriami indikatoriai prognozuoja trumpuoju laikotarpiu, todėl pelningumas yra fiksuojamas 2, 4 ir 8 savaitės po indikatoriaus signalo susiformavimo

A. Indikatoriaus Program Trading Buy/Sell Spread tyrimas

Indikatorius lyginamas su S&P 500 indeksu.

Tiriamas laikotarpis nuo 2007 metų sausio 01 dienos iki 2013 metų sausio 01 dienos. Indikatorius apskaičiuojamas vienos savaitės duomenimis.

Indikatoriaus signalai: indikatorius turi tiksliai apibrėžtus pirkimo ir pirkimo skolon signalus. Indikatoriaus reikšmės svyruoja apie nulį ir tiriamu laikotarpiu nebuvo nutolę toliau negu -30,75 ir 71,975. Aiškus pirkimo signalas atsiranda ties indikatoriaus reikšme -10, pirkimo skolon – ties 35.

Tiriamu laikotarpiu indikatoriaus pagalba buvo nustatyti 7 pirkimo ir 3 pirkimo skolon signalai. Susidarius aiškiam pirkimo signalui, fiksuojama S&P 500 indekso reikšmė signalo susidarymo metu ir po nustatyto laiko. Antros indekso reikšmės fiksavimas yra kintanti indikatoriaus tyrimo reikšmė.

Program trading indikatoriaus pelningumo palyginimas	***Pardavimo arba atpirkimo reikšmės fiksavimas po:***		
	2 savaičių	***4 savaičių***	***8 savaičių***
Pirkimo signalų pelningumas	-56,91 indekso punktų	-80,78 indekso punktų	263,5 indekso punktų
Pirkimo skolon signalų pelningumas	-31,18 indekso punktų	-10.07 indekso punktų	489,69 indekso punktų
VISO	-88,09 indekso punktų	-90.85 indekso punktų	753.19 indekso punktų

Nors buvo gauti 7 pirkimo signalai, o pirkimo skolon tik 3, galima teigti, kad stipresni yra pirkimo skolon signalai. Taip pat tyrimo metu nustatyta, kad indikatorius veiksmingesnis ilgesniu laikotarpiu, indekso reikšmių fiksavimas po 8 savaičių nuo signalo suformavimo „uždirbo“ netgi 753 indekso punktų lyginamąjį pelną.

B. Indikatoriaus Put / Call Ratio tyrimas

Indikatorius lyginamas su S&P 500 indeksu.

Tiriamas laikotarpis nuo 2007 metų sausio 01 dienos iki 2013 metų sausio 01 dienos. Indikatorius apskaičiuojamas vienos savaitės duomenimis.

Indikatoriaus signalai: kadangi indikatoriaus reikšmė yra santykis, indikatoriaus reikšmė negali būti mažesnė už nulį ir tik retais atvejais viršija vienetą. Pirkimo signalas atsiranda ties indikatoriaus reikšme 0,7, o pirkimo skolon – ties 0,55.

Tiriamu laikotarpiu indikatoriaus pagalba buvo nustatyti 24 pirkimo ir 19 pirkimo skolon signalai. Susidarius aiškiam pirkimo signalui, fiksuojama S&P 500 indekso reikšmė signalo susidarymo metu ir po nustatyto laiko. Antros indekso reikšmės fiksavimas yra kintanti indikatoriaus tyrimo reikšmė.

Put / Call Ratio indikatoriaus pelningumo palyginimas	***Pardavimo arba atpirkimo reikšmės fiksavimas po:***		
	2 savaičių	***4 savaičių***	***8 savaičių***
Pirkimo signalų pelningumas	-35,97 indekso punktų	102,96 indekso punktų	16,71 indekso punktų
Pirkimo skolon signalų pelningumas	-29,55 indekso punktų	78,49 indekso punktų	27,24 indekso punktų
VISO	-65,52 indekso punktų	181,45 indekso punktų	43,95 indekso punktų

Tiriamu laikotarpiu indikatorius suformavo žymiai daugiau signalų negu Program Trading Buy/Sell Spread indikatorius. Tyrimo metu nustatyta, kad indikatorius veiksmingesnis vidutiniu 4 savaičių laikotarpiu, indekso reikšmių fiksavimas po 2 ir 8 savaičių nuo signalo suformavimo „uždirbo“ mažiau lyginamojo pelno.

C. *Indikatoriaus Smart Money Flow tyrimas*

Indikatorius lyginamas su S&P 500 indeksu.

Tiriamas laikotarpis nuo 2007 metų sausio 01 dienos iki 2013 metų sausio 01 dienos. Indikatorius apskaičiuojamas vienos savaitės duomenimis.

Indikatoriaus signalai: indikatorius parodo dalinę prekybos informaciją, todėl dažniausiai koreliuoja su lyginamo indekso reikšme. Tuo atveju, kai indikatoriaus ir indekso reikšmės prasilenkia daugiau nei 10%, tikimasi krypties pasikeitimo.

Tiriamu laikotarpiu indikatoriaus pagalba buvo nustatyti 2 pirkimo ir 1 pirkimo skolon signalai. Susidarius pirkimo signalui, fiksuojama S&P 500 indekso reikšmė signalo susidarymo metu ir po nustatyto laiko. Antros indekso reikšmės fiksavimas yra kintanti indikatoriaus tyrimo reikšmė.

Smart Money Flow indikatoriaus pelningumo palyginimas	***Pardavimo arba atpirkimo reikšmės fiksavimas po:***		
	2 savaičių	***4 savaičių***	***8 savaičių***
Pirkimo signalų pelningumas	7,10 indekso punktų	19,37 indekso punktų	33,63 indekso punktų
Pirkimo skolon signalų pelningumas	-14,75 indekso punktų	24,58 indekso punktų	27,35 indekso punktų
VISO	-7,65 indekso punktų	43,95 indekso punktų	60,98 indekso punktų

Tiriamu laikotarpiu indikatorius iš viso suformavo tik 3 signalus. Labiausiai pelningas 8 savaičių laikotarpis. Galima teigti, kad indikatoriaus signalai yra aiškūs.

D. *Globalaus ateities sandorių dugno indikatoriaus tyrimas*

Indikatorius lyginamas su S&P 500 indeksu.

Tiriamas laikotarpis nuo 2007 metų sausio 01 dienos iki 2013 metų sausio 01 dienos. Indikatorius apskaičiuojamas vienos savaitės duomenimis.

Indikatoriaus signalai: indikatorius parodo tik pirkimo signalus. Pirkimo signalas suformuojamas, kai indikatoriaus reikšmė yra lygi arba mažesnė už 5.

Tiriamu laikotarpiu indikatoriaus pagalba buvo nustatyti 2 pirkimo signalai. Susidarius pirkimo signalui, fiksuojama S&P 500 indekso reikšmė signalo susidarymo metu ir po nustatyto laiko. Antros indekso reikšmės fiksavimas yra kintanti indikatoriaus tyrimo reikšmė.

Globalaus ateities sandorių dugno indikatoriaus pelningumo palyginimas	***Pardavimo arba atpirkimo reikšmės fiksavimas po:***		
	2 savaičių	***4 savaičių***	***8 savaičių***
Pirkimo signalų pelningumas	-10,97 indekso punktų	-5,63 indekso punktų	13,69 indekso punktų

Tiriamu laikotarpiu indikatorius iš viso suformavo tik 2 pirkimo signalus. Labiausiai pelningas 8 savaičių laikotarpis. Indikatorius tiksliai nurodė indekso mažiausią reikšmę, po kurios preasidėjo indekso augimas.

VIII. IŠVADOS

Atlikus rinkų nuotaikų indikatorių lyginamąją analizę, buvo išrinkti indikatoriai, kurie bus naudojami tyrime ir gali būti lyginami su S&P500 indeksu.

Eksperimentinio tyrimo metu atliktas 4 indikatorių pelningumo tyrimas fiksuojant tris laikotarpius – po 2, 4 ir 8 savaičių. 2 savaičių laikotarpiu visi indikatoriai parodė neigiamą pelningumą. Labiausiai pelninga investavimo strategija gali būti sudaryta iš 4 ir 8 savaičių trukmės investavimo laikotarpių.

Efektyviausias ilgojo laikotarpio indikatorius – Program Trading indikatorius.

Indikatoriai Smart Money Flow ir Put / Call ratio kartu naudojant gali sudaryti pelningą vidutinės trukmės strategiją.

Globalaus ateities sandorių dugno indikatorius gali būti naudojamas bet kurioje strategijoje kaip indikatorius, patvirtinantis kitų indikatorių suformuotus signalus.

IX. LITERATŪRA

[1] Malcolm Baker, Jeffrey Wurgler (2006), Investor Sentiment and the Cross-Section of Stock Returns, Journal of Finance, Vol. 61 Iss. 4 pp. 67–80.

[2] DeLong, J. Bradford, Andrei Shleifer, Lawrence H. Summers, Robert J. Waldmann. (1990). Noise Trader Risk in Financial Markets. Journal of Political Economy, Vol. 98 Iss. 4: pp. 23–38.

[3] D'Avolio, Gene. 2002. "The Market for Borrowing Stock." Journal of Financial Economics, 66(2–3): 271–306.

[4] Miller, Edward M. 1977. "Risk, Uncertainty, and Divergence of Opinion." Journal of Finance, 32(4): 1151–68.

[5] Qiu, Lily X., and Ivo Welch. (2006). "Investor Sentiment Measures." [žiūrėta 2011-12-01] Prieiga per internetą http://ssrn.com/abstract589641.

[6] Kamstra, Mark J., Lisa A. Kramer, and Maurice D. Levi. 2003. "Winter Blues: A SAD Stock Market Cycle." American Economic Review, 93(1): 1257–63.

[7] Frazzini, Andrea, and Owen Lamont. 2006. "Dumb Money: Mutual Fund Flows and the Cross-section of Stock Returns." National Bureau of Economic Research Working Paper 11526.

[8] Ritter, Jay (1984). The ‚Hot Issue' Market of 1980. The Journal of Business, Vol. 57 Iss. 2 pp. 35–40.

[9] Malcolm Baker, Jeffrey Wurgler (2007), Investor Sentiment in the Stock Market, Journal of Economic Perspectives, Vol. 21 Iss. 2 pp. 129–151.

[10] Wall Street Courier Indicator Dashboard. Prieiga per internetą http://www.wallstreetcourier.com/spro/sp-indicator-dashboard.htm

[11] The American Association of Individual Investors sentment survey. Prieiga per internetą http://www.aaii.com/sentimentsurvey

[12] Zhi Da, Joseph Engelberg and Pengjie Gao (2009) The Sum of All FEARS: Investor Sentiment and Asset Prices. Available at SSRN: http://ssrn.com/abstract=1509162

INFORMACINĖS TECHNOLOGIJOS ● IVUS 2013 ● ISSN 2029-249X ● eISSN 2029-4824

Formalizuotas judriųjų metodikų aprašymas

Aurelija Orlova
Programų sistemų katedra,
Matematikos ir informatikos fakultetas
Vilniaus universitetas
Didlaukio g. 47, Vilnius, Lietuva
El. paštas: aurelija.orlova@mif.stud.vu.lt

Santrauka. **Šiuo metu pradėta plačiai taikyti judriąsias metodikas, kadangi priešingai nei taikant tradicines programų sistemų kūrimo metodikas, yra sudaroma galimybė vykdyti projektą su nuolat besikeičiančiais reikalavimais [1, 2]. Įmonės, siekdamos sėkmingai įgyvendinti programų kūrimo projektus – pateikti rezultatus laiku, neviršijant biudžeto, ir kad klientai būtų patenkinti jų kokybe – taip pat renkasi programų kūrimo procesų vertinimo modelius, tokius kaip ISO/IEC 15504 ir CMMI, pagal kuriuos gali įvertinti savo programų kūrimo procesą ir jį gerinti. Skirtingos metodikos gali užtikrinti skirtingą proceso brandą, o bet kuris proceso keitimas yra sudėtingas ir brangus, todėl išankstinis metodikos įvertinimas prieš įdiegimą leistų įmonei sutaupyti daug laiko ir lėšų [3]. Tačiau kiekviena judrioji metodika, tokia kaip Scrum, DSDM Atern, XP, Crystal ir kt., yra aprašyta savaip, todėl sunku jas lyginti tarpusavyje ir įvertinti, kokį gebėjimą/brandą pagal programų kūrimo procesų vertinimo modelius gali užtikrinti pasirinkta judrioji metodika. Problemai spręsti, straipsnyje yra pateikiama tyrimo metu sukurta judriųjų metodikų struktūrizavimo metodika, pagal kurią yra sudaromi vienodai formalizuoti judriųjų metodikų aprašymai. Turint tokius aprašymus, palengvėja skirtingų judriųjų metodikų tarpusavio lyginimas, sudaroma galimybė sukurti vertinimo metodiką, pagal kurį nors programų kūrimo procesų vertinimo modelį, tinkančią visoms judriosioms metodikoms. Remiantis sukurta metodika buvo sudarytas formalizuotas DSDM Atern aprašymas, naudojant EPF įrankį.**

Reikšminiai žodžiai: judriosios metodikos, judriųjų metodikų ontologija, judriųjų metodikų struktūrizavimas, SPEM, EPF, DSDM Atern.

I. ĮVADAS

Judriosios (angl. agile) programų sistemų kūrimo metodikos, atsiradusios 90-ųjų metų pabaigoje, greitai išpopuliarėjo. Po 2001 metų, kai buvo išleistas judriųjų metodikų manifestas (angl. Agile manifesto) [1], apibrėžęs pagrindinius šių metodikų principus, pradėta plačiai jas taikyti [2]. Priešingai nei taikant tradicines programų sistemų kūrimo metodikas, judriųjų metodikų manifesto apibrėžti principai skelbia, kad sistema turi būti kuriama taip, kad galėtų prisitaikyti prie nuolat besikeičiančių reikalavimų ir atitiktų kiekvieną verslo poreikį.

Įmonėms taip pat tikslinga žinoti, kokį programų kūrimo proceso brandą/gebėjimą pagal programų kūrimo procesų vertinimo modelius, tokius kaip CMMI ir ISO/IEC 15504, gali pasiekti, įsidiegus konkrečią judriąją metodiką. Kuo proceso branda/gebėjimas didesnis, tuo įmonės projektuose mažiau aptinkama defektų, mažėja darbo sąnaudos, tiksliau prognozuojami projekto terminai ir biudžetas [3]. Metodikų įvertinimas taip pat gali padėti tobulinti metodiką, siekiant pasiekti aukštesnį gebėjimo/brandos lygį.

Šiuo metu egzistuoja daug skirtingų judriųjų metodikų, tokių kaip Scrum, Dynamic Systems Development Method (DSDM) Atern, Crystal Clear, Extreme Programming (XP) ir kt. [1]. Deja, kiekviena metodika aprašyta skirtingai, todėl sunku jas lyginti tarpusavyje ir įvertinti, kokį gebėjimą/brandą pagal programų kūrimo procesų vertinimo modelius gali užtikrinti pasirinkta judrioji metodika. Be to, skirtingose metodikose tas pats terminas gali turėti skirtingą prasmę arba gali būti naudojami skirtingi terminai, aprašantys tą pačią semantinę koncepciją. Pavyzdžiui, Agile UP metodo sąvoka disciplina (angl. discipline) atitinka XP metodo veiklos (angl. activity) sąvoką [4].

Norint palengvinti skirtingų judriųjų metodikų panašumų ir skirtumų radimą bei vertinimą pagal programų kūrimo procesų vertinimo modelius, judriosios metodikos turėtų būti formalizuotai aprašytos.

Vis labiau pradėta metodikų aprašymui naudoti SPEM (Software & Systems Process Engineering Metamodel) [5], kuris leidžia struktūrizuoti metodikas pakankamai, kad būtų galima lengviau su jomis susipažinti. Šiuo metu SPEM yra aprašytos Scrum, Agile Business Rules Development, OpenUp, EPF Practices ir XP metodikos, tačiau kiekviena jų aprašyta skirtingai [6].

Norint vienodai aprašyti judriąsias metodikas, turi būti remiamasi vieninga judriųjų metodikų ontologija, pagal kurią būtų galima susieti skirtingų judriųjų metodikų terminiją. Apibrėžus tokią ontologiją, būtų aiškus ir vienareikšmiškas aprašymų atspindėjimas skirtingose judriosiose metodikose pagal apibrėžtus terminų ryšius.

Šio darbo tikslas – pateikti sukurtą formalizuoto judriųjų metodikų aprašymo metodiką, pagal kurią aprašius judriąsias metodikas, palengvėtų skirtingų judriųjų metodikų tarpusavio lyginimas ir vertinimas pagal programų kūrimo procesų vertinimo modelius.

Tyrimo metu buvo papildyta ir patikslinta judriųjų metodikų ontologija, sukurta formalizuoto judriųjų metodikų

aprašymo metodika bei pagal ją sudarytas formalizuotas DSDM Atern aprašymas.

II. SUSIJĘ DARBAI

Atlikus literatūros analizę, rasta darbų, kuriuose pateikiami konkrečių judriųjų metodikų vertinimai pagal programų kūrimo proceso modelius [7, 8]. Dažniausiai sutinkami yra vertinimai pagal CMMI. Vienas iš tokių darbų yra N. Potter ir M. Sakry 2009 metais publikuotas darbas [7], kuriame yra pateikiama kaip Scrum metodo naudojimas gali įgyvendinti 2 ir 3 brandos lygio CMMI procesų sričių specifines praktikas. Tuo tarpu 2011 metais SPICE konferencijoje buvo pristatytas tyrimas [8], kuris parodo, kaip Scrum metodas padeda įgyvendinti ISO/IEC 12207:2008 procesų bazines praktikas. Deja, šiuose darbuose nėra pateikiama metodika kaip atliekami vertinimai, kad būtų galima pritaikyti kitų metodikų vertinimui.

Norint vertinti judriąsias metodikas reiktų jas formalizuotai ir vienodai aprašyti. Tai įgyvendinti padėtų judriųjų metodikų ontologija. Yra darbų, pateikiančių programų kūrimo procesų ontologijas [3, 9] , tačiau mažai darbų, kalbančių apie vieningą judriųjų metodikų ontologiją [4].

D. Parsons straipsnyje [4] yra pateikiama judriųjų metodikų ontologija, sukurta išnagrinėjus septynių skirtingų judriųjų metodikų terminus, radus sinonimus bei jiems parinkus vieną terminą. Deja, [4] straipsnyje nėra paaiškinama, kaip buvo identifikuoti sinonimai.

I lentelėje išdėstomi D. Parsons siūlomi ontologijos terminai ir judriųjų metodikų sinonimai. Jie identifikuoti, remiantis straipsnyje naudojamais Agile Microsoft Solutions Framework, Agile UP, Crystal Clear, DSDM, eXtreme Programming (XP), Feature Driven Development, Scrum judriųjų metodikų terminais ir jų pavyzdžiais.

I LENTELĖ. SĄRYŠIAI TARP D. PARSONS SUKURTOS ONTOLOGIJOS BEI JUDRIŲJŲ METODIKŲ TERMINŲ (REMIANTIS[4])

D. Parsons sukurtos ontologijos terminai	Judriųjų metodikų terminai
Darbo metodas (angl. technique)	Darbo metodas (angl. technique)
Fazė (angl. phase)	Fazė (angl. phase)
Principas (angl. principle)	Principas (angl. principle), ypatybė (angl. property), vertybė (angl. value)
Veikla (angl. activity)	Veikla (angl. activity), disciplina (angl. discipline)
Praktika (angl. practice)	Praktika (angl. practice), mąstysena (angl. mindset), filosofija (angl. philosophy), strategija (angl. strategy)

Dėl nežinomų priežasčių D. Parsons pateikta ontologija neįtraukia darbo produkto ir vaidmens terminų bei šių terminų sąryšių su kitais ontologijos terminais.

S. Jeners, H. Lichter, C. G. Rosenkranz [9] straipsnyje pateikiamas metamodelis, leidžiantis efektyviai įsisavinti ir įvertinti skirtingus procesų modelius. Proceso modeliais įvardijami programų kūrimo proceso modeliai (pvz., CMMI, ISO/IEC 15504) bei kiti modeliai/standartai (pvz., ITIL, COBIT). Tačiau straipsnyje nėra kalbama apie jokias metodikas, taip pat ir apie judriąsias.

Metodikoms aprašyti gali būti naudojami skirtingi aprašymo būdai. Vienas iš galimų būdų yra metodikas aprašyti SPEM [5]. SPEM yra naudojamas aprašyti programinės įrangos ir sistemos kūrimo procesus ir jų komponentus. Šiuo metu plačiai pradėta naudoti SPEM judriąsias metodikas, tokias kaip Scrum, Agile Business Rules Development, OpenUp, EPF Practices ir XP, aprašymams [6, 10].

Kita alternatyva – RDF (Resource Description Framework) [11]. RDF naudojama aprašyti interneto išteklių informaciją, pasitelkiant įvairius sintaksės formatus. Šis resursų aprašymo karkasas taip pat gali būti paimtas metodikoms aprašyti, pavyzdžiui, [12] straipsnyje pateikiama XP metodikos ontologija, atvaizduota panaudojus RDF.

Skirtingai negu RDF, SPEM įgalina pavaizduoti sudėtingus programinės įrangos ir sistemos kūrimo procesų veiksmus ir sąryšius diagramomis, todėl sumodeliuoti procesai yra lengviau suprantami ir išmokstami, gali būti vertinami ir analizuojami.

III. PAPILDYTA IR PATIKSLINTA JUDRIŲJŲ METODIKŲ ONTOLOGIJA

Judriųjų metodikų ontologija ir jos terminų sąryšiai, apibrėžti D. Parsons [4] straipsnyje, tyrimo metu buvo atnaujinti pagal naują DSDM metodikos versiją – DSDM Atern [13]. Tikslinant ir papildant ontologiją, nuspręsta remtis S. Jeners, H. Lichter, C. G. Rosenkranz straipsniu [9] ir SPEM specifikacija [5] dėl aiškiai apibrėžtų ir aprašytų terminų tarpusavio sąryšių. SPEM specifikacija taip pat turėjo įtakos dėl to, kad nuspręsta taikyti šį metamodelį metodikoms aprašyti. Toks sprendimas buvo priimtas dėl aiškaus grafinio atvaizdavimo, nemokamų įrankių, tokių kaip EPF (Eclipse Process Framework), bei jau egzistuojančių kelių judriųjų metodikų aprašymų (nors ir ne vienodų). Be to, SPEM yra charakterizuojamas kaip procesų metamodelis, kai tuo tarpu RDF – kalba, naudojama aprašyti interneto išteklių informaciją.

II lentelėje pavaizduoti sąryšiai tarp ontologijos bei judriųjų metodikų terminų (pakeitimai užrašyti pasviruoju tekstu). Pagrindinis sąryšių šaltinis yra [4] straipsnyje pateikti sąryšiai, kiti ontologijos terminai – vaidmuo, darbo produktas – buvo pridėti, remiantis S. Jeners, H. Lichter, C. G. Rosenkranz [9] straipsniu, SPEM specifikacija [5] ir papildomai išsinagrinėjus [4] straipsnyje nagrinėjamų septynių judriųjų metodikų terminus, radus sinonimus bei jiems parinkus vieną terminą.

Skirtingai nuo [4] straipsnio, pakeistas ontologijos veiklos (angl. activity) terminas į disciplinos (angl. discipline) terminą tam, kad vėliau nebūtų klaidinama su SPEM proceso terminu – veikla (angl. activity). SPEM [5] aprašo tiek disciplinos, tiek veiklos sąvokas. Abi jos yra skirtingos. Disciplina yra darbo skirstymo kategorija, o veikla yra pats darbas, kuris gali būti skaidomas į smulkesnius darbus. Disciplina siejama su „problemine sritimi“, tokia kaip reikalavimai, testavimas, projektavimas. SPEM įvardija, kad užduočių skirstymas į disciplinas yra ne kas kita kaip projekto supratimas iš tradicinio „krioklio“ (angl. waterfall) modelio perspektyvos. Kaip tik toks variantas yra nagrinėjamų judriųjų metodikų veiklos ir disciplinos atveju.

Remiantis nauja DSDM Atern versija [13], disciplinos terminas susietas su dar vienu sinonimu – sritis.

Reikia atkreipti dėmesį, kad kai kuriose judriosiose metodikose darbo metodai taip pat vadinami praktikomis. Pavyzdžiui, MoSCoW prioritizavimas DSDM Atern vadove

[13] vadinamas tiek darbo metodu, tiek praktika. Tokiais atvejais metodikos terminas *praktika* siejamas su ontologijos darbo metodo terminu.

Į judriųjų metodikų ontologiją buvo įtraukti dar trys terminai – *veikla*, *užduotis* ir *žingsnis*, kur *veikla* yra išskaidoma į smulkesnes *veiklas* ir/ar *užduotis*, o *užduotis*, priklausomai nuo užduoties aprašymo detalumo, gali būti skaidoma į *žingsnius*. Šiuo atveju, *veikla* atitinka SPEM [5] veiklos terminą – darbas, suskaidytas į smulkesnius darbus. II lentelėje *veiklos*, *užduoties* ir *žingsnio* terminai nėra įtraukti, kadangi judriosiose metodikose šie terminai ne visada yra išimtinai išskirti, o jei ir išskirti, tai nėra aiškiai parodomi jų tarpusavio sąryšiai. Pavyzdžiui, Agile Unified Process (AUP) metodikos fazių aprašymuose yra išskiriamos veiklos su trumpais jų aprašymais, tačiau nepateikiami veiklų sąryšiai su vaidmenų užduotimis. Tokiu atveju, reikia gerai išmanyti metodiką, norint susieti veiklas su vaidmenų užduotimis.

II LENTELĖ. SĄRYŠIAI TARP ONTOLOGIJOS BEI JUDRIŲJŲ METODIKŲ TERMINŲ

Ontologijos terminas	**Judriųjų metodikų terminai**
Darbo metodas	Darbo metodas, *praktika (pvz. DSDM Atern atveju)*
Fazė	Fazė
Principas	Principas, ypatybė , vertybė
Disciplina	Veikla, disciplina , *sritis*
Praktika	Praktika, mąstysena, filosofija , strategija
Vaidmuo	*Vaidmuo*
Darbo produktas	*Darbo produktas, artefaktas, rezultatas*

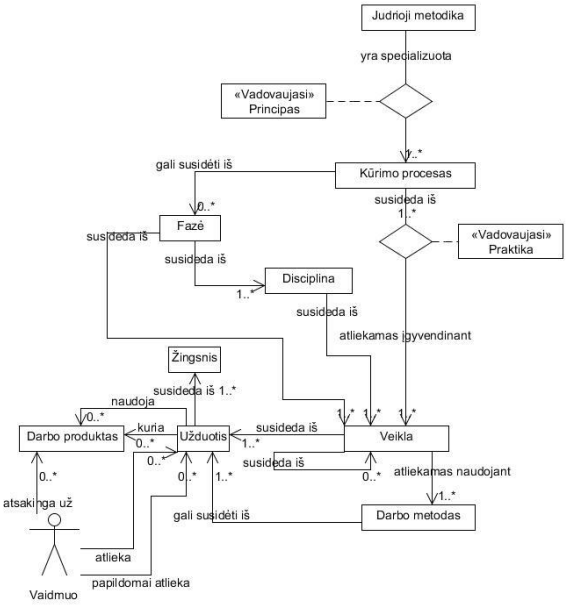
Norint sudaryti nuoseklų ir aiškų judriosios metodikos, kurioje *veiklos*, *užduotys* ir *žingsniai* nėra išimtinai išskirti, kūrimo procesą, reikalinga detali judriosios metodikos literatūros analizė. Tokiose metodikose nėra aiškiai pateikiami jokie atitinkami veiklų *užduočių* ir/ar *žingsnių* terminai bei jų tarpusavio sąryšiai, todėl sudarant judriosios metodikos kūrimo procesą reiktų vadovautis fazių tikslais, disciplinų aprašymais, vaidmenų atsakomybėmis, darbo produktų sukūrimo tikslais, darbo metodų aprašymais bei visu kitu metodikos aprašymu. Šiuose aprašymuose neretai yra pateikiama, kaip ir kas turi būti atliekama ir kuris vaidmuo turi būti atliekamas.

Atlikus literatūros analizę, yra nustatyti judriųjų metodikų ontologijos terminų sąryšiai (1 pav.).

Judrioji metodika vadovaujasi apibrėžtais principais. Procesas gali būti sudarytas iš fazių, o pastarosios įtraukia disciplinas. Disciplinos turi sąrašą veiklų, kurios gali būti skaidomos į smulkesnes veiklas ir/ar užduotis. Veikloms įgyvendinti gali būti naudojami įvairūs darbo metodai. Vieni iš darbo metodų pavyzdžių yra programavimas poromis (angl. pair programming), programavimas greta (angl. Side-by-Side programming), aktyvi suinteresuotoji šalis (angl. active stakeholder), nuoseklus tempas (angl. sustainable pace), kasdieninis susitikimas (angl. stand up meeting), retrospektyva (angl. retrospective) ir kt.

Remiantis S. Jeners, H. Lichter, C.G. Rozenkranz straipsniu [9], SPEM specifikacija [5] identifikuoti vaidmenų ir darbo produktų bei užduočių sąryšiai. Vaidmuo gali būti atsakingas už darbo produktą, kuris gali būti kuriamas ar naudojamas atliekant darbo užduotį, atliekamų pagrindinio arba papildomo vykdytojo. Visi šie sąryšiai pavaizduoti 1 pav.

Remiantis SPEM specifikacija [5], atsakomybės gali būti įvairios: pavyzdžiui, sukurti, peržiūrėti, patvirtinti, priimti ar bendradarbiauti kuriant produktą. Visos atsakomybės identifikuojamos priskiriant darbo produktams būsenas. Pastebėtina, kad [9] straipsnyje nėra identifikuoti jokie darbo produkto ir vaidmens tarpusavio sąryšiai.



1 pav. Judriųjų metodikų ontologijos terminų tarpusavio sąryšiai

IV. FORMALIZUOTO JUDRIŲJŲ METODIKŲ APRAŠYMO METODIKA

Naudojantis apibrėžta judriųjų metodikų ontologija, galima formalizuotai aprašyti judriąsias metodikas taip, kad visų apibrėžimai būtų tokios pačios formos. Pasirinkta aprašyti SPEM [5]. Pagrindinė pasirinkimo priežastis yra ta, kad SPEM skirtas aprašyti programinės įrangos ir sistemų kūrimo procesus ir jų komponentus bei sąryšius tarp jų vartotojui lengvai suprantamu būdu. SPEM įgalina sudėtingus veiksmus bei sąryšius išreikšti diagramomis, todėl sumodeliuoti procesai yra lengviau suprantami ir išmokstami, gali būti vertinami ir analizuojami. Kitas SPEM naudojimo privalumas – įrankių pasirinkimo galimybė, Šiuo metu SPEM yra palaikomas kelių įrankių, tokių kaip Enterprise Architect (leidžia kurti SPEM diagramas) ar Eclipse, kaip Eclipse proceso karkaso projekto (angl. Eclipse Process Framework Project – EPF) dalis. Pastarasis yra nemokamas ir plačiai naudojamas – yra sukurta ir publikuota keleta judriųjų metodikų aprašymų, tokių kaip Scrum, OpenUp, XP ir kiti [6]. Kaip tik šis įrankis buvo pasirinktas aprašyti metodikas SPEM.

SPEM [5] įgalina aprašyti programų kūrimo procesą ar metodiką skirtingais abstrakcijos lygiais, išskirdamas tokius 3 pagrindinius paketus: proceso struktūra (angl. Process structure), valdymo turinys (angl. Managed Content), metodo

turinys (angl. Method Content). Formalizuoto judriųjų metodikų aprašymo SPEM metodika sudaryta iš trijų etapų. Šie etapai išdėstyti taip, kad pirma yra sukuriamas valdymo turinys, po to – metodo turinys ir galiausiai – proceso struktūra. Tokia seka buvo parinkta todėl, kad kiekvieno kito etapo elementams sukurti reikalingi praėjusiuose etapuose sukurti elementai. Atlikus kiekvieną naują etapą, judrioji metodika yra vis labiau detalizuojama ir struktūrizuojama.

1 etapas. Judriųjų metodikų struktūrizavimas turi būti pradedamas sukuriant EPF valdymo turinio elementus. Remiantis III lentele, šiame etape turi būti aprašomi judriųjų metodikų principai, praktikos ir darbo metodai. Kiekviena praktika ir darbo metodas gali vadovautis metodo principais, todėl patartina pirma sukurti principus, o vėliau kuriant praktikas ir darbo metodus juos susieti su sukurtais principais.

Sukūrus EPF gairės elementą, darbo metodas aprašomas gairės elemento visame aprašyme taip, kaip yra aprašyta judriosios metodikos dokumentacijoje. Analogiškai aprašomi principai ir praktikos. Svarbu neišskirti jokių užduočių, kadangi tai bus atliekama kituose etapuose.

Pabaigus šį etapą, metodika yra dokumentuojama natūralia kalba aprašytu tekstu.

III LENTELĖ. JUDRIŲJŲ METODIKŲ ONTOLOGIJOS SUSIEJIMAS SU EPF VALDYMO TURINIU

Judriųjų metodikų ontologijos terminas	**EPF valdymo turinys**	
	Elementas	**Piktograma**
Principas	Koncepcija (angl. concept)	
Praktika	Praktika (angl. practice)	
Darbo metodas	Gairė (angl. guideline)	

2 etapas. Šiame etape turi būti kuriami metodo turinio elementai – darbo produktai, užduotys ir jų žingsniai bei vaidmenys (IV lentelė). Taip pat turi būti sukuriamos disciplinos, kurios yra EPF metodo turinio elementų – užduočių – skirstymo kategorijos.

Judriųjų metodikų darbo produktai turėtų būti aprašomi sukuriant EPF rezultato (angl. deliverable) elementus. Rezultato darbo produktas gali būti pavienis (pvz., specifikacija, darbo planas) arba jį gali sudaryti keli darbo produktai, taip išskiriant darbo produktų dalis, už kurias atsakingi skirtingi vaidmenys.

Šiame etape kiekvienas vaidmuo turi būti siejamas su darbo produktais, už kuriuos jis yra atsakingas. Atsakomybės (pvz., sukurti, priimti, peržiūrėti, patvirtinti) turi būti siejamos, darbo produktams priskiriant būsenas (pvz., sukurtas, priimtas, peržiūrėtas, patvirtintas).

Kiekviena užduotis turi būti siejama su užduotį atliekančiais vaidmenimis bei naudojamais ir kuriamais darbo produktais. Užduotys taip pat siejamos su pirmame etape sukurtais darbo metodais. Užduotį susiejus su darbo metodu, yra parodoma, kad jis yra naudojamas šiai užduočiai atlikti. Jeigu metodikoje užduotis yra išskaidyta į žingsnius, pastaruųjų EPF elementai taip pat turi būti sukuriami.

Antrame etape užduotys turi būti susietos su judriosios metodikos disciplinomis.

Pabaigus šį etapą, yra sukuriama struktūrizuota ir dokumentuota judrioji metodika su apibrėžtais vaidmenimis ir atsakomybėmis, užduotimis ir naudojamais/kuriamais darbo produktais, kiekviena užduotis priskiriama atitinkamai disciplinai.

IV LENTELĖ. JUDRIŲJŲ METODIKŲ ONTOLOGIJOS TERMINŲ SUSIEJIMAS SU EPF METODO TURINIU

Judriųjų metodikų ontologijos terminas	**EPF metodo turinys**	
	Elementas	**Piktograma**
Disciplina	Disciplina (angl. discipline)	
Užduotis	Užduotis (angl. task)	
Žingsnis	Žingsnis (angl. step)	-
Darbo produktas	Rezultatas (angl. deliverable)	
Vaidmuo	Vaidmuo (angl. role)	

3 etapas. Šiame etape yra kuriami EPF procesų struktūros elementai (V lentelė). Remiantis SPEM specifikacijos [5] rekomendacijomis, kuriant judriosios metodikos programų kūrimo procesą, proceso struktūros elementai – gebėjimo šablonas (angl. capability pattern), darbo produkto deskriptorius, užduoties deskriptorius, vaidmens deskriptorius – atitinkamai siejami su metodo turinio elementais – disciplinomis, darbo produktais, užduotimis ir vaidmenimis (2 pav.). Tokiu atveju procesas yra sukuriamas panaudojant jau sukurtus elementus.

Struktūrizavimas turi būti pradedamas nuo gebėjimo šablonų, sudarytų iš veiklų, būdingų tai disciplinai, kurią atitinka proceso šablonas, sukūrimo, kadangi judriosios metodikos proceso modeliui sukurti yra reikalingi jau sukurti gebėjimo šablonų elementai. Kuomet judriosios metodikos proceso modelis yra sudarytas iš proceso šablonų veiklų, ateityje, prireikus atnaujinti veiklas, tereikės atnaujinti jas gebėjimo šablonuose. Atlikus pakeitimus, automatiškai yra atnaujinami atitinkami judriosios metodikos proceso modelio elementai.

Kiekviena gebėjimo šablonui priskirta veikla turi būti suskaidoma į kitas veiklas ir/ar užduotis, kurios sukuriamos panaudojant EPF užduočių deskriptorius, susietus su metodo turinio užduotimis. Detalumo lygis priklauso nuo judriosios metodikos darbo užduočių aprašymo.

Galiausiai kiekvienas gebėjimo šablonas turi būti susietas su atitinkama metodo turinio disciplina, joje nurodant pavyzdinę darbo eigą (angl. reference workflow), susietą su proceso šablonu.

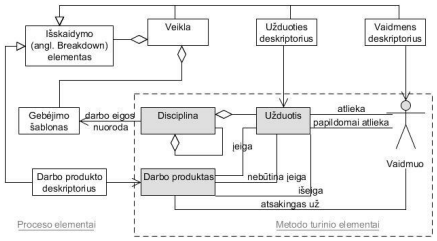
Turint iki galo aprašytus proceso šablonus, gali būti pradedamas kurti judriųjų metodikų proceso modelis. Sukuriamas EPF galutinis procesas (angl. delivery process), priskiriamos judriosios metodikos gyvavimo ciklo fazės (jei yra). Kiekviena fazė (jei jų nėra – procesas) turi būti sudaroma dinamiškai panaudojant gebėjimo šablonuose aprašytas veiklas. Jeigu tam tikra veiklų/užduočių grupė vykdoma iteracijomis, jas reikia sugrupuoti, sukuriant EPF iteracijos elementą.

Šiame etape darbo metodai taip pat gali būti detalizuojami jiems sukuriant darbo užduotis. Antrame etape darbo metodai nebuvo detalizuojami į užduotis dėl to, kad dažniausiai darbo metoduose aprašomos užduotys nėra siejamos su vaidmenimis, šios užduotys yra bendresnės. Pavyzdžiui, DSDM Atern [13] įvardijamas *darbo metodas* – iteratyvus kūrimas (angl. iterative

development) – sako, kad kiekviena iteracija turi būti atliekama perėjus per 4 ciklo fazes – identifikavimas, planavimas, kūrimas ir peržiūra. Šias keturias ciklo fazes galima traktuoti kaip iteratyvaus kūrimo metodo užduotis, ir jos nėra siejamos su vaidmenimis. Kiekvienas procesų struktūros elementas turi būti siejamas su pirmame etape sukurtais principais, darbo metodais ir praktikomis, jei tik judriosios metodikos dokumentacijoje parodomas aiškus elementų sąryšis. Pabaigus šį etapą, judrioji metodika bus visiškai struktūrizuota.

V LENTELĖ. JUDRIŲJŲ METODIKŲ ONTOLOGIJOS TERMINŲ SUSIEJIMAS SU EPF PROCESO STRUKTŪRA

Judriųjų metodikų ontologijos terminas	EPF proceso struktūra	
	Elementas	Piktograma
Programų kūrimo procesas	Galutinis procesas (angl. delivery process)	
Disciplina	Gebėjimo šablonas (angl. capability pattern)	
Fazė	Fazė (angl. phase)	
-	Iteracija (angl. iteration)	
Veikla	Veikla (angl. activity)	
Užduotis	Užduoties deskriptorius (angl. task descriptor)	
Darbo produktas	Darbo produkto deskriptorius (angl. work product descriptor)	
Vaidmuo	Vaidmens deskriptorius (angl. role descriptor)	



2 pav. EPF procesų struktūros ir metodo turinio elementų sąryšiai

V. FORMALIZUOTAS DSDM ATERN APRAŠYMAS

Formalizuotas DSDM Atern [13] metodikos aprašymas atliktas keliais etapais.

1. DSDM Atern terminų susiejimas su apibrėžta judriųjų metodikų ontologija (III skyrius).
2. DSDM Atern struktūrizavimas, remiantis sukurta metodika (IV skyrius).

A. *DSDM Atern terminų susiejimas su apibrėžta judriųjų metodikų ontologija*

Judriųjų metodikų ontologijos terminai – kūrimo procesas, fazė, principas, disciplina, darbo produktas – atitinkamai siejami su DSDM Atern terminais: Atern projekto gyvavimo ciklas, fazė, principas, sritis ir darbo produktas.

Ontologijos terminas *darbo metodas* atitinka DSDM Atern terminus: praktika, darbo metodas. DSDM Atern vadove šie du terminai naudojami kaip sinonimai. DSDM Atern vadove nėra iškirtas joks kitas terminas, kurį galima būtų susieti su ontologijos terminu *praktika*, todėl šis ontologijos terminas neturi DSDM Atern termino atitikties.

DSDM Atern nėra konkrečiai išskirtų veiklų. Pagal judriųjų metodikų ontologiją, proceso fazių tikslai yra pasiekiami, atlikus tam tikrų veiklų rinkinį. Remiantis DSDM Atern gyvavimo ciklo fazių aprašymu ir nurodytais fazių tikslais, identifikuotos veiklos.

Kaip ir veiklos atveju, DSDM Atern nėra konkrečiai išskiriamos užduotys, tačiau detaliai aprašomų vaidmenų atsakomybės, darbo produktai siejami su konkrečiais vaidmenimis – nurodoma, koks vaidmuo kuria, priima, patvirtina, peržiūri ar bendradarbiauja kuriant darbo produktą. Darbo metodų ar sričių aprašymuose taip pat galima rasti vaidmenų susiejimą su tam tikromis užduotimis, pavyzdžiui, reikalavimų srities aprašyme yra nurodoma, kad norint pakeisti reikalavimo aukščiausią prioritetą į žemesnį, turi būti gaunamas verslo vizijos kūrėjo (angl. business visionary) vaidmens patvirtinimas. Vadinasi, viena iš verslo vizijos kūrėjo užduočių yra tvirtinti/atmesti aukščiausių prioritetų keitimą žemesniais. Remiantis vaidmenų atsakomybėmis, darbo produktų aprašymais ir kitu DSDM Atern vadovo [13] tekstu, buvo identifikuotos užduotys ir jų žingsniai.

B. *DSDM Atern struktūrizavimas*

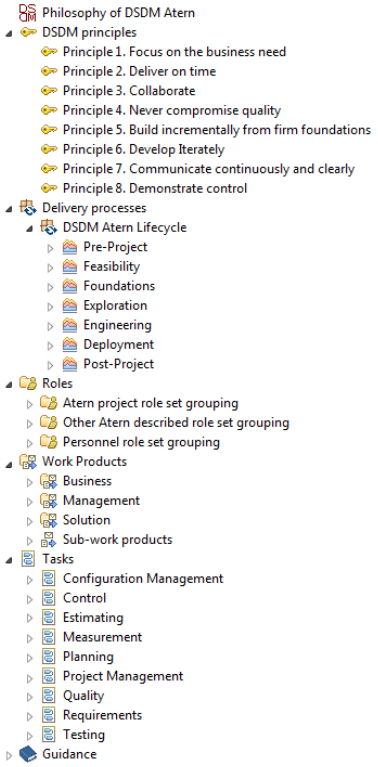
3 pav. pavaizduota EPF įrankiu sukurtos struktūrizuotos DSDM Atern metodikos hierarchija.

Sukurti DSDM Atern metodikos elementai susideda iš pavadinimo, tikslo, trumpo ir viso aprašymo, sąryšių su kitais elementais ir grupavimo (jei toks buvo nurodytas). Struktūrizuotą DSDM Atern metodiką sudaro gyvavimo ciklas, darbo metodai, suskirstyti į grupes vaidmenys, darbo produktai, principai, vaidmenų atliekamos užduotys ir veiklos. Daugumai darbo produktų yra pateikiami produktų šablonai, kurie yra siūlomi DSDM konsorciumo. Išskiriant veiklas, užduotis ir užduočių žingsnius, buvo susidurta su sunkumais jas grupuojant, kadangi šie elementai buvo sudaromi, remiantis DSDM Atern vadovu [13], kuriame elementai yra išsibarstę po visą tekstą. Darbas buvo atliktas, išrašius visus įmanomus aprašytus darbus DSDM vadove, susiejus juos su vaidmenimis bei darbo produktais, priskyrus disciplinai, kuriai priklauso, ir vėliau sugrupavus į veiklas, užduotis ir žingsnius.

Galutinis struktūrizavimo rezultatas yra vertingesnis už originalią dokumentaciją, nes pateikia tikslius sąryšius tarp visų elementų, leisdamas matyti kiekvieno elemento naudojimą bei atskleisdamas metodikos dokumentacijos trūkumus, pavyzdžiui, egzistuoja aprašytos užduotys, bet neaišku kuris vaidmuo vykdo; egzistuoja darbo produktų, bet neaišku, kada jie sukuriami – kokią užduotį atliekant. Sąryšių tarp elementų nebuvimas parodo dokumentacijos trūkumą.

Struktūrizuojant DSDM Atern metodiką buvo rasta neatitikimų. Pavyzdžiui, vienas iš darbo produkto – sprendimo peržiūros strategija (angl. Solution review strategy) – sukūrimo tikslų yra nustatyti, kokie testavimo ir/ar peržiūros darbo metodai (angl. technique) turi būti naudojami visose projekto fazėse. Pats dokumentas yra sukuriamas trečioje gyvavimo ciklo fazėje – įsteigimo fazėje, todėl akivaizdu, kad ne visose gyvavimo fazėse testavimo ir/ar peržiūros darbo metodai (angl. technique), apibrėžti šiame dokumente, turi būti naudojami. Taip pat to paties darbo produkto aprašyme yra nurodyta, kad

šis dokumentas naudojamas/keičiamas tik trijose iš septynių projekto gyvavimo fazių. Kita problema – DSDM vadovas yra ne visas ir netikslus. Pavyzdžiui, rasta darbo produkto sąvoka – *sprendimo architektūros apibrėžimas* (angl. Solution architecture definition), kuri panaudota visame dokumente du kartus, bet niekur nėra apibrėžiamas toks darbo produktas. Galima spėti, kad čia yra kalbama apie sistemos architektūros dokumentą (angl. System architecture definition). Esant tokiems neatitikimams, nėra aišku, ar struktūrizuojant metodiką palikti taip kaip yra – su rastais neatitikimais, ar remiantis visa kita vadovo [13] informacija, ją reiktų keisti, o pakeitus, ar galima teigti, kad aprašymas pilnai atspindi visą DSDM Atern metodiką.



3 pav. Struktūrizuoto DSDM Atern hierarchija

VI. IŠVADOS IR TOLESNI TYRIMAI

Išankstinis metodikos įvertinimas pagal programų kūrimo proceso modelius prieš įdiegimą leistų įmonei sutaupyti daug laiko ir lėšų. Tyrimo metu sukurta formalizuoto judriųjų metodikų aprašymo metodika, kuria remiantis aprašius judriąsias metodikas, palengvėja skirtingų judriųjų metodikų tarpusavio lyginimas ir vertinimas pagal programų kūrimo procesų vertinimo modelius. Formalizuotas judriųjų metodikų aprašymas yra vertingesnis už originalią dokumentaciją, nes pateikia tikslius sąryšius tarp visų elementų, leisdamas matyti kiekvieno elemento naudojimą, o sąryšių tarp elementų nebuvimas atskleidžia metodikos dokumentacijos trūkumus.

Formalizuoto judriųjų metodikų aprašymo metodika sukurta, remiantis patikslinta ir papildyta judriųjų metodikų ontologija. Pateikti sąryšiai tarp judriųjų metodikų ontologijos ir SPEM elementų. Išskiriami trys formalizuoto aprašymo sudarymo etapai. Nagrinėjamas formalizuotas DSDM Atern aprašymas, sudarytas EPF įrankiu pagal sukurtą metodiką. Aptariami iškilę sunkumai dėl DSDM Atern aprašymo trūkumų.

Ateityje planuojama įvertinti DSDM Atern pagal programų kūrimų proceso vertinimo modelius, tokius kaip CMMI ir ISO/IEC 15504, identifikuojant, kokį aukščiausią proceso brandą/gebėjimą įmonė galės įgyti visiškai įdiegus DSDM Atern Taip pat numatoma patobulinti judriųjų metodikų aprašymo metodiką įtraukiant ne tik judriuosius, bet ir kitus programų sistemų kūrimo metodikas. Šiam tikslui įgyvendinti reiktų išnagrinėti kitas metodikas, atnaujinti ontologiją ir atitinkamai keisti metodiką. Atlikus pakeitimus, bus sudaryta galimybė sukurti vienodai formalizuotus aprašymus ne tik judriųjų, bet ir kitų metodikų, pvz., RAD, RUP. Turint tokius aprašymus palengvės judriųjų ir kitų metodikų tarpusavio lyginimas, bus sudaryta galimybė sukurti vertinimo metodiką pagal kurį nors programų kūrimo procesų vertinimo modelį, tinkančią visoms metodikoms.

LITERATŪRA

[1] "Manifesto for Agile Software Development". Agile Alliance, 2001. [žiūrėta 2013-02-17]. Prieiga per internetą: http://agilemanifesto.org/.

[2] T. Dingsøyr, S. Nerur, V. Balijepally, N. B. Moe. A decade of agile methodologies: Towards explaining agile software development. Journal of Systems and Software, Vol. 85, Issue 6, 2012, pp. 1213–1221. [Žiūrėta 2013-04-07]. Prieiga per internetą: http://www.sciencedirect.com/science/article/pii/S0164121212000532#

[3] S. Peldžius, S. Ragaišis. Reikalavimai tarpiniam programų kūrimo proceso modeliui. Informacijos mokslai, Vilniaus universitetas, 2011, p. 138-145.

[4] D. Parsons. An Ontology of Agile Aspect Oriented Software Development. Res. Lett. Inf. Math. Sci., 2011, Vol. 15, 11 p. [Žiūrėta 2013-02-04]. Prieiga per internetą: http://www.massey.ac.nz/~dpparson/AgileDPy.pdf.

[5] Software & Systems Process Engineering Meta-Model Specification. Versija 2.0, Object management group. 2008, 236 p. [žiūrėta 2013-02-04]. Prieiga per internetą: http://www.omg.org/spec/spem/2.0/pdf.

[6] EPF wiki internetinis puslapis. [Žiūrėta 2013-04-07] Nuoroda per internetą: http://epf.eclipse.org/

[7] N. Potter, M. Sakry. Implementing Scrum (Agile) and CMMI Together. The Process Group, Vol. 16 No. 2, 2009, 6 p. [Žiūrėta 2013-02-04]. Prieiga per internetą: http://www.processgroup.com/pgpostmar09.pdf

[8] E. Irrazabal1, F. Vásquez, R. Díaz, J. Garzás. Applying ISO/IEC 12207:2008 with SCRUM and agile methods. SPICE 2011, CCIS 155, 2011, pp. 169-180.

[9] S. Jeners, H. Lichter, C. G. Rosenkranz. Efficient Adoption and Assessment of Multiple Impovement Reference Models. e-Informatica Software Engineering Journal, Vol. 7, Issue 1, 2013, pp. 15–24. [Žiūrėta 2013-04-06] Prieiga per internetą: http://www.e-informatyka.pl/attach/e-Informatica_-_Volume_7/eInformatica2013Art2.pdf.

[10] I. A. Zualkernan. An ontology-driven approach for generating assessments for Scrum software process. New Trends in Software Methodologies, Tools and Techniques. IOS Press, 2008, pp. 190-205.

[11] Resource Description Framework (RDF). World Wide Web Consortium (W3C) Recommendation, 2004. [žiūrėta 2013-02-17] Prieiga per internetą: http://www.w3.org/TR/rdf-concepts.

[12] P. Ceravolo, E. Damiani, M. Marchesi, S. Pinna, F. Zavatarelli. A Ontology-based Process Modelling for XP. Software Engineering Conference, 2003. Tenth Asia-Pacific , 2003, pp. 236-242.

[13] DSDM Atern, The Handbook, DSDM Consortium, 2008, 201 p.

Vietovės nustatymo (GPS) paklaidų atsiradimo priežastys, jų mažinimo metodai ir tyrimai

Nerijus PAŽERECKAS
Programų inžinerijos katedra
Informatikos fakultetas
Kauno technologijos universitetas
Kaunas, Lietuva
Nerijus.pazereckas@ktu.lt

Santrauka. Orientuotis aplinkoje galima įvairiai, tačiau ne visi būdai pritaikomi mechanikoje. GPS įrenginiai ir metodai plačiai paplitę pasaulyje ir naudojami įvairiose srityse, tarp jų ir autonominiams įrenginiams. GPS įrenginių yra keletas rūšių, visi jie veikia tuo pačiu principu, tačiau jų tikslumas skiriasi. Pastaruoju metu mokslininkai šiai nagrinėjamai sričiai skiria didesnį dėmesį ir atlieka tyrimus, kurių pagrindinis tikslas - išsiaiškinti netikslumų atsiradimų priežastis ir sumažinti GPS įrenginių paklaidas iki minimalesnių. Straipsnyje bus pateikiami jau esami ir naujai ištirti programiniai ir matematiniai metodai, padedantys sumažinti GPS koordinačių paklaidas.

Išsamiau su platesniais tyrimų aprašymais ir jų rezultatais galima susipažinti straipsnio autoriaus magistriniame darbe.

Reikšminiai žodžiai *— GPS, koordinatės, tyrimas, paklaidos, programavimas, metodai.*

I. ĮŽANGA

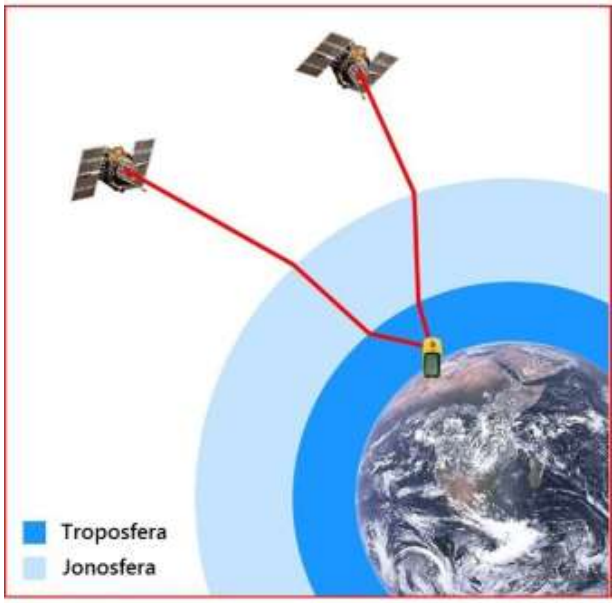
Rinkoje yra nemaži kiekis vietovės nustatymo prietaisų ir jau egzistuojančių įgyvendintų sprendimų, padedančių iš taško A nuvykti į tašką B. Vieni prietaisai yra labai tikslūs – tikslumas siekia iki 1 cm, o kiti, namų vartotojams skirti, prietaisai koordinates nurodo iki kelių dešimčių metrų tikslumu. GPS įrenginių tikslumą riboja ir vyriausybinės saugumo įstaigos vengdamos teroristinių aktų, todėl specialiai iškraipo GPS koordinačių tikslumą.

Norint gauti tikslesnes vietovės koordinates dažniausiai reikia naudoti atitinkamą GPS prietaisą. Norint gauti tikslesnes koordinates naudojantis namų vartotojams skirtu pigiu GPS prietaisu, reikia kitais būdais mažinti paklaidas. Vienas iš būdų yra naudoti bent tris GPS imtuvus, nutolusius vienas nuo kito per tam tikrą atstumą, tačiau toks būdas tinkamas ne visiems atvejams. Kitas būdas sumažinti paklaidas – naudoti tam tikrus matematinius, statistinius ar programinius metodus, kurių pagalba būtų patikslinamos jau gautos koordinatės.

II. GPS PAKLAIDOS IR JŲ ATSIRADIMO PRIEŽASTYS

A. *Atmosferos trikdžiai ir kliūtys žemėje*

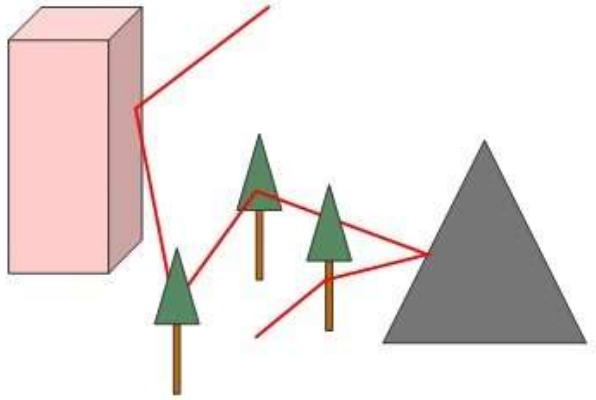
Didžiausią įtaką GPS imtuvų tikslumui turi atmosferos trikdžiai ir įvairios kliūtys bei veiksniai, esantys žemėje. Signalas sklisdamas iš palydovo turi nukeliauti labai didelį atstumą. Signalas sklisdamas link žemės turi pereiti per kosmoso sferą, kirsti ozono sluoksnį, o tada dar sklisti iki GPS imtuvo oro terpėje, kurioje taip pat yra įvairių trikdžių, kaip, pavyzdžiui, kitų diapazonų dažniai. Taip sklisdamas signalas sulėtėja ir susilpnėja, todėl atsiranda netikslumai. Tyrimų parodyta, kad jonosferos įtaka signalui yra gan didelė, dėl šių trikdžių atsiranda 3,6 m. paklaida. Egzistuoja įvairūs algoritmai ir priemonės, kurios šiuos trikdžius sumažina [1].



1 pav. Atmosferos trikdžiai

Imtuvų tikslumą įtakoja ir objektai, esantys netoliese imtuvo. Jeigu imtuvą naudojantis vartotojas yra kalnuose arba tankiai apgyvendintame plote, atsiranda atspindžio efektas, pavaizduotas antrame paveiksle. Signalą įtakoja bet kokie trikdžiai, o jeigu signalas negali sklisti tolyn troposferoje, jis arba būna sugeriamas, arba atsimuša į kokį nors objektą ir atsispindi. GPS imtuvai dažnai priima ne iš palydovų siunčiamą signalą, o nuo kokio nors objekto (kalno, pastato ar pan.) atspindėtą signalą. Toks signalas jau būna iškraipytas ir trikdo imtuvo darbą. Į signalo iškraipymą įeina kelios dedamosios, tačiau pagrindinė yra laikas ir palydovo padėtis horizonte. Jeigu

signalas buvo daug kartų atspindėtas, palydovo padėtis ir signalo perdavimo laikas bus žymiai pasikeitę ir todėl atsiras paklaida. Labai svarbu, kad imtuvas būtų ne uždaroje patalpoje ir kad imtuvo antenai ar pačiam imtuvui (jei antena integruota) būtų matomas horizontas [2].



1 pav. GPS signalo atspindys žemėje

B. Klaidos, atsirandančios dėl palydovų ir jų išsidėstymo

Kartais GPS koordinačių netikslumai atsiranda ir nuo palydovų kaltės. GPS imtuvai veikia, kai horizonte mato ir signalus gauna ne mažiau kaip iš 4 palydovų. Jeigu nors vienas iš šių palydovų nustoja siųsti signalus arba atsiranda trikdžių, imtuvas pasimeta ir neberodo vartotojo padėties arba rodo neapibrėžtą padėtį. Palydovuose gali atsirasti įvairių trikdžių, vienas iš tokių yra efemeridės anomalija. Šios anomalijos metu signalas dėl neaiškių priežasčių, vykstančių atmosferoje, būna iškraipomas arba apskritai nutraukiamas. Kitos priežastys gali būti palydovo laikrodžio klaida, kuri yra labai maža, nes naudojamas atominis laikrodis, tačiau paklaida visada yra ir daugiapaketinio signalo klaida. Pastaroji klaida atsiranda, kai signalas GPS imtuvo anteną pasiekia vėliau nei tikėtasi ir atsiranda koordinačių paklaida. Taip pat šią klaidą gali iššaukti ir pats imtuvas, jeigu nespėjo laiku apdoroti visų gautų signalo paketų. Palydovinio laikrodžio gedimas ar netikslumas koordinačių tikslumą įtakoja apie 0,6 m [3].

Kartais palydovai orbitoje išsidėsto blogai (kaip pavaizduota trečiame paveiksle), tada jų siunčiama vietos informacija būna klaidinga. Tai reiškia, kad palydovai yra išsidėstę per arti vienas kito ir siunčią skirtingas judėjimo koordinates, tada imtuvas gauna ne panašias, o visiškai skirtingų vietų koordinates, ir atsiranda paklaida nuo kelių iki keliasdešimt metrų. Kai palydovai išsidėstę toliau vienas nuo kito (yra skirtinguose žemės vietose), jų siunčiama informacija būna tikslesnė. Gerą palydovų išsidėstymą vaizduoja ketvirtas paveikslas.

3 pav. Blogas palydovų išsidėstymas

4 pav. Geras palydovų išsidėstymas

C. Kiti trikdžiai

Be prieš tai išvardintų trikdžių atsiradimo priežasčių, yra ir kitų trikdžių sukėlėjų. Kadangi signalai sulėtėja, būna sunkiau nustatyti palydovo padėtį ir tokio trikdžio reikšmė yra 0,6m koordinačių tikslumo iškraipymas. Taip pat nemažą įtaką tikslumui turi ir patys imtuvai, kurie ilgesnį laiką veikdami dažnai perpildo savo atmintį, arba atsiranda kitokių gedimų ir jų tikslumas iškraipomas apie 1,2 m paklaida. Bet be šių paklaidų atsiradimo priežasčių, didžiausią paklaidą sukelia patys žmonės, JAV gynybos ministerija teigia, kad jie specialiai iškraipo signalą sukeldami 7,5 m paklaidą.

III. GPS KOORDINAČIŲ TIKSLUMO DIDINIMO BŪDAI

GPS koordinates galima tikslinti keliais būdais: galima stiprinti GPS imtuvo signalą, naudoti papildomus prietaisus - pavyzdžiui, giroskopus ar magnetometrus, arba galima koordinates patikslinti, panaudojus matematinius metodus.

A. GPS signalo stiprinimo galimybės

Vienas plačiausiai naudojamų GPS signalo stiprinimo priemonių yra signalo priėmimas iš WAAS ir EGNOS palydovų.

Tai speciali GPS matavimų paklaidos mažinimo sistema. EGNOS signalai priimami iš 3-4 geostacionarinių palydovų, esančių virš Atlanto ir Indijos vandenynų. WAAS palydovai yra virš Amerikos, ir WAAS sistema naudojama tik JAV. Šie palydovai žemės atžvilgiu yra fiksuotoje padėtyje. Signalai bus priimami automatiškai (jei tik imtuvas turi tokią funkciją), nieko papildomai įsidiegti nereikia. Svarbu, kad matavimų metu EGNOS palydovų niekas neužstotų, nes jų padėtis atmosferoje yra gan žema, tik apie 15-20 laipsnių virš horizonto pietvakarių ir pietryčių kryptimis.

Jeigu imtuvas sugeba priimti signalus iš EGNOS ar WAAS palydovų ir horizonte mato daugiau nei 10 aktyvių kitų GPS palydovų, tai vietos koordinačių tikslumas gali būtų padidintas ir paklaida gali sumažėti net iki 1,5 – 1 metro. EGNOS ir WAAS palydovai yra išsidėstę taip, kad padengtų atitinkamą žemyną. Imtuvai, gebantys priimti šiuos signalus, gauna ne vien patikslintas koordinates, bet ir turi programinę įrangą, kuri atitinkamus signalų blokus apdoroja naudojant sferinių harmoninių, Teiloro eilutės arba trimatės jonosferinės eilutės išskleidimo metodus. Panaudojus šiuos metodus, atliekami atitinkami skaičiavimai ir taip dar labiau patikslinamos vartotojo GPS koordinatės [4].

Galima gerinti GPS signalą panaudojant papildomą aparatūrinę įrangą. Papildomai prie GPS imtuvo galima naudoti vieną ar kelis sensorius, skirtus taip pat nustatyti koordinates:

- giroskopus,
- žvaigždžių sensorius,
- Saulės sensorius,
- Žemės horizonto sensorius,
- magnetometrus.

Naudojant iš palydovų gaunamus signalus ir papildomai įdiegiant vieną iš šių sensorių, koordinates galima nustatyti kur kas tiksliau. Jeigu iš palydovų siunčiamų signalų sužinome apytikslį buvimo plotą, tai papildomai panaudoję vieną iš sensorių, žinome įvairių objektų (žvaigždžių, saulės) padėtį orbitoje duotuoju laiko momentu ir gauname atitinkamas koordinates. Iš palydovų ir sensorių gautas koordinates palyginame, atliekame skaičiavimus matricų pagalba ir objekto žemėje padėties plotą sumažiname iki kelių kartų [5].

IV. KOORDINAČIŲ TIKSLINIMAS MATEMATINIAIS METODAIS

Neturint papildomų prietaisų, tik įprastą GPS imtuvą programinėje įrangoje, galime naudoti įvairius papildomus matematinius skaičiavimus. Atliekant tam tikrus matematinius skaičiavimus, iš GPS signalo galima eliminuoti trikdžius, gautus jonosferoje, arba dalinai pašalinti atspindžio efektą. Taip pat galima patikslinti koordinates atliekant skaičiavimus su gautais apytiksliais duomenimis.

A. Pozicijos ir aukščio nustatymas naudojant pradinius duomenis

GPS imtuvai, veikiantys EGNOS sistemoje, turi dvi antenas, kurios iš palydovų priima esamą poziciją ir nukrypimą nuo kurso ir aukštį jūros lygyje. Jeigu priimtume, kad pradinis vietos etalonas yra iš pirmosios antenos gauta koordinatė, o antrosios antenos priimtą informaciją naudotume parametrizuoti nukrypimui, tai antrosios antenos gaunamą informaciją būtų galima išskaidyti į du elementus: antrąją pozicijos koordinatę ir aukštį. Krypties kintamasis ψ ir vertikalioji projekcija θ, yra skaičiuojami, panaudojus bE, bN ir bU koordinates ir įstačius į formulę (1):

$$\psi = tan^{-1}\left(\frac{b_E}{b_N}\right),\ \theta = tan^{-1}\left(\frac{b_U}{\sqrt{b_N^2+b_E^2}}\right) \qquad (1)$$

Formulėje (1) bE, bN ir bU yra išskaidytos GPS koordinatės į sferinius kintamuosius, o apskaičiuoti kintamieji naudojami patikslinti koordinatėms, kai naudojamasi kokia nors transporto priemone, arba GPS imtuvas sumontuotas aukščiau virš horizonto. Dažniausiai tai naudojama panaikinti paklaidas, atsiradusias dėl automobilio aukščio ir dydžio [6].

B. Klaidų išskleidimas iš GNSS kompaso

Klaidų išskleidimas iš GPS kompaso yra vienos epochos ir susideda iš dviejų dalių: pradinio vektoriaus klaidos ir pradinės koordinatės klaidos. Apskaičiuota dispersinė matrica iš pradinio vektoriaus parodo pradinės vietos klaidas, kurias įtakoja palydovų išsidėstymas ir vietovės matavimų tikslumas. Pozicijos transformavimo klaidos atsiranda dėl pradinio vektoriaus ilgio ir jo parametrų.

Pradinis vektorius Rytai-Šiaurė-Aukštis rėmuose gali būti išreiškiamas formule:

$$b = \begin{bmatrix} b_E \\ b_N \\ b_U \end{bmatrix} = \begin{bmatrix} l\sin\psi\cos\theta \\ l\cos\psi\cos\theta \\ l\sin\theta \end{bmatrix} = \Omega(\Lambda) \qquad (2)$$

Čia Ω - nelinijinis operatorius ir $\Lambda = (\psi\ \theta\ l)^T$. Šios lygtys gali būti pateikiamos nauja išraiška:

$$p = \begin{bmatrix} l_0 cos\psi_0 cos\theta_0 & -l_0 sin\psi_0 sin\theta_0 & sin\psi_0 cos\theta_0 \\ -l_0 sin\psi_0 cos\theta_0 & -l_0 cos\psi_0 sin\theta_0 & cos\psi_0 cos\theta_0 \\ 0 & l_0 cos\theta_0 & sin\theta_0 \end{bmatrix}$$

$$\begin{bmatrix} \delta b_E \\ \delta b_N \\ \delta b_U \end{bmatrix} = p \times \begin{bmatrix} \delta\psi \\ \delta\theta \\ \delta l \end{bmatrix} \qquad (3)$$

Trečioje formulėje pateikta Teiloro taškų lygtis $\Lambda_0 = (\psi_0\ \theta_0\ l_0)^T$. Panaudojant dispersijos ir kovariacijos taisyklę, matricą Λ galima apskaičiuota pagal (4) lygtį:

$$Q_\Lambda = J^T Q_{b(a)} J \qquad (4)$$

Čia J yra matrica, kuri rodo pradinio klaidos vektoriaus parametrus trečiojoje lygtyje. Svarbu žinoti, kad $Q_{b(a)}$ parametrų niekaip negalima įtakoti, nes jie priklauso tik nuo horizonte matomų palydovų išsidėstymo ir signalui sukeliamų trikdžių (iškraipymų) [6].

Panaudojus ketvirtąją (4) formulę, galima patikslinti gaunamas koordinates, pašalinus iš jų triukšmus ir iškraipymus, gaunamus iš palydovų. Pritaikius šį metodą, koordinačių tikslumas padidėja iki 50%.

C. Koordinačių korekcija taikant statistinius matematinius metodus

Dažnai GPS imtuvas geba priimti daugiau nei vieną koordinatę tuo pačiu metu. Pavyzdžiui, tyrimui naudojamas GPS imtuvas, turintis Sirf star III mikroprocesorių, atnaujina informaciją kas sekundę ir kiekvieno atnaujinimo metu priima nuo vienos iki trijų koordinačių. Tiriant koordinačių išsisklaidymą horizonte buvo pastebėta, kad tarp tuo pačiu momentu gautų koordinačių dažnai būna gan didelis atstumas ir nemažas atstumas nuo pradinės fiksuotos koordinatės.

Pastebėta, kad tuo pačiu laiko momentu gaunama daugiau nei viena koordinatė, todėl buvo nuspręsta patikrinti, kiek pasikeis paklaida pritaikant aritmetinio vidurkio formulę (5):

$$x = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (5)$$

Čia n – kintamųjų skaičius (mūsų atveju koordinačių skaičius), x_i – koordinatės, x – naujai išvesta koordinatė.

Taip paprastai aritmetinio vidurkio formulės, skaičiuojant vidurkius erdvėje, pritaikyti negalima, todėl buvo išvesta GPS koordinačių vidurkio formulė.

Pirmiausiai koordinatės išskaidomos į atskiras dedamąsias:

$$\begin{aligned} x &= \cos(lat)\cos(lon) \\ y &= \cos(lat)\sin(lon) \\ z &= \sin(lat) \end{aligned} \qquad (6)$$

Čia x – aukštis erdvėje, y – ilgis, z – vektorius erdvėje, lat – platumos laipsniai, lon – ilgumos laipsniai. Norint skaičiuoti koordinatės vidurkį, reikia skaičiuoti kiekvieno erdvėje esančio kintamojo x, y, z vidurkius, įstačius į vidurkio formulę (5).

Gautus tiesinius duomenis reikia paversti atgal į sferines koordinates:

$$\begin{aligned} lon &= \arctan(y, x) \\ hyp &= \sqrt{x^2 + y^2} \\ lat &= \arctan(z, hyp) \end{aligned} \qquad (7)$$

Čia hyp – normalizuotas vektorius, lon – ilgumos sferinė koordinatė, lat – platumos sferinė koordinatė.

Šis metodas gali būti taikomas tik statinių koordinačių patikslinimui, arba kai jau žinomas tikslus maršruto kelias, nes kitu atveju gautos koordinatės gali neatitikti fiksuotos koordinatės pradinio. Judant gaunamos koordinatės priklauso nuo laiko, jeigu judėjimo greitis didelis, duotuoju momentu (vienos sekundės intervale) gautos koordinatės gali būti fiksuojamos judėjimo metu, kai objektas jau bus kitoje vietoje nei priimtos koordinatės. Tačiau jeigu žinomas maršrutas ir judėjimo greitis, galima gautas koordinates lyginti su nustatyto maršruto koordinatėmis ir, panaudojus judėjimo greitį, apytiksliai apskaičiuoti vietą šiuo metu judančio objekto kuriam fiksuojamos koordinatės.

D. Koordinačių tikslinimas taikant maršruto žingsninio projektavimo metodą

Šis metodas gali būti taikomas, kai žinomas judėjimo maršrutas ir visi judėjimo maršrutu taškai (koordinatės). Metodo esmė - projektuoti kitą judėjimo GPS koordinatę toliau reikiamos nuvykti koordinatės. Tarkime, jeigu žinome, kad reikia judėti tiesiai 10 metrų, tai kitą koordinatę nustatome ta pačia kryptimi, tik kelis kartus toliau. Projektuojamas judėjimo vektorius didesnis nei reikia todėl, kad žinoma (maršruto kelyje nustatyta) koordinatė nepapultų į GPS signalo paklaidos lauką. Vektoriumi judame ta pačia kryptimi, tik ne iki vektoriaus pabaigos, o tiek metrų, kiek reikia (šiuo atveju 10).

Pavyzdžiui, jeigu žinoma, kad reikia nuo taško A judėti n metrų kryptimi d iki koordinatės B, tai kelią n padauginame iš nusistatyto kintamojo k ir suskaičiuojame naują koordinatę. ir gauname naują vektorių. Tai galima išreikšti formulėmis:

$$\begin{aligned} \varphi_2 &= \text{asin}(\sin(\varphi_1) * \cos\left(\frac{d}{R}\right) + \cos(\varphi_1) * sin\left(\frac{d}{R}\right) * \cos(\theta)) \\ \text{a}_1 &= \sin(\theta) * sin\left(\frac{d}{R}\right) * \cos(\varphi_1) \\ \text{a}_2 &= \cos\left(\frac{d}{R}\right) - \sin(\varphi_1) * \sin(\varphi_2) \\ \lambda_2 &= \lambda_1 + \text{atan2}(a_1, a_2) \end{aligned} \qquad (8)$$

Formulėje (8) φ – platumos koordinatės, λ – ilgumos koordinatės, θ – maršruto kryptis laipsniais, kai atskaitos koordinatė šiaurės kryptimi, d – vektoriaus (kelio) ilgis, R – žemės spindulys ($\frac{d}{R}$ – kampinis vektoriaus atstumas radianais).

Taikant šį metodą galima sumažinti nuokrypį nuo maršruto kelio.

V. STATISTINIŲ MATEMATINIŲ IR MARŠRUTO ŽINGSNINIO PROJEKTAVIMO METODŲ PRITAIKYMO TYRIMO REZULTATAI

Tyrimui naudojamas GPS imtuvas su Sirf star III mikroprocesorių 20 palydovų. Tyrimas buvo atliktas didmiesčio teritorijoje.

Koordinatės renkamos pasirinktos vietos fiksuotame taške, pagal pasirinktą fiksuotą atskaitos koordinatę, kurioje stacionariai laikomas GPS imtuvas renka duomenis iš palydovų.

1 lentelė. Pradiniai GPS paklaidų duomenys

Bandymo n.	Koord. skaičius	Maks. Paklaida (m)	Min. Paklaida (m)	Vid. Paklaida (m)
1	55	0.402	0,188	0,290
2	55	2,659	0,567	1,670
3	55	0,960	0,092	0,467
4	53	0,979	0,093	0,448
5	55	1,432	0,106	0,823
6	145	1,126	0,093	0,682
7	109	3,146	0,188	1,479
8	118	1,126	0,093	0,633
9	163	3,659	0,188	2,061
10	275	3,659	0,188	2,162

Lentelėje „Pradiniai GPS duomenys“ koord. skaičius – bandymo metu fiksuotų koordinačių skaičius, maks. Paklaida – maksimali koordinačių paklaida nuo pradinės fiksuotos koordinatės; min. paklaida – minimali paklaida; vid. paklaida – vidutinė paklaida, gauta aprašytojo bandymo metu.

Pritaikius koordinačių korekcijos formulę (7), gautos naujai sudarytos koordinatės, kurių nukrypimų nuo pradinio taško duomenys aprašyti antroje lentelėje „Paklaidos po korekcijos“:

2 lentelė. GPS duomenų paklaidos, pritaikius koordinačių korekcijos algoritmą

Bandymo n.	Koord. skaičius	Maks. Paklaida (m)	Min. Paklaida (m)	Vid. Paklaida (m)
1	28	0,308	0,108	0,201
2	27	2,015	0,071	1,112
3	28	0,922	0,054	0,442
4	28	1,375	0,213	0,785
5	28	0,474	0,013	0,181
6	73	1,081	0,054	0,638
7	55	3,047	0,108	1,412
8	59	1,081	0,054	0,588
9	82	3,547	0,108	1,980
10	137	3,547	0,108	2,147

Iš antroje lentelėje gautų duomenų matoma, kad pritaikius matematinį vidurkio modelį, koordinačių paklaidos sumažėjo. Maksimalios paklaidos sumažėjo nuo 0,1 iki 2,5 metro, o vidutiniškai paklaidos buvo sumažintos pakoreguojant jas 0,25 m.

Žingsninio projektavimo metodas (8) buvo pritaikytas judant sudarytu maršrutu. Pritaikius šį metodą, tikimybė išklysti iš sudaryto maršruto sumažėja 20%. Nenaudojant metodo buvo atlikta 10 bandymų, iš kurių 6 kartus buvo išklysta už sudaryto maršruto ribų. Pritaikius žingsninio projektavimo metodą ir atlikus kitus 10 bandymų, už maršruto ribų buvo išklysta tik 4 kartus.

VI. IŠVADOS

Pagal atliktus tyrimus ir gautus rezultatus galima matyti, kad matematiniai statistiniai metodai gali padėti sumažinti koordinačių paklaidas nuo 0,1 iki 2,5 metro. Koordinačių korekcijos dydis priklauso nuo gautų pradinių koordinačių sklaidos ir koordinačių kiekio duotuoju laiko momentu (vienos sekundės intervale). Pradėtas tirti žingsninio projektavimo metodas, kuris teoriškai turi padėti pakoreguoti maršruto kryptį ir judėjimą maršrutu neiškrypstant iš kelio, tačiau tai dar reikia įrodyti, atlikus detalesnius skaičiavimus. Pradiniai skaičiavimai rodo, kad tiriant metodo naudą, iš sudaryto maršruto koordinačių buvo išklysta 20% mažiau nei netaikant metodo.

Toliau planuojama tirti statistinių metodų pritaikomumą GPS koordinačių tikslumui padidinti ir bandymų metu ištirti pasirinktus metodus atskirai, ir panaudojant kelis metodus kartu.

LITERATŪRA

[1] J. Blanch, Using kriging to bound satellite ranging errors due to the ionosphere, A Dissertation submitted to the department of Aeronautics and astronautics and the committee on graduate studines of Stanford university, 2003 December. per internetą: <http://waas.stanford.edu/~wwu/papers/gps/PDF/Thesis/JuanBlanchThesis03.pdf>.

[2] T. H. Dixon, An introduction to the global positioning system and some geological applications, *Reviews of Geographysics, 29, 2 / May 1991, P.N. 91RG00152, 249-276.*

[3] S. Clark, GPS error - an exciting technical report, *BYU (Brigham Young University), Computer Science, 2010, per internetą: <https://facwiki.cs.byu.edu/HCMI/images/a/ac/Gpstech-Feb_4.pdf>*.

[4] T. H. Witte, A. M. Wilson, Accuracy of WAAS-enabled GPS for the determination of position and speed over ground, *Journal of Biomechanics, Volume 38, Issue 8, August 2005, 1717–1722.*

[5] M. Jensen, T. Bak. High Precision GPS Attitude Determination using the AAU Testbed, *STVF, Doc. No.: I-2000-602801, 2000-06-27.*

[6] W. Chen, H. Qin, Y. Zhang, T. Jin, Accuracy assessment of single and double difference models for the single epoch GPS compass, *Advances in Space Research, 49* (2012) *725–738.*

Disko analizei atsparus duomenų slėpimo metodas

Grigas Petraitis, Nerijus Morkevičius
Kompiuterių katedra
Kauno technologijos universitetas
Kaunas, Lietuva
grigas.petraitis@stud.ktu.lt

Santrauka. Šiame darbe analizuojama slaptos informacijos slėpimo diskiniuose kompiuterių kaupikliuose problema. Straipsnyje pasiūlytas ir išanalizuotas naujas metodas, skirtas informacijos slėpimui klasterinėse failų sistemose. Metodas naudoja kelis dengiančiuosius failus ir informaciją saugo keisdamas šių failų klasterių tarpusavio padėtį. Eksperimento rezultatai patvirtina, jog pasiūlytas metodas turi dvigubo pagrįsto išsigynimo savybę.

Reikšminiai žodžiai: slaptas kanalas, steganorgrafija, duomenų slėpimas, FAT, failų sistema, klasteris.

I. Įvadas

Duomenų saugumas bei skaitmeninės informacijos saugus perdavimas yra labai aktuali tema. Dažniausiai slapta informacija yra užšifruojama panaudojant kriptografinius algoritmus, kurie yra pastoviai tobulinami ir papildomi naujais metodais [1, 2]. Visgi kai kuriais atvejais duomenų šifravimas panaudojant kriptografinius metodus neužtikrina pakankamos jų saugos. Tradiciškai užšifruotų duomenų egzistavimo faktas nėra slepiamas arba yra nesunkiai aptinkamas, todėl jų savininkas gali būti priverstas atskleisti šifravimo raktą [3, 4]. Vienas iš tokios situacijos pavyzdžių yra B. W. Lampson darbe [5] aprašyta „kalinių problema“. Norint išspręsti šią problemą, galima panaudoti duomenų slėpimo metodus, tokius kaip steganografija ar slaptas kanalas [3]. Steganografiniai metodai paslepia duomenis „įprastoje“ dengiančioje terpėje taip, kad nesukeltų galimo stebėtojo įtarimo [4], pvz., slaptą tekstinį pranešimą galima paslėpti skaitmeniniame vaizde. Slapti kanalai leidžia paslėpti duomenis terpėje, kuri nėra skirta duomenų perdavimui ar saugojimui. Tikslui pasiekti jie dengiančią informaciją perduoda tam tikru būdu išdėstytą laike, laikmenoje ar kitoje terpėje taip, kad gavėjas tik pagal jos išdėstymą gali išgauti slaptą pranešimą [6, 7]. Dažniausiai duomenų slėpimo metodai skirti naudoti informacijos perdavimo protokoluose [8, 9, 10], tačiau egzistuoja daugelis darbų [11, 12, 13, 14, 15, 16], kurie aprašo efektyvius duomenų slėpimo metodus kietuosiuose kompiuterių diskuose.

H. Khan et al. savo darbe [7] įveda du saugos metodų, skirtų duomenų slėpimui, efektyvumo įvertinimo lygius – tai pagrįstas išsigynimas ir dvigubas pagrįstas išsigynimas. Šios sąvokos apibūdina kaip efektyviai metodas gali paslėpti duomenis. Pagrįstas išsigynimas yra tokia saugos metodo savybė, kuri vienai iš šalių leidžia pagrįstai teigti kitoms šalims (pvz., teisėjui arba tyrėjui), kad paslėpti duomenys neegzistuoja [14]. Dažnai šia sąvoka apibūdinami ir metodai, leidžiantys šaliai atskleisti tik dalį paslėptų ir/ar užšifruotų duomenų ir sėkmingai paneigti bet kokių kitų duomenų egzistavimą. Šiuo atveju pats paslėptų duomenų egzistavimo faktas nėra efektyviai paslepiamas. Pavyzdžiui, didžioji dalis steganografinių metodų slepia duomenis nenaudojamose kietojo disko vietose [14, 15, 16, 17, 18]. Tyrėjas, atliekantis disko analizę, gali lengvai rasti nenaudojamus disko sektorius, užpildytus tam tikrais duomenimis, bet nežinodamas rakto, jis negali įrodyti, jog tai paslėpti svarbūs duomenys, o ne atsitiktiniai duomenys įrašyti anksčiau dėl intensyvaus disko naudojimo. Visi steganografiniai metodai turi tik pagrįsto išsigynimo savybę.

Dvigubo pagrįsto išsigynimo savybę turi tokie duomenų slėpimo metodai, kurie užtikrina, kad šalis gali du kartus pagrįstai paneigti paslėptų duomenų egzistavimą. Pirmuoju bandymu galima neigti, kad išvis egzistuoja bet kokie paslėpti duomenys. Nepavykus pirmajam bandymui, šie metodai palieka galimybę dar kartą paneigti duomenų (ar svarbesnės jų dalies) egzistavimą taip pat kaip viengubo pagrįsto išsigynimo atveju. Šią, dvigubo pagrįsto išsigynimo, savybę turi tik slapto kanalo metodai, kurie slėpdami informaciją į terpę neįrašo jokių papildomų duomenų. Du slapto kanalo metodai, tinkami naudoti kietuosiuose diskuose ir turintys dvigubo pagrįsto išsigynimo savybę, aprašyti [7].

Šiame straipsnyje aprašytas naujas, efektyvus ir lengvai įgyvendinamas metodas jautrios informacijos paslėpimui diskiniuose kaupikliuose, naudojančiuose klasterinę FAT [19] failų sistemą. Pasiūlytas metodas buvo išanalizuotas saugos požiūriu. Straipsnyje pateikti atliktų eksperimentinių tyrimų rezultatai. Eksperimentų rezultatų analizė leidžia daryti išvadą, jog siūlomas metodas pasižymi dvigubo pagrįsto išsigynimo savybe.

II. Egzistuojantys duomenų paslėpimo klasterinėse failų sistemose metodai

Keletas steganografinių metodų, slepiančių duomenis neužimtose failų sistemos struktūrose, aprašyti [11, 12, 17] darbuose. Anderson et al. [16] darbe pasiūlė steganografinę failų sistemą, naudojančią du skirtingus metodus informacijos paslėpimui, užpildant dengiančius failus atsitiktiniais duomenimis. Kiti autoriai [14, 15] dar labiau patobulino šiuos metodus. Visose metodo modifikacijose į įvairias failų sistemos duomenų struktūras rašoma informacija, kuri gali būti lengvai aptikta žemame lygyje analizuojant diską, todėl visi šie metodai pasižymi tik viengubo pagrįsto išsigynimo savybe.

H. Khan et al. pasiūlė [7] du slapto kanalo metodus, pasižyminčius dvigubo pagrįsto išsigynimo savybe ir skirtus paslėpti duomenis FAT failų sistemoje. Šie metodai į diską

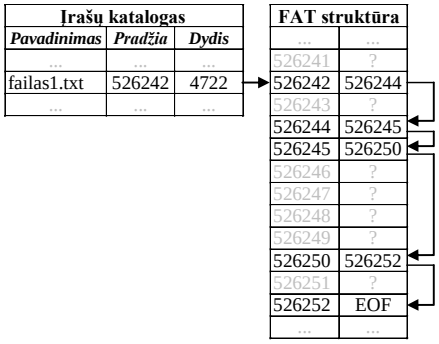
neįrašo jokių papildomų duomenų, o informacijos saugojimui naudoja FAT failų sistemos klasterių, priklausančių dengiančiajam failui, išdėstymą (tarpus tarp gretimų failo klasterių). Kadangi tam tikrai reikšmei paslėpti būtinas tam tikro dydžio tarpas, kyla kolizijų problema, kurios sprendimai yra gana sudėtingi. Be to, kolizijų kiekis stipriai išauga, kai failų sistema yra užpildyta duomenimis. Dėl to neįmanoma iki galo užpildyti diskų. Abejų metodų talpa (matuojama bitais, kuriuos galima paslėpti vienu dengiančiojo failo klasteriu) labai krenta, kai failų sistema yra artipilnė. Iš anksto įvertinti, kiek naudingų duomenų bus galima paslėpti duotajame diske, neįmanoma, nes tai taip pat priklauso net ir nuo pačių duomenų.

III. FAT FAILŲ SISTEMŲ STRUKTŪRA

FAT šeimos failų sistemos yra palaikomos daugumoje OS, dėl paprastumo ir suderinamumo jos yra naudojamos daugumoje šiuolaikinių prietaisų (televizorių, vaizdo grotuvų ir t. t.). FAT yra labai populiari nedidelės talpos nešiojamuose atmintinėse, dėl to kompanija Microsoft neseniai sukūrė naują FAT išplėtimą (exFAT), skirtą naudoti didesnės talpos diskuose.

FAT failų sistemas [19] sudaro tik 2 iš 5 kitose failų sistemose aptikamų duomenų struktūrų – įrašų katalogo struktūra bei failų paskirstymo lentelės (FAT). Grafiškai ryšys tarp šių struktūrų pavaizduotas 1 pav. Schemoje matome, kad kiekvienas failas turi įrašą katalogo struktūroje, kurioje, be failo vardo, dar saugomas jo dydis, pirmojo klasterio adresas bei kiti metaduomenys. Failų bei katalogų duomenys yra saugomi duomenų blokuose, vadinamuose klasteriais. Jei failas ar katalogas užima daugiau nei vieną klasterį, kiti klasteriai yra randami naudojant struktūrą, vadinamą FAT. Ji yra naudojama kito, failui priklausančio, klasterio radimui bei klasterių vietų užimtumo nustatymui.

Klasterių paskirstymo failams algoritmas priklauso nuo operacinės sistemos, tačiau dauguma sistemų naudoja „pirmas laisvas“ algoritmą, kai imamas pirmas pasitaikęs laisvas klasteris po jau užimtojo. Taip daroma todėl, kad diskiniai kaupikliai daug sparčiau dirba su klasteriais, kai jie yra šalia vienas kito. Tačiau dėl aktyvaus darbo su failų sistema dažnai nepavyksta rezervuoti visų klasterių iš eilės – susidaro failų sistemos fragmentacija. Tokiu atveju visus failo klasterius rasti pavyksta naudojant FAT struktūrą, panašiai kaip dinaminį sąrašą (1 pav.).



1 pav. Ryšys tarp įrašų katalogo ir FAT struktūros

IV. SIŪLOMAS METODAS

Pagrindinė siūlomo duomenų slėpimo FAT failų sistemoje metodo idėja - panaudoti daugiau nei vieną dengiantį failą ir slaptą žinutę perduoti panaudojant dengiančių failų klasterių tarpusavio išsidėstymą. Tokiu atveju į kitus failų sistemos duomenis galima nekreipti dėmesio ir analizuoti tik dengiančių failų klasterius. Galima sudaryti didėjimo tvarka surikiuotą masyvą su dviejų dengiančių failų klasterių numeriais ir stebėti, kaip vieno failo klasteriai „pinasi“ su klasteriais, priklausančiais kitam failui. Jei kažkurioje masyvo vietoje po pirmojo failo klasterio eina kito failo klasteris, tai atitinkamoje paslėptos žinutės vietoje yra bitas „1“. Jei kažkurioje masyvo vietoje yra du iš eilės to paties failo klasteriai, tai reiškia, jog paslėptoje žinutėje yra „0“. Tokiu būdu galima paslėpti visą slaptą žinutę neįrašant nei vieno bito į diską.

Norint padidinti slapto kanalo talpą, galima panaudoti daugiau dengiančiųjų failų. Tada kiekvienam failui reikia skirti numerius ir kiekvienam dengiančių failų klasterių masyvo elementui apskaičiuoti skirtumą tarp einamojo ir prieš tai buvusio failo numerių. Šis skirtumas ir bus atitinkamų paslėptos žinutės bitų reikšmė. Naudojant tokią taktiką galima paslėpti daugiau nei vieną slaptą bitą viename dengiančio failo klasteryje. Tokiu atveju dengiančių failų vardai bei jų pateikimo eilės tvarka tampa slaptojo rakto dalimi, nes juos būtina žinoti, norint nuskaityti paslėptus duomenis.

V. DUOMENŲ PASLĖPIMO ALGORITMAS

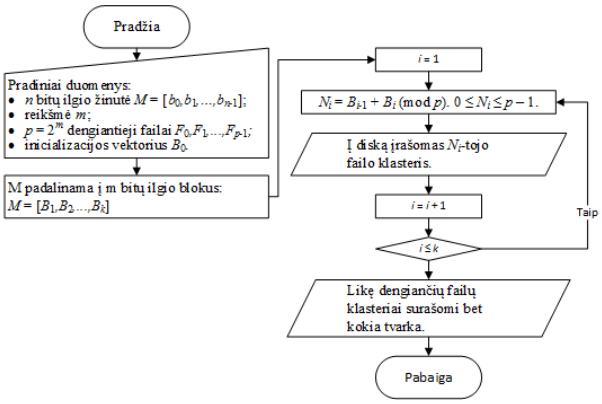
Tarkime, norime paslėpti n bitų ilgio žinutę $M=[b_0,b_1,\ldots,b_{n-1}]$. Pasirenkame $p=2^m$ dengiančiųjų failų $F_0,F_1,\ldots,F_{p-1}$. Kiekvienam iš šių failų priskiriame sveiką skaičių (failo eilės numerį). Įrašydami vieną reikiamo failo klasterį į diską galėsime paslėpti m žinutės bitų, todėl žinutę reikia suskaidyti į m bitų ilgio blokus $M=[B_1,B_2,\ldots,B_k]$ ir juos apdoroti atskirai, pradedant nuo pirmojo. Jei žinutės ilgis nesutampa su sveiku blokų skaičiumi, tai žinutę reikia papildyti nuliniais bitais iki užpildyto bloko. Kiekvienas bitų blokas B_i, $i=0,1,\ldots,k$ interpretuojamas kaip natūralusis skaičius. Prieš pradedant slėpimą, reikia laisvai pasirinkti nulinio bloko reikšmę B_0. Jei slepiamos kelios žinutės iš eilės, tai šia reikšme galima pasirinkti paskutinio prieš tai apdorotos žinutės bloko reikšmę.

Iš eilės peržiūrimi visi pradinės žinutės blokai B_i $i=1,2,\ldots,k$, ir dengiančiojo failo eilės numeris N_i apskaičiuojamas pagal formulę:

$$N_i := B_{i-1} + B_i \pmod p,\ 0 \le N_i \le p-1. \qquad (1)$$

Pradedant nuo pirmojo laisvo disko klasterio, į diską įrašomas eilinis dar neįrašytas N_i-tojo failo klasteris. Procedūra kartojama tol, kol paslepiama visa žinutė.

Duomenų paslėpimo algoritmo blokinė diagrama pateikta 2 pav.



2 pav. Duomenų paslėpimo algoritmo blokinė diagrama

Norint atkurti pradinę žinutę, paslėptą diske, būtina žinoti tokią informaciją:

- reikšmę m;
- slaptos žinutės ilgį bitais n;
- inicializacijos vektorių B_0;
- visų dengiančiųjų failų $F_0, F_1, \ldots, F_{p-1}$, vardus ir jų pateikimo algoritmui tvarką.

Visa ši atkūrimui būtina informacija sudaro siūlomo slapto kanalo metodo slaptąjį raktą, be kurio neįmanoma atkurti pradinės žinutės net ir turint diską, kuriame ji įrašyta.

Duomenų nuskaitymo algoritmas yra analogiškas. Tarkime, kad žinoma visa informacija, sudaranti algoritmo slaptąjį raktą. Tuščias bitų masyvas M suskaidomas į k blokų po m bitų, $M = [B_1, B_2, \ldots, B_k]$. Kiekvienas blokas B_i interpretuojamas kaip natūralusis skaičius.

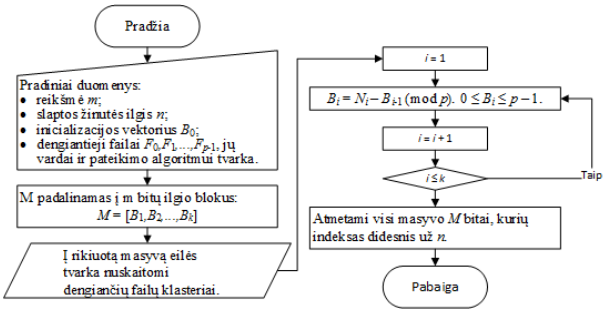
Disko, kuriame paslėpta žinutė, failų sistema analizuojama pradedant nuo mažiausio klasterio numerio. Tiriami tik tie klasteriai, kurie priklauso vienam (bet kuriam) iš dengiančiųjų failų. Iš tokių klasterių numerių galima sudaryti surikiuotą masyvą.

Kiekvienam atkuriamos žinutės blokui B_i, $i = 1,2\ldots,k$ pagal surikiuoto klasterių masyvo i-tąjį elementą, panaudojant FAT failų sistemos struktūras, randamas atitinkamo dengiančiojo failo numeris N_i. Bitų bloko B_i reikšmė apskaičiuojama pagal formulę:

$$B_i := N_i - B_{i-1} \pmod p,\ 0 \le B_i \le p-1. \qquad (2)$$

Procedūra kartojama tol, kol apdorojami visi blokai. Rezultato masyvo M bitai, kurių indeksas didesnis už n, atmetami.

Duomenų nuskaitymo blokinė diagrama pateikta 3 pav.



3 pav. Duomenų nuskaitymo algoritmo blokinė diagrama

Reikėtų atkreipti dėmesį į dengiančiųjų failų ilgiams taikomus apribojimus. Visų pirma, maksimalią visų dengiančių failų dydžio sumos reikšmę apibrėžia likusi laisva failų sistemoje vieta. Antra, galimos dvi strategijos, naudojamos rašyti dengiančius failus – pabaigus rašyti slaptam pranešimui paslėpti reikalingus klasterius, likusius dengiančiųjų failų klasterius tiesiog „išmesti“, t. y. neberašyti į diską. Antra strategija – įrašyti visus likusius dengiančiųjų failų klasterius.

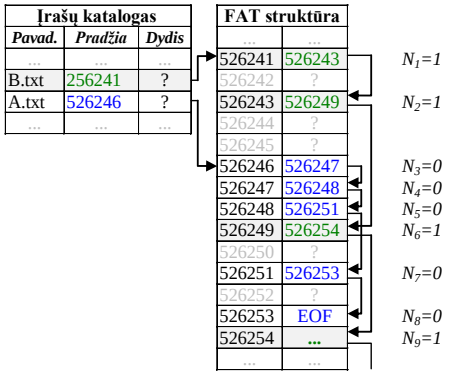
Dengiančiųjų failų pakankamo ilgio problema kyla todėl, kad siūlomas metodas, priklausomai nuo slepiamų duomenų, gali nevienodai „panaudoti“ skirtingų failų klasterius. Egzistuoja išskirtinės situacijos, kai visas pranešimas gali būti paslepiamas naudojant tik vieno failo klasterius, nors dengiančiųjų failų reikia daug daugiau. Dėl tokio galimo failų panaudojimo asimetriškumo, jei nenorime naudoti pirmosios strategijos, būtina užtikrinti, kad kiekvienas failas būtų pakankamo ilgio paslėpti visą pranešimą arba prieš jį slepiant apskaičiuoti kiekvieno failo reikiamą dydį atliekant žinutės paslėpimo simuliaciją, t.y. suskaičiuoti, kiek kiekvieno failo klasterių reikės, nerašant jų į failų sistemą.

VI. Pavyzdys

Tarkime, norime paslėpti dvejetainę žinutę $M = [1,0,0,0,0,1,1,1,0]$. Parenkame duomenis slaptam raktui: $B_0 = 0$, $m = 1$, $p = 2^m = 2$. Tai reiškia, kad mums reikės dviejų dengiančių failų – imame „A.txt“ ir „B.txt“, jų indeksai masyve bus atitinkamai 0 ir 1. Kadangi $m = 1$, t. y. slėpsime po vieną bitą vienu dengiančių failų klasteriu, tai žinutės M blokas B_i yra lygus vienam žinutės bitui.

Pagal (1) formulę paskaičiuojame, kurio failo klasterį rašysime pirmą: $N_1 := B_0 + B_1 \pmod p = 0 + 1 \pmod 2 = 1$. Į pirmą laisvą disko klasterį įrašome failo su pirmu indeksu („B.txt“) pirmąjį klasterį ir algoritmą vykdome toliau. Analogiškai $N_2 = 1 + 0 \pmod 2 = 1$, todėl į antrąjį nuo disko pradžios laisvą klasterį rašysime pirmojo failo antrąjį klasterį. Nesunku apskaičiuoti, kad $N_3 = 0$, $N_4 = 0$ ir t. t. Paslėpus visus bitų blokus, pabaigiame rašyti dengiančius failus bet kokia tvarka (arba likusių jų dalių visai neberašome į diską).

4 pav. pateiktoje schemoje matoma kaip atrodo FAT struktūra, paslėpus duotąją žinutę.



4 pav. FAT struktūra po duomenų paslėpimo operacijos

Norėdami nuskaityti slaptą pranešimą, turime žinoti paslėpimui naudotus parametrus. Kadangi žinome, kurių failų klasteriai slepia duomenis, išrenkame juos į atskirą masyvą ir žiūrime, kuriam failui priklauso pirmasis masyvo elementas. Tuomet pagal (2) formulę paskaičiuojame pirmojo bitų bloko reikšmę $B_1 = N_1 - B_0 \pmod p = 1 - 0 \pmod 2 = 1$. Kitų blokų reikšmės paskaičiuojamos analogiškai: $B_2 = N_2 - B_1 \pmod p =$ $= 1 - 1 \pmod 2 = 0$, $B_3 = 0 - 0 \pmod 2 = 0$, $B_4 = 0$, $B_5 = 0$, $B_6 = 1$, $B_7 = 1$, $B_8 = 1$, $B_9 = 0$. Bitų blokus surašius į masyvą, gauname nuskaitytą pranešimą $M = [1,0,0,0,0,1,1,1,0]$

VII. Metodo talpos įvertinimas

Siūlomas metodas pasižymi pastovia $C = \log_2 p$ talpa. Čia C yra slapto kanalo talpa matuojant bitais, paslepiamais vienu dengiančio failo klasteriu, p – dengiančių failų skaičius. Ši išraiška teisinga tik tuo atveju, kai nebūtina į diską pilnai įrašyti kiekvieno dengiančiojo failo. Vaizdesnė algoritmo talpos priklausomybė nuo naudojamų dengiančių failų skaičiaus pavaizduota 1 lentelėje.

1 lentelė. Talpos priklausomybė nuo dengiančių failų skaičiaus

Naudojamų dengiančių failų skaičius p	Slaptų bitų skaičius viename dengiančio failo klasteryje C
2	1
4	2
8	3
16	4
32	5
64	6
128	7
256	8

VIII. Kokybinių charakteristikų įvertinimas

Metodo kokybinių charakteristikų analizės metu įvertintas paslėptų duomenų atsparumas pokyčiams failų sistemoje ir dengiančiuose failuose. Taip pat, remiantis įvairiais kriterijais, įvertintas metodo saugumas.

A. *Paslėptų duomenų stabilumas*

Steganografiniai metodai, kurie remiasi slaptos informacijos saugojimu nenaudojamose failų sistemos vietose, yra neatsparūs paslėptų duomenų sugadinimui, kai su failų sistema dirba apie slaptus duomenis nieko nežinanti OS ar kitos programos. Pavyzdžiui, naujai sukūrus failą, jo duomenys gali perrašyti prieš tai „laisvose“ disko vietose paslėptą informaciją.

Mūsų siūlomas metodas nesaugo jokios papildomos informacijos failų sistemoje, todėl yra atsparus įprastam darbui su failų sistema (failų kūrimui, ne dengiančių failų keitimui, ištrynimui). Visgi kai kurie žemo lygio darbui su disku skirti įrankiai, tokie kaip disko defragmentavimo programos, keičia jau esančių failų klasterių išsidėstymą ir gali sunaikinti paslėptą informaciją. Dengiančių failų papildymas ar kitoks keitimas gali sukelti klasterių išdėstymo pakeitimą, o tuo pačiu ir slaptų duomenų vientisumo pažeidimą. Galiausiai vieno iš dengiančių failų pervardinimas, perkėlimas, nukopijavimas ar ištrynimas garantuos paslėptų duomenų sunaikinimą.

B. *Saugumo įvertinimas*

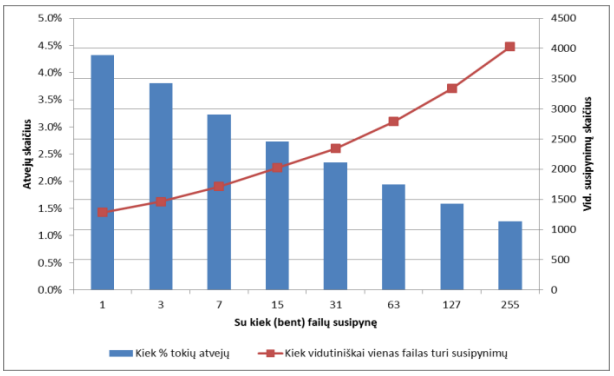
Siūlomas slapto kanalo metodas yra atsparus žemo lygio disko analizei. Diską analizuojantis tyrėjas neras jokių papildomų duomenų „nestandartinėse“ disko vietose, visos failų sistemos struktūros bus korektiškos ir visiškai atitiks FAT specifikaciją. Vienintelis šalutinis metodo naudojimo efektas yra padidėjusi dengiančių failų fragmentacija. Tačiau fragmentacija aptinkama ir natūraliomis sąlygomis, kai failų sistema yra naudojama ilgą laiką ir joje vykdomas aktyvus darbas su failais. Taigi, norint įvertinti pasiūlyto metodo saugumą, reikėtų atsakyti į klausimą: su kokiais parametrais turėtų būti naudojamas pasiūlytasis metodas, kad sukeliama fragmentacija būtų tokia pati, kaip ir kylanti įprasto failų sistemose naudojimo atveju?

IX. Metodo saugumo eksperimentinis tyrimas

Eksperimentinio tyrimo metu buvo išanalizuoti universiteto laboratorijose naudojami kompiuterių diskai norint nustatyti, kaip failai fragmentuojasi kasdienio naudojimo sąlygomis. Pradinių duomenų apie failų sistemą nuskaitymui naudotas „The Sleuth Kit“ [20, 21]. Vėliau šie duomenys buvo įkelti į duomenų bazę, kurioje SQL užklausų pagalba buvo siekiama nustatyti, kiek tiriamose failų sistemose buvo fragmentuotų ir susipynusių failų ir kokiais atvejais jie susipina. Norint suprasti eksperimentų rezultatus, reikia formaliai apibrėžti susipynusių failų sąvoką: du failai F_1 ir F_2 laikomi tarpusavyje susipynusiais, jeigu egzistuoja bent viena situacija, kai tarp failo F_1 dviejų gretimų klasterių yra bent vienas failo F_2 klasteris ir atvirkščiai (tarp dviejų gretimų failo F_2 klasterių yra bent vienas failo F_1 klasteris). Jei ši situacija diske aptinkama keletą kartų, sakoma, kad failai susipynę atitinkamą skaičių kartų. Suprantama, failas gali būti susipynęs ne tik su vienu, bet su keliais kitais failais. Gretimais failo klasteriais laikomi iš eilės einantys to failo klasteriai (pagal failo turinį) ir turintys didėjančius klasterių numerius (dengiantiems failams ši sąlyga bus tenkinama visiems klasteriams, kitiems failams – ne visada).

Atlikta diskų analizė parodė, kad 4,33% visų failų yra susipynę bent su vienu kitu failu (šie failai turėjo vidutiniškai po 1284 susipynimų), 3,80% su bent trim kitais failais (vid. po 1461 susipynimų), 3,23% su bent septyniais kitais failais (vid.

po 1715 susipynimų), o 1,26% analizuotų failų buvo susipynę su 255 ir daugiau kitų failų, ir kiekvienas failas turėjo vid. po 4030 susipynimų (5 pav.). Šie rezultatai parodo, kad panaudojant didesnį dengiančių failų skaičių ne tik padidiname metodo talpą (1 lentelė), bet taip pat gauname didesnį susipynimų skaičių, kas leidžia prisidengiant „natūraliomis sąlygomis“ paslėpti daugiau duomenų.



5 pav. Susipynimo atvejų ir susipynimų priklausomybė nuo failų skaičiaus

Kitas eksperimentinių tyrimų tikslas buvo nustatyti, kokiose failų sistemose, operacinėse sistemose bei kokio tipo ir dydžio failai yra labiausiai linkę susipinti.

Atlikta analizė atskleidė, kad, lyginant su FAT32 tipo failų sistemomis, susipynusių failų daugiau turi NTFS tipo FS (2 lentelė). Taip pat 3 lentelėje matome, kad daugiausia susipynusių failų bei susipynimo atvejų aptikta failų sistemose, naudotose Windows XP bei Windows Vista operacinėse sistemose.

2 lentelė. FAILŲ SUSIPYNIMO PRIKLAUSOMYBĖ NUO FS TIPO

FS tipas	Vid. failo dydis, klast.	Su kiek vid. kitų failų susipynęs vienas failas	Vid. susipynimo atvejų skaičius
NTFS	31,45	31,68	115,20
FAT32	42,38	1,53	6,41

3 lentelė. FAILŲ SUSIPYNIMO PRIKLAUSOMYBĖ NUO OS

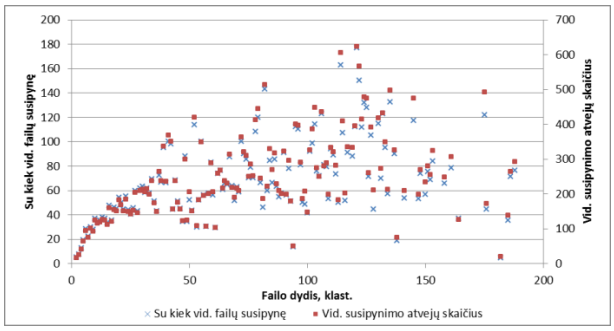
OS	Vid. failo dydis, klast.	Su kiek vid. kitų failų susipynęs vienas failas	Vid. susipynimo atvejų skaičius
Windows XP	39,25	19,79	70,21
Windows Vista	26,20	15,66	61,61
Windows 98	35,12	2,01	9,05

4 lentelėje pateikti 10 daugiausiai susipynimo atvejų turinčių failų tipų. Matome, kad „natūraliomis sąlygomis“ labiausiai susipina archyvo bei įrašų (duomenų bazės, žurnalo) tipo failai. Taigi norint saugiai paslėpti duomenis geriausia naudoti tokių tipų dengiančiuosius failus.

4 lentelė. FAILŲ SUSIPYNIMO PRIKLAUSOMYBĖ NUO FAILO TIPO

Failo tipas	Vid. failo dydis, klast.	Su kiek vid. kitų failų susipynęs vienas failas	Vid. susipynimo atvejų skaičius
.msi	892,87	326,31	1349,40
.zip	2152,61	126,90	589,36
.jar	204,50	133,79	500,25
.msp	2319,98	114,41	458,41
.log	40,51	102,39	458,12
.msf	39,50	106,04	396,75
.cab	884,87	80,78	367,79
.otp	46,48	100,27	326,33
.db	67,38	75,85	274,27
.frx	7,37	61,34	202,82

6 pav. pateiktame grafike matome kaip keičiasi vidutiniai supynusių failų bei susipynimo atvejų skaičiai didėjant failą sudarančių klasterių skaičiui.



6 pav. Susipynusių failų bei susipynimo atvejų skaičiaus priklausomybė nuo failo dydžio

Kita eksperimentinių tyrimų dalis buvo skirta nustatyti, ar įprastai dirbant su failų sistemomis (naudojant tik standartines OS funkcijas ir tiesiogiai nekeičiant žemo lygio disko ir failų sistemų struktūrų) susidaro situacijos, kai du ar daugiau failų tarpusavyje susipina daugelį kartų. Tam tikslui buvo sumodeliuotos trys situacijos Windows 7, Windows XP ir Ubuntu 11.04 operacinėse sistemose bei FAT32 ir NTFS failų sistemose:

1. Du failai buvo vienu metu kopijuojami į tą pačią failų sistemą OS priemonėmis. Šiam bandymui įgyvendinti buvo parašyti scenarijai, kurie vienu metu paleisdavo OS suteikiamas programas (Ubuntu atveju „cp“, Windows – „copy“) dviejų 100 MB dydžio failų kopijavimui.
2. Du failai vienu metu siunčiami į tą pačią failų sistemą, naudojant BitTorrent protokolo programą. Buvo surasti du nuorodas į apie 100 MB failus turintys „.torrent“ failai ir, panaudojant atitinkamas programas, siunčiami į tą pačią vietą.
3. Duomenys buvo lygiagrečiai rašomi į du failus, esančius toje pačioje failų sistemoje, papildymo būdu. Šio eksperimento įgyvendinimui sukurta programa Java kalba. Programa naudojo atskiras gijas kiekvieno failo kopijavimui, o kiekvienos gijos veikimas atitiko šį scenarijų: atidaromas failas papildymui, įrašoma 1 KB duomenų, failas uždaromas. Scenarijus vykdomas, kol failo dydis pasiekia 100 MB.

Kiekvienu atveju po atlikto duomenų įrašymo failų sistema buvo analizuojama saugumo įvertinimo skyriuje (žr. VIII skyriaus VIII.B poskyrį) aprašyta metodika.

Pirmos dvi situacijos nesukėlė jokios išskirtinės fragmentacijos ar failų susipynimo. Iš 5 lentelėje pateiktų rezultatų matoma, jog beveik visais atvejais rašant į failus vienu metu papildymo būdu sukeliama labai stipri fragmentacija ir failų tarpusavio susipynimas. Lentelė taip pat parodo, kad failų klasterių išsidėstymas stipriai priklauso nuo operacinės sistemos bei jos tvarkyklių. Labiau optimizuotos sistemos sugeba išvengti stiprios fragmentacijos (Ubuntu 11.04 OS, NTFS failų sistemos atvejis).

5 lentelė. RAŠYMO PAPILDYMO BŪDU SUKELTA FRAGMENTACIJA

	Windows 7		**Windows XP**		**Ubuntu 11.04**	
	FAT32	***NTFS***	***FAT32***	***NTFS***	***FAT32***	***NTFS***
Vid. fragmentų skaičius faile	25580	19869	16458	24675	24774	7
Vid. vieno failo susipynimų skaičius	51157	29684	14812	49341	49546	12

Reikia pabrėžti, kad rašant į failus buvo naudotos tik OS suteikiamos aukšto lygio funkcijos. Sukurta rašymo programa prie failų sistemos struktūrų priėjimo neturėjo. Tai parodo, kad įprastos programos gali sukelti labai stiprų failų tarpusavio susipynimą, jei rašoma į juos papildymo būdu. Dažnai tokiu būdu rašoma į žurnalo tipo failus ir tą patvirtina 4 lentelėje matomas 5 vietoje pagal susipynimo atvejus esantis žurnalo failo tipas.

X. IŠVADOS

Straipsnyje buvo pristatytas naujas metodas saugiam duomenų paslėpimui klasterinėje FAT failų sistemoje. Siūlomas metodas naudoja keletą dengiančių failų ir yra pagrįstas tų failų klasterių tarpusavio išsidėstymu failų sistemoje. Metodas yra paprastai įgyvendinamas, nes darbo metu nesukelia kolizijų problemos [7]. Siūlomo metodo talpumas nepriklauso nuo failų sistemos dydžio ar likusios laisvos vietos joje. Metodas užtikrina pastovų paslėptų bitų, tenkančių vienam dengiančiųjų failų klasteriui, skaičių ir leidžia visiškai užpildyti failų sistemą. Dengiantys failai, jų skaičius bei jų eilės tvarka sudaro slaptąjį raktą, kuris yra reikalingas paslėptiems duomenims nuskaityti. Dengiančių failų skaičiaus parinkimas leidžia keisti metodo talpos charakteristiką (kiek slaptos informacijos paslepiama vienu dengiančio failo klasteriu), kuri yra panaši, o kai kuriais atvejais netgi geresnė negu kitų panašių metodų. Siūlomo metodo naudojimas sukelia failų fragmentaciją bei ypatingą jos tipą – failų susipynimą. Eksperimentuose nustatyta, kad toks klasterių susipynimas yra dažnai sutinkamas senose aktyviai naudotose failų sistemose, be to, labai didelis kelių failų tarpusavio susipynimas gali būti sukeltas programų, kurios dažnai rašo į failus papildymo būdu. Eksperimentiniai tyrimai patvirtina, jog pasiūlytas metodas pasižymi dvigubo pagrįsto išsigynimo savybe.

LITERATŪRA

[1] Toldinas J., Stuikys V., Damasevicius R., Ziberkas G., Banionis M. Energy Efficiency Comparison with Cipher Strength of AES and Rijndael Cryptographic Algorithms in Mobile Devices. // Electronics and Electrical Engineering, 2011, Vol. 108, No. 2, P. 11–14.

[2] Luksys K., Sakalauskas E., Venckauskas A. Implementation Analysis of Matrix Power Cipher in Embedded Systems. // Electronics and Electrical Engineering, 2012, Vol. 118, No. 2, P. 95–98.

[3] Petitcolas F. A. P., Anderson R. J., Kuhn M. G. Information hiding-a survey. // Proceedings of the IEEE: Special Issue on Protection of Multimedia Content, 1999, Vol. 87, No. 7, P. 1062–1078.

[4] Provos N., Honeyman P. Hide and Seek: An Introduction to Steganography. // IEEE Security and Privacy, 2003 May, Vol. 1, No. 3, P. 32–44.

[5] Lampson B. W. A note on the confinement problem. // Communications of the ACM, 1973 Oct, Vol. 16, No. 10, P. 613–615.

[6] Simmons G. J. The Prisoners' Problem and the Subliminal Channel. // Advances in Cryptology, Proceedings of Crypto 83, New York: Plenum Press, 1984, P. 51–67.

[7] Khan H., Javed M., Khayam S. A., Mirza F. Designing a cluster-based covert channel to evade disk investigation and forensics. // Computers & Security, 2011, Vol. 30, No. 1, P. 35–49.

[8] Hanaoka G., Hanaoka Y., Hagiwara M., Watanabe H., Imai H. Unconditionally secure chaffing-and-winnowing: a relationship between encryption and authentication. // Proceedings of the 16th international conference on Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, Berlin, Heidelberg: Springer-Verlag, 2006, P. 154–162.

[9] Dakhane D. M., Patil S., Patil M. Detection and elimination of covert communication in Transport and Internet layer - A Survey. // IJCA Proceedings on International Conference on Recent Trends in Information Technology and Computer Science (ICRTITCS-2011), New York, USA: Foundation of Computer Science, 2012, P. 36–41.

[10] Bellare M., Boldyreva A. The Security of Chaffing and Winnowing. // Proceedings of the 6th International Conference on the Theory and Application of Cryptology and Information Security: Advances in Cryptology, London, UK, UK: Springer-Verlag, 2000, P. 517–530.

[11] Piper S., Davis M., Manes G., Shenoi S. Detecting Hidden Data in Ext2/Ext3 File Systems. // IFIP Int. Conf. Digital Forensics, Springer, 2005, P. 245–256.

[12] Eckstein K., Jahnke M. Data Hiding in Journaling File Systems. // Refereed Proceedings of the 5th Annual Digital Forensic Research Workshop, New Orleans, Louisiana, USA, 2005, P. 1–8.

[13] Anderson R. J. Information Hiding Terminology - Results of an Informal Plenary Meeting and Additional Proposals. // Information Hiding, Springer, 1996, P. 347–350.

[14] McDonald A. D., Kuhn M. G. StegFS: A Steganographic File System for Linux. // Information Hiding, Springer, 1999, P. 462–477.

[15] Pang H., Tan K. L., Zhou X. StegFS: A steganographic filesystem. // Proceedings of the 19th International Conference on Data Engineering, IEEE, 2003, P. 657–668.

[16] Anderson R. J., Needham R. M., Shamir A. The Steganographic File System. // Proceedings of the Second International Workshop on Information Hiding, London, UK: Springer-Verlag, 1998, P. 73–82.

[17] Huebner E., Bem D., Wee C. K. Data hiding in the NTFS file system. // Digital Investigation, 2006 Dec, Vol. 3, No. 4, P. 211–226.

[18] Czeskis A., Hilaire D. J. S., Koscher K., Gribble S. D., Kohno T., Schneier B. Defeating Encrypted and Deniable File Systems: TrueCrypt v5.1a and the Case of the Tattling OS and Applications. // HotSec, USENIX Association, 2008, P. 1–7.

[19] Carrier B. File System Forensic Analysis. Addison-Wesley Professional, 2005, 382 p.

[20] Garfinkel S. L. Automating Disk Forensic Processing with SleuthKit, XML and Python. SADFE, IEEE Computer Society, 2009, P. 73–84.

[21] Forte D. V. Open Source Forensics: The PTK: An alternative advanced interface for Sleuth Kit. // Network Security, 2008 Apr, Vol. 2008, No. 4, P. 10–13.

INFORMACINĖS TECHNOLOGIJOS ● IVUS 2013 ● ISSN 2029-249X ● eISSN 2029-4824

Development of Robot Programming Ontology

Ignas Plauska
Centre of Real Time Computer Systems
Kaunas University of Technology
Studentu 50, Kaunas, Lithuania
ignas.plauska@ktu.lt

***Abstract* — The robotics domain is inherently complex and requires knowledge of artificial intelligence, real-time systems, embedded software, mechatronics, kinematics and dynamics, communication protocols, behavior architectures and control algorithms. The complexity of the domain can be addressed by using domain ontology, i.e., a formal model that describes domain concepts and their relations. In this paper we describe the development of ontology for robot programming domain, describe its structure (sub-ontologies) and main concepts.**

Keywords — ontology engineering, robotics, programming.

I. INTRODUCTION

Robotics is a complex domain that includes both hardware and software parts and requires deep knowledge of artificial intelligence, embedded software, real-time systems, communication and control protocols, kinematics, navigation, sensors, mechanics and robot programming languages. Complexity of the robotics domain underscores the need for representation and management of semantic knowledge. The advantage of having a formalized description of semantic knowledge about some domain is facilitating reuse of domain artefacts (best practices, designs, algorithms, etc.) and easing knowledge sharing among domain actors (robot constructors, programmers, do-it-yourself users). Many robotic systems exploit semantic knowledge, but it is often hard-coded in their implementation and can be learned only implicitly.

Ontology is a common tool for expressing and formalizing semantic knowledge about domain concepts and their relationships, which seeks to provide a definitive and exhaustive classification of entities in a domain [1]. Ontology aims to define an unambiguous vocabulary of terms shared across domain experts and taxonomy of concepts to model a domain of interest. It covers classes and sets (groups or types of concepts), relations (how concepts are related to each other), individuals (instances of concepts), and attributes (data attached to individual concepts). OWL is a standard language for describing ontologies [2]. The advantage of having a machine-readable and interpretable ontology specification in OWL is that it can be validated using ontology validation tools and additional knowledge can extracted and reused using logical inference tools.

The context of this paper is the development of a visual programming language for the robotics domain. Our idea is to develop a language that is 1) is grounded semantically (a formal ontology of robot programming domain), and 2) uses a familiar and easily recognizable set of graphical symbols reused from a road sign collection of symbols, and 3) allows to develop programs at a higher abstraction level using programming idioms and robot missions rather than simple commands as in well-known visual programming languages for robotics such as Microsoft's Visual Programming Language [3] and Lego NXT-G [4].

For this language, an extensive, general-purpose robot programming ontology is necessary. Existing ontologies analysed in this paper only focus on some sub-domains of robotics and therefore are not sufficient for a visual programming language for the robotics domain. The ontology proposed in this paper aims to cover the general domain of robot programming.

II. RELATED ONTOLOGIES

A robot is an integrated set of mechanical parts, sensors, actuation servomotors and hardware, sensor signal processing software, robot behaviour control software, and servomotor control software with "appropriate" interfaces [5]. Kuipers [6] identified the following levels of the robotics domain: topological level (the categories of places, paths and regions, with their associated connectivity and containment relations), temporal level (time, period), causal level (views, actions, events and the causal relations among them), control level, sensory level, and metrical level (units of measurement), where each level has its own sub-ontology of concepts and problem-solving methods. Further, Paull *et al.* [7] listed sub-domains of robotics: power (power supply, batteries), actuators (servo motors), sensors (infrared, light, ultrasound, GPS, temperature, touch, etc.), control algorithms (adaptive, neural network, PID, hybrid, etc.), dynamics (principles of motion), kinematics (forward/inverse kinematics laws, position, orientation, direction, types of manoeuvres), locomotion (driving and steering systems, legged locomotion), localization (mapping), planning (decision theory), communications (media, protocol).

First, the robotics ontologies *per se* are of interest when searching for related ontologies for analysis and possibly reuse of concepts described there. Second, the ontologies of related domains which share the same layers of semantic knowledge such as automation, heterogeneous environments, Internet of Things, navigation, programming, Sensor Networks, and topology. Finally, the so-called upper level ontologies which describe general concepts that are the same across many knowledge domains such as spatial dimensions, geographical locations or temporal relations should be considered.

TABLE I. COMPARISON OF ONTOLOGIES

Ontology	Concepts of ontology	Description
A3ME [19]	Devices, sensor, actuator, computing, energy, communication, storage	Ontology for self-description and discovery of devices and their capabilities in heterogeneous mobile networks including resource constrained sensor nodes
COPS [20]	Artefacts of computation, Actions, Programming languages	General ontology of programs including sub-ontology of programming languages
CSIRO [21]	Sensor, Operation Model and Process	Ontology for describing and reasoning about sensors, observations and scientific models
Global Ontology for IoT [22]	Devices, Estimation models, Physics of things, Mathematics	Multi-faceted ontology that describes device metadata, and mathematical models of IoT
IoT Ontology [23]	Devices and their functionalities	Representation of interconnected and semantically coordinated smart entities in the IoT
Ontonym [24]	Time, location, people, events, sensing, provenance, device, resource	Set of lower-level ontologies that represent core concepts in pervasive computing
OntoSensor [25]	Sensor Capabilities, Measures	Knowledge base of technical specifications of sensors
OntoSTIT+ [26]	Action	Ontology for describing different types of action
PROTEUS [9]	Robots, Missions, Environments	Ontology aiming to model the mobile robot mission scenarios
Robot description ontology [16]	Environments and their conditions, Disaster situations, Robot features	Ontology for describing rescue robot features and for describing disaster scenarios in the context of search and rescue effort.
Robotics Science Ontology [10]	Agent, model. context	Semantics of meta-level concepts and domains of robotics using the robot agent and robot system models and the Context Space
SDO [27]	Sensors and their parameters	Ontology that associates sensor information taxonomy for searching and interpreting raw data streams
SensorML [28]	Measurement devices, Measurement systems	Provides a powerful and modular way of describing data processing chains by allowing the definition of reusable processing components
SORA [8]	Devices, control strategies, motion planning, perception, knowledge representation	Ontology and associated methodology for knowledge representation and reasoning in robotics and automation, together with concepts in an initial set of application domains. Represents all parts of a robot from low level devices, such as sensors and actuators; to the knowledge required to solve a problem
SSN [29]	Sensors, Stimuli	Describes sensors, their accuracy and capabilities, observations and methods used for sensing, sensor networks
Spatial Ontology [9]	Spatial Relations, topology, locations regions	DOLCE extension particularly for space and navigation
SWAMO [8]	Sensor systems, Autonomous agent control systems	Describes autonomous agents for system-wide resource sharing, decision making, autonomic operations
WISNO [27]	Wireless sensors, data types and processing	Ontology for describing and flexible modelling of sensor data.

The robotics ontologies include IEEE Standard Ontology for Robotics and Automation (SORA) [8], PROTEUS ontology devoted to the description of mobile robotic scenarios [9], Robotics Science Ontology [10], Sensor Web for Autonomous Mission Operations (SWAMO) ontology [11]. Other robotics-related ontologies include Spatial Ontology for navigation [12], knowledge representation and action generation [13], [14], route instruction [15], UGVs [16], data representation [17], and robot behaviour [18]. The ontologies of related domains include Agent-based Middleware Approach for Mixed Mode Environments (A3ME) [19], Core Ontology of Programs and Software (COPS) [20], CSIRO Sensor Ontology [21], Global Ontology for the Internet of Things [22], Internet of Things (IoT) Ontology [23], Ontonym ontology for pervasive systems [24], OntoSensor ontology for sensor networks [25], OntoSTIT+ action ontology [26], Sensor Data Ontology (SDO) [27], SensorML Processes [28], W3C Semantic Sensor Net (SSN) [29], and Wireless Sensor Networks Ontology (WISNO) [30]. In Table I, we present the comparison of ontologies.

III. DEVELOPMENT OF THE ONTOLOGY

Rigorous construction of ontology requires the use of ontology development methods. The development of ontology can be performed using a variety of ontology engineering methodologies and techniques: 1) Domain analysis and classification of concepts and extraction of relationships in a domain; 2) Reverse engineering of existing robot programming languages and environments; 3) Ontology reuse and integration of existing ontologies or parts thereof (sub-ontologies, concepts) into new ontology.

We follow the methodology proposed by Uschold *et al.* [31] to define ontology, which consists of three stages. 1) *Brainstorming:* organize brainstorming session to identify all potential concepts and phrases in the domain of interest. 2) *Grouping:* structure the terms into provisional categories. 3) *Refinement*: refine the grouping and identify the semantic cross-reference between areas. According to [32], the following strategies for building ontologies can be applied: 1) *Application-dependent:* the ontology is built based on an application knowledge base using the abstraction process; 2) *Application-semidependent:* ontology is identified during software specification stage; 3) *Application-independent*: ontology is developed independently upon its future use. Strategies for identifying ontology concepts are [31]: 1) *bottom-up:* from the most concrete concepts to the most abstract one, 2) *top-down:* from the most abstract to the most concrete, and 3) *middle-out:* from the most relevant to the most abstract and most concrete.

In the development of our ontology, we follow the application-dependent strategy (focusing on the future development of a visual robot programming language), and after brainstorming and identification of most relevant concepts apply grouping, abstraction and refinement processes simultaneously to derive concept hierarchy and link it with existing concept hierarchies reused from other ontologies.

IV. DESCRIPTION OF THE ONTOLOGY

A. Structure and concepts of ontology

The developed ontology has a modular structure and consists of sub-ontologies at *Action*, *Algorithm*, *Communication*, *Grapheme*, *Location*, *Mission*, *Sensor* levels, which are combined at the *Core* level. The structure of ontology is given in Fig. 1 (we use OWLGrEd [33] for visualization). The summary of ontology concepts is presented in Table II. Current version of the ontology is available online at http://proin.ktu.lt/~damarobe/RobotOntology/ .

TABLE II. STRUCTURE OF ONTOLOGY

Ontology level	Concepts	Description
Action	ChangeLocation, ChangeSpeed, ChangeDirection	In current version, only simple moving concepts are described
Algorithm	OneInside, TwoInside, OneBounce, TwoBounce, FollowGray, Straddle	In current version, only line following algorithms described
Commun-ication	Language, Interface, Channel, Protocol, Command	Means and types of communication
Core	Thing, Physical Thing, SmartThing, Robot, MobileRobot	Hierarchical description of robot concept
Grapheme	Symbol, Image, Figure, Colour	Categories of graphical data representation
Location	Location	Spatial properties and relations
Mission	PushBox, FollowLine, FollowWall, AvoidObstaccle	Missions of mobile wheel robot
Sensor	Touch, Light, Sound, Distance, Temperature	Different kind of robot sensors
Time	Time, TimeRange, Interval, Instant	Temporal properties and relations

B. Core sub-ontology

The core ontology has high-level concepts that roughly define the robot and its environment. The "Robot" concept is a subclass of "SmartThing", which is in turn a subclass of "PhysicalThing". This hierarchy helps to separate the properties of a robot as a "thing", e.g., mass, shape, location, etc., from its unique properties as a smart thing (e.g., unique identification, services) and finally, a robot (e.g., means of powering, operational capabilities, degree of autonomy, sensory capabilities). Fig. 1 shows visualization of Core sub-ontology.

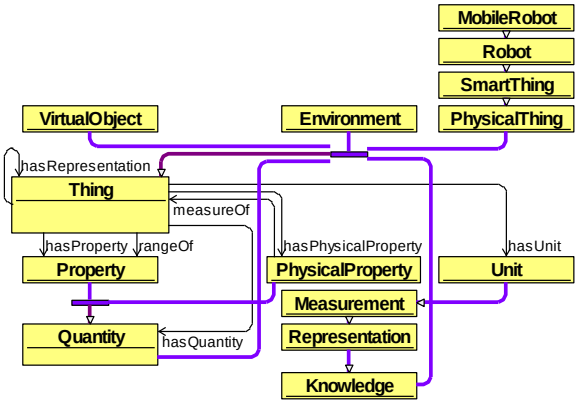


Fig. 1. Core ontology concepts

C. Sub-ontology of algorithms

This sub-ontology defines various algorithms that the robot can use to accomplish its mission. In the current version of the ontology, only line following algorithms are described. Another important concept in this ontology is "Data", which defines types of data and data structures which can be used in described algorithms. Data is further divided into "Input" and "Output" data. Fig. 2 shows visualization of sub-ontology of algorithms.

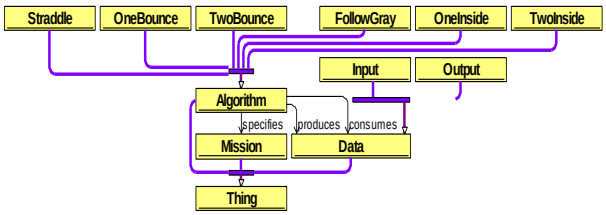


Fig. 2. Sub-ontology of algorithms

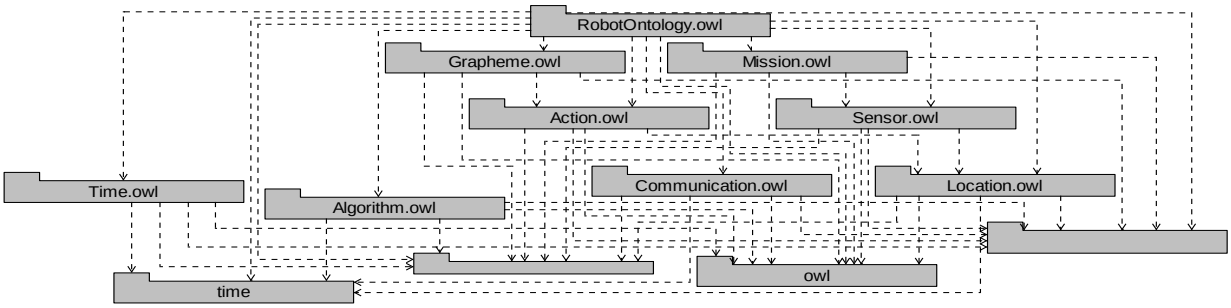


Fig. 3. Structure of proposed ontology

D. Sub-ontology of actions

The sub-ontology of actions is a direct subontology of the "Robot" class. It defines types of actions that the robot can perform. Currently defined are simple moving concepts that any mobile robot should be able to perform. Again, hierarchy is introduced in order to separate more abstract properties of actions. For example, a "Change location" class has "Drive backward" and "Drive forward" subclasses, which may not be suitable for legged robots. However, new subclasses can be easily introduced, under the same concept of "Change location". Fig. 4 shows visualization of sub-ontology of actions.

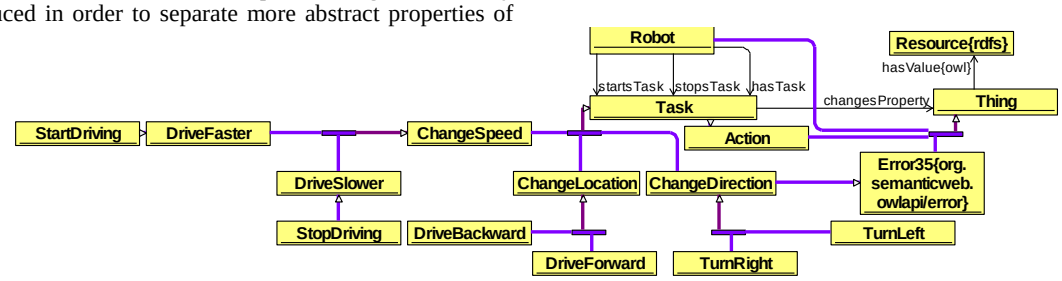


Fig. 4. Sub-ontology of actions

E. Sub-ontology of communication

Communication sub-ontology defines basic means and types for robot communication. The main concepts in this sub-ontology are "Communication", "Protocol" and "Interface". Fig. 5 shows visualization of sub-ontology of communication.

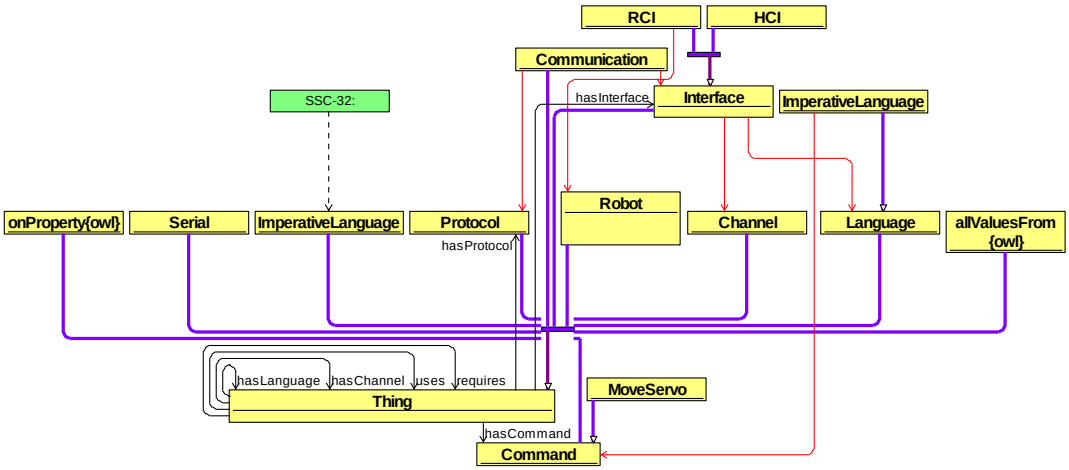


Fig. 5. Sub-ontology of communication

F. Sub-ontology of sensors

The last sub-ontology provides a description of the sensors that could be attached to the robot. Again, this is a sub-ontology of a "Thing" class. This means that the sensors are meant to be used not only by robots, but also by generic autonomous agents (of which robots are a subclass) for system-wide resource sharing, distributed decision making, autonomic operations. Fig. 6 shows visualization of sub-ontology of sensors.

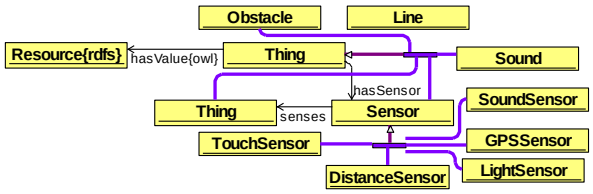


Fig. 6. Sub-ontology of sensors

G. Sub-ontology of graphemes

This sub-ontology describes simple objects of graphical data representation (e.g., figure, image, symbol, etc.). Fig. 7 shows visualization of sub-ontology of graphemes.

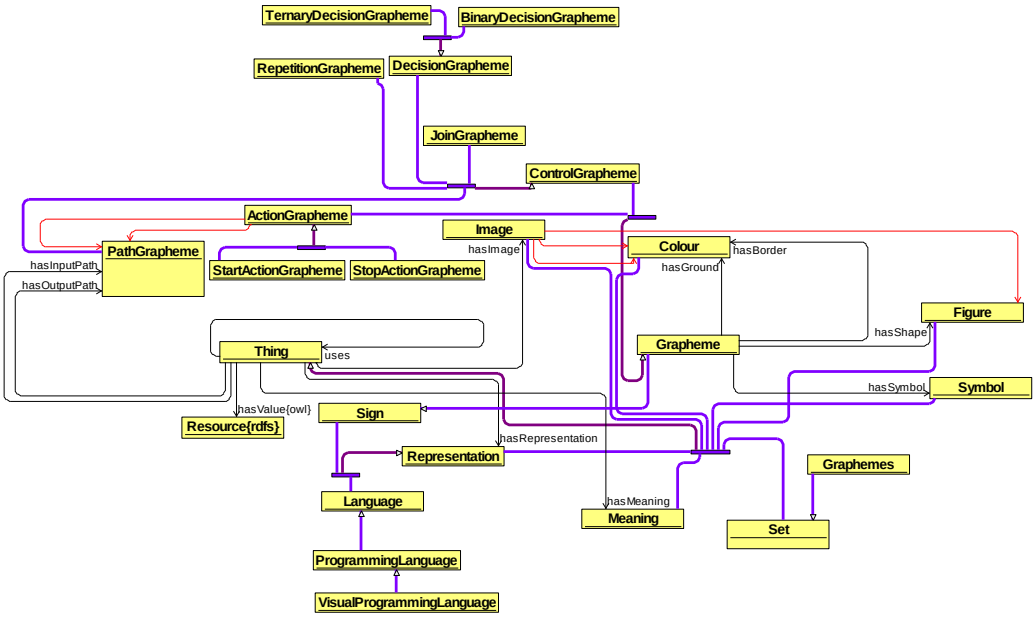


Fig. 7. Sub-ontology of graphemes

H. *Sub-ontology of missions*

Robots need to have a mission in which they can demonstrate their usefulness. Currently defined are simple demonstrative missions of a four-wheeled robot. Fig. 8 shows visualization of sub-ontology of missions.

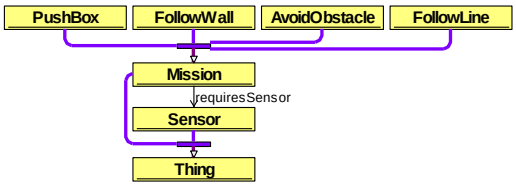


Fig. 8. Sub-ontology of missions

I. *Sub-ontology of location*

The location sub-ontology is an abstract sub-ontology, derived from "Physical Thing". It defines spatial properties relations of a robot as a "physical thing". The sub-ontology includes classes "Location", "Orientation" and distance. Fig. 9 shows visualization of sub-ontology of location.

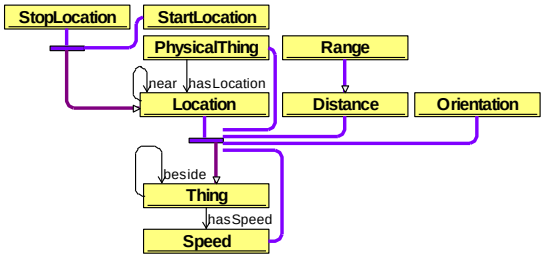


Fig. 9. Sub-ontology of location

J. *Sub-ontology of time*

Much like the sub-ontology of location, the sub-ontology of time defines basic temporal properties and relations of a robot as a "thing", such as "Instant" and "Interval". Fig. 10 shows visualization of sub-ontology of time.

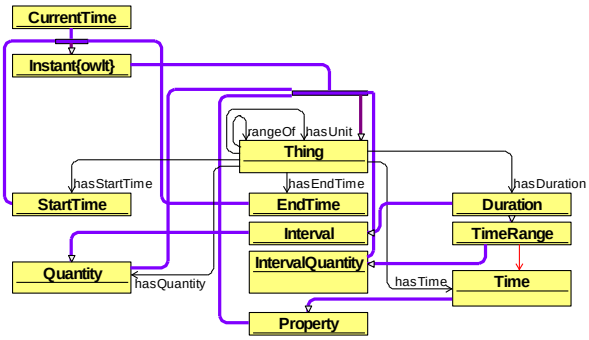


Fig. 10. Sub-ontology of time

K. *Summary*

The characteristics of the proposed ontology and its comparison with other ontologies are summarized in Table III. We used University of Manchester's Ontology Metrics tool (http://owl.cs.manchester.ac.uk/metrics).

Although not as big as the other compared ontologies, the proposed ontology covers the robotics domain more evenly.

TABLE III. COMPARISON WITH OTHER ONTOLOGIES

Ontology	Classes	Object properties	Data properties	Object property axioms	Data property axioms	Logical axioms
PROTEUS	404	214	42	545	430	2674
Dolce-LITE	37	70	0	255	140	349
Robot Programming Ontology	93	48	19	22	10	37

V. CONCLUSIONS AND FUTURE WORK

To the best of our knowledge so far there has been no ontology developed specifically for the robot programming domain. We have developed a draft version of ontology for the robot programming domain. The ontology is still in development and will be extended as required. It can be considered as a first step towards development of a comprehensive ontology of robotics programming domain. The ontology integrates concepts from third party robotics, Internet of Things, sensor ontologies as well as relevant concepts (time, location, etc.) from upper-level ontologies and additional concepts required to describe robot programming.

The ontology has 93 classes, 48 object properties, 19 data properties and 37 logical axioms. Other analysed ontologies are larger and more mature, but they are not focused directly on the robotics programming domain. The developed ontology will be used a semantic foundation for the design of a visual robot programming language. The ontology is freely available and other researchers will be able to use it to develop Semantic Web applications for the robotics domain.

Future work will include application of the ontology for the development of visual robot programming language.

REFERENCES

[1] B. Smith, "Ontology," in The Blackwell guide to philosophy of computing and information, L. Floridi, Ed., Malden: Blackwell, 2004, pp. 153-166.

[2] D. L. McGuinness and F. Van Harmelen, "OWL web ontology language overview," W3C recommendation, vol. 10, no. 2004–03, p. 10, 2004.

[3] Surhone, L.M., Tennoe, M.T., Henssonow, S.F, "Microsoft Visual Programming Language," Betascript Publishing, 2010.

[4] Kelly, J. F, "LEGO MINDSTORMS NXT-G Programming Guide," Apress, 2007.

[5] J. Hallam, H. Bruyninckx. "An Ontology of Robotics Science," in European Robotics Symposium 2006, H. Christensen, Ed., Springer, 2006, pp. 1-14.

[6] B. Kuipers, "The spatial semantic hierarchy," Artificial Intelligence, vol. 119, no. 1, pp. 191–233, 2000.

[7] L. Paull, G. Séverac, G.V. Raffo, J.M. Angel, et al., "Towards an Ontology for Autonomous Robots," in Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems, 2012, pp.1359-1364.

[8] C. Schlenoff, E. Prestes, R. Madhavan, P. Goncalves, H. Li, S. Balakirsky, et al, "An IEEE standard Ontology for Robotics and Automation", Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems, 2012, pp. 1337-1342.

[9] S. Dhouib, N. Du Lac, J.-L. Farges, S. Gerard, M. Hemaissia-Jeannin, J. Lahera-Perez, et al., "Control Architecture Concepts and Properties of an Ontology Devoted to Exchanges in Mobile Robotics," in 6th National Conference on Control Architectures of Robots, 2011.

[10] J. Hallam and H. Bruyninckx, "An ontology of robotics science," in European Robotics Symposium 2006, 2006, pp. 1–14.

[11] K.J. Witt, J. Stanley, D. Smithbauer, „Enabling Sensor Webs by Utilizing SWAMO for Autonomous Operations," in NASA Earth Science Technology Conference (ESTC2008), 2008.

[12] J.A. Bateman, S. Farrar," Modelling Models of Robot Navigation Using Formal Spatial Ontology," Spatial Cognition, 2004, pp. 366-389.

[13] M. MacMahon, B. Stankiewicz, B. Kuipers, "Walk the talk: Connecting language, knowledge, and action in route instructions," Proceedings of the 21st national conference on Artificial intelligence, Vol. 2. AAAI Press, 2006, pp. 1475-1482.

[14] S. Coradeschi, A. Saffiotti, "Perceptual anchoring of symbols for action," Proc. of the 17th International Joint Conference on Artificial Intelligence, Seattle, WA, 2001, pp. 407-412.

[15] S. Lauria, T. Kyriacou, G. Bugmann, J. Bos, E. Klein, "Converting natural language route instructions into robot executable procedures," Proc. of IEEE Int. Workshop on Robot and Human Interactive Communication, Berlin, Germany, 2002, pp. 223-228.

[16] R. Chatterjee, F. Matsuno, "Robot description ontology and disaster scene description ontology: analysis of necessity and scope in rescue infrastructure context," Advanced Robotics, 19(8), pp. 839–859, 2005.

[17] L. N. Soldatova, A. Clare, A. Sparkes, and R. D. King, "An ontology for a Robot Scientist," Bioinformatics, 22(14), pp. e464–e471, 2006.

[18] J. P. Diprose, B. Plimmer, B. A. MacDonald, J. G. Hosking. How People Naturally Describe Robot Behaviour, Proc. of Australasian Conf. on Robotics and Automation, Wellington, New Zealand, pp. 3-5, 2012.

[19] A. Herzog, D. Jacobi, and A. Buchmann, "A3ME-an Agent-Based middleware approach for mixed mode environments," in Second Int. Conf. on Mobile Ubiquitous Computing, Systems, Services and Technologies, 2008, pp. 191-196.

[20] P. Lando, A. Lapujade, G. Kassel, and F. Fürst, "Towards a general ontology of computer programs," in Proc. of the 2nd Int. Conference on Software and Data Technologies, ICSOFT, 2007, pp. 25.

[21] H. Neuhaus, M. Compton, "The Semantic Sensor Network Ontology: A Generic Language to Describe Sensor Assets," in AGILE Workshop Challenges in Geospatial Data Harmonisation, 2009.

[22] K. Kotis and A. Katasonov, "Semantic Interoperability on the Web of Things: The Semantic Smart Gateway Framework," in Proc. of 6th Int. Conf. on Complex, Intelligent, and Software Intensive Systems, CISIS, 2012, pp. 630–635.

[23] S. Hachem, T. Teixeira, and V. Issarny, "Ontologies for the Internet of Things," Proc. of the 8th Middleware Doctoral Symposium, 2011, p. 3.

[24] G. Stevenson, S. Knox, S. Dobson, P. Nixon, "Ontonym: a collection of upper ontologies for developing pervasive systems," Proc. of 1st Workshop on Context, Information and Ontologies, CIAO 2009, pp.1-8.

[25] C. Goodwin, D. J. Russomanno, "An ontology-based sensor network prototype environment," in 5th Int. Conf. on Information Processing in Sensor Networks, 2006.

[26] R. Trypuz. Formal Ontology of Action: a Unifying Approach. Doctoral dissertation, DIT - University of Trento, 2007.

[27] M. Eid, R. Liscano, A. El Saddik, "A Novel Ontology for Sensor Networks Data," Proc. of IEEE Int. Conf. on Computational Intelligence for Measurement Systems and Applications, pp. 75-79, 2006.

[28] A. Robin, M. E. Botts, "Creation of Specific SensorML Process Models", White paper, University of Alabama in Huntsville, 2006.

[29] L. Lefort, C. Henson, K. Taylor, P. Barnaghi, M. Compton, O. Corcho, et al, "Semantic sensor network xg final report," W3C Incubator Group Report, 2011.

[30] Y. Hu, Z. Wu, M. Guo, "Ontology driven adaptive data processing in wireless sensor networks," Proc. of the 2nd international conference on Scalable information systems, InfoScale '07, pp. 1-2, 2007.

[31] U. Uschold, M. Gruninger, "Ontologies: Principles, methods and applications", Knowledge Engineering Review, 11(2), pp. 93-155, 1996.

[32] M. Fernández-López, A. Gómez-Pérez, "Overview and analysis of methodologies for building ontologies", The Knowledge Engineering Review, 17(2), 129-156, 2002.

[33] J. Bārzdiņš, G. Bārzdiņš, K. Čerāns, R. Liepiņš, A. Sproģis, "UML Style Graphical Notation and Editor for OWL 2", Lecture Notes in Business Information Processing, Vol. 64, Part 2, pp. 102-114, 2010.

INFORMACINĖS TECHNOLOGIJOS ● IVUS 2013 ● ISSN 2029-249X ● eISSN 2029-4824

Atstumų skaičiavimas iki ieškomų žodžių ištisinėje šnekoje

V. Pnevski
Matematikos ir informatikos institutas, Vilniaus universitetas,
A. Goštauto g. 12, LT-01108 Vilnius, Lietuva
vladislav.pnevski@mii.vu.lt

Santrauka — Šio darbo tikslas - parodyti vieną iš galimų būdų, kaip galima rasti žodį kalbos signale. Straipsnyje nagrinėjamas kalbos signalo analizės, palyginimo, žodžio radimo iš ilgo laiko trunkančio signalo metodas. Aprašomas visas algoritmas, teorija, kokiu būdu buvo pasiektas teigiamas rezultatas. Iliustracijos parodo vaizdinę informaciją apie tos dalies sugalvotą sprendimo būdą. Pasirinktas dinaminio programavimo algoritmas, kepstro analizė ir palyginimo metodas. Atlikti eksperimentai parodė metodo veikimą ir gautą teigiamą rezultatą.

Reikšminiai žodžiai — Kalbos apdorojimas, signalo analizė, atstumų skaičiavimas, kalbos analizė.

I. ĮVADAS

Pasaulio duomenų bazėse yra daugybė šnekos įrašų, kuriuose reikia susigaudyti, perklausyti tam tikrą pateiktą informaciją, ją atrinkti ar panaudoti reikalingiems darbams atlikti. Tam tikslui yra kuriamos programos, tokios kaip „Žodžių atpažintuvas“, „Atpažinimas“, kitos, pagrįstos paslėptaisiais Markovo modeliais ir t.p., kuriuose atpažįstama reikiama informacija. Mano dėmesį patraukė programa „Atpažinimas“ ir jos autorius G. Tamulevičius, savo disertacijoje aprašęs jos išskirtinę veikimo principą [1]. Todėl mano darbo tikslas yra rasti ieškomąjį žodį kalbos signale, pasinaudojus vienų iš analizės metodų ir signalo palyginimo būdų. Tam tikslui pasiekti reikia ištirti metodus ir priemones lietuvių šnekos žodžiams atpažinti nepalankioje akustinėje aplinkoje.

Kaip tyrimo objektą pasirinkau nežinomąją ištisinę šneką, kurioje reikia surasti reikiamą žodį ir nurodyti laiką tos vietos, kur yra ieškomas žodis, galintis pasirodyti kelis kartus tam tikru laiku.

Šiame darbe pateiktas sukurtas metodas, algoritmas ir eksperimentai, parodantys kaip galima rasti reikiamąjį žodį ar frazę kalbos signale.

Šiems eksperimentams atlikti buvo pasirinktas dinaminis laiko skalės kraipymo metodas. Metodas dažniausiai buvo taikomas pavieniams žodžiams atpažinti, tačiau buvo bandymų pritaikyti žodžių junginiams ir net ištisinei kalbai atpažinti. Iškyla vienas klausymas - kaip palyginti dvi lygias dalis ir nurodyti, jog jos yra panašios. Panašumui apibrėžti naudosim jų elementų atstumus. Kuo atstumas tarp jų bus mažesnis, tuo jie tampa panašesni.

II. TEORIJA

Signalų apdorojimas. Kadangi signalai dažnai būna įrašyti skirtinga įranga, tai signaluose gali atsirasti nuolatinė dedamoji. Dar ji gali atsirasti dėl išorinių triukšmų įrašant balsą, todėl signalas išsikraipo. Kad tokie iškraipymai netrukdytų signalo analizei ir signalų palyginimui, reikia nuolatinę dedamąją pašalinti. Apdorojimas vyksta dviem etapais.

Pirmuoju veiksmu pašaliname nuolatinę dedamąją iš $A(n)$ ir $B_k(n)$ signalų pagal tokią išraišką:

$$\begin{aligned}\tilde{A}(n) &= A(n) - 1/N \sum_{n=1}^{N} A(n),\\ \tilde{B}_k(n) &= B_k(n) - 1/N \sum_{n=1}^{N} B_k(n),\end{aligned} \quad (1)$$

kai $n = 1,2,...,N$, N – signalo ilgis atskaitomis, k – signalo atkarpos numeris, $B_k(n)$ - signalo dalis, $A(n)$ - etalonas.

Antruoju veiksmu apskaičiuojamas I eilės RIR filtras, kurio išraiška yra:

$$\begin{aligned}R(n) &= \tilde{A}(n) - \alpha\tilde{A}(n-1),\\ Z(n) &= \tilde{B}_k(n) - \alpha\tilde{B}_k(n-1),\end{aligned} \quad (2)$$

kai $n = 1,2,...,N$, N – signalo ilgis atskaitomis, k – signalo atkarpos numeris, α – svorinis koeficientas, kuris pasirenkamas nuo 0 iki 1, $R(n)$ – etalonas, $Z(n)$ – signalo atkarpa. Jis reikalingas sumažinti nestabilumų tikimybę [9], taip pat leidžia sumažinti klaidos signalo svyravimus, gaunamus tiesinės prognozės analizei naudojant autokoreliacinį metodą. Teigiama, kad autokoreliacijos metodo stabilumui palaikyti 10–12 tiesinės prognozės eilė yra pakankama [2,5].

Analizė. Atliekama TPMK analizė. Tiesinės prognozės modelio kepstro analizė leidžia pasiekti didžiausią atpažinimo tikslumą [3,7]. Signalo analizė vyksta kadrais, kiekvieną langą dauginant iš Hamming lango [9]. Analizės kadro ilgis lygu

250 atskaitų (22,6 ms), postūmis lygu 125 atskaitų (11,3 ms) [1]. Apskaičiuojami etalono R ir kiekvienos Z atkarpos elementai pagal žemiau parašytą algoritmą. Atitinkamai vienu atveju apskaičiuojant etalono R elementų vektorius yra r_i, kuriam priskiriamas c_m kepstro vektorius, kitu atveju apskaičiuojant Z atkarpos elementų vektorius z_j, kuriam priskiriamas c_m kepstro vektorius.

$$c_0 = \log G; \quad (3)$$

$$c_m = a_m + \sum_{k=1}^{m-1}(k/m)c_k a_{m-k}, \text{ kai } 1 \le m \le p; \quad (4)$$

$$c_m = \sum_{k=1}^{m-1}(k/m)c_k a_{m-k}, \text{ kai } m > p. \quad (5)$$

Čia a_m – m-asis autokoreliacijos koeficientas, c_m – m-asis kepstro koeficientas, p – autokoreliacijos eilė, c_0 – nulinis kepstro koeficientas, G – pasirengta pirminė funkcija, priklausanti nuo signalo, o $a_m = a_j$ apskaičiuojamas pagal Levinsono-Durbino algoritmą [5,6]:

Inicializacija

$$E^{(0)} = r(0). \quad (6)$$

Iteracija, kai $1 \le i \le p$;

$$k_i = (r(i) - \sum_{j=1}^{i-1} a_j^{(i-1)} r(i-j))/E^{(i-1)}; \quad (7)$$

$$a_i^{(i)} = k_i; \quad (8)$$

$$a_j^i = a_j^{(i-1)} - k_i a_{i-j}^{(i-1)}, \text{ kai } 1 \le j < i; \quad (9)$$

$$E^{(i)} = (1-k_i^2)E^{(i-1)}. \quad (10)$$

Užbaigimas:

$$a_j = a_j^{(p)}, \text{ kai } 1 \le j \le p. \quad (11)$$

Atliekama kepstro analizė su vidurkio atėmimu $\overline{(TPMK)}$ [1]:

$$c_m = c_m - \sum_{k=1}^{K} c_k / k, \text{ kai } m = 1,2,...,K, \quad (12)$$

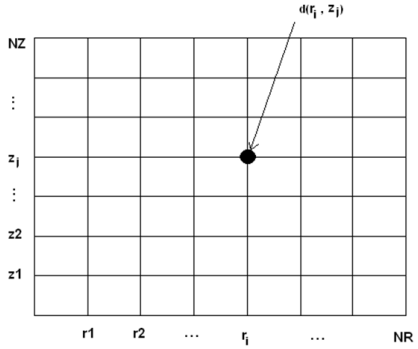
čia c_m – m-asis kepstro vektorius, K – kepstro koeficientų skaičius. Vienu atveju apskaičiuojant etalono R elementų vektorius yra r_i, kitu atveju apskaičiuojant Z atkarpos elementų vektorius z_j, kur $i, j = 1,...,K$, r_i – kepstrinių koeficientų etalono vektorius, z_j – kepstrinių koeficientų atkarpos vektorius.

Paliginimas. Lyginant du signalus reikia lyginti jų kepstrinių koeficientų vektorius [1]. Visų vektorių apskaičiuoti euklidiniai atstumai surašomi į matricą, kaip parodyta 1 pav..

Atstumai tarp etalono R elementų $i = 1,..., NR = N$ ir signalo Z atkarpų elementų $j = 1,..., NZ = N$ skaičiuojamas pasitelkiant išraišką:

$$d(r_i, z_j) = \sqrt{(r_i - z_j)^2}, \quad (13)$$

kur r_i, z_j kepstrinių koeficientų vektorių reikšmės.



1 pav. B_k elemento $d(r_i, z_j)$ atstumas tarp vektorių r_i ir z_j

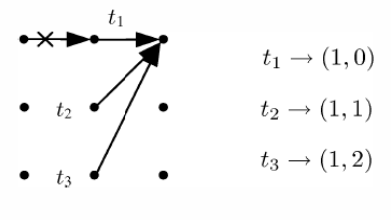
Paveikslėlyje parodytas vieno elemento $d(r_i, z_j)$ tarp r_i ir z_j vektorių, kitaip tariant, tarp kepstro požymių koeficientų etalono ir kepstro požymių koeficientų nagrinėjamos atkarpos.

Po visos matricos užpildymo (kurį toliau vadinsime tinkleliu), trumpiausi atstumai tarp etalonų R ir Z atkarpų paskaičiuojami pagal formulę:

$$D_{RZ}(\phi) = \sum_{i,j=1}^{N} \phi(i,j), \quad (14)$$

kur $D_{RZ}(\phi)$ – suminis atstumas tarp pavyzdžių, $\phi(i,j)$ - atstumas, skaičiuojamas pagal Itakura taisyklę, parodytą 2 pav., kur $i, j = 1,..., N$ – tinklelio koordinačių taškai, N – tinklelio ilgis, $\phi = 1,..., n$ - visi galimi perėjimo variantai, n -perėjimo variantų skaičius,

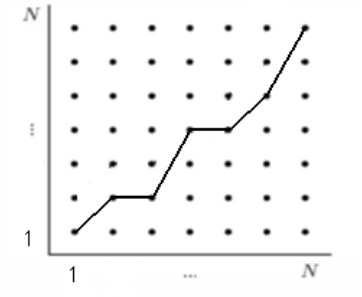
$$\phi(i,j) = \sum_{i,j=1}^{N} d(r_i, z_j). \quad (15)$$



2 pav. Itakura perėjimo taisyklė

Trajektorijos paieška vykdoma pagal dinaminio programavimo algoritmą [10,11].

Optimalios trajektorijos paieška pradedama taške (1,1), pasibaigia taške (N,N). Nagrinėjant tinklelio taškus, atsižvelgiama į globalius ir lokalius apribojimus, taigi nagrinėjamų trajektorijų skaičius sumažėja. Pasiekus galutinį tinklelio tašką, lieka tik viena trajektorija, kuri garantuoja minimalų atstumą tarp lyginamųjų pavyzdžių [1]. 3 pav. parodyta vienintelė galima trajektorija.



3 pav. Vienintelė galima trajektorija

$$D_{RZ} = \min_{\phi} D_{RZ}(\phi). \qquad (16)$$

Čia D_{RZ} – trumpiausias atstumas tarp etalono ir signalo atkarpos, ϕ – visos galimos trajektorijos.

Taigi atstumas gaunamas minimizuojant uždavinį, radus trumpiausią kelią nuo pirmojo iki paskutiniojo taško.

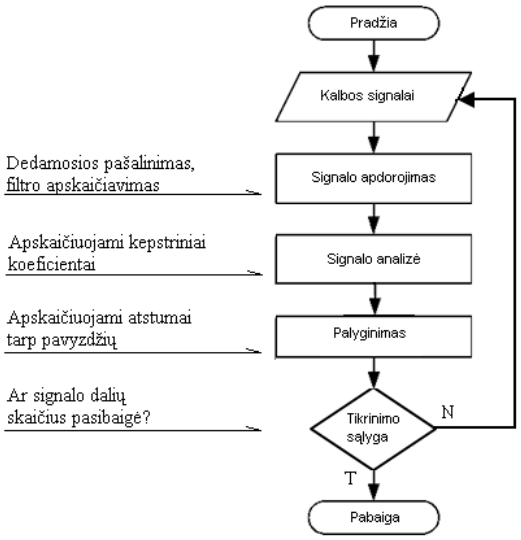
III. ALGORITMAS

Kaip atpažinimo metodą pasirinkome dinaminį laiko skalės kraipymą. Pasirinkimą lėmė įsitikinimas, kad pavieniams žodžiams atpažinti dinaminis laiko skalės kraipymo metodas yra tinkamesnis dėl savo paprastumo ir efektyvumo [1]. Signalui analizuoti panaudojome tiesinės prognozės modelio kepstro analizės metodus. Toliau 4 pav. yra pavaizduojama algoritmo schema.

Šnekos signalo ir etalono parinkimui reikia teisingai parinkti garso įrašus. Kaip nežinomas šnekos signalas, taip ir etalonas įrašomi iš anksto per garso redaktorių. Signalai turi būti įrašyti vieno tipo failų (pavyzdžiui: „Windows“ operacinėje sistemoje – wav) ir atitikti šiuos techninius parametrus:

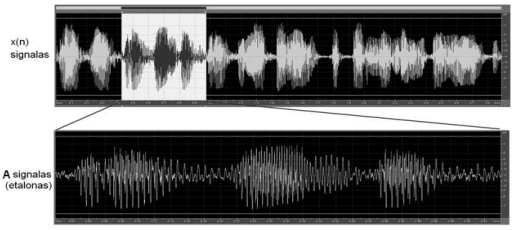
- garso kokybė 11025 Hz;
- skiltiškumas 16 bitų;
- vieno kanalo garsas.

Nežinomas garso įrašas turi būti ne trumpesnis nei sukurtas etalonas.



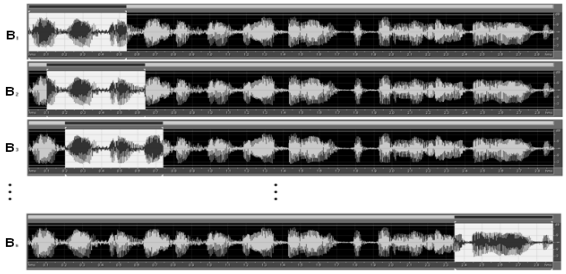
4 pav. Algoritmo schema

Turime diskretinio laiko šnekos signalą, pažymėsime jį $x(n)$, n =1,...,N. Iš jo dalies padaromas etalonas arba etalonas sukuriamas įrašant balsą per mikrofoną, pasinaudojus garso redaktorium, kurį žymėsime A.



5 pav. Etalono A sukūrimas iš $x(n)$ signalo

$x(n)$ signalas yra dalinamas į $B_k(n)$ atkarpas, kurios lygios etalono ilgiui. $B_k(n)$ – tai $x(n)$ signalo dalis, kuri bus lyginama su etalonu. $A(n)$, kur $k = 1,..., N$, $n = 1,..., N$. Atkarpų skaičius priklauso nuo $x(n)$ signalo ilgio. Kad gautume kitą $B_k(n)$ signalo atkarpą, signalas $B_{k-1}(n)$ pastumiamas žingsniu l (kaip parodyta 2 pav.), l gali būti pasirinktas bet koks.



6 pav. B_k signalų sukūrimas

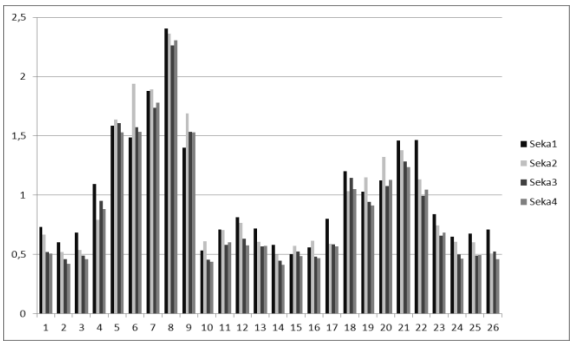
Įkrauname etaloną ir kalbos įrašą į programos atmintį, kurioje vyksta signalų apdorojimas. Vėliau atliekama signalų analizė, apskaičiuojamos etalono kepstro vektorių reikšmės ir kalbos įrašo atkarpų vektorių reikšmės slenkant nurodytu atstumu, kiekvieną kartą iš naujo apskaičiuojant kepstro reikšmes. Visų gautų etalono ir kalbos signalo atkarpų vektorių reikšmių euklidiniai atstumai apskaičiuojami ir surašomi į matricos pavidalo tinklelį, kuriame pagal Itakura taisyklę vykdomas optimizavimo uždavinys. Jo rezultatas yra reikšmė – skaičius, nurodantis atstumą tarp etalono ir kalbos įrašo dalies. Tokių būdų gaunami visi atstumai. To atstumo, kuris bus trumpiausias reikšmė mažiausia, ir galima teigti, jog tai yra ta kalbos įrašo dalis (žodis), kuris ieškomas. Galima sakyti, kad suradome ieškomo žodžio vietą (laiką), sudėjus visus postūmio žingsnius (laiką) nuo įrašo pradžios iki trumpiausiojo atstumo.

IV. EKSPERIMENTAI

Preliminarūs eksperimentai buvo atlikti panaudojant skirtingų žmonių frazes, iškirptas iš kalbos įrašų (žinių), taip pat sugalvotų frazių. Ieškota įvairiausių žodžių (prezidentas, namas, ...) tose frazėse, su sąlyga, jog jos iš tikrųjų ten yra. Patikrinus, jog metodas veikia, pradėti rimtesni tyrimai.

Eksperimentams atlikti buvo panaudoti 4–i sukurti etalonai „banko“ ir viena frazė, ištarta 4 kartus skirtingų žmonių. Dalyvavo 4 žmonės – trys vyrai ir viena moteris, kurie ištarė po vieną kartą sugalvotą sakinį: „Namas šalia banko stovi“. Šiame sakinyje buvo ieškotas žodis „banko“. Palyginti visi etalonai su frazių dalimis ir surašyti visi atstumai tarp etalonų ir frazių dalių. Gautame rezultate, kuriame buvo taikyti trys etalonai ieškant žodžio vietos ištartoje frazėje, iš keturių bandymų trijose frazėse buvo rasta žodžio vieta, viename bandyme neteisingai nustatyta žodžio vieta. Žodžio vieta nustatom pagal trumpiausią atstumą tarp etalono ir lyginamosios dalies.

7 pav. parodytas vienas iš bandymų rezultatų, kuriame matomi keturių etalonų atstumai iki kiekvienos frazės dalies. Vienas stulpelis atitinka vieną atstumą. Horizontaliai sugrupuoti po keturis etalonai rodo laiką, vienas skaičius lygus 0,1 sekundės. Vertikaliai atidėtos atstumų reikšmės tarp etalonų ir frazės dalies. Pagal rezultatus galima atsakyti, kurioje vietoje gali būti ieškomasis žodis, tai yra 14 numeriu pažymėti stulpeliai. Tikroji žodžio vieta yra ties numeriu 14. Tačiau yra nemažai panašių rezultatų, todėl reikia sugalvoti būdą, kuris atsakytų į klausimą – kuriuos rezultatus priimti, kuriuos ne.



7 pav. Etalonų atstumai iki kiekvienos frazės dalies

V. IŠVADOS

Naudotas metodas veikia, taikant dinaminio programavimo algoritmą. Taip pat pavaizduotas variantas, kai kalbos signalas suskaldomas į mažesnes dalis, lygias etalono ilgiui, ir juos lyginame surašant visus rezultatus. Atlikus eksperimentus, buvo gauti teigiami ir neigiami rezultatai. Atsirasdavo tokių vietų, kur žodžio vieta buvo nustatyta klaidingai arba netiksliai. Taip pat pastebėta, jog ieškant žodžio naudojant kelis etalonus tiksliau nustatoma ieškomo žodžio vieta nei su vienu etalonu.

LITERATŪRA

[1] G. Tamulevičius. Pavienių žodžių atpažinimo sistemų kūrimas. VGTU leidyklos technika, 1495-m mokslo literatūros knyga, isbn 978-9955-28-279-2, 2008, p. 39–46.

[2] Atal, b. S.; hanauer, s. L. Speech analysis and synthesis by linear prediction of the speech wave. The journal of the acoustical society of america, 1971, t. 50, nr. 2, p. 637-655.

[3] Davis, s. B.; mermelstein, p. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. Ieee transactions on acoustics, speech and signal processing, 1980, t. 28, nr. 4, p. 357-366.

[4] Gold, b.; morgan, n. Speech and audio signal processing: processing and reception of speech and music. John wiley & sons, inc., new york, 2000. Isbn 0-471-35154-7.

[5] Makhoul, j. Linear prediction: a tutorial review. Proceedings of the ieee, 1975, t. 63, nr. 4, p. 561-580.

[6] Markel, j. D.; gray, a. H. Linear prediction of speech. Springer-verlag, berlin, 1976. Isbn 3-540-07563-1.

[7] Paliwal, k. K.; rao, p. V. S. Evaluation of various linear prediction parametric representations in vowel recognition. Signal processing, 1982, t. 4, nr. 4, p. 323-327.

[8] Rabiner, l. R.; atal, b. S.; sambur, m. R. Lpc prediction error - analysis of its variation with the position of the analysis frame. Ieeeassp, 1977 t. 25, nr. 5, p. 434-442.

[9] Rabiner, l. R.; schafer, r. W. Digital processing of speech signals. Prentice-hall, new jersey, 1978. Isbn 0-13-213603-1.

[10] Sakoe, h.; chiba, s. Dynamic programming algorithm optimization for spoken word recognition. Ieee transactions on acoustics, speech and signal processing, 1978, t. 26, nr. 1, p. 43-49.

[11] Винцюк, т. К. Распознавание слов устной речи методами динамического программирования. Кибернетика, 1968, t. 1, p. 81-88.

[12] White, g. M.; neely, r. B. Speech recognition experiments with linear predication, band pass filtering, and dynamic programming. Ieee transactions on acoustics, speech and signal processing, 1976, t. 24, nr. 2, p. 183-188.

INFORMACINĖS TECHNOLOGIJOS ● IVUS 2013 ● ISSN 2029-249X ● eISSN 2029-4824

Balso komandų aptikimas kalbos signale naudojant naivų Bajeso klasifikatorių

Tomas Rasymas
Kauno Humanitarinis fakultetas
Vilniaus Universitetas
Kaunas, Lietuva
tomas.rasymas@gmail.com

Santrauka: Straipsnyje nagrinėjamas metodas, padedantis automatinei kalbos atpažinimo sistemai išskirti kalbą iš triukšmingo signalo. Metodas pagrįstas tokiomis, iš atskirų signalo pjūvių išskiriamomis, kalbos savybėmis: garso slėgio lygis (dB), trumpalaikė energija, nulių kirtimų dažnis. Šneka klasifikuojama į dvi klases: triukšmas ir kalba. Klasifikavimas atliekamas naudojant naivų Bajeso klasifikatorių.

Raktiniai žodžiai: naivus Bajeso klasifikatorius, kalbos atpažinimas, kalbos aptikimas signale, kalbos savybės, garso slėgio lygis, nulio kirtimų dažinis, trumpalaikė kalbos energija.

I. ĮŽANGA

Viena iš svarbiausių kalbos atpažinimo proceso dalių yra šnekos aptikimas kalbos signale. Netaikant papildomų metodų kalbos atpažinimo sistemai sunku identifikuoti, kuri informacija kalbos signale yra naudinga (kalba), o kur yra tik triukšmas. Naudojant kalbos aptikimo algoritmą yra sumažinamas klaidingai atpažintų žodžių procentas bei pasiekiamas didesnis greitis ir efektyvumas, nes sistemai nereikia analizuoti neinformatyvių (triukšmingų) kalbos signalo dalių.

Dažnai kalbos aptikimo problema sprendžiama analizuojant kalbos energiją: jei ji viršija tam tikrą dydį, identifikuojama kalbos pradžia, o jeigu energija nukrenta žemiau tam tikro dydžio – identifikuojama, kad tai yra kalbos pabaiga [1] [2]. Tokie metodai, kuriuose kalbos aptikimui naudojami statiniai dydžiai, dažnai neduoda gerų rezultatų [1]. Tad šiame darbe bus tiriama galimybė atskirti kalbą nuo triukšmo pakeičiant statinius kalbos pradžios ir pabaigos dydžius apmokytu naiviu Bajeso klasifikatoriumi.

II. KALBOS POŽYMIŲ APRAŠYMAS

A. Pradinis kalbos signalo apdorojimas

Prieš pradedant vykdyti požymių išskyrimą, kalbos signalas yra išskaidomas į nepersidengiančius kadrus, kurių ilgis yra 25 ms. Tada kiekvienam kadrui yra atliekamas pradinis kalbos signalo apdorojimas.

Pradinis kalbos signalo apdorojimas taikomas kalbos signalo spektro išlyginimui – siekiama, kad signalo spektro energija būtų pasiskirsčiusi visame dažnių diapazone tolygiai [1]. Todėl šiam tikslui pasiekti yra naudojamas pirmos eilės ribotos impulsinės reakcijos filtras. Šis filtras laiko srityje išreiškiamas:

$$S(n) = s(n) – as(n - 1) \quad (1)$$

Šioje formulėje S(n) – apdorojamo signalo išėjimo signalas, s(n) – diskretizuoto kalbos signalo reikšmė, a – koeficientas, nurodantis spektro išlyginimo laipsnį (dažniausiai taikoma 0,90–0,98). Iš esmės šis filtras naudojamas dėl to, jog naudingą informaciją sauganti kalbos signalo dalis turi neigiamą nuožulnumą ir yra susikoncentravusi žemų dažnių srityje, o tai apdorojant signalą sukelia didesnes tikslumo paklaidas [1].

B. Nulio kirtimų dažnis

Nulio kirtimų dažnis (NKD) rodo, kiek kartų analizuojamame kalbos signalo kadre buvo kirstas nulio lygis [3].

$$NKD = \frac{\sum_{n=1}^{N} \left| \operatorname{sgn} s(n) - \operatorname{sgn} s(n-1) \right|}{2N} \quad (2)$$

Formulėje N – diskretinio signalo ilgis, s(n) – diskretinis signalas, sgn – ženklo funkcija (signum) [3].

Šis dydis skiriasi, kai kalbos signale yra tik triukšmas arba tik kalba [3], todėl jis buvo pasirinktas kaip vienas iš požymių, pagal kuriuos bus bandoma atskirti kalbą nuo triukšmo.

C. Trumpalaikė energija

Trumpalaikė kalbos signalo kadro energija parodo, kiek energijos yra kiekviename kadre. Tai leidžia identifikuoti kadrą kaip triukšmą arba kaip šneką [4]. Trumpalaikė energija (E) skaičiuojama pagal formulę:

$$E = \sum_{n=1}^{N} s(n)^2 \quad (3)$$

Formulėje N – diskretinio signalo ilgis, s(n) – diskretinis signalas [4].

D. Garso slėgio lygis (sound pressure level)

Garso slėgio lygis (garso galia) yra matuojamas decibelais (dB). Šis dydis rodo garso bangos intensyvumą [5]. Garso slėgio lygio (GSL) formulė:

$$GSL = 20\log_{10}(x) \tag{4}$$

Formulėje x – reikšmė, kuriai bus skaičiuojama GSL, o x skaičiuojamas pagal formulę:

$$x = \frac{\sum_{n=1}^{N}|s(n)|}{N} \tag{5}$$

Formulėje N – diskretinio signalo ilgis, s(n) – diskretinis signalas.

III. NAIVUS BAJESO KLASIFIKATORIUS

Klasifikavimo pagrindinis tikslas yra pagal turimus duomenis apmokyti klasifikatorių taip, kad gavus naują, nežinomą objektą su tam tikrais parametrais, jį būtų galima priskirti tam tikrai klasei. Šiuo atveju klasifikavimo problemai spręsti naudojamas tiesinis klasifikatorius – naivus Bajeso klasifikatorius. Jis bus apmokytas atskirti dvi klases: triukšmą ir kalbą. Bajeso formulė:

$$P(w \mid x) = \frac{P(x \mid w)P(w)}{P(x)} \tag{6}$$

Formulėje P(w) – apriorinė w klasės tikimybė, P(w|x) – sąlyginė w klasės tikimybė, kai žinomi x duomenys, P(x|w) – sąlyginė x duomenų tikimybė, kai žinoma, kad jie yra w klasės, P(x) – apriorinė x duomenų tikimybė [6].

Pagal naivų Bajeso klasifikatorių, visos duomenų komponentės yra nepriklausomos, todėl pirmiausia apskaičiuojamos kiekvieno įvykio tikimybės atskirai , o iš jų gaunama bendra įvykio tikimybė. Tikimybė, jog tai buvo triukšmas (triuk) ir tikimybė, jog tai buvo kalba (klb), bus skaičiuojamos atitinkamai:

$$P(triuk \mid x) = \frac{P(triuk)\,p(E \mid triuk)\,p(GSL \mid triuk)\,p(NKD \mid triuk)}{P(x)} \tag{7}$$

$$P(klb \mid x) = \frac{P(klb)\,p(E \mid klb)\,p(GSL \mid klb)\,p(NKD \mid klb)}{P(x)} \tag{8}$$

$$P(x) = P(triuk)\,p(E \mid triuk)\,p(GSL \mid triuk)\,p(NKD \mid triuk) + P(klb)\,p(E \mid klb)\,p(GSL \mid klb)\,p(NKD \mid klb) \tag{9}$$

Formulėse p(E|triuk) – sąlyginė tikimybė, kad triukšmo atkarpa signale gali „pasireikšti“ trumpalaikės energijos reikšme E; p(GSL|triuk) – sąlyginė tikimybė, kad triukšmo atkarpa signale gali „pasireikšti“ garso slėgio lygio reikšme GSL; p(NKD|triuk) – sąlyginė tikimybė, kad triukšmo atkarpa signale gali „pasireikšti“ nulio kirtimų dažinio reikšme NKD. P(triuk) ir P(klb) – atitinkamos klasės tikimybė (turimų apmokymo duomenų dalis), P(triuk|x) – tikimybė, kad analizuojamas objektas priklauso triukšmo klasei, P(klb|x) – tikimybė, jog objektas priklauso kalbos klasei.

Daroma prielaida, jog duomenys bus pasiskirstę pagal normalųjį skirstinį:

$$p(x) = \frac{1}{\sqrt{2\pi}\ \sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{10}$$

Formulėje μ – aritmetinis vidurkis, σ – standartinis nuokrypis, x – kintamasis, kuriam atliekamas skaičiavimas [6]. Aritmetinis vidurkis ir standartinis nuokrypis bus apskaičiuojami iš atitinkamos klasės surinktų apmokymų duomenų.

Kuriai klasei priklauso objektas, bus sprendžiama pagal tai, kurio įvykio tikimybė bus didesnė, t.y. jeigu

$$P(triuk \mid x) > P(klb \mid x) \tag{11}$$

objektas bus priskiriamas triukšmo klasei, o jeigu

$$P(triuk \mid x) < P(klb \mid x) \tag{12}$$

priskiriamas kalbos klasei.

IV. BALSO KOMANDŲ APTIKIMO ALGORITMAS

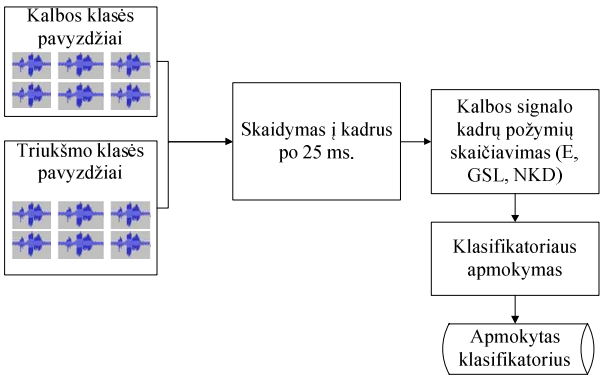
Balso komandų aptikimo algoritmas susideda iš dviejų pagrindinių dalių:

1) Klasifikatoriaus mokymas – surenkami kiekvienos klasės pavyzdiniai duomenys ir jais apmokomas klasifikatorius.
2) Balso komandų aptikimas kalbos signale – turint apmokytą klasifikatorių – nežinomas kalbos signalas yra klasifikuojamas kaip triukšmas arba kaip kalba.

Vykdant klasifikatoriaus apmokymą, arba balso komandų aptikimą, yra atliekami tam tikri bendri kalbos signalo apdorojimo veiksmai:

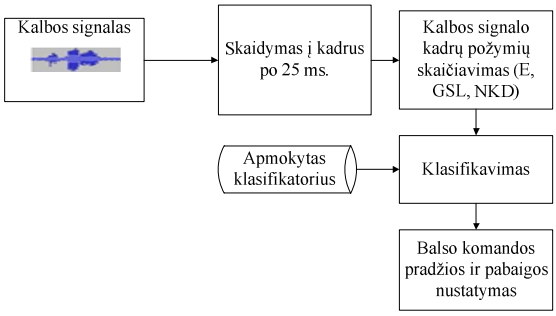
1) Kalbos signalo skaidymas į kadrus.
2) Požymių skaičiavimas kiekviename kalbos signalo kadre.

Klasifikatoriaus apmokymo algoritmo blokinė diagrama:



1 pav. Klasifikatoriaus apmokymo algoritmas

Balso komandų aptikimo algoritmo veikimo schema.

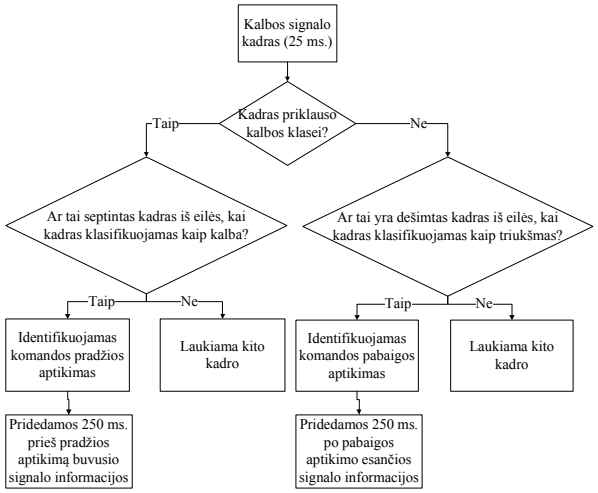


2 pav. Balso komandų aptikimo kalbos signale algoritmas

Antrame paveikslėlyje, „Balso komandos pradžios ir pabaigos nustatymas“ dalyje, yra naudojami trys papildomi apribojimai.

1) Komandos pradžia aptinkama tik tada, kai iš eilės einantiems septyniems kalbos signalo kadrams klasifikatorius grąžina reikšmę, jog tai kalba. Tai reiškia, jog bent 175 ms. turi vykti kalba, kad ji būtų aptikta.
2) Komandos pabaiga aptinkama tik tada, kai dešimčiai kalbos signalo kadrų iš eilės klasifikatorius grąžina reikšmę, jog tai triukšmas. Tai reiškia, jei yra nutylama bent 250 ms. – identifikuojama komandos pabaiga.
3) Identifikavus šnekos pradžią ir pabaigą, visada komandos pradžioje ir pabaigoje papildomai pridedama 250 ms. kalbos signalo informacijos.

Šie apribojimai leidžia sumažinti klaidingai aptiktų komandų skaičių.



3 pav. Apribojimo taisyklių detalus atvaizdavimas

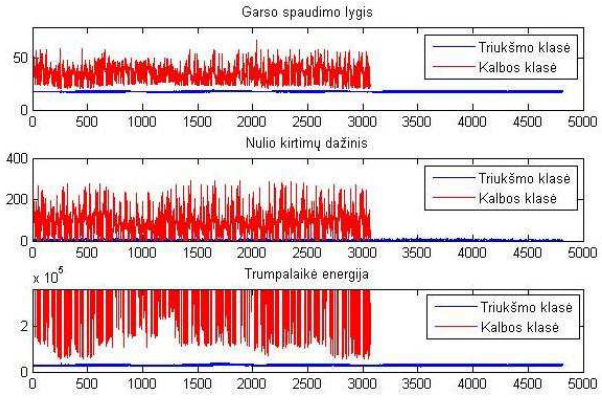
Laiko intervalai, naudojami šiose taisyklėse, yra gauti eksperimentiškai. Tyrime dalyvavo penki žmonės, kurie po penkis kartus diktavo komandas: „ir“, „į“, „i“, „e“, „t“, „b“, „p“, „c“, „o“, „nulis“, „saulė“, „mėnulis“, „šaukštas“, „du“, „Onutė“, „septyni“. Tokios komandos pasirinktos dėl to, jog jos yra skirtingų ilgių („į“, „ir“, „nulis“, „šaukštas“ ir pan.), prasideda arba baigiasi garsais, kurie generuoja skirtingo ilgumo ir stiprumo energijos reikšmes (sprogstamieji priebalsiai, frikatyviniai priebalsiais ir pan.), tariant kai kurias komandas jų viduryje padaroma maža pauzė („sep“ – pauzė – „tyni“, „šauk“ – pauzė – „štas“). Šios pasirinktų komandų savybės leidžia sudaryti taisykles, kurios taikomos šiame balso komandų aptikimo metode.

Išanalizavus gautus rezultatus nustatyta – jeigu kalba identifikuojama mažesniame negu 175 ms. intervale, aptinkamas pašalinis triukšmas, nes trumpiausių diktuotų komandų ilgis yra apie 180–200 ms. Kartais diktuojant komandas žodyje padaromos pauzės, kurių ilgis svyruoja nuo 100 ms. iki 200 ms. Tad norint, kad komanda būtų aptinkama, nebūtų suskaidoma, yra įvestas apribojimas: jei 200 ms. atkarpoje neaptinkama kalbos, tik tada identifikuojama, jog tai yra žodžio pabaiga. Komandos pradžioje ir pabaigoje papildomai pridedama 250 ms. informacijos, nes tokių garsų kaip frikatyvinis priebalsis „s“ arba balsis „i“ aptikimas yra sudėtingas, kadangi laiko kitimo skalėje jie savo energiją didina arba mažina palaipsniui. Taip pat naudojant tokį apribojimą yra sumažinamas prarandamų garsų skaičius ir padidinamas aptiktų komandų atpažinimo procentas.

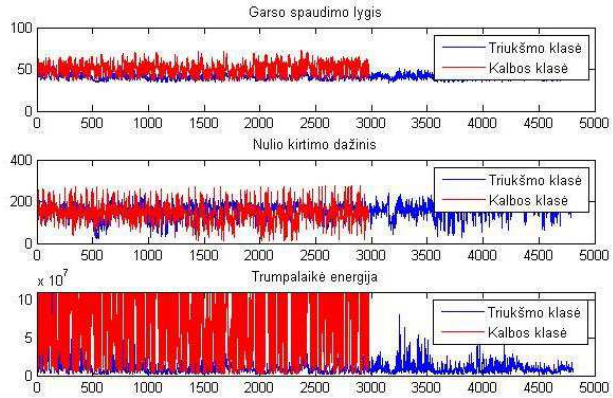
V. METODO TYRIMAS

Norint ištirti metodą, pirmiausia buvo sudaryta viena aibė komandų – klasifikatoriaus kalbos klasės apmokymo aibė: „nulis“, „vienas“, „du“, „trys“, „keturi“, „penki“, „šeši“, „septyni“, „aštuoni“, „devyni“. Triukšmo klasės apmokymui buvo įrašytas aplinkos triukšmas. Metodo tyrimas vyko trijose aplinkose: su mažu triukšmu 10 – 30 dB, vidutiniu triukšmu 30 – 50 dB ir dideliu triukšmu 50 – 70 dB. Šiame tyrime dalyvavo penki asmenys: trys vyrai ir dvi moterys. Tyrimo dalyviai kiekvienoje aplinkoje po tris kartus diktavo kalbos klasės apmokymo aibės komandas. Taip pat kiekvienoje aplinkoje buvo padarytas dviejų minučių trukmės triukšmo įrašas, skirtas triukšmo klasės apmokymui. Balso įrašymui naudotas mikrofonas su ausinėmis Creative HS-330.

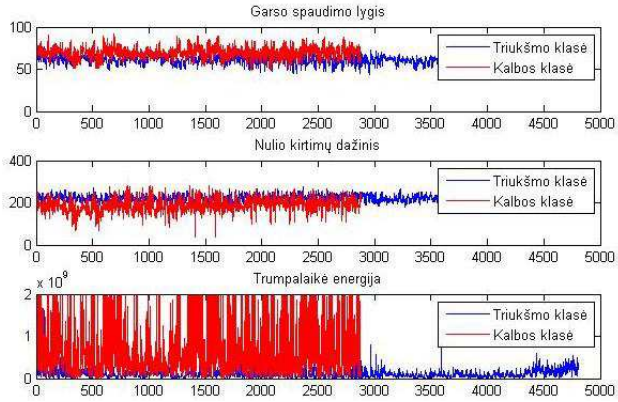
Kiekvienoje aplinkoje atlikus klasifikatoriaus apmokymus, gauti žemiau pateikti rezultatai. Grafikų abscisių ašyse vaizduojami kalbos signalo kadrų eilės numeriai, ordinačių ašyje – analizuojamo kalbos signalo požymio (pateikta kiekvieno grafiko viršuje) dydis.



4 pav. Klasifikatoriaus apmokymo duomenys esant mažam triukšmo lygiui

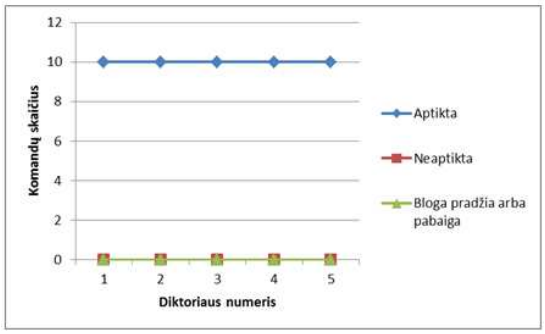


5 pav. Klasifikatoriaus apmokymo duomenys esant vidutiniam triukšmo lygiui

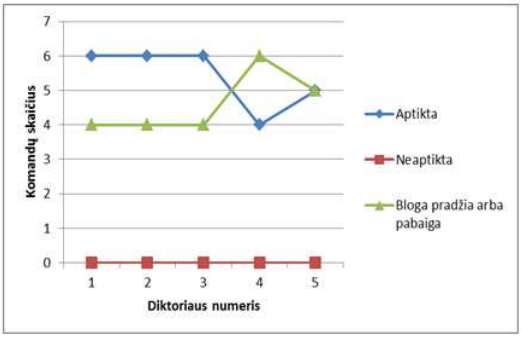


6 pav. Klasifikatoriaus apmokymo duomenys esant dideliam triukšmo lygiui

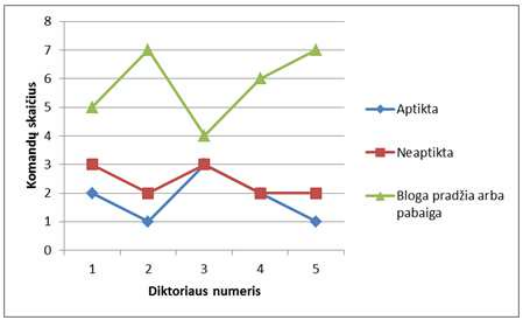
Metodo efektyvumui ir tikslumui ištirti buvo sudaryta kita penkių žmonių grupė. Kiekvienas tiriamasis atsitiktinai diktavo po dešimt bet kokių komandų kiekvienoje iš tiriamų aplinkų. Pagal gautus rezultatus buvo nubraižyti metodo daromų klaidų grafikai.



7 pav. Metodo efektyvumas esant mažam triukšmo lygiui

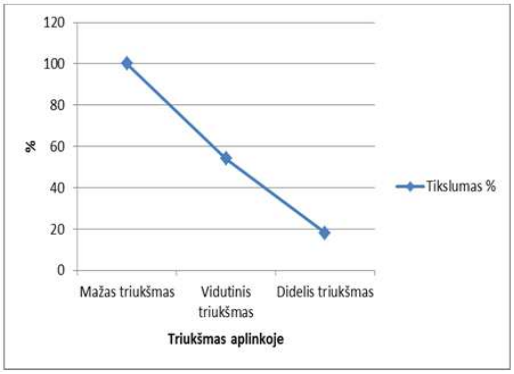


8 pav. Metodo efektyvumas esant vidutiniam triukšmo lygiui



9 pav. Metodo efektyvumas esant dideliam triukšmo lygiui

Remiantis gautais rezultatais galime teigti, jog mažai triukšmingoje aplinkoje metodas veikia labai tiksliai, o vidutinio triukšmingumo aplinkoje metodo veikimas yra patenkinamas, nes visos komandos yra aptiktos, tačiau kai kurių komandų pradžia arba pabaiga yra identifikuotos netiksliai. Triukšmingoje aplinkoje metodo veikimas yra netikslus ir dėl didelio klaidų skaičiaus yra netinkamas kalbos atpažinimo sistemai.



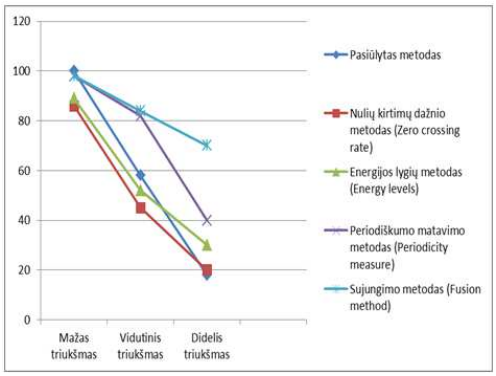
10 pav. Metodo tikslumas pagal veikimo aplinkas

Kaip matome grafike, kuo didesnis triukšmas, tuo metodo tikslumas mažesnis.

Šiuo metu pasaulyje yra pasiūlyta nemažai balso komandų aptikimo metodų. Daugumoje jų naudojami dažnių srities požymiai (melo dažnių skalės kepstriniai koeficientai ir pan.), kurie yra labiau atsparūs triukšmui negu laiko srities požymiai.

Tokių balso komandų aptikimo metodų tikslumas triukšmingose aplinkose yra geresnis [6] [7] [8].

Pagal [9] literatūros šaltinyje pateiktus balso komandų aptikimo algoritmus, sudarytas jų ir pasiūlyto metodo tikslumo palyginimo grafikas.



11 pav. Balso komandų aptikimo kalbos signale metodų tikslumo palyginimas

Kaip matyti iš grafiko, pasiūlyto metodo tikslumas mažo ir vidutinio triukšmingumo aplinkose lenkia kitus metodus, kurie naudojasi laiko skalės parametrais. Didelio triukšmingumo aplinkoje pasiūlyto metodo tikslumas yra mažiausias.

VI. IŠVADOS

Atlikus tyrimą nustatyta, jog metodas yra neefektyvus esant dideliam triukšmui – tiriant metodą gauta 18 % klaidų. Mažo ir vidutinio triukšmingumo aplinkose pasiektas 79 % tikslumas, todėl esant panašioms sąlygoms metodą galima naudoti praktikoje.

Kaip šio metodo privalumus galima išskirti: nesudėtingi ir daug resursų nereikalaujantys skaičiavimai, greita realizacija, reikalingas mažas apmokymo duomenų kiekis. Kadangi metodas naudoja signalo laiko skalės požymius, kurie nėra atsparūs triukšmui, jo efektyvumas triukšmingose aplinkose yra mažas.

VII. ATEITIES DARBAI

Šiame darbe kalbos signalo požymiai yra skaičiuojami naudojant tik signalo laiko skalės požymius, todėl ateityje planuojama atlikti tyrimą su signalo dažnių skalės požymiais, atlikti geriausių požymių kombinacijų paiešką ir rasti tokius kalbos signalo požymius ar jų rinkinius, kurie padidintų metodo tikslumą labai triukšmingose aplinkose.

LITERATŪROS SĄRAŠAS

[1] T. Rasymas, „Balso komandų aptikimas triukšmingame kalbos signale, skirtas automatinei kalbos atpažinimo sistemai", Informacinės technologijos 2012, Nr. 4(2012), ISSN 2029-249X, p. 19–22.

[2] S. Gokhum Tanyers, H. Ozer, „Voice Activity Detection in Nonstationary Noise", IEEE Transactions on Speech and Audio Processing, Vol. 8, No. 4, July 2000, p. 478–482.

[3] A. U. Khan, L. P. Bhaiya, S. K. Banchhor, „Hindi Speaking Person Identification Using Zero Crossing Rate", International Journal of Soft Computing and Engineering, ISSN: 2231-2307, Vol. 2, Issue 3, July 2012.

[4] S. K. Banchhor, A. Khan, „Musical Instrument Recognition using Zero Crossing Rate and Short-time Energy", International Journal of Applied Information Systems, ISSN: 2249-0868, Vol. 1, No. 3, 2012.

[5] S. W. Smith, „The Scientist and Engineer's Guide to Digital Signal Processing", Second Edition, ISBN 0-9660176-4-1, California Technical Publishing, USA, 1999.

[6] A. Martin, D. Charlet, L. Mauuary, „Robust Speech/Non-Speech Detection Using LDA to MFCC", France Telecom R&D, 2010.

[7] J. Ramirez, P. Yelamos, J. M. Gorriz, J. C. Segura, „SVM-based speech endpoint detection using contectual speech features", Electronics Letters, 2006, Vol. 42, No. 7.

[8] J. Ramirez, J. M. Gorriz, J. C. Segura, „Voice Activity Detection. Fundamentals and Speech Recognition System Robustness", Robust Speech Recognition and Understending, ISBN 987-3-90213-08-0, inTech, Austria, 2007.

[9] S. Gokhum Tanyer, H. Ozer, „Voice Activity Detection in Nonstationary Noice", IEEE Transactions on Speech and Audio Processing, 2000, Vol. 8, No. 4, p. 478– 482.

[10] S. Theoderidis, K. Koutroumbas, „Pattern Recognition, 4 edition", ISBN 978-1-59749-272-0, Elsevier Inc., USA, 2009.

Augmented Reality in Radiofrequency Ablation of the Liver Tumours

Francesco Ricciardi
Information Engineering Department
Polytechnic University of Marche
Ancona, Italy

Cosimo Luigi Manes, Lucio Tommaso De Paolis
Department of Innovation Engineering
University of Salento
Lecce, Italy

Abstract. The advantages of the Minimally Invasive Surgery are evident for the patients, but these techniques have some limitations for the surgeons. In medicine, the Augmented Reality (AR) technology allows surgeons to have a sort of ñX-rayò vision of the patientôs body and can help them during the surgical procedures. In this paper we present two applications of Augmented Reality that could be used as support for a more accurate preoperative surgical planning and also for an image-guided surgery. The AR can support the surgeon during the needle insertion for the Radiofrequency Ablation of the liver tumours in order to guide the needle and to have a precise placement of the instrument within the lesion. The augmented visualization can avoid as much as possible to destroy healthy cells of the liver.

Keywords: Augmented Reality, medical images, minimally invasive surgery, RF ablation.

I. INTRODUCTION

One trend in surgery is the transition from open procedures to minimally invasive interventions, where visual feedback to the surgeon is only possible through the laparoscope camera and direct palpation of organs is not possible. Minimally Invasive Surgery (MIS), such as laparoscopy or endoscopy, has changed the way to practice the surgery and, as a promising technique, the use of this surgical approach is nowadays widely accepted and adopted as an alternative to classical procedures.

These techniques offer the possibility to surgeons of reaching the patient's internal anatomy in a less invasive way and causing only a minimal trauma. The diseased area is reached by means of small incisions in the body and specific instruments and a camera are inserted through these ports; during the operation a monitor shows what is going on inside the body.

Shorter hospitalizations, faster bowel function return, fewer wound-related complications and a more rapid return to normal activities have contributed to accept these surgical procedures and, if the advantages of this surgical method are evident on the patients, these techniques involve some limitations for the surgeons. In particular, the imagery is in 2D and the surgeon needs to develop new skills and dexterity in order to estimate the distance from the anatomical structures and work in a very limited workspace.

The acquisition of medical images (CT or MRI) associated to the latest medical image processing provides an accurate knowledge of the patient's anatomy and pathologies and could lead to an improvement in patient care by guiding of the surgeons during the surgical procedure.

The emerging Augmented Reality (AR) technology has the potential to bring the direct visualization advantages of open surgery back to minimally invasive surgery and can increase the physician's view with information gathered from the patient's medical images.

AR technology refers to a perception of a physical real environment whose elements are merged with virtual computer-generated objects in order to create a mixed reality. In addition to a mixture of real and virtual information, an AR application has to run in real time and its virtual objects have to be aligned (registered) with real world structures. Both of these requirements guarantee that the dynamics of real world environments remain after virtual data has been added [1].

In medicine the AR technology makes it possible to overlay virtual medical images of the organs on the real patient and it allows the surgeon to have a sort of "X-ray vision" of the patient's internal anatomy. In order to register the data and fuse virtual and real imagery in real time, special devices are used in the AR platform.

An advantage of using AR in surgery could be found in a better spatial perception and in the duration of the surgical procedure that could be shorter than in conventional way.

The aim of this paper is to present an AR system that could be used as support for a more accurate surgical preoperative planning and also for an image-guided surgery. The application can support the surgeon during the needle insertion in the radiofrequency ablation of the liver tumours.

II. PREVIOUS WORKS

Motivated by the benefits that MIS can bring to patients, many research groups are now focusing on the development of systems in order to assist the surgeons during the minimally invasive surgical procedures.

Furtado and Gersak [1] present some examples of how AR can be used to overcome the difficulties inherent to MIS in the cardiac surgery.

Samset et al. [3] present some decision support tools based on concepts in visualization, robotics and haptics and provide tailored solutions for a range of clinical applications.

Bichlmeier et al. [4] focus on the problem of misleading perception of depth and spatial layout in medical AR and present a new method for medical in-situ visualization.

Navab et al. [5], [6], [7] present a new solution for using 3D virtual data in many AR medical applications and introduce the concept of a laparoscopic virtual mirror, a virtual reflection plane within the live laparoscopic video.

De Paolis et al. [8] present an AR system that can guide the surgeon in the operating phase in order to prevent erroneous disruption of some organs during surgical procedures. They provide distance information between the surgical tool and the organs and they use a sliding window in order to obtain a more realistic impression that the virtual organs are inside the patient's body.

Nicolau et al. [9] present a real-time superimposition of virtual models over the patient's abdomen in order to have a three dimensional view of the internal anatomy. The authors have used the developed system in an operating room and to reduce the influence of the liver breathing motion they have tried to simulate the patient's breathing cycle.

LiverPlanner [10], [11] is a virtual liver surgery planning system developed at Graz University of Technology that combines image analysis and computer graphics in order to simplify the clinical process of planning liver tumor resections. The treatment planning stage enables the surgeon to elaborate a detailed strategy for the surgical intervention and the outcome of pre-operative planning can then be used directly for the surgical intervention.

Maier-Hein et al. [12] present a system developed for computer-assisted needle placement that uses a set of fiducial needles to compensate for organ motion in real time; the purpose of this study was to assess the accuracy of the system in vivo.

Stüdeli et al. [13] present a system that provides surgeon, during placement and insertion of RFA needle, with information from pre-operative CT images and real-time tracking data.

III. THE USED TECHNOLOGIES

A reconstruction of the 3D model of the anatomical and pathological structures of the patient is required in order to improve the standard slice view.

An efficient 3D reconstruction of the patient's organs is generated by applying some segmentation and classification algorithms to medical images (CT or MRI) of the patient. The grey levels in the medical images are replaced by colours that are associated to the different organs [14].

Nowadays there are different software used in medicine for the visualization and the analysis of scientific images and the 3D modelling of human organs; Mimics [15], 3D Slicer [16], and OsiriX [17] play an important role among these tools.

In our application we have used 3D Slicer for the building of the 3D model of the patient's organ; 3D Slicer is a multi-platform open-source software package for visualization and image analysis and in Fig. 1 is shown a 3D model built from CT images of the patient with the tumours in green.

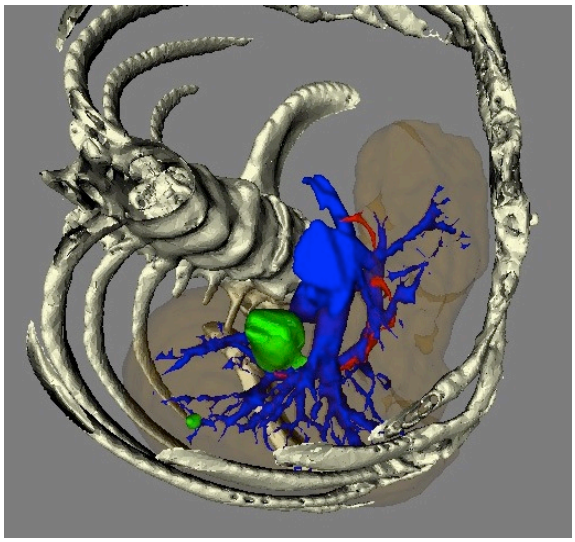


Fig. 1. 3D model from the CT dataset

For the visualization the Image-Guided Software Toolkit (IGSTK), a set of high-level components integrated with low-level open source software libraries and application programming interfaces, has been used [18].

In the developed AR platform it is necessary to use an optical tracker that is able to detect some intentionally introduced retro-reflective spheres (fiducials) placed on the surgical tools and provide within a defined coordinate system the real-time spatial measurements of the location and orientation of the surgical instruments used during the surgical procedure.

The tracking system is also used in order to permit the overlapping of the virtual organs on the real ones providing an augmented visualization of the scene. This registration phase is one of the most delicate step in an AR system for surgery.

In the first prototype of the system we have used the Polaris Vicra optical tracker [19]; in the second protopype, developed after the first test in operating theater, we used the Bonita Vicon tracking system [20].

A specific library has been developed and tested for interfacing with the Bonita Vicon tracker.

At the base of the functioning of IGSTK there is the use of a state machine that allows to increase the safety and robustness of the toolkit [21].

The use of a state machine, in fact, allows to limit and

control the possible behaviors of the application in order to permit that it is always in a valid state already planned in the design phase; in this way a reproducible and deterministic behavior is guaranteed.

The separation between public and private interface of a component IGSTK is shown in figure 2.

When a user send a request for an action through a call of a public method, this request is translated into an input for the state machine that, taking into account the current state and the sent input, changes its state as expected in the design phase. In any case, the class will always be in a valid state where each type of behavior has been programmed.

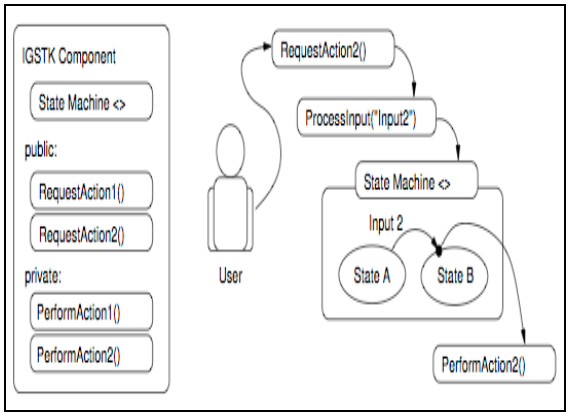


Fig. 2. Working modality of a component IGSTK

The component IGSTK Tracker communicates with the tracker to obtain the transformation relative to each tool present in the acquisition volume. Then this information is transmitted to other components IGSTK who request it.

The two classes of IGSTK that provide for the management of the trackers are:

- igstk :: Trackers that serves to manage the status information of the tracking system;
- igstk :: TrackerTool that manages the information associated with each tool present on the scene.

The use of the class igstk :: Communication for the management of communications allows to prevent blockage of the application in the case of any lack of connection with the tracker; in this case an error event is generated (non-blocking mode). The acquired data are stored into a memory buffer.

The class igstk :: PulseGenerator generates the clock used by the application in order to synchronize all events. In this way, reading and updating of the data are asynchronous operations allowing high standards of safety and performance.

To calibrate the instrument used in the surgical application and obtain the transformation elapsing between the tool and the point of interest has been implemented the PivotCalibration algorithm that permits to store the position and orientation of the tool just holding fixed the position of the instrument tip and moving the tool in the space [22].

IV. AR FOR THE RF ABLATION OF THE LIVER TUMOUR

Hepatic cancer is one of the most common solid cancers in the world. Hepatocellular carcinoma (HCC) is the most common primary hepatic cancer. The liver is often the site of metastatic disease, particularly in patients with colorectal adenocarcinoma.

The use of chemotherapy for malignant form of liver cancer rarely led to good results in long-term survival rate. Today surgery is the best approach to avoid the death of the patient and the reversion of hepatic cancer. Unfortunately only from 5 to 15 per cent of HCC or hepatic metastasis patients diagnosed again undergo a potentially curative resection of the liver cancer [23].

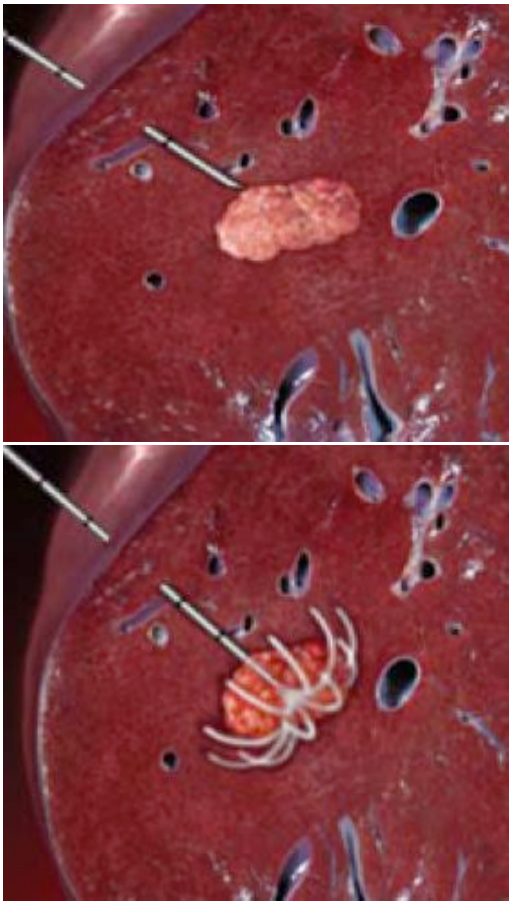


Fig. 3. The effect of the RFA technique at the liver tumour

Patients with confined disease of the liver could not be candidates to resection because of multifocal disease, proximity of tumor to key vascular or biliary structures that

preclude a margin-negative resection potentially unfavourable in case of presence of multiple liver metastases. Very often the tumor is also associated to a pre-existent cirrhosis that can reduce resection margins.

Liver transplant is the only radical therapy that eliminates the risk of recurrence but it can't be always used. So, since most of patients with primary or malignancies confined metastatic at the liver are not candidates for surgical resection, we should use different approaches to control and potentially cure liver diseases.

The Liver Radiofrequency Ablation (RFA) is a minimally invasive treatment for cancer used since 1980's and consists in the placement of a needle inside the liver parenchyma in order to reach the centre of the tumor lesion. The aim of this image-guided technique is to heat and destroy the cancer cells.

When the lesion is reached, an array of electrodes is extracted from the tip of the needle and it is expanded in the tumor tissue. From these electrodes is injected in the tumor tissue a radiofrequency current that causes tumor cell necrosis for hyperthermia (the local temperature is higher than 60 °C and cancer cells are more sensitive to heat than normal cells).

In Figure 3 is shown the effect of the RFA technique at the liver tumour.

One problem in using radiofrequency tumor ablation technique is the correct placement of the needle that should reach the tumor lesion.

Today surgeons use ultrasound, CT or MNR acquired during the needle placement in order to correctly direct the needle to the tumor. The use of these two-dimensional images makes the procedure very difficult and requires sometimes more than one insertion.

In addition, the surgeon, in order to destroy all tumoral cells, applies the RFA on an extended area of the liver, so even a large number of healthy cells are also destroyed. This practice can cause to the patient a number of other different consequences.

There is a need of enhancing the prognosis by a better control of the area interested by the RFA and the accuracy of the needle insertion in order to avoid some areas (important vessels) during the intervention.

A sort of guidance of the needle in tumour ablation procedures can be obtained using Augmented Reality technology. With the superimposition of the virtual models of the patient's anatomy (liver, cancer, etc) exactly where are the real ones, it is possible to make the needle placement less difficult.

The purpose of this AR application is to provide a guidance system that can help the surgeon during the needle insertion in liver RFA. The position and orientation of the real tool is tracked using some reflective spheres detected by means of an optical tracker and used in IGSTK to set the position and the orientation of the virtual tool.

To achieve a correct augmentation it is necessary to have a correspondence between the virtual organs and the real ones. This task is very difficult in an image-guided surgery application because a very small error in the registration phase can cause a serious consequence on the patient.

The registration process is carried out just before starting the surgical procedure and the applied method is based on the placement of fiducial points on the patient's body before the CT scanning and the detection of these in the 3D model built from the acquired medical images. The optical tracker by means of a tool placed on the patient's body detects some possible movements of the patient over the operating table.

The application is provided of an user interface designed to be simple and functional at the same time. In the left side of that interface we have put the application control bottons; in the right-top window we show the 3D model and the augmented reality scene and in the right-bottom the three smaller windows where are placed the axial, coronal and sagittal views of CT dataset.

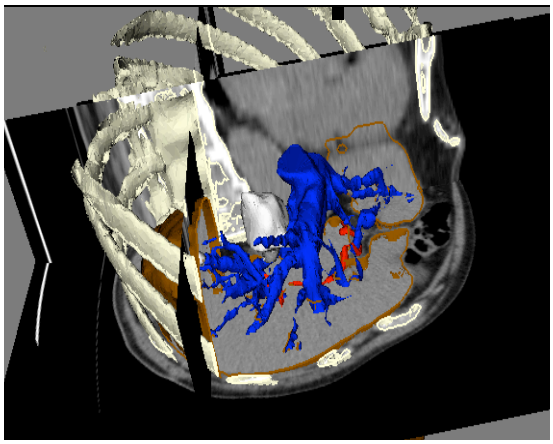


Fig. 4. Clipping visualization applied to the liver and thoracic cage

There is the possibility to apply the clipping modality that permits the surgeon to dissect the model and to study its internal structure changing the opacity of the organs (Figure 4). The dissection could be made along the three principal axes.

The application features described till now are part of what we consider the pre-operative planning task. During this task the surgeon can use the application to study the pathology in a more simple and natural way than that provided by simple CT slice visualization.

For the navigation and augmentation task the surgeon needs to use the optical tracker and to carry out the registration task. When the registration process is complete, the virtual tool is shown in 3D view and is coupled with the real one so it follows the movements that real tool makes.

In Figure 5 is shown the augmented visualization of the 3D virtual model over the patient's body (a dummy). This visualization can guide the surgeon during the needle insertion in the radiofrequency ablation of the liver tumour.

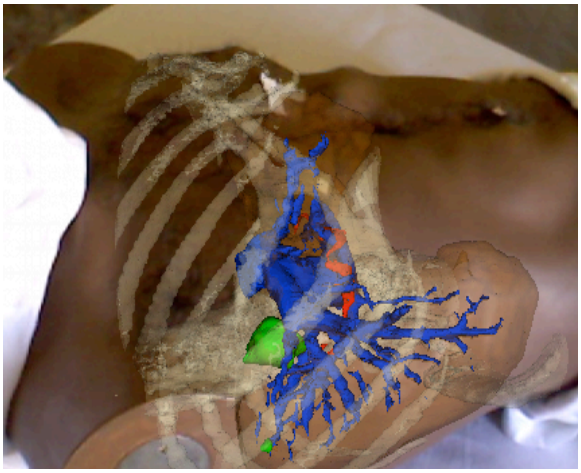


Fig. 5. Augmented visualization of the 3D virtual model over the patient's body (a dummy)

V. CONCLUSIONS AND FUTURE WORK

In this paper we present a guidance system for needle placement in radiofrequency ablation of the liver tumour.

The application has been tested previously on a dummy and afterwards, in order to start the validation of the system, a first test has been carried out in the operating room during an open surgery procedure for the liver tumour resection.

Operating rooms are overloaded of systems and devices and an efficient use of space is mandatory. In this first test the correct position of the devices close to the operating table has been chosen and a precise definition of the more appropriate fiducial points used for the registration phase has been decided.

Figure 6 is shown a phase of the first test in operating room.

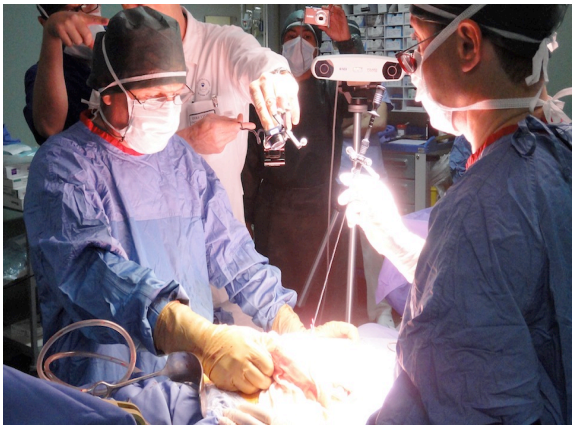


Fig. 6. First test in the operating room

In order to take into account the latest suggestions of the surgeons, a second prototype has been realized using a different optical tracker; a specific library has been developed and tested for interfacing with this tracker.

As future work we are planning to add some improvements of the visualization of the organs taking into account the latest suggestions of the surgeons. The next test will be done on a pig liver in order to measure the precision of the image-guided application.

We are also taking into account the possibility to include in the system the simulation of the virtual model deformations due to the breathing of the patient.

REFERENCES

[1] S. Maad, "Augmented Reality. The Horizon of Virtual And Augmented Reality: The Reality of the Global Digital Age", Intech, January 2010, ISBN 978-953-7619-69-5.

[2] H. Furtado and B. Gersak, "Minimally Invasive Surgery and Augmented Reality. New Technology Frontiers in Minimally Invasive Therapies", 2007, pp. 195-201.

[3] E. Samset, D. Schmalstieg , J. Vander Sloten, A. Freudenthal, J. Declerck, S. Casciaro, Ø. Rideng, and B. Gersak, "Augmented Reality in Surgical Procedures", SPIE Human Vision and Electronic Imaging XIII, 2008.

[4] C. Bichlmeier, F. Wimmer, H. S. Michael, and N. Nassir, "Contextual Anatomic Mimesis: Hybrid In-Situ Visualization Method for Improving Multi-Sensory Depth Perception in Medical Augmented Reality", Proc. Sixth IEEE and ACM Int. Symposium on Mixed and Augmented Reality (ISMAR '07), Nara, Japan, 2007, pp. 129-138.

[5] N. Navab, M. Feuerstein, and C.Bichlmeier, "Laparoscopic Virtual Mirror - New Interaction Paradigm for Monitor Based Augmented Reality", Proc. IEEE Virtual Reality Conf. 2007 (VR 2007), Charlotte, North Carolina, USA, 2007, pp. 10-14.

[6] C. Bichlmeie, S. M. Heining, M. Rustaee, and N. Navab, "Laparoscopic Virtual Mirror for Understanding Vessel Structure: Evaluation Study by Twelve Surgeons", Proc. 6th IEEE International Symposium on Mixed and Augmented Reality, Nara, Japan, 2007, pp. 1-4.

[7] C. Bichlmeie, F. Wimmer, S. M. Heining, and N. Navab, "Contextual Anatomic Mimesis: Hybrid In-Situ Visualization Method for Improving Multi-Sensory Depth Perception in Medical Augmented Reality", IEEE Proc. Int. Symposium on Mixed and Augmented Reality, Nara, Japan, 2007.

[8] L. T. De Paolis, M. Pulimeno, M. Lapresa, A. Perrone, and G. Aloisio, "Advanced Visualization System Based on Distance Measurement for an Accurate Laparoscopy Surgery", Proc. Joint Virtual Reality Conf. of EGVE - ICAT - EuroVR, Lyon, France, 2009, pp. 17-18.

[9] S. Nicolau, A. Garcia, X. Pennec, L. Soler, X. Buy, A. Gangi, . N. Ayache, and J. Marescaux, "An Augmented Reality System for Liver Thermal Ablation: Design and Evaluation on Clinical Cases", Elsevier, 2009.

[10] LiverPlanner, http://liverplanner.icg.tu-graz.ac.at

[11] B. Reitinger, A. Bornik, R. Beichel, G. Werkgartner, and E. Sorantin, "Tools for Augmented Reality Based Liver Resection Planning", Proceedings of the SPIE Medical Imaging 2004: Visualization, Image-Guided Procedures, and Display, pages 88-99, San Diego, February 2004.

[12] L. Maier-Hein, A. Tekbas, A. Seitel, et al., "In Vivo Accuracy Assessment of a Needlebased Navigation System for CT-Guided Radiofrequency Ablation of the Liver", Medical Physics, 2008, Vol. 35, No. 12, 5385-5396, 0094-2405.

[13] T. Stüdeli, D. Kalkofen, P. Risholm, et al., "Visualization Tool for Improved Accuracy in Needle Placement During Percutaneous Radio-

Frequency Ablation of Liver Tumors", Medical Imaging 2008: Visualization, Image-Guided Procedures, and Modeling, Pts 1 and 2, Vol. 6918, B9180-B9180, 0277-786X.

[14] T. S. Yoo, "Insight into Images: Principles and Practice for Segmentation, Registration, and Image Analysis", A K Peters, Ltd., 2004.

[15] Mimics Medical Imaging Software, Materialise Group, http://www.materialise.com/.

[16] 3D Slicer, http://www.slicer.org.

[17] OsiriX Imaging Software, http://www.osirix-viewer.com .

[18] K. Clearya, L. Ibanez, and S. Ranjan, "IGSTK: a Software Toolkit for Image-Guided Surgery Applications", Conference on Computer Aided Radiology and Surgery, Chicago, USA, 2004.

[19] NDI Polaris Vicra, http://www.ndigital.com.

[20] Vicon Bonita, http://www.vicon.com/products/bonita.html .

[21] Kevin Cleary, Patrick Cheng, Andinet Enquobahrie, Ziv Yaniv, "IGSTK: The Book", May 29, 2009.

[22] A. Lorsakul, J. Suthakorn and C. Sinthanayothin, "Point-Cloud-to-Point-Cloud Technique on Tool Calibration for Dental Implant Surgical Path Tracking", Proc. SPIE 6918, Medical Imaging 2008: Visualization, Image-Guided Procedures, and Modeling, March 17, 2008 .

[23] D. Nagorney, J. Van Heerden, D. Ilstrup, et al., "Primary Hepatic Malignancy: Surgical Management and Determinants of Survival", Surgery, 1989, vol. 106, pp. 740-.748.

INFORMACINĖS TECHNOLOGIJOS ● IVUS 2013 ● ISSN 2029-249X ● eISSN 2029-4824

Nusikalstamos veiklos „skaičiavimų debesies“ saugyklose atkūrimo metodas

Nerijus Saikauskas
Kauno technologijos universitetas
Kompiuterių katedra, Studentų g. 50
Kaunas, Lietuva
nerijus.saikauskas@stud.ktu.lt

Santrauka. „Skaičiavimų debesies“ (angl. cloud computing) technologijos sukūrimas suteikė galimybę padidinti kompanijų veiklos efektyvumą, tačiau sukėlė ir naujų problemų, viena kurių – skaitmeninės teismo ekspertizės (angl. digital forensics) atlikimas nutolusioje aplinkoje. Apskritai teigiama, kad jeigu „skaičiavimų debesies“ paslauga nefiksuoja tinkamų audito įrašų, nustatyti įkalčius tampa sunku arba tiesiog neįmanoma. Deja, paprastai šiam tikslui siūlomas funkcionalumas yra gana ribotas arba iš viso neegzistuoja. Šiame straipsnyje siūlome naują metodą-įrankį – žurnalizavimo paramos sistemą (ŽPS), padedančią užfiksuoti ir atkurti vartotojų veiksmus kaip įkalčius „skaičiavimų debesies“ saugyklose. ŽPS įgyvendina kitų autorių pasiūlytą unifikuotą audito įrašų formatą tokio pobūdžio aplinkoms ir sukuria save aprašančių duomenų efektą, kuris, manoma, yra svarbus žingsnis siekiant efektyviai tirti nusikaltimus „skaičiavimų debesies“ saugyklose. Metodo efektyvumas yra įvertinamas apskaičiuojant ŽPS atkuriamų veiksmų skaičiaus priklausomybę nuo virtualių mašinų (angl. virtual machine) kopijų kūrimo dažnumo ir vartotojų atliekamų veiksmų skaičiaus.

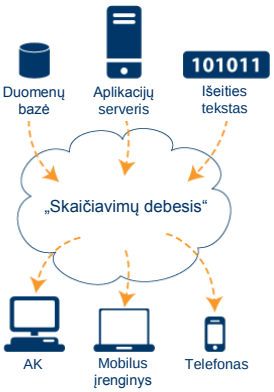
Reikšminiai žodžiai: „skaičiavimų debesis“, virtuali mašina, skaitmeninė teismo ekspertizė, saugykla, veiksmų atkūrimas, paramos sistema.

I. ĮŽANGA

Paskutiniais keleriais metais įvyko reikšmingų IT pasikeitimų, vienas kurių – „skaičiavimų debesies“ (angl. *cloud computing*) sukūrimas. Šios technologijos atsiradimas ėmė keisti būdus, kaip organizacijos įsigyja kompiuterijos. Patikimi tiekėjai, tokie kaip „Microsoft“, „Amazon“, „Google“, „Yahoo“ ir kiti, pasiūlė kompanijoms efektyvias galimybes reikiamų IT išteklių pirkti iš trečiųjų šalių ir tokiu būdu pagrindinį dėmesį sutelkti į savo esminį verslą [1].

„Skaičiavimų debesis“ – tai „modelis, suteikiantis visur pasiekiamą, patogią, „pagal pareikalavimą“ tinklo prieigą prie bendrų, konfigūruojamų kompiuterių išteklių (tinklų, serverių, saugyklų, vartotojo programinės įrangos, paslaugų ir kt.), kurie gali būti sparčiai keičiami minimaliomis valdymo pastangomis ar paslaugos tiekėjo įsikišimu“ [2] (1 pav.).

„Skaičiavimų debesies“ paslauga sparčiai populiarėja. Prognozuojama, kad metinis globalus jos IP srautas išaugs 3,5 karto – nuo 1,2 zetabaitų 2012 m. iki 4,3 zetabaitų (tai sudaro maždaug 4 600 000 000 terabaitų) 2016 m. ir sudarys du trečdalius visų duomenų centrų srauto [4].

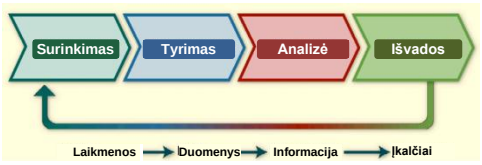


1 pav. „Skaičiavimų debesies“ modelis [3]

Nežiūrint į privalumus, „skaičiavimų debesis“ kelia įvairių abejonių ir rūpesčių. Ypač baiminamasi dėl esamų galimybių paslaugos tiekėjui būti užpultam elektroninės erdvės nusikaltėlių, kurie gali pavogti klientų informaciją. Galima tiek tiesioginė, tiek netiesioginė (pasinaudojant kliento puse) „debesies“ infrastruktūros kompromitacija, egzistuoja įvairios „debesies“ pažeidžiamos vietos, kyla paslaugos tiekėjo šnipinėjimo grėsmė, be to, dėl nutolusių išteklių, sunkiai sprendžiamas audito ir skaitmeninės teismo ekspertizės atlikimo klausimas [5].

Skaitmeninė teismo ekspertizė (angl. *digital forensics*) turi daug apibrėžimų, bet plačiąja prasme tai yra „mokslo panaudojimas siekiant identifikuoti, surinkti, ištirti ir išanalizuoti duomenis išlaikant jų vientisumą ir išsaugant griežtą „arešto grandinę“ (angl. *chain of custody*) (2 pav.). Per pastarąjį dešimtmetį labai išaugo nusikaltimų, atliekamų elektroninėje erdvėje, skaičius. Todėl, įsikūrė daug kompanijų, kurios siūlo produktų, padedančių teisėsaugos organams nustatyti, kas, ką, kur, kada ir kaip padarė nusikaltimus kompiuterinėmis priemonėmis. Skaitmeninė teismo ekspertizė, kaip sfera, išsivystė būtent tam, kad būtų užtikrintas tinkamas elektroninių įkalčių pristatymas teismui [6].

Straipsnio II skyriuje aprašoma veiksmų kaip įkalčių atkūrimo „skaičiavimų debesyje“ problematika, III skyriuje nusakomas mūsų siūlomas metodas – žurnalizavimo paramos



2 pav. Skaitmeninės teismo ekspertizės procesas

sistema (ŽPS), – kurią pasitelkus įvykus nusikaltimui būtų galima atkurti vartotojų veiksmus „skaičiavimų debesies“ saugyklose pasinaudojant esamomis atvirojo kodo programinėmis priemonėmis, IV skyriuje įvertinamas metodo efektyvumas ir pateikiamos išvados.

II. VEIKSMŲ KAIP ĮKALČIŲ ATKŪRIMO „SKAIČIAVIMŲ DEBESYJE“ PROBLEMATIKA

A. *Skaitmeninė teismo ekspertizė „skaičiavimų debesies“ kontekste*

Kyla daug klausimų, ar „skaičiavimų debesis“ gali būti tinkamai ištirtas skaitmeninės teismo ekspertizės tikslams. Viena vertus, šios technologijos tinkamumas tokioms užduotims nėra pakankamai išanalizuotas. Kita vertus, toks gebėjimas nėra esminis jos reikalavimas – kadangi tarptautinės teisinės bazės šiuo klausimu nėra, tai laikoma prabanga, kurią gali arba negali būti įsidiegę paslaugos tiekėjai savo nuožiūra [7].

Kalbant apie techninius skaitmeninės teismo ekspertizės atlikimo „skaičiavimų debesyje“ aspektus, yra galimi 3 įkalčių šaltiniai:

1) „Skaičiavimų debesies“ dalis (vadinamoji virtuali instancija), kurią naudoja klientas. Yra galimybės „gyvai“ tirti vis dar veikiančias virtualias instancijas [8, 9] arba padaryti jų kopijas tam tikru laiko momentu (angl. *snapshot*) ir išanalizuoti jas atskirai.

2) Kompiuterių tinklas. Skirtingi tinklo OSI[1] „sluoksniai“ galėtų suteikti įvairios naudingos informacijos apie juose veikiančius protokolus, kurie užtikrina duomenų srautų tarp „skaičiavimų debesies“ ir kliento apsikeitimą. Deja, tipiškas „skaičiavimų debesies“ paslaugų tiekėjas nesuteikia jokių tinklo komponentų audito įrašų (angl. *logs*).

3) Kliento sistema. Paprastai tai yra tik naršyklė, veikianti kliento kompiuteryje, kuri leidžia pasiekti visas reikiamas „skaičiavimų debesies“ paslaugas [10].

Vertinant iš skaitmeninės teismo ekspertizės pusės, pagrindinis „skaičiavimų debesies“ privalumas yra centralizuoti duomenys: vienoje vietoje esančią informaciją lengviau kontroliuoti. Kitas teigiamas dalykas – milžiniški kompiuterijos ištekliai, galintys saugoti neribotą tiriamos informacijos kiekį ir daug sparčiau atlikti sudėtingų skaičiavimų reikalaujančius veiksmus, pavyzdžiui, nulaužti slaptažodžius. Be to, „skaičiavimų debesies“ mastas suteikia neribotas galimybes sukurti ir suderinti audito įrašų kūrimo mechanizmus taip, kaip to pageidauja esama kliento situacija, kad ir kokia kompleksiška ji būtų ir reikalautų išteklių [1]. Kaip parodė [11], „skaičiavimų debesies“ paslauga gali tapti nusikaltimo tyrimo įrankiu netgi pati savaime.

Vis dėlto „skaičiavimų debesis“, vertinant jį iš skaitmeninės teismo ekspertizės pozicijų, turi ir esminių trūkumų, kurių pagrindinis – nutolę duomenys, kurių fizinė buvimo vieta nėra tiksliai žinoma [12]. Tiriant tokiomis aplinkybėmis, yra didelė tikimybė prarasti svarbius nusikaltimo artefaktus. Pavyzdžiui, gali nebūti galimybių pasiekti „skaičiavimų debesies“ duomenų centruose esančius registro įrašus (angl. *registry entries*), laikinuosius failus (angl. *temporary files*) ir laikinąją atmintį (angl. *memory*). Jeigu duomenys yra atsisiunčiami iš „skaičiavimų debesies“, gali būti prarasta ir metainformacija.

Yra ir daugybė kitų problemų, susijusių su teismo ekspertizės atlikimu „skaičiavimų debesyje“, tokių kaip įrankių nebuvimas [1], teisiniai aspektai (duomenų centrų išsidėstymas keliose šalyse kelia jurisdikcijos klausimą), įkalčių stabilumas ir kitos, kurias aktyviai siūloma spręsti mokslininkų bendruomenei [13]. Tai yra vis dar besivystanti sritis, ir aktyviai kuriami įvairūs metodai jai patobulinti [1].

B. *Audito įrašai kaip priemonė atkurti veiksmus*

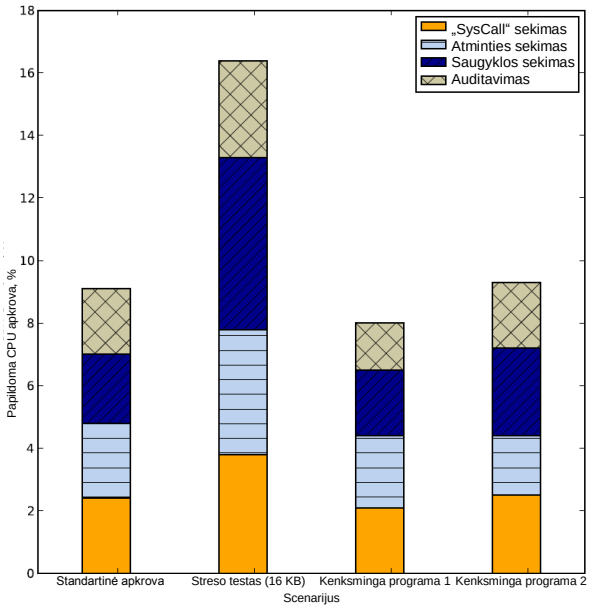
Jeigu „skaičiavimų debesies“ paslauga nefiksuoja tinkamų audito įrašų, nustatyti įkalčius tampa sunku arba tiesiog neįmanoma [7]. Šiam klausimui skiriama daugiausiai dėmesio.

Siūloma pereiti nuo apsaugos, kuri veikia iš sistemos perspektyvos, prie efektyvesnės, kuri veikia atskaitos tašku pasirinkdama failą. Tai reiškia, kad duomenys, saugomi „skaičiavimų debesyje“, turi būti patys save apibūdinantys, kad būtų galima atsekti visą veiklą nuo jų sukūrimo iki sunaikinimo, nepriklausomai nuo aplinkos apribojimų [14]. W. Zhou ir kt. [15] apibūdina pagrindinius iššūkius, susijusius su šios idėjos, dar vadinamos duomenų centriškumu (angl. *data-centric view*), įgyvendinimu, ir siūlo galimą platformą (angl. *framework*) apsikeisti tokia informacija „skaičiavimų debesyje“ tarp skirtingų sistemų, kuri padėtų išsaugoti reikiamus metaduomenis.

Pabrėžiama, kad dabartinių pasyvių „skaičiavimų debesies“ apsaugos metodų nepakanka siekiant atsekti su pačiais save apibūdinančiais duomenimis susijusius veiksmus. Todėl būtina aktyviai daryti audito įrašus, tenkinant tokius reikalavimus:

- *Sekti failus.* Stebimi sisteminiai failų sistemos skaitymo / rašymo kreipiniai (angl. *call*) padėtų atsekti ir susieti virtualias ir fizines failų atminties vietas. (Tyrimais įrodyta, kad dėl šios papildomos funkcijos virtualizacijos platformos procesoriaus apkrova padidėtų iki 20 % (3 pav.) [16].) Failų vietos keitimo sekimas (įskaitant tinklo audito įrašus) sugeneruotų papildomos naudingos informacijos skaitmeninės teismo ekspertizės tyrėjui.
- *Sekti duomenis.* Šie dažnai būna išsidėstę keletame failų sistemos objektų (failų, katalogų), tad integralesnis (bendresnis) audito įrašų darymas padėtų sukaupti reikiamų įrodymų platesniame kontekste.

[1] Atvirųjų sistemų sujungimo tinklu modelis, apibrėžiantis duomenų perdavimą atviraisiais nevienalyčiais tinklais [24].

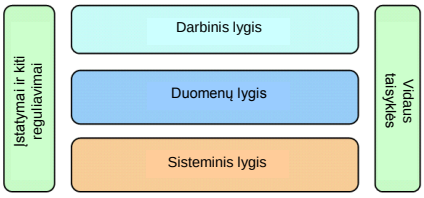


3 pav. Procesoriaus apkrovos padidėjimas papildomai apdorojant sisteminius failų sistemos skaitymo / rašymo kreipinius skaitmeninės teismo ekspertizės tikslams

- *Sekti informaciją.* Kadangi būtent informacija, o ne duomenys ar failai yra svarbiausia organizacijos vertybė, užfiksuoti įrašai, parodantys, kaip failai ir duomenys transformavosi į informaciją, yra neįkainojamos vertės tiriant saugumo incidentus.
- *Sekti informaciją ir duomenų srautus.* Aukšto lygio rizika, tokia kaip sprendimų priėmimo informacija, taip pat turėtų būti kontroliuojama.

Šiuos reikalavimus autoriai išdėstė vadinamojo „TrustCloud“ koncepcijoje, kuri vaizdžiai apibrėžia centralizuotą požiūrį į audito įrašų darymą (4 pav.) [14].

Sunkumų kelia tai, kad nėra oficialaus standarto, kokiu formatu įrašai turėtų būti fiksuojami siekiant vėliau efektyviai ištirti nusikaltimą „skaičiavimų debesyje“. R. Marty [17] siūlo priimti bendras gaires, kuriose apibrėžiama konkreti sintaksė. Autorius mano, kad audito įrašais turėtų būti fiksuojamas bent įvykio laikas, susijusi aplikacija, vartotojas, sesijos ID, svarbumo reitingas (angl. *severity*), priežastis ir kategorija, nes tai padėtų atsakyti į klausimus: kada, kas, ką ir kodėl. O įrašų sintaksę siūloma grįsti laukų grupavimu (angl. *key-value pair*), nes tai leistų efektyviai apdoroti audito įrašus papildomomis programinėmis priemonėmis (5 pav.).



4 pav. „TrustCloud“ koncepcija

```
time=2010-05-13 13:03:47.123231PDT,
session_id=08BaswoAAQgAADVDG3IAAAAD,
severity=ERROR,user=pixlcloud_zrlram,
object=customer,action=delete,status=failure,
reason=does not exist
```

5 pav. Siūloma sintaksė audito įrašų „skaičiavimų debesyje“ fiksavimui

„Skaičiavimų debesies“ audito įrašai negalės būti panaudoti teismo procese, jeigu nebus užtikrintas jų integralumas. Ši problema, iškilusi „skaičiavimų debesies“ saugyklų kontekste, yra sprendžiama siūlymais įtraukti trečiąją šalį, kuri, apsikeisdama specialiomis žinutėmis su vartotoju ir „skaičiavimų debesimi“, papildomai išsaugotų ir įrašų privatumą [18]. Tuo užsiimti, manoma, galėtų oficialiai pripažinta audito kompanija (6 pav.), tačiau [20] bei [21] nurodoma ir metodų, kurie užtikrina paskirstytose sistemose (angl. *distributed systems*) esančių įrašų integralumą, neįsikišant pašaliniams.

III. Žurnalizavimo paramos sistema

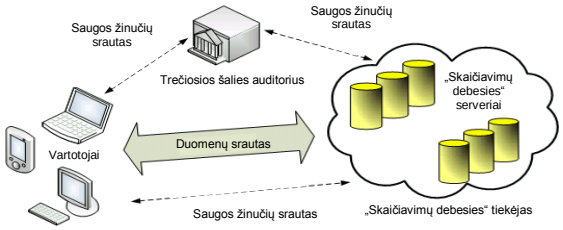
Siūlome naują metodą-įrankį – *žurnalizavimo paramos sistemą*, kuri, pasinaudodama atvirojo kodo priemonėmis, padeda pirmiausia užfiksuoti, o vėliau – atkurti vartotojų veiksmus kaip įkalčius „skaičiavimų debesies“ saugyklose. ŽPS skirta fiksuoti nustatytiems duomenims, nepriklausomai nuo to, ar egzistuoja koks nors „skaičiavimų debesies“ saugyklų paslaugos funkcionalumas šiuo klausimu, ar ne. ŽPS vieta nagrinėjamos problemos kontekste pavaizduota 7 pav.

Sistema suprojektuota taip, kad veiktų kaip atskiras komponentas, savarankiška programinė įranga, kurią savo pasirinkimu galėtų įdiegti bet kokia privati kompanija, kuri naudojasi lokaliomis „skaičiavimų debesies“ saugyklomis. Ją funkcionalumo prasme sudaro 2 dalys:

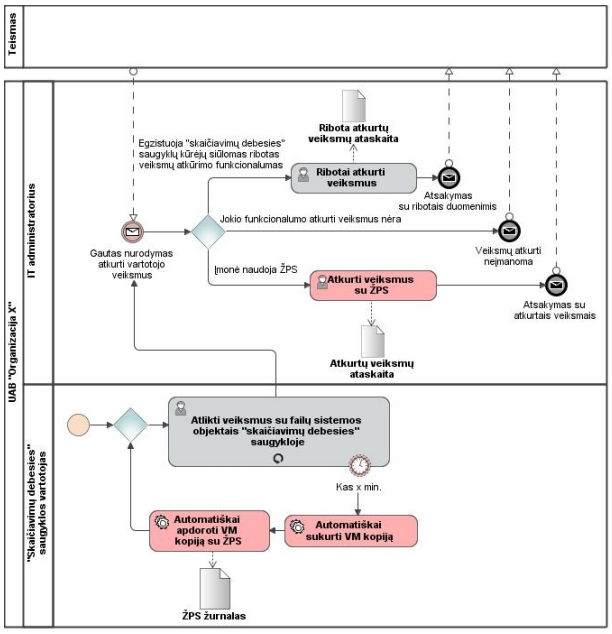
1) Veiksmus fiksuojanti. Apdoroja naujas virtualios mašinos (angl. *virtual machine*) kopijas ir papildo ŽPS audito žurnalą.

2) Veiksmus atkurianti. Atkuria pasirinkto laikotarpio veiksmus.

Tiek veiksmus fiksuojanti, tiek veiksmus atkurianti dalis gali būti įdiegta į bet kokią pasirinktą „Windows“ arba „Linux“ kompiuterinę sistemą pagal patogumą – svarbiausia, kad iš jų būtų galima pasiekti paruoštas VM kopijas ir vėliau ŽPS sukurtą audito žurnalą (norint atkurti veiksmus). Suteikiamas tinkamai paruoštų ir į nustatytą katalogą įdėtų „aktyvių“ virtualių mašinų kopijų (angl. *snapshot*) apdorojimo funkcionalumas, atsiribojant nuo kopijų kūrimo klausimo ir su juo susijusios problematikos. VM kopijos turi tenkinti šiuos reikalavimus:



6 pav. Audito įrašų integralumo užtikrinimas pasinaudojant trečiąja šalimi [19]



7 pav. Žurnalizavimo paramos sistemos vieta problemos kontekste (BPMN modelis)

- Turi būti 2-u failai – virtualaus HDD kopija ir virtualios RAM kopija.
- Tiek HDD, tiek RAM kopija turi būti išsaugota (esant reikalui, konvertuota) vadinamuoju „grynu“ (angl. *raw*) formatu.

Žurnalizavimo paramos sistema yra pagrįsta „Python“ programavimo kalba jungiant dvi nemokamas atvirojo kodo programines priemones: „The Sleuth Kit“ (TSK) [22] ir „The Volatility Framework“ (TVF) [23]. TSK leidžia iš virtualaus HDD nuskaityti šiuos failų sistemos objektų (failų ir katalogų) metaduomenis:

- sukūrimo data;
- redagavimo data;
- ištrynimo data.

TVF suteikia galimybę surasti reikiamus artefaktus operatyviojoje atmintyje. ŽPS atveju iš virtualios RAM kopijos yra nuskaitomi prie VM esamu momentu prisijungę IP adresai. Virtualių mašinų kopijos yra kuriamos nustatytu periodu (žr. IV skyrių), ir tiek TSK, tiek TVF sugeneruoti duomenys išsaugomi ŽPS audito žurnaluose, kurių formatas atitinka [17] reikalavimus ir kuriuose dėl patogumo naudojamas „Unix“ laiko formatas (8, 9 pav.).

8.

```
time=1358303434,object=./data/folder1/,action=created
time=1358303438,object=./data/folder1/file3,action=created
time=1358303441,object=./data/folder1/file3,action=modified
time=1358303444,object=./data/folder1/folder4/,action=created
time=1358303447,object=./data/folder1/file3,action=deleted
```

8 pav. Audito žurnalo „zps-fs.log“ turinys

```
time=1358289391,object=[10.0.2.2,65.7.11.15]
time=1358293119,object=[192.168.14.22]
time=1358299579,object=[10.0.2.15,192.168.0.14,213.15.26.5]
```

9 pav. Audito žurnalo „zps-ip.log“ turinys

Kadangi „The Sleuth Kit“ nuskaito failų sistemos objektų metaduomenis, visais atvejais yra pateikiama galutinė ir užtikrinta informacija apie su objektais atliktus veiksmus. Turint omenyje, kad vartotojams paprastai yra suteikiama prieiga ne prie absoliučiai visų, o tik specialiai tam skirtų failų sistemos išteklių (atskirų skaidinių arba katalogų), žurnalizavimo paramos sistemos darbas yra optimizuojamas konfigūracijoje nurodant atitinkamas failų sistemos dalis, kurių veiksmus reikia fiksuoti. O „The Volatility Framework“ išsaugo visus prie VM prisijungusius tinklo IP adresus, nepriklausomai nuo to, ar jie gali būti susiję su TSK dalies užfiksuotais failų sistemos objektų pokyčiais, ir pateikia tai kaip papildomą informaciją bylos nagrinėtojams. Aptarta ŽPS veikimo schema pateikiama 10 pav.

Veiksmams atkurti ŽPS įrankis nuskaito audito žurnalą, kuriame yra išsaugota ŽPS surinkta informacija. Galima surasti įrašus pagal:

- datą;
- IP adresą;
- failų sistemos objektą;
- veiksmą (sukūrimas, redagavimas, ištrynimas).

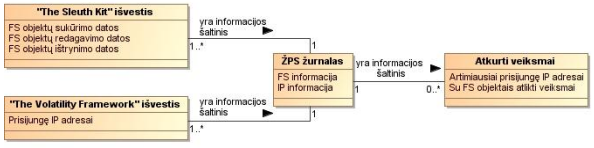
11 pav. pateiktas atkurtų veiksmų, susijusių su failu „file1“, audito įrašų pavyzdys. Išvestyje įrašomi prieš failų sistemos objektų pokyčius arčiausiai užfiksuoti IP adresai. Gauname atkuriamąją informaciją apie objektą, nepriklausomai nuo to, kokiose sistemose ar priemonėse jis buvo naudojamas. Tokiu būdu įgyvendiname [14] ir [15] pasiūlytą save aprašančių duomenų koncepciją, kuri, kaip buvo aptarta II skyriuje, yra svarbus žingsnis siekiant efektyviai tirti nusikaltimus „skaičiavimų debesies“ saugyklose.

Kitame skyriuje pateikiamas žurnalizavimo paramos sistemos efektyvumo įvertinimas.

IV. ŽPS EFEKTYVUMO ĮVERTINIMAS

Siūlomas ŽPS metodas-įrankis buvo tiriamas šiais aspektais:

- Atkurtos informacijos kiekio priklausomybė nuo VM kopijų kūrimo (ir atitinkamai apdorojimo pasitelkiant žurnalizavimo paramos sistemą) dažnumo.



10 pav. ŽPS veikimo schema (koncepcinis duomenų modelis)

```
time=1358225932,object=[10.0.2.15]
time=1358228301,object=./data/file1,action=created
time=1358229629,object=./data/file1,action=modified
time=1358233118,object=[10.0.2.15,192.168.0.147,213.15.26.5]
time=1358234629,object=./data/file1,action=modified
time=1358234660,object=./data/file1,action=deleted
```

11 pav. Atkurti veiksmai, susiję su failų sistemos objektu „file1“

- Atkurtos informacijos kiekio priklausomybė nuo vartotojų veiksmų dažnumo.

Buvo naudojamas kompiuteris „HP nx6125“ (I lent.) su „Lubuntu 12.10“ 64 bitų operacine sistema („Linux“ branduolio versija – „3.5.0-21-generic“), kurioje įdiegta „Oracle VirtualBox 4.1.18“ virtualizacijos platforma. Sukurtos 6-ios virtualios mašinos su „Debian 4.3.5-4“ 32 bitų operacine sistema („Linux“ branduolio versija – „2.6.32-5-686“). Viena buvo naudojama kaip tyrimo objektas – būtent jos kopijos ir buvo kuriamos bei analizuojamos, o likusios buvo panaudotos kaip skirtingų 5-ių vartotojų kompiuterių imitavimas, kuriais buvo jungiamasi prie pagrindinės VM ir atliekami veiksmai su failų sistemos objektais. ŽPS sistema buvo leidžiama iš pagrindinės kompiuterio OS.

I LENTELĖ. TYRIME NAUDOTAS KOMPIUTERIS

Komponentas	Techniniai duomenys
Procesorius	AMD *Turion* 64 *Mobile Technology* ML-30 (1.6-GHz, 1-MB L2 *cache*)
Pagrindinė plokštė	ATI RADEON XPRESS 200M Chipset
RAM	2048-MB 333-MHz DDR SDRAM
HDD	50-GB 5400 rpm
Vaizdo korta	*Integrated* ATI MOBILITY RADEON X300 (128-MB *allocated system memory*)

Siekiant imituoti vartotojų veiksmus realiu laiku, „Python“ programavimo kalba buvo sukurtas scenarijus (angl. *script*):

1) Laukia atsitiktinai sugeneruotą sekundžių skaičių (nustatoma maksimali riba).
2) Prisijungia prie nurodytos virtualios mašinos.
3) Atsitiktinai pasirenka vietą egzistuojančiame failų sistemos medyje.
4) Atsitiktinai pasirenka atlikti vieną iš šių veiksmų:
a) Sukurti naują katalogą.
b) Ištrinti esamą katalogą.
c) Sukurti naują failą.
d) Modifikuoti esamą failą.
e) Ištrinti esamą failą.
5) Laukia (vis dar prisijungęs prie VM) atsitiktinai sugeneruotą sekundžių skaičių (nustatoma maksimali riba).
6) Atsijungia nuo VM.
7) Kartoja viską iš pradžių.

Šis scenarijus paleidžiamas tuo pačiu metu iš 5-ių virtualių mašinų su skirtingais IP adresais ir jungiasi prie tos pačios pagrindinės VM (jos parametrai: 5 GB HDD, 256 MB RAM), kuri bus tiriama. Prisijungimas vyksta pagal SSH protokolą, veiksmai atliekami naudojant standartines „Linux“ komandinės eilutės priemones: „mkdir“, „touch“, „echo“ ir kt. Kad būtų galima palyginti, scenarijus visus faktiškai atliktus veiksmus ir įvykusių prisijungimų laikus išsaugo audito žurnaluose.

Virtualių mašinų kopijos nustatytu periodu buvo kuriamos rankiniu būdu, pasinaudojant virtualizacijos platformos „VirtualBox“ tam skirtu funkcionalumu ir dedamos į pasirinktą katalogą tyrime naudojamame kompiuteryje, iš kurio buvo automatiškai apdorojamos iš anksto paleistos vykdyti žurnalizavimo paramos sistemos.

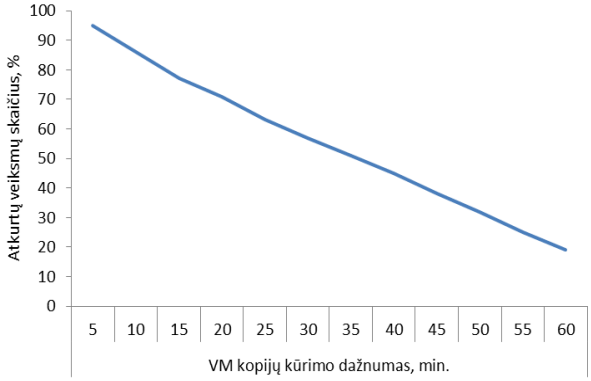
Tiriant atkurtos informacijos kiekio priklausomybę nuo VM kopijų kūrimo dažnumo, kopijos buvo daromos kas 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55 ir 60 min. Maksimalus vartotojų veiksmus imituojančių scenarijų laukimo laikas buvo nustatytas vienodas – po 15 min. Kiekvienas kopijos kūrimo laikas buvo tiriamas kartojant bandymą po 10 kartų. Rezultatai pateikiami 12 pav.

Siekiant nustatyti, kaip atkurtos informacijos kiekis priklauso nuo vartotojų veiksmų kiekio, buvo nustatytas 5 min. VM kopijų kūrimo dažnumas, o vartotojų veiksmus imituojantys scenarijai buvo koreguojami po kiekvienos kopijos sukūrimo iteracijos taip, kad atliekamų veiksmų skaičius didėtų. Tai buvo įgyvendinta keičiant maksimalų laukimo laiką į atitinkamai 15, 13, 10, 9, 8, 7, 6, 5, 3 ir 1 min. Atkurtų veiksmų skaičius buvo nustatomas po kiekvieno atlikto bandymo, rezultatai pateikiami 13 pav.

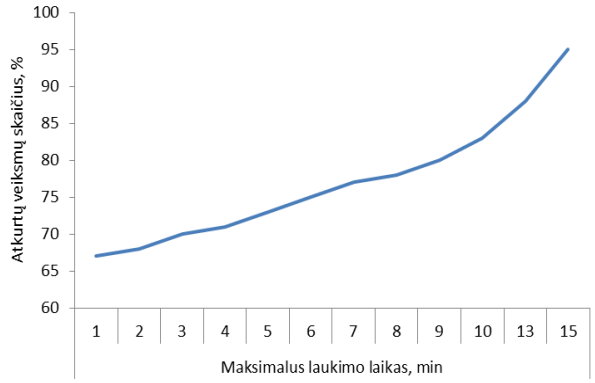
Tyrimo rezultatuose atsispindi tiesinė virtualių mašinų kopijų kūrimo dažnumo ir atkurtos informacijos kiekio priklausomybė. Kopijas kuriant kas 5 min., naudojant žurnalizavimo paramos sistemą pavyksta atkurti net 95 % vartotojų veiksmų. O priklausomybė tarp vartotojų veiksmų ir atkurtos informacijos kiekio yra eksponentinė – net ir labai didinant veiksmų skaičių, ŽPS sugeba atkurti gana daug informacijos (daugiau kaip 65 %; su sąlyga, kad VM kopijos yra kuriamos kas 5 min.).

V. IŠVADOS

Šiame straipsnyje pristatėme naują metodą-įrankį – žurnalizavimo paramos sistemą, padedančią užfiksuoti ir atkurti vartotojų veiksmus kaip įkalčius „skaičiavimų debesies“



12 pav. Atkurtos informacijos kiekio priklausomybė nuo VM kopijų kūrimo dažnumo



13 pav. Atkurtos informacijos kiekio priklausomybė nuo maksimalaus vartotojų laukimo laiko (sąlygojančio vartotojų veiksmų skaičių)

saugyklose. „Python“ programavimo kalba jungianti kelias atvirojo kodo programines priemones skaitmeninei teismo ekspertizei atlikti, ŽPS įgyvendina kitų autorių pasiūlytą unifikuotą audito įrašų formatą tokio pobūdžio aplinkoms ir sukuria save aprašančių duomenų efektą, kuris, manoma, yra svarbus žingsnis sprendžiant veiksmų atkūrimo „skaičiavimų debesies“ saugyklose problemą. Pateiktas metodo efektyvumo įvertinimas atskleidžia atkuriamų veiksmų skaičiaus priklausomybę nuo virtualių mašinų kopijų kūrimo dažnumo ir vartotojų atliekamų veiksmų skaičiaus: analizuojant virtualių mašinų kopijas kas 5 min., žurnalizavimo paramos sistema leidžia atkurti net 65–95 % veiksmų, priklausomai nuo juos sukūrusių vartotojų aktyvumo.

Ateityje planuojama išmatuoti ŽPS veikimo trukmės priklausomybę nuo virtualaus HDD ir RAM dydžio.

LITERATŪROS SĄRAŠAS

[1] D. Reilly, C. Wren, T. Berry, „Cloud Computing: Pros and Cons for Computer Forensic Investigations“, International Journal Multimedia and Image Processing (IJMIP), Volume 1, Issue 1, March 2011.

[2] P. Mell, T. Grance, „The NIST Definition of Cloud Computing: Recommendations of the National Institute of Standards and Technology“, Special Publication 800-145, September 2011. http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf.

[3] Access US. http://www.accessus.net/business-services/cloud-computing.

[4] Cisco Global Cloud Index: Forecast and Methodology, 2010–2016, White Paper.
http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns1175/Cloud_Index_White_Paper.html.

[5] R. Chow, et al., „Controlling Data in the Cloud: Outsourcing Computation without Outsourcing Control“, CCSW'09, November 13, 2009, Chicago, Illinois, USA, Copyright 2009 ACM.

[6] K. Kent, S. Chevalier, T. Grance, H. Dang, „Guide to Integrating Forensic Techniques into Incident Response: Recommendations of the National Institute of Standards and Technology“, Special Publication 800-86, August 2006. http://csrc.nist.gov/publications/nistpubs/800-86/SP800-86.pdf.

[7] M. Taylor, J. Haggerty, D. Gresty, R. Hegarty, „Digital Evidence in Cloud Computing Systems“, Computer Law & Security Review 26, 2010, pp. 304-308.

[8] J. C. F. Cruz, T. Atkinson, „Digital Forensics on a Virtual Machine“, 49th ACM Southeast Conference, March 24-26, 2011, Kennesaw, GA, USA, 2011 ACM.

[9] J. C. F. Cruz, T. Atkinson, „Evolution of Traditional Digital Forensics in Virtualization“, ACM Southeast Conference, March 29-31, 2012, Tuscaloosa, AL, USA, 2012 ACM.

[10] D. Birk, „Technical Challenges of Forensic Investigations in Cloud Computing Environments“, Systematic Approaches to Digital Forensic Engineering (SADFE), 2011, IEEE Sixth International Workshop.

[11] X. Fu, Zh. Ling, W. Yu, J. Luo, „Cyber Crime Scene Investigations (C2SI) through Cloud Computing“, Proceedings of the 2010 IEEE 30th International Conference on Distributed Computing Systems Workshops.

[12] S. L. Garfinkel, „Digital Forensics Research: The Next 10 Years“, Elsevier Ltd. Digital Investigation Magazine. No. 7, 2010.

[13] S. D. Wolthusen, „Overcast: Forensic Discovery in Cloud Environments“, 2009 Fifth International Conference on IT Security Incident Management and IT Forensics, 2009 IEEE.

[14] R. K. L. Ko, et al., „Trustcloud: Framework for Accountability and Trust in Cloud Computing“, 2011 IEEE.

[15] W. Zhou, et al., „Towards Data-Centric View of Cloud Security“, CloudDB 2010, October 30, 2010, Toronto, Ontario, Canada, 2010 ACM.

[16] S. Krishnan, K. Snow, F. Monrose, „Trail of Bytes: Efficient Support for Forensic Analysis“, CCS'10, October 4–8, 2010, Chicago, Illinois, USA, Copyright 2010 ACM.

[17] R. Marty, „Cloud Application Logging for Forensics“, SAC'11 March 21-25, 2011, TaiChung, Taiwan, Copyright 2011 ACM.

[18] Y. Zhu, H. Wang, Z. Hu, G. Ahn, H. Hu, S. S. Yau, „Dynamic Audit Services for Integrity Verification of Outsourced Storages in Clouds“, SAC'11 March 21-25, 2011, TaiChung, Taiwan, 2011 ACM.

[19] C. Wang, Q. Wang, K. Ren, W. Lou, „Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing“, IEEE INFOCOM 2010.

[20] A. A. Yavuz, P. Ning, „BAF: An Efficient Publicly Verifiable Secure Audit Logging Scheme for Distributed Systems“, 2009 IEEE.

[21] R. Lu, X. Lin, X. Liang and Sh. Shen, „Secure Provenance: The Essential of Bread and Butter of Data Forensics in Cloud Computing“, ASIACCS'10 April 13-16, 2010, Beijing, China, 2010 ACM.

[22] The Sleuth Kit. http://www.sleuthkit.org.

[23] The Volatility Framework. https://code.google.com/p/volatility.

[24] V. Dagienė, G. Grigas, T. Jevsikova, „Enciklopedinis kompiuterijos žodynas“, 2-as papildytas leidimas, Vilnius: TEV, 2008.

Wireless Sensor Networks for Environment Monitoring

Zhun Shen
IBM,
Suzhou, China
shenzhun@outlook.com

David Olalekan Afolabi, Ka Lok Man, Hai-Ning Liang, Nan Zhang, Sheng-Uei Guan
Xi'an Jiaotong-Liverpool University, Suzhou, China
david.afolabi09@student.xjtlu.edu.cn,
{ka.man, haining.liang}@xjtlu.edu.cn,
{nan.zhang, steven.guan}@xjtlu.edu.cn

Prudence W. H. Wong
University of Liverpool,
Liverpool, UK
pwong@liverpool.ac.uk

Abstract. This paper is about two projects in the Wireless Sensor Network Laboratory. One project mainly focus on how to use Hough transform (HT) to detect and track angular changes between a series of images acquired by vision sensors attached to unmanned aerial vehicles (UAV) used in traffic surveillance. The other project is about porting Loosely-coupled Component Infrastructure to Zigduino platform, which is an AVR based microcontroller.

Keywords: Wireless Sensor Network, Loosely-coupled Component Infrastructure, Contiki, Ardunio, Zigduino, Ttraffic surveillance, Vibration reduction.

I. INTRODUCTION

Wireless Sensor Networks are composed of tiny embedded computers known as 'motes' that are equipped with sensors and low-power radios. These motes self-organize to form networks that are capable of sensing physical phenomena in the real world. Because WSNs are flexible enough to deploy in various situations, WSN lies at the center of solving 21st century challenges such as healthcare, environmental protection, and ensuring continued economic competitiveness. They have been widely used in different areas, such as environmental monitoring, infrastructure monitoring, manufacturing monitoring, health care monitoring and home monitoring [5].

The Wireless Sensor Network Laboratory: WSN-Lab (http://datics.nesea-conference.org/WSN-Lab) at the Department of Computer Science and Software Engineering of Xi'an Jiaotong-Liverpool University (XJTLU) – China focuses on enabling the creation of large scale WSNs that are capable of supporting applications such as industrial process optimization, energy conservation, water management and realizing the vision of the 'smart city'. Specifically, we provide development tools, middleware support and develop cutting-edge demonstration applications for WSN [5].

The WSN-Lab has a particular focus on important real world applications. For example, middleware technologies developed by the WSN-Lab have been applied to create pollution and flood monitoring system for the city of Sao Carlos in Brazil [5].

The remainder of this paper is structured as follows. Section II provides related projects of WSN Lab. Section III summarizes and discusses directions for future work.

II. RELATED PROJECTS

A. Real-time Vibration Reduction in UAV's image sensors using efficient Hough Transform

This project explores the use of Hough transform (HT) to detect and track angular changes between a series of images acquired by vision sensors attached to unmanned aerial vehicles (UAV) used in traffic surveillance. The small size and atmospheric conditions can lead to unstable flying, and this can induce vibration in the pictures [3]. Therefore with the aid of Wireless Vision Sensor Network images acquired from the field are transmitted to a base station where necessary information needed to improve the flight that is computed.

In this research we proposed that the vibration in the acquired images can be reduced by using the angular variation as feedback to stabilize the motion of the UAV as it captures subsequent frames [3]. After the application of an edge detection technique on the initial image, we compute the HT and select the most prominent line in the scene and continuously track its geometrical distortion for angular variations. The detected angular displacement is used as a feedback to the servomotors of the UAV to fine tune its motion

while acquiring the next image [3]. Other image registration algorithms where tested such as the MATLAB transform estimation which applies RANdom Sample Consensus (RANSAC) to match random data points in the image (see Figure 1). Both scale and rotation distortions can be extracted using this robust algorithm. On the other hand, it is not able to register images containing limited information and cannot be optimized to compute the distortion in a subsection of the image and therefore requires more time to process the whole image.

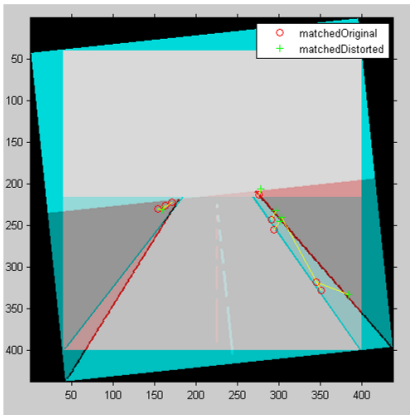


Fig. 1. Application of MATLAB image registration function

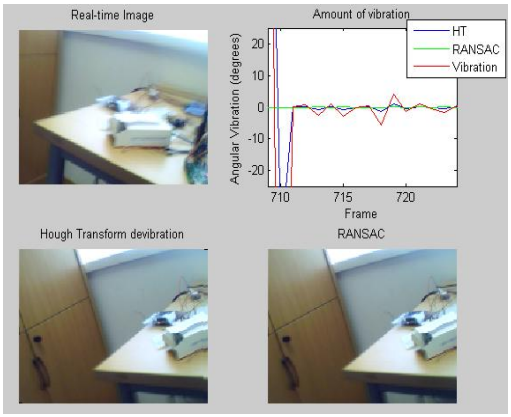


Fig. 2. Application of HT and RANSAC function

A CPU time profile of the implementation shows that at6 frames per second with a total of 554 calls to both the HT and RANSAC functions it took 31.47s and 40.93s respectively. In order to have a real-time feedback mechanism, the HT function is adequate, but calculating the Hough Transform requires the input of an edge-detected image. In this research we applied the canny edge detection which took 11.06s of the CPU time in the aforementioned test. The Zigduino board used for the WVSN is an Arduino microcontroller which contains an 802.15.4 radio and it supports a number of protocols such as Zigbee and 6LoWPAN. With a baud rate of 250 Kbit/s the low power consumption microcontroller is sufficiently able to transmit the flight control feedback data of 32 bits per frame, even at a significantly higher frame rate. Figure 2 shows the amount of vibration (in red – largest angular deviation). HT and RANSAC functions display the corrected image and output the amount rotation needed to register the next frame (in blue and green lines).

From the simulation, the angular vibration of the image is reduced by a damping factor of 0.75. Further experimentation will be undertaken to detect and remove translational vibration in the UAV. This research demonstrates how HT can be used to reduce image vibration in unmanned aerial vehicles to support surveillance of road traffic with the aid of a WVSN.

B. Porting LooCI to Zigduino

In this project, a suitable approach to port Loosely-coupled Component Infrastructure (LooCI) into the Zigduino platform is implemented and tested step by step.

LooCI comprises of a runtime reconfigurable application level component model, a hierarchical type system and a distributed event bus. Its features promote safe and efficient application development, management and reconfiguration [1]. LooCI cleanly separates distributed concerns from component implementation, supports application-level interoperability between heterogeneous WSN platforms, and provides compatibility testing of bindings at runtime [1]. LooCI runs on the Contiki operating system that provides a specialized set of abstractions that can be used to build highly efficient embedded software. Specifically, Contiki provides dynamic loading and unloading of individual programs and services [10].

Arduino is one of the most common hardware platforms because of its small size, low cost and modularity; it is used not only for prototyping, but also for creating interactive applications. Despite its many advantages, the basic Arduino platform lacks wireless connectivity, which makes it unsuitable for supporting WSN applications [11]. The Zigduino platform is an Arduino-compatible microcontroller platform that addresses this problem by integrating an 802.15.4 radio. Zigduino offers a reverse polarity SMA connector (RP-SMA) for an external antenna. All I/O pins on Zigduino are 5V compatible and can also run at 3.3V. Zigduino is based around ATmega128RFA1, and has 128 KB of flash memory of which 2 KB is occupied by the boot loader. It also has 16 KB of SRAM and 4 KB of EEPROM, which can be accessed through the EEPROM library [12]. The picture below shows a production Zigduino kit with all components.

Fig. 3. A Zigduino kit with all components

According to existing work, LooCI runs on AVR Raven. The aim of this research is to identify a viable approach to port the LooCI component to the Zigduino platform. It presents several challenges and we have tackled them in a sequential manner [13]. First, we need to install Contiki on Zigduino, which can be achieved using the Contiki port for Zigduino that is available on Github [14].

The next challenge is to write the required code to show LEDs blinking on Zigduino platform. To overcome this challenge, we can rely on the Contiki 2.5 doc, use and adapt both etime.h and process.h (C header files) to create the new code. Then the shell function is also used to check whether the network can run using two Zigduino boards. Another challenge we encounter in the process is to have a LooCI environment built into the elf file with a blink component. In order to evaluate the LooCI environment running well on Zigduino, shell modules in Contiki OS are also used to start and stop the blink component. The final challenge is to get the program work for the LooCI's wireless communication.

III. FUTURE WORK

There are two plans for our future work. The first one is to do numerical evaluation of LooCI on multiple different platforms, such as mobile phones, web pads and sensor nodes. Currently, we focus on the following three research directions. The first research direction is to evaluate the timing performance. We aim to test that how long it takes to install, remove, bind or unbind a component. The second research direction is about energy consumption, we plan to measure how many Joules it consumes for each fundamental execution in LooCI. The third research direction is about memory usage. We pay attention to determine the ROM and RAM usage for a number of benchmark applications with LooCI running on it.

The other plan is to develop applications of LooCI for real-world applications such as energy conservation and pollution measurement.

ACKNOWLEDGEMENTS

The research work presented in this paper was partially supported by Xi'an Jiaotong-Liverpool University (XJTLU) Research Development Fund, XJTLU Final Year Project Fund, XJTLU SURF Fund and Transcend Epoch International Co., Ltd, Hong Kong.

REFERENCES

[1] Danny Hughes, K. L. Man, Zhun Shen, and Kyung Ki Kim, "A Loosely-Coupled Binding Model for Wireless Sensor Networks", IEEE Proceedings of the 9th International SoC Design Conference – ISOCC 2012, Jeju, South Korea, November, 2012.

[2] Zhun Shen, Ka Lok Man, Chi-Un Lei, Eng Gee Lim and Joongho Choi, "Assuring System Reliability in Wireless Sensor Networks Via Verification and Validation", IEEE Proceedings of the 9th International SoC Design Conference – ISOCC 2012, Jeju, South Korea, November, 2012.

[3] David Olalekan Afolabi, Ka Lok Man, Hai-Ning Liang, Eng Gee Lim, Zhun Shen, Chi-Un Lei, Tomas Krilavičius, Yue Yang, Lixin Cheng, Vladimir Hahanov, and Igor Yemelyanov, "A WSN Approach to Unmanned Aerial Surveillance of Traffic Anomalies: Some Challenges and Potential Solutions", IEEE Proceedings of the 10th East-West Design and Test International Symposium - EWDTS'12, Kharkov, Ukraine, September, 2012.

[4] Daniel Hughes, Hai-Ning Liang, Nan Zhang, Kaiyu Wan, Ka Lok Man and Eng Gee Lim, Need for an Evolvable and Adaptive Wireless Sensor Network Software Stack, in the International Journal of the Applied Mechanics and Materials, 2012.

[5] Wireless Sensor Network laboratory,Online: http://datics.nesea-conference.org/WSN-Lab/[accessed on March 2013].

[6] David Olalekan Afolabi, Zhun Shen, K.L. Man, Hai-Ning Liang, Nan Zhang, E.G. Lim, Modelling and Analysis of LooCI Models in Zigduino, in Proceedings of the IAENG International MultiConference of Engineers and Computer Scientists - IMECS'13, Hong Kong, March, 2013.

[7] Chi-Un Lei, J.K. Seon, Zhun Shen, Ka Lok Man, Danny Hughes, Youngmin Kim, Building a Laboratory Surveillance System via a Wireless Sensor Network, in Proceedings (Lecture Notes in Computer Science Series Volume) of the 8th International Conference on Grid and Pervasive Computing - GPC 2013, Korea, May, 2013.

[8] Chi-Un Lei, Kaiyu Wan, Ka Lok Man, Developing a Smart Learning Environment in Universities Via Cyber-Physical Systems, in Proceedings (Elsevier Procedia Computer Science series) of the First International Conference on Information Technology and Quantitative Management - ITQM 2013, Suzhou, China, May, 2013.

[9] Zhun Shen, Ka Lok Man, Hai-Ning Liang, Nan Zhang, David Afolabi and Eng Gee Lim, Porting LooCI Component into Zigduino, Proceedings (Elsevier Procedia Computer Science series) of the First International Conference on Information Technology and Quantitative Management - ITQM 2013, Suzhou, China, May, 2013.

[10] A. Dunkels, B. Gronvall, and T. Voigt, Contiki - A Lightweight and Flexible Operating System for Tiny Networked Sensors, 29th Annual IEEE International Conference on Local Computer Networks 2004, pp. 455- 462.

[11] V. Georgitzikis, O. Akribopoulos, I. Chatzigiannakis, Controlling Physical Objects via the Internet using the Arduino Platform over 802.15.4 Networks, IEEE Latin America Transactions (Revista IEEE America Latina), 2012, vol. 10, no. 3, pp. 1686-1689.

[12] Logos-electro, Online: http://logos-electro.com/zigduino/ [accessed on March 2013].

[13] S. Alexandru, Porting the Core of the Contiki, 2007, Online: http://www.eecs.iu-bremen.de/archive/bsc-2007/stan.pdf [accessed on March 2013].

[14] Manizcbug, "Zigduino platform for Contiki OS", 2013, Online: https://github.com/maniacbug/contiki-avr-zigduino [accessed on March 2013].

INFORMACINĖS TECHNOLOGIJOS ● IVUS 2013 ● ISSN 2029-249X ● eISSN 2029-4824

Analysis and Modelling of Dynamic Spectrum Access Information Infrastructure (DSA II)

Gintarė Sukarevičienė*, Vladislav V. Fomin**
Department of Applied Informatics
Vytautas Magnus University
Kaunas, Lithuania
gintare.sukareviciene@gmail.com*, vvfomin@gmail.com**

Abstract. **Due to the fact that we are running out of usable radio frequencies, contemporary spectrum management focuses on more and more improved spectral efficiency, better utilization of the radio frequencies by intelligent sensing the surrounding radio environment. This approach is referred to as Cognitive Radio (CR) – a new way of radio communications, which promises to become a stepping stone to the new paradigm of wireless services. However, to date the principle of management of radio frequencies is mostly based on a licensing regime which only allows strict frequency allocation to specific services and technologies. This means that spectrum scarcity problem is as much a result of limited spectrum resources, as it is of inefficient spectrum management policy. Given the technological and policy limitations of the current wireless communications paradigm, Dynamic Spectrum Access (DSA) became an important concept in the research seeking to find ways to boost efficiency of radio spectrum utilization. In this paper we explore the concept of DSA, aiming to make sense of the complex phenomena and thus develop a foundation for the further research.**

Keywords: Cognitive Radio, Dynamic Spectrum Access, Information Infrastructure, System, White Space.

I. INTRODUCTION

During the last century wireless communication technologies have become one of the most exciting and the most promising technology bringing substantial benefits to the industry and creating a need for radio spectrum as a medium for wireless communications for the use of wide range of new emerged opportunities. The life of human becomes more and more technology life but, unfortunately, today's increasing variety of wireless communication systems is not a sign that radio spectrum is being used efficiently and this must be taken into account and the relationships between policy, practices and technology must be analyzed.

In wireless communication systems, efficient spectrum access since long has been recognized as an essential element of organizing telecommunications services. To overcome the limiting rigidities of the existing wireless communications infrastructure, a new kind of infrastructure is needed, one where previously unrelated and often untried technological solutions would be meshed to give birth to a new system [10]. Development of the next generation Dynamic Spectrum Access Information Infrastructure (DSA II) requires versatile and comprehensive modelling, encompassing such broadly defined variables as regulations, standards, new protocols, business models, licensed rights (if any), rules and principles for effective radio spectrum usage, and other.

This paper focuses on describing a number of qualitative and quantitative constraints (or variables), which are likely to shape the future development of ground wireless communications services and their enabling infrastructures. The purpose of this paper is to give a momentum to the analysis of future DSA II in Lithuania, laying foundation for the further research.

II. THE CONCEPT OF DYNAMIC SPECTRUM ACCESS INFRASTRUCTURE

Dynamic Spectrum Access Information Infrastructure (DSA II) has a large potential to make spectrum reform to current static spectrum management policy for particular users in particular geographic areas by creating new business models, spectrum allocation policy and even certification processes not only in the Lithuania but also in the other countries because the development and use of DSA II is considered around the globe. Unfortunately DSA II development has too high complexity due to many variables and it requires discovering and describing a number of qualitative and quantitative constraints, which are likely to shape the future development of ground wireless services and infrastructures.

Thus, to start modelling DSA II, the relationships between different variables and their impact to each other within DSA infrastructure must be outlined (see Fig. 1 for the concept).

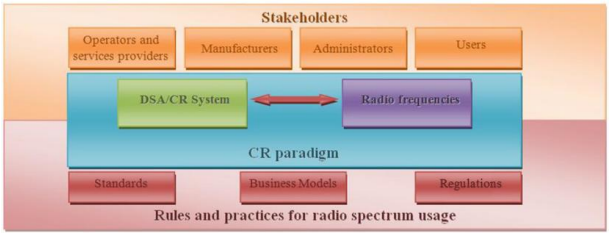


Fig. 1. DSA infrastructure (elaborated by the authors)

A. Stakeholders

At the top of Fig. 1, possible stakeholders and their roles within the wireless telecommunication services are enlisted. Analysis of the relationship between different stakeholders requires a close look at each stakeholder in order to understand their preferences, impact or involvement by answering such questions as e.g., who stands to benefit from the innovation and in what way? What are the perceptions, expectations and understanding of DSA II of different stakeholders, how different stakeholders can/should contribute to the formation of the DSA II?

Stakeholders who are considered here affect or are affected by DSA II and its development, could be divided into two broad groups:

- Directly impacted by DSA II, such as operators and services providers, manufacturers, administrators. The stakes of these stakeholders are not equal within wireless telecommunication services, e.g., the stake of existing and new operators' as well as service providers' is to reach the most positive value through DSA II enabled services that promise a wide range of new opportunities to consumers. Likewise, the stake of manufacturers' is to benefit from developing new technologies, new infrastructure equipment that supports regulatory and technical conditions and can enable new services to customers. Finally administrators' stake is to gain through logistics management of the physical network (e.g., while developing and operating geo-location database (GDB) in the most proper way in order to improve spectrum utilization and benefit from enabling the innovation in the market).
- Indirectly impacted by DSA II, such as users, whose stake is to benefit from the advanced features of DSA (e.g., on-demand switching between multiple operators in order to ensure the best quality of service (QoS) and the lowest access/use prices).

The main difference between these two groups of stakeholders is that directly impacted stakeholders are the key stakeholders, as they can influence the development of DSA II much more than indirectly impacted stakeholders.

B. Rules and practices for radio spectrum usage

At the bottom of Fig. 1., elements defining the rules and practices (principles) for radio spectrum usage are listed:

- Standards. Analysis of standards should answer such questions as: which technologies must be standardized? What is the environment of complementary and competing standards? For example, at the moment 802.11af (WiFi extension to TV White Space) is under development [13]. It will be less demanding and focusing on geo-location only. Likewise, here are many others standards that are already existing or will emerge in due time [3].
- Business Models. Business Models should help better understand how DSA II can be deployed [11], [12], [16], [18]. They should be acceptable not only for government, but also for the industry and the consumers, because all players should benefit from having knowledge on what expect from DSA II, how the innovation will proceed, etc.
- Regulations. Analysis of the regulation environment should answer such a question as: what are all the relevant rules and how stakeholders could benefit from playing by these rules? [19] Basically, the US (FCC) and European (ECC) authorities play the main role in setting the rules, but in Europe also national/ regional authorities are playing a role. Due to this, region- or country-specific regulations in Europe must be developed or approved by the individual European member states, depending on the national interests, wireless telecommunications market structure, but aslo taking into account decisions made by the neighboring countries (as radio waves do not recognize the national borders).

C. Cognitive Radio (CR) paradigm

The CR paradigm in the centre of Fig. 1 is a conceptual construct of the problem domain, consisting of enabling technology infrastructure and the radio spectrum resources (frequencies). CR paradigm is expected to help provide new innovative ways for finding more spectrum for any given service by suggesting more flexible/intelligent regimes of spectrum allocation and adressing constraints for current and future radio communications [9]. Broadly, the main elements of the paradigm could be:

1) Radio frequencies – a pre-defined set of radio frequency bands and rules for accessing (and using) them.

Generally, radio frequencies can be regarded as limited resource in nature due to the existing practice of spectrum allocation. Increasingly, radio spectrum becomes a scarce resource, as wireless bandwidth is increasingly becoming crowded due to the growing demand for wireless communications. Thus, there is a growing risk of different devices interfering with each other, and, as a result, a diminishing QoS. According to the FCC estimates, the surplus turns into a deficit as early as this year in providing voice, text and Internet services to its customers by the U.S. mobile phone industry [2]. This brings to the conclusion that radio spectrum is not well managed and this problem should be solved. One possible solution to the problem – reallocating frequency bands from obsolete or less popular services to the new ones for which there is a high demand.

Some initial prime candidates of frequencies where CR paradigm could be considered for deployment:

- TVWS – the range of spectrum available between the signals used by television stations in UHF bands (470-790 MHz), because after the digital switch over not all channels in this range are occupied at each location. It is challenging to think of TVWS use because one of the ways to get more spectrum is to reallocate it or to allocate unused spectrum. Overall, access to TVWS could be a stepping stone to facilitate future dynamic access to WS in other bands;

- 1452-1492 MHz and VHF T-DAB band. 1452-1492 MHz refers to so-called L band and T-DAB refers to the bands that are allocated for public DAB services;
- Radar bands, fixed service bands that refers to the wide band of transmitted frequencies in which radars system and fixed services works;
- 2.3-2.4 GHz frequencies.

2) DSA/CR System – a technology toolbox to allow dynamic access to spectrum by a number of users.

Here DSA/CR System could be understood as a set of different technology, business and policy elements, detailed possible development scenarios, models, future technology trajectories, spectrum allocation policy and even certification processes that are organized based on certain relationships between each other for one purpose – to enable DSA. This is future environment that require taking into account technology, business and policy related aspects (e.g. GDB approach in the context of DSA) and determining relationships between them (e.g. the form of white space devices (WSDs) Internet connectivity).

To start with, all possible DSA strategies could be broadly categorized under three models [14]:

- Dynamic Exclusive Use Model that introduces flexibility in order to improve spectrum efficiency using two different approaches: spectrum property rights and dynamic spectrum allocation.
- Open Sharing Model (Spectrum Commons Model) that addresses technological challenges using two different spectrum sharing strategies: centralized and distributed.
- Hierarchical Access Model that refers to the hierarchical licensed spectrum access for primary and secondary users using two different approaches of spectrum sharing: spectrum underlay and spectrum overly.

The comparison of these three models shows that hierarchical model is the most suitable for current spectrum management policies and legacy wireless systems, because it could eliminate White Space (WS) in spectrum and further improve spectrum efficiency by imposing restrictions on when and where secondary users may transmit. This could bring to the conclusion that DSA/CR System should be based on Hierarchical Access Model which the basic idea is to open licensed spectrum to secondary users while limiting the interference perceived by primary users [14].

III. THE CASE OF LITHUANIA

In the case of Lithuania, developing DSA II requires modelling the co-dependence of the aforementioned elements/variables (see Fig. 1).

One of the examples that corresponds to "Operators and service providers" variable under the "Stakeholders" element in Fig. 1 could be the number of operators. Current number of three mobile operators in Lithuania must be examined taking into account competitive environment, needed investments, possible return and other factors by answering such a questions as: does Lithuania need the fourth mobile operator and could the fourth mobile operator benefit although in Lithuania the market is enough small? For example, in Latvia there are also three mobile operators (such as: Bite, LMT and Tele2) but, despite the case of Lithuania and Latvia, in some countries there are four and even more mobile operators [15]:

- Russian Federation: MTS, VimpleCom, MegaFon, Tele2 Russia, Rostelecom, other;
- Finland: DNA, Elisa, TeliaSonera, other;
- Other countries (see Fig. 2 for more information).

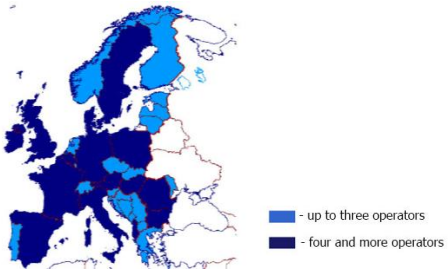


Fig. 2. Different number of operators in different countries within the Europe [15]

With regards to all these factors above, it becomes clear that "Operators and service providers" variable in Fig. 1 has some co-dependence with other variables and elements within DSA II, e.g.: "Business Models" under "Rules and practices" and "Users" under "Stakeholders". These co-dependencies must also be examined in order to get better view of the whole DSA II.

Another example could be that in Lithuania there also are a number of national as well as local considerations, limitations, requirements and cross-border frequency coordination issues for the use of radio frequencies for electronic communication services depending on the situation of the neighbouring countries such as Latvia, Belorussia, Russian Federation, as well as Finland [15]:

1) Few limitations and requirements from Latvia:
- The use of digital TV BC stations (end 2015);
- The impact of DVB-T stations on Lithuanian mobile services stations, e.g.: impact of 61-64 TV channels on mobile stations and 66-69 TV channels on base stations.

2) Few limitations and requirements from Belorussia:
- Mobile service stations must be implemented more than 100 km to the border.

3) Few limitations and requirements from Russian Federation:
- Protection of aeronautical radionavigation (ARNS) systems in the 790-862 MHz band: RSBN (in 790.5-793.5 MHz, in 798.5-801.5 MHz and in 806.5-809.5 MHz), RLS1 (in 842.5-845.5 MHz, in 851.5-854.5 MHz and in 857.5-860.5 MHz), RLS2 (836.0-839.0 MHz);

- Protection of FX systems in 820-821 MHz band;
- Mobile service stations must be implemented more than 100 km to the border;
- SAT system;
- Digital switchover 2015.

4) Few limitations and requirements from Finland:

- The impact of DVB-T2.

All these constraints listed above belong to the "Regulations" variable under "Rules and Practices for radio spectrum usage" element in Fig. 1, but also has implications for "DSA/CR System" and "Radio frequencies" variables under the "CR Paradigm" element as well as for majority of variables under "Stakeholders" element. All together these limitations and requirements could lead to different possible technology and administrative development trajectories, for example, due to the requirements of protection of Russian FX system there could be few different versions with three different protection zones in Lithuania, as shown in Fig. 3, which could impact the number of mobile operators (3 or 4) in the country:

- protection zone a) defined by calculating ≤10 dBμV/m electromagnetic field within the boarder from one station.
- protection zone b) defined by International Telecommunication Union (ITU) conditions [8].
- protection zone c) defined by considering the influence of Lithuanian stations operating within 64 channel to Russian systems that are registered by ITU.

Fig. 3. Different protections zones [15]

In any case, if there are not enough details, it could be impossible to make good modelling. Thus, more detailed analysis of the regulatory environment not only in Lithuania but also in neighbouring countries could help to develop more feasible models for DSA in Lithuania.

Besides this, DSA infrastructure must be examined not only within the system sense but also in a broader sense (for example, viability of merging LTE and DVB-T2 from the point of view of the needed information infrastructure to support/enable DSA within the merged system-space) by reducing factors and administrative development trajectories what are surely already known and examining these factors and possible trajectories that are not surely known depending on the situation of different European countries.

A. Methods to be applied to DSA/CR system analysis and modelling

Techno-economic modelling could be the first future-oriented step for analysing, evaluating and assessing economic feasibility of complex DSA technical system [17]. Analysis and modelling should combine a number of methods of technology analysis, cost-benefit analysis, scenarios, trend analysis, and expert opinion to quantitative modelling.

Results gained from techno-economic modelling studies could include: market and services description, stakeholders' models and strategies, GDB and other technology definitions and deployment scenarios, technology relationships with the market, revenue/cost modelling, discounted cash flow analysis, sensitivity analysis [17].

IV. CONCLUSIONS AND EXPECTED CONTRIBUTIONS

While future business models and regulatory frameworks are under discussion it becomes very important to realize the economic potential of future DSA II architectures and services. Especially, due to the complexity of the problem area, further research on Dynamic Spectrum Access Information Infrastructure (DSA II) related issues, such as cost driving technical elements, the timing influence and administrative and selling expenses has to be undertaken.

The aim of this research is to develop a preliminary model for DSA II that could enhance our understanding of the ways in which new information infrastructures are being formed, and serve as valuable source of knowledge to relevant bodies in Lithuania and elsewhere, thus facilitating decision making process on how DSA-supporting information infrastructure build-out should unfold.

ACKNOWLEDGMENT

This work was supported in part by COST Action IC0905 "Techno-Economic Regulatory framework for Radio spectrum Access for Cognitive Radio/Software Defined Radio" (TERRA) (www.cost-terra.org).

REFERENCES

[1] Europe's Information Society: Thematic Portal; find your operator: http://ec.europa.eu/information_society/activities/roaming/consumer/operators/index_en.htm.

[2] FCC. Mobile Broadband: The Benefits of Additional Spectrum. OBI Technical Paper Series. October 2010.

[3] G. Baldini, O. Holland, V. Stavrouslaki, K. Tsagkaris, P. Demestichas, A. Polydoros, S. Karanasios, and D. Allen. "The Evolution of Cognitive Radio Technology in Europe: Regulatory and Standardization Aspects," Telecommunication Policy, Elsevier 2012.

[4] G. Sukarevičienė, "Developing Business Model for Geo-Location Database for Operations of Cognitive Radio in the TV White Space Bands,", unpublished manuscript.

[5] G. Sukarevičienė, D. Jankūnas, and V. V. Fomin, "On the Use of a Controlling Database for Cognitive Radio in Wireless Body Area Network," Proceedings of the 7th International Conference on Body Area Networks, pp. 221-224, 2012, ICST.

[6] G. Sukarevičienė, M. Barrie, S. Delaere, P. Ballon, and V. V. Fomin, "Identifying Geolocation Database Requirements for Distinct Wireless Applications," unpublished manuscript.

[7] G. Sukarevičienė, and V. V. Fomin, "Business Model Analysis: The Case of TV White Space Database," ISSN 2029-249X, pp. 125-138, 2012.

[8] ITU home page: http://www.itu.int/en/Pages/default.aspx.

[9] J. Mitola, "Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio," Sweden: May 2000.

[10] K. Lyytinen, and V. V. Fomin. "Achieving High Momentum in the Evolution of Wireless Infrastructures: The Battle over the 1G Solutions", Telecommunications Policy 26 (2002) pp. 149-170.

[11] M. Barrie, S. Delaere, G. Sukarevičienė, J. Gesquiere, and I. Moerman, "Geolocation Database beyond TV White Spaces? Matching Applications with Database Requirements,", DySPAN 2012.

[12] P. Ballon, and R. Hawkins. "Standardization and Business Models for Platform Competition: The Case of Mobile Television," International Journal of IT Standards & Standardization Research (JITSR) 7 (2009): 1–12.

[13] P. Talmola, "Cognitive Radio/TV White Space Standardisation," January 2013.

[14] Q. Zhao, and B. M. Sadler, "A Survey of Dynamic Spectrum Access: Signal Processing, Networking, and Regulatory Policy," IEEE Signal Processing Magazine, May 2007.

[15] Ryšių reguliavimo tarnyba, "Radijo ryšio plėtros 790-862 MHz radijo dažnių juostoje planas," gruodis, 2012.

[16] S. Delaere, and P. Ballon. "Multi-Level Standardization and Business Models for Cognitive Radio: The Case of the Cognitive Pilot Channel," The 3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN'08). Chicago, IL: IEEE, 2008. doi:10.1109/DYSPAN.2008.67.

[17] T. Simura, "Techno-econimic Modelling of Wireless Network and Industry Architectures," Finland: March 2012.

[18] V. V. Fomin, A. Medeišis, and D. Vitkutė-Adžgauskienė. "In Search of Sustainable Business Models for Cognitive Radio Evolution," Technological and Economic Development of Economy Vol.18, No. 2 (2012): pp. 230–247. doi:10.3846/20294913.2012.663415.

[19] V. V. Fomin, A. Medeišis, and D. Vitkutė-Adžgauskienė. "The Role of Policy in the Development of Cognitive Radio Systems: Co-Evolutionary Perspective," The 7th International Conference on Standardisation and Innovation in Information Technology (SIIT 2011), pp. 79–90. Berlin: Mainz Publishing, Aachen, 2011.

Statistinių kalbos modelių taikymas natūralios kalbos dialoguose

Toma Šalnaitė
KHF Informatikos katedra
Vilniaus Universitetas
Kaunas, Lietuva
t.salnaite@gmail.com

Santrauka. Šiame straipsnyje kalbama apie dialogų modeliavimo problematiką, apžvelgiama dialogų modelių reikšmė dialogų sistemose, supažindinama su labiausiai paplitusiais statistiniai kalbos modeliais, kurie yra naudojami kuriant dialogų modelius. Aprašoma natūralių dialogų rinkimo procesas, pristatomi pirminiai tokių dialogų tyrimai ir analizės rezultatai.

Reikšminiai žodžiai: dialogų modeliavimas, dialogų sistemos, neredaguoti dialogai, kalbos modeliavimas, statistiniai kalbos modeliai.

I. ĮŽANGA

Informacinių technologijų plėtra leidžia vis plačiau kompiuterizuoti tekstinės informacijos, t. y. išreikštos natūralia kalba, apdorojimo uždavinius. Dėl to yra kuriami įvairūs statistiniai kalbos apdorojimo modeliai, kurių duomenis naudojant būtų galima pereiti į aukštesnius kalbos kompiuterizavimo lygmenis ir sukurti naujas informacines paslaugas.

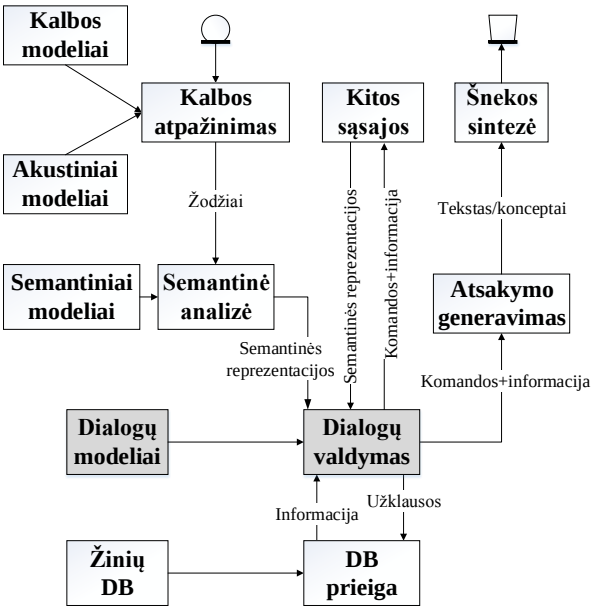
Statistiniai kalbos modeliai svarbūs ir šnekamosios kalbos atpažinimo sistemose, kai jie panaudojami atpažintų sakinių ir frazių hipotezėms patikrinti. Tačiau taikant statistinius kalbos modelius dialogų atpažinimo ir valdymo uždaviniuose susiduriama su problema, kad dauguma statistinių modelių parametrų būna nustatyti iš tekstynų, kuriuose saugomas redaguotas, t. y. apdorotas tekstas. O realioje šnekoje naudojamos frazės ir sakiniai dažnai ženkliai skiriasi nuo tų sakinių, kurie saugomi tekstynuose. Tokiuose sakiniuose gana dažnai sutinkami pasikartojimai, kurie būna pašalinti redaguojant, gramatiškai netaisyklingi sakiniai, negramatiniai leksiniai elementai ir pan. Todėl taikant statistinių modelių įverčius, gautus iš redaguotų tekstynų, stebimas neatitikimas taikant juos šnekamosios kalbos dialogų apdorojimo procese. Dėl to labai svarbu, kad tokiu atveju būtų naudojami statistinių modelių įverčiai, gauti iš kalbos pavyzdžių, kurie yra panašiausi į realiuose dialoguose naudojamos kalbos pavyzdžius.

Siekiant šio tikslo susiduriama su kliūtimi – turima labai mažai tokio tipo tekstinių išteklių (lietuviškų neredaguotų dialogų pavyzdžių beveik nėra sukaupta). Šiame tyrime buvo nuspręsta surinkti lietuviškų maitinimo įstaigos užsakymų dialogų pavyzdžius, juos išanalizuoti, nustatyti, kuo dialogų tekstas skiriasi nuo redaguoto teksto, ir įvertinti natūralių dialogų ir panašių tekstynų dialogų parametrų įverčių skirtumus.

II. DIALOGŲ MODELIŲ REIKŠMĖ DIALOGŲ SISTEMOSE

Komunikacija tarp žmogaus ir kompiuterio nėra identiška žmonių bendravimui [1], dėl to kalba modeliuojama, siekiant sukurti dialogų modelius, kurie formuotų efektyvią žmogaus ir kompiuterio sąveiką.

1 paveiksle galima matyti, kad dialogų sistemos yra sudarytos iš skirtingų modulių (tarp kurių yra ir dialogų modeliai) ir, nors daugumos jų metodai yra apibrėžti, galutinė sistema yra sudėtinga.



1 pav. Dialogo sistemos struktūra

Dialogo sistemas gali sudaryti fiksuotosios sekos apdorojimo etapai, pradedant nuo kalbėjimo atpažinimo, apdorojimo / analizės, dialogo valdymo ir atsakymo generavimo [2].

Dialogų modeliai yra skirstomi į tris grupes [3]:

- Kompiuteriniai dialogų modeliai[1], kurie yra orientuoti į paprastų dialogų sistemas ir dialogų valdymą.
- Paprasti dialogų modeliai[2] – baigtinių būsenų skaičiaus automatai ir dialogų gramatikos.
- Dinaminiai dialogų modeliai[3] – planais grindžiami koperatyviniai modeliai, agentų teorija grindžiami modeliai.

Kompiuteriniai dialogų modeliai labiausiai susiję su informacijos pasikeitimu tarp žmogaus ir kompiuterio, nes šio tipo dialogų modeliuose vienas iš dalyvių yra dirbtinis (kompiuteris).

III. KALBOS MODELIAVIMAS

Tekstinės informacijos kompiuterizavimo pagrindas yra kalbos modeliavimas, kurio uždavinys yra sukurti kalbos modelius, kurie gali užfiksuoti natūralios kalbos dėsningumus.

Kalbų modeliai prognozuoja žodžio informacijos įvykių tikimybę pateiktame kontekste, kuriame šis žodis turėtų atsirasti, nes bet kokioje kalboje egzistuoja sintaksiniai, semantiniai, pragmatiniai apribojimai, kurių poveikis kai kurioms žodžių sekoms labiau tikėtinas. Įvairios tam tikros dalykinės srities kalbos sistemos turi taip pat tam tikrų apribojimų dėl žodyno dydžio ir sintaksės.

Bet kokioje kalboje egzistuoja daug įvairių galimų žodžio kombinacijų, dėl to tam tikra dalis sąlyginių tikimybių galėtų būti reikalingos. Nėra įmanoma apibrėžti visų šitų tikimybių individualiai, dėl to būtinos kai kurios modeliavimo prielaidos, pagal kurias būtų galima apibrėžti kai kurias sąlygines, labiau valdomas tikimybes [4].

Kalbos modeliavimo metodai gali būti padalyti į dvi pagrindines grupes [5]:

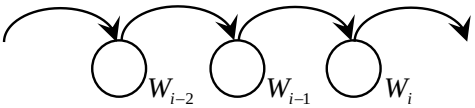
- žiniomis pagrįsti kalbos modeliai yra suprojektuoti kalbininkų rankiniu būdu, remiantis jų kalbos žiniomis ir intuicija apie geriausią būdą pateikti formalius kalbos objektus ir ryšius;
- statistiniai kalbos modeliai, priešingai, atsiskleidžia kaip neprižiūrimo kalbos modelio mokymo rezultatas.

Kadangi siekiama tirti neredaguotą dialogo kalbą, plačiau apžvelgiami statistiniai kalbos modeliai.

IV. STATISTINIAI KALBOS MODELIAI

Statistiniai kalbos modeliai suteikia galimybę atstovauti žodžiams, frazėms ir sakiniams pagal tikimybių padalijimą, taip supaprastinant ir tiksliai suformuojant išvadas apie konkrečios srities kalbą [6]. Statistiniai kalbos modeliai modeliuoja žodžio pasirodymo tekste tikimybinius skirstinius pagal gretimus žodžius. Yra išskiriami pagrindiniai du metodai: Bajeso taisyklė ir N-gramų metodas [5].

N-gramos yra vienos paprasčiausių, tačiau geriausių ir populiariausių kalbos modelių, kuriuose įvertinama žodžio sekos tikimybė. Taikant n-gramas daroma prielaida, kad žodis priklauso nuo ankstesnio žodžio (n-1) [7][8].



2 pav. N-gramų modelių veikimo esmė

Žodžio sekos tikimybei apskaičiuoti naudojama bendroji formulė:

$$P(W) = \prod_{i=1}^{K} P(\omega_i | \omega_{i-n+1}, ..., \omega_{i-1}) \quad (1)$$

Čia K yra žodžių seka, $P(\omega_i | \omega_{i-n+1}, ..., \omega_{i-1})$ yra tikimybė, kad bus ištartas žodis ω_i, jei prieš tai buvo pasakyta $|\omega_{i-n+1}, ..., \omega_{i-1}$, n žymi modelio eilę.

Dažniausiai yra išskiriami trys n-gramų modeliai: vieno žodžio seka (n=1) vadinama unigrama, dviejų (n=2) – bigrama, trijų (n=3) – trigrama. Gali būti naudojama ir ilgesnė žodžių seka, kai n=4, n=5 ar dar daugiau, tačiau tai – rečiau naudojami atvejai.

Paprasčiausias n-gramų variantas yra unigrama. Ji (n=1) neįvertina ankstesnių žodžių pasirodymų, tačiau gali įvertinti tikimybes ir žodžių dažnumą dialoguose tam tikra tematika.

$$P(W) = \prod_{i=1}^{K} P(\omega_i) \quad (2)$$

Bigrama (n=2) – dviejų žodžių kombinacija. Naudojant bigramas dažnumui ir tikimybėms skaičiuoti yra atsižvelgiama į ankstesnių žodžių pasirodymus.

$$P(W) = \prod_{i=1}^{K} P(\omega_i | \omega_{i-1}) \quad (3)$$

Trigrama (n=3) – trijų žodžių kombinacija. Naudojant trigramas dažnumui ir tikimybėms skaičiuoti yra atsižvelgiama į ankstesnių žodžių pasirodymus.

$$P(W) = \prod_{i=1}^{K} P(\omega_i | \omega_{i-2}, \omega_{i-1}) \quad (4)$$

[1] Angl. *Computational Dialogue Models.*
[2] Angl. *Simple Dialogue Models.*
[3] Angl. *Dynamic Dialogue Models.*

V. TAIKYMO PROBLEMATIKA

Teoriniu požiūriu daugelis statistinių modelių yra iš principo lygiaverčiai [9], o jų efektyvumą konkrečioje situacijoje labiausiai veikia mokymo duomenų kiekis ir jų parametrai. Tačiau ir tai nėra pati svarbiausia ypatybė ir visų metodų sunkumas.

Šiuolaikinėse kalbos apdorojimo sistemose sintaksiniai modeliai yra įvertinami naudojant tekstynus (atsitiktinius arba specialiai paruoštus (pvz., tematinius) tekstų rinkinius). Paprastai šie rinkiniai gaunami iš jau paruoštų tekstinių duomenų rinkinių – knygų, laikraščių, stenografuotų posėdžių medžiagos ir pan. [10]. Visiems šitiems tekstų rinkiniams yra būdinga tai, kad jie yra specialiai apdoroti. Knygos rašomos gražia literatūrine kalba, laikraščių, žurnalų ir posėdžių publikuojama medžiaga yra redaguojama ir pataisoma. Pataisymai apima linksnių suderinimą, pasikartojimų pašalinimą, neliteratūrinių žodžių pašalinimą ir panašias veiklas. O realioje šnekamojoje kalboje ir realiuose kalbos dialoguose, kuriuos turi apdoroti kompiuterinės sistemos, dažnai susiduriama su visai kitais kalbos pavyzdžiais. Juose gausu pasikartojimų, linksniai dažnai nesuderinti, žodžių tvarka sakiniuose įvairuoja, vartojami įvairūs neleksiniai žodžiai ir pan. [11].

Akivaizdu, kad statistinių kalbos modelių parametrai, nustatyti iš tekstynų ir gauti iš realios šnekamosios kalbos pavyzdžių, turėtų skirtis, – tuo ir grindžiama tiriamoji sritis.

VI. DUOMEMŲ DIALOGŲ MODELIAMS RINKIMAS IR APDOROJIMAS

Šiuo metu yra surinkta medžiaga (dialogai), kuri bus skirta tolimesniems uždaviniams įgyvendinti t. y. pagal surinktus duomenis bus siekiama įvertinti statistinių modelių efektyvumą ir praktinio panaudojimo galimybes dialoguose.

Reikia pripažinti, kad šiuo metu dar nėra surinkti visi numatomi panaudoti teksto dialogai, tačiau pradinis duomenų rinkinys leidžia atlikti kai kuriuos pirminius tyrimus. Kita vertus, rinkti natūralius dialogų pavyzdžius nėra lengva, o kartais gali būti netgi draudžiama juridiškai.

Bandomajam tyrimui atlikti buvo surinkti 42 natūralios kalbos dialogai lietuvių kalba. Tyrimui pasirinkti lietuviškų maitinimo įstaigų užsakymų dialogai, nes jie dažniausi ir juos rinkti lengviau. Renkant teksto dialogų pavyzdžius dialogo dalyviai buvo įspėjami apie jų pokalbio įrašymą, siekiant atspindėti realią užsakymo situaciją. Pokalbio metu buvo tiksliai užfiksuotas kiekvienas jo metu pavartotas žodis, jų tvarka ir t. t. Todėl pokalbis neretai buvo įrašomas ir tik po to tiksliai transkribuojamas. Šiuo metu yra surinktas teksto dialogų žodžių žodynas, kurį sudaro apie 1 200 žodžių, kuriuos vartoja klientai. Žodyne nurodomas skirtingų žodžių formų kiekis (lemos), t. y. pavyzdžiui, žodžiai „noriu“ ir „norėčiau“ vertinami kaip skirtingi.

VII. DUOMENŲ ANALIZĖ

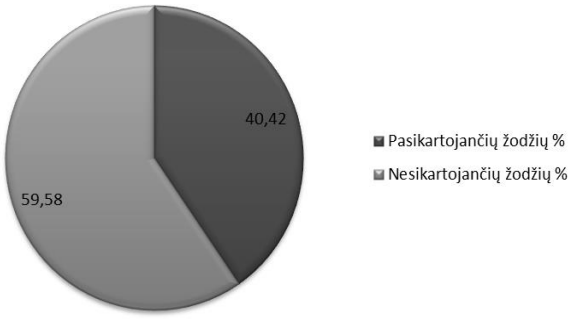
Pirmiausia buvo įvertinta, kokio ilgio sakiniai vyrauja pasirinktuose natūralios kalbos dialoguose. Šis tyrimas reikalingas įvertinti, ar dialoguose vyrauja pakankami ilgi sakiniai, kurie leis nustatyti bigramų ir trigramų įverčius.

I LENTELĖ. SAKINIŲ ILGIS IR JŲ SKAIČIUS PROCENTAIS PAGAL JUOSE PANAUDOTŲ ŽODŽIŲ SKAIČIŲ

Žodžių skaičius sakinyje	Sakinių skaičius procentais
1 žodis	8,09 %
2 žodžiai	14,04 %
3 žodžiai	14,47 %
4 žodžiai	14,89 %
5 žodžiai	11,06 %
6 žodžiai	13,19 %
7 žodžiai	4,68 %
8 žodžiai	5,53 %
9 žodžiai	2,98 %
10 žodžių	2,98 %
11 žodžių	1,70 %
12 žodžių	2,98 %
13 žodžių	0,85 %
14 žodžių	2,13 %
15 žodžių	0,43 %

Surinktų dialogų įrašų analizė parodė, kad šiuose dialoguose vyrauja įvairaus ilgio sakiniai (nuo vieno iki penkiolikos žodžių sakinyje). Iš lentelėje pateiktų duomenų matyti, kad daugiau nei po 14 proc. visų sakinių yra sudaryti iš dviejų, trijų ir keturių žodžių. Taip pat daugiau nei 11 proc. sakinių sudaryti iš penkių žodžių, o daugiau nei 13 proc. – iš šešių žodžių. Likusių rūšių sakinių, pagal juose vartojamų žodžių skaičių, yra mažiau nei po 10 proc.

Vertinant teksto dialogų žodžių žodyną, buvo atsižvelgta į žodžių pasikartojimą. Atlikus skaičiavimus paaiškėjo, kad įvertinus žodžių pasikartojimus žodyno imtis sumažėjo beveik per pusę.



3 pav. Pateiktų atsakymų ilgis pagal juose pavartotų žodžių skaičių

Įvertinant pasikartojimus galima pastebėti, kad daugiau nei 40 proc. žodžių kartojasi, o likusioji dalis, kurią sudaro beveik 60 proc., nesikartoja.

Kitas bandomasis tyrimas buvo atliktas siekiant įvertinti įvairių negramatinių įvykių skaičių dialoguose. Kadangi negramatinių įvykių įvairovė yra labai didelė, buvo pasirinkti keli lengviau identifikuojami negramatiniai įvykiai – pasikartojimų įvykiai, prieštaravimų įvykiai, žargonizmai surinktuose natūralios kalbos dialoguose.

II LENTELĖ. NEGRAMATINIŲ ĮVYKIŲ SKAIČIUS IR PROCENTAS DIALOGUOSE

Negramatinis įvykis	Negramatinių įvykių skaičius	Skaičius procentais
Pasikartojimas	4	2 %
Prieštaravimas	16	8 %
Žargonizmai	17	4 %

Pagal pradinius duomenis gauti rezultatai rodo, kad apie 14 proc. žodžių, aptinkamų realioje šnekamojoje kalboje, yra negramatiniai reiškiniai (pagal kalbos gramatikos taisykles jų neturėtų būti). Šiuo bandomuoju tyrimu siekiama ne gauti statistiškai patikimą rezultatą, o numatyti galimas tendencijas.

Dar vieno bandomojo tyrimo metu buvo atlikti pradiniai skaičiavimai trims pagrindiniams n-gramų tipams: unigramai, bigramai, trigramai įvertinti. Įvertintos dažnausiai pasikartojančios žodžių kombinacijos ir jų tikimybės.

Atlikus unigramų skaičiavimus, gauti duomenys leido įvertinti žodžių dažnumą dialogo modeliuose. Dažniausiai pasitaikantys žodžiai: „ir“ (46 k.), „su“ (26 k.), „norėčiau“ (23 k.), „noriu“ (19 k.), „man“ (18 k.) ir kt.

Buvo atliktas bigramų skaičiavimas, formuojant 2 žodžių kombinaciją. Pavyzdžiui, žodžių kombinacija „Aš norėčiau” buvo panaudota 6 kartus, o tikimybė lygi 0,462. Žodžių kombinacija „Man prašyčiau” buvo panaudota 5 kartus, tikimybė 0,278.

Kitame žingsnyje buvo atliktas trigramų skaičiavimas, formuojant 3 žodžių kombinaciją. Kadangi trigramos, kaip ir bigramos, atsižvelgia į ankstesnių žodžių pasirodymus, dėl to pagal gautus pradinius duomenis galima įvertinti tiek žodžių dažnumą, tiek tikimybes teksto dialoguose. Pavyzdžiui, žodžių kombinacija „Pas jus galima” buvo panaudota 3 kartus, o tikimybė lygi 0,375.

Kadangi surinktų ir apdorotų dialogų kol kas nėra pakankamai daug, nėra galimybės pateikti išsamesnių rezultatų. Kai bus surinktas planuojamas dialogų skaičius, bus atliekama detalesnė dažniausiai pasitaikančių žodžių porų ir žodžių trejetų analizė, kurios metu bus bandoma nustatyti tikslius jų įverčius, kurie bus palyginti su didžiojo lietuviško tekstyno, sukaupto VDU, nustatytais tų pačių žodžių bigramų ir trigramų įverčiais.

VIII. IŠVADOS

Pastaruoju metu pastebima, kad, siekiant tiksliau nustatyti šnekamosios kalbos sintaksinių apribojimų modelių parametrus, reikia rinkti specialius tekstus ir kurti kalbos modelius, remiantis juose esančiais sintaksiniais ir prasminiais žodžių ryšiais.

Tyrimams atlikti buvo surinktas ir apdorotas pirminis dialogų rinkinys lietuvių kalba. Bandomųjų tyrimų metu buvo išanalizuota kokio ilgio sakiniai vyrauja surinktuose dialoguose. Taip pat įvertinti žodžių pasikartojimai, atsižvelgiant į esamą žodyno dydį.

Buvo atliktas tyrimas, kurio metu buvo įvertintas negramatinių įvykių skaičius. Negramatiniai reiškiniai buvo vertinami atsižvelgiant į pasikartojimų ir prieštaravimų įvykius bei žargonizmus. Remiantis gautais duomenimis galima teigti, kad daugiau apie 14 proc. realiojoje šnekoje pasitaikančių įvykių yra vadinamieji negramatiniai įvykiai.

Dar vienas tyrimas buvo atliktas siekiant įvertinti skirtingus n-gramų tipus. Buvo įvertinamas žodžių dažnumas ir įverčiai surinktuose dialoguose, kurie vėliau turės būti santykinai palyginti su didžiojo lietuviško tekstyno, sukaupto VDU, bigramomis ir trigramomis.

Toliau bus siekiama surinkti dar daugiau natūralios kalbos dialogų, didinti žodyną ir gauti statistiškai patikimesnius rezultatus.

Taip pat planuojama papildomai apdoroti atnaujintą žodyną, žodžiu laikant tai, kas sudaryta iš tos pačios šaknies, t. y. pavyzdžiui, žodžiai „noriu“ ir „norėčiau“ bus vertinami kaip identiški. Pagal tokio žodyno struktūrą bus atlikti panašūs skaičiavimai.

Vėliau pagal gautus rezultatus bus įvertinamas modelių efektyvumas šnekamosios kalbos atpažinimo sistemose.

IX. LITERATŪRA

[1] J. Benesty, M. M. Sondhi, Y. Huang, „Springer Handbook of Speech Processing. Natural Language Understanding”, pp. 617, ISBN: 978-3-540-49125-5. 2008.

[2] C. Munteanu, M. Boldea, „A Description Language for Dialog Modeling and Management,“ Proceedings of the Fourth International Conference on Technical Informatics – CONTI, 2000.

[3] V. Pallota, „Computational Dialogue Models”, Multimodal Dialogue Management. pp. 12-18, 2003.

[4] J. Holmes, W. Holmes, “Speech Synthesis and Recognition,” pp. 196 – 20, ISBN 0-203-48468-1 Master e-book ISBN. 2003.

[5] I. Oparin, „Statistical Language Models for Dialogue Systems”, St.Petersburg State University (Russia), Dept. of Phonetics, University of West Bohemia (Czech Rep.), Dept. of Computer Science and Engineering, Speech Technology Center (Russia), 2006.

[6] T. Liu, „Analysis of Smoothing Techniques in Language Model Synthesis”. University of California at Berkeley.

[7] A. Vaičiūnas, „Lietuvių kalbos statistinių modelių ir jų taikymo šnekos atpažinimui tyrimas, kai naudojami labai dideli žodynai“, Daktaro disertacija, pp. 27–30, Kaunas: VDU, 2006.

[8] Y. Gotoh, S. Renals, „Statistical Language Modelling”, To appear in: Text and Speech Triggered Information Access, S. Renals, and G.Grefenstette (eds.), Springer, 2003.

[9] T. Schultz, K. Kirchhoff, „Multilingual Speech Processing,“ ISBN 13:978-0-12-088501-5, Academic Press, 2006.

[10] Ch. Maning, P. Raghavan, H. Schutze, „Introduction to Information Retrieval”, ISBN 978-0-521-86571-5, Camridge University Press, 2008.

[11] B. Moore, L. Tyler, W. Maslen-Wilson, „The Perception of Speech: from Sound To Meaning”, ISBN 978-0-19-956131-5, OUP Press, 2009.

INFORMACINĖS TECHNOLOGIJOS ● IVUS 2013 ● ISSN 2029-249X ● eISSN 2029-4824

Saityno informacinių sistemų vartotojo sąsajos testavimo metodika

Edita Šarapnickytė
Informacijos sistemų katedra
Informatikos fakultetas
Kauno technologijos universitetas
Kaunas, Lietuva
El. paštas: edita.sarapnickyte@stud.ktu.lt

Kęstutis Kapočius
[1]Informacijos sistemų katedra
[2]Informacinių sistemų projektavimo technologijų centras
Informatikos fakultetas
Kauno technologijos universitetas
Kaunas, Lietuva
El. paštas: kestutis.kapocius@ktu.lt

Santrauka. **Svarbus veiksnys, lemiantis vartotojų požiūrį į sistemą ir jos efektyvumą – tai aukštas sistemos panaudojamumas, kurį ženkliai veikia vartotojo sąsajos charakteristikos. Dėl milžiniškos konkurencijos pranašumas panaudojamumo srityje ypač aktualus saityne, todėl akivaizdu, jog saityno informacinių sistemų sąsają tikslinga nuodugniai testuoti. Vis dėlto atlikus žinomų testavimo būdų analizę nustatyta, jog tokiam saityno sistemų sąsajos testavimui keliami reikalavimai gali būti tenkinami tik iš dalies. Sprendžiant šią problemą buvo sukurtas specializuotas saityno IS vartotojo sąsajai testuoti pritaikytas sprendimas, jungiantis ir pagal poreikį praplečiantis egzistuojančius metodus. Atlikta preliminari sukurtosios metodikos taikymo galimybių analizė leidžia daryti prielaidą, jog metodika turėtų leisti ne tik aptikti daugiau ir esminių vartotojo sąsajos klaidų, bet ir operatyviai priimti sprendimus dėl aptiktų defektų šalinimo.**

Reikšminiai žodžiai: testavimas, grafinė vartotojo sąsaja, informacinės sistemos.

I. ĮVADAS

Vis didesnę įtaką organizacijų veiklos efektyvumui ir populiarumui darančių saityno informacinių sistemų (IS) efektyvumas ir kokybė tampa vis aktualesni ir vertinami ne tik informacinių sistemų savininkų, jų kūrėjų, bet ir eilinių vartotojų. Nesunku nuspėti, jog tai lemia milžiniška saityne egzistuojanti konkurencija ir galimybė vienu pelės paspaudimu iš nepatikusios svetainės persikelti į patrauklesnę konkurentų sistemą. Ypač svarbus veiksnys, darantis įtaką vartotojų požiūriui į sistemą ir jos efektyvumą – tai didelis sistemos panaudojamumas, kurį labai veikia (grafinės) vartotojo sąsajos (GVS) charakteristikos. Mažas panaudojamumas gali lemti tai, jog sistema bus tiesiog ignoruojama vartotojų, todėl efektyvus saityno informacinių sistemų vartotojo sąsajos testavimas yra ne mažiau svarbus nei funkcinis šių sistemų testavimas.

Testuojant tokio tipo informacines sistemas, taikyti tik funkcinio testavimo metodikas negana, tačiau žinomi tradicinės programinės įrangos vartotojo sąsajos testavimo būdai nėra visiškai tinkami testuoti IS sąsajai [2]. Saityno informacinių sistemų testavimas apskritai traktuotinas kaip besiformuojanti, neišplėtota informatikos sritis. Nors saityno IS vartotojo sąsajai testuoti galima pasirinkti šiuo metu žinomus vartotojo sąsajos testavimo ir vertinimo metodus, atlikta analizė atskleidė, kad šiam uždaviniui įgyvendinti taikant tradicinius testavimo būdus testavimo procesui keliami reikalavimai bus tenkinami tik iš dalies. Atkreiptinas dėmesys į tai, jog saityno svetainių ir informacinių sistemų sąsaja pasižymi keletu darbo būdo (saityne *naršoma*) ir techninės aplinkos lemiamų ypatybių [12]. Dėl to formuojasi specifinė, tik tokio pobūdžio sistemoms būdingų sąsajos defektų ir galimų jų sprendimo būdų aibė [7] [8].

Sukūrus specializuotas, saityno informacinių sistemų vartotojo sąsajai testuoti pritaikytas priemones, pavyktų užtikrinti aukštą produkto kokybę, išvengiant dažniausiai pasitaikančių vartotojo sąsajos klaidų ir trūkumų. Straipsnyje pristatoma būtent tokia metodika, skirta testuoti saityno informacinių sistemų vartotojo sąsają. Metodika grindžiama standartiniu testavimo procesu, rankinio testavimo būdu ir euristinio vertinimo metodu. Vadovaujantis sukurta metodika ne tik galima tikėtis aptikti daugiau ir esminių vartotojo sąsajos klaidų, bet ir gauti parengtus defektų šalinimo sprendimus, kurie gali būti panaudoti dar kartą arba papildyti.

Antroje straipsnio dalyje apžvelgiami su šia tema susiję sprendimai, aptariami testavimo tipų, metodų privalumai ir trūkumai. Čia taip pat pateikiama testavimo procesų bei priemonių, kurias galima taikyti testavimo metu, apžvalga. Trečia straipsnio dalis skirta siūlomai saityno informacinės sistemos grafinės vartotojo sąsajos testavimo metodikai pristatyti. Ketvirtoje straipsnio dalyje pateikiamos metodikos įgyvendinimo ir taikymo galimybės. Straipsnis baigiamas išvadomis.

II. ŽINOMŲ VARTOTOJO SĄSAJOS TESTAVIMO SPRENDIMŲ ĮVERTINIMAS

Testavimas – tai bet kokia veikla, kurios tikslas yra įvertinti programinės sistemos charakteristikas ar galimybes bei nustatyti, ar ji tinkama reikalaujamiems rezultatams gauti [11]. Tarptautinio programinės įrangos testavimo standarto apibrėžtame testavimo proceso modelyje vienas iš esminių testavimo proceso lygių yra dinaminiai testavimo procesai [5].

Dinaminių testavimo procesų lygis yra sudėtinis, jį sudaro keturi žemesnio lygio procesai:

- Testo kūrimas.

- Testavimo aplinkos parinkimas ir konfigūravimas.
- Testo vykdymas.
- Testavimo rezultatų fiksavimas.

Šių sudėtinių testavimo procesų vykdymas ir veiklos sudėtis priklauso nuo pasirinktos testavimo technikos, metodų ir priemonių, kurios gali būti parenkamos atsižvelgiant į testuojamą objektą ir siekiamus rezultatus.

Sistemų vartotojo sąsajai testuoti reikalinga unikali metodika, tačiau jos įgyvendinimas turi būti grindžiamas tradicinės programinės įrangos funkcinio testavimo eiga [2]. Galima daryti prielaidą, jog, remiantis standartiniu programinės įrangos testavimo proceso modeliu, gali būti atliktas ir saityno informacinės sistemos tiek funkcinis, tiek vartotojo sąsajos ir panaudojamumo testavimas, tačiau būtina šį procesą pritaikyti ir detaliau apibrėžti. Specializavus procesą saityno informacinės sistemos grafinės vartotojo sąsajos testavimui, bus sukurtos prielaidos tokio tipo sistemų testavimo efektyvumui (rastų defektų skaičius/esamų defektų skaičius) ir produktyvumui (defektų skaičiaus radimas per kuo trumpesnį laiką) padidinti.

Specializuojant testavimo procesą, gali būti parenkamos skirtingos testavimo priemonės, būdai ir technikos, atsižvelgiant į tai, ką siekiama pasiekti testavimu. Taigi, būtina gerai suprasti prieinamus sprendimus ir jų privalumus bei trūkumus. Egzistuojančias programinės įrangos testavimo metodikas pagal testavimo vykdymo būdą galima suskirstyti į dvi pagrindines grupes:

1) Rankinis testavimas. Seniausiai taikomas, tradicinis ir labiausiai nuo testuotojo patirties ir asmeninių savybių priklausantis testavimo būdas [1] [10]. Rankiniu būdu dažniausiai testuojama siekiant aptikti sistemos vartojimo patogumo klaidas, nes pagrindinės ir kritiškiausios vartotojo sąsajos klaidos randamos tik tuomet, kai testuoja realūs žmonės.

2) Automatinis testavimas. Testavimo būdas, kai testavimo atvejams vykdyti naudojamos programos bei įrankiai, skirti programų ar modulių, kurių kokybę norima įvertinti, darbui patikrinti [10].

Siekiant palyginti ir įvertinti, kuris metodas gali būti tinkamesnis saityno informacinės sistemos grafinei vartotojo sąsajai testuoti, galima pasinaudoti K. Smilgytės ir J. Nenortaitės (2011) atlikta analize [9]. Autorės palygino rankinio ir automatinio testavimo būdų privalumus ir trūkumus pagal tokius pagrindinius kriterijus: kainos, testų parengimo ir vykdymo laiko, pritaikymo prie besikeičiančių reikalavimų, išmokstamumo, testavimo aprėpties ir klaidų aptikimo.

Nors rankinių testavimo metodų taikymas testuoti grafinei vartotojo sąsajai yra lėtas, testavimo atvejų padengiamumas gali būti didesnis, nei taikant automatinius testus. Automatiniai testai labiau tinka testuoti paviršutiniškai, o ne identifikuoti kritines ir specifines klaidas [1]. Automatinis testavimas negali pakeisti rankinio testavimo metu taikomo analitinio mąstymo. Taip pat svarbu įvertinti ir tai, kad atlikus automatinį testavimą vis vien būtina analizė ir sprendimų priėmimas, kuris testuojant rankiniu būdu yra atliekamas iš karto. Taigi, tikslinga testavimo etape rinktis rankinius testus, kartais kombinuojant juos su automatiniais, didelio įdirbio nereikalaujančiais testais.

Automatiniai testai padeda aptikti ir eliminuoti tik esmines klaidas, pasitaikančias kūrimo etape. Pagrindinės ir kritiškiausios grafinės vartotojo sąsajos klaidos randamos testuojant rankiniu būdu. Rankinio testavimo metodus galima taikyti jau ankstyvose sistemos kūrimo stadijose, kai vartotojo sąsaja dar nėra pakankamai stabili, automatiniai testai dar nėra veiksnūs ir jų vykdymas neįmanomas arba neturi prasmės [3][10]. Nereikia pamiršti, jog defektų aptikimas ankstyvose kūrimo stadijose padeda sumažinti jų sprendimo sąnaudas [4].

Analizuojant automatinių testavimo metodų greitį, būtina pažymėti, jog šis pranašumas prieš rankinius metodus išryškėja tik testų vykdymo etape. Testavimo automatizavimas leidžia ženkliai padidinti testų skaičių panaudojant testus keletą kartų, vykdant juos reguliariai. Tai nereikalauja papildomų išteklių, tačiau įvertinant reikalavimų pasikeitimus testus kaskart tenka koreguoti, o jiems parengti reikia laiko. Papildomas pasiruošimas ir žinios reikalingos ne tik pirmą kartą kuriant testus, bet ir nuolat juos atnaujinant ir tobulinant. Automatiniams testams kurti būtinos išskirtinės programavimo ir modeliavimo žinios [10], todėl tokio tipo metodų taikymas neretai pareikalauja papildomų išteklių ir gali sukelti diskomfortą testavimo komandai.

Testavimo įrankiai, naudojami įvertinti sistemos panaudojamumui skirtingais būdais, pagal taikymo principus gali būti suskirstyti į keturias grupes [12]:

- stebėjimo;
- palyginimo;
- klausymo;
- įvertinimo.

Atsižvelgiant į tai, kad defektų aptikimas ir jų pataisymas ankstyvose kūrimo stadijose padeda sumažinti sistemos kūrimo išlaidas, galima išskirti tokius pagrindinius sąsajos testavimo ir įvertinimo metodus, taikytinus nuo ankstyvos testavimo stadijos [12]:

- Pažintinė peržvalga.
- Atitikties rekomendacijoms ir standartams vertinimas (formalus).
- Euristinis vertinimas.

Šie metodai nereikalauja pašalinių vertintojų dalyvavimo ir gali būti atlikti pačių kūrėjų (ar kuriančios kompanijos testuotojų) rankiniu būdu. Tai lemia santykinai neaukštą tokio testavimo kainą.

Pažintinė peržvalga – tai vartotojo sąsajos įvertinimo metodas, kurio centre yra specifikuojant reikalavimus apibrėžtų vartotojų užduočių vykdymas [12]. Šis būdas leidžia kuo anksčiau gauti dizaino sprendimo įvertinimus [3] ir jau projektavimo etape įvertinti specifikuotas užduočių vykdymo sekas. Kuo anksčiau dizaino klaidos aptinkamos, tuo lengviau jas spręsti. Vis dėlto taikant metodą ankstyvose sistemos kūrimo stadijose, naudojant grafinius ar programinius sąsajos prototipus, neatsižvelgiama į galimus interaktyvius vartotojo veiksmus [3]. To-

dėl iškyla poreikis imituoti interaktyvius procesus, įtraukiant ir specifines užduotis. Toks pasiruošimas testavimui ir testavimas vadovaujantis formaliu metodo aprašu reikalauja intensyvaus darbo. O taikant tik supaprastintas priemones gaunama paviršutiniška ir siaura analizė, apimanti tik teksto ir grafinio dizaino, sąsajos elementų išdėstymo problemas.

Atitikties rekomendacijoms ir standartinio vertinimo metodas – tai sąsajos peržiūra, kuri atliekama vadovaujantis iš anksto griežtai apibrėžtomis rekomendacijomis. Rekomendacijose labai tiksliai nurodomi reikalavimai arba taisyklės, kurių reikėtų laikytis kuriant konkrečius sąsajos elementus [12]. Vadovaujantis šiomis rekomendacijomis, aptinkamos bendrinės ir dažniausiai pasikartojančios klaidos, taip pat įvertinami ne tik teksto ir grafinių vaizdų defektai, bet ir dizaino ar pačių objektų išdėstymo problemos. Defektų šalinimo sprendimų priėmimas taip pat vykdomas vadovaujantis rekomendacijomis. Vis dėlto ne visos situacijos gali būti numatytos ar apibrėžtos rekomendacijose, tarptautiniuose arba įmonių naudojamuose standartų rinkiniuose.

Euristinis vertinimas – tai kruopštus sistemos vertinimas, kurį atlieka vartotojo sąsajos specialistų ar kvalifikuotų IS kūrimo komandos nariai [12]. Taikant šį metodą vertinama pagal asmeninę patirtį, bet dažniau pagal tai, kaip sistema atitinka žinomus panaudojamumo principų rinkinius – euristikas. Euristikos dažniausiai pateikiamos kaip atvirų kontrolinių klausimų ar pasiūlymų sąrašas. Šis sąrašas nėra veiksmų, kuriuos turi atlikti testuotojas, sąrašas; jis naudojamas ne kontroliuoti atliekamiems veiksmams, o leidžia įvertinti daugiau galimų ir įdomių problemų aspektų [6]. Tokiu būdu gali būti aptikta daug įvairaus tipo defektų, įskaitant ir panaudojamumo problemas, tačiau neretai gali būti sunku aptikti globalias sistemos problemas ir trūkstamus sąsajos elementus. Vadovaujantis euristikomis ne visada lengva spręsti iškilusias problemas, todėl problemų sprendimų paieška neretai tampa dideliu kūrybiniu iššūkiu testavimo, analitikų ar programuotojų komandai.

Apibendrinti testavimo metodų ypatumų analizės rezultatai pateikiami 1 lentelėje.

Remiantis atlikta analize galima daryti prielaidą, jog egzistuojantys metodai netenkina visų saityno informacinės sistemos grafinės vartotojo sąsajos testavimo procesui keliamų reikalavimų. Idealiu atveju testavimo metodas turėtų ne tik leisti aptikti esminius saityno informacinės sistemos grafinės vartotojo sąsajos defektus, bet ir padėti juos efektyviai šalinti. Tokį sprendimą galima tikėtis sukurti sujungus ir pagal poreikį praplėtus egzistuojančius metodus.

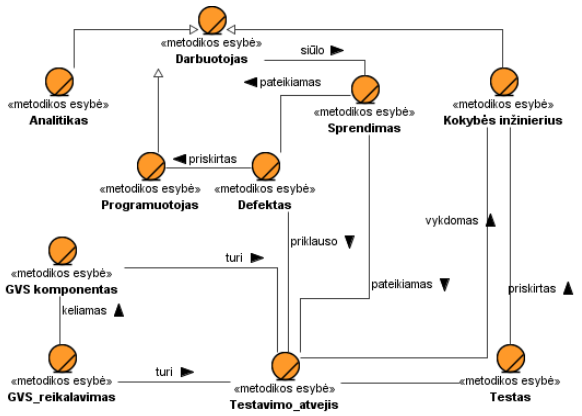
I LENTELĖ. TESTAVIMO METODŲ PALYGINIMAS

	Pažintinė peržvalga	**Atitikties rekomendacijoms ir standartams vertinimas**	**Euristinis vertinimas**
Testavimo atvejai	Užduotys, kurias reikia atlikti	Reikalavimų pagrindu sukurtos rekomendacijos arba taisyklės	Euristikos pagal žinomus panaudojamumo principus
Neaptinkamos klaidos	Bendros ir dažnos klaidos	Rimtos koncepcinės, navigacijos ir operacinės klaidos	Globalios problemos ir trūkstami sąsajos elementai
Aptinkamos	Išskirtos problemos gali būti ne esminės	Bendrinės ir pasikartojančios objektų išdėstymo ir dizaino klaidos	Daug įvairių, įskaitant ir panaudojamumo klaidas
Problemų sprendimai	Nėra (sprendimus priima projektuotojai)	Yra (vadovaujamasi priimtomis rekomendacijomis ir standartais)	Ne visada apibrėžtas būdas problemų sprendimams gauti

III. SIŪLOMA SAITYNO IS VARTOTOJO SĄSAJOS TESTAVIMO METODIKA

Kuriant čia pristatomą metodiką buvo remiamasi tarptautinio programinės įrangos testavimo standarto apibrėžtu testavimo proceso modeliu, jis detalizuotas ir pritaikytas pasirinktam testavimo tipui ir metodui. Kaip jau buvo minėta, šio tyrimo metu buvo siekiama sukurti sprendimą, kuris būtų specializuotas būtent saityno informacinių sistemų (grafinei) vartotojo sąsajai testuoti. Pasirinktas rankinio testavimo principas, kuriam būdingas dideles testavimo laiko sąnaudas buvo siekiama sumažinti. Metodika kurta euristinio vertinimo metodo pagrindu, tačiau sprendžiant šiam metodui būdingas trūkstamų sąsajos elementų nepastebimumo ir sprendimų priėmimo problemas.

Sudarant teorinį testavimo metodikos modelį (1 pav.) išskirti pagrindiniai konceptai ir jų tarpusavio ryšiai saityno informacinės sistemos grafinės vartotojo sąsajos testavimo metodikoje.



1 pav. Koncepcinis testavimo metodikos modelis

Atlikus testavimo proceso dalyvių analizę išskirti programuotojo, analitiko ir kokybės inžinieriaus konceptai. Šių konceptų ryšys su kitais konceptais identifikuotas pagal testavimo proceso dalyvių atsakomybių sritis ir jų vykdomas užduotis.

Vienas iš dažniausiai nurodomų euristinio vertinimo metodo trūkumų yra tai, kad nepateikiami defektų šalinimo sprendimų būdai. Čia pristatomoje metodikoje numatoma, kad kuriamas testavimo atvejis (euristika, TA) arba testavimo atvejo defektas (TD) gali turėti sprendimą. Kiekvieną kartą iš naujo panaudojant universalias euristikas skirtingoms vartotojo sąsajoms testuoti, aptikus defektą galima pritaikyti anksčiau taikytą

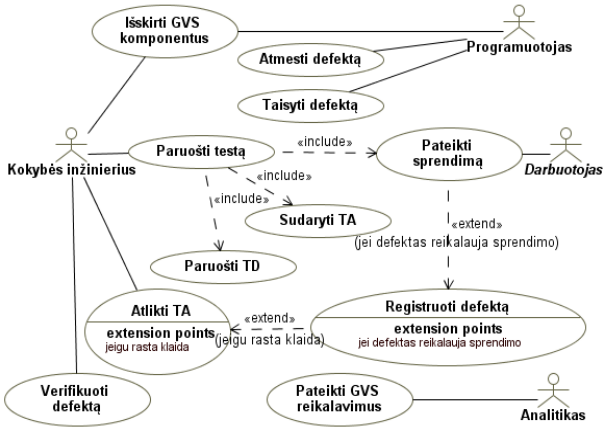
konkretaus defekto šalinimo sprendimą, jį modifikuoti arba pridėti naują.

Kitas euristinio metodo trūkumas – testuojant nepastebimi trūkstami vartotojo sąsajos elementai. Siūloma metodika grindžiama grafinės vartotojo sąsajos išskaidymu į komponentus. Sudaromus testavimo atvejus (euristikas) siūloma susieti su vartotojo sąsajos komponentais. Tokio ryšio užmezgimas leidžia ne tik testuoti kiekvieną komponentą, bet ir kaupti jiems pritaikytus testavimo atvejus.

Metodikos taikymas iliustruojamas panaudojimo atvejų modeliu (2 pav.), kuriame matyti tipiniai metodikos vartotojai ir metodikos panaudojimas.

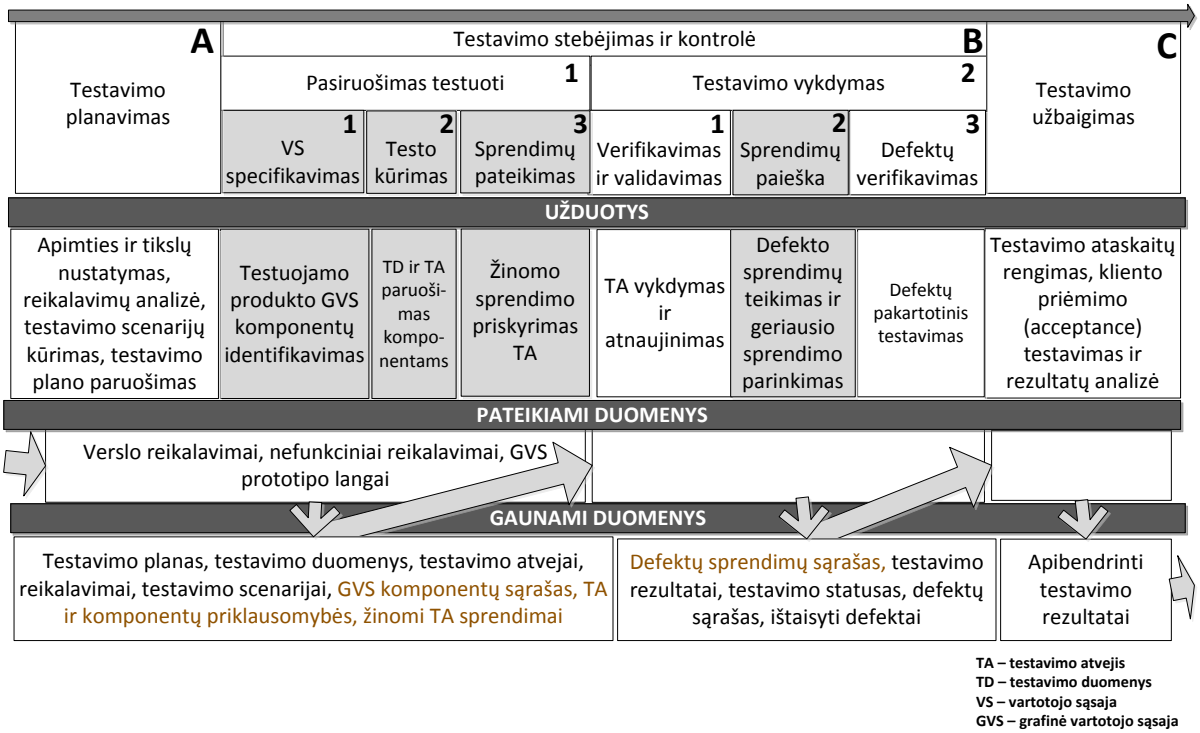
Remiantis tarptautinio programinės įrangos testavimo standarto „ISO/IEC 29199 Software testing“ [5] siūlomu testavimo proceso modeliu ir išlaikant jo struktūrą, išskirti panaudojimo atvejai buvo suskirstyti į grafinės vartotojo sąsajos testavimo metodikos etapus (3 pav.). Be to, procesai detalizuoti, išskaidant juos į smulkesnius procesus ir pritaikant saityno informacinių sistemų vartotojo sąsajai testuoti.

Sudarytas vartotojo sąsajos testavimo metodikos etapų modelis, palyginti su tipiniu programinės įrangos testavimo procesų modeliu, išlaiko nepakitusį testavimo valdymo procesų lygį, t. y. tokį, koks yra apibrėžtas tarptautiniame programinės įrangos testavimo standarte. Testavimo valdymo procesų lygį sudaro testavimo planavimas (A), testavimo stebėjimas ir kontrolė (B) bei testavimo užbaigimas (C).



2 pav. Metodikos taikymo veiklos

Dinaminis testavimo proceso lygis, kurį standartiškai sudaro testo kūrimo (1.2), testo vykdymo (2.1) ir defektų registravimo etapai (2.3), metodikoje papildytas trimis saityno informacinės sistemos grafinės vartotojo sąsajai testuoti pritaikytais, specializuotais etapais (3 pav. pažymėti pilka spalva). Tai vartotojo sąsajos analizė (1.1), sprendimų pateikimo (1.3) ir sprendimų paieškos (2.2) etapai.



3 pav. Vartotojo sąsajos testavimo metodikos etapai (pilkesni elementai – naujai įvesti etapai, pritaikyti vartotojo sąsajai testuoti)

Taigi, siūlomas vartotojo sąsajos testavimo procesas susideda iš keturių etapų:

A. Testavimo planavimas

B. Pasiruošimas testuoti

- Vartotojo sąsajos analizė.
- Testų kūrimas.
- Sprendimų pateikimas.

C. Testavimo vykdymas

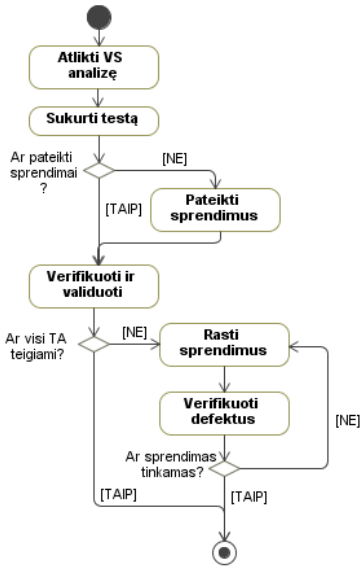
- Validavimas ir verifikavimas.
- Sprendimų paieška.
- Defektų verifikavimas.

D. Testavimo užbaigimas (schemoje nerodomas)

Dinaminiai sudarytosios metodikos etapai – tai pasiruošimas testuoti ir testavimo vykdymas. Šie etapai buvo pritaikyti grafinės vartotojo sąsajos testavimui.

Pasiruošimo ir testavimo vykdymo subetapų tvarka nėra griežtai apibrėžta. Etapai gali iš dalies sutapti, būti vykdomi kelis kartus pakartotinai pagrindiniame etape arba, nesant būtinybės, praleidžiami. Pasiruošimo testuoti subetapai gali būti vykdomi lygiagrečiai, tačiau tikslingiausia būtų juos pirmiausiai pradėti nuo vartotojo sąsajos analizės, o vykdant testų kūrimo etapą pateikti ir žinomus sprendimus.

Taigi, apibendrintas saityno IS vartotojo sąsajos testavimo procesas, nevertinant planavimo ir užbaigimo, susideda iš šešių žingsnių (žr. 4 pav.). Kaip matome, identifikuoti trys sprendimo taškai – ar pateikti sprendimai, ar visi TA teigiami ir ar sprendimas tinkamas. Šie sprendimo taškai lemia, ar visi grafinės vartotojo sąsajos testavimo metodikos etapai ir procesai bus vykdomi ir kiek kartų jie bus vykdomi.

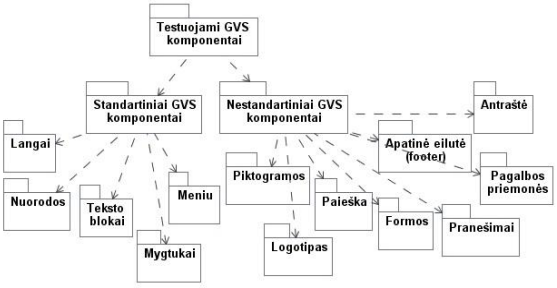


4 pav. Bendrinis testavimo metodikos procesas

IV. METODIKOS ĮGYVENDINIMO IR TAIKYMO GALIMYBĖS

Vienas iš pirmųjų ir svarbiausių pasiruošimo etapų, numatytų sudarytoje saityno grafinės vartotojo sąsajos testavimo metodikoje, yra vartotojo sąsajos analizės etapas. Šio etapo metu pagal turimus dizaino šablonus ar prototipus yra sudaroma saityno grafinės vartotojo sąsajos komponentų aibė. Pagal šią aibę nusprendžiama, kokie testavimo atvejai bus reikalingi testuojamos saityno informacinės sistemos grafinei vartotojo sąsajai.

Saityno informacinės sistemos skiriasi savo struktūra ir turiniu, tačiau pasitaiko standartinių, daugelio saityno IS vartotojo sąsajai būdingų komponentų. Tyrimo metu rekomendacinis universalus standartinių komponentų sąrašas (žr. 5 pav.) sudarytas atsižvelgiant į dažniausiai pasitaikančių saityno informacinių sistemų vartotojo sąsajos klaidų grupes, apibrėžtas J. Johnsono [7] [8]. Klaidų grupės turi joms būdingą klaidų pasitaikymo vietą saityno informacinėse sistemose, kurios gali būti traktuojamos kaip komponentai.



5 pav. Rekomenduojama komponentų aibė

Kita stambi komponentų grupė yra nestandartiniai grafinės vartotojo sąsajos komponentai. Tokie komponentai gali būti būdingi ne visų saityno informacinių sistemų grafinei vartotojo sąsajai. Siekiant sudaryti tokį specifinių komponentų sąrašą yra tikslinga pasirinkti tikslinį metodikos sprendimo poaibį.

Rekomenduojamas standartinių ir nestandartinių komponentų rinkinys (žr. 5 pav.) gali būti panaudotas kaip aibė, su kuria galima dirbti pagal sukurtą metodiką.

Išskirtiems grafinės vartotojo sąsajos komponentams turi būti sudaryti ir priskirti testavimo atvejai (euristikos). Statiniai (universalūs) testavimo atvejai turi būti įtraukti į duomenų bazę kaip tipiniai testavimo atvejai, tinkami visoms sistemoms. Prasidėjus naujam projektui, jeigu reikalinga, testuojamam produktui būdingi dinaminiai testavimo atvejai gali būti numatomi ir įtraukiami testų kūrimo etape.

V. IŠVADOS

Vartotojo sąsajos testavimo ir vertinimo metodų analizė parodė, kad šiuo metu vyraujantys metodai nėra pakankamai tinkami saityno informacinių sistemų vartotojo sąsajai testuoti. Vis dėlto žinomos metodikos pasižymi tam tikrais privalumais, kuriais verta pasinaudoti. Šio tyrimo metu buvo sukurta metodika, kurioje remiamasi testavimo standartais, panaudojamas rankinio testavimo principas bei pritaikytas euristinio vertinimo metodas.

Preliminarus metodikos tyrimas parodė, jog sprendimas gali būti realizuotas ir turėtų leisti pagerinti ne tik sąsajos defektų aptikimo efektyvumą, bet ir jų sprendimo greitį, pritaikytų pataisymų kokybę. Vienas svarbiausių sukurtosios metodikos ypatumų – galimybė kaupti žinias ne tik apie galimas saityno IS klaidas, bet ir apie šių klaidų taisymo galimybes.

Artimiausiuose planuose numatoma sukurti metodika grindžiamą testavimo sistemos prototipą. Sukurtąjį prototipą bus siekiama panaudoti metodikos tyrimui realiuose projektuose, o gauti rezultatai bus panaudoti tolimesniam metodikos vystymui.

LITERATŪRA

[1] Leitner, I. Ciupa, B. Meyer, M. Howard, „Reconciling Manual and Automated Testing: The Autotest Experience“, Proceedings of the 40th Hawaii International Conference on System Sciences (HICSS-40 2007), pp. 261-261a, 2007.

[2] M. Memon, „GUI Testing: Pitfalls and Process“, Computer, vol. 35, no. 8, pp. 87-88, 2002.

[3] Lewis, P. Polson, C. Wharton, J. Rieman. „Testing a Walkthrough Methodology for Theory-Based Design of Walk-Up-and-Use Interfaces“, CHI '90 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM New York, pp. 235-242, 1990.

[4] G. Mujtaba, T. Mahmood, Z. Nasir, „A Holistic Approach to Software Defect Analysis and Management“, Australian Journal of Basic and Applied Sciences, 5(6), pp. 1632-1640, 2011.

[5] ISO/IEC JTC1/SC7, ISO/IEC 29119 „Software Testing Standard“, prieiga per internetą: http://www.softwaretestingstandard.org/.

[6] J. Bach, „Heuristic Risk-Based Testing“, Software Testing and Quality Engineering Magazine, 11/1999, prieiga per internetą: http://nilachakra.org/documents/material/L%20-%20RiskAnalysis.pdf

[7] J. Johnson, „Web Bloopers: 60 Common Web Design Mistakes, and How to Avoid Them“. Morgan Kaufman, 2003.

[8] J. Johnson, „GUI Bloopers 2.0, Second Edition: Common User Interface Design Don'ts and Dos“. Morgan Kaufman, 2007.

[9] K. Smilgytė, J. Nenortaitė, „Dirbtiniais neuroniniais tinklais paremtas testavimo tipo pasirinkimo metodas“, Proceedings of the 15th Master and PhD conference “Information Society and University studies” (IVUS 2010), pp. 111-116, 2011.

[10] OTS Solutions, „Manual Testing v/s Automated Testing“, prieiga per internetą: http://www.slideshare.net/Softwarecentral/manual-testing-vs-automated-testing.

[11] W. C. Hetzel, „The Complete Guide to Software Testing“, 2nd ed. QED Information Sciences, 1988.

[12] W. O. Galitz, „The Essential Guide to User Interface Design“. Third Edition. Wiley Publishing, 2007.

Mažmeninės bankininkystės savitarnos terminalų paslaugų vertinimo modelis

Brigita Šilanskaitė
KHF Informatikos katedra
Vilniaus universitetas
Kaunas, Lietuva
brigita.silanskaite@khf.vu.lt

Darius Dilijonas
KHF Informatikos katedra
Vilniaus universitetas
Kaunas, Lietuva
darius.dilijonas@khf.vu.lt

***Santrauka.* — Straipsnyje yra pristatomas mažmeninės bankininkystės paslaugų optimizavimo modelis, sudarytas remiantis Lietuvos bankų patirtimi. Modelis sukurtas remiantis STTQUAL (savitarnos terminalų kokybės valdymo metodika), SERVQUAL (paslaugų kokybės valdymo metodika), Savitarnos paslaugų sistemų efektyvumo vertinimo ir tobulinimo bei Savitarnos paslaugų sistemų kokybės vertinimo kriterijų modeliais. Siūlomo modelio pagrindinės dalys: vartotojų paslaugų kokybės vertinimas, tiekėjų paslaugų vertinimas, ATM (automatinių kasininko pagalbininkų) efektyvumo vertinimas. ATM teikiamų paslaugų kokybei įvertinti naudojami vartotojų apklausos duomenys, efektyvumui įvertinti - išvestiniai veiklos rodikliai. Mažmeninės bankininkystės savitarnos terminalų paslaugų vertinimo modelyje siūloma derinti SOM (save organizuojančiais tinklais) ir ekspertinę sistemą.**

Reikšminiai žodžiai: savitarnos paslaugų kokybė, mažmeninė bankininkystė, ATM

I. ĮVADAS

M. Hensmanas pateikia Anglijos bankuose vyravusį požiūrį: „Klientas turi laimėti ne palūkanų dydžiu, bet paslaugų įvairove ir standartais (Clegg, 1969)“ [10]. Taikant naujausias technologijas finansinėms savitarnos paslaugoms teikti bankininkystė tapo tuo ką mes darome bet kur ir bet kada, o ne kur mes einame [12]. Paslaugų teikėjai, siekdami užtikrinti vientisą paslaugų teikimą, naudoja SST (angl. *self service technologies*) kombinacijas [5], bankuose dažniausiai naudojama: internetas, mobilieji įrenginiai, ATM. Klientai palankiai vertina jiems suteiktas galimybes naudotis paslaugomis patogiu metu ir būdu [5]. Bankomatai (savitarnos įrenginiai) mažmeninėje bankininkystėje pradėti naudoti prieš 40 metų Londone. Lietuvoje internetinė bankininkystė ir bankomatai pradėti diegti 2000 metais. ATM „Marketplace“ ataskaitoje [1] nurodo, kad ATM yra vienas iš pačių svarbiausių kanalų teikiant paslaugas klientams. Visame pasaulyje yra įdiegta apie 1,58 mln. ATM [1]. Didžiojoje Britanijoje yra įdiegta daugiau nei 60 000 ATM, JAV - 250 000 ATM, prognozuojama, kad iki 2017 metų bus įdiegta 3,4 mln. ATM.

Išsamiau tirti paslaugas pradėta aštuntame dešimtmetyje, tyrimų sritis yra nauja ir palyginti sudėtinga, dėl paslaugų gausos ir teikimo specifikos, kuri skiriasi nuo produktų. Paslaugos pelningumą rinkos ekonomikoje lemia jos kokybė ir vartotojo suvokiama nauda, todėl daugiausiai mokslinių tyrimų yra atliekama siekiant sukurti paslaugų kokybės modelį [2, 3]. Literatūroje yra išskiriama keletas paslaugų vertinimo modelių, kurie pateikia vertinimo modelius. Klasikiniai paslaugų kokybės modeliai [4, 5, 6, 23] ne iki galo vertina elektroninių paslaugų aspektus. Tiriant elektroninių paslaugų kokybę dažniausiai kuriamos vertinimo metodologijos pritaikytos „Web“ tinklalapiams [4, 5] (Gounaris ir kt. (2005), Yoo ir kt. (2001), Lociacono, Watson, Goodhue (2002) (WebQual)), mažmeninei e-prekybai ir e-parduotuvėms (Dabholkar (1996), Lee ir kt. (2005), Sohn ir kt. (2008), Janda, Trocchia ir kt. (2002), Collier (2006)). Vystantis naujoms SST technologijoms svarbu tirti, ko klientas tikisi, kai paslauga teikiama SST [5]. Pastaraisiais metais informacinių technologijų kaita turėjo didelę įtaką paslaugų teikimo procesams: atsirado galimybės efektyviau rinkti informaciją apie paslaugų teikimą, stebėti vartotojo elgseną, elektroninės paslaugos nuolat populiarėja ir tampa neatskiriama kasdieninio gyvenimo dalimi. Finansų srityje didžiausi pokyčiai įvyko sukūrus ATM. Šiuo metu bankomatai teikia daugialypes paslaugas: pinigų išgryninimo, elektroninių mokėjimų, išankstinių papildymų ir informacijos.

Mokslinėje literatūroje yra daug straipsnių, nagrinėjančių bankų paslaugų kokybę, mokslininkai pateikia bankų paslaugų vertinimus. K. O. Siddiqi (2011) tyrinėjo Bangladešo mažmeninės bankininkystės sektorių, A. Abedniya ir M. N. Zaeim (2011) pristatė paslaugų kokybės tyrimų rezultatus Islamiškoje bankininkystės sistemoje, K. H. Hanzaee ir S. A. S. Salehi (2011) ištyrė Irano mažmeninės bankininkystės rinką, kokybės suvokimo skirtumus islamiškuose ir tradiciniuose bankuose Pakistane tyrinėja H. D. Aslam ir kt. (2012) [6-9]. Dažnai yra vertinamos visos teikiamos banko paslaugos neišskiriant finansinių savitarnos paslaugų, tyrimai atliekami apklausiant bankų klientus, todėl paslaugos įvertinamos tik iš kliento pusės. Internetinės bankininkystės paslaugų kokybės kriterijus ištyrė M. Jun (2001), elektroninių finansinių paslaugų teikimo kanalus ir kokybės dimensijas išnagrinėjo O. T. Ombati ir kt. (2010) [11], ATM teikiamų paslaugų kokybės empirinį tyrimą Pakistano bankuose pristatė Khan (2010), savitarnos paslaugų sistemų kokybės vertinimo kriterijų modelį pristatė D. Dilijonas ir kt. (2009) [15]. D. Dilijonas ir kt. (2011) kitame tyrime pristatė efektyvumo vertinimo įrankius. C. Gronroos išnagrinėjo paslaugų produktyvumą, išskyrė išorinį, vidinį ir apimčių našumą.

Paslaugų kokybės vertinimo modeliai dažnai nurodo, ką reikia vertinti, bet nepateikia metodų, kaip tai įvertinti. Formuojant ir teikiant paslaugas reikia įvertinti savo paslaugų kokybės teikimo procesą. Vertinant paslaugas sunkiausia tai, kad paslaugos kokybės negalima įvertinti prieš ją parduodant klientui.

Straipsnio tikslas yra pristatyti sukurtą savitarnos sistemų paslaugų vertinimo modelį, kuris leidžia finansinių paslaugų institucijoms įvertinti savo teikiamų paslaugų kokybę ir suformuoti tvarią plėtros strategiją bei padidinti teikiamų paslaugų pelningumą.

Antrame skyriuje nagrinėjami paslaugų kokybės kriterijai bei paslaugų kokybės ir efektyvumo vertinimo modeliai. Trečiame skyriuje pristatoma siūloma paslaugų kokybės ir efektyvumo vertinimo metodologija. Ketvirtame skyriuje pateikiami eksperimentinio tyrimo rezultatai, penktame – straipsnio išvados.

II. SAVITARNOS PASLAUGOS IR JŲ VERTINIMO MODELIŲ ANALIZĖ

XXI a. ekonomikai įtaką darančios technologijos yra informacinės technologijos, nanotechnologijos, biotechnologijos [24]. Pagrindinis paslaugos skirtumas nuo produkto yra kliento dalyvavimas kūrimo procese. Informacinės technologijos leidžia mažinti sąnaudas, suteikia daugiau asmeniškumo, ir klientas gali visapusiškai kontroliuoti paslaugos kūrimą. Tokios paslaugos lėmė naujų informacinių technologijų atsiradimą: į paslaugas orientuota architektūra, „Web“ paslaugos ir paslaugų kompiuterija (angl. *service computing*) [2], „debesų kompiuterija“ (angl. *cloud computing*).

Paslauga yra veikla, tenkinanti vartotojo poreikius, neturi apčiuopiamo materialaus rezultato, procesas, kurio metu sąveikauja teikėjas ir vartotojas [22].

J. Nollet ir J. Haywood-Farmer (1992) teigia, kad paslaugos koncepcija jungia pagrindinę ir papildomas paslaugas, jų teikimo sistemą ir procesą. Teikėjams tokia koncepcija turėtų leisti įgyti vartotojų pasitikėjimą, suvokti jų įpročius, patikrinti naujas paslaugos teikimo procedūras ir fizines priemones [21]. Klientų aptarnavimo sistema yra technologijų rinkinys ir organizacinis tinklas, pritaikytas paslaugų teikimui. J. B. Quinnas ir P. C. Paquette (1990) paslaugų sistemas apibrėžia kaip paslaugų technologines sistemas, kurios skirtos tam, kad nepatyrę žmonės galėtų greičiau teikti paslaugas [25]. E -paslaugas galima apibrėžti, kaip elgesį, pastangas ar veiklą, kuri teikiama per informacines technologijas. Apskritai e -paslaugas galima apibrėžti kaip interaktyvias, tikslinio turinio, internetu teikiamas paslaugas, kurias iniciuoja klientas, o integruotos technologijos ir palaikymo sistemos priklauso paslaugos teikėjui [4]. Technologijomis grįstos savitarnos paslaugos yra visiškai automatizuotos, teikimo procesas perduotas klientui, kuris naudojasi savitarnos terminalais.

Pagrindines paslaugų charakteristikas apibrėžė Zeithaml, Parasuraman ir Berry (1985) remdamiesi 33 mokslininkų paslaugas analizuojančiais darbais (1975 - 1983). Pagrindinės charakteristikos gali būti trumpinamos IHIP [26]: neapčiuopiamumas (angl. *intangibility*), heterogeniškumas (angl. *heterogenety*), neperskiriamumas (angl. *inseparability*), teikimui nebūtinas inventorius.

Pagal ekonomikos sektorius galima išskirti tokius paslaugų tipus: verslas verslui (B2B) paslaugos, verslas klientui (B2C) paslaugos, vidinės paslaugos, viešojo sektoriaus paslaugos (G2C), neorientuotos į pelną [3].

A. Bankininkystėje naudojami savitarnos įrenginiai

Finansinių paslaugų sektoriuje revoliucija įvyko 1990 metais, kai bankų finansinės paslaugos buvo pradėtos teikti naudojant internetą. Informacinių technologijų pažanga leidžia bankams kurti naujus produktus ir juos teikti taip, kad kuo geriau būtų patenkinami klientų poreikiai. Bankų tinklalapiai, internetinė bankininkystė, savitarnos terminalai ir mobilūs įrenginiai suformavo į klientą orientuotas paslaugas [12].

Mažmeninės bankininkystės sektoriuje savitarnos paslaugos, teikiamos savitarnos įrenginiais, yra esminis veiksnys, lemiantis banko konkurencingumą, nes klientai nori gauti jiems reikalingas paslaugas nesikreipdami į banko skyrių, valdyti paslaugos teikimo procesą. Klientai įprato valdyti paslaugas, kai jos buvo pradėtos teikti internetu, išmanieji įrenginiai ir mobilieji mokėjimai suformavo įprotį naudotis paslaugomis bet kuriuo metu. Savitarnos paslaugos sudaro 90 % - 95 % kasdienių transakcijų [12]. Įvertinant, kiek informacijos pateikiama bankų tinklalapiuose, galima teigti, kad šioms paslaugoms neteikiama pakankamai svarbos, o tai tampa priežastimi, dėl kurios klientai jomis nesinaudoja arba pasirenka kitą banką.

Finansinės institucijos nurodo, kad 56 % klientų aptarnaujami skyriuose, 20 % - savitarnos įrenginiais, o mobilioji bankininkystė sudarė tik 5 %. 70 % bankų teigia, kad ateityje ATM svarba aptarnaujant klientus didės [1]. Klientai įprato naudoti savitarnos įrenginiais, o mobilioji bankininkystė, klientų suvokimu, stokoja saugumo.

Siekiant suprojektuoti labiausiai vartotojo poreikius tenkinančias savitarnos paslaugas, jų teikimo procese įtraukiami skirtingi savitarnos įrenginiai (informacijos terminalai, bankomatai, automatiniai seifai). Savitarnos sistemų darbui užtikriniti būtinos įvairaus pobūdžio programinės įrangos aplikacijos ir technologijos: verslo intelektikos, ESB architektūra, interneto protokolai (XML/XHTML, CSS, J2EE, Ajax ir RSS), mobilieji protokolai (SMS, MMS, LBM) [12]. ATM galima integruoti su kitais paslaugų teikimo kanalais diegiant M2M (angl. *machine to machine*) bevieles komunikavimo technologijas [17]. Bankuose dažniausiai naudojami savitarnos įrenginiai yra finansiniai terminalai, automatiniai kasininkų seifai, pinigus išduodantys bankomatai, daugiafunkciniai bankomatai, priimti ir išduoti pinigus galintys bankomatai.

Dauguma mokslininkų (Majkgard, 1998; Berry, 1995; Barczyk, 1999) teigia, kad vienintelis paslaugos vertintojas gali būti vartotojas. Paslaugų, kurias įvertnti vartotojas neturi pakankamai kompetencijos, vertinimui svarbus ir teikėjo kokybės vertinimas [21]. B. Edvardsson ir J. Juran savo paslaugų kokybės apibrėžime teigia, kad kokybę lemia tik vartotojo poreikių atitikimas, t. y. psichologiniai veiksniai. Ph. Crosby, E. Gummensson, J. E. Swan ir L. J. Combs (1976) į paslaugų kokybės apibrėžimą įtraukia ne tik psichologinių, bet

ir techninių parametrų vertinimą. R. Normann (1984) paslaugos kokybės įvaizdžio formavimą siūlo kontroliuoti naudojant tiesos momentus [21, 22]. Mokslininkai išskiria panašius elektroninių paslaugų kokybės vertinimo kriterijų rinkinius, tačiau jiems suteikia skirtingą svarbą. B. Thomasson, Ch. Gronroos, R. Johnston, R. Silvestro, Dabholkar, Zeithaml ir kt., Madu (2002), Santos (2003), Kim, Stoel (2004), Jun, Cai (2001) išskiria tokius kriterijus: patikimumas, prieinamumas, pagalba ir informacija, saugumas, draugiška vartotojo sąsaja, paslaugos formavimo proceso paprastumas [4, 21]. Klientų pasitenkinimas paslaugomis yra svarbus siekiant jų lojalumo. Svarbu susieti paslaugų kokybę, klientų pasitenkinimą ir lojalumą [6]. Paslaugų kokybės ryšį su klientų lojalumu nagrinėjo: Kumar ir kiti (2009), Heskett ir kt. (1997), Parasuraman ir kiti (1988), Naeem ir Saif (2009). Daugiausiai technologijomis grįstų paslaugų kokybės tyrimų atliekama apsiribojant paslaugomis, teikiamomis internetu [5].

ATM paslaugoms svarbūs šie aspektai: patogi ATM įrengimo vieta, paslaugos saugumas, suasmeninta vartotojo sąsaja, draugiška ir daugiakalbė vartotojo sąsaja, vartotojo poreikių nuspėjamumas [11, 14] .

1) SSTQUAL skalė

SSTQUAL pateikia keleto dimensijų skalę, skirtą įvertinti pagrindinius veiksnius, darančius SST paslaugų kokybei įtaką [5]. SSTQUAL skalė apima funkcionalumą, naudojimąsi, saugumą, užtikrintumą, dizainą, patogumą ir pritaikymą. Skalė parodo psichologines savybes, kurios yra nustatytos atlikus tyrimus, kaip klientai elgiasi naudodamiesi skirtingais SST pavyzdžiais. Meuter ir kiti (2000) apibrėžė, kad SST yra technologinė sąsaja, leidžianti klientui sukurti paslaugą, nepriklausomai nuo aptarnaujančio personalo įsikišimo [5]. Funkcionalumas aprašo SST funkcines charakteristikas: reagavimas, patikimumas, paprasta vartotojo sąsaja. Pasitenkinimas apibūdina suvoktą pasitenkinimą paslaugos gavimo metu ir gautais rezultatais. Saugumas / privatumas žymi suvokiamą saugumą nuo vagysčių, įsibrovimų ir asmeninės informacijos praradimo. Užtikrinimas yra susijęs su SST teikėjo kompetencija ir reputacija. Dizaino dimensija apibūdina viską, kas susiję su SST paslaugų sistemos projektavimu, patogumas apibūdina SST paslaugos prieinamumą. Pritaikomumas parodo, kiek SST paslaugos gali atitikti kliento reikalavimus ir jo transakcijų istoriją [5] .

2) Kokybės spragų modelis SERVQUAL

Vienas iš anksčiausiai sukurtų paslaugų kokybės vertinimo modelių yra kokybės spragų modelis [27]. SERVQUAL metodika buvo sukurta 1985 metais mokslininko Parasuraman, V. A. Zeithaml ir L. L. Berry ir yra skirta matuoti paslaugų kokybei, tai yra kokybės spragų modelio išplėtimas. SERVQUAL – tai paslaugų kokybės vertinimo priemonė, skirta įvertinti tiek paslaugos suvokimą, tiek lūkesčius pagal paslaugų charakteristikas. Suvokiama paslaugos kokybė nustatoma kaip balų, kuriais įvertinama laukiama ir patirta paslaugos kokybė, skirtumas. SERVQUAL metodiką sudaro 5 kriterijai: apčiuopiamumas, patikimumas, reagavimas, užtikrintumas, įsijautimas [4, 23].

Metodikos autorių teigimu, ji gali būti taikoma įvairių paslaugų kokybei matuoti, nes tik nedaugelis paslaugų sferos organizacijų pasižymi tokiomis išskirtinėmis savybėmis, kurių šios metodikos parametrai negali atspindėti. SERVQUAL metodika galima išmatuoti [21]:

- lūkesčių ir patirtos kokybės bendrąjį nuokrypį;
- atskiro parametro ir patirtos kokybės nuokrypį;
- paslaugos kokybės indeksą, t. y. lūkesčių balų ir patirtos kokybės balų sumų santykį.

Atsižvelgdami į elektroninių paslaugų ir paslaugų skirtumus, mokslininkai R. Suomi ir H. Li pasiūlė SERVQUAL metodikoje pridėti papildomas tris dimensijas: puslapio dizainas, saugumas, informacija [4].

3) Savitarnos paslaugų sistemų kokybės vertinimo kriterijų modelis

Klientai yra patenkinti SST, kai jos atitinka jų poreikius, veikia greitai ir yra lankstesnės nei alternatyvios paslaugos. Savitarnos paslaugų sistemų našumas gali būti apibendrinamas kaip greitis, prieinamumas ir patikimumas [15].

Paslaugų teikimo sisteminimas ATM tinkle užtikrina veiklos kokybę ir gerą paslaugų valdymą. ATM įrenginių teikiamų paslaugų kokybės tobulinimas yra susijęs su būtinais ištekliais: pakankamu skaičiumi ATM įrenginių tinkle, saugia įrengimo vieta (aplinka), vartotojui draugiška sistema, plačiu funkcionalumu [15].

4) Savitarnos paslaugų sistemų efektyvumo vertinimo ir tobulinimo modelis

Modelis leidžia užtikrinti vidinį, išorinį ir pajėgumo našumą [14]. Produktyvumas susideda iš kelių komponentų, kurie daro įtaką visos organizacijos našumui [14]. Kainos našumas apima reikalingas sąnaudas, kurios užtikrina reikiamą kokybę už mažiausią kainą. Įrengimų našumas - tai optimalus sąnaudų kiekis, naudojamas paslaugoms teikti. Techninis neefektyvumas – kai galima atlikti daugiau paslaugų su naudojamais ištekliais, nei yra atliekama. Apimčių efektyvumas – komponentas, kuris nukreipia optimalų veiklos apimties lygį, atliekant daugiau ar mažiau paslaugų nei optimalus lygis, įtraukiant išteklius, susijusius tik su paslaugų apimtimi [28]. Modelyje naudojami įrankiai yra klasifikuojami į tris kategorijas: kokybės (K), kiekybės (S), nepažymėti apima abu (K/S). Modelio linija, jungianti metodologiją ir veiklos komponentus, rodo, kurios technikos yra tinkamos specifinių komponentų veiklos vertinimui. Išorinis produktyvumas susideda iš pelningumo ir kokybės vertinimo kriterijų [14].

5) Paslaugų produktyvumo modelis

Pagal paslaugų produktyvumo modelį paslaugos produktyvumas yra funkcija, rodanti, efektyviai paslaugos procese naudojami ištekliai paverčiami į paslaugos rezultatą (vidinis efektyvumas), kiek gerai yra suvokiama paslaugos kokybė (išorinis efektyvumas), paslaugos proceso apimties efektyvumas. Svarbu nustatyti paslaugų apimtį dėl to, kad jų negalima sandėliuoti kaip produktų [13]. Kaip paslaugų proceso rezultatas gali būti įvardijama tik rezultato kiekybė ar rezultato kokybė (procesas ir rezultatas). Paslaugų produktyvumas apibrėžiamas formule [13]:

paslaugų produktyvumas = f (vidinis našumas, išorinis našumas, apimties našumas) (1)

SSTQUAL yra naudinga nustatant paslaugoms teikti reikalingas technologijas ir parodo paslaugų tobulinimų rezultatus (stiprius ir silpnus taškus). Skalė gali vertinti tiek kokybinius tiek kiekybinius parametrus, segmentuoti organizacijos klientus, nustato paslaugos teikėjo privalumus ir silpnus taškus [5].

SERVQUAL vertina paslaugos teikimo procesą, tačiau nevertina techninių paslaugos kriterijų bei ne visada leidžia įvertinti klientų suvoktą kokybę [23]. Analizė parodė, kad savitarnos paslaugų sistemų kokybės vertinimo kriterijų ir savitarnos paslaugų sistemų efektyvumo vertinimo ir tobulinimo modeliai pateikia vertinimo kriterijus, bet nenurodo paslaugų vertinimo įrankių. Paslaugų kokybę dažniausiai siūloma vertinti atliekant vartotojų apklausą.

III. SIŪLOMAS SAVITARNOS TERMINALŲ PASLAUGŲ VERTINIMO MODELIS

Literatūroje dažnai siūloma sprendimų, kaip vertinti paslaugas kokybiškai iš vartotojo pusės. Tačiau paslaugų kokybę vertėtų vertinti ne tik iš kliento, bet ir tiekėjo pusės, be to, analizuojant savitarnos terminalų duomenis rasti problemas ir jas šalinti (planuoti profilaktinius patikrinimus ir pan.). Pagal paslaugų produktyvumo vertinimo modelius yra išskiriamos trys našumo sritys, kiekvienai jų sudaromi kokybinių ir kiekybinių vertinimo kriterijų rinkiniai. Bankomatų duomenų analizė atliekama naudojant SOM, taip nustatoma klasė ir savybės, toliau tęsiamas vertinimas su ekspertine sistema, kuri pateikia rekomendacijas. ATM teikiamų paslaugų vertinimo algoritmas pateikiamas 1 pav. Vertinant ATM veiklą yra išskiriami pagrindiniai žingsniai:

1. Paslaugų procesų aprašymas ir analizė.
2. ATM vertinimo kriterijų nustatymas.
3. ATM veiklos duomenų surinkimas ir indeksų skaičiavimas.
4. ATM veiklos duomenų pateikimas SOM tinklui.
5. Rezultatų pateikimas ekspertinei sistemai.

Pagal eksperimentinės sistemos rekomendacijas optimizuojamas paslaugų teikimo procesas. Ekspertinės sistemos tikslas – vertinti ATM tinklus ir teikti rekomendacijas, susijusias su tinklo valdymu ir savitarnos paslaugų kokybės gerinimu.

ATM sistemos vertinimo procese galima išskirti tokias dalyvių grupes: bankas, klientas, ATM įrenginys, komunikacijų paslaugų teikėjai, paslaugų teikėjai. ATM tinklo kokybės vertinimą gali atlikti bankas arba organizacija, teikianti įrenginių priežiūros ir monitoringo paslaugas.

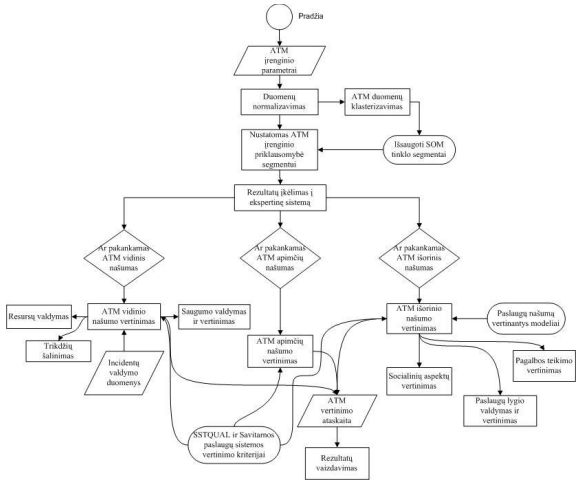
SOM tinklo esmė – duomenų topografijos išlaikymas. Taškai, esantys arti vieni kitų įėjimo vektorių erdvėje, yra vaizduojami arti vieni kitų SOM tinkle. SOM žemėlapiai gali būti naudojami siekiant vizualiai pateikti duomenų klasterius [18]. Tai galima interpretuoti kaip didelio matavimo duomenų vizualizavimą [19].

Pagal sudarytus kriterijus formuojamos ekspertinės sistemos taisyklės. Vertinamos tiekėjų ir vartotojų paslaugos. Naudojamos tokios ekspertinės sistemos klausimų grupės: vartotojų kokybės vertinimas, prieinamumo vertinimas, efektyvumo vertinimas, paslaugų funkcionalumo vertinimas. Kiekvienoje taisyklių grupėje yra suformuoti rekomendacijų rinkiniai.

IV. EKSPERIMENTINIAI SAVITARNOS TERMINALŲ PASLAUGŲ VERTINIMO MODELIO TYRIMAI

Eksperimentiniam tyrimui naudoti 5 daugiavaliučių ATM veiklos duomenys: išgryninamos sumos, inkasuojamos sumos, išinkasuojamos sumos ir gedimo trukmė. Bendriems rodikliams apskaičiuoti buvo naudojami Lietuvos centrinio banko statistiniai duomenys [29], apie visų Lietuvos bankų veiklą: pinigus išduodančių bankomatų skaičius, daugiafunkcinių ATM skaičius, atliktų operacijų vertė. Tirti 2008 - 2012 metų duomenys. Naudojant šiuos duomenis sudaryta ATM veiklos efektyvumo vertinimo lentelė pagal savitarnos paslaugų sistemos efektyvumo vertinimo ir tobulinimo modelį.

Sudarant SOM tinklą naudoti duomenys: ATM ID, inkasuota suma, išinkasuota suma, gedimo trukmė. Tikslas – taikant klasifikavimo metodą (SOM) ištirti, kaip pinigų kiekis bankomate ir bankomato gedimai lemia paslaugos kokybę (naudojimąsi bankomatu), kur dažniausiai sutrinka bankomatų veikla.



[a] Šaltinis sudarytas autorių.

1 pav. ATM teikiamų paslaugų vertinimo algoritmas

Naudojant SOM tinklą įrenginiai suskirstomi į klasterius pagal jų bendrus bruožus, įkėlus naujus įrenginio duomenis galima identifikuoti pakitusias savybes, netipines situacijas, įvertinti našumą. Stebint įrenginių našumą, galima identifikuoti sritis, kurios yra mažai apkrautos, o kurios - per daug, pašalinant ir perkeliant ATM galima optimizuoti ATM tinklą ir jam naudojamus išteklius. Subalansavus tinklą gerės paslaugų kokybė, nes klientams nereikės laukti eilėse, įrenginiai dirbs visu pajėgumu. Kokybiškos paslaugos klientus skatina naudotis visomis teikiamomis savitarnos paslaugomis.

Naudojant ATM veiklos duomenis sukurtas SOM tinklas, jo mokymui ir testavimui sudaromos imtys, naudojant

kryžminio perskirstymo procedūrą, kai kelis kartus yra perskirstomas duomenų rinkinys, taip pat apmokymui ir testavimui naudojant skirtingas rinkinio dalis. Duomenų normalizavimas atliekamas kuriant SOM tinklą. Kuriant tinklą duomenų grupėms suteikiami prioritetai padidinant jų koeficientus.

Išanalizavus ATM pasiskirstymą miestuose galima teigti, kad visuose miestuose vidutiniškai 1 ATM tenka 3 000 klientų. Duomenys surinkti iš bankų tinklalapių. Nuo šalies bendrojo vidurkio nagrinėjamų įrenginių apkrova labai skiriasi, daroma išvada, kad įrenginio apkrova priklauso nuo jo konkrečios įrengimo vietos. Bendras šalies ATM sistemų išdavimo operacijų standartinis nuokrypis yra 150 000 litų.

I LENTELĖ. ATM GRYNŲJŲ PINIGŲ OPERACIJŲ VERČIŲ PALYGINIMAS

Įrenginio ID	Vidurkis	Standartinis nuokrypis	Valiuta
1004	83	17	Litai
	120	28	Eurai
	497	109	Bendra
1045	21	7	Litai
	64	49	Eurai
	241	174	Bendra
2001	34	9	Litai
	42	20	Eurai
	179	73	Bendra
5011	34	9	Litai
	74	57	Eurai
	413	191	Bendra
7012	30	7	Litai
	180	97	Eurai
	650	330	Bendra
Bendra sistemos	58	48	Litai
	65	81	Eurai
	281	277	Bendra
Šalies bendra sistema	350	10	Litai

I lentelėje palyginama, kaip vienos operacijos vertė skiriasi nuo ATM įrengimo vietos. Didžiuosiuose miestuose operacijų vertę didina operacijos užsienio valiuta, litais išgryninamos mažos pinigų sumos. Rekomendacijos atvirose erdvėse įrengtiems ATM yra pateikiamos remiantis bendros šalies sistemos duomenimis. Vertinimo indeksai ATM, įrengtiems uždarose patalpose, apskaičiuoti pagal nagrinėtų įrenginių duomenis. Apkrova yra skaičiuojama įrenginių, kurie veikia visą parą. Apkrovos indeksas parodo, kaip efektyviai dirba įrenginys. Skaičiuojamas vieno įrenginio paslaugų teikimo našumas imant savaitės trukmės duomenis. Išteklių našumą siūloma vertinti ir kiekvieno įrenginio, ir bendrą visos sistemos. Laikas skaičiuojamas valandomis per mėnesį ir parą. Vertinimo indeksai pateikti II lentelėje.

Bankai Lietuvoje smulkių transakcijų neapmokestina, tačiau yra papildoma paslauga – SMS siuntimas pasikeitus sąskaitos likučiui. Paslauga apmokestinama 0,3-0,45 ct skirtinguose bankuose. Kaštai apskaičiuoti remiantis savitarnos įrenginių veiklos duomenimis, bankuose taikomomis palūkanų normomis, ATMIA (ATM Industry Association) pateikta gerąja praktika. ATM veiklos kaštai skaičiuojami savaitei, nes dažniausiai ATM inkasuojami kartą per savaitę.

Antroje lentelėje yra pateikiami kiekybinių vertinimo indeksai, kurie naudojami teikiant veiklos vertinimo ir gerinimo rekomendacijas. Skaičiuojant paslaugų teikimo našumą laikoma, kad pusė transakcijų buvo atlikta siunčiant ir žinutę. Vieno terminalo įšaldytieji ištekliai apskaičiuojami naudojant ATM sistemos inkasuotų ir išinkasuotų sumų kvartilius, mažesnė rodiklio reikšmė rodo geriau panaudojamus išteklius. Nagrinėjant vieno iš Lietuvos komercinių bankų ATM duomenis buvo nustatyta, kad ATM yra neprieinami 1,73 % viso darbo laiko per mėnesį. Pagal šiuos duomenis apskaičiuojamas įrenginio prieinamumas. Vertinimo rodikliai apskaičiuoti remiantis savitarnos paslaugų sistemų efektyvumo vertinimo ir tobulinimo modeliu.

II LENTELĖ. ATM KIEKYBINIŲ ĮVERTINIMŲ REKOMENDACIJŲ LENTELĖ

Dimensija	Nepakankamas		Pakankamas		Geras	
	Mėnesio	*Dienos*	*Mėnesio*	*Dienos*	*Mėnesio*	*Dienos*
Įdėjimo operacijų atlikimas (tūks.)	2-2,2	60-76	2,2-2,5	76-87	2,6-2,8	87-95
Pinigų pervedimo operacijų atlikimas	3		4		5	
Išdavimo operacijų atlikimas (tūks.)	3-3,7	96-124	3,7-4,7	125-155	4,7-5,5	156-185
Išdavimo operacijų vertė (tūkst., litais) A[1]	980-1370	30-21	1375-1799	31-38	1800-2150	38-45
Išdavimo operacijų vertė (tūkst.. Lt) P[2]	75-145	2-3	150-610	3-20	615-1100	20-36
Įdėjimo operacijų vertė P[3]	1615-2090	50-71	2090-2566	72-94	2566-3041	95-116
Bendra operacijų vertė	2737-3540		3547-4357		4400-5167	
ATM apkrova = atliktų transakcijų skaičius/ darbo laikas A	0,5-0,7	4-5	0,74-0,93	5,2-6,4	0,93-1,1	6,5-7,75
ATM apkrova = atliktų transakcijų skaičius/ darbo laikas P	6,9 -8,19	6,5-8,3	8,2 -10	8,7-10,1	10,1-11,5	10-11,7
Apkrautumas = atliktas transakcijų skaičius/darbo laikas	6,6-8,2	8-10	8,2-10,4	10-13	10,4-12	13-16
Prieinamumas = gedimo trukmė (neveikimo)/visos darbo valandos	24		12,45		5,4	
Visi įšaldytieji ištekliai = grąžinti ar prarasti ištekliai/visi ištekliai (visų terminalų)	0,7		0,18-0,6		0,14	
Iššaldyti ištekliai = grąžinti ar prarasti ištekliai/visi ištekliai (vieno terminalo)	0,4		0,2		0,1	
Sąnaudų našumas= transakcijų sk./išteklių sąnaudos	2-2,4		2,5-3		3-3,5	

[a.] 1 - vertinant ATM įrengtus atvirose erdvėse; 2 - vertinant ATM įrengtus uždarose patalpose; 3 - visi banknotus priimantys ATM, įrengti patalpose, skaičiuojama pagal bendrus šalies duomenis.

Inkasavimo sąnaudos didėja, jei bankomatas priima pinigus, nes tai - nepanaudojami ištekliai, išauga išinkasuojamų sumų skirtumas palyginti su inkasuojamomis. Grąžinamų pinigų kiekis yra apie 40 %, jį sudaro likę ir įdėti

pinigai. Galima teigti, kad išinkasuojamos sumos dydžiui didesnę įtaką turi inkasuojama suma nei neveikimo trukmė, kuri dažniausiai būna labai trumpa. Vertinant savitarnos įrenginių ir ATM tinklo našumą reikia atsižvelgti į įrenginiams tenkančių klientų srautą, įrengimo vietą, funkcionalumą. ATM veiklos vertinimo indeksai yra apskaičiuoti ir naudojami sukurtame ATM vertinimo modelyje.

V. IŠVADOS

Išaugus paslaugų sektoriui mokslininkai pradėjo intensyviai tirti paslaugų kokybę. Paslaugų kokybės modeliuose pabrėžiami skirtingi kokybės vertinimo kriterijai. Sparčiai diegiant SST aptarnavimo sferoje atsirado poreikis naujų kokybės vertinimo modelių, kurie vertinti paslaugų sistemų efektyvumą ir kokybę. Dažnai finansinių paslaugų kokybės tyrimuose mokslininkai apsiriboja klientų apklausomis.

Straipsnyje pristatytas mažmeninės bankininkystės savitarnos terminalų paslaugų vertinimo modelis, kuris leidžia įvertinti paslaugų kokybę iš kliento pusės ir efektyvumą - iš tiekėjo pusės. Atliekant eksperimentinius tyrimus apskaičiuotos ATM veiklos sąnaudos, ATM veiklos efektyvumo vertinimo indeksai, vidutinis transakcijų skaičius ir vertė. Vertinant savitarnos terminalus svarbu atsižvelgti, kiek klientų tenka įrenginiui ir kokia darbo trukmė. Eksperimentinių tyrimų rezultatai parodė, kad derinant intelektines sistemas (intelektinių agentų, SOM, ekspertinių sistemų) ir taikant savitarnos paslaugų vertinimo procese galima sukurti efektyvų savitarnos paslaugų vertinimo modelį.

SST sparčiai diegiamos gydymo įstaigose (medicinos savitarnos terminalai) ir prekybos centruose (savitarnos kasos). Tęsiant eksperimentinius tyrimus mažmeninės bankininkystės savitarnos paslaugų modelis gali būti pritaikytas kitų sričių savitarnos terminalams vertinti, nes visos SST turi panašių savybių, modifikuojamos šios dalys: efektyvumo parametrai, papildomo funkcionalumo vertinimas ir ekspertinės sistemos rekomendacijos.

LITERATŪRA

[1] J. Grove, R. Slawsky, "2012 ATM Software Trends and Analysis", 5th edition. Published by Networld Media Group, pp. 13.

[2] J. Spohrer, Paul P. Maglio, "The Emergence of Service Science: Toward Systematic Service Innovations to Accelerate Co-creation of Value", pp. 7, 8, 11. IBM Almaden Research Center, 2008.

[3] R. Jonston, "Service Operations Management: From the Roots up". International Journal of Operations and Production Management, Vol. 25, No. 12, pp. 1300-1302.

[4] L. Hongxiu, R. Suomi, "A Proposed Scale for Measuring E-service Quality". Journal of u- and e-Service, Science and Technology, 2009, pp. 3-4.

[5] C. Lin, P. Hseih, "Assessing Self-Service Technology Encounters: Development and Validation of SSTQUAL Scale", Journal of Retailing, 87, 2011, pp. 196-206.

[6] K. O. Siddiqi, „Interrelations between Service Quality Attributes, Customers Satisfaction and Customers Loyalty in the Retail Banking Sector in Bangladesh", IJBM, Vol. 6, No. 3. March 2011, pp. 5, 3.

[7] A. Abedniya, M. N. Zaeim, „Measuring the Perceive Service Quality in the Islamic Banking System in Malaysia", International Journal of Business and Social Science, Vol. 2, No. 13, July 2011.

[8] K. H. Hanzaee, S. A. S. Salehi, "A Perceived Service Quality Measurement Scale in Iran's Retail Banking Market", AJBM Vol. 5 (5), pp. 1804-1810, 4 March, 2011.

[9] M. F. Shabbir, H. D. Aslam, "Perceived Service Quality of Islamic and Non Islamic Banks Operating in Pakistan", American Journal of Scientific Research, ISSN 1450-223 Issue 51, pp. 27-36. Euro Journals Publishing, Inc. 2012.

[10] M. Hensmans, "A Respublican Settlement Theory of the Firms Applied to Retails Banks in England and the Nytherlands (1930-2007)". 214, 253 psl. ISBN 978-90-5892-235-9. 2010.

[11] T. O. Ombati, P. O. Magutu, S. O. Nyamwange, R. Nyaoga, "Technology and Service Quality in the Banking Industry". 10 p. *AJBUMA* pp. 16 Vol. 1, 2010.

[12] B. King, "Bank 2.0: How Customer Behavior and Technology Will Change the Future of Financial Services". Marshall Cavendish Reference (July 15, 2010). ISBN-13: 978-9814302074.

[13] C. Gronroos, K. Ojasalob, "Service Productivity Towards a Conceptualization of the Transformation of Input into Economics Results in Services". Journal of Business Research 57, 2004, 414-423.

[14] D. Dilijonas, V. Sakalauskas, "Self-Service Systems Performance Evaluation and Improvement Model". International Federation for Information 2011.

[15] D. Dilijonas, D. Krikščiūnienė, V. Sakalauskas R. Simutis, "Sustainability Based Service Quality Approach for Automated Teller Machine Network". 5th International Vilnius Conference EURO Mini Conference September 30–October 3, 2009, VILNIUS, Lithuania. ISBN 978-9955-28-482-6.

[16] M. Jun, S. Cai, "The Key Determinants of the Internet Banking Service Quality: a Content Analysis", International Journal of Bank Marketing 19/7, ISSS0265-2323, 2001.

[17] G. Wollenhaup, "Convert ATMs to Wireless Communication". ATM Market Place White Paper. 2012.

[18] V. Medvedev, „Tiesioginio sklidimo neuroninių tinklų taikymo daugiamačiams duomenims vizuolizuoti tyrimai", daktaro disertacija, Vilnius: Technika, 2007 ISBN 978-9955-28-211-2.

[19] D. Plikynas, P. Daniušis, „Kompiuterinis intelektas ir daugiagentės sistemos socialinių mokslų srityje", p. 85. Monografija, 2009.

[20] C. Hsieh, " Implementing Self-Service Technology To Gain Competitive Advantages ". Communications of the IIMA Vol. 5 Issue 1. 2005.

[21] L. Bagdonienė, R. Hopenienė, "Paslaugų marketingas ir vadyba". 16, 107, 137, 141, 251 p. Kaunas 2005: Technologija. ISBN 9955-09-579-2.

[22] N. Lengvinienė, B. Vengrienė, „Paslaugų teorija ir praktika". Pp. 13, 20-22, 2005 Kaunas: Technologija, ISBN 9955-09-924-0.

[23] K. Ravichandran, B. T. Mani, S. A. Kumar, S. Prabhakaran, "Influence of Service Quality on Customer Satisfaction Application of SERVQUAL Model", International Journal of Business and Management, vol. 5, no. 4; April 2010.

[24] J. M. Tien, D. Berg, "A case for service systems engineering". Journal of Systems Science and Systems Engineering, Vol. 12, No. 1, pp. 13-38, March 2003.

[25] J. B. Quinn, P. C. Paquette, "Technology in Services: Creating Organizational Revolutions". MIT Sloan Management Review. 31(2), 1990.

[26] C. Lovelock, "The Future of Services Marketing". Presentation. Frontiers in Services Conference University of Miami, October 31, 2004.

[27] G. Salvendy, W. Karwowski, "Introduction to Service Engineering", ISBN 987-0-470-38241-7TK 5105.5.1685, 2010.

[28] H. D. Sherman, J. Zue, "Service Productivity management, Improving Service Performance Using Data Envelopment Analysis (DEA)", Springer, 2006, pp. 3-7.

[29] LIETUVOS RESPUBLIKOS CENTRINIS BANKAS, Šalies bankomatai ir kortelių skaitytuvai bei juose atliekamos operacijos, 2012.

INFORMACINĖS TECHNOLOGIJOS ● IVUS 2013 ● ISSN 2029-249X ● eISSN 2029-4824

Automatinis citatų atpažinimas lietuviškuose naujienų portaluose

Raimondas Širvinskas
Vytauto Didžiojo universitetas
Kaunas, Lietuva
raimis.sirvis@gmail.com

Žygimantas Medelis
UAB „Tokenmill“
Vilnius, Lietuva
zygimantas.medelis@tokenmill.lt

Tomas Krilavičius
Vytauto Didžiojo universitetas, Baltijos pažangių technologijų institutas,
Kaunas ir Vilnius, Lietuva
t.krilavicius@gmail.com

Santrauka. Šiame straipsnyje pristatoma automatizuoto citatų atpažinimo lietuviškuose interneto portaluose metodika. Tyrime išanalizuotos lietuvių kalbos citavimo taisyklės. Apžvelgti automatinio citatų atpažinimo metodai naudojant GATE ir JAPE užrašytas taisykles citatų atpažinimo, taip pat apibrėžti kokybės kriterijai rezultatams įvertinti. Rezultatai įvertinti naudojant anotuotą tekstyną, sukurta demonstracinė sistema, kuri automatiškai skenuoja tekstus ir identifikuoja citatas.

Reikšminiai žodžiai — citata, kalbos technologijos, kompiuterinė lingvistika, esybių atpažinimas.

I. ĮVADAS

Staigus informacijos kiekio internete augimas [1] sukelia vis daugiau keblumų vartotojams, kuriems reikia iš didelio kiekio informacijos atsirinkti tik dalį, kuri jiems aktualiausia ir reikalingiausia. Naujienų portalai kiekvieną dieną sugeneruoja labai daug straipsnių, todėl, norint surasti reikiamą informaciją juose, tenka ilgai naršyti ir skaityti. Šiuo metu pasaulyje informacijos kiekis auga labiau nei turimų saugyklų kiekis, todėl pradėta informaciją apdoroti ir struktūrizuoti atrenkant tik vertingą.

Citatų atpažinimo tema pasirinkta dėl to, kad ji buvo mažai tyrinėta tiek Lietuvoje, tiek pasaulyje. Pačios citatos reikalingos politikoje rašant kalbas, žiniasklaidoje rašant straipsnius, versle atsižvelgiant į stambių korporacijų vadovų pasisakymus bei prognozes. Pavyzdžiui, politikų rinkiminiai pažadai, cituoti naujienų portaluose, gali būti peržvelgiami prieš naujus rinkimus ir sulyginti su atliktais darbais. Taip pat citatas galima naudoti klasifikuojant straipsnių, priskiriant juos kuriai nors temai pagal straipsniuose rastas citatas.

Egzistuojantys citatų gavybos rezultatai: australų Timo O'Keefe, Irenos Koprinskos ir Jameso Currano darbas „Quotation extraction and attribution“ [2], atliktas Sidnėjaus universitete, Australijoje. Jų taisyklėmis grindžiamas citatų identifikavimo metodas identifikuoti 99,64% tiesioginių citatų iš jų tekstyno su mažai citatų klaidingai atpažintų dėl klaidų tekste pasitaikiusių klaidų (rašybos klaidos, trūkstamos kabutės ir pan.). Citatos išskyrimas buvo kitoks nei lietuvių kalboje, tačiau jų taikytas metodas buvo labai tikslus, todėl šiame darbe buvo nuspręsta remtis taisyklėmis pagrįstais metodais.

Šiame darbe sukurta ir pristatyta automatizuoto citatų atpažinimo lietuviškuose interneto portaluose metodika. Išanalizuotos lietuvių kalbos taisyklės skirtos išskirti tekste citatoms. Užrašytos taisyklės kaip turi būti atpažįstama citata. Apibrėžti kokybės kriterijai rezultatams įvertinti. Sukurta demonstracinė sistema.

Antrajame skyriuje nurodoma kaip turėtų būti cituojama lietuvių kalboje. Trečiajame skyriuje pateikiami esybių nustatymo metodai. Ketvirtajame skyriuje aprašomas eksperimentinis įvertinimas. Penktajame skyriuje detaliai aprašytos sukurtos taisyklės šiam darbui. Šeštajame skyriuje aprašomas viso proceso automatizavimas. Pabaigoje pateikiami rezultatai ir išvados.

II. CITAVIMAS LIETUVIŲ KALBOJE

Valstybinės lietuvių kalbos komisijos nutarime [3] apibrėžti keli citatos atskyrimo tekste būdai:

1. Citatos išskiriamos kabutėmis. (Kai citatos išskiriamos šriftu ar kitokiu grafiniu būdu, kabutės gali būti nerašomos.)
2. Kai citatą sudaro atskiras sakinys ar keli sakiniai, po įvadinių žodžių dedamas dvitaškis, o citata paprastai rašoma toje pačioje eilutėje.
3. Ilgesnė citata gali būti pradedama nauja pastraipa, pvz.:

   Dar Aristotelis rašė: „Retorika apima visas žmogaus gyvenimo sritis.“ Apie perdėtą kuklumą pasakyta taip: „Jei esi nekompetentingas ar nepasiruošęs, kaip sakaisi, kam eini kalbėti? O jei tas netiesa, kam pradėti savo kalbą melu?“ Štai koks gyvas kvieslio paveikslas: „Savaitės viduryje į namus atjojo kvieslys kviesti į vestuves. Arklio galva ir jo paties skrybėlė buvo išpuošta margais kaspinais. <...> Jis sustojo vidury aslos įsitempęs, kaip kareivis rikiuotėje, pakėlė ranką ir pakratė lazdą, ant kurios suplevėsavo kaspinai ir rūtos ir sutilindžiavo varpelis.“
4. Po citatų esantys ir į jas įsiterpę autoriaus žodžiai rašomi taip, kaip autoriaus žodžiai prie tiesioginės kalbos (12.6 punktas iš [3]), pvz.:

   „Ar visuomet vaizdingas žodis geriau už neutralų?“ – svarstydamas žodžių vartosenos tikslingumą klausia J. Pikčilingis. „Stinga išraiškos

priemonės, vartok tarmybę“, – sakė Rygiškių Jonas. „Pareina namo, – toliau rašo skaitytoja, – ir prasėdi visą laiką prie televizoriaus.“

Netaisyklingai išskiriamų citatų, atliekant šį tyrimą, dažniausiai pasitaikė dėl skirtingų kabučių citatų pradžioje ar gale, taip pat dažnai nesilaikoma lietuvių kalbos taisyklių. minėtų II skyriuje ir rašybos klaidų.

I lentelė. Kabučių įvairovė

Kabutė vizualiai	Kabutės kodas („Unicode“ formatu)
“	201C
″	05F4
"	0348
̏	030F
”	201E
”	201D
“	201C
″	05F4
˝	02DD
"	0022
″	2033
“	201F

A. *Kabučių vartosena naujienų portaluose*

Kabutės vartojamos tiesioginei kalbai, citavimui, netikra reikšme vartojamiems žodžiams, prekių pavadinimams išskirti. Todėl pati pirmoji ir pagrindinė problema buvo atskirti citavimą nuo tiesioginės kalbos, netikra reikšme vartojamų žodžių bei kitų kabučių vartojimo atvejų, kurie yra plačiai apžvelgti Renatos Vyšniauskaitės magistro darbe [4].

III. ESYBIŲ NUSTATYMAS

A. *Esybių atpažinimo užduotis*

Citatas galima priskirti vienai iš esybių klasių, todėl buvo sėkmingai pritaikyti esybių atpažinimo metodai, leidžiantys atpažinti citatas ir autorius.

Esybių atpažinimas (angl. *named entity eecognition* – NER) [5] yra informacijos gavybos užduotis, kuria siekiama identifikuoti ir klasifikuoti elementus tekste, kurie priklauso iš anksto numatytoms kategorijoms, pavyzdžiui, vardai, organizacijos, vietos, laikas, kiekį bei piniginės išraiškos.

Nesudėtingos sistemos naudojasi taisyklėmis besiremiančiais algoritmais. Šiuo metu modernios sistemos telkiasi į pagalbą mašininio mokymo algoritmus, tokius kaip Paslėptas Markovo modelis (angl. *hidden Markov model*) [6], Sprendimų medžiai (angl. *Decision tree*) [7], maksimalios entropijos modeliai (angl. *maximum entropy modeling*)[8], sąlyginiai atsitiktiniai laukai (angl. *conditional random fields*) [9] bei atraminių vektorių metodas (angl. *support vector machine*) [10].

Prižiūrimi algoritmai (angl. *Supervised Learning*) [11] naudoja bendrinę procedūrą (paprastai vadinamą mokymosi), pagal kurią susikuria klasifikatorių. Šiame etape keletas tekstų rankiniu būdu apdorojami identifikuojant (ir klasifikuojant) esybes, iš kurių algoritmas daro išvadą apie charakteristikas, apibrėžiančios šias esybes. Vėlesniame klasifikavimo etape klasifikatoriui pateikiami nauji tekstai, kuriuose tikimasi, kad bus atpažintos esybės.

B. *Esybių atpažinimo taikymas*

Esybių atpažinimas tokiose teksto žanrų srityse kaip žurnalistika, moksliniai straipsniai, neformalūs tekstai ir kituose žanruose (sodininkystė, sportas, verslas ir kituose) buvo mažai naudojamas. Keli tyrimai buvo specialiai skirti įvairių žanrų ar sričių esybių atpažinimui.

C. *Įrankiai*

1) GATE [12] (*General Architecture for Text Engineering*) – teksto analizės sistema, sukurta „Java“ kalba Sheffieldo universitete ir prieinama kaip atvirojo kodo programinė įranga. *GATE* yra infrastruktūros kūrimo ir diegimo komponentai, kurie apdoroja žmogaus kalbą. Jau beveik septyniolika metų ji yra aktyviai naudojama žmogaus kalbai apdoroti.
2) „LingPipe“ – 2003 m. kompanijos „Alias-i“ sukurtas įrankis tekstui apdorojimui [13]. Šio įrankio pagrindinė paskirtis – identifikuoti ir išskirti žmonių, organizacijų ar vietovių pavadinimus iš naujienų srautų.
3) „Tesla“ [14] – teksto inžinerijos programinė įranga, sukurta Kelno universitete, Vokietija. Galima naudoti sukurtus eksperimentus, lingvistinius tekstynus ir plėtoti juos naujais natūralios kalbos apdorojimo algoritmais. „Tesla“ PĮ veikimo pobūdis yra klientas-serveris, kuris gali būti naudojamas vienas ar darbui grupėmis. „Tesla“ atpažįsta .xml tipo failus ir su jais dirba.

D. *Paieška pagal sąrašus*

Vienas iš standartinių būdų atpažinti esybes yra sąrašų (angl. *list, gazetter, dictionary*) naudojimas. Šis būdas gana efektyvus dirbant su tikriniais daiktavardžiais, rūšių vardais ir pan. Sąrašai padeda atskirti ne citatas, o organizacijas, išskirtinius žodžius, netiesiogine kalba pavartotus žodžius, kurie tekste yra dažniausiai išskiriami kabutėmis.

E. *Taisyklėmis grįsti metodai*

Taisyklėmis grįsti metodai šiuo metu yra vienas populiariausių metodų. Šiuo atveju užrašomos taisyklės, skirtos atpažinti tekste tam tikroms esybėms, pvz., datoms, įmonių pavadinimams, pavardėms ir pan. Dažnai taisyklėmis grįsti metodai derinami su žodžių sąrašais.

Dažniausiai tam gali būti naudojamos reguliarios išraiškos (angl. *regexp*). Tarkime, norint atpažinti elektroninio pašto adresą, gali būti naudojama išraiška [15]:

```
[A-Za-z0-9](([_\.\-]?[a-zA-Z0-
9]+)*)@([A-Za-z0-9]+)(([\.\-]?[a-zA-Z0-
9]+)*)\.([A-Za-z]{2,})
```

Taip pat galima naudoti JAPE [16] (*Java Annotation Patterns Engine*). Tai – GATE modulis, kuris veikia labai panašiai kaip reguliariosios išraiškos, tačiau kalba yra lengviau skaitoma, efektyvesnė.

Kaip jau minėta, šie metodai leidžia pasiekti gana gerų rezultatų, įvertinti visas išimtis ir į jas atsižvelgti. Tačiau rašyti taisyklėms reikia daug laiko, būtina įvertinti visus variantus, o norint patobulinti sistemą – gerai suprasti, kaip visa tai veikia.

JAPE buvo pasirinkta darbui todėl, kad jos gramatikos taisyklės susideda iš modelių, veiksmų, taisyklių. Gramatika visada turi dvi puses: kairę ir dešinę. Taisyklėmis yra nustatomas anotacijos modelis, kurio kairioji pusė gali būti reguliarios išraiškos operatoriai (pvz.: *, ?, +), o dešinioji pusė apibrėžia, kokių veiksmų reikia imtis, kad anotacija atitiktų peržiūrint aptiktos anotacijos manipuliavimo modelį. JAPE taisyklės yra lengvai skaitomos ir suprantamos žmogui, todėl lengva aprašyti ilgas ir sudėtingas taisykles.

IV. EKSPERIMENTINIS ĮVERTINIMAS

A. *Eksperimentinis tekstynas*

Testavimui buvo surinkta 500 straipsnių iš interneto, straipsniai buvo anotuoti rankiniu būdu. Naudojant GATE „Teamware“ anotavimo sistemą, kurios pagalba buvo suanotuotos 207 citatos. Pasitelkiant įrankį GATE „Developer“ ir taisykles JAPE iš viso buvo atpažinta 190 citatų. Testavimui buvo pasirinkti straipsniai tik lietuvių kalba, nes anglų kalba cituojama kitaip bei lietuvių kalboje dar nėra atliktų jokių tyrimų šia tema.

Naudotas tekstynas:

1) 200 straipsnių iš www.delfi.lt;
2) 200 straipsnių iš http://www.culture.lt/satenai/;
3) 100 straipsnių iš įvairių tinklaraščių.

Tekstyno apimtis – 516 336 žodžių arba 3 680 117 simbolių (3 197 497, neįskaitant tarpo simbolio).

Tekstynui anotuoti prireikė daug laiko. Dirbant beveik aštuoniasdešimties valandų 500-uose anotuotų dokumentų buvo rastos 207 citatos.

B. *Klaidos vertinimo kriterijai*

Esybių atpažinimo sistema paprastai yra vertinama remiantis tuo, kaip jos išvestis siejasi su žmogaus kalba. Tai gali lemti penkias skirtingas klaidas:

- Sistema spėja esybę, kurios nebuvo.
- Esybė buvo visai klaidingai atpažinta sistemos.
- Sistema atpažino esybę, bet suteikė jai neteisingą etiketę.
- Sistema atpažino esybę, bet klaidingai atpažino jos ribas.
- Sistema suteikė esybei neteisingą etiketę ir klaidingai atpažino jos ribas.

Sistemos teikiamas rezultatas priklauso nuo to, kaip šios klaidos yra skaičiuojamos. Todėl lyginant rezultatus reikia tiksliai išsiaiškinti, kaip skaičiuojamas tikslumas ir išsamumas [17] (atitinkamai 1, 2 ir 3 formulės skaičiuoti įverčiams pateikiamos toliau).

C. *Atlikti eksperimentai*

Iš 207 citatų buvo atpažinta 190 citatų. 16 citatų nebuvo atpažinta dėl to, jog prie citatos nebuvo rastas jos autorius, dėl citavimo ne pagal valstybinės lietuvių kalbos komisijos nustatytus šablonus bei įvairių klaidų citatoje ar jos aprašyme. Taip sunkumų kyla dėl kabučių, nes jų yra daugybė variantų, kurie pateikiami I lentelėje. Reiktų juos visus surasti ir pritaikyti taisyklę JAPE. Be to, viena citata buvo atpažinta klaidingai. Klasifikuoti rezultatai pateikiami II lentelėje. Pagal 1 formulę apskaičiuotas taikyto metodo tikslumas, pagal 2 formulę paskaičiuotas išsamumas ir pagal 3 formulę paskaičiuotas – F įvertis. Gauti rezultatai pateikti III lentelėje.

II lentelė. Rezultatų klaidų tipai (I ir II tipo klaidos) [18]

tt – 190 (teisingai teigiami) Teisingai atpažintos citatos	kt – 1 (klaidingai teigiami) Tekstas priskirtas ne esamai citatai
kn – 16 (klaidingai neigiami) Tekstas, kuris yra citata, neatpažintas	tn – 0 teisingai neigiami Tekstas nepriskirtas ne esamai citatai

$$\text{Tikslumas [19]} = \frac{tt}{tt + kt} \quad (1)$$

$$\text{Išsamumas [19]} = \frac{tt}{tt + kn} \quad (2)$$

$$\text{F įvertis [19]} = 2 * \frac{tikslumas * išsamumas}{tikslumas + išsamumas} \quad (3)$$

III lentelė. Rezultatų įverčiai

F įvertis	0.957
Tikslumas	0.995
Išsamumas	0.922

V. CITATŲ ATPAŽINIMAS NAUDOJANT GATE IR JAPE

Norint atpažinti citatą reikia aprašyti taisykles, pagal kurias vėliau programa GATE atpažintų tekste esančias citatas.

Taisyklę JAPE galima išskirstyti į tokią seką:

1. Aprašoma makro taisyklė, kuri apibrėžia citatą.
2. Aprašoma makro taisyklė, kuri apibrėžia citatos autorių.
3. Aprašoma bendra citatos ir autoriaus atpažinimo taisyklė, kurioje bandoma atpažinti citatos ribas,

iškviečiamos makro taisyklės, aprašytos ankstesniuose punktuose.

```
Macro: CONTENT
(
  (({Token.orth==upperInitial})          |
({Token.kind==number}))
  ({Token.length>=2})[3, 5]
  ({Token})[4, 100]
)
```

Aprašomas citatos turinys: citatos turinys turi prasidėti didžiąja raide arba skaičiumi. Citatą sudaro nuo 4 iki 100 leksemų.

```
   Macro: PERSON
   (
     (({Token.kind==word,Token.length>=3,
Token.orth==upperInitial})[2])
     |
     ({Token.kind==word,Token.length==1,
Token.orth==upperInitial}
     {Token.string ==".")
     {Token.kind==word,
Token.orth==upperInitial})
   )
```

Aprašomas citatos autoriaus atpažinimas. Tai gali būti du žodžiai, ilgesni nei 3 simboliai, kurie prasideda didžiąja raide, arba tai gali būti autoriaus vardo pirmoji raidė ir autoriaus pavardė. JAPE kalboje aprašoma taip, kad pirmiausia eina leksema, kurios ilgis lygus vienam ir kuri yra didžioji raidė; toliau iškart po raidės eina taško simbolis, kuris reiškia vardo sutrumpinimą, o po jo leksema, prasidedanti didžiąja raide.

```
   Rule: Citata
   Priority:50
   (
     (
       (
         ({Token.string == "."}) |
         ({Token.string == "“"}) |
         ({Token.string == ")"}) |
         ({Token.string == "!"}) |
         ({Token.string == "?"})
       )
       {Token.string == „"}
       (
         (CONTENT)
       ):body
       {Token.string == "“"}
       {Token.string == ","}
       {Token.string == "-"}
       (
         ({Token})[0,20]):tmp
       )
   (PERSON):person
      (
        ({Token})[0,3]
        {Token.string ==","}
       )
   ):all
   -->
   :all.Quote = {},
     :person.Author = {rule = "Citata"},
     :body.Body = {rule = "Body"}
```

Šiame etape viskas sujungiama, ir pradedama analizuoti tekstą. Citata prasideda pasibaigus sakiniui (lietuvių kalboje citatos pradžios atskyrimo kabutė yra „), tuomet pritaikoma anksčiau aprašyta makro komanda, kuri atpažįsta visą citatos turinį ir laiko tai citatos tekstu iki citatos pabaigos kabučių (lietuvių kalboje citatos pabaigos atskyrimo kabutė yra “), tačiau to dar nėra gana, nes lietuvių kalboje citatos užbaigiamos tokia simbolių seka: “, − po kurios dažniausiai įterpiamas teksto autoriaus žodis, kuriame dažniausiai paminimas citatos autorius, todėl po citatos pabaigos neieškoma arba ieškomi keli autoriaus žodžiai bei po jų ieškoma citatos autoriaus vardo bei pavardės pritaikant ankčiau aprašytą makro komandą. Po autoriaus citatos gali eiti keli teksto autoriaus žodžiai. Citatos atpažinimo procesas vyksta iki sakinio pabaigos.

VI. CITATŲ ATPAŽINIMO AUTOMATIZAVIMAS

Iki šiol visas anotavimo procesas būdavo iš dalies rankinis:

1. Paleisti anotavimo įrankį (mūsų atveju GATE „Developer“ arba GATE „Teamware“),
2. Įsikelti norimus anotuoti dokumentus,
3. Pasirinkti anotavimo papildinius,
4. Pasirinkti taisyklę JAPE,
5. Paleisti programą dirbti,
6. Laukti, kol visas procesas baigsis ir bus matyti rezultatas.

Duomenų apdorojimas buvo automatizuotas. Šiam tikslui buvo pasitelkta nemokamai platinama biblioteka „gate-clj“ [20] biblioteka, kuria buvo galima paleisti GATE „Developer“ įrankį ir komandomis pagalba susikurti tekstynus (angl. *corpus*), į tekstynus įkelti virtualius dokumentus, laikinai sukurtus atmintyje iš straipsnių, esančių duomenų bazėje, juos apdoroti su ANNIE *with defaults* [21] bei su anksčiau aprašytu taisyklių JAPE rinkiniu, pagal kurį atpažįstamos citatos.

Duomenys (citata, autorius, šaltinis) saugomi paprastos struktūros duomenų bazėje.

Programos veikimo principas gana paprastas, inicijuojama virtuali GATE aplikacija, vėliau sukuriamas tekstynas. Imama po 500 straipsnių iš duomenų bazės (po 500 dėl to, jog norint įsikelti daugiau dokumentų nepakanka darbinės atminties), iš straipsnių kuriami virtualūs dokumentai ir priskiriami tekstynui, pradedamas GATE anotavimas, ir gauti duomenys įrašomi į duomenų bazės lentelę. Iš duomenų bazės suindeksuojamos citatos į „Apache Solr” indeksą, tekstiniame faile pažymimas paskutinio apdoroto įrašo numeris, kad kitą kartą būtų apdorojami dar neapdoroti duomenys.

Buvo atliekami eksperimentai su periodiniu skenavimu ir pastebėta, jog interneto skeneris (angl. *web crawler*) baigia savo darbą vidutiniškai kas penkias valandas, o naujienos atsinaujina dažniau, todėl buvo nuspręsta skenuoti du kartus per dieną, kad būtų surenkama daugiau medžiagos apdoroti.

Surinkta medžiaga kas valandą buvo apdorojama anksčiau minėta programa ir atsimenama kiek duomenų apdorota, kad vėliau darbas būtų tęsiamas.

VII. REZULTATAI IR IŠVADOS

Naudojant JAPE ir GATE eksperimentinį tekstyną buvo gauti citatų išskyrimo kokybės rezultatai testuojant tekstyną su 500 dokumentų ir su 207 tiesioginėmis citatomis. Kokybę būtų galima pagerinti bandant aprašyti taisykles, kurios atpažintų net ir ne visiškai taisyklingai išskirtą citatą. Taip pat reiktų visame tekste ieškoti kas yra citatos autorius, kai prie citatos jis nėra įvardijamas. Remiantis rezultatais sukurtas įrankis, kuris periodiškai savaitę skenavo ir sužymėjo 2 279 citatų, dėl to paaiškėjo, jog įrankį dar reiktų tobulinti.

Norint pagerinti sistemos veikimo tikslumą reikėtų:

1) mėginti pritaikyti mašininio mokymo metodus ir palyginti su jau turimais rezultatais;
2) Sukurti metodą, skirtą, autorių atpažinti iš konteksto, kadangi šalia beveik 8 proc. esamų citatų tekste nėra paminėtas autorius.

LITERATŪROS SĄRAŠAS

[1] Cloudtweaks, „Cloud Computing Storage and the Exponential Growth of Data," [Tinkle]. Prieiga per internetą: http://www.cloudtweaks.com/2010/09/cloud-computing-storage-and-the-exponential-growth-of-data/. [Žiūrėta 2013 02 22].

[2] T. O. –. S. I. K. A. J. Curran. [Tinkle]. Prieiga per internetą: http://web.it.usyd.edu.au/research/conversazione_2011/okeefet.pdf. [Žiūrėta 2012 05 07].

[3] „Nutarimas „Dėl Privalomosios skyrybos taisyklių", [Tinkle]. Prieiga per internetą: http://www.vlkk.lt/lit/10118. [Žiūrėta 2012 05 09].

[4] R. Vyšniauskaitė, „LABA", 2009. [Tinkle]. Prieiga per internetą: http://vddb.laba.lt/fedora/get/LT-eLABa-0001:E.02~2009~D_20090629_124212-09058/DS.005.0.02.ETD. [Žiūrėta 2013 04 08].

[5] L. Thierry Poibeau, „Proper Name Extraction from Non-Journalistic Texts, Computational Linguistics in the Netherlands", 2001.

[6] „Hidden Markov model − Wikipedia, the Free Encyclopedia", [Tinkle]. Prieiga per internetą: http://en.wikipedia.org/wiki/Hidden_Markov_model. [Žiūrėta 2012 05 17].

[7] „Decision Tree Learning − Wikipedia, the Free Encyclopedia", [Tinkle]. Prieiga per internetą: http://en.wikipedia.org/wiki/Decision_tree_learning. [Žiūrėta 2012 05 17].

[8] „Maximum Entropy Modeling", [Tinkle]. Prieiga per internetą: http://homepages.inf.ed.ac.uk/lzhang10/maxent.html. [Žiūrėta 2012 06 19].

[9] „Conditional Random Field − Wikipedia, the Free Encyclopedia", [Tinkle]. Prieiga per internetą: http://en.wikipedia.org/wiki/Conditional_random_field. [Žiūrėta 2012 05 17].

[10] „Support Vector Machine − Wikipedia, the Fee Encyclopedia", [Tinkle]. Prieiga per internetą: http://en.wikipedia.org/wiki/Support_vector_machine. [Žiūrėta 2012 05 17].

[11] S. B. Kotsiantis, „http://www.informatica.si/", [Tinkle]. Prieiga per internetą: http://www.informatica.si/PDF/31-3/11_Kotsiantis%20-%20Supervised%20Machine%20Learning%20-%20A%20Review%20of...pdf. [Žiūrėta 2012 05 18].

[12] „GATE.ac.uk - family/developer.html", [Tinkle]. Prieiga per internetą: http://gate.ac.uk/family/developer.html. [Žiūrėta 2012 05 13].

[13] „Prudure Online Writing Lab", [Tinkle]. Prieiga per internetą: http://owl.english.purdue.edu/owl/resource/577/01/. [Žiūrėta 2012 05 05].

[14] „Tesla", Text Engineering Software Laboratory, [Tinkle]. Prieiga per internetą: http://tesla.spinfo.uni-koeln.de/nightlies/index.html. [Žiūrėta 2012 05 18].

[15] „Regular Expression Library", [Tinkle]. Prieiga per internetą: http://regexlib.com/DisplayPatterns.aspx. [Žiūrėta 2012 03 01].

[16] „GATE.ac.uk", [Tinkle]. Prieiga per internetą: http://gate.ac.uk/sale/thakker-jape-tutorial/GATE%20JAPE%20manual.pdf. [Žiūrėta 2012 03 03].

[17] R. T. Giuseppe Rizzo, „Evaluating Named Entity Recognition," [Tinkle]. Prieiga per internetą: http://nerd.eurecom.fr/ui/paper/Rizzo_Troncy-wekex11.pdf. [Žiūrėta 2012 03 22].

[18] The Stanford Natural Language Processing Group, „Evaluation in Information", [Tinkle]. Prieiga per internetą: http://nlp.stanford.edu/IR-book/pdf/08eval.pdf. [Žiūrėta 2012 05 02].

[19] „Precision and Recall - Wikipedia, the Free Encyclopedia," [Tinkle]. Prieiga per internetą: http://en.wikipedia.org/wiki/Precision_(information_retrieval). [Žiūrėta 2012 02 14].

[20] „Clojure Interface to GATE Project Functions", [Tinkle]. Prieiga per internetą: https://github.com/tokenmill/clojure-libs/tree/master/gate. [Žiūrėta 2012 04 02].

[21] GATE.ac.uk, „ANNIE: A Nearly-New Information Extraction System", [Tinkle]. Prieiga per internetą: http://gate.ac.uk/sale/tao/splitch6.html#chap:annie. [Žiūrėta 2013 02 25].

Reliacinių duomenų bazių duomenų integravimas taikant ontologijas

Tomas Trainys
Informacijos sistemų katedra
Kauno technologijos universitetas
Kaunas, Lietuva
tomas.trainys@stud.ktu.lt

Virginija Uzdanavičiūtė
Informacijos sistemų katedra
Kauno technologijos universitetas
Kaunas, Lietuva
virginija.uzdanaviciute@stud.ktu.lt

prof. Rimantas Butleris
Informacijos sistemų katedra
Kauno technologijos universitetas
Kaunas, Lietuva
rimantas.butleris@ktu.lt

Santrauka. **Kuriant naujas ar plečiant esamas informacijos sistemas (IS) susiduriama su heterogeninių duomenų bazių jungimo, integravimo bei duomenų pakartotinio panaudojimo problemomis. Yra atlikta tyrimų, eksperimentų, sukurta programinių prototipų, tačiau integravimo uždavinys vis dar iki galo nepritaikytas / neišspręstas. Konceptualaus integruojamų schemų suderinamumo modeliui įgyvendinti siūloma taikyti ontologijas, nes jos skatina dalytis žiniomis. OWL modelis pasižymi išraiškingumu, yra standartizuotas (W3C rekomendacija), vienareikšmiškai suprantamas, formalus, jam interpretuoti yra sukurta daug programinės įrangos. Straipsnyje aprašoma metodika, kaip integruoti duomenų šaltinius taikant ontologijas. Išanalizuotos kalbos ontologijoms sudaryti, aptartos galimos integravimo architektūros, integravimo procesas bei semantiniai konfliktai ir jų sprendimo būdai.**

Reikšminiai žodžiai: ontologija, OWL, duomenų šaltinių integravimo metodai, duomenų modelis, semantiniai konfliktai, heterogeniniai duomenų šaltiniai.

I. ĮVADAS

Per pastaruosius kelis dešimtmečius įmonės duomenims saugoti sėkmingai pritaikė įvairias duomenų bazių valdymo sistemas (DBVS). Plačiai paplito įvairių gamintojų DBVS – *MS SQL Server*, *Oracle*, *MySQL* ir kt. Istoriškai susiklostė, kad pirmiausia duomenų bazės (DB) buvo diegiamos laikantis decentralizuoto požiūrio, o įmonėms plečiantis, didėjant padalinių skaičiui ir jungiantis į didesnes korporacijas bei sindikatus, atsirado poreikis įmonių duomenis integruoti ir paruošti toliau naudoti. Integravimo klausimas tapo svarbus daugiau nei prieš trisdešimt metų ir iki šiol yra aktualus.

Reikia paminėti, kad pastaruoju metu dideliu tempu besikeičiant technologinei aplinkai bei plečiantis ir bendradarbiaujant įmonėms duomenų integracijos klausimas išlieka aktualus ir turi įtakos galutinei kuriamo produkto kainai. Be to, atsižvelgiant į didėjančią telekomunikacinių tinklų spartą, kuri globaliuose ir vietiniuose kompiuterių tinkluose gali siekti iki kelių GB/s, ir įvertinant šiuo metu eksploatuojamos techninės įrangos našumą (procesorinę galią), taip pat atminčių spartą, galima teigti, kad techninės įrangos sluoksnis yra pakankamai išvystytas didėlės apimties heterogeniniams duomenų šaltiniams integruoti. Todėl yra racionalu geografiškai paskirstytus duomenų šaltinius integruoti, tam pritaikant centralizuotos duomenų bazės (dar vadinamos federaline duomenų baze) metodus. Toks sprendimas leistų sumažinti DBVS administravimo kainą ir pagreitinti taikomosios programinės įrangos kūrimo procesą.

Centralizuotą duomenų bazę galima apibūdinti kaip sistemą, suteikiančią naudotojams galimybę, naudojant *SQL / SPARQL* užklausas, manipuliuoti geografiškai paskirstytų duomenų šaltinių duomenimis. Yra galimybių manipuliuoti ir heterogeninių duomenų šaltinių duomenimis, tačiau dažniausiai susiduriama su struktūriniais ir semantiniais konfliktais duomenų šaltinių schemos bei duomenų lygmenyse.

Heterogeniškumą galima apibūdinti keliais aspektais [13]: *struktūrinis* – apima skirtingus duomenų modelius; *sintaksinis* – apima skirtingas kalbas ir duomenų vaizdavimą; *sisteminis* – apima operacines sistemas ir techninę įrangą; *semantinis* – apima semantinį duomenų modelių suderinamumą.

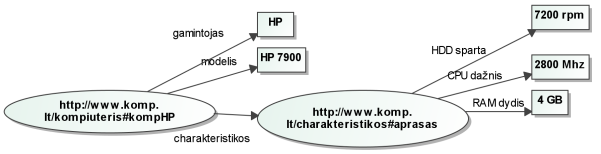
Darbas yra tiriamojo pobūdžio, apima tik semantinio suderinamumo klausimus heterogeninių duomenų šaltinių integravimo procese. Straipsnyje aptariami semantinio suderinamumo uždavinių sprendimo būdai, heterogeniniams duomenų šaltiniams taikant ontologijas, kaip aukštesnio lygmens duomenų šaltinių aprašymo modelį. Plačiai nagrinėjamas ontologijų pagrindu vykdomas integravimo procesas, aptariamos architektūros ir metodai, leidžiantys užtikrinti duomenų bazių homogeniškumą bei informacijos integralumą. Taip pat aprašomi įrankiai procesui įgyvendinti, pateikiama sprendimo metodika ir integravimo sistemos architektūros prototipas.

II. ONTOLOGIJOS IR JŲ APRAŠYMO KALBOS

„Ontologijos“ sąvoka informatikos srityje vartojama dalykinės srities konceptų specifikacijai apibūdinti [12]. Ontologija – aukšto semantinio lygmens dalykinės srities aprašas, kuris susideda iš dalykinės srities *konceptų, ryšių tarp konceptų, aksiomų ir apribojimų*. Tai yra duomenų modelis, apimantis klases, objektus, jų savybes ir ryšius [11]. Modelis suteikia galimybę detaliau nei *ER* modelis (angl. *entity relationship*) konceptualizuoti dalykinę sritį, sudaryti klasių hierarchiją bei sukurti ryšius tarp skirtingų dalykinių sričių. Vienas iš reikšmingų ontologijos modelio privalumų – išraiškingumas, žmogui ir kompiuterių sistemai suprantamas, teikiantis galimybę sudaryti dalykinių sričių žodynus. Formaliai ontologiją *O* galima aprašyti grafu [10] *G*,

turinčiu *Mo* baigtinę mazgų aibę ir baigtinę briaunų aibę *Bo*. Briaunos gali būti aprašomos trejetu m_1, m_2, p_1, kai m_1, m_2 yra *Mo* konceptai, o p_1 yra briaunos pavadinimas. Formaliai ontologijos koncepcinį duomenų modelį galima užrašyti $G_O = (M_O, G_O)$.

Ontologijoms sudaryti yra naudojami kelių kalbų poaibiai: *RDF*, *RDF shema* (angl. *resource description framework*). *RDF* modelis sudarytas iš trijų pagrindinių dalių: *subjekto* – bet kokio daikto, aprašyto *RDF*; *predikato* – *savybės*, charakteristikos, atributo ar ryšio, skirto ištekliui apibūdinti; *objekto* – *savybės reikšmės*, tai gali būti kitas išteklius arba duomenų tipo reikšmė, kurie kartu vadinami teiginiu. *RDF* grafo pavyzdys pateiktas 1 pav., jo tekstinis aprašas – 1 lentelėje.



1 pav. RDF grafas

1 lentelė. RDF GRAFO APRAŠAS

Subjektas	Predikatas	Objektas
http://www.komp.lt/kompiuteris#kompHP	Gamintojas	HP
http://www.komp.lt/kompiuteris#kompHP	Modelis	HP 7900
http://www.komp.lt/kompiuteris#kompHP	Charakteristikos	http://www.komp.lt/kompiuteris#aprasas
http://www.komp.lt/kompiuteris#aprasas	HDD sparta	7200 rpm
http://www.komp.lt/kompiuteris#aprasas	CPU dažnis	2800 Mhz

OWL (angl. *web ontology language*) – *RDF* ir *RDF-S* pagrindu sukurta ontologijų aprašymo kalba. *OWL* ontologija susideda iš konceptų ir ryšių tarp jų, apribojimo ir aksiomų. Toliau detaliai aptarsime *OWL* paketo struktūrą: pagrindinis *OWL* konstrukcinis elementas yra klasė, kalboje ji aprašoma *owl:Class* elementu. *OWL* klasė teikia galimybę grupuoti panašias charakteristikas turinčius išteklius. Savybe *rdfs:subClaccOf* yra aprašomi ryšiai tarp klasių, o savybė *rdfs:subPropertyOf* naudojama klasių hierarchijai sudaryti.

Ekvivalenčios klasės nustatomos naudojant konstrukcinį elementą *owl:equivalentClass*. Tuo pažymima, kad keli ištekliai (klasės, egzemplioriai, savybės) aprašo tą patį objektą. *OWL* klasė yra asocijuojama su klasės egzemplioriais, kurie yra vadinami klasės išplėtimais. Klasių egzemplioriai yra išvardijami naudojant konstrukcinį elementą *owl:oneOf*. Pagrindinės kalbos savybės yra: *owl:DatatypeProperty* ir *owl:ObjectProperty*, šiomis savybėmis yra sudaromi ryšiai tarp klasės egzempliorių bei priskiriami duomenų tipai.

OWL kalba yra aprašomos specifinės savybių charakteristikos, tokios kaip tranzityvumas, simetriškumas, inversija ir funkcinė priklausomybė. Šios charakteristikos aprašomos naudojant: *owl:SymmetricProperty*, *owl:InverseFunctionalProperty*, *owl:FunctionalProperty* ir *owl:TransitiveProperty* kalbos elementus.

Ryšių kardinalumas aprašomas naudojant elementus: *owl:cardinality*, *owl:minCardinality* ir *owl:maxCardinality*.

Sudėtinės klasės sudaromos naudojant savybes: *owl:intersectionOf*, *owl:unionOf* ir *owl:complementOf*, kuriomis vykdant sankirtos, sąjungos ir papildymo operacijas gaunamos sudėtinės klasės (dar vadinamos anoniminėmis klasėmis).

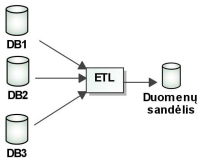
SPARQL (angl. *simple protocol and RDF query language*) – tai užklausų formulavimo kalba [22], skirta duomenų gavimui iš *RDF* grafo poaibių bei naujų RDF grafų formavimui. Pagrindiniai kalbos operatoriai: *SELECT* – grąžina rezultatus (pagal užklausoje pateiktus kintamuosius); *ASK* – identifikuoja informacijos egzistavimą; *DESCRIBE* – pateikia išteklius aprašantį *RDF* grafą; *CONSTRUCT* – naudojamas *RDF* grafui konstruoti.

III. DUOMENŲ ŠALTINIŲ INTEGRAVIMO ARCHITEKTŪROS IR METODAI

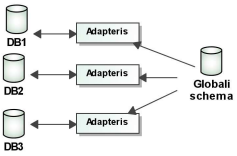
Duomenų, priklausančių skirtingiems duomenų šaltiniams, surinkimo, integravimo bei pateikimo uždaviniai gali būti sprendžiami taikant tradicinius metodus, tokius kaip duomenų sandėlių ir federalinių duomenų bazių [7], [13], [17] bei ontologijomis grindžiamus metodus. Toliau juos aptarsime detaliau.

A. Tradiciniai metodai

Duomenų sandėlio (angl. *warehouse*) architektūros (2 pav.) atveju duomenys iš duomenų šaltinių yra surenkami, transformuojami ir perkeliami (*ETL* – angl. *extract, transform, load*) į centralizuotą saugyklą [13]. Šios architektūros privalumas – duomenų gavybos operacijų (informacijos analizė, žinių išgavimas ir kt.) vykdymas nedarant įtakos paskirstytų duomenų šaltiniams. Naudotojai gali siųsti užklausas į didelius duomenų rinkinius. Paprastai tokios sistemos yra labai patikimos, užklausos greitai įvykdomos. Šios architektūros trūkumas – duomenų naujumas, kadangi naujausi duomenys būna paskirstytuose duomenų šaltiniuose ir turi būti sinchronizuojami su duomenų sandėliu.



2 pav. Duomenų sandėlio architektūra



3 pav. Globalios schemos architektūra

Globalios schemos architektūroje, skirtingai nei duomenų sandėlių sistemose, duomenys nėra perkeliami į pagrindinę duomenų saugyklą [13]. Autonominiai duomenų šaltiniai yra integruojami schemos lygmenyje. Tam yra sudaroma globali (3 pav.), integruojamų šaltinių schema (dar vadinama federaline schema arba virtualia duomenų baze). Integracija vykdoma susiejant duomenų šaltinių schemas su globalia schema. Sudarant globalią schemą svarbu suvienodinti šaltinių schemas, susieti schemų elementus ir išspręsti konfliktus tarp jų (pvz., skirtingi schemos elementų pavadinimai ir kt.).

Sistema, kurioje yra įdiegiama globali schema, vadinama globalia ar federaline sistema.

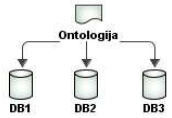
Federalinė sistema leidžia siųsti užklausas į visus integruotus duomenų šaltinius. Užklausos į duomenų šaltinius yra siunčiamos per adapterius (angl. *wrapper*), kurie atlieka pateiktų užklausų transformavimo ir rezultatų paketų sudarymo (angl. *encapsulation*) operacijas bei pateikia naudotojams rezultatus. Lyginant duomenų sandėlio ir federalines sistemas paminėtina, kad federalinės sistemos architektūros privalumas – duomenų naujumas. Pagrindinis architektūros trūkumas – ilgesnis užklausų siuntimo laikas.

B. Ontologijomis grindžiamas duomenų šaltinių integravimas

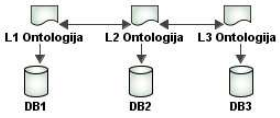
Ontologijos teikia galimybę detaliai konceptualizuoti dalykinę sritį, todėl šis semantinių technologijų metodas gali būti taikomas informacijos šaltinių integravimo uždaviniams spręsti. Dalykinės srities konceptualizavimo procese sudaromi ryšiai tarp dalykinės srities konceptų ir nustatomi semantiškai nesuderinami objektai. Paminėtina, kad pats konceptualizavimo procesas padeda semantiškai suderinti skirtingas duomenų šaltinių schemas. Taikant ontologijų metodą galima išspręsti skirtingų schemų terminų daugiaprasmiškumus bei terminų sąryšių konfliktus [4], [17], pavyzdžiui, nustatyti ekvivalenčius schemos elementus tarp duomenų bazių ir kt. Tokios problemos apskritai apibūdinamos terminu „semantinis heterogeniškumas“. Ontologijų metodas leidžia naudotojams siųsti užklausas į skirtingus duomenų šaltinius taikant semantinį modelį. Semantinis modelis – globalios schemos atitikmuo (lyginant su tradiciniais metodais). Literatūros šaltiniuose [31], [7], [17] yra pateikiami skirtingi minėtos integracijos įgyvendinimo metodai. Trumpai aptarsime kiekvieną iš jų.

Globalios ontologijos metodas. Taikant šį metodą sudaroma viena ontologija (4 pav.), kurioje aprašomi visi duomenų šaltiniai ir tokiu būdu gaunamas integruotas dalykinės srities vaizdas, kuris dar gali būti apibūdinamas kaip semantinis modelis. Šis metodas leidžia siųsti užklausas per vieną ontologiją. Kai kuriais atvejais gali būti naudojamos kelių specializuotų ontologijų kombinacijos (dar vadinamos modulinėmis arba agreguotomis ontologijomis). Globalios ontologijos metodas gali būti taikomas integruojant konceptualiai panašius informacijos šaltinius [3]. Nors metodo architektūros sudėtis gali atrodyti paprasta, tačiau ontologijai sudaryti reikia dalykinės srities eksperto (dažnai kelių), kuris / kurie žinotų visų duomenų šaltinių semantiką. Pagrindinis metodo trūkumas – semantinio modelio priklausomybė nuo informacijos šaltinių pakeitimų. Pavyzdžiui, vieno informacijos šaltinio schemos pakeitimas gali paveikti visą ontologiją bei kitų informacijos šaltinių ryšius.

Taikant *sudėtinių ontologijų metodą* yra sudaroma kiekvieno duomenų šaltinio atskira ontologija (lokali ontologija) bei pavaizduojami ontologijų ryšiai (angl. *mapping*), (5 pav.). Vaizduojant šiuos ryšius svarbu įvertinti integruojamų šaltinių ontologijų terminų semantinį heterogeniškumą. Šioje architektūroje [31] [7] užklausos yra siunčiamos per lokalias ontologijas, o vaizdavimas yra naudojamas lokalių užklausoms siųsti į kitus duomenų šaltinius per jų lokalias ontologijas. Šio metodo privalumas – nepriklausomumas nuo informacijos šaltinių pakeitimų. Pavyzdžiui, pakeitus vieno informacijos šaltinio ontologiją, ji nepaveiktų kitų, pavaizduotų informacijos šaltinių ontologijų. Literatūroje minima [7], kad gana sudėtinga taikyti šį metodą dėl tarpusavio ontologijų semantinio heterogeniškumo aspektų.

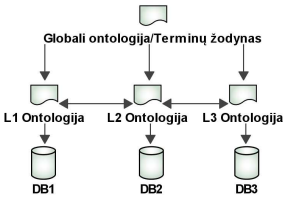


4 pav. Globalios ontologijos architektūra



5 pav. Sudėtinių ontologijų architektūra

Mišrus ontologijų metodas buvo pasiūlytas [7] prieš tai aptartų metodų trūkumams pašalinti. Šio metodo atveju duomenų šaltiniai, kaip ir sudėtinio metodo atveju, yra aprašomi lokalia ontologija (6 pav.), o visų duomenų šaltinių konceptai aprašomi bendra ontologija (dar vadinama žodynu [7]). Žodynas apima bazinius dalykinės srities terminus. Lokalios ontologijos sudaromos naudojant žodyne aprašytų terminų kompleksinius junginius. Ontologijos konstruojamos vartojant bazinius žodyno terminus, kurie leidžia semantiškai suderinti lokalias ontologijas. Naudotojai užklausas į šaltinius pateikia per bendrą žodyną. Šio metodo privalumai, palyginanti su prieš tai aptartais – paprastesnis papildomų informacijos šaltinių integravimas į esamą struktūrą bei geresnės sąlygos ontologijoms vystyti / plėsti.



6 pav. Mišrių ontologijų architektūra

Apibendrinant aptartus metodus reikėtų pažymėti, kad ontologijos ir duomenų šaltinių schemos yra panašios [9], žiūrint į abu modelius kaip į terminų žodyną, kurie abiejų modelių atveju aprašo dalykinę sritį. Be to, abu modeliai yra suvaržyti dalykinės srities žodyno terminų prasmių / reikšmių.

IV. SEMANTINIAI KONFLIKTAI, INTEGRAVIMO PROCESO PROBLEMOS IR METODAI JOMS SPRĘSTI

Skirtingų duomenų šaltinių integravimo klausimai plačiai analizuojami jau daugiau nei trisdešimt metų. Esami metodai ir algoritmai yra taikomi sistemų diegimo ir palaikymo etapuose. Daugeliu atvejų literatūroje minimų [3], [7], [8], [17], [23] integravimo problemų ir jų sprendimo būdų spręsti duomenų šaltinių heterogeniškumo klausimą yra panašūs. Integravimo procese susiduriama su semantiniais konfliktais, kurie charakterizuojami kaip skirtingos schemų ir duomenų lygmens loginės struktūros. Šie pagal tipą yra skirstomi į: sinonimus, homonimus, duomenų tikslumo bei kt. Minėtiems konfliktams nustatyti galima naudoti jau aptartos *OWL* kalbos savybes, kurios pateiktos 1 lentelėje. Toliau aptarsime pagrindinius konfliktų tipus bei jų nustatymo galimybes naudojant *OWL* savybes [5], [4], [6].

2 lentelė. OWL / RDF SAVYBĖS KONFLIKTAMS NUSTATYTI

Savybė	Panaudojimo aprašymas
owl:equivalentClass	Nustato ekvivalenčias klases.
owl:euqivalentProperty	Nustato ekvivalenčias ypatybes.
owl:sameIndividualAs	Nustato ekvivalenčius egzempliorius.
owl:disjointWith	Nustato semantiškai ekvivalenčius egzempliorius.
owl:differentFrom	Nustato, kad vienas egzempliorius skiriasi nuo kito.
owl:inverseOf	Pažymi, kad viena ypatybė yra priešinga kitai ypatybei.
owl:unionOf	Pažymi, kad nauja klasė yra gauta iš kitų klasių (sąjungos būdu).
owl:intersectionOf	Pažymi, kad nauja klasė yra gauta iš kitų klasių (sankirtos būdu).
owl:complementOf	Pažymi, kad nauja klasė yra egzistuojančių klasių papildinys.

A. Semantinių konfliktų nustatymas naudojant OWL

1) Pavadinimų konfliktai. Tokio tipo konfliktai įvyksta, kai du ir daugiau semantiškai ekvivalenčių /neekvivalenčių arba vienodą identifikatorių turinčių konceptų reprezentuoja tą patį objektą. Šie konfliktai įvardijami kaip sinonimai, antonimai ir homonimai:

a) Sinonimų konfliktai išsprendžiami suvienodinant ekvivalenčių konceptų pavadinimus. Ekvivalentūs konceptai yra nustatomi *OWL* kalbos savybėmis: *owl:equivalentClass, owl:equivalentProperty* ir *owl:sameInvidualAs.* Panagrinėkime pavyzdį (7 pav.), kuriame konceptai *Apskritis* ir *Provincija* yra ekvivalentūs; naudojant *owl:equivalentClass* elementą yra nustatomi tokio tipo konfliktai.

b) Antonimų konfliktai išsprendžiami išskiriant semantiškai neekvivalenčius konceptus. Tokio tipo konfliktai nustatomi *OWL* kalbos elementais *owl:disjointWith ir owl:differentFrom*. Pavyzdžiui (8 pav.), konceptai *BendrasĮvertinimas* ir *Pažymys* nėra ekvivalentūs. OWL elementu *owl:disjointwiths* yra išskiriami du neekvivalentūs konceptai, ir tokiu būdu išsprendžiamas antonimų konfliktas.

```
<owl:Class
rdf:ID="Provincija">
<rdfs:label>Provincija
</rdfs:label>
<owl:equivalentClass
rdf:resource="#Apskritis"/>
</owl:Class>
```

7 pav. Sinonimų konfliktų nustatymo aprašas

```
<owl:Class
rdf:ID="BendrasĮvertinimas">
<rdfs:label>Provincija</rdfs:label>
< owl:disjointwith
rdf:resource="#Pažymys"/>
</owl:Class>
```

8 pav. Antonimų konfliktų nustatymo aprašas

c) Homonimų konfliktai įvyksta, kai konceptai turi skirtingus identifikatorius, tačiau yra žymimi tokiais pačiais ar panašiais pavadinimais. Tokio tipo konfliktai dar yra vadinami identifikavimo konfliktais [5]. Juos reikia spręsti suteikiant atributams skirtingus pavadinimus [5]. Aptarkime pavyzdį (9 pav.): sakykime, turime dvi duomenų šaltinių schemas *A* ir *B*, kurių abiejų atributas yra *tvarkaraščio_Data*, tačiau ontologijose jie yra pavaizduoti skirtingai: *A* ontologijoje *studijuTvarkaraščioData*, o *B* ontologijoje *darboTvarkaraščioData*, tačiau abiejose ontologijose abu atributai žymimi žyme *tvarkaraščioData* (naudojamas elementas *rdfs:label*). Šiame pavyzdyje norint išvengti homonimų konfliktų, *A* ir *B* duomenų šaltinių schemose atributams reikia suteikti skirtingus pavadinimus.

```
<rdf:Property rdf:ID=" studijuTvarkaraščioData ">
<rdfs:label> tvarkaraščio_Data </rdfs:label>
</rdf:Property>
<rdf:Property rdf:ID=" darboTvarkaraščioData ">
<rdfs:label> tvarkaraščio_Data </rdfs:label>
</rdf:Property>
```

9 pav. Homonimų konfliktų nustatymo aprašas

2) Kompleksinių klasių konfliktai. Tokio tipo konfliktai įvyksta tarp klasių, kurios buvo sudarytos naudojant sankirtos / sąjungos operacijas, kai vieno koncepto elementų priklausomybė nuo kito koncepto elementų yra visiška arba dalinė. Aptarsime du tokių konfliktų atvejus:

a) Visiškos priklausomybės konfliktų gali įvykti, kai vienas konceptas yra visiškai priklausomas (įeina į kitą konceptą) nuo kito koncepto [15], [5]. Tokio tipo konfliktai yra išsprendžiami klasės atributus išskaidant į dalinius atributus. Šie konfliktai gali būti nustatomi konstrukciniu elementu *owl:unionOf.* Panagrinėkime pavyzdį (10 pav.), kuriame turime du konceptus *Vardas* ir *Pavardė*, ir *owl:unionof* elementas pažymi, kad minėti konceptai yra gauti sąjungos būdu. Kitaip sakant, jei klasė *A* gauta sąjungos būdu iš klasių *B* ir *C*, o *A* klasė nėra ekvivalenti *B* ir *C* klasėms, tada sujungiant *A* ir *B* arba *A* ir *C* klases gali įvykti visiškos priklausomybės konfliktas.

```
<rdf:Property rdf:ID="Pavardė">
< owl:unionOf
rdf:resource="#Vardas">
</rdf:Property>
```

10 pav. Visiškos priklausomybės konfliktų nustatymo aprašas

```
<owl:Class
rdf:ID="Asistentas">
<owl: intersectioOf
rdf:resource="#Absolventas"/>
</owl:Class>
```

11 pav. Dalinės priklausomybės konfliktų nustatymo aprašas

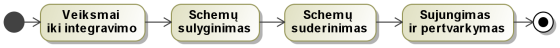
b) Dalinės priklausomybės konfliktų įvyksta tarp dviejų iš dalies sutampančių konceptų. Tokio tipo konfliktai nustatomi per *OWL* ypatybę owl:*intersectioOf*. Šie konfliktai gali būti išsprendžiami išskiriant iš dalies sutampančias konceptų dalis [5], [4]. Pavyzdyje, 11 pav., matyti dviejų schemų integracija, kur vienoje yra *Asistentas*, o kitoje – *Absolventas*. Čia abu konceptai (asistentas ir absolventas) iš dalies sutampa, nes kai kurie asistentai gali būti ir absolventai, bet ne visi asistentai yra absolventai ir ne visi absolventai yra asistentai. Šiuo atveju norint išvengti konfliktų, konceptai turi būti išskaidomi, sudarant atskiras klases: studijų nebaigusius asistentus, studijas baigusius asistentus ir studijuojančius asistentus.

B. Integravimo procesas

Heterogeninių duomenų šaltinių integravimo ontologijų pagrindu metodai yra įgyvendinami pagal procesą, pateiktą 12 pav.

Išskiriami šie proceso etapai: *pasiruošimas arba veiksmai iki integravimo.* Šiame etape yra atliekama integruojamų schemų analizė, sudaromas integravimo algoritmas ir nustatomas integravimo veiksmų eiliškumas; *schemų sulyginimo* etape suderinami schemų konceptai ir nustatomi

jų konfliktai; *schemų suderinimo* etape sprendžiami schemų konfliktai; *jungimo ir pertvarkymo* etape schemos sujungiamos ir pertvarkomos pagal iš anksto apibrėžtus kriterijus.



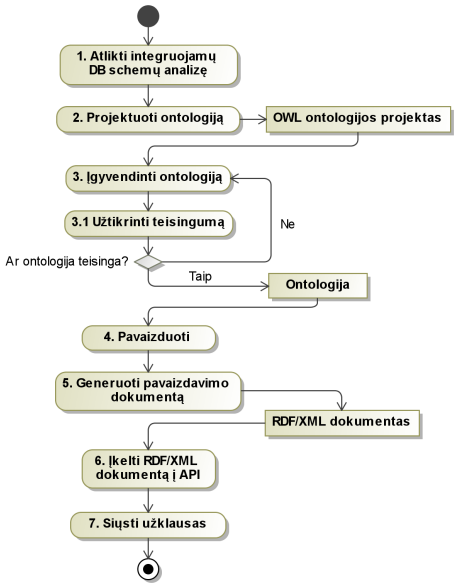
12 pav. Integravimo proceso diagrama

C. *Duomenų šaltinių integravimo ontologijų pagrindu metodika*

Šioje dalyje aptarsime siūlomą schematinio integravimo procesą, taikant globalios ontologijos metodą.

Heterogeninių duomenų šaltinių integravimo uždavinių procesas pavaizduotas 13 pav. Proceso diagramoje kiekvienas uždavinys pažymėtas skaitmenimis, kuriuos detaliai aptarsime.

Integravimo procesas pradedamas nuo integruojamų duomenų šaltinių schemų analizės, kurios metu nagrinėjami schemų konceptai, jų egzemplioriai ir ryšiai tarp jų.



13 pav. Duomenų šaltinių integravimo proceso diagrama

Antrame etape, remiantis integruojamų duomenų šaltinių analizės duomenimis, sudaromas ontologijos projektas. Tam naudojami programiniai įrankiai, tokie kaip *MagicDraw*, *Altova Semantic Works*, *NeOnToolKit* ir kiti bei ontologijų sudarymo metodologijos, pvz., *TOVE*, *UPON* ir kt. Metodologijos aprašo, kokio proceso reikėtų laikytis sudarant, jungiant ar pakartotinai naudojant ontologijas, taip pat aptariami ontologijų evoliucionavimo aspektai ir jų pritaikymas taikomajai programinei įrangai.

Trečiame etape įgyvendinama ontologija. Tam taip pat rekomenduojama naudotis ontologijų sudarymo metodologijomis bei programiniais įrankiais, pvz.: *Protege*, *NeOnToolKit*, *Altova Semantic Works* ir kt. Ontologijų konstravimas yra iteracinis procesas, kuris pradedamas nuo dalykinės srities nustatymo, po to išvardijami terminai, sukuriamos klasės ir jų egzemplioriai, sudaromos ypatybės (ryšiai) tarp klasių ir jų egzempliorių, klasių egzemplioriams priskiriamos reikšmės ir galiausiai sudaromos aksiomos. Ontologijos teisingumui nustatyti, yra naudojami loginių išvadų varikliai, tokie kaip *Palet*, *F++*, *HermiT* ir kt., kurie iš sudarytų faktų leidžia atlikti logines išvadas ir automatizuotai sudaro klasių hierarchiją. Kai kurie įrankiai (*Protege*, *NeOnToolKit*) turi integruotus išvadų variklius.

Ketvirtame etape pavaizduojamos ontologijos ir duomenų šaltinių schemos (angl. *mapping*). Tai daroma šiais būdais: pirmas – naudojant įrankį *NeONToolKit su ODE Mapster* plėtiniu. Įrankis leidžia grafiniu režimu sukurti duomenų šaltinių ir ontologijos vaizdavimo dokumentą; antras – naudojant *2DR* programinį paketą [18], kuriuo galima sugeneruoti vaizdavimo dokumentą. *2DR* programinis paketas naudoja *D2RQ* vaizdavimo kalbą, kurios aprašas pateiktas [19] šaltinyje; trečias – naudojant *W3C* [20], [21] tiesioginio vaizdavimo standartą, kuris leidžia vaizduoti duomenų šaltinius ir globalią ontologiją. Sudarant atvaizdą, gali būti naudojamas vienas iš antrame etape paminėtų ontologijų konstravimo įrankių.

Penktame etape iš sudaryto vaizdavimo modelio sugeneruojamas *RDF*, XML ar kito specifinio formato dokumentas. Paminėtina, kad sugeneruotas vaizdavimo dokumentas (jo standartas) turi būti suderintas su integravimo platforma, kurioje jis bus naudojamas.

Šeštame etape sugeneruotas modelis įdiegiamas taikomajai programinei įrangai, turinčiai *OWL / RDF* bibliotekas, arba gali būti naudojamas integravimui skirtuose technologiniuose sprendimuose, pvz., *2DR*, *Virtuoso* ir kt.

Septintame etape suderinama sistema ir siunčiamos semantinės užklausos į ontologijoje aprašytus duomenų šaltinius.

V. INTEGRAVIMO SISTEMOS PROTOTIPO KOMPONENTAI

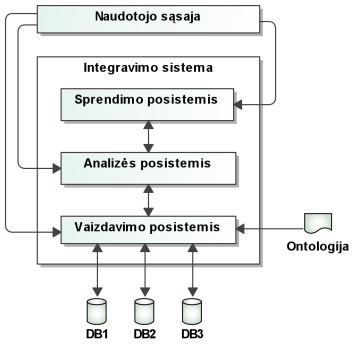
Aprašytų ontologijomis grindžiamų duomenų šaltinių integravimo metodams bei jų procesui įgyvendinti literatūros šaltiniuose [2], [16], [18], [19] pateikiamos įvairios galimos integravimo sistemų komponentų kompozicijos bei galimi, semantinio duomenų šaltinių integravimo sprendimai. Atsižvelgiant į technologinius pasiekimus ir prieinamas eksperimento priemones, šiame darbe aprašoma architektūra, kurią galima apibūdinti kaip susietų komponentų sistemą, kurios pagrindinės dalys nurodytos 14 pav. Toliau detaliai aptarsime kiekvieną iš jų.

Vaizdavimo posistemis vykdo atvaizdavimą tarp duomenų šaltinių ir globalios ontologijos bei leidžia vykdyti semantines užklausas. Vaizdavimui gali būti naudojami tokie įrankiai: *NeonToolKit*, *2RD*, *Protege* (naudojant *DataMaster* išplėtimą), *2DRQ* ir *W3C* kalbos [20] ir *Jena* ar *Sesam* adapteriai.

Analizės posistemis vykdo semantinių konfliktų nustatymo operacijas, o aptiktus konfliktus perduoda *sprendimo posistemiui*, kuris atlieka objektų transformavimo operacijas. Literatūroje [11] aptariami įvairūs galimi šio komponento įgyvendinimo ir diegimo atvejai. Tai gali būti *SOA* (angl. *service oriented architecture*). Vienas iš *SOA* atvejų *XML* technologijos pagrindu veikiantis servisas, kuriuo yra vykdomos objektų transformavimo operacijos. Kitas *RDF / OWL* pagrindu veikiantis servisas *SWS* (angl. *semantic web service*) [11]. Konfliktai paprastai nėra sprendžiami

visiškai automatizuotai, dažniausiai – pusiau automatizuotai, leidžiant naudotojui priimti konkrečiam atvejui tinkamą sprendimą.

Naudotojo sąsaja skirta sistemai konfigūruoti, ontologijai įkelti, konfliktams valdyti ir užklausoms į duomenų šaltinius siųsti.



14 pav. Integravimo sistemos architektūra

Komponentai gali būti sujungiami naudojant programavimo technologijas (pvz., *JAVA, .Net ir kt.*), kurios turi bibliotekas dirbti su *RDF / OWL*, pvz., *Java.API, DotNet.RDF* ir kt.

VI. IŠVADOS

Šiame straipsnyje išanalizuotos ontologijų panaudojimo galimybės duomenų bazių integravimo uždaviniams įgyvendinti. Aprašyti semantinių konfliktų tipai bei jų nustatymo galimybės taikant *OWL* kalbą. Automatizuotas reliacinių duomenų šaltinių integravimas yra komplikuotas uždavinys, nes integruojant schemas reikalinga nustatyti neatitikimus (sinonimai, antonimai, homonimai) tarp schemos elementų ir duomenų. Taikant vien schematinio integravimo metodą, be duomenų suderinimo, galima suderinti skirtingas schemas, kurios gali būti semantiškai skirtingos, bet struktūriškai panašios; tačiau daugelyje taikomųjų uždavinių svarbu pasiekti abiejų lygių suderinamumą. Todėl prieš integruojant duomenis tikslinga identifikuoti ir išspręsti semantinius konfliktus. Tam tikslui pasiūlyta semantinius konfliktus ir jų tipus aprašyti *OWL* kalba, duomenims suderinti – įvykdyti duomenų transformacijas *XML* ar *SWS* servisais, panaudojant *SOA* architektūrą.

Tolesniuose tyrimuose numatoma išanalizuoti duomenų lygmens semantinius konfliktus, jų sprendimų metodus, panaudoti prototipą duomenų šaltiniams sujungti, taikant globalios schemos architektūrą, bei sudaryti algoritmą automatizuotai nustatyti schemos elementus ir iš dalies automatizuotai juos spręsti.

[1] S. Ram, J. Park, „Semantic Conflict Resolution Ontology (SCROL): An Ontology for Detecting and Resolving Data and Schema-Level Semantic Conflict“, IEEE Transactions on Knowledge and Data Engineering, vol. 16, no. 2, pp. 189-202, 2004.

[2] F. Giunchiglia, A. Autayeu, J. Pane, „S-MATCH: „An Open Source Framework for Maching Lightweight Ontologines“, 2010, DISI - Via Sommarive 14 - 38123 Povo - Trento (Italy).

[3] M. Gagnon, „Ontology-Based Integration of Data Sources“, Information Fusion, 10th International Conference, ISBN: 978-0-662-45804-3, pp. 1-8, 2007.

[4] J. Wen, S. Zhang, Z. Yan, „SLCO and DLCO: Two Ontologies for Detecting and Resolving Schema and Data-Level Semantic Conflicts“, Information and Automation, 2009. ICIA '09. International Conference, pp. 637-642, 2009.

[5] C. Li, T. Wang Ling, „OWL-Based Semantic Conflicts Detection and Resolution for Data Interoperability“, Conceptual Modelling for Advanced Application Domains Lecture Notes in Computer Science, vol. 3289, pp. 266-277, 2004.

[6] Y. Zhao, S. Zhang, „Ontology-based Model for Resolving the Data-Level and Semantic-Level Conflicts“, Information and Automation, 2009. ICIA '09. International Conference, pp. 455-459, 2009; M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[7] H. Wache, T. Vogele, U. Visser, H. Stuckenschmidt, G. Schuster, H. Neumann, S. Hubner, „Ontology-Based Integration of Information – A Survey of Existing Approaches“, Proceedings of IJCAI-01 Workshop: Ontologies and Information Sharing. Seattle, WA, pp. 108-117, 2001.

[8] J. Wang, J. Lu, Y. Zhang, et al., „Integrating Heterogeneous Data Source using Ontology “, Journal of Software, vol. 4, no. 8, pp. 843-850, 2009.

[9] P. Shvaiko, J. Euzenat, „A Survey of Schema-Based Matching Approaches“, Journal on Data Semantics, vol. IV, pp.146-171, 2005.

[10] X. Liu, S. Akram, A. Bouguettaya, Change Management for Semantic Web Services. Springer Science+Business Media. LLC, pp. 14-30, 2011.

[11] D. Gaševič, D. Djurič, V. Devedžič, Model Driven Engineering and Ontology Development. 2nd Edition, Springer, pp. 110-138; 133-141, 2009.

[12] D. Allemanng, J. Hendler, Semantic Web for Working Ontologist, Efective Modeling in RDFS and OWL. Second Edition, Elsevier Inc., pp. 13-22, 125-279, 2011.

[13] S. Busse, R. D. Kutsche, U. Leser, H. Weber, Federated Information Systems: Concepts, Terminology and Architectures. Technische Universität Berlin, Fachbereich 13-Informatik, 1999.

[14] W3C, OWL 2 Web Ontology Language Document Overview (Second Edition). Prieiga per internetą: http://www.w3.org/TR/owl2-overview/.

[15] W3C, RDF Vocabulary Description Language 1.0: RDF Schema. Prieiga per internetą: http://www.w3.org/TR/2003/PR-rdf-schema-20031215/.

[16] V. Uzdanavičiūtė, R. Butleris, „Oracle DBVS galimybės ontologijomis grindžiamo duomenų integravimo procese“. Proceedings of the 15th Master and PhD conference „Information Society and University studies“ (IVUS 2010), 13th May 2010, Kaunas, ISSN-4824, pp. 169-174, 2011.

[17] V. Uzdanavičiūtė, R. Butleris, „Ontology-Based Foundations for Data Integration“. BUSTECH' 2011: Proceedings of The First International Conference on Business Intelligence and Technology, 2011.

[18] D2R Server: Accessing Databases with SPARQL and as Linked Data. Prieiga per internetą: http://d2rq.org/d2r-server.

[19] The D2RQ Mapping Language. Prieiga per internetą: http://d2rq.org/d2rq-language.

[20] OWL 2 Web Ontology Language Mapping to RDF Graphs. Prieiga per internetą: http://www.w3.org/TR/owl2-mapping-to-rdf/.

[21] J. F. Sequeda, M. Arenas, D. P. Miranker, „On Directly Mapping Relational Databases to RDF and OWL“, Proceedings of the 21st International Conference on World Wide Web, April 16-20, Lyon, France, pp. 649-658, 2012.

[22] SPARQL Query Language for RDF. Prieiga per internetą: http://www.w3.org/TR/rdf-sparql-query/.

[23] C. E. Naiman, A. M. Ouksel, „A Classification of Semantic Conflicts in Heterogeneous Database Systems“, Journal of Organizational Computing, vol. 5, no. 2, pp. 167-193, 1995.

Lietuvos Seimo narių balsavimų tyrimas pasitelkiant socialinių tinklų analizę: tinklo konstravimo metodologiniai aspektai

Rūta Užupytė
Vytauto Didžiojo universitetas
Baltijos pažangių technologijų institutas
r.uzupyte@gmail.com

Vaidas Morkevičius
Kauno technologijos universiteto
Politikos ir viešojo administravimo institutas
vaidas.morkevicius@ktu.lt

Santrauka. Kompiuterinių technologijų pažanga įgalina vis plačiau taikyti duomenų gavybos metodus socialiniuose moksluose. Šiame straipsnyje analizuojamos galimybės pritaikyti socialinių tinklų analizės metodus Lietuvos Seimo narių balsavimų tyrimams. Aptariami svarbiausi su šia dalykine sritimi susiję tinklo konstravimo etapai: balsavimų kodavimas, balsavimo panašumo įvertinimo matai, grafo (tinklo) vizualizavimo metodai ir skirtingų panašumų slenksčių parinkimas. Straipsnyje parodoma, kokią įtaką šie veiksniai daro galutiniam rezultatui.

I. Įvadas

Parlamento narių balsavimų analizė nėra nauja sritis politikos moksluose, pirmieji bandymai pateikti [1], [2]. Tačiau įvairių institucijų, pavyzdžiui, teismų, Parlamento, tarptautinių organizacijų balsavimų analizėje vis dažniau taikomi įvairūs duomenų gavybos metodai [3], [4], [5], [6], [7], [8], [9]. Naujų metodų taikymą politologijos srityje įgalino informacinių technologijų plėtra ir kompiuterių pajėgumo augimas. Iki šiol Lietuvos Seimo narių balsavimų tyrimuose buvo taikoma homogeniškumo analizė [10], [11], [12], daugiamačių skalių analizė [12], [13] bei klasterizavimas [13], [14]. Kadangi vizualizacija yra vienas iš būdų efektyviai perduoti informaciją, šiam tyrimui buvo pasirinktas socialinių tinklų metodas kaip nauja alternatyva Parlamento balsavimų analizei. Vienas iš pagrindinių tikslų buvo sukurti tinklą (grafą), vaizduojantį Lietuvos Respublikos Seimo narių balsavimų panašumus (skirtumus). Socialiniam tinklui kurti buvo panaudoti 2008–2012 m. kadencijos duomenys.

Taigi straipsnio II skyriuje apibrėžiama „socialinių tinklų“sąvoka bei jų interpretacija Lietuvos Seimo balsavimų analizei. III skyriuje aprašomi tyrime naudoti duomenys ir jų paruošimas grafo konstravimui. IV skyriuje analizuojami įvairūs socialinių tinklų vizualizavimo aspektai. Tyrimo rezultatai pristatomi V, o išvados pateiktos - VI skyriuje.

R. Užupytė ir V. Morkevičius dėkoja už Lietuvos mokslo tarybos projekto „Studentų mokslinės veiklos skatinimas“(VP1-3.1-ŠMM-01-V-02-003) paramą. Projektas yra finansuojamas pagal Žmogiškųjų išteklių plėtros veiksmų programos 3 prioritetą „Tyrėjų gebėjimų stiprinimas“Europos socialinio fondo ir Lietuvos Respublikos valstybės biudžeto lėšomis.

II. Socialiniai tinklai

Socialinis tinklas yra teorinis modelis, naudojamas socialiniuose moksluose siekiant tyrinėti grupių, organizacijų ar net atskirų visuomenės narių ryšius. Literatūros šaltiniuose pateikiamas tikslesnis šios sąvokos apibrėžimas:

> **Socialinis tinklas** – socialinė struktūra, susidedanti iš individų (ar organizacijų), vadinamųjų „viršūnių“(angl. *node*), kurie yra susiję vienu ar keliais tarpusavio ryšiais [15].

Pagrindinė šio metodo idėja – išnagrinėti socialinę sąveiką analizuojant tinklo vienetų ryšių savybes, užuot analizavus pačių vienetų savybes. Šis požiūris yra naudingas nagrinėjant tam tikrus realaus pasaulio reiškinius, tačiau reikėtų nepamiršti, jog metodas nepalieka daug vietos individualių atvejų analizei. Nepaisant to, tai – vienas iš būdų, leidžiančių atlikti santykių (ryšių) analizę vaizdiniame ir matematiniame lygmenyje.

Socialinis tinklas (grafas) apibrėžiamas kaip aibių pora $G = (V, E)$, kurią sudaro:

V – viršūnių aibė;
E – viršūnių porų $e := (x, y) =: xy,\ x, y \in V$ arba lankų aibė.

Grafas, kurio briaunos, jungiančios viršūnes, neturi krypties, vadinamas neorientuotu: $e := (x, y) =: xy = yx,\ x, y \in V$. Priešingu atveju (briaunos turi kryptis) grafas yra vadinamas orientuotu. Tinklo briaunai galima suteikti ne tik kryptį, bet ir skaitinę reikšmę (svorį): $e \in E$ priskiriamas skaičius $w(e)$ vadinamas briaunos svoriu, $w : E \to \mathbb{R}_+$[1] – svorio funkcija. Tokiu atveju socialinis tinklas vadinamas svertiniu (svoriniu) (angl. *weighted*).

A. Parlamento socialinis tinklas

Lietuvos Seimą [2] sudaro 141 parlamentaras (šis skaičius gali kisti dėl parlamentarų „migracijos“). Vienu iš parlamentarus siejančių ryšių galima laikyti panašų balsavimą. Tokiu atveju Seimo narius galima laikyti grafo viršūnėm, o jų balsavimų panašumus grafo lankais. Tokiu atveju reikia atsakyti į tokius klausimus:

[1] Teigiamų realiųjų skaičių aibė.
[2] Žr. www.lrs.lt.

- Kaip paversti balsavimus panašumais (ryšiais)?
- Kaip skaičiuoti panašumus?
- Kada panašumas tampa pakankamai didelis ir galima sakyti, kad du panašiai balsuojantys Seimo nariai yra „susiję“ (pagal balsavimą)?

III. Duomenų paruošimas

Tyrime naudojami 2008–2012 metų kadencijos Lietuvos Respublikos Seimo (LRS) balsavimai – iš viso 4 951 balsavimas. Duomenys saugojami matricoje D_{ij}. Šios matricos eilutes sudaro Seimo nariai, o stulpelius – įvairių teisės aktų projektai. Elementas d_{ij} yra i-ojo Seimo nario balsavimo reikšmė dėl j-ojo teisės akto projekto. Balsavimo profiliu vadinamas vektorius, kurį sudaro atitinkamo Seimo nario balsavimų rezultatai, t. y. matricos eilutė D_i.

A. Balsavimų kodavimas

Seimo narių balsavimų rezultatų duomenys yra kategorinio tipo, o jų galimų reikšmių aibė: už – balsavo už įstatymo projektą ar pasiūlymą; prieš – balsavo prieš įstatymo projektą ar pasiūlymą; susilaikė – balsavimo metu susilaikė; nebalsavo – užsiregistravo balsavime, bet nebalsavo; nedalyvavo – balsavimo metu nedalyvavo posėdyje[3]. Kadangi įvairūs metodai reikalauja kiekybinių reikšmių, duomenys yra perkoduojami. Atliekant tyrimą buvo pasirinkti trys perkodavimo variantai:

1,0,0,0,0,0	Už - 1	Prieš – 1
Už – 1	Prieš – -1	Susilaikė – 0
Prieš – 0	Susilaikė – 0	Nedalyvavo – -1
Susilaikė – 0	Nedalyvavo – 0	Nebalsavo – -2
Nedalyvavo – 0	Nebalsavo – 0	Nėra
Nebalsavo – 0	Nėra	informacijos –
Nėra	informacijos – 0	-10
informacijos –0	**2,1,0,-1,-2,-10**	
1,-1,0,0,0,0	Už – 2	

Kodavimo **2,1,0,-1,-2,-10** tikslas – kiekvienai reikšmei priskirti skirtingą įvertį (duomenų paruošimas Hemingo matui).

B. Panašumo matai

Norint balsavimus paversti panašumais (ryšiais) reikia įvertinti atstumus tarp balsavimų vektorių. Šiam tikslui buvo pasirinkti trys matai: normalizuoti ir transformuoti į panašumus Minkowskio ir Hemingo atstumai.

1) **Minkowskio atstumas**
Atstumas tarp dviejų taškų apibrėžiamas formule:

$$d(x,y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^p .$$

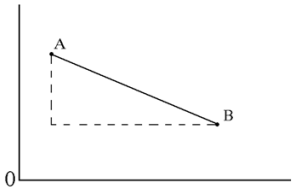
Atliekant tyrimą buvo naudojami atskiri šio mato atvejai: $p = 2$ – Euklido atstumas, $p = 1$ – Manheteno atstumas.

a) **Euklido atstumas**
Tai – trumpiausias atstumas tarp dviejų taškų:

$$d(x,y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

Siekiant normalizuoti atstumus gauta reikšmė yra dalijama iš balsavimų skaičiaus kvadratinės šaknies ir atimama iš vieneto:

$$s(x,y) = 1 - \frac{d(x,y)}{\sqrt{n}}.$$



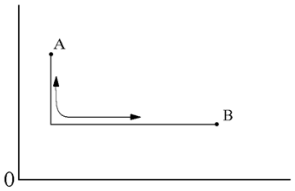
1 pav. Euklido atstumas dvimatėje plokštumoje

b) **Manheteno atstumas**
Šis atstumas, dar vadinamas miesto kvartalo metrika, yra lygus absoliučių skirtumų tarp visų stebėjimo porų reikšmių sumai:

$$d(x,y) = \sum_{i=1}^{n} |x_i - y_i|.$$

Siekiant normalizuoti atstumus gauta reikšmė yra dalinama iš balsavimų skaičiaus ir atimama iš vieneto:

$$s(x,y) = 1 - \frac{d(x,y)}{n}.$$



2 pav. Manheteno atstumas dvimatėje plokštumoje

2) **Hemingo atstumas**
Šis atstumo matas gali būti naudojamas tik vienodo ilgio vektoriams. Jis yra lygus pozicijų, kuriose

[3] Be to, kai nėra informacijos apie parlamentaro dalyvavimą/nedalyvavimą balsavime ir pan. (pavyzdžiui, jis dar nebuvo Seimo narys arba jo jau nebėra Seime), tokie variantai žymimi kodu „nėra informacijos“.

vienas vektorius skiriasi nuo kito, skaičiui:

$$d(x,y) = \sum_{i}^{n} d_i(x_i, y_i),$$

$$d_i(x_i, y_i) = \begin{cases} 1, & \text{jei } x_i \neq y_i \\ 0, & \text{jei } x_i = y_i \end{cases}$$

Pavyzdžiui, atstumas tarp vektoriaus $A = (110010)$ ir vektoriaus $B = (111011)$ yra lygus 2.
Siekiant normalizuoti atstumus gauta reikšmė yra dalijama iš balsavimų skaičiaus ir atimama iš vieneto:

$$s(x,y) = 1 - \frac{d(x,y)}{n}.$$

IV. Socialinių tinklų vizualizavimas

Kitas svarbus etapas atliekant socialinių tinklų analizę yra tinkamų socialinio tinklo vaizdavimo metodų ir būdų parinkimas. Grafo išdėliojimas yra gana subjektyvus dalykas, priklausantis tiek nuo konkrečios situacijos, tiek nuo grafo kūrėjo, tačiau tinklas turi atitikti tam tikrus bazinius reikalavimus: lankų susikirtimų skaičius turi būti minimalus; braižant lankus turi būti vengiama netiesių linijų; vaizdavimo erdvė turi būti minimali. Šie grafo išdėstymo uždaviniai įgyvendinami pasitelkiant programinių paketų vizualizavimo metodus.

A. *Grafų vizualizavimo metodų palyginimas*

Grafų braižymo algoritmų tikslas – išdėlioti viršūnes dvimatėje ar trimatėje erdvėje taip, kad lankai būtų apytikriai vienodo ilgio, o susikirtimų skaičius būtų minimalus. Tai aktualu siekiant efektyviai vizualizuoti Seimo narių balsavimų panašumus (skirtumus) bei skaičiuojant socialinių tinklų charakteristikas. Šiam tikslui buvo atlikta pagrindinių tinklo konstravimo metodų analizė atsižvelgiant į tokius kriterijus:

- panašumų (lankų) kaičių[4];
- objektų (viršūnių) skaičių;
- objektų (viršūnių) tarpusavio padėtį.

Grafo (tinklo) vizualizavimo metodai:

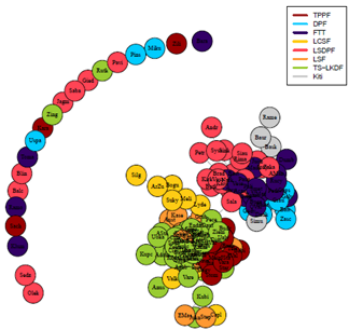
DrL – tai jėgos laukų pagrindu veikiantis (angl. *force-based*) algoritmas, kuris yra taikomas, kai socialinį tinklą sudaro daugiau nei 1 000 viršūnių. Taigi Seimo narių balsavimams vizualizuoti šis metodas netinka.

LGL – tai dar vienas tinklo vizualizavimo metodas, skirtas grafams turintiems daug tarpusavyje sujungtų viršūnių. Šis algoritmas gali būti naudojamas vaizduoti išskirtims, nes viršūnės, neturinčios lankų su kitomis viršūnėmis, išdėstomos grafo nuošalyje.

Atsitiktinis algoritmas pirmiausia atsitiktinai išdėlioja grafo viršūnes, po to sujungia jas lankais. Kadangi Seimo narių balsavimų analizei viršūnių tarpusavio padėtis yra labai svarbi, šis metodas netinka.

Skritulinis grafų konstravimo metodas leidžia vienodu atstumu išdėlioti viršūnes ant vienetinio apskritimo krašto ir sujungti jas lankais. Šis metodas neatsižvelgia į viršūnių tarpusavio padėtį, todėl balsavimų analizei netinka.

Kamada ir Kawai algoritmas veikia jėgos laukų pagrindu. Šį metodą rekomenduojama taikyti grafams, turintiems ne daugiau 100 viršūnių, kurios yra tarpusavyje susietos. Išbandžius šį metodą Seimo narių balsavimų analizėje gauti rezultatai parodė, kad algoritmas susidoroja ir su didesniu skaičiumi viršūnių (≈141), tačiau atvejais, kai grafas turi išskirčių, taikant metodą jos išdėstomos grafo nuošalyje, o pats grafas pavaizduojamas nedidelėje erdvės dalyje, todėl sunku detaliai analizuoti jo struktūrą (žr. 3 pav.).



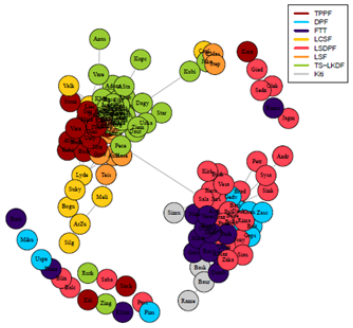
3 pav. Panašumų skaičius – 10 proc.; panašumo matas – Hemingo; grafo konstravimo metodas – Kamada ir Kawai).

Fruchtermano ir Reingoldo algoritmas taip pat veikia jėgos laukų pagrindu. Jis naudojamas vaizduoti grafams, turintiems ne daugiau 1 000 viršūnių. Gauti tyrimo rezultatai parodė, kad šis algoritmas labiausiai tinka Seimo narių balsavimams vizualizuoti – net ir išskirčių atvejais grafas yra paskleidžiamas erdvėje, ir yra įgalinama detali jo struktūros analizė (žr. 4 pav.).

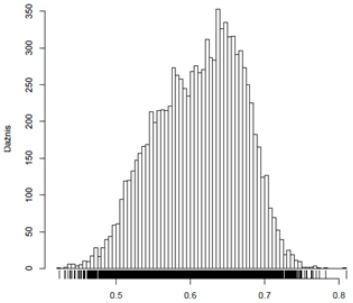
B. *Panašumų skaičiaus slenkstis*

Analizuojant Lietuvos Seimo narių socialinio tinklo konstravimo galimybes taip pat reikia įvertinti, kada balsavimų panašumas tampa pakankamai didelis, jog būtų galima teigti, kad du Lietuvos Seimo nariai balsuoja panašiai (žr. 5 pav.). Panašumo mato reikšmė, kuri nurodo ribą tarp reikšmingų ir nereikšmingų ryšių, vadinama panašumo slenksčiu. Konstruojant socialinį tinklą vaizduojami tik didelį panašumą turintys ryšiai, t. y. panašumo matai, viršijantys panašumo slenkstį.

[4] *igraph* paketas leidžia, nustačius slenksčius, parinkti lankų skaičių, kurie bus naudojami sudarant grafą. Atliekant Seimo narių balsavimų analizę, lankų skaičius (proporcija) grafuose buvo keičiamas pagal balsavimų panašumų įverčių dydžius.



4 pav. Panašumų skaičius – 10 proc.; panašumo matas – Hemingo; grafo konstravimo metodas – Fruchtermano ir Reingoldo



5 pav. 2008−2012 m. kadencijos Lietuvos Seimo narių balsavimų panašumų histograma (panašumo matas – Euklido; kodavimas **1,-1,0,0,0,0**).

Balsavimo panašumo slenksčiams nustatyti buvo apskaičiuoti šie kvantiliai: $1/10, 1/5, 1/4, 1/3, 2/5, 1/2$ (mediana),$3/5, 2/3, 3/4, 4/5, 9/10$.

V. Tyrimo rezultatai

A. Programinės įrangos apžvalga

Programinė įranga, skirta socialinių tinklų analizei, gerokai palengvina tinklų modeliavimą, vertinimą bei vizualizaciją: įvairūs išdėstymo algoritmai įgalina prasmingai vaizduoti socialinę struktūrą, o funkcijos užtikrina grafo statistikų apskaičiavimą. Programinės įrangos parinkimo kriterijai:

- duomenų paruošimas socialiniam tinklui braižyti;
- vizualizacijos galimybės;
- socialinių tinklų analizės metodai.

Aptarsime keletą populiariausių specializuotų socialinių tinklų analizės programų:

„Gephi“ – programa, skirta braižyti grafams, tačiau jai trūksta kai kurių analizės metodų: klikų radimo, daugiamačių skalių analizės. Svarbus trūkumas – įvesties duomenys jau turi būti paruošti grafui kurti.

„Pajek“ – paketas, skirtas dirbti su dideliais grafais (viršūnių ir lankų skaičius neribojamas). Pasižymi galimybe vizualizuoti tinklą trimatėje erdvėje, tačiau turi tiktai keletą pagrindinių analizės metodų.

„Ucinet“ – programa, išsiskirianti grafo charakteristikų ir statistinių metodų gausa, tačiau neturinti tinklų vizualizavimo procedūros.

Kita įrankių klasė yra programų bibliotekos. Pagrindinis privalumas, palyginti su specializuotomis programomis yra beveik neribotos duomenų paruošimo socialinio tinklo braižymui (duomenų (per)kodavimo, atstumų matricos sudarymo ir pan.) galimybės. Aptarsime dvi bibliotekas: **„NetworkX“** ir **„igraph“**. Šios bibliotekos skirtos socialinių tinklų analizei programos **R** aplinkoje, o **„NetworkX“** taip pat gali būti naudojama ir **„Python“** programoje. Lyginant šias bibliotekas išskirti keli **„igraph“** bibliotekos pranašumai:

- daugiau bendruomenių aptikimo algoritmų;
- galimybė vizualizuoti grafus ne tik dvimatėje, bet ir trimatėje erdvėje;
- galimybė atlikti klasterinę analizę;
- pasižymi greitesniu kompiliavimu.

Taigi, socialinių tinklų analizei atlikti buvo pasirinkta **R** programos biblioteka **„igraph“**.

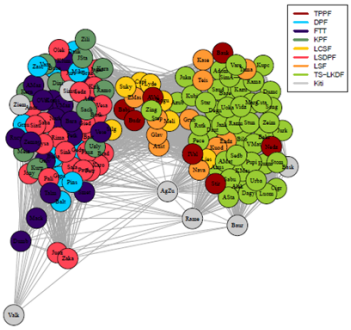
B. Balsavimų kodavimo analizė

Siekiant įvertinti skirtingų balsavimų kodavimų įtaką rezultatams buvo palyginta dviejų socialinių tinklų, gautų naudojant skirtingus duomenų kodavimo atvejus, struktūra (žr. 6 pav.). Akivaizu, jog naudojant **1,0,0,0,0,0** ir **1,-1,0,0,0,0** koduotes gaunami skirtingi, tačiau iš esmės panašūs rezultatai. Taigi kodavimo parinkimas neturi didelės įtakos socialinių tinklų analizės rezultatams. Kiti autoriai taip pat buvo atlikę balsavimų kodavimo įtakos analizę, taikydami daugiamačių skalių analizės metodą [12]. Tyrimų išvada analogiška: skirtingi balsavimų kodavimo būdai turi tam tikrą poveikį rezultatams, tačiau skirtumas labai nedidelis.

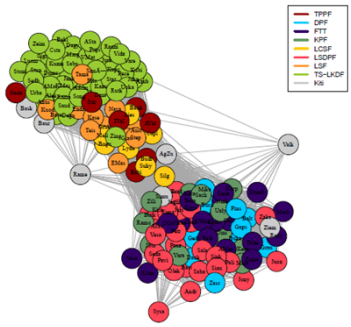
C. Panašumo matų įtakos analizė

Balsavimo profilių panašumo matų įtakos vertinimui tarpusavyje buvo palyginti du socialiniai tinklai, sukonstruoti naudojant Euklido ir Manheteno atstumus (žr. 7 pav.). Nesunku pastebėti, kad abiem atvejais rezultatai labai panašūs. Kiti autoriai taikydami, klasterizavimą balsavimų analizėje, taip pat tyrė panašumo matų įtaką rezultatams [13]. Naudojant tiek Manheteno, tiek Euklido matus buvo gauti beveik identiški rezultatai. Minkowskio panašumo matai buvo nagrinėti ir balsavimų analizėje, grindžiamoje daugiamačių skalių analize [12]. Šiuo atveju gauti skirtingi rezultatai taikant Euklido ir Manheteno atstumų matus.

Vertinant Lietuvos Seimo narių balsavimo panašumus pagal Hemingo matą (žr. 8 pav.) gaunami panašūs rezultatai kaip ir taikant Minkowskio atstumus. Tačiau naudo-

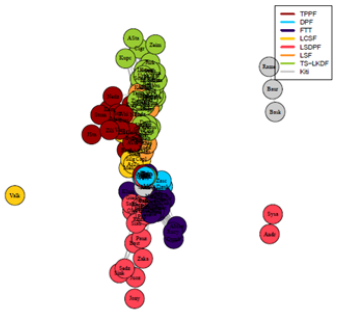


(a) kodavimas: **1,0,0,0,0,0**

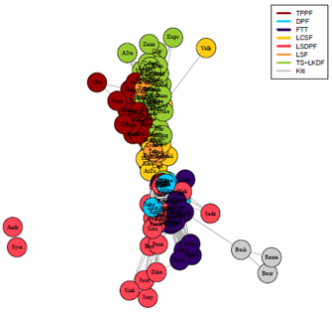


(b) kodavimas: **1,-1,0,0,0,0**

6 pav. Panašumų skaičius – 50 proc.; panašumo matas – Euklido; grafo konstravimo metodas – Fruchtermano ir Reingoldo

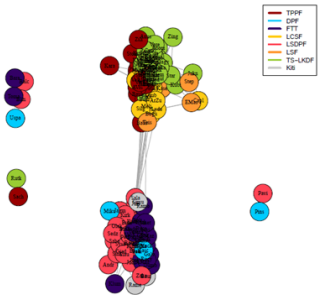


(a) Panašumo matas – Euklido



(b) Panašumo matas – Manheteno

7 pav. Panašumų skaičius – 25 proc.; kodavimas: **1,-1,0,0,0,0**; grafo konstravimo metodas – Fruchtermano ir Reingoldo



8 pav. Panašumų skaičius – 25 proc.; panašumo matas – Hemingo; kodavimas: **2,1,0,-1,-2,-10**; grafo konstravimo metodas – Fruchtermano ir Reingoldo

jant Hemingo atstumą perkoduoti nereikia, taigi nereikia interpretuoti balsavimo elgsenos prasmės. Nors šis metodas pasižymi aiškiausia interpretacija, nereikėtų pamiršti pavojaus, kad jį taikant skirtumams tarp Seimo narių nebalsavimo, nedalyvavimo ar susilaikymo balsuojant bus suteikta per daug įtakos, nors teisės aktų priėmimo procese praktiškai jie turi vienodą reikšmę.

D. Panašumų slenksčio įtakos analizė

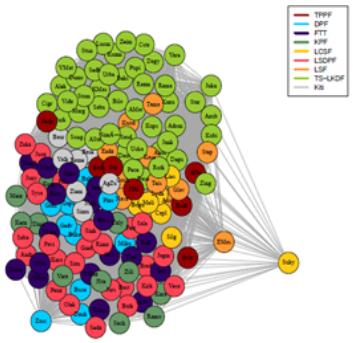
Siekiant įvertinti panašumų skaičiaus slenksčio įtaką buvo braižomas socialinis tinklas, vis mažinant didžiausių panašumų skaičių. Keičiant balsavimo panašumų skaičiaus slenkstį buvo pastebėta, kad balsavimų panašumų skaičius turi įtakos ne tik grafo vaizdavimui, bet ir gaunamiems rezultatams. (žr. 9 pav.).

Grafe, vaizduojančiame didelį panašumų skaičių (pvz., 80 %, žr. 9(a) pav.), t. y. ir santykinai nedidelius panašumus laikant „reikšmingais“, pozicijos ir opozicijos nariai tiesiog išsidėsto skirtingose grafo pusėse. Pavaizdavus grafą su dvigubai mažesniu panašumų slenksčiu (40 %, žr. 9(b) pav.) pozicijos ir opozicijos nariai aiškiai atsiskyrė vieni nuo kitų, tačiau viduje dar liko gana „susimaišę“. Sumažinus į grafo konstravimo procesą įtraukiamų panašumų skaičių iki minimalaus slenksčio (10 %, žr. 9(c) pav.) gaunami politologiškai bene įdomiausi rezultatai: opozicijoje išsiskiria skirtingoms frakcijoms priklausantys Lietuvos Seimo nariai, o pozicija išlieka gana homogeniška.

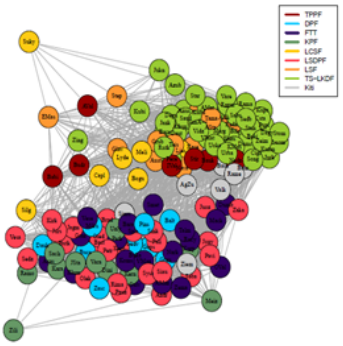
VI. Išvados

Pritaikius socialinių tinklų analizę Lietuvos Seimo narių balsavimams tyrinėti, prieita prie tokių išvadų:

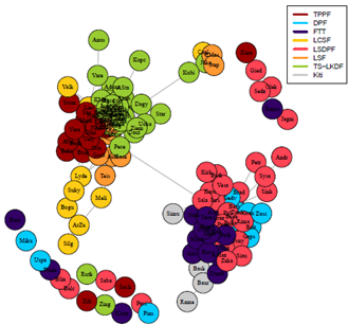
1) Naudojant skirtingus balsavimų profilių panašumo matus ir balsavimų kodavimo schemas, gaunami iš



(a) panašumų skaičius – 80 proc.;



(b) panašumų skaičius – 40 proc.;



(c) panašumų skaičius – 10 proc.;

9 pav. Panašumo matas – Hemingo; grafo konstravimo metodas – Fruchtermano ir Reingoldo

esmės panašūs rezultatai. Balsavimo profilių panašumui apskaičiuoti rekomenduojama naudoti Hemingo atstumą, nes jis pasižymi aiškiausia interpretacija ir šiuo atveju išvengiama subjektyvių sprendimų koduojant duomenis.

2) Mažinant į grafo konstravimo procesą įtraukiamų panašumų skaičių (taigi didinant balsavimo profilių panašumų dydžio slenkstį) grafuose išsiskyrė skirtingoms frakcijoms priklausantys Lietuvos Seimo nariai, kurie pasidalijo į dvi priešingas „stovyklas“ – pozicijos ir opozicijos.
3) Fruchtermano ir Reingoldo metodas labiausiai tinka konstruoti Lietuvos Seimo narių balsavimų panašumų socialiniam tinklui. Šis algoritmas efektyviai pavaizduoja socialinį tinklą, jį paskleisdamas erdvėje ir įgalindamas detalią jo struktūros analizę.

Išanalizavus socialinių tinklų formavimą, kitas etapas Lietuvos Seimo narių balsavimų tyrimų srityje – gautų rezultatų analizė. Vizualizavus duomenis daug informacijos galima gauti iš grafo, tačiau efektyvesnei analizei užtikrinti rekomenduojama įvertinti grafo statistikas. Tinklo charakteristikos užtikrina trijų lygių analizę: individualios viršūnės padėties tinkle, bendro pasiskirstymo tarp viršūnių ir paties tinklo kaip visumos atžvilgiu. Taigi, siekiant gauti naujų įžvalgų parlamentarų balsavimų analizėje tikslinga tęsti tyrimus šioje srityje.

Literatūra

[1] L. F. Anderson, M. W. Watts, A. R. Wilcox, Legislative Roll-Call Analysis. Evanston: Northwestern University Press, 1966.
[2] D. MacRae, Issues and Parties in Legislative Voting. New York: Harper, 1970.
[3] S. Hix, A. Noury, G. Roland, „Dimensions of Politics in the European Parliament“. American Journal of Political Science, vol. 50, no. 2, pp. 494-520, 2006.
[4] K. T. Poole, Spatial Models of Parliamentary Voting. New York: Cambridge University Press, 2005.
[5] J. A. Segal, A. D. Cover, „Ideological Values and the Votes of U.S. Supreme Court Justices“. American Political Science Review, vol. 83, no. 2, pp. 557–565, 1989.
[6] A. D. Martin, K. M. Quinn, „Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999“. Political Analysis, vol. 10, no. 2, pp. 134–153, 2002.
[7] J. Clinton, S. Jackman, D. Rivers, „The Statistical Analysis of Roll Call Data“. American Political Science Review, vol. 98, no. 2, pp. 355-370, 2004.
[8] M. A. Bailey, „Comparable Preference Estimates across Time and Institutions for the Court, Congress, and Presidency“. American Journal of Political Science, vol. 51, no. 3, pp. 433–448; 2007.
[9] E. Voeten, Data and Analyses of Voting in the UN General Assembly (July 17, 2012). Prieiga per internetą: http://ssrn.com/abstract=2111149.
[10] V. Morkevičius, „Neideologinis Seimas? Statistinė svarbių 2004–2008 m. kadencijos Lietuvos Seimo balsavimų analizė. Partinės demokratijos pabaiga? Politinis atstovavimas ir ideologijos“. A. Ramonaitė (sud.), pp. 53–87. Vilnius: Versus aureus. 2009.
[11] T. Krilavičius, V. Morkevičius, „Lietuvos parlamentarų ideologinės pozicijos ir jų raida 2008–2012 m. kadencijos Seime: statistinė Seimo narių balsavimų analizė“. Review of the Political Ideas in Lithuania, Europe and the World: Political Reality and Tendencies of It's Analysis. LPA, VU TSPMI 2010, Vilnius, November 19, 2010.
[12] T. Krilavičius, V. Morkevičius, „Mining Social Science Data: A Study of Voting of the Members of the Seimas of Lithuania by Using Multidimensional Scaling and Homegeneity Analysis“. Intellectual Economics, vol. 5, no. 2 (10), pp. 224–243, Vilnius, 2011.
[13] T. Krilavičius, A. Žilinskas, „On structural analysis of parliamentarian voting data. Informatica“, vol. 19, no. 3, 2008.
[14] T. Krilavičius, V. Morkevičius, „Kaip analizuoti Seimo narių balsavimus? Metodologinės pastabos“. Pranešimas Lietuvos soc. mokslų forume 2009.
[15] L. R. Izquierdo, R. A. Hanneman, „Introduction to the formal analysis of social networks using mathematica“, Version 2, p. 3, 2006.

INFORMACINĖS TECHNOLOGIJOS ● IVUS 2013 ● ISSN 2029-249X ● eISSN 2029-4824

Nuolaidų ir lojalumo kortelių perkėlimo į išmaniuosius įrenginius tyrimas

Vilius Vaičiūnas
Vilniaus universitetas,
Kauno humanitarinis fakultetas, Informatikos katedra
Kaunas, Lietuva
vilius.vaiciunas@khf.stud.vu.lt

Santrauka. Straipsnyje siekiama supažindinti su šiandieninėmis lojalumo / nuolaidų kortelių sistemomis, identifikuoti plastikinių kortelių keliamus nepatogumus. Taip pat supažindinama su galimomis plastikinių kortelių alternatyvomis išmaniuosiuose telefonuose. Pirmoje straipsnio dalyje apžvelgiamos esamos šios srities problemos, pateikiami pavyzdiniai integracijos į mobiliuosius įrenginius būdai. Toliau lyginami įgyvendinti programiniai sprendimai, sudaromas jų bendrinis modelis, pateikiamas patobulintas nuolaidų kortelių integracijos į mobiliuosius įrenginius modelio prototipas.

Reikšminiai žodžiai: nuolaidos, nuolaidų kortelės, mobiliosios aplikacijos, išmanieji telefonai, išmanieji įrenginiai.

I. ĮVADAS

Lojalumo ir nuolaidų kortelės – dažniausiai plastikinės kreditinės kortelės dydžio vienetas, pirkėjui suteikiantis privilegijų perkant prekių ar paslaugų. Daugelis mūsų jų turi ne vieną ir ne dvi, kadangi skirtingos prekybos vietos turi savo atskiras lojalumo ir nuolaidų programas. Lojalumo sistema prekybininkui padeda išskirti individualų klientą ir taip suteikia galimybę jį „prisirišti“ bei palaikyti ryšį dažniausiai tiesioginėmis priemonėmis – el. paštu, telefonu, laiškais.

Ši abipusiška nauda žvelgiant iš pirkėjo perspektyvos sukelia keletą problemų. Pirmoji – nešiotis daug nuolaidų kortelių nepatogu, kadangi jos užima nemažai vietos, o atsirinkti konkrečią kortelę iš daugelio neretai užtrunka daug laiko ir yra nepatogu, nes dažnai reikia atsiskaityti greitai. Antroji problema – kortelės pametamos, dažnai jų pasigendama konkrečiose situacijose, kadangi gali būti pamirštos pasiimti su savimi, paskolintos ar susidėvėjusios. Nešiojantis visas turimas korteles, susiduriama su piniginės svorio / dydžio problema, o neturint kortelės su savimi prekę ar paslaugą tenka pirkti be nuolaidos.

Vokietijos nepriklausomų tyrimų kompanijų duomenimis [1], šioje šalyje įvairių kuponų ir nuolaidų kortelių skaičius viršijo 10 mlrd. Taip pat prognozuojamas įvairių nuolaidų sistemų pritaikymas naudojant mobiliuosius įrenginius. Sparčiai auganti išmaniųjų telefonų rinka atveria ir mobiliųjų aplikacijų panaudojimo galimybes, todėl naujas būdas turėti visas nuolaidų ir lojalumo korteles su savimi ir nejausti dėl to jokio diskomforto – turėti jas išmaniajame telefone.

Norint atskleisti šiandienines lojalumo / nuolaidų kortelių programas ir jų perkėlimo į mobiliuosius įrenginius mechanizmus, buvo apžvelgtos kelių populiarių Lietuvos prekybos vietų siūlomos lojalumo programos, sudarytas bendras jų veikimo modelis. Taip pat buvo išanalizuoti jau įgyvendinti programiniai sprendimai – kitaip tariant, jau veikiančios alternatyvos plastikinėms nuolaidų kortelėms.

Apžvelgus galimus programinio įgyvendinimo būdus, buvo išskirta keletas pagrindinių būdų. Kadangi naujasis modelis turėjo būti naudingas tiek prekybininkui, tiek pirkėjui, atsisakyta didesnių investicijų prekybininkams – naujų terminalų, skaitytuvų keitimo ar kt. techninių sprendimų ir ieškota naudos pirkėjui.

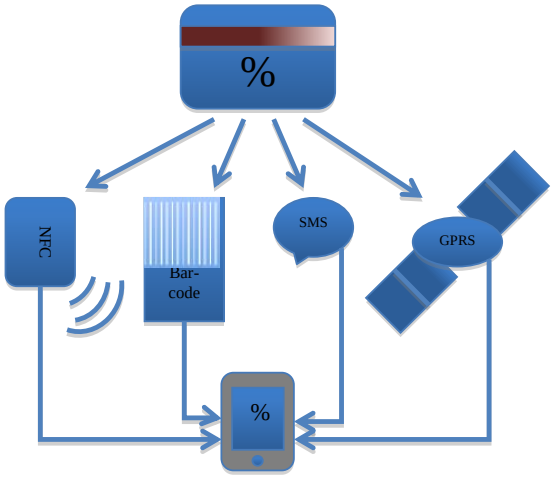
Naujojo nuolaidų ir lojalumo kortelių išmaniajame įrenginyje modelio pasiūlymas – apklausų, eksperimentų, esamų programinių sprendimų bei mokslinės literatūros, susijusios su tiriama sritimi, analizės rezultatas.

II. LOJALUMO / NUOLAIDŲ PROGRAMŲ INTEGRACIJOS Į MOBILŲJĮ ĮRENGINĮ GALIMYBIŲ ANALIZĖ

A. Preliminarūs problemos sprendimo būdai

Kaip minėta anksčiau, vienas iš galimų įvade minėtų problemų sprendimo būdų – lojalumo ir nuolaidų programų integravimas į mobilųjį įrenginį. Sparčiai auganti išmaniųjų telefonų plėtra didina ir mobiliųjų aplikacijų panaudojimo galimybes, todėl šioje terpėje nuolaidos gali būti saugojamos bei lojalumo programa – veikti taikant įvairius metodus – NFC (angl. *near field communication*), barkodus, SMS ar EDGE/3G būdais. Populiarėjantys išmanieji telefonai kelia naujus iššūkius integruojant įvairias gyvenimo sritis.

Atliktas tyrimas parodė, kad populiariausi yra 4 būdai (1 pav.) perteikti nuolaidų kortelių ir lojalumo programų sistemas mobiliaisiais įrenginiais: tai – NFC (angl. *near field communication*), barkodai, SMS ar EDGE/3G.



1 pav. Galimi nuolaidų kortelių perkėlimo į mobilųjį įrenginį būdai

B. Trumpa pateiktų būdų apžvalga

NFC. Pasitelkus identifikacijos radijo bangomis (RFDI) technologiją ir nekontaktinės daugiafunkcės kortelės infrastruktūrą, NFC technologija veikia mažu, dažniausiai kelių centimetrų, spinduliu [2]. Naudojant NFC technologiją mobilusis telefonas gali nuskaityti nedidelį informacijos kiekį nuo RFDI paslaugų žymelių ir palaikyti ryšį su kitais prietaisais. Todėl priliečiant mobilųjį telefoną prie RFDI žymelės, aktyvinama toje žymelėje esanti paslaugos nuoroda [3].

Barkodas. Kadangi dauguma nuolaidų kortelių sistemų turi įdiegtą brūkšninį kodą, lengviausias būdas perkelti informaciją į telefoną – pasinaudojus integruota kamera nuskaityti kortelės kodą. Tuomet jį juodai perpiešus baltame fone, atsiranda galimybė lengvai nuskenuoti su skeneriu prekybos vietose. Atrodytų, jog tai – nesudėtingas metodas perkelti nuolaidų kortelę į išmanųjį telefoną, tačiau susiduriama su problemomis bandant nuskaityti brūkšninį kodą tiesiai iš ekrano. Barkodas iš telefono ekrano lengvai nuskaitomas tik su senesnio tipo „Infrared“ ir rečiau paplitusiais „Conventional“ tipo vaizdų skaitytuvais, tačiau iškyla sunkumų bandant nuskaityti kodą populiariausiais lazeriniais skeneriais, kurie prasiskverbia į gilesnį sluoksnį nei pats ekranas, kuriame ir pavaizduojamas barkodas.

Šis būdas, mano manymu, yra vienas patraukliausių, kadangi prekybininkui nereiktų diegti naujų technologinių sprendimų prekybos vietoje, o barkodų pavaizdavimas ir nuskenavimas – nesudėtingos procedūros.

SMS. SMS (angl. *short message service*) – vienas primityviausių būdų nuolaidų pasiūlymams platinti. SMS tekstu gavėjas gali gauti specialių pasiūlymų, kuriuos gali tiesiog parodyti prekybos vietoje. Pavyzdys – prekybos centrų tinklo „IKI“ trumposiomis žinutėmis platinami pasiūlymai, kuriuose įterpiamas slaptažodis, kurį reikia pasakyti atsiskaitant už prekes.

Sistemos trūkumai:

- nemažai kainuoja – paslaugą teikia ryšio operatoriai; toliau pateikiame vienos SMS marketingo Lietuvoje įmonės įkainius verslo klientams siunčiant SMS į įvairias šalis (2 lentelė);
- galima lengvai platinti nuolaidos žinutes – užtenka persiųsti į kitą mobilųjį telefoną ir pakeisti SMS siuntėjo vardą į, šiuo atveju, „IKI“.

Sistemos privalumai [4]:

- nesudėtingas procesas;
- galimybė labai tiksliai segmentuoti reklamos gavėjus pagal tam tikrus kriterijus (amžių, lytį, miestą ir pan.);
- SMS žinutės yra siunčiamos tik sutinkantiems gauti reklamą abonentams. Šiuo metu sutikimą davusių abonentų Lietuvoje yra per 300 000;
- galimybė naudotis SMS reklamos paslaugomis įvairiose šalyse;
- tikėtinas „voratinklio“ plitimo principas, kai lengvu būdu išplatinamos nuolaidos; pavyzdžiui, nuolaida įėjimo į naktinį klubą mokesčiui gaunama SMS žinute, greitu būdu išplinta, todėl didesnė tikimybė, kad naktinis klubas sulauks daugiau lankytojų.

QR/EDGE/3G. Nuolaidų sistemų perkėlimo į mobilųjį telefoną naudojant internetą sistemų pavyzdžiu galime laikyti QR kodų naudojimą, kuriuos nuskenavus specialia aplikacija, gaunamas nuolaidos kodas, kuris išsaugomas telefone. Nuskaičius QR kodą, dažniausiai aptinkama nuoroda į tam tikrą internetinį portalą. Nuorodai panaudoti reikalingas tam tikras interneto ryšys – dažniausiai EDGE, 3G arba Wi-fi. Jei pateikiamas tik tam tikras kodas, jį galima pateikti prekybos vietoje kaip nuolaidos kuponą ar kortelę [5].

III. ESAMŲ PLASTIKINIŲ NUOLAIDŲ IR LOJALUMO KORTELIŲ SISTEMŲ ĮVERTINIMAS

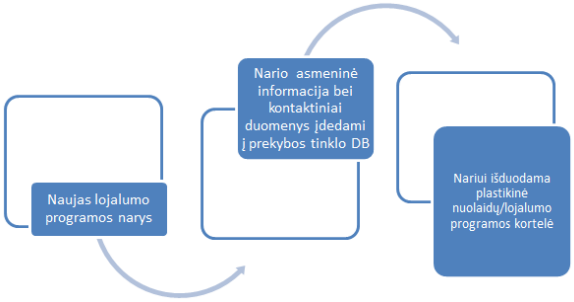
Atliekant kiekybinį tyrimą, buvo siekiama išsiaiškinti, kokią informaciją renka prekybos vietos Lietuvoje, suteikdamos klientams teisę naudotis nuolaidomis ar dalyvauti lojalumo programoje. Eksperimentui, kurio tiksline grupe galima įvardyti Lietuvos prekybininkus, atlikti buvo pasirinktos trijų skirtingų prekybos centrų – „Drogas“, „Senukai“ ir „Rimi“ – lojalumo programos.

Atlikus eksperimento metu gautų rezultatų analizę, galima teigti, kad trijų atsitiktinai pasirinktų didelių prekybos centrų bei prekybos vietų vykdomos lojalumo programos schemos praktiškai sutampa. Visi trys centrai turi savo atskiras lojalumo programų bei nuolaidų sistemas, sujungtas į vieną kortelę, skirtą vienam prekybos tinklui. Naujo lojalumo programos nario asmeninės informacijos bei kontaktinių duomenų kategorijos taip pat sutampa: reikalaujama vardo, pavardės, gimimo datos, „Senukai“ papildomai prašo pažymėti labiausiai dominančių prekių kategorijas, o „Rimi“

nuolaidų kortelės anketoje privalomas laukas yra „šeiminė padėtis“, – tai nėra privaloma kitose dviejose minėtose anketose.

Kortelės kaina sutampa „Senukuose“ ir „Droge“ – 4,99 Lt. Naujos „Rimi“ kortelės gamyba šiek tiek pigesnė – 4,44 Lt.

Nepaisant šių nedidelių anketos privalomų laukų skirtumų bei nežymaus naujos kortelės gamybos kainos skirtumo viename iš prekybos tinklų, visų jų lojalumo programų naujų narių įtraukimą iliustruoja bendros schemos principas (2 pav.):



2 pav. Prekybos tinklų lojalumo programų principinė schema

IV. Šiuo metu įgyvendinti programiniai sprendimai

Tyrimo pradžioje nebuvo nė vieno Lietuvoje įgyvendinto programinės įrangos sprendimo, leidžiančio išspręsti mūsų tiriamą problemą. Šiuo metu jau egzistuoja „mobiliosios piniginės“, kurių veikimo principas apima ir mūsų minimų kortelių integraciją.

Pasauliui rudenį pristatytas daug skirtingų atgarsių susilaukęs „Apple“ išmanusis telefonas „iPhone 5“ kartu pateikė ir programinės įrangos iOS 6 naujovių, tarp kurių – mobilioji piniginė „PassBook“, apimanti ne tik nuolaidų, lojalumo korteles bei kuponus, tačiau ir lėktuvo bilietus, kreditus, kuriais galima atsiskaityti prekybos vietose, renginio bilietus ir pan. „Passbook“ aplikacija ekrane vaizduoja trijų tipų 2D barkodus: „Aztec“, PDF417 ir QR. Kiekvienas skaitmeninis kuponas ar bilietas programinėje įrangoje apibrėžiamas kaip „pass“ (3 pav.). Vartotojui pirmąkart įjungus „Passbook“ aplikaciją, atidaromas įžanginis langas, kuriame taip pat yra aplikacijų krautuvės mygtukas, siūlantis naršyti bei įsigyti „Passbook“ integraciją turinčios programinės įrangos. „Passes“ taip pat gali būti platinami „Safari“ naršykle arba siunčiami vartotojams el. paštu. „Passes“ yra sinchronizuojami tarp iOS platformos pagrindu veikiančių išmaniųjų įrenginių, naudojant „debesų kompiuterijos“ sistemą „iCloud“. Nors „Passbook“ aplikacija atsiradusi kaip standartinė programinė įranga su iOS 6 versija, ją palaiko tik „iPhone“ ir „iPod Touch“ įrenginiai, planšetiniai kompiuteriai „iPad“ nepalaiko.

3 pav. „Apple“ „Passbook“ pagrindinės savybės: a) visų „passes“ išrikiavimas; b) pasirinkto bilieto / kupono informacija bei QR kodas; c) telefoną suaktyvinantys pranešimai

Pagrindinės „Passbook“ ypatybės:

- pavaizduoja trijų tipų 2D barkodus: „Aztec“, PDF417 ir QR;
- suaktyvina mobilųjį telefoną, priklausomai nuo lokacijos;
- suaktyvina mobilųjį telefoną, priklausomai nuo bilieto / kupono galiojimo (aktyvumo) laiko;
- vienas bilietas / kuponas gali saugoti iki 35 kalbų vertimų;
- „passai“ yra aktyvūs – „Apple“ „Push Notification“ servisas automatiškai praneša apie kupono / bilieto pasikeitimus.

Deja, ši sistema kol kas veikia tik Jungtinėse Amerikos Valstijose.

Lietuvoje pamažu prisitaiko „Cloudi“ programinė įranga, skirta tiek „Android“, tiek iOS platformoms bei leidžianti tiesiog perkelti plastikinę kortelę į telefoną, taip sugeneruojant barkodą, kurį vartotojas turi pateikti prekybos vietoje.

UAB „Claudis“[1] Lietuvoje pradėjo diegti naują išmaniųjų telefonų programėlę „Cloudi“, kuri, pasak kūrėjų, turėtų palengvinti naudojimąsi nuolaidų kortelėmis. Programėlėje vaizdo kamera leidžia išsaugoti ant nuolaidų kortelių esančius brūkšninius kodus, kuriuos klientas naudoja kaip įprastas korteles. Pasak kūrėjų, įmonė planuoja toliau tobulinti programėlę, perkelti nuolaidų kuponus į išmaniuosius telefonus, pasitelkiant programėlę organizuoti apklausas ir gauti vartotojų atsiliepimus apie parduotuvių pasiūlymus. Įmonė taip pat siūlys visiškai naujas lojalumo programas, kurioms nereikės plastikinių nuolaidų kortelių.

„Apple“ mobiliųjų aplikacijų talpykloje per „App Store“ paiešką radus „Cloudi“ aplikaciją, iškart matoma 19 (2013-01-10) vartotojų atsiliepimų bei 1,5 iš 5 žvaigždučių jų įvertinimas. Žemas įvertinimo balas bei neigiami atsiliepimai

[1] Pranešimas spaudai – 2012-10-25. Prieiga internete: [http://www.technologijos.lt/n/pranesimai_spaudai/S-29085/straipsnis/Lietuviai-siulo-perkelti-nuolaidu-korteles-i-ismaniuosius-telefonus

leidžia spręsti, jog aplikacija turi kol kas neišspręstų problemų: vartotojai skundžiasi nuolaidų kortelių įdėjimo į programėlę problemomis, barkodo nuskaitymo vaizdo kamera sunkumais, neintuityvia grafine sąsaja bei kitomis iškylančiomis problemomis.

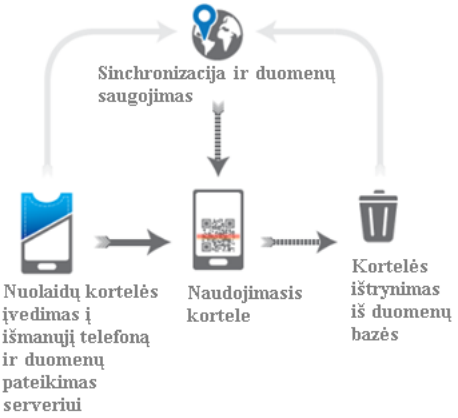
Įsidiegus aplikaciją, norintis ja naudotis vartotojas turi užsi registruoti. Taigi yra esminis skirtumas tarp egzistuojančio produkto ir mūsų siūlomo modelio – registracijai pakanka tik el. pašto bei sugalvoto slaptažodžio – nerenkami jokie kontaktiniai duomenys ar asmeninė informacija apie vartotoją.

Sėkmingai užsiregistravus (jokių patvirtinimų el. paštu nereikalaujama, – taip pat viena iš spragų, kadangi praktiškai galima užsiregistruoti su bet kokio asmens elektroninio pašto adresu), pereinama prie pagrindinio aplikacijos lango. Jame yra 5 (penki) pagrindiniai meniu punktai, išdėstyti ekrano apačioje – „Šalia manęs“, „Kategorijos“, „Kortelės“, „Pasiūlymai“ ir „Nustatymai“. Grafiškai viskas atrodo gana gražiai ir iš pirmo žvilgsnio atitinka šiuolaikinių aplikacijų dizaino reikalavimus – pastelinės spalvos, gražūs logotipai ir fono paveikslėliai. Tačiau intuityvumo, kaip ir teigė anksčiau minėti komentatoriai, išties trūksta. Nėra aišku, nei ką atlieka parinktis „Šalia manęs“ (kadangi rodomas tuščias langas pirmo paleidimo metu), nei kam skirtos visos kategorijos, esančios meniu punkte „Kategorijos“, kadangi čia randame tiesiog išvardytas sugrupuotas prekybos vietas ir jokių kortelės pridėjimo nuorodų. Apskritai neaptikta nuorodų, patarimų (angl. *tooltips*), todėl pirmąkart įsijungus programinę įrangą nėra aišku, ką ir kokios funkcijos atlieka.

Likę du meniu punktai – „Pasiūlymai“ ir „Nustatymai“, kol kas taip pat nėra labai funkcionalūs. „Nustatymuose“ galima tik pasikeisti prisijungimo slaptažodį, įvesti telefono numerį arba atsijungti. Meniu punkte „Pasiūlymai“ pateikiamos vykstančios akcijos tam tikrose prekybos vietose Lietuvoje. Eksperimento metu buvo galima paskaityti apie vienintelę vykstančią akciją vienoje iš juvelyrinių parduotuvių bei jos interneto adresą, kurį patariama aplankyti, norint sužinoti daugiau informacijos.

4 pav. „Cloudi“ aplikacijos pradinio naudojimosi langai

Pati programinė įranga turi nemažai loginių spragų. Nuolaidų kortelių, kurias galima įdėti į programėlę, duomenų bazė kol kas nėra didelė. Tačiau, jei projektas turi potencialą, sparčiai vystomas bei atsižvelgiama į vartotojų pastabas, šie trūkumai gana greitai ištaisomi. Populiarumo augimo sparta taip pat priklauso nuo prekybininkų geranoriško bendradarbiavimo, palaikymo ir investicijų į naujų technologinių sprendimų plėtrą. Plastikinių kortelių integravimo į išmaniuosius įrenginius programinės įrangos „Cloudi“ veikimo schema pateikiama 5 pav.:



5 pav. „Cloudi“ aplikacijos veikimo schema

V. SIŪLOMO NUOLAIDŲ BEI LOJALUMO PROGRAMŲ KORTELIŲ PERKĖLIMO Į IŠMANIUOSIUS ĮRENGINIUS PRINCIPINĖ SCHEMA

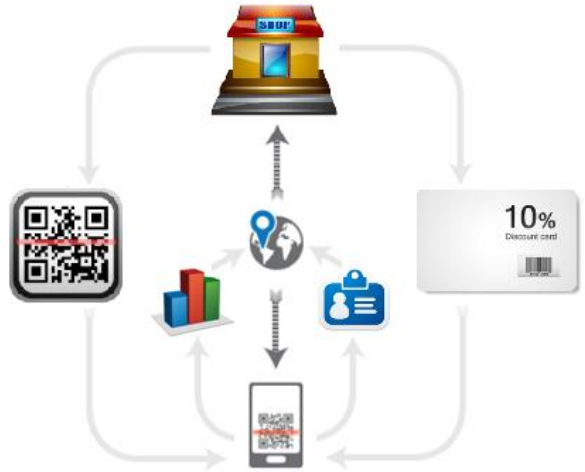
Išanalizavus eksperimentų analizės preliminarias išvadas, 6 pav. pateikiamas patobulintas modelis, sujungiantis lojalumo, nuolaidų bei pirkėjų atsiliepimų analizės sistemą. Jis taip pat aprėpia esamus programinius sprendimus bei išsprendžia plastikinių kortelių gamybos, popierinių anketų rinkimo, sudėtingo duomenų įvedimo bei analizės problemas.

Naujo modelio veikimą iliustruojantis pavyzdys: tarkime, egzistuoja parduotuvė ABC, vartotojas (pirkėjas) X ir parduotuvėje dirbantis pardavėjas Y. Vartotojas X, atsisiuntęs mūsų modeliu grįstą mobiliąją aplikaciją, įsidiegia ją savo išmaniajame telefone. Pirmąkart įjungęs programėlę, vartotojas privalo pateikti asmeninę ir kontaktinę savo informaciją – ji gali būti tokia pati, kaip ir anksčiau analizuotose prekybos centrų anketose – t. y. vardas, pavardė, amžius, lytis ir t. t., kuri yra įdedama į naujosios aplikacijos serverius. Vartotojui X tai padaryti užtenka vieną vienintelį kartą. Jam užsukus į parduotuvę ABC, galimi trys variantai: X dar neturi ABC nuolaidų kortelės; X turi plastikinę nuolaidų kortelę; X turi nuolaidų kortelę naujojoje aplikacijoje. O šia sistema nusprendę naudotis prekybininkai savo informaciją taip pat įdeda į serverį, jiems sugeneruojamas QR kodas, kurį prisiklijuoja, tarkime, ant prekystalio.

- Pirmasis variantas: X dar neturi ABC nuolaidų kortelės. Atėjęs į parduotuvę ABC ir nuskaitęs QR kodą, iš serverio gauna būtent ABC nuolaidų kortelę į išmanųjį įrenginį (jei, tarkime, šį asmenį patvirtina prie sistemos prisijungęs pardavėjas Y).
- Antrasis variantas: X turi nuolaidą plastikinėje kortelėje. X gali elgtis pagal pirmąjį variantą arba įkelti kortelę į aplikaciją pagal „Cloudi“ veikimo schemą.

- Trečiasis variantas: X ateina į parduotuvę ABC ir nori už prekes atsiskaityti su nuolaida, kurią jau turi savo išmaniajame telefone. Tuomet prototipas gali veikti dvejopai: pasirinkęs turimą kortelę aplikacijoje, vartotojas gali atsisiųsti barkodą, kurį pateiktų nuskenuoti pardavėjui Y arba, jei nuskaitymas nesėkmingas, pažymėtų 2 parinktį – PIN kodą, galiojantį, tarkime, dvi minutes ir susidedantį iš 3 skaičių. Atsisiunčiant PIN pardavėjas Y iš serverio gautų tokio paties PIN kodo reikšmę, – tai reikštų, jog vartotojas X tikrai turi teisę į nuolaidą parduotuvėje ABC.

Toks sistemos principas būtų naudingas tiek pirkėjams, tiek prekybininkams. Pirkėjai išvengtų diskomforto nešiotis daugybę plastikinių kortelių, visuomet turėtų nuolaidą savo telefone, nereikėtų kaskart pildyti naują anketą naujoje prekybos vietoje, o prekybininkas išvengtų didelių lojalumo sistemų organizavimo, plastikinių kortelių gamybos, anketų teksto atpažinimo ir sudėtingo pirkėjų duomenų analizavimo. Siūlomas patobulintas sistemos modelis pavaizduotas 6 pav.



6 pav. Siūlomo mobiliosios nuolaidų ir lojalumo programų, integruotų į mobiliąją aplikaciją, sistemos veikimo schema

VI. IŠVADOS

1. Apžvelgus esamą padėtį, akivaizdu, jog už plastikinėmis kortelėmis teikiamų nuolaidų ir lojalumo programų slypi gana komplikuotos, nemažų sąnaudų reikalaujančios sistemos. Pradinis jų etapas – informacijos apie pirkėją kaupimas popierinėmis ar internetinėmis priemonėmis. Tolesni etapai – šių duomenų išsaugojimas, plastikinės kortelės pardavimas, jos susiejimas su naujuoju lojalumo programos dalyviu.
2. Atliekant analizę ištirta, kad šiuo metu įgyvendinti plastikinių kortelių perkėlimo į mobiliuosius (išmaniuosius) telefonus sprendimai turi nemažai trūkumų ir pirkėjams, ir pardavėjams ir kol kas nėra abipusiškai naudingi.
3. Pasiūlytas patobulintas nuolaidų ir lojalumo sistemų kortelių perkėlimo į mobiliuosius įrenginius modelis, naudingas tiek pirkėjams, tiek prekybininkams.

VII. LITERATŪRA

[1] C. Jokisch, Smartphone Instead of Discount Card. [Interaktyvus]. Ezine articles. [Žiūrėta 2012 balandžio 14 d.]. Prieiga per internetą: http://www.eurocis.com/cgi-bin/md_eurocis/custom/pub/content.cgi?lang=1&oid=4474&ticket=g_u_e_s_t&ca_page=de%2Fpress-releases_634.htm

[2] K. Kurran, Near Field Communication. [Interaktyvus]. Aberdeen Group 2008. [Žiūrėta 2012 gegužės 12 d.]. Prieiga per internetą: http://www.iaesjournal.com/online/index.php/IJECE/article/view/234.

[3] J. P. Mangalindan, NFC Coming to the Nook, 2011. [Interaktyvus]. i-Trend-Madison. [Žiūrėta 2012 balandžio 13 d.]. Prieiga per internetą: http://tech.fortune.cnn.com/2012/05/01/nook/.

[4] J. Hart, M. Hannan, The Future of Mobile Technology and Mobile Wireless Computing. Campus-Winde Information Systems, 2004, vol. 21, Issue: 5, pp.1017–1040. DOI:10.1108/03030560310477645.

[5] D. Blumberg, The Next Generation of Business CARDS. [Interaktyvus]. Investopedia. [Žiūrėta 2012 gegužės 12 d.]. Prieiga per internetą: http://articles.businessinsider.com/2011-03-21/tech/30015797_1_qr-codes-paper-card-business-cards.

Roboto valdymo sistemos neuroninės kompiuterio sąsajos tyrimas

Mindaugas Vasiljevas[1], Kęstutis Šidlauskas[2]
Kauno technologijos universitetas
Programų inžinerijos katedra
Kaunas, Lietuva
vasiljevasm@gmail.com[1], kestusissid@gmail.com[2]

Santrauka. Šiame straipsnyje nagrinėjama neuroninė kompiuterio sąsaja, skirta valdyti vežimėlio tipo robotui. Pateikiamas paprastas ir nesudėtingai taikomas metodas – valdymas per paviršinės elektromiogramos signalo lygį. Taip pat šiame straipsnyje trumpai aprašoma mūsų sukurta signalo lygio valdymo sistema bei pateikiami roboto valdymo eksperimento, atlikto su šia sistema, rezultatai.

Reikšminiai žodžiai. žmogaus ir mašinos sąsaja, smegenų ir kompiuterio sąsaja, neuroninė kompiuterio sąsaja, robotų valdymas.

I. ĮVADAS

Šiuo metu mokslo pasaulyje vis aktyviau tyrinėjami alternatyvūs kompiuterio valdymo būdai. Nagrinėjami įvairūs žmogaus ir mašinos sąsajos metodai, tokie kaip akies vyzdžio sekimo sistemos, gestų valdymo sąsajos, valdymas balsu, smegenų ir kompiuterio sąsaja bei neuroelektroninės sąsajos (valdymas raumenų judesiais). Šie metodai taikomi ne tik paprasto kompiuterio valdymui, bet ir kitų skaitmeninių įrenginių valdymui (pvz., egzoskeletų, invalidų vežimėlių, galūnių protezų, įvairių robotų ir pan.). Smegenų ir kompiuterio sąsajoje kaip valdymo signalai naudojamos smegenų bangos – smegenų elektroencefalograma. Savo ruožtu ši sąsaja gali būti apibrėžta kaip komunikavimo sistema, kuri smegenų generuojamas bangas verčia į kompiuterio ar kito skaitmeninio įrenginio komandas [1]. O neuroelektroninėse sąsajose dažniausiai naudojami paviršinių raumenų generuojami elektriniai signalai – paviršinė elektromiograma.

Šiame darbe tyrinėjame žmogaus ir mašinos sąsają, kuri yra valdoma per paviršinės galvos raumenų elektromiogramą. Šiais signalais valdomas „Arduino“ 4WD robotas [2]. Tai – vežimėlio tipo robotas, turintis keturis ratus ir valdomas bevieliu ryšiu. Elektromiogramos signalams nuskaityti naudojamas OCZ NIA EEG ir EMG skaitytuvas [3]. Tokia sistema gali būti laikoma sumažinta išmaniojo invalido vežimėlio [4, 5, 6] versija. Panašias sistemas galima plėtoti ir pasitelkiant žmogaus okulogramą (akių judesių generuojamus signalus) [7]. Taip pat išmaniajam vežimėliui valdyti gali būti taikoma smegenų kompiuterio sąsaja [8].

Pagrindiniai išmaniojo invalido vežimėlio naudotojai yra žmonės, turintys rimtų motorikos sutrikimų, negalintys valdyti ne tik kojų, bet ir rankų. Panašias sistemas galima taikyti ir pramonėje, kur reikia „trečiosios rankos“. Taip pat šie metodai gali būti taikomi protezų valdymui ir kompiuteriniams žaidimams.

Panašiose robotų valdymo sistemose [9] taikomi įvairūs duomenų gavybos ir klasifikavimo metodai. Bene labiausiai paplitę yra įvairių tipų neuroniniai tinklai [10]. Taip pat dažnai naudojamas atraminių vektorių metodas (SVM) [11], pagrindinių komponentų analizė, nepriklausomų komponentų analizė ir „Random Forest“ algoritmas [12].

Mūsų sukurtoje sistemoje naudojamas nesudėtingas ir patogus klasifikavimo būdas – komandų klasifikavimas pagal signalo lygį. Šis metodas patrauklus tuo, kad nereikalauja daug išteklių ir gali būti taikomas nedidelės atminties mikroprocesorių platformose (pvz., „Arduino“), taip pat reikalui esant – programuojamos logikos įrenginiuose ir integrinėse schemose.

Antrame šio straipsnio skyriuje aprašomas valdymo naudojant signalo lygį metodas. Pateikiama mūsų sukurtos sistemos principinė schema su paaiškinimais. Trečiame šio straipsnio skyriuje aptariamas eksperimentas, atliktas su mūsų sukurta sistema, bei jo rezultatai. Ketvirtame skyriuje pateikiamos išvados, apžvelgiamos eksperimento metu iškilusios problemos ir nusakomi ateityje numatomi tyrimai.

II. VALDYMAS NAUDOJANT SIGNALO LYGĮ

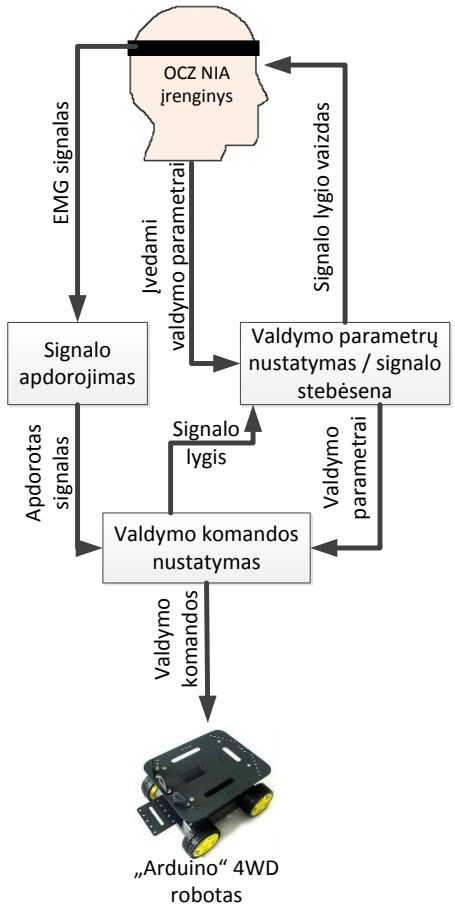
Valdymas naudojant paviršinės elektromiogramos signalo lygį yra nesudėtingas, kadangi operuojama tik su momentine signalo lygio reikšme. Signalo pokytis laike nėra svarbus. Pagrindinė efektyvaus valdymo šiuo principu sąlyga yra pakankamai aukšta signalo amplitudė. Signalo nuskaitymo būdas turėtų leisti generuoti pakankamai aukštas amplitudes, tuomet signalo amplitudę galima suskirstyti į intervalus, žyminčius atitinkamas valdymo komandas.

Savo sistemoje taikome metodą, pagal kurį išskiriamos keturios valdymo komandos: „judėti pirmyn“, „judėti kairėn“, „judėti dešinėn“ ir „stovėti“. Kuo didesnis valdymo komandų skaičius, tuo mažesnis sistemos valdymo tikslumas. Siekdami didesnio valdymo tikslumo, atsisakėme komandos „judėti atgal“. Norint su robotu judėti atgal, galima apsisukti ir judėti pirmyn.

Signalo amplitudė skaidoma į penkis intervalus (žr. 3 pav.). Intervalų ilgiai yra individualūs kiekvienam žmogui, todėl sistema leidžia reguliuoti signalo amplitudės intervalų ilgius. Intervalų ilgius gali tekti pakeisti net ir tam pačiam vartotojui skirtingais laiko momentais, nes valdymas remiasi veido raumenų elektromiogramos nuskaitymu. Šiuos signalus gali labai paveikti nuovargis.

Sistemoje naudojamas OCZ NIA [3] EEG ir EMG signalų skaitytuvas, su kompiuteriu komunikuojantis per USB sąsają. Šis skaitytuvas yra tvirtinamas prie vartotojo kaktos ir yra labai jautrus akių mirksniams bei kaktos raumenų įtempimui. Vartotojo inicijuotas komandas vykdo „Arduino“ 4WD įrenginys [2]. Tai – vežimėlio tipo robotas, su kompiuteriu komunikuojantis bevieliu ryšiu.

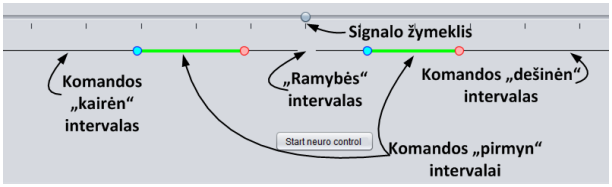
Sistemos principinė schema pateikta 1 pav. Kaip joje matyti, skiriami trys pagrindiniai etapai: signalo apdorojimas, valdymo komandos nustatymas ir valdymo parametrų nustatymas bei signalo stebėsena.



1 pav. Sistemos principinė schema

Apdorojant signalą jo reikšmių skaičius yra sumažinamas. Taip pat signalo reikšmė kiekvieno laiko momento metu yra suvidurkinama su prieš tai buvusia signalo reikšme.

Valdymo parametrų nustatymas ir signalo stebėsena yra sujungti, kad būtų patogiau valdyti sistemą. Šį etapą iliustruoja sistemos valdymo sąsaja, pateikta 2 pav. Signalo žymeklis nurodo momentinę signalo reikšmę. Kai šis žymeklis yra vidurinėje padėtyje, tai signalo reikšmė lygi nuliui. Į kairę nuo nulinės vertės yra neigiamos signalo reikšmės, o dešinėje – teigiamos reikšmės. „Ramybės“ intervalas žymi intervalą, kai jokie veiksmai neatliekami. Kai signalo žymeklis atsiduria šiame intervale, robotas stovi vietoje. Kai žymeklis atsiduria komandos „kairėn“ intervale, robotas juda kairėn. Žymekliui atsidūrus komandos „dešinėn“ intervale, robotas juda dešinėn. Judėjimui į priekį yra skirti 2 intervalai (žr. 2 pav. komandos „pirmyn“ intervalus). Signalo žymekliui atsidūrus viename iš šių intervalų, robotas juda pirmyn.

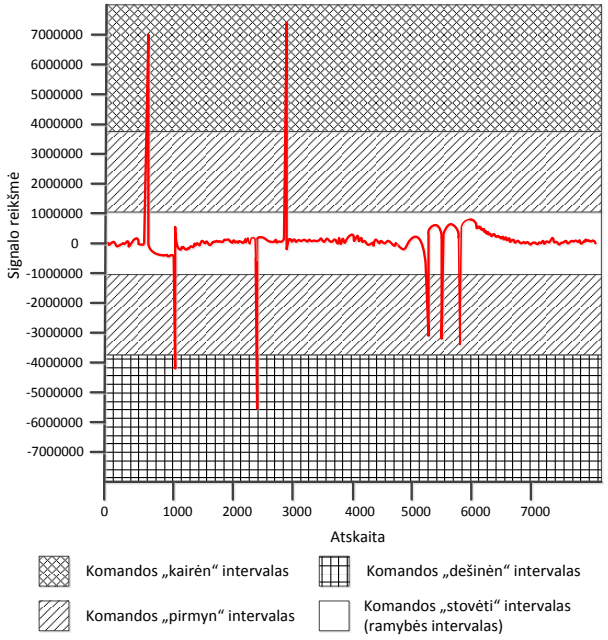


2 pav. Valdymo parametrų nustatymas ir stebėsena

Sistemos valdymo parametrai yra komandų intervalų ilgiai. Šie ilgiai gali būti keičiami kompiuterio pelyte tempiant intervalų ribų taškus (2 pav. pateiktoje vartotojo sąsajoje jie skiria atskirus intervalus).

Valdymo komandos nustatymo etape užfiksuojama momentinė roboto valdymo komanda. Užfiksuota komanda siunčiama robotui. Robotas vykdo gautą komandą tol, kol negauna kitos valdymo komandos.

3 pav. pateiktas pavyzdys, kuriame matyti, kaip signalo amplitudė sudalijama į atskirus intervalus. Konkretus intervalas žymi konkrečią valdymo komandą (žr. 3 pav.). Kad signalo reikšmė būtų viename ar kitame intervale, sistemos vartotojas turi ritmiškai įtempinėti savo kaktos raumenis ir mirksėti akimis. Merkiant kairę akį signalas įgauna teigiamą reikšmę (atsiduria komandos „kairėn“ intervale), o merkiant dešinę – neigiamą (atsiduria komandos „dešinėn“ intervale). Viduriniame intervale signalas būna tada, kai nėra įtempiami jokie raumenys, todėl šis intervalas žymi ramybės būseną. Komandos „pirmyn“ intervalai yra išsidėstę tarp komandos „kairėn“ intervalo ir ramybės intervalo bei tarp komandos „dešinėn“ intervalo ir ramybės intervalo (žr. 3 pav.). Bandymai parodė, kad lengviausiai į šiuos intervalus patenkama įtempiant kaktos raumenis. Tuo metu generuojama vidutinio dydžio signalo amplitudė. Priklausomai nuo žmogaus, ji gali būti teigiama arba neigiama, todėl judėjimui pirmyn ir buvo išskirti 2 intervalai – teigiamas ir neigiamas (žr. 3 pav.).



3 pav. Signalo amplitudės padalijimas į intervalus

Veiksmai, kuriuos reikia atlikti norint įvykdyti konkrečią valdymo komandą, yra sąlyginiai. Jie labai priklauso nuo konkretaus žmogaus dominuojančių veido raumenų, todėl galimi ir kiti komandų iniciavimo būdai (pvz., lėtas ir greitas akies primerkimas ir pan.).

III. ROBOTO VALDYMO PER SIGNALO LYGĮ TYRIMAS

A. *Eksperimento aprašymas*

Eksperimentas buvo atliekamas su 3 sveikais asmenimis, kurių amžius – tarp 20 ir 30 metų. Buvo sugeneruotas 120 užduočių sąrašas. 40 sąrašo punktų buvo skirta komandai „kairėn“, 40 punktų – komandai „dešinėn“, 40 punktų – komandai „pirmyn“. Komanda „stovėti“ eksperimento metu nebuvo registruojama, nes ji vykdoma subjektui esant ramybės būsenoje (neįtempiant jokių veido raumenų). Ši būsena yra natūrali daugumai žmonių, todėl jos generuojamos komandos tikslumas būtų beveik 100 %. Komandos „kairėn“, „dešinėn“ ir „pirmyn“ sąraše buvo išmėtytos atsitiktine tvarka. Prieš atliekant eksperimentą buvo atliekamas sistemos kalibravimas – sistemos pritaikymas eksperimento dalyviui. Tada subjektui buvo liepiama atlikti iš eilės einančias komandas iš sąrašo, tarp atskirų komandų vykdymo darant 5 sekundžių pertrauką. Šis eksperimentas buvo kartojamas 5 kartus. Tarp kiekvieno bandymo buvo daroma 15–30 min. pertrauka. Subjektui inicijavus nurodytą komandą, buvo stebima, ar „Arduino“ 4WD robotas elgiasi taip, kaip numatyta (ar vykdo tinkamą komandą). Stebėjimo duomenys buvo registruojami, ir pagal šiuos stebėjimus buvo apskaičiuotos roboto valdymo tikslumo metrikos (žr. I, II, III lenteles).

4 pav. OCZ NIA įrenginio montavimas

4 pav. parodyta, kokiu būdu eksperimento metu buvo uždėtas paviršinės elektromiogramos skaitymo įrenginys OCZ NIA.

Eksperimento metu subjektui buvo siūloma atlikti tokius veiksmus:

- Komanda „kairėn“ – mirksėti kaire akimi.
- Komanda „dešinėn“ – mirksėti dešine akimi.
- Komanda „pirmyn“ – ritmiškai įtempti kaktos raumenis.

Roboto valdymo tikslumas buvo skaičiuojamas pagal tokią formulę:

$$N = \frac{n-f}{n} \cdot 100\ \%. \qquad (1)$$

Čia N – valdymo tikslumas, n – sąraše esančių komandų skaičius, f – neteisingai atliktų komandų skaičius.

B. *Eksperimento rezultatai*

I, II ir III lentelėse pateikti visų trijų eksperimento dalyvių eksperimento rezultatai.

I lentelė. Pirmojo subjekto eksperimento rezultatai

Bandymo numeris	**Komandos „kairėn“ tikslumas**	**Komandos „dešinėn“ tikslumas**	**Komandos „pirmyn“ tikslumas**	**Bendras tikslumas**
1.	95 %	100 %	95 %	96,67 %
2.	95 %	95%	97,5 %	95, 83 %
3.	90 %	100 %	97,5 %	95,83 %
4.	80 %	97,5 %	97,5 %	91,67 %
5.	70 %	95 %	90 %	85 %
Bendras komandos tikslumas	86 %	97,5 %	95,5 %	93 %

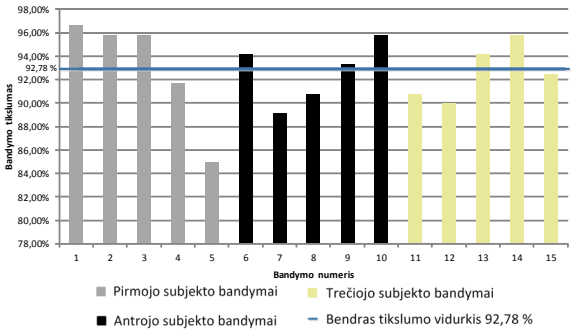
II lentelė. Antrojo subjekto eksperimento rezultatai

Bandymo numeris	**Komandos „kairėn“ tikslumas**	**Komandos „dešinėn“ tikslumas**	**Komandos „pirmyn“ tikslumas**	**Bendras tikslumas**
1.	100 %	95 %	87,5 %	94,17 %
2.	100 %	97,5%	70 %	89,17 %
3.	100 %	95 %	77,5 %	90,83 %
4.	95 %	97,5 %	87,5 %	93,33 %
5.	100 %	97,5 %	90 %	95,83 %

Bandymo numeris	Komandos „kairėn“ tikslumas	Komandos „dešinėn“ tikslumas	Komandos „pirmyn“ tikslumas	Bendras tikslumas
Bendras komandos tikslumas	99 %	96,5 %	82,5 %	92,67 %

III lentelė. Trečiojo subjekto eksperimento rezultatai

Bandymo numeris	Komandos „kairėn“ tikslumas	Komandos „dešinėn“ tikslumas	Komandos „pirmyn“ tikslumas	Bendras tikslumas
1.	97,5 %	95 %	80 %	90,83 %
2.	100 %	82,5%	87,5 %	90 %
3.	90 %	100 %	92,5 %	94,17 %
4.	92,5 %	97,5 %	97,5 %	95,83 %
5.	92,5 %	100 %	85 %	92,5 %
Bendras komandos tikslumas	94,5 %	95 %	88,5 %	92,67 %



5 pav. Bandymų tikslumų stulpelinė diagrama

5 pav. pateikta visų 15 bandymų stulpelinė diagrama. Skirtingomis spalvomis išskirti atskirų eksperimentų subjektų bandymai. Vertikalioje ašyje nurodomas eksperimentų tikslumas. Horizontali linija ties 92,78 % atžyma žymi bendrą visų bandymų tikslumų vidurkį.

IV. IŠVADOS

Atlikus eksperimentą galima teigti, kad signalo klasifikavimas pagal signalo lygį yra pakankamai efektyvus metodas. Visų trijų subjektų eksperimento metu bendras valdymo tikslumas buvo ~93 %.

Šiame straipsnyje aprašyti eksperimentai buvo atliekami su 5 sek. pertrauka tarp valdymo komandų. Taikant sistemą praktikoje, pavyzdžiui, valdant invalido vežimėlį, šis intervalas tarp valdymo komandų sukeltų nepatogumų. Eksperimento metu toks laiko intervalas buvo naudojamas todėl, kad priešingu atveju susiduriama su problema: jeigu komandos yra vykdomos iš karto viena po kitos, signalo lygis pradeda smarkiai svyruoti į priešingą pusę. Pasireiškia signalo „drebėjimo“ efektas. Tam gali turėti įtakos mūsų naudota EMG nuskaitymo įranga. Ateityje tyrimų metu bus mėginama šį neigiamą efektą pašalinti ar bent jau minimizuoti naudojant įvairias glotninimo funkcijas ar netiesinius operatorius ir tokiu būdu trumpinti laiko intervalą tarp valdymo komandų.

Taip pat svarbu didinti valdymo komandų skaičių. Tai būtų galima pasiekti formuojant nuskaitomų signalų šablonus ir naudojant tokius klasifikavimo algoritmus kaip neuroniniai tinklai, „Random Forest“ ir pan. Komandų skaičių ir (arba) valdymo tikslumą būtų galima padidinti naudojant EMG/EEG/EOG skaitymo įrenginį, kuris išskiria daugiau signalo kanalų. Kitas būdas išgauti daugiau valdymo komandų yra panaudoti keletą EMG nuskaitymo įrenginių, signalus nuskaitant nuo skirtingų kūno vietų.

LITERATŪRA

[1] U., Hoffmann, J.M., Vesin, T., Ebrahimi. „Recent Advances in Brain-Computer Interfaces“. IEEE 9th Workshop on Multimedia Signal Processing, vol. 17, 2007.

[2] Arduino 4 WD robotas. 2013-04-04. Prieiga internete: http://www.robotshop.com/dfrobot-4wd-arduino-mobile-platform-3.html.

[3] Neural Impulse Actuator. 2013-04-04. Prieiga internete: http://www.bcinet.com/products/.

[4] Y., Zhang, L., Dai, Y., Luo, H., Hu. „Design of a Surface EMG – Based Human – Machine Interface for an Intelligent Wheelchair“. Electronic Measurement & Instruments (ICEMI), 2011 10th International Conference, vol. 3, pp. 132-136, 2011.

[5] L., Wei, H., Hu, K. Yuan. „Use of Forehead Bio-Signals for Controlling an Intelligent Wheelchair“. Robotics and Biomimetics, 2008. ROBIO 2008. IEEE International Conference, pp. 108-113, 2009.

[6] M., Hashimoto, K., Takahashi, M. Shimada. „Wheelchair Control Using an EOG- and EMG-Based Gesture Interface“. Advanced Intelligent Mechatronics, 2009. AIM 2009. IEEE/ASME International Conference, pp. 1212-1217, 2009.

[7] P., Swami, R., Singh, T., Gandhi, S. Anand. „A Novel Embedded Approach for the Development of Wireless Electro-oculogram Based Human-Computer Interface“. 2012 2nd IEEE International Conference on Parallel, Distributed and Grid Computing, pp. 252-255, 2012.

[8] L., Jingkun, X., Guizhi, H., Minchai, L. Lijuan. „Application of Wireless Energy Transfer Based on BCI to the Power System of Intelligent Wheelchair“. Networking and Digital Society (ICNDS), 2010 2nd International Conference, pp. 2, 35-38, 2010.

[9] I., Martisius, M., Vasiljevas, K., Sidlauskas, R., Turcinas, I., Plauska, R. Damasevicius. „Design of a Neural Interface Based System for Control of Robotic Devices“. ICIST 2012, pp. 297-311, 2012.

[10] K., Šidlauskas, I. Martisius. „EEG duomenų klasifikavimas naudojant balsavimo ir daugiasluoksnius perceptronus“. Informacinės technologijos: 17-oji tarpuniversitetinė magistrantų ir doktorantų konferencija: konferencijos pranešimų medžiaga, p. 39-42, 2012.

[11] M., Vasiljevas, R., Turčinas, I. Martišius. „Netiesinių operatorių taikymas EEG duomenų apdorojimui“. Informacinės technologijos: 17-oji tarpuniversitetinė magistrantų ir doktorantų konferencija: konferencijos pranešimų medžiaga, p. 23-26, 2012.

[12] K. Q., Shen, C. J., Ong, X. P., Li, Z., Hui, E. P., Wilder-Smith. „A Feature Selection Method for Multilevel Mental Fatigue EEG Classification“. Biomedical Engineering, IEEE Transactions, vol. 54, no. 7, pp. 1231-1237, 2007.

Tarpusavio bendradarbiavimu pagrįstas nuotolinio mokymosi paramos sistemos modelis

Žygimantas Veiverys
Programų inžinerijos katedra
Kauno technologijos universitetas
Kaunas, Lietuva
Zygimantas.Veiverys@stud.ktu.lt

Ramūnas Kubiliūnas
Programų inžinerijos katedra
Kauno technologijos universitetas
Kaunas, Lietuva
Ramunas.Kubiliunas@ktu.lt

Santrauka. Sparti informacinių technologijų plėtra, nuolatos besikeičiančios darbo sąlygos ir priemonės, auganti konkurencija dėl darbo yra esminiai šiandieninės informacinės visuomenės bruožai. Siekdami neprarasti darbo, kiekvienos įmonės ir institucijos darbuotojai turi nuolatos kelti kvalifikaciją ir išlaikyti ją aukšto lygio. Šiame straipsnyje analizuojami kvalifikacijos kėlimo kursus baigusių dalyvių poreikiai ir įgytos kvalifikacijos palaikymo galimybės. Įvertinami besimokančiųjų dalijimosi patirtimi ir žiniomis privalumai bei priemonės. Pristatomas tarpusavio bendradarbiavimu pagrįstas nuotolinio mokymosi paramos sistemos modelis.

Reikšminiai žodžiai: nuotolinis mokymasis, paramos sistema, nuotolinio mokymosi paramos sistema, dalijimasis žiniomis, dalijimasis patirtimi, bendradarbiavimas.

I. ĮVADAS

„Informacinės visuomenės, žinių visuomenės formavimas – ne tik viena aktualiausių šiandienos suaugusiųjų švietimo sistemos krypčių, bet ir socialinė būtinybė. Tam, kad užtikrinti visuotinį informacinių resursų prieinamumą plačiausiems visuomenės sluoksniams, sėkmingai spręsti mokymosi visą gyvenimą uždavinius, daugelio šalių vyriausybės bei tarptautinės organizacijos vysto techninę infrastruktūrą, automatizuoja valstybines tarnybas ir jų valdymą, diegia elektroninę raštvedybą ir interaktyvios sąveikos su piliečiais sistemas. Tačiau vien techninėmis priemonėmis šis tikslas negali būti įgyvendintas – moderniai sistemai reikalingas toks naudotojas, kuris sugebėtų efektyviai išnaudoti jam suteikiamas galimybes“ [3]. O kiekvienai įmonei ir institucijai reikalingas toks darbuotojas, kurio kvalifikacija visada būtų aukšto lygio.

Kvalifikacijai kelti organizuojami įvairūs mokymai ir kvalifikacijos kėlimo kursai. Juose gali dalyvauti visi norintys: privatūs asmenys, privačių įmonių, taip pat valstybinių institucijų darbuotojai. Paprastai privačios įmonės savo darbuotojus siunčia į kvalifikacijos kėlimo kursus pagal poreikį, todėl jų darbuotojai būna gana motyvuoti. Naujos žinios ir įgūdžiai jiems reikalingi kokybiškai atlikti jiems pavestiem darbam. O valstybinės institucijos reguliariai organizuoja savo darbuotojų tobulinimo procesą, todėl neretai jie būna ne tokie motyvuoti. Gali būti, kad naujos žinios ir įgūdžiai jiems dar nėra reikalingi ir pravers tik ateityje, todėl jų mokymosi rezultatai būna prastesni.

Bet kuriuo atveju už įgytos kvalifikacijos palaikymą po mokymų yra atsakingi patys darbuotojai. Mokymai bei kvalifikacijos kėlimo kursai yra trumpalaikiai. Kursų metu besimokantieji turi mokymosi medžiagą, gali kreiptis pagalbos į dėstytoją. Pasibaigus kursams dėstytojas nebegali suteikti pagalbos arba gali tai daryti tik išskirtiniais atvejais. Vienintelis informacijos šaltinis mokymų dalyviams lieka mokymosi medžiaga. Tačiau, mokantis suaugusiesiems dažniau reikalingas periodinis kartojimas, medžiagos peržiūrėjimas, informacijos atnaujinimas (skirtingai nei moksleivių ar studentų mokymo procese). Per kvalifikacijos kėlimo kursus medžiaga įsisavinama trumpesniam laikui ir tik siaurai darbo specifikai panaudoti. Todėl net ir po kvalifikacijos kėlimo kursų besimokantiesiems būtina padėti. Tam būtų galima panaudoti NM paramos sistemą, kurioje kursų dalyviai galėtų kreiptis pagalbos vieni į kitus kompiuterinio raštingumo ar kitais klausimais, padėti vieni kitiems ir taip palaikyti mokymuose įgytą kvalifikaciją.

Tyrimo tikslas – sukurti ir ištirti nuotolinio mokymosi paramos sistemą, pagrįstą tarpusavio bendradarbiavimu ir užtikrinančią darbuotojų kvalifikacijos palaikymą kompiuterinio raštingumo kursams pasibaigus.

Darbo uždaviniai:

1. Įvertinti besimokančiųjų dalijimosi patirtimi ir žiniomis privalumus bei priemones nuotoliniame mokymesi.
2. Ištirti kvalifikacijos kėlimo kursus baigusių dalyvių poreikius.
3. Sudaryti ir išanalizuoti tarpusavio bendradarbiavimu pagrįstą nuotolinio mokymosi paramos sistemos modelį.

II. DALIJIMASIS ŽINIOMIS PASITELKIANT NUOTOLINĮ MOKYMĄSI

Tobulėjančios informacinės ir komunikacinės technologijos išplečia galimybes išsilavinimą įgyti nuotoliniu būdu. Nuotolinis mokymasis leidžia spręsti netolygaus mokslinio

pedagoginio personalo išsidėstymo keliamas problemas, pakelti mokymosi kokybę, kai nėra sąlygų pasitelkti kvalifikuotų pedagogų, todėl gerėja mokslo pasiekiamumas, masiškumas, dažnai atpinga išsilavinimo teikimas [2]. Nuotolinis mokymasis patogus tuo, kad nereikalauja kur nors vykti, mokytis galima patogiu laiku ir patogioje vietoje.

„Nuotolinio mokymosi“ sąvoka neatsiejama nuo „mokymosi visą gyvenimą“ sąvokos. Mokymosi visą gyvenimą paradigma atveria plačias mokymosi galimybes suaugusiesiems, kartu siūlo šiuolaikinių, į besimokantįjį nukreiptų mokymosi formų, leidžiančių mokytis pagal galimybes, tempą ir turimus laiko išteklius. Ypač svarbus efektyvus nuotolinis mokymasis visą gyvenimą. Mokymasis visą gyvenimą yra geriausias būdas sutikti visus technologinius pokyčius, integruoti visus visuomenės narius į socialinį gyvenimą ir aktyvų pilietiškumą, padėti jiems susikurti produktyvų ir juos patenkinantį gyvenimą.

Nuotolinio mokymo ar jo elementų taikymas suaugusiųjų švietimo sistemoje praplėstų mokymo programų sudarymo ir šiuolaikiškų mokymo metodų taikymo galimybes, pagerintų mokymosi prieinamumą ir padėtų priartinti mokymąsi prie besimokančiojo. Vis dėlto nuotoliniam mokymuisi reikia ir tinkamų sąlygų, ir asmens nusiteikimo bei gebėjimų, tokių kaip techniniai mokymosi proceso įgūdžiai, studijuojamo dalyko pažinimas, savęs, kaip besimokančiojo, suvokimas ir įsipareigojimas mokytis pasirinktu laiku [9].

Įgyvendinti nuotolinį mokymąsi labai padeda nuotolinio mokymosi paramos sistemos. Jų būna labai įvairių – tam tikrai kvalifikacijai palaikyti, žinioms įsisavinti, iškilusioms problemoms spręsti. Tokių sistemų pagrindiniai dalyviai yra dėstytojas (kuratorius) ir studentai bei mokomoji medžiaga. Nuotolinio mokymosi (NM) paramos sistemos kvalifikacijai palaikyti pagrindiniai dalyviai yra studentai, sistemos administratorius bei mokomoji medžiaga. Tokios NM paramos sistemos principas yra dalijimasis žiniomis ir praktine patirtimi, kai studentai, kurie turi tam tikrų klausimų kompiuterinio raštingumo srityje, gali kreiptis pagalbos į kitus studentus ir dalytis savo sukauptomis žiniomis bei patirtimi, diskutuoti.

Dalijimasis patirtimi ir žiniomis yra viena iš mokymosi apraiškų, nes padeda pasiekti ne tik gerų veiklos rezultatų, bet ir įtvirtinti sukauptą patirtį. Ši veikla prisideda prie švietimo bendruomenės kvalifikacijos ugdymo bei išlaikymo, ir drauge – prie švietimo sistemai keliamų tikslų.

Dalijimosi žiniomis procesas susijęs su informacijos aprūpinimu ir praktine patirtimi įgyvendinimo kaip padėti kitiems ir bendradarbiaujant spręsti iškylančias problemas, kelti naujas idėjas ar atlikti tam tikras procedūras. Žiniomis dalijamasi susirašinėjant ar tiesiogiai palaikant ryšius su kitais ekspertais, t. y. studentais [1]. Pačią sąvoką „dalijimasis žiniomis” mokslininkai apibūdina įvairiai. Vieni teigia, kad tai – procesas, kuris vienu metu apima kelis mainus tarp asmenų, kurie siunčia ir (arba) priima žinias [8]. Kiti teigia, kad tai – abipusis keitimasis žiniomis ir kartu, naujų žinių kūrimas, o tai reiškia sinerginį asmenų, kurie siekia bendro tikslo, bendradarbiavimą [6]. Dar dalijimasis žiniomis apibūdinamas kaip pagrindinė priemonė generuoti naujom idėjom ir plėtoti naujom galimybėm socializacijai ir mokymuisi [4]. Pasak L. Rubikio, norint sukurti aplinką, kurioje greitai suklestėtų dalijimasis žiniomis, reikia kelių komponentų [7]:

- tinkamų sąlygų – tai bendra ir patikima infrastruktūra;
- tinkamų priemonių – bendro modelio, priemonių ir procesų;
- tinkamų veiksmų – tai darbuotojų nuolatinis žinių ieškojimas, dalijimasis jomis ir pritaikymas.

Dalijimosi patirtimi veiklai keliamų tikslų siekiama įgyvendinant šias priemones [1]:

- *Besimokantiesiems sudaromos sąlygos susirasti bendraminčių.*

Šios priemonės tikslas – koncentruoti ir skleisti sukauptą patirtį. Papildomos galimybės ieškoti bendraminčių sudaro prielaidas aktyviems specialistams burtis į specializuotas grupes, skleisti patirtį ir kelti kvalifikaciją.

- *Besimokantieji įtraukiami į konsultavimo veiklą.*

Šios priemonės tikslas – paskatinti dalijimąsi patirtimi, suteikti besimokančiajam karjeros galimybių. Dalijimosi patirtimi veiklai svarbu, kad būtų sutelkta geriausia teigiama patirtis. Savo srities specialistų įtraukimas į konsultavimo veiklą padėtų skleisti gerąją patirtį.

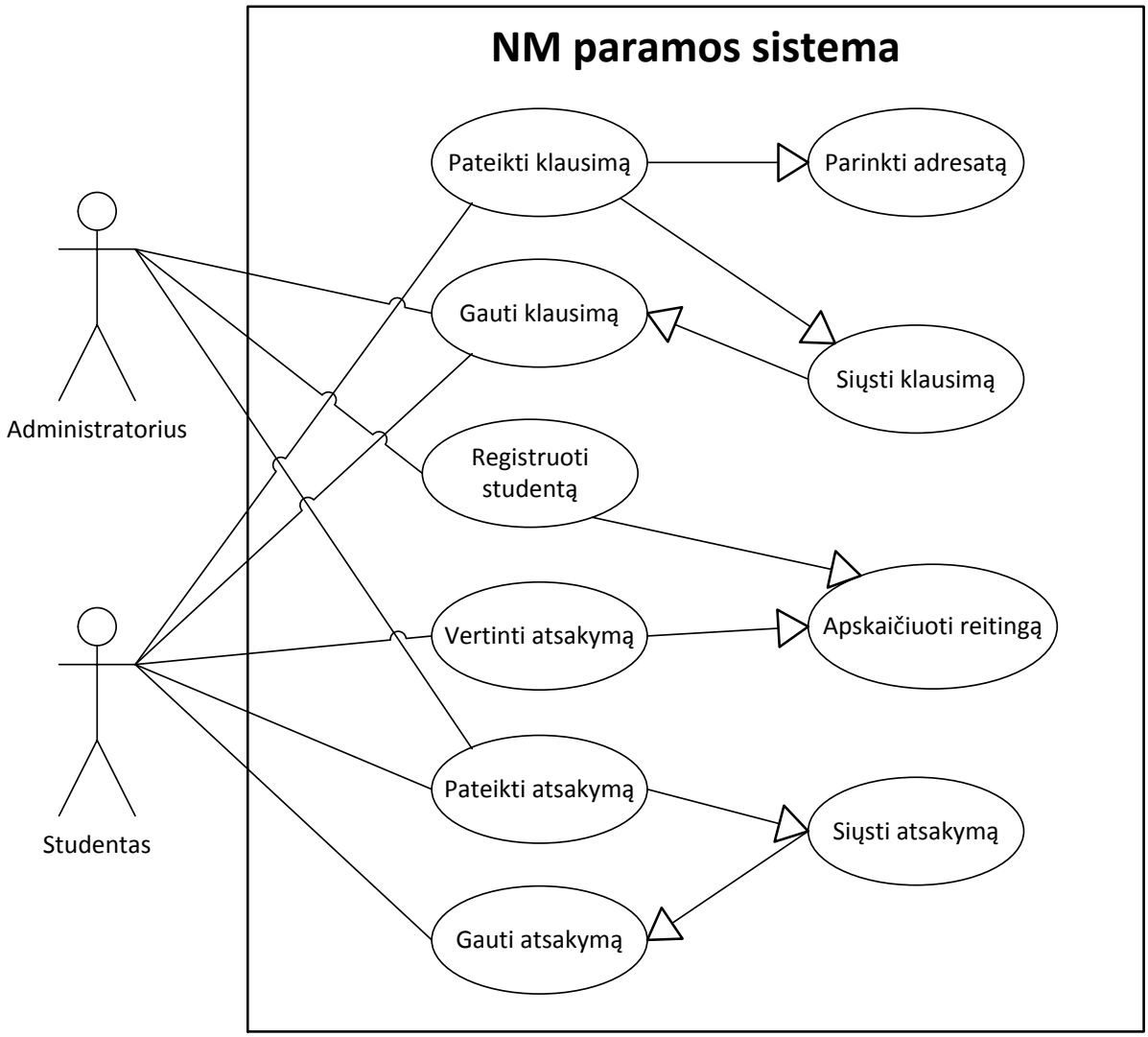
- *Besimokantieji kaupia gerąją praktiką.*

Šios priemonės tikslas – sudaryti sąlygas dalytis idėjomis, patirtimi, gerosios praktikos pavyzdžiais. Gerajai praktikai kaupti sukuriamas idėjų bankas. Jame bendruomenės nariai gali registruoti kylančias idėjas dėl mokymo veiklos tobulinimo, pateikti nuorodas į efektyvių veiklos pavyzdžių aprašymus, aprašyti savo teigiamą patirtį. Idėjų banko turinį galima naršyti, taip pat reitinguoti, tuomet populiariausios idėjos yra lengvai surandamos. Švietimo bendruomenės nariai gali susirasti ir pritaikyti jiems tinkančias kolegų siūlomas idėjas.

- *Besimokantieji naudojasi šiuolaikiškomis komunikacijos priemonėmis.*

Šios priemonės tikslas – naudojant technologijos pasiekimus padaryti komunikaciją efektyvesnę.

Dalijimosi patirtimi priemones būtų galima naudoti tarpusavio bendradarbiavimu pagrįstoje NM paramos sistemoje. Sistema vienytų kvalifikacijos kėlimo kursuose dalyvavusius besimokančiuosius. Paprastai tokiuose kursuose užsimezga draugiški ryšiai tarp kurso dalyvių, tačiau jie pasimiršta kursams pasibaigus. Iškilus klausimui sistemos dalyviai galėtų kreiptis pagalbos vieni į kitus – į tuos, kurie galbūt geriau įsisavino kursuose pateiktą mokymosi medžiagą ar geriau išlavino tam tikrus savo įgūdžius. Tad besimokantiesiems būtų ne tik suteikiama pagalba po kvalifikacijos kėlimo kursų, bet ir būtų skatinama bendradarbiauti ir palaikyti kvalifikacijos lygį. Taip pat būtų sumažinamas krūvis, tenkantis dėstytojui, jeigu tokia sistema būtų pradėta naudoti jau per kvalifikacijos kėlimo kursus, arba besimokančiųjų klausimų skaičius po kvalifikacijos kėlimo kursų.



1 pav. Tarpusavio bendradarbiavimu pagrįstos NM paramos sistemos modelis

III. POTENCIALIŲ NM PARAMOS SISTEMOS NAUDOTOJŲ POREIKIŲ TYRIMAS

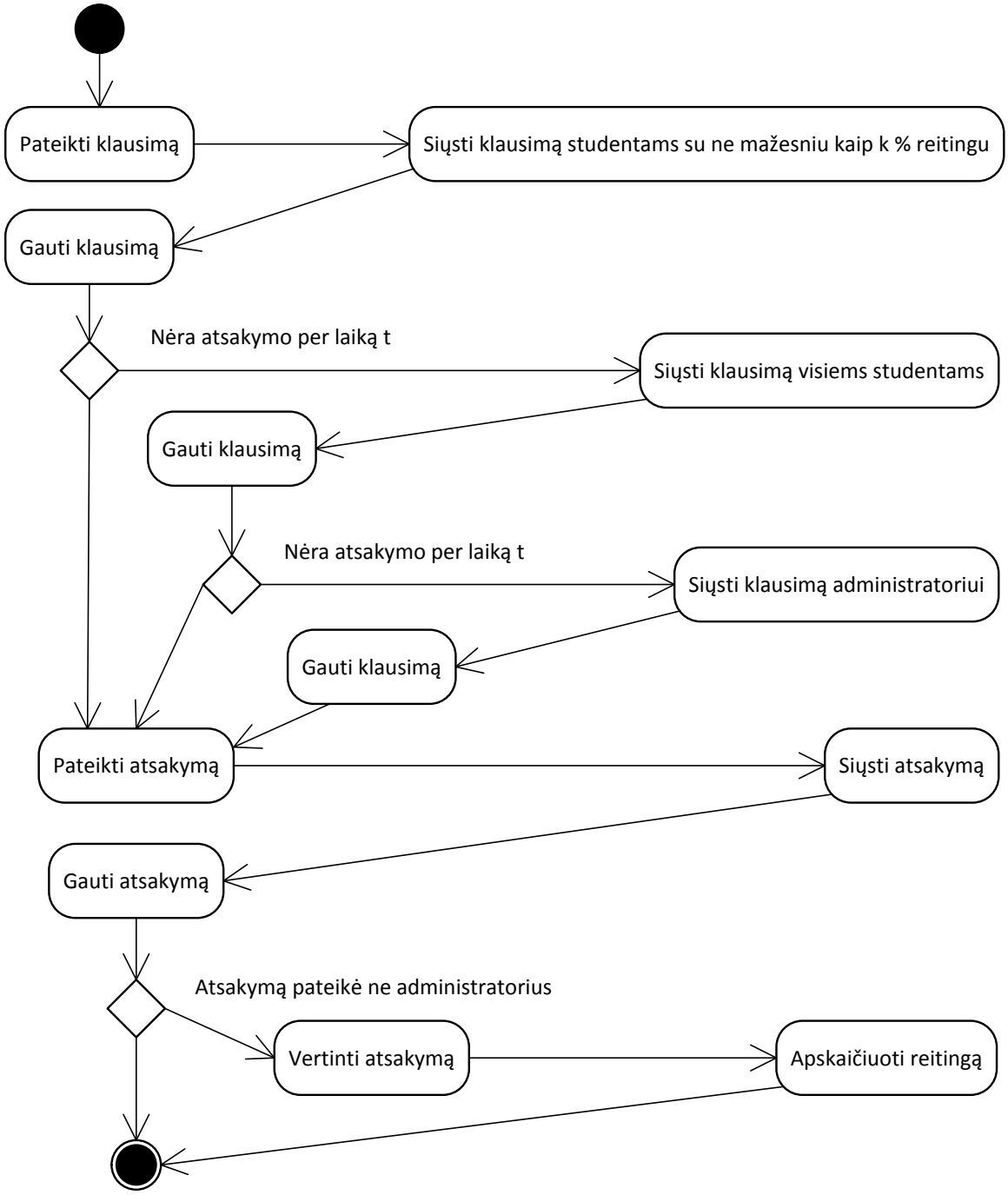
Tarpusavio bendradarbiavimu pagrįstos NM paramos sistemos funkcionalumui apibrėžti atliktas potencialių sistemos naudotojų poreikių tyrimas. Toks tyrimas atskleidžia naudotojų požiūrį, poreikius, problemas, su kuriomis tenka susidurti, ir kartu suteikia pagrindą NM paramos sistemai įgyvendinti. Tyrimui atlikti buvo sudaryta anketa ir parengt internetine apklausos priemone (http://www.apklausa.lt). Taip duodama daugiau laiko atsakyti į klausimus ir užtikrinamas didesnis anonimiškumas. Be to, internete respondentai yra atviresni nei atsakinėdami į tiesioginio interviu klausimus.

Potencialių NM paramos sistemos naudotojų poreikių tyrimo apklausoje dalyvavo 19 respondentų, baigusių kompiuterinio raštingumo kursus. Apklausoje buvo pateikta 12 klausimų, kurių tikslas buvo išsiaiškinti respondentų poreikius, įpročius ir galimybes suteikti paramą kasdieniame jų darbe, kai iškyla problemų susijusių su kompiuteriniu raštingumu. Respondentų pateikti atsakymai į anketoje numatytus klausimus leido identifikuoti sistemos poreikį ir funkcionalumą, kuris yra svarbus diegiant tarpusavio bendradarbiavimu pagrįstą NM paramos sistemą kvalifikacijai palaikyti:

1. Kompiuterinio raštingumo žinių pakartojimas yra labai reikalingas.
2. Kasdieniame darbe gana dažnai susiduriama su tam tikromis problemomis kompiuterinio raštingumo srityje.
3. Iškilus problemai pagalbos dažniausiai kreipiamasi į kolegas arba ieškoma informacijos internete.

4. Problemos sprendimą tikimasi gauti kuo greičiau, tačiau tai dažniausiai užtrunka.
5. Nėra svarbu, kas suteikia atsakymą į iškilusį klausimą (ar kolega, ar dėstytojas), jeigu tas atsakymas yra teisingas.

Pateikti duomenys rodo, kad asmenys, baigę kompiuterinio raštingumo kursus (ECDL), kasdieniniame darbe gana dažnai susiduria su įvairiomis problemomis. NM paramos sistema padėtų išspręsti iškilusias problemas greičiau ir veiksmingiau. Tokia sistema taip pat galėtų naudotis bet kokios srities kvalifikacijos kėlimo kursus baigę asmenys.



2 pav. Iškilusios problemos sprendimo algoritmas

IV. TARPUSAVIO BENDRADARBIAVIMU PAGRĮSTOS NM PARAMOS SISTEMOS MODELIS

Baigus tam tikrus kursus ir išlaikius egzaminus, dažniausiai įgytos žinios greitai užsimiršta, jeigu jos nenaudojamos praktikoje. Todėl kvalifikacijos kėlimo kursus ar kitus mokymus baigę asmenys, darbe taikydami sukauptas žinias, susiduria su įvairiais klausimais. Tarpusavio bendradarbiavimu pagrįstoje NM paramos sistemoje, sukurtoje pasitelkus informacines technologijas, jie galėtų pasidalyti sukauptomis žiniomis ir patirtimi bei taip išspręsti iškilusius klausimus.

Tarpusavio bendradarbiavimu pagrįstos NM paramos sistemos dalyviai ir naudotojai yra administratorius ir studentai (1 pav.). Studentai – tai kvalifikacijos kėlimo kursuose dalyvavę besimokantieji. Sistemos administratorius – asmuo, prižiūrintis sistemos veiklą ir padedantis išspręsti klausimus, į kuriuos studentai nežino arba neranda atsakymų. Studentas paramos sistemoje gali pateikti klausimą ir gauti atsakymą, gauti klausimą ir pateikti atsakymą, įvertinti atsakymą į savo užduotą klausimą. Atlikus vertinimą sistema perskaičiuoja atsakymą pateikusio studento reitingą. Administratorius sistemoje gali registruoti naudotojus, taip pat kaip studentai gauti klausimą bei pateikti atsakymą. Registruodamas studentus po kvalifikacijos kėlimo kursų, administratorius taip pat suveda kursuose dalyvavusių besimokančiųjų – studentų įvertinimus, pagal kuriuos apskaičiuojamas pradinis studentų reitingas. Kuo aukštesnis buvo įvertinimas kursuose, tuo aukštesnis studento reitingas. Gali būti, kad kai kuriems studentams vienos temos sekėsi geriau, o kitos – blogiau. Atitinkamai pastarųjų temų kategorijose studento reitingas yra žemesnis.

Tarpusavio bendradarbiavimu pagrįsta NM paramos sistema veikia pagal klausimo – atsakymo algoritmą (2 pav.). Iškilus klausimui studentas NM paramos sistemoje užduoda – pateikia jam rūpimą klausimą. Klausimas pirmiausia siunčiamas ir matomas tik tiems studentams, kurie turi ne mažesnį kaip *k* procentų reitingą (pvz., didesnis nei vidutinis kurso dalyvių įvertinimas). Jei atsakymas gaunamas per laiką *t* (pvz., panašiose sistemose numatyta reikšmė yra 30 minučių) ir yra tinkamas užklausėjui, jis gali būti vertinamas 100 % skalėje. Tokiu būdu studentai yra reitinguojami. Pagal reitingą atrenkami daugiausiai žinių turintys studentai bei tie, kurie potencialiai greitai ir tiksliai atsako į klausimus. Jei atsakymas nepateikiamas per laiką *t*, klausimas siunčiamas visiems studentams. Jeigu atsakymas per papildomą laiką *t* negaunamas ir iš kitų studentų, galiausiai jis persiunčiamas administratoriui.

Klausimai iš karto nėra siunčiami visiems studentams, siekiant gauti kuo tikslesnį atsakymą. Be to, taip neapkraunami klausimais tie studentai, kurie greičiausiai net nežino atsakymo į pateiktą klausimą, bei administratorius (gali būti dėstytojas), kuris paprastai ir taip yra gana užsiėmęs. NM paramos sistema parodo klausimus ir atsakymus visiems sistemos naudotojams tuomet, kai atsakymas yra įvertintas, t. y. užklausėjui gautas atsakymas buvo priimtinas. Klausimai yra išsiunčiami ne tik sistemoje, bet ir į išorinį naudotojų elektroninį paštą, kadangi ne visi asmenys tuo metu gali būti prisijungę prie NM paramos sistemos. Jeigu atsakymo nesulaukiama nei iš aukščiausius įvertinimus turinčių asmenų, nei apskritai iš visų naudotojų, klausimą gauna administratorius, kuris privalo padėti išspręsti iškilusią problemą.

Tarpusavio bendradarbiavimu pagrįsta NM paramos sistema skiriasi nuo įprasto diskusijų forumo tuo, kad klausimus pirmiausia gauna tik tie sistemos dalyviai, kurie gali žinoti atsakymus, o ne visi. Be to, studentai skatinami patys spręsti iškilusias problemas. Taip jiems ne tik suteikiama pagalba, bet ir padedama išlaikyti įgytą kvalifikaciją.

V. IŠVADOS

Pasibaigus kvalifikacijos kėlimo kursams besimokantieji dažnai susiduria su kvalifikacijos palaikymo problema. Mažai naudojant ir nekartojant įgytų žinių, jos greitai pasimiršta. Todėl po kursų besimokantiesiems taip pat reikalinga parama, kurią galėtų suteikti kiti besimokantieji, pasitelkdami į pagalbą tarpusavio bendradarbiavimu, dalijimosi patirtimi ir žiniomis pagrįstus metodus bei priemones. Tokiu būdu skatinama bendradarbiauti, savarankiškai spręsti problemas tarpusavyje, kurti ir atnaujinti naujas žinias ir geriau palaikyti įgytą kvalifikaciją.

Besimokantiesiems visiškai nėra svarbu, kas atsako į jiems iškilusius klausimus – kolega ar dėstytojas. Pasibaigus kursams, dažniausiai kreipiamasi pagalbos į kolegas darbe. Svarbiausia, kad į iškilusį klausimą būtų gautas tikslus atsakymas ir per kuo trumpesnį laiką.

Tarpusavio bendradarbiavimu pagrįsta NM paramos sistema veikia pagal klausimo – atsakymo principą ir atsakymų reitingavimą. Nuo panašių, pagal forumo principą veikiančių sistemų ji skiriasi tuo, kad leidžia ne tik tarpusavyje dalytis žiniomis ir praktine patirtimi, bet ir padeda gauti kokybišką problemos sprendimą greitai, neapkraudama klausimais visų sistemos dalyvių.

LITERATŪRA

[1] A. Rutkienė, E. Trepulė, „Nuotolinis suaugusiųjų mokymas(is) mokymosi visą gyvenimą kontekste“, Acta Paedagogica Vilnensia, t. 23, pp. 29–42, 2009.

[2] D. Rutkauskienė, A. Targamadzė ir kt. Nuotolinis mokymasis. Mokomoji knyga. Kaunas: Technologija, 2003. 256 p. ISBN: 9955-09-321-8.

[3] G. Gedvilienė, S. Karasevičiūtė, „Suaugusiųjų požiūris į IKT žinių visuomenėje“. Aukštųjų mokyklų vaidmuo visuomenėje: iššūkiai, tendencijos ir perspektyvos, t. 1, Nr.1, pp. 67–74, 2012.

[4] H. Lin, „Knowledge Sharing and Firm Innovation Capability: An Empirical Study“, International Journal of Manpower, vol. 28, no. 4, pp. 315−332, 2007.

[5] J. N. Cummings, „Work Groups, Structural Diversity, and Knowledge Sharing in a Global Organization“, Management Science, vol. 50, no. 3, pp. 352−364, 2004.

[6] M. Gagne, „A Model of Knowledge-Sharing Motivation“, Human Resource Management, vol. 48, no. 4, pp. 571– 589, 2009.

[7] L. Rubikis, „Žinios kaip konkurencinis pranašumas“, Verslas, 2011.

[8] S. Michailova, D. B. Minbaeva, „Organizational Values and Knowledge Sharing in Multinational Corporations, The Danisco Case“, International Business Review, vol. 21, pp. 59–70, 2012.

[9] Švietimo ir mokslo ministerija. „Virtualios aplinkos koncepcija“. Lyderių laikas, pp. 1–137.

Semantinių informacinių sistemų kūrimas virtualizuotų reliacinių duomenų bazių pagrindu

Tomas Vileiniškis, Rita Butkienė

Kauno technologijos universitetas, Informacijos sistemų katedra, Studentų g. 50, Kaunas, Lietuva, tomas.vileiniskis@ktu.lt, rita.butkiene@ktu.lt

Santrauka. Straipsnyje nagrinėjamos semantinio tinklo technologijų ir virtualizuotų reliacinių duomenų bazių panaudojimo kuriant informacines sistemas (IS) galimybės. Apžvelgiami tam skirti įrankiai, technologijos, išryškinami jų privalumai ir trūkumai. Straipsnyje pateikiama formalizuota semantinių IS kūrimo metodika naudojant D2RQ platformą. Metodika nusako įrankio panaudojimo žingsnius reliacinėms duomenų bazėms virtualizuoti, jų turiniui skelbti tinkle ir toliau pritaikyti IS kūrimo procese. Taip pat pasiūlomas semantinių IS architektūrinis sprendimas, grindžiamas MVC paradigma. Siūlomi sprendimai įvertinami eksperimentiškai, įgyvendinant virtualizuota reliacine duomenų baze grindžiamą semantinę IS.

Reikšminiai žodžiai*: *semantinis tinklas, duomenų virtualizavimas, D2RQ, RDF, RDB.

I. ĮVADAS

Pastaruosius keletą metų vis plačiau kalbama apie pasaulinį semantinį tinklą (angl. *semantic web*) – tinklą, kuriame kompiuteriai geba suvokti duomenų prasmę ir efektyviau juos apdoroti. Dabartinį pasaulinį tinklą (angl. *world wide web*) (toliau – WWW) sudaro įvairūs dokumentai, tarpusavyje susieti kontekstinėmis nuorodomis. O semantinio tinklo koncepcija yra grindžiama globalios duomenų bazės egzistavimu, kurioje duomenys tarpusavyje susieti prasmingais semantiniais ryšiais. Viešai prieinami semantiniais ryšiais susieti duomenys ne tik suteiktų galimybę kompiuteriams efektyviau juos apdoroti, bet ir atvertų plačias IS funkcionalumo praplėtimo galimybes, tokias kaip duomenų integravimas su išoriniais duomenų šaltiniais, sudėtingų taikomųjų uždavinių sprendimas, įvairiapusė informacijos paieška.

Nors semantinį tinklą kuriančios technologijos vystomos gana sparčiai, jų praktinio taikymo principai dažnai lieka migloti. Vienos iš pagrindinių to priežasčių yra neaiški semantinių sistemų architektūra ir semantiniais ryšiais susietų duomenų trūkumas. Pastarajai problemai spręsti būtina identifikuoti didžiausius WWW tinklo duomenų šaltinius. Remiantis 2007 m. atliktu tyrimu [1], neindeksuojamose duomenų bazėse laikoma beveik 500 kartų daugiau duomenų negu indeksuojamame statiniame tinkle, taip pat net 70 % tinklalapių pagrindas yra reliacinės duomenų bazės (toliau – RDB). RDB duomenims virtualizuoti ir skelbti semantiniame tinkle egzistuoja keletas įrankių, tačiau skurdžios jų specifikacijos, neretai ir ribotos funkcinės galimybės apsunkina jų praktinį taikymą semantinėse IS. Formali semantinių IS kūrimo naudojant RDB virtualizavimo įrankius metodika sumažintų atotrūkį tarp teorinio ir praktinio jų taikymo aspektų, o IS kūrimo procesą padarytų aiškesnį ir nuoseklesnį.

II. SEMANTINIO TINKLO TECHNOLOGIJOS IR STANDARTAI

Naujosios kartos tinklo architektūrą atspindi semantinio tinklo technologijų dėklas [2]. Jame hierarchiškai, skirtinguose sluoksniuose pateikiamos technologijos ir standartai būtini semantinio tinklo įgyvendinimui. Kiekvienas sluoksnis išnaudoja hierarchiškai žemesniame sluoksnyje esančių technologijų teikiamas galimybes. Kadangi semantinis tinklas yra WWW tinklo plėtinys, dalis technologijų ir standartų semantinio tinklo dėkle naudojami būtent iš dabartinio tinklo. Tai – URI (angl. *uniform resource identifier*) [3] tinklo išteklių identifikavimo standartas, XML (angl. *extensible markup language*) [4] duomenų struktūrizavimo kalba ir Unikodo (angl. *Unicode*) [5] standartas simboliams koduoti kompiuteriuose.

Informacijos aprašymo ir pateikimo pagrindu semantiniame tinkle laikomas RDF (angl. *resource description framework*) [6] standartas, kuris remiasi grafo tipo duomenų modeliu. Kiekvienas RDF teiginys aprašomas trilypiu duomenų rinkiniu (angl. *triple*), sudarytu iš subjekto (angl. *subject*), predikato (angl. *predicate*) ir objekto (angl. *object*). Šio standarto dėka struktūriškai aprašomiems duomenims suteikiami semantiniai ryšiai, kurie yra naujosios kartos tinklo esminis bruožas. Baziniai semantiniai ryšiai nusakomi RDF žodynų aprašymo kalbos RDFS (angl. *RDF schema*) [7] elementais (klasėmis ir savybėmis). O OWL (angl. *web ontology language*) [8] ontologijų užrašymo kalba praplečia RDFS galimybes, suteikdama kur kas aukštesnį išraiškingumo laipsnį. Pagal ekspresyvumą skiriami trys OWL kalbos lygiai (nuo mažiausio iki didžiausio): OWL „Lite“, OWL DL (angl. *description logics*) ir OWL „Full“. Skirtingai nei RDFS, OWL kalba leidžia nusakyti ne tik hierarchines klasių priklausomybes, sąsajas tarp klasių ir jų savybių, bet ir tranzityvias, atvirkštines klasių priklausomybes. SPARQL (angl. *SPARQL protocol and RDF query language*) [9] užklausų protokolas ir kalba yra semantinio tinklo standartas, skirtas informacijai išrinkti bei įterpti RDF grafuose. SPARQL užklausos taip pat gali būti vykdomos kelių RDF šaltinių atžvilgiu vienu metu.

III. RDB TURINIO PUBLIKAVIMO SEMANTINIAME TINKLE METODAI

Didėjantis RDB turinio publikavimo semantiniame tinkle poreikis paskatino ne tik spartų šiam procesui reikalingų įrankių bei metodų kūrimą, bet ir W3C konsorciumo RDB2RDF [10] darbo grupės įkūrimą. Šios grupės pagrindiniai tikslai yra tirti egzistuojančius RDB2RDF metodus ir sukurti standartizuotą RDB2RDF kalbą. RDF duomenims iš RDB turinio išgauti skiriami du pagrindiniai metodai [11]:

- Esamos RDB turinys gali būti fiziškai transformuojamas į RDF formatą, pasitelkiant ETL (angl. *extract-transform-load*) metodiką, ir įdedamas į RDF saugyklą (angl. *RDF store*). RDF saugykla gali saugoti daugybę RDF duomenų ir taip palengvinti jų tarpusavio integravimo uždavinius. Tačiau bene didžiausias šio metodo trūkumas yra visiškas RDB atsiejimas nuo RDF kopijos – dinamiškai kintant RDB turiniui, RDF kopija išlieka nepakitusi, t. y. ji vaizduoja pirminę, seną RDB kopiją.
- Alternatyvus metodas, nereikalaujantis fiziškai transformuoti RDB duomenų į RDF formatą, vadinamas virtualiu vaizdavimu (angl. *Mapping*). Šis metodas remiasi RDB elementų vaizdavimo į ontologijos elementus taisyklių sukūrimu, kurių pagrindu vykdomos SPARQL užklausos išverčiamos į standartines RDB SQL užklausas. RDB turinio vaizdavimas į virtualius RDF grafus realiu laiku panaikina pirmojo metodo trūkumą – virtualus RDF grafas visuomet pateiks „šviežiausius“ RDB duomenis. Šio metodo silpnoji pusė yra sudėtingesnis duomenų integravimo procesas, kuriam įgyvendinti neretai tenka naudoti jungtines (angl. *federated*) SPARQL užklausas.

Nors abu aprašyti metodai turi savų trūkumų, virtualaus vaizdavimo metodas daugeliu atveju yra aktualesnis, ypač kai RDB naudojama aktyvioje sistemoje, o jos turinys sparčiai kinta. Fizinio RDF duomenų transformavimo į RDF metodas pagal ETL taikytinas tais atvejais, kai semantizuojami duomenys iš liktinių RDB, prieš tai įsitikinus, kad RDB nebėra pildoma naujais duomenimis arba duomenų atnaujinimo dažnis nedaro įtakos anksčiau suformuotų semantinių ryšių korektiškumui.

IV. RDB TURINIO VIRTUALIZAVIMO IR PUBLIKAVIMO SEMANTINIAME TINKLE ĮRANKIAI

RDB turiniui virtualizuoti ir skelbti semantiniame tinkle egzistuoja keli įrankiai: D2RQ platforma [12], „SquirrelRDF“ [13], „Triplify“ [14], „Virtuoso RDF Views“ [15]. Pagal W3C konsorciumo RDB2RDF darbo grupės sudarytas rekomendacijas, kiekvieno iš RDB2RDF įrankių tiesioginis taikymas turėtų suteikti galimybę:

- RDF duomenis pasiekti per SPARQL užklausas;
- RDF išteklius (jų URI) pasiekti HTTP protokolu;
- generuoti RDF fizines kopijas (angl. *dump*) ir užtikrinti prieigą prie jų HTTP protokolu.

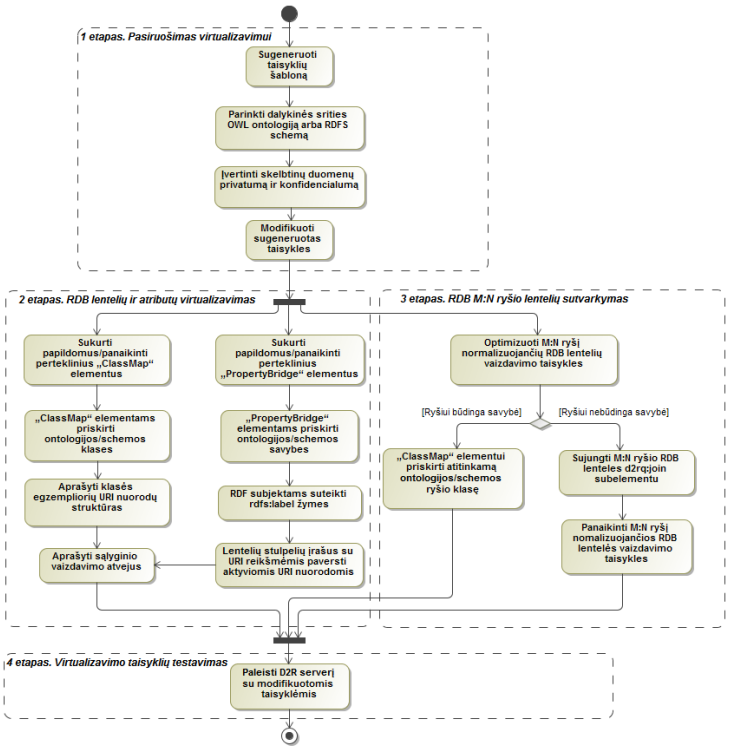
RDB2RDF darbo grupės aprašyti reikalavimai labiau atspindi gautų RDF duomenų panaudos atvejus viso semantinio tinklo kontekste. Siekiant įvertinti RDB2RDF įrankių taikymo galimybes semantinių IS kūrimo atžvilgiu, kiekvienas įrankis įvertintas pagal atitinkamus kriterijus, tokius kaip suderinamumas su skirtingomis duomenų bazių valdymo sistemomis (DBVS), virtualaus vaizdavimo taisyklių failo generavimo ir modifikavimo galimybės, SPARQL užklausų palaikymas, programinių karkasų API (angl. *application programming interface*) prieigos užtikrinimas ir atvirojo kodo filosofijos laikymasis. Apibendrinta įrankių galimybių suvestinė pateikta 1 lentelėje. Plačiausiomis teorinio pritaikymo kuriant semantines IS galimybėmis pasižymi D2RQ platforma. Įrankį sudarantys komponentai (D2RQ variklis ir D2R serveris) sudaro mechanizmą daugeliui minėtų funkcinių reikalavimų patenkinti.

V. SEMANTINIŲ IS NAUDOJANT D2RQ PLATFORMĄ KŪRIMO METODIKA

Metodika skirta nusakyti esminiam žingsniam, kaip kurti semantines IS naudojant D2RQ platformą. Metodiką sudaro dvi pagrindinės dalys: pirmoji dalis nukreipta į D2RQ taikymą RDB duomenims virtualizuoti, o antroji – D2RQ platformos (kartu ir D2R serverio) bei programinių karkasų integravimą IS. Formalizuojant semantinių IS kūrimo žingsnius, siekiama visą procesą padaryti nuoseklesnį ir aiškesnį. Metodika turėtų būti traktuojama kaip pagalbinė priemonė RDB administratoriams skelbti RDB turinį semantiniame tinkle, o IS architektams / programuotojams – bazinis algoritmas semantinėms IS kurti.

1 lentelė. RDB2RDF įrankių galimybių suvestinė

Kriterijus / Įrankis	Skirtingos DBVS	Automatinis virtualizavimo taisyklių generavimas	Taisyklių failo modifikavimas	SPARQL palaikymas	Programinių karkasų API prieiga	Atvirojo kodo filosofija
D2RQ	+	+	+	+	+	+
„SquirrelRDF“	+	+	+	+	-	+
„Virtuoso RDF Views“	-	+	+	+	+	-
„Triplify“	-	-	+	-	-	+



1 pav. RDB duomenų virtualizavimo algoritmas

A. RDB duomenų virtualizavimas naudojant D2RQ virtualizavimo kalbą

Pagrindiniai RDB duomenų virtualizavimo etapai, kurie matyti 1 pav., yra:

1. Pasiruošimas virtualizavimui.
2. RDB lentelių ir atributų virtualizavimas.
3. RDB M:N ryšio lentelių sutvarkymas.
4. Virtualizavimo taisyklių testavimas.

1 etapas. Pasiruošimas virtualizavimui

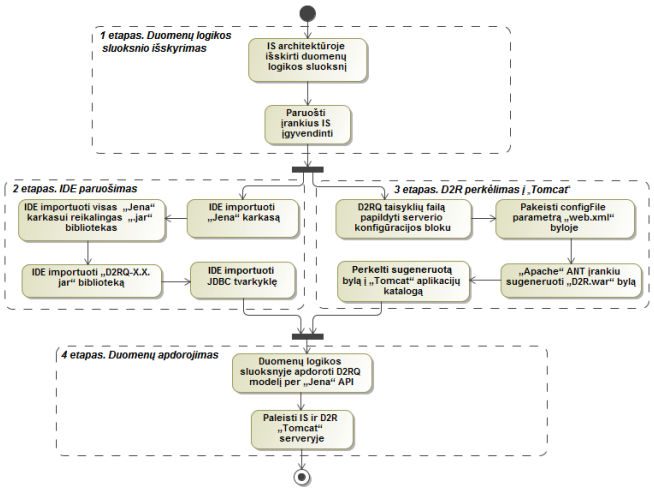
Pirmajame etape automatiškai sugeneruojamas virtualizavimo taisyklių šablonas, kurio turinį nulemia RDB schemos struktūra. Prieš modifikuojant sugeneruotą šabloną, reikia pasirinkti dalykinės srities ontologiją(-as) ar RDFS schemą(-as), kurių pagrindu RDB elementams bus suteikiami semantiniai ryšiai. Itin svarbu nuspręsti, kokius RDB duomenis norime ir galime skelbti. Tai neturėtų būti vartotojų asmeniniai duomenys, slaptažodžiai ar verslo komercinės paslaptys. Konfidencialių, privačių duomenų virtualizavimo taisyklės iš šablono turi būti pašalinamos tolimesniuose etapuose.

2 etapas. RDB lentelių ir atributų virtualizavimas

RDB elementai virtualizuojami modifikuojant sugeneruotą taisyklių šabloną. D2RQ automatinio taisyklių generavimo skriptas, išanalizavęs RDB struktūrą, sugeneruoja virtualizavimo taisykles, sudarytas iš šių pagrindinių elementų:

- „ClassMap“ – elementas, nusakantis, kokios ontologijos klasės egzemplioriais virsta RDB lentelės eilutės. „ClassMap“ elemente taip pat nurodoma sugeneruotų egzempliorių URI struktūra.
- „PropertyBridge“ – elementas, nusakantis, kokiomis ontologijos savybėmis atvaizduojami RDB lentelės atributai.

Sugeneruotose šabloninėse taisyklėse „ClassMap“ ir „PropertyBridge“ elementams priskiriamos numatytosios fiktyvios ontologijos klasės ir savybės, kurios atlieka tik laikiną (angl. *placeholder*) vaidmenį. Ontologijos klasės ir savybės, kartu ir jų vardų sritis (angl. *name space*), pakeičiamos egzistuojančiomis iš pasirinktos dalykinės srities ontologijos. Jei dalykinės srities ontologija yra skurdi, nepasižyminti pakankamu savybių skaičiumi, RDF ištekliams rekomenduojama suteikti savybes iš bendrosios



2 pav. D2RQ platformos parengimo taikymui IS algoritmas

paskirties ontologijų / žodynų (RDFS, FOAF, „vCard" ir kt.). D2RQ kalbos specifikacija [16] taip pat aprašo subelementus sąlyginio virtualizavimo ir kitiems nestandartiniams atvejams.

3 etapas. RDB M:N ryšio lentelių sutvarkymas

Trečiasis etapas gali būti vykdomas lygiagrečiai antrajam. D2RQ taisyklių generavimo skriptas automatiškai sukuria virtualizavimo taisykles ir tarpinėms M:N ryšį normalizuojančioms lentelėms. Jeigu pastarosios lentelės neturi jokių kitų atributų (išskyrus normalizuojamų lentelių pirminio rakto atributus), kitaip tariant, jeigu M:N ryšiui nebūdinga jokia savybė, tarpinės lentelės virtualizavimo taisyklės tampa perteklinės. Tokiu atveju M:N ryšio lentelės sujungiamos tam skirtu D2RQ subelementu (*d2rq:join*), o perteklinės taisyklės panaikinamos. Jeigu M:N ryšiui yra būdinga savybė, sugeneruotos tarpinės lentelės virtualizavimo taisyklės paliekamos, o „ClassMap" elementui priskiriama atitinkama ontologijos ryšio klasė.

4 etapas. Virtualizavimo taisyklių testavimas

Modifikuotoms taisyklėms testuoti pasitelkiama D2RQ platformos sudedamoji dalis – D2R serveris. D2R serveris suteikia prieigą prie virtualizuotų RDB duomenų SPARQL klientams, semantinio tinklo bei paprastoms interneto naršyklėms. Analizuojant virtualių RDF išteklių HTML aprašus, galima aiškiai matyti, ar virtualizavimo taisyklės veikia taip, kaip numatyta, t. y. ar teisingai sukuriami išteklių URI adresai, ar teisingai RDB lentelės atributams priskiriamos ontologijų / RDFS žodynų savybės ir ar gaunami visi taisyklėse numatyti RDF trejetai.

B. D2RQ platformos ir programinių karkasų integravimas į IS

Pagrindiniai D2RQ platformos parengimo taikyti IS etapai, kurie matyti 2 pav., yra:

1. Duomenų logikos sluoksnio išskyrimas.
2. IDE (angl. *integrated development environment*) aplinkos paruošimas.
3. D2R serverio aplikacijos perkėlimas į „Tomcat" serverį.
4. Duomenų apdorojimas.

1 etapas. Duomenų logikos sluoksnio išskyrimas

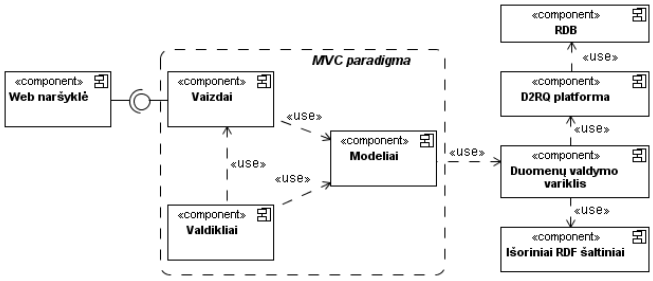
Kaip ir standartinėse IS, semantinėse IS taip pat rekomenduotina atskirti duomenų valdymo logiką nuo jų pateikimo. Kadangi semantinio tinklo technologijos suteikia galimybę sistemose naudoti paskirstytus duomenų šaltinius, nesant atskiro duomenų valdymo komponento IS architektūroje būtų sunkiau palaikyti sistemos funkcionalumą.

2 etapas. IDE aplinkos paruošimas

Semantinėms IS kurti rekomenduojama naudoti „Java" platformą, kadangi dauguma programinių karkasų, bibliotekų yra sukurti ir pritaikyti būtent „Java" kalbai. Didžiausiomis funkcinėmis galimybėmis pasižymi „Jena" [17] karkasas, skirtas RDF, OWL, RDFS apdorojimui, SPARQL užklausoms vykdyti ir kitoms funkcijoms semantiniame tinkle atlikti. Nepaisant pasirinktos IDE aplinkos, į IS projektą reikia importuoti „Jena" karkasą, visas karkaso naudojamas bibliotekas, D2RQ biblioteką ir JDBC tvarkyklę. Siekiant išvengti versijų suderinamumo problemų, vertėtų naudoti bibliotekas, esančias D2RQ programinio paketo kataloge.

3 etapas. D2R perkėlimas į „Tomcat"

Pagal numatytąjį scenarijų D2R serverio „Java" aplikacija veikia „Jetty" [18] HTTP serveryje. Kuriant didelės apimties semantines IS, tampa aktualus sistemos patikimumo lygis. Šiam lygiui pasiekti D2R serverio aplikaciją vertėtų perkelti į „Tomcat" [19] aplinką. Perkėlimo žingsniai yra virtualizavimo taisyklių papildymas konfigūracijos bloku, virtualizavimo taisyklių failo katalogo nurodymas *web.xml* byloje ir aplikacijos vykdomojo failo sugeneravimas.



3 pav. MVC paradigmos taikymas semantinių IS architektūroje

4 etapas. Duomenų apdorojimas

Išskirtame duomenų logikos sluoksnyje IS duomenys apdorojami naudojant „Jena“ API metodus. Metodai taikomi „Jena“ modeliui (šiuo atveju virtualiam RDF grafui). Alternatyvus duomenų išgavimo iš virtualaus RDF grafo būdas yra SPARQL užklausos, kurios vykdomos per ARQ užklausų varikliuką. Užklausos gali būti vykdomos ne tik virtualaus RDF grafo, bet ir išorinių SPARQL prieigos taškų (angl. *SPARQL endpoint*) atžvilgiu. Patenkinus sistemai keliamus funkcinius reikalavimus, IS paketas kartu su D2R serveriu paleidžiami „Tomcat“ aplinkoje, suteikiant prieigą ne tik prie IS, bet ir prie virtualizuoto RDB turinio.

VI. MVC PARADIGMOS TAIKYMAS SEMANTINIŲ IS ARCHITEKTŪROJE

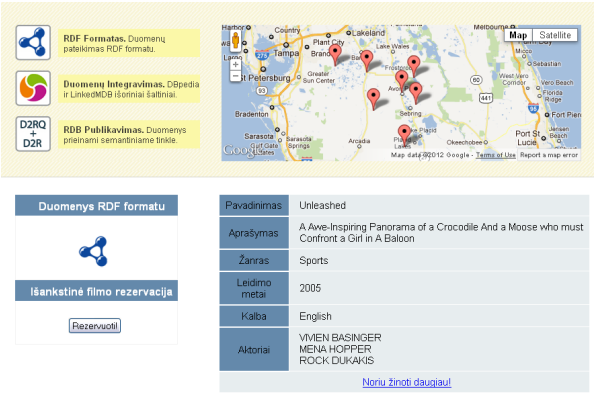
Semantinio tinklo koncepcija sąlygoja sparčiai kintančią informacinę aplinką: plintant naujosios kartos tinklo idėjai, daugėja naujų duomenų šaltinių, keičiasi egzistuojančių šaltinių struktūra ir pan. Semantinių IS architektūriniu požiūriu būtina užtikrinti mechanizmą, padedantį prisitaikyti prie greitai kintančių funkcinių reikalavimų. Siūlomas sprendimas – MVC (angl. *model-view-controller*) paradigmos taikymas semantinių IS architektūroje.

Veiklos logikos sluoksnį atskyrus nuo pateikimo ir duomenų logikos sluoksnių, sistema tampa lankstesne – tai supaprastina jos priežiūrą. 3 pav. matomas „Duomenų valdymo variklis“ atlieka ne tik dalinę DAO (angl. *data access object*) funkciją, bet sprendžia ir duomenų integravimo su išoriniais šaltiniais uždavinius. Semantinio tinklo kontekste iš esmės keičiasi tik MVC modelio komponentai, tiksliau tariant, praplečiamos jų užpildymo duomenimis galimybės. D2RQ platforma šioje architektūroje atlieka tarpinės programinės įrangos vaidmenį: virtualus RDB duomenų RDF atvaizdas suteikia galimybę pateikimo logikos sluoksnyje išvesti prasmingą informaciją ne tik vartotojui, bet ir kompiuteriams suprantamu formatu.

VII. SEMANTINĖS IS PROTOTIPAS

Remiantis 5 skyriuje pateikta metodika ir 6 skyriuje pasiūlyta architektūra, eksperimentiniams tikslams buvo sukurtas demonstracinės semantinės IS prototipas. IS yra nukreipta į filmų nuomos dalykinę sritį, kuri pasirinkta pagal MySQL mokomąją RDB „Sakila“ [20].

Vykdant metodikos žingsnius, RDB duomenys buvo virtualizuoti ir paskelbti tinkle per D2R serverį. IS išskirtame duomenų valdymo variklyje virtualizuoti RDB duomenys per SPARQL prieigos taškus yra integruojami su išoriniais „LinkedMDB“ [21] ir „Dbpedia“ [22] šaltiniais. Užklausų vykdymas išorinių prieigos taškų atžvilgiu gali užtrukti, ypač jeigu užklausa sudėtinga ar šaltinio serveris neužtikrina reikiamo pralaidumo. Todėl duomenų integravimo su išoriniais šaltiniais procesas vykdomas asinchroniškai, pasitelkiant AJAX technologiją.



4 pav. Semantinės IS prototipas

VIII. IŠVADOS

Plintant semantinio tinklo idėjai aktualu ne tik vystyti jį kuriančias technologijas, bet ir perteikti naujosios kartos tinklo teikiamą naudą eiliniam interneto vartotojui. Dėl pernelyg didelio atotrūkio tarp teorinio ir praktinio technologijų taikymo aspektų, semantinės informacinės sistemos kuriamos ganėtinai vangiai. Straipsnyje pasiūlyta semantinių IS architektūra bei jų kūrimo metodika siekiama šį procesą padaryti aiškesnį ir nuoseklesnį. Kuriant RDB grindžiamos semantinės IS prototipą, pademonstruotos praktinio sprendimų taikymo galimybės. Sukurta semantinė IS jos vartotojui leidžia pasiekti norimą informaciją kur kas mažesniu žingsnių skaičiumi, o paieškos „voriukams“ pateikia duomenis jiems suprantamu formatu (RDF). D2R serveriu paskelbtas, viešai prieinamas RDB turinys ir automatiškai sukurtas SPARQL prieigos taškas atveria galimybes spręsti įvairius taikomuosius dalykinės srities uždavinius.

LITERATŪRA

[1] B. He, M. Patel, Zh. Zhang, K. Chen-Chuan Chang, Accessing the Deep Web: A Survey, 2007.

[2] M. R. Koivunen, E. Miller, W3C Semantic Web Activity. Prieiga per internetą: <http://www.w3.org/2001/12/semwebfin/w3csw>.

[3] L. Sauermann, R. Cyganiak, Cool URIs for the Semantic Web. W3C Interest Group Note, 03 December 08. Prieiga per internetą: <http://www.w3.org/TR/cooluris/>.

[4] T. Bray, J. Paoli, C. M. Sperberg-McQueen, E. Maler, Extensible Markup Language (XML) 1.0 (Fifth Edition). W3C Recommendation, 26 November 2008. Prieiga per internetą: <http://www.w3.org/TR/xml/>.

[5] Unicode Consortium. Prieiga per internetą: <http://unicode.org/>.

[6] G. Klyne, J. J. Carroll, B. McBride, Resource Description Framework (RDF): Concepts and Abstract Syntax. W3C Recommendation, 10 February 2004. Prieiga per internetą: <http://www.w3.org/TR/rdf-concepts/>.

[7] D. Brickley, R.V. Guha, RDF Vocabulary Description Language 1.0: RDF Schema. W3C Recommendation, 10 February 2004. Prieiga per internetą: <http://www.w3.org/TR/rdf-schema/>.

[8] D. L. McGuinness, F. van Harmelen, OWL Web Ontology Language Overview. W3C Recommendation, 10 February 2004. Prieiga per internetą: <http://www.w3.org/TR/owl-features/>.

[9] E. Prud'hommeaux, A. Seaborne, SPARQL Query Language for RDF. W3C Recommendation, 15 January 2008. Prieiga per internetą: <http://www.w3.org/TR/rdf-sparql-query/>.

[10] S. Auer, L. Feigenbaum, D. Miranker, A. Fogarolli, J. Sequeda, Use Cases and Requirements for Mapping Relational Databases to RDF. W3C Working Draft, 8 June 2010. Prieiga per internetą: <http://www.w3.org/TR/rdb2rdf-ucr/>.

[11] J. Sequeda, Relational Database and the Semantic Web. Prieiga per internetą: <http://semanticweb.com/relational-database-and-the-semantic-web_b16083>.

[12] Ch. Bizer, A. Seaborne, D2RQ – Treating Non-RDF Databases as Virtual RDF Graphs. Prieiga per internetą: <www4.wiwiss.fu-berlin.de/bizer/pub/bizer-d2rq-iswc2004.pdf>.

[13] SquirrelRDF. Prieiga per internetą: <http://jena.sourceforge.net/SquirrelRDF/>.

[14] Triplify. Expose Semantics. Prieiga per internetą: <http://triplify.org/Overview>.

[15] Mapping Relational Data to RDFwith Virtuoso's RDF Views. Prieiga per internetą: <http://virtuoso.openlinksw.com/whitepapers/relational%20rdf%20views%20mapping.html>.

[16] R. Cyganiak, C. Bizer, J. Garber, O. Maresch, Ch. Becker, The D2RQ Mapping Language. Prieiga per internetą: <http://d2rq.org/d2rq-language>.

[17] Apache Jena. Prieiga per internetą: <http://jena.apache.org/>.

[18] Jetty – Jetty WebServer. Prieiga per internetą: <http://jetty.codehaus.org/jetty/>.

[19] Apache Tomcat. Prieiga per internetą: <http://tomcat.apache.org/>.

[20] Sakila Sample Database. Prieiga per internetą: <downloads.mysql.com/docs/sakila-en.a4.pdf>.

[21] Linked Movie Data Base. Prieiga per internetą: <http://linkedmdb.org>.

[22] DBpedia. Prieiga per internetą: <http://dbpedia.org>.

Stereoscopic Human-Computer Interface for Network Centric Operations

Alessandro Zocco, Davide Cannone, Andrea Sindico
Product Innovation & Advanced EW Solutions
Elettronica S.p.A.
Rome, Italy
{alessandro.zocco, davide.cannone, andrea.sindico}@elt.it

Lucio Tommaso De Paolis
Department of Engineering for Innovation
University of Salento
Lecce, Italy
lucio.depaolis@unisalento.it

***Abstract*— Network Centric Warfare can be accomplished thanks to a network of geographically distributed forces, granting a flow of increased contents, quality and timeliness of information, building up a shared situational awareness. When this flow is displayed to an operator, there is the possibility of reaching a state of information overload. To avoid this situation, new ways to conceive the interface between human and computer must be evaluated. This paper proposes an investigation about the usability of stereoscopic vision for network centric operations, using LOKI, a command and control system developed by Elettronica S.p.A.**

Keywords*— *Stereoscopic Vision, Human Computer Interface, Network Centric Operations, Command and Control System

I. INTRODUCTION

Network Centric Warfare (NCW) [21] [3], and consequently Network Centric Operations (NCO), can be accomplished thanks to a network of geographically distributed forces. The network, directly connected to the platforms by means of sensing, commanding, controlling and engaging systems, increases contents, quality, and timeliness of information between nodes enhancing the *situational awareness*. In this context it is more appropriate talking of shared situational awareness, because all the network elements can have access at the same up-to-date information.

Fig. 1 shows an example of a Network Centric scenario: two different platforms (e.g., ships), have the capability of sensing some limited areas and each one has a personal limited awareness of its proximity (a); each platform sends collected data (e.g., electromagnetic tracks) (in (a) depicted as dotted arrows starting from ships) to a specific platform, known as Command and Control (C2), that has the special task to fuse data, in a manual or automatic way; then, the C2 sends the fused data (in (b) depicted as solid arrows starting from C2) to the platforms; in this way they will share the same enhanced situational awareness (b); this process is continuously repeated during the military operation.

As described in the example, the C2 holds an important role in these networks; it is the system devoted to the decision-making process of the operational aspects of the warfare. Such systems are operated by commanders by means of a Human Computer Interface (HCI), in order to get access to information gained by the other platforms, that act like sensors in the network, and in order to make decisions (e.g., sharing the commander's perception of the situation, manifesting command decisions) [1].

Increasing the number of commanded platforms, an operator of the C2 can easily reach a state of *information overload*, where information flow rate is greater than the operator's processing rate; this situation could cause a wrong mental model of the mission scenario and, consequently, the making of wrong decisions that could lead to catastrophic situations [19].

Thus, the HCI becomes a key factor when developing the architecture of a C2.

The focus of this paper is on the display issues of a HCI and how it can be improved in order to reduce information overload and enhance the usability of information. In particular, we evaluate usability of an immersive synthetic environment in the understanding of a NCW scenario.

This research is part of the LOKI Project, a Command and Control (C2) system for Electronic Warfare (EW) developed

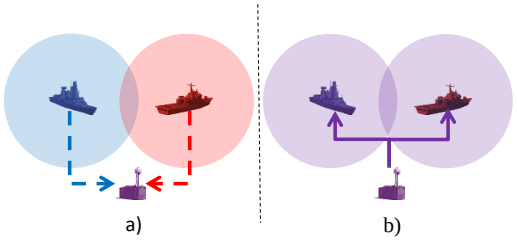


Fig.1. Example of a Network Centric scenario. In (a) each naval platform is collecting local data and send them to the C2 platform. In (b), C2 platform is sharing fused data with the naval platform, building up their shared situational awareness.

by Elettronica S.p.A. (ELT). Despite this work focuses on warfare topic, we believe that any time-pressure system operated by a human (e.g., HCI for network Intrusion Detection System), for network operations can benefit from this research [4].

II. Related Works

Gaining a detailed understanding of the modern battle space is essential for the success of any military operation.

In these applications, the main function of a human-computer interface is to display the current situation and the relevant information and intentions to the operator (e.g., location of own forces, reconnoitered opponent troops and facilities, commands and order from the superiors, platform status); these information are generally displayed on scaled maps with regional properties of the mission area.

Several research groups have focused their activities on the design and development of new display paradigms and technologies for advanced information visualization.

Dragon [11] has been one of the first research projects in formalizing requirements for systems with the need to visualize a huge amount of information on tactical maps for real-time applications. A real-time situational awareness virtual environment for battlefield visualization has been realized with an architecture composed of interaction devices, display platforms and information sources.

Other solutions have been proposed by [17] and [2]. In the former, the proposed visualization environment is based on the projection of four independent stereoscopic image pairs at full resolution upon a custom designed optical screen. This system suffers from apparent crosstalk between stereo images pairs. The latter presents some examples of Augmented Reality and Virtual Reality technologies, showing benefits and flaws, and the results of the experiments regard the evaluation of visibility and interactivity performances.

Kapler and Wright [12] have developed a novel visualization technique for displaying and tracking events, objects and activities within a combined temporal and geospatial display. The events are represented within an X, Y, T coordinate space, in which the X, Y plane shows flat geographic space and the T-axis represents time into the future and past. This technique is not adequate for an immersive 3D virtual environment because it uses an axis to describe the time evolution constrains the spatial representation on a flat surface; the altitude information, that is an important information in avionics scenarios, can't be displayed. However, it is remarkable that the splitting-up of geographical and logical information (e.g., health of a platform) can enhance the usability of the system.

III. Stereoscopic Vision

The stereoscopic vision can improve the understanding of a modern battle space by providing the depth perception and enhancing the level of realism and the sense of presence.

Different technologies have been developed for generating 3D stereoscopic visualization. Some of these are related to entertainment such as cinema [13] and video games [15], as well as to other serious/work-related applications such as medical interventions and telerobotics [7] [14].

Stereoscopic visualization, or simply stereo, can be active or passive [5]. In short, passive stereo is a solution where light is polarized differently for left and right eyes. The polarization can be obtained in various ways; the most known is the colour polarization, used in cinemas in the 1950s for the first time. Nowadays, the most used polarizations within virtual reality applications are the linear polarization or the circular one. The latter has more degrees of freedom than the former when the viewer moves the head in relation to the image, avoiding effects that may degrade the stereo perception.

Passive stereo only needs a pair of low-cost LCD projectors, a pair of light cardboard glasses, both with built-in polarization filters and a good reflecting screen that assures that the light beam isn't omnidirectional reflected, but straight back to the viewer.

The active stereo requires glasses synchronized with the projectors; so, the right and left lens are blackened out alternately at the same rate of the images projected on the screen for the left and right eyes. Active stereo is generally more expensive than passive one because the former has the requirement of a 120 Hz image frequency and the normal projectors are usually not built for this frequency.

The stereoscopic visualization artificially reproduces the mechanisms that govern the binocular vision and it is closer to the way we naturally see the world [8].

The stereoscopy leads to several improvements: comprehension and appreciation of presented visual inputs, perception of structure in visually complex scenes, spatial localization, motion judgment, concentration on different depth planes and perception of material surfaces.

However, a stereo vision could be hard to get right at first attempt because the hardware could cause crosstalk, misalignment, image distortion (due to lens, displays or projectors), and all these situations can cause eye strain, a double image perception, depth distortion, look around distortion (typical for head-tracked displays). These drawbacks prevented a large application of stereoscopic visualization [18].

IV. Proposed Investigation

The main goal of the research presented in this paper is the design and implementation of a visualization system for NCW scenarios (e.g., displaying symbols and logical information on tactical maps) by creating a stereoscopic 3D synthetic environment aimed at a total immersion of the operator. This facility is part of the LOKI Project, a C2 system for Electronic Warfare.

A. High-level architecture of LOKI

Fig. 2 shows the high-level architectural view of the LOKI system.

The *LOKI Core* component continuously executes an advanced multi-sensor data fusion process on the data retrieved from cooperating systems. Once these data are properly fused, the system is capable to infer new important information such as a better localization of emitters and countermeasures strategy. This information is transferred to the *LOKI HCI* using a communication middleware based on Data Distribution Service (DDS) paradigm [16].

The *HCI Manager* component provides a persistence mechanism to decouple the presentation layer from the core application logic. It is responsible for the communication with the core (i.e., receiving input data by the core and sending operator commands to the core) and for the translation of received data in a model understandable by the presentation layer.

The *HCI Display* component contains the elements that implement and display the User Interface (UI) and manage user interaction. It provides a high definition view of a realistic geographic environment. Platforms are positioned on the scene according to their geographic coordinates and are represented according to the Common Warfighting Symbology MIL-STD-2525C standard [6].

B. *Design choices for HCI*

The HCI has been designated with high modularity applying UI Design Patterns (UIDP). Using these patterns helps to ensure that key human factors concepts are quickly and correctly implemented within the code of advanced visual user interfaces [9]. In addition, structural patterns Composite and Decorator were used: the former allows to dynamically add properties (e.g., borders around a window) or behaviors (e.g., scrolling) to any component of the interface; the latter allows to compose interfaces as tree structures, to show part-whole structures, and lets to equally handle single objects and compositions [10].

The software has been developed in Java language~~,~~ using an OpenGL binding in order to talk to the OpenGL runtime installed on the underlying operating system.

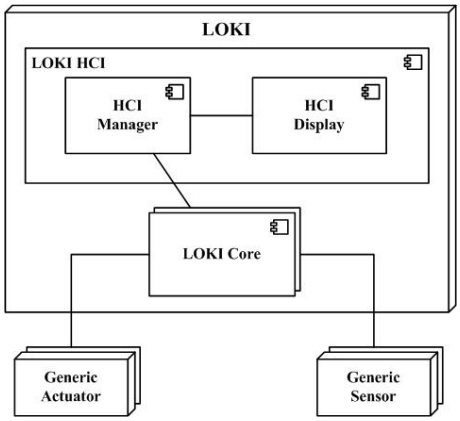


Fig.2. LOKI architecture in the large.

Fig. 3 shows a sample of the interface, built up in two different layers, as inspired by the research of [12]. A 3D terrain map, in the bottom part of the screen, is used to show both features of the selected terrain and geographic data of the elements of the scenario (e.g., real position, past track) and for elements that are not grounded, a transparent curtain is used to indicate their altitude. A parallel layer, that hosts the so-called "logical view" of the scenario, displayed above is used to represent other relevant non-geographic information (e.g., health status, lethality); it can be also used to visualize connections between the elements and to show elements that are outside the area that the operator is currently viewing in the geographic layer below. The geographical reference is maintained through connections between the two layers, using an algorithm of forces that avoids most possible crossing between lines. This separation, with the use of colours to show different levels of alerts, grants the operator the possibility to focus on geographic locations avoiding the overloading of symbols and text on the terrain.

C. *Hardware setup*

The stereo vision setup includes:

- a PC equipped with a Nvidia Quadro graphics card;
- a Digital Light Processing (DLP) projector with a WUXGA (1920x1200px) resolution and a brightness of 7000 ANSI lumens;
- a special eyewear comprising two infrared controlled Liquid Crystal Display (LCD) light shutters working in synchronization with the projector (Fig. 4).

When the projector displays the left eye image, the right eye shutter of the active stereo eyewear is closed, and vice versa. The projector is capable of displaying at a refresh rate high enough (greater than 120 Hz) that the viewer does not perceive a flicker between alternate frames. We decided to choose an active stereo system because more light is projected to each eye and therefore the 3D image appears brighter. We adjust stereo parameters (i.e., separation and convergence) in order to obtain a negative screen parallax. So, when stereo pair is viewed through shutter glasses, the 3D objects of the scene

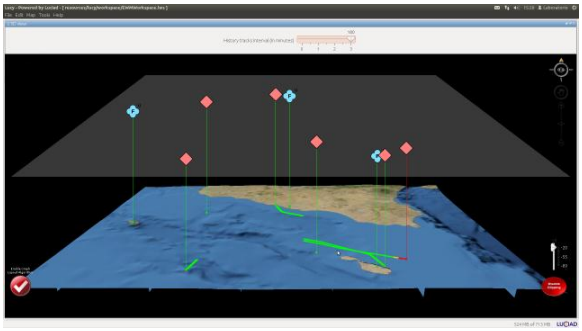


Fig.3. Visualization of EW entities with their geographic location.

appear out of the screen (Fig. 5).

D. Evaluation setup

The preliminary evaluation study took place at the facilities of ELT in Rome and involved 12 users. The case study required around half hour per participant to be completed.

To evaluate the proposal, we developed a realistic scenario based on a coastal sea surveillance task. The user supervises, using both mono and stereo visualization, a relatively big area of sea, the Strait of Sicily, where there is a large volume of traffic, generated by different types of vessels. The operator is able to see the trajectory generated by each track and intelligence regarding each vessel.

The simulator, where the scenario is developed and executed, is an integration with Commercial-Off-The-Shelf (COTS) products and proprietary software and is based on the principles of distributed and live simulation [20].

The following qualitative data were collected using questionnaires and interviews: the realism of the visual feedback, the sense of presence, the depth impression and the user's viewing comfort.

E. Results

Most of the participants had no doubts that the depth impression and the sense of presence are higher in case of stereo visualization. With complex electronic warfare scenarios, when monocular depth cues are ambiguous, the stereo viewing enhances spatial judgments: it is possible to detect very closely spaced icons on the screen (representing platforms with installed active emitters).

All users find that stereo visualization provides more realism than mono-viewing. The image resulting from the fusion of the stereoscopic pair is very clear and natural looking because surface properties such as luster, scintillation, and sheen are different in luminance and colour between the left and right retinal images. This allows the viewer to perceive the differences between successive frames (e.g., platforms' positions) at a glance.

There is no significant difference in viewing comfort between stereo and mono visualization. This result contradicts the general assumption of stereo viewing causes some problems such as visual fatigue and headache.

V. Conclusions and Future Work

In the design and development of the C2 systems for NCO, a key element is the HCI. Bad assumptions may lead to bad design choices; those in turn may lead the operator to an information overload state. To avoid this situation, new ways to conceive HCI must be explored.

In this paper we evaluated the use of a stereoscopic 3D synthetic environment, aiming at a total immersion of the operator. Preliminary evaluation shows the relevant role played by stereo visualization and its advantages in terms of sense of depth, presence and realistic viewing perception. The results presented are not authoritative in terms of metrics; however they represent the initial experimentation phase of continuing research into user interface measurement for military purposes.

In the next future, we are interested in performing a formal user study aimed at improving and extending previous evaluations. We have designed a test plan according to recommendation gathered from the literature.

The test procedure will start with a brief presentation of the project and the purpose of the evaluation study. Then a visual attention test will be performed to classify the participants' level of selective visual attention. Each user will be involved in the understanding of a complex NCW scenario within an interactive test, during which quantitative data (e.g. errors made while estimating the distance, number and percentage of tasks completed correctly) will be recorded. The last step consists in the completion of pre-designed questionnaires to acquire qualitative data referring to the users' experience with stereovision technology.

We will put special attention on the counterbalancing of the tasks as well as the sequence during the entire user study to avoid fatigue and learning effects. This aim will require the participants to perform the tests according to a precise

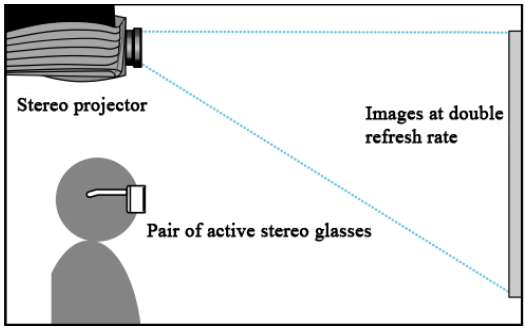


Fig. 4. Active 3D display system.

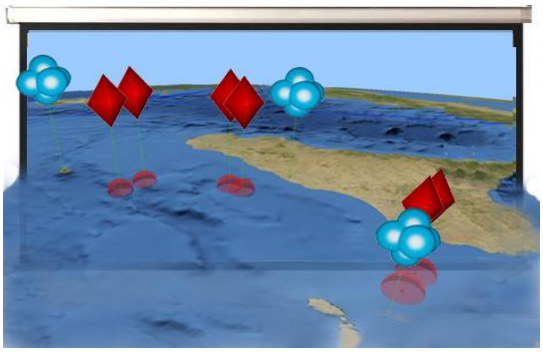


Fig. 5. Binocular fusion of the stereo pairs.

schedule.

The collected evaluation measures will be analyzed through inferential and descriptive statistics and the results will be graphically represented by means of diagrams.

We expect that formal test results will clearly confirm the benefits of the stereoscopic vision in the understanding of a NCW scenario.

REFERENCES

[1] D. S. Alberts, J. J. Garstka, and F. P.Stein. Network Centric Warfare: Developing and Leveraging Information Superiority. 2nd ed. United States: CCRP Publication Series, 1999.

[2] T.Alexander, H.Renkewitz, J.Conradi. Applicability of Virtual Environments as C4ISR Displays. United States of America, 2006, [online]. Available at: http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA473296 [accessed on 03 September 2012].

[3] T. K. Braulinger. Network Centric Warfare Implementation and Assessment. Master's Thesis. U.S. Army Command and General Staff College, 2005.

[4] K. C. Cox, S. G. Eick, and T. He. "3D Geographic Network Displays". ACM Sigmod Record, vol. 25, no. 4, pp. 50-54, 1996.

[5] B. Cyganek, and J. P. Siebert,. An Introduction to 3D Computer Vision Techniques and Algorithms. Great Britain: John Wiley & Sons, Ltd., 2009.

[6] Department of Defense. MIL-STD-2525C: Common Warfighing Symbology. United States of America, 2008, [online]. Available at: http://www.mapsymbs.com/ms2525c.pdf [accessed on 30 July 2012].

[7] D. Dey, D. G. Gobbi, P. J. Slomka, K. J. M. Surry, and T. M. Peters. "Automatic Fusion of Freehand Endoscopic Brain Images to Three-Dimensional Surfaces: Creating Stereoscopic Panoramas". IEEE Transaction on Medical Imaging, vol. 21, no. 1, pp. 23-30, 2002.

[8] D. Drascic. "Skill Acquisition and Task Performance in Teleoperation using Monoscopic and Stereoscopic Video Remote Viewing". Human Factors and Ergonomics Society Annual Meeting, vol. 35, no. 19, pp. 1367-1371, 1991.

[9] S. Feng, M. Liu, J. Wan. "An Agilely Adaptive User Interface Based on Design Pattern". IEEE, 6th International Conference on Intelligent Systems Design and Applications, Jinan, China, 16-18 October 2006. Available on: IEEE Xplore Digital Library [accessed on 5 March 2013].

[10] E. Gamma, R. Helm, R. Johnson, and J. Vlissides. Design Patterns: Elements of Reusable Object-Oriented Software. Massachussets (USA): Addison-Wesley, 2005.

[11] S. Julier, R. King, B. Colbert, J. Durbin, and L. Rosenblum. "The Software Architecture of a Real-Time Battlefield Visualization Virtual Environment". IEEE, Proceedings of Virtual Reality. Houston, Texas, USA, 13-17 March 1999. Available on: IEEE Xplore Digital Library [accessed on 26 July 2012].

[12] T. Kapler, and W. Wright. Geotime Information Visualization. Information Visualization, vol. 4, no. 2, pp. 136-146, 2005.

[13] L. Lipton. "The Last Great Innovation: The Stereoscopic Cinema". SMPTE Motion Imaging Journal, vol. 116, no. 11-12, pp. 518-523, 2007.

[14] S. Livatino, G. Muscato, S. Sessa, C. Koffel, C. Arena, A. Pennisi, D. Di Mauro, and E. Malkondu. "Mobile Robotic Teleguide Based on Video Images". IEEE Robotics and Automation Magazine, vol. 15, no. 4, pp. 58-67, 2008.

[15] N. Mahoney, A. Oikonomou, and D. Wilson. "Stereoscopic 3D in Video Games: A Review of Current Design Practices and Challenges". IEEE, 16th International Conference on Computer Games. Louisville, KY, USA, 27-30 July 2011. Available on: IEEE Xplore Digital Library [accessed on 26 July 2012].

[16] OMG, 2007. DDS, http://portals.omg.org/dds/ [accessed on 08 April 2013].

[17] L. W. Pettersson, U. Spak, and S. Seipel. "Collaborative 3D Visualizations of Geo-Spatial Information for Command and Control". : Proceedings of SIGRAD, Gävle, Sweden, 24-25 November 2004. Available on: LiU Electronic Press [accessed on 26 July 2012].

[18] I. Sexton, and P. Surman. "Stereoscopic and Autostereoscopic Display Systems". IEEE Signal Processing Magazine, vol. 16, no. 3, pp. 85-89, 1999.

[19] T. Shanker, and M. Richtel. "In New Military, Data Overload can be Deadly". The New York Times, [online] 17 January 2011. Available at: http://www.umsl.edu/~sauterv/DSS4BI/links/17brain.pdf [accessed on 27 July 2012]

[20] A. Sindico, D. Cannone, S. Tortora, G. F. Italiano, M. Naldi. "Distributed Simulation of Electronic Warfare Command and Control Scenarios". Proceedings of Symposium on Navy Operational Research and Logistics (SPOLM). Rio de Janeiro, Brazil, 23-24 August 2012.

[21] U. S. Navy. Copernicus… Forward C4I for the 21st Century. United States: Office of Chief of Naval Operations (OPNAV), 1995. Available at: http://www.navy.mil/navydata/policy/coperfwd.txt [accessed on 26 June 2012].

INFORMACINĖS TECHNOLOGIJOS ● IVUS 2013 ● ISSN 2029-249X ● eISSN 2029-4824

Statinės kodo analizės įrankio „Gendarme“ tobulinimas

Andrius Zonys
Kompiuterių katedra
Kauno technologijos universitetas
Kaunas, Lietuva
a.zonys@gmail.com

Jonas Čeponis
Kompiuterių katedra
Kauno technologijos universitetas
Kaunas, Lietuva
jonas.ceponis@ktu.lt

Santrauka. **Šiame darbe aptariama statinė ir dinaminė kodo analizė, analizės įrankių tipai bei jų paskirtis. Nagrinėjami įrankiai, skirti .NET programų analizei. Plačiau analizuojamas laisvai platinamas statinės kodo analizės įrankis „Gendarme“. Pateikiamos spragos, rastos esamose „Gendarme“ taisyklėse bei siūlomi jų patobulinimai, kurie turėtų padėti aptikti daugiau klaidų ir padidinti programų spartą, stabilumą, saugumą ir kodo skaitomumą.**

Reikšminiai žodžiai: statinė analizė, kodo tikrinimas, kodo tikrinimo taisyklės.

I. ĮVADAS

Kompiuteris tampa neatskiriama mūsų gyvenimo dalimi, o viena svarbiausių kompiuterio dalių yra programinė įranga. Programų skaičius sparčiai auga, nuolat kuriama daug specializuotų ir nespecializuotų programų verslui ir žmonėms. Kai kurioms programoms saugumas yra nelabai svarbus, tačiau daugeliui, ypač verslui skirtoms programoms, jis yra labai svarbus. Saugumas bei kokybė yra labai brangus dalykas, ypač, jei programinis kodas yra ilgas [1]. Apie 50 % programų kūrimo kainos sudaro testavimas ir derinimas [2], todėl ne visos, ypač mažosios įmonės, gali sau tai leisti. Jos taupo saugumo sąskaita ir tikisi, jog į jų sistemas nebus įsibrauta. Tačiau yra gana nebrangių ar net nemokamų programinių priemonių, padedančių užtikrinti geresnį kodo saugumą ir kokybę. Jos atlieka statinę ir dinaminę kodo analizę. Šiai analizei yra sukurta nemažai įrankių, tačiau joks įrankis ar būdas negali užtikrinti visiško saugumo.

Šiame darbe analizuojami nemokami statinės kodo analizės įrankiai .NET programoms tikrinti, ir siūlomi pasirinkto įrankio patobulinimai, kurie padės aptikti daugiau kodo klaidų ir pagerinti kodo saugumą.

II. PROGRAMINIO KODO ANALIZĖ

Kodo analizės metodus galima suskirstyti į tris pagrindines kategorijas: tai rankinis metodas, kai kodas peržiūrimas žmogaus, statinis – kai kodas tikrinamas automatizuotai jo nevykdant, ir dinaminis – kai kodas tikrinamas veikiant programai.

A. Statinė kodo analizė

Tai – programinės įrangos kodo peržiūra jo nevykdant. Peržiūros metu nagrinėjamas programos kodas, ir pagal žinomus klaidų šablonus bandoma aptikti klaidingas jo vietas.

Statinės kodo analizės įrankiai turi aptikti programinio kodo klaidas ir pranešti apie jų vietas arba turėti priemonių, leidžiančių ignoruoti aptiktas klaidas. Dauguma klaidų yra suskirstytos į grupes. Įrankiai gali ne tik pranešti apie klaidą, bet ir pasiūlyti, kaip ją ištaisyti. Kai kurie iš jų gali pateikti išsamias ataskaitas apie aptiktas klaidas [3].

Statinė kodo analizė gali būti atlikta gana anksti programavimo procese, kadangi ji atliekama nevykdant kodo. Programuotojai, naudodami statinę kodo analizę, iš karto po to, kai parašomas kodas, gali aptikti ir ištaisyti klaidas dar prieš modulių (angl. *unit*) ir integracijos testavimą. Kuo anksčiau aptinkama klaida, tuo pigiau kainuoja ją pataisyti. Tai – pagrindinis statinės kodo analizės privalumas [4].

Statinės kodo analizės įrankiai būna skirtingų tipų. Vieni analizuoja tiesioginį kodą, kiti – sukompiliuotą baitų kodą. Ir vieni, ir kiti turi savų pliusų. Kai analizuojamas tiesioginis programuotojo parašytas kodas, jis prieš analizę neiškraipomas. O – kompiliatoriaus sukompiliuotas ir optimizuotas dvejetainis kodas gali skirtis nuo originalaus. Tačiau dvejetainio kodo analizė yra kur kas spartesnė, o sparta labai svarbi tiriant daug kodo eilučių [5].

Nors įrankių tipai ir skiriasi, jie turi tuos pačius pagrindinius bruožus: skaito kodą ir sukuria jo abstraktų modelį, pagal klaidų sąrašus ir jų šablonus ieško atitinkančių vietų. Taip pat įrankiai analizuoja duomenų srautus, bando numatyti reikšmes, kurias kintamasis gali įgyti tam tikrose kodo vietose.

Dauguma produktų gali atlikti neišbaigto kodo analizę (nepritaikyti visi metodai ar nesukurtos visos klasės). Kuo kodas išbaigtesnis, tuo geriau įrankis gali aptikti klaidas [6].

Įrankiai gali aptikti retai pasitaikančias klaidas ar slaptas duris (angl. *hidden back doors*). Jie analizuoja kodą jo nevykdant, dėl to gali išskaičiuoti kur kas daugiau ryšių tarp skirtingų modulių ir komponentų.

Testuoti paprasta, tačiau parašyti testą, kuris iki galo patikrintų kokį nors modulį, nėra lengva. Kai aptinkama naujų spragų, turi būti parašyti nauji testai joms aptikti. Jei

aptinkamų klaidų sąrašas bus neatnaujinamas, įrankis neaptiks naujausių saugumo spragų.

B. Dinaminė kodo analizė

Priešingai nei statinė kodo analizė, dinaminės kodo analizės metu tiriamas kodas jį vykdant. Nors ši technika nepateikia išsamių duomenų, tačiau – turi didelį privalumą – vienintelė analizuoja kodą jį vykdant. Tokiu būdu dinaminė kodo analizė padeda aptikti bandymus pakenkti programai ir aptikti besimodifikuojančias programas. Taip pat šio tipo įrankiai gali išnagrinėti programą įvairiose aplinkose [7]. Kad dinaminė kodo analizė būtų efektyvi, reikia ganėtinai daug testų.

Dinaminės kodo analizės įrankius galima suskirstyti į tris pagrindinius tipus: tai analizuojantys atmintį, kodo padengimą ar veikimo spartą. Atminties klaidas aptikti gana sunku. Šio tipo įrankius geriausia naudoti, kai problemą ne visada įmanoma atkartoti ar atminties naudojimas per daug ar per greit išauga. Atminties spragas dažnai naudoja programišiai perimdami programų kontrolę. Įrankiai apskaičiuoja, kokia kodo dalis ar kokios funkcijos buvo patikrintos. Pasitelkiamas baltos dėžės testavimas. Naudojant kodo padengimą, lengviau aptikti nepatikrintas kodo vietas ir užtikrinti geresnę testavimo kokybę. Šio tipo įrankiai skaičiuoja, kiek kartų buvo iškviesta viena ar kita funkcija bei kiek laiko funkcija buvo vykdoma. Naudojant šio tipo įrankius lengviau aptikti silpnąsias vietas ir pagerinti programos spartą.

Dinaminės kodo analizės privalumas yra tas, kad kodas analizuojamas programai veikiant ir dėl to galima aptikti daugiau įvairių klaidų. Taip pat kai kurie dinaminės kodo analizės įrankiai leidžia pasirinkti ką analizuoti, o ką praleisti. Šio tipo analizę galima atlikti įvairioms programoms. Tačiau pagrindinė šio tipo analizės problema yra ta, kad sunku nusakyti tikslią pažeidžiamą kodo vietą. Dėl to tai reikalauja daugiau programuotojo pastangų ir ilgiau trunka ištaisyti rastą klaidą.

III. STATINĖS KODO ANALIZĖS ĮRANKIAI

Statinei .NET programų kodo analizei yra sukurta įvairių įrankių – tiek mokamų, tiek nemokamų. Šiame darbe nagrinėjame nemokamus statinės kodo analizės įrankius „FxCop“, „Smokey“ ir „Gandarme“.

„FxCop“ [8] – tai programa, kuri analizuoja kodą, parašytą .NET CLR formatu, ir ataskaitose pateikia informaciją apie aptiktas projektavimo, lokalizacijos, našumo ir saugumo problemas. Dauguma problemų susijusios su programavimo ir projektavimo taisyklių pažeidimais, išdėstytais „Microsoft“ projektavimo rekomendacijose. Jos skirtos padėti rašyti patikimą ir lengvai palaikomą kodą naudojant „.NET Framework“.

„FxCop“ yra suprojektuota visiškam integravimui į programinės įrangos kūrimo ciklą. Ji platinama ir kaip programa, kuri turi grafinę vartotojo sąsają interaktyviam darbui, ir kaip komandinės eilutės sąsaja grindžiamas įrankis, kuris gali būti integruotas į kompiliavimo procesą ar integruotas į „Microsoft Visual Studio“ programavimo aplinką.

„Smokey“ [9] yra atvirojo kodo komandinės eilutės sąsaja grindžiamas įrankis, skirtas .NET ir „Mono“ programų analizei. Tai panašus įrankis į „FxCop“ ir „Gendarme“. „Smokey v1.4“ versija palaiko .NET 3.5 ir „Mono 2.0“ versijas ir turi 233 taisykles. Jis gali sugeneruoti aptiktų klaidų ataskaitas HTML, XML ar teksto formatais.

„Gendarme“ [10] yra plečiamas, taisyklėmis grindžiamas įrankis, skirtas aptikti klaidoms .NET programose bei bibliotekose. Įrankis analizuoja programas ir bibliotekas, parašytas ECMA CIL [11] formatu, ir ieško dažniausiai pasitaikančių programavimo klaidų bei tokių, kurių kompiliatorius įprastai neaptinka ar nebando aptikti. „Gendarme“ kodui analizuoti naudoja biblioteką „Cecil“. Ši biblioteka, kaip ir pats įrankis, parašyta C# programavimo kalba.

„Gendarme“ taisyklės, kurios skirtos kodo analizei, yra naudojamos per paleidėjus (angl. *runners*). Tai – programos dalis atsakinga už taisyklių paleidimą, analizę ir klaidų pateikimą. „Gendarme“ šiuo metu turi du skirtingus paleidėjus. Vienas grindžiamas komandinės eilutės sąsaja, kitas – SWF grafine sąsaja. Komandinės eilutės paleidėjas gali išvesti rezultatus į komandinę eilutę, XML failus ar tinkamai suformuotus HTML failus. SWF grindžiamas paleidėjas leidžia pažingsniui pasirinkti modulius (angl. *assembly*), taisykles, nustatymus ir pavaizduoti analizės rezultatus kaip ir komandinės eilutės paleidėjas. Kad būtų paprastesnis, jis turi mažiau nustatymų, nei komandinės eilutės paleidėjas.

Toliau tobulinti buvo pasirinkta įrankį „Gendarme“ įrankis, kadangi jis yra atvirojo kodo, parašytas C# programavimo kalba, turi aiškias taisykles, kurios svarbios testavimo metu. Nesuprasta taisyklė dažnai reiškia ignoruojamą klaidą arba dar blogiau – rasta klaida traktuojama kaip netikro pavojaus signalas [12]. „FxCop“ nėra atvirojo kodo, o tai sumažina tobulinimo galimybės. „Smokey“ yra atvirojo kodo, tačiau turi tik komandinės eilutės sąsają ir paskutinį kartą buvo atnaujinta tik 2008 m., dėl to nepalaiko naujausių .NET ir „Mono“ versijų.

IV. ĮRANKIO „GENDARME“ TAISYKLIŲ PATOBULINIMAS

Šioje dalyje aprašomi siūlomi esamų įrankio „Gendarme“ taisyklių patobulinimai. Šie patobulinimai turėtų padėti aptikti daugiau klaidų ir padidinti programų spartą, stabilumą, saugumą ir kodo skaitomumą. Prie kiekvienos taisyklės pateikiami netinkamo ir gero programinio kodo pavyzdžiai. Netinkamas kodas – tai kodas, kuris turi saugumo spragą ar tiesiog nėra optimalus variantas ir kurį reikia taisyti. Gero programinio kodo pavyzdys parodo, kaip vieną ar kitą problemą siūloma išspręsti.

A. ReviewLockUsedOnlyForOperationsOnVariablesRule

Ši taisyklė priklauso „Gendarme.Rules.Concurrency“ kategorijai ir tikrina, ar `lock` operatorius yra naudojamas tik operacijoms su klasės ar lokaliais kintamaisiais atlikti. Jei vienintelis tikslas kritinėje dalyje yra užtikrinti, kad kintamasis bus pakeistas automatiškai, tuomet geriau naudoti `System.Threading.Interlocked`.

Netinkamo programinio kodo pavyzdys, kuriame programa neaptikdavo klaidos:

```
lock (m_lockObject)
{
    count++;
    m_someSharedObject = anotherObject;
    Interlocked.Add (pages, 5);
}
```

Gero programinio kodo pavyzdys:

```
Interlocked.Increment (count);
Interlocked.Exchange (m_someSharedObject,
anotherObject);
Interlocked.Add (pages, 5);
```

Taisyklė tikrina visas `lock` viduje esančias operacijas, kol aptinka tokią, kuri nėra automatinė operacija su kintamuoju. Jei tokia operacija neaptinkama visoje zonoje, tuomet registruojama klaida. Patobulinta taisyklė dabar atpažįsta visus `System.Threading.Interlocked` klasės metodus kaip automatinius metodus su kintamaisiais. Taigi algoritmas nebenutraukiamas aptikus `Iterlocked` klasės metodą.

B. *CheckParametersNullityInVisibleMethodsRule*

Taisyklė priklauso „Gendarme.Rules.Correctness“ kategorijai. Ji sako, jog matomų metodų parametrai turėtų būti patikrinti, ar ne `null` prieš juos naudojant.

Netinkamo programinio kodo pavyzdys, kai programa neaptikdavo, jog nėra `null` patikrinimo:

```
public void CheckParametersNullity (string name)
{
    string nameCopy = string.Copy (name);
    if (name.Length > 10)
        Console.WriteLine ("Name: " + name);
}
```

Gero programinio kodo pavyzdys:

```
public void CheckParametersNullity (string name)
{
    string nameCopy = string.Copy (name);
    if (name != null && name.Length > 10)
        Console.WriteLine ("Name: " + name);
}
```

Tačiau ši taisyklė nebeaptikdavo klaidos, jei prieš tai būdavo taikomas bet koks metodas ir jam kaip parametras buvo perduodamas šio metodo parametras. Patobulinta taisyklė šio veiksmo nebetraktuoja kaip `null` reikšmės patikrinimo. Tačiau išlieka dar viena problema: algoritmas tik aptinka, ar yra patikrinimas `param != null` ar `param == null`, tačiau netikrina, kas vyksta toliau. O po to gali eiti toks veiksmas: `if (param == null) param.Update ()`, kurį vykdydama programa savo darbą baigs nenumatytu būdu.

C. *ConsiderAddingInterfaceRule*

Taisyklė priklauso „Gendarme.Rules.Design“ kategorijai. Ji praneša apie klaidą, jei klasė įgyvendina visus sąsajos narius, tačiau nepaveldi pačios sąsajos. Patobulinta taisyklė nebeatsižvelgia į didžiąsias ir mažąsias raides. Ieškant metodo su atitinkamu pavadinimu dabar naudojamas `StringComparison.Ordinal` palyginimo tipas. Taip pat patobulinta taisyklė palyginimuose dabar taiko ir statinius metodus.

Netinkamo programinio kodo pavyzdys, kai programa nesiūlydavo paveldėti sąsajos:

```
public interface ITest
{
    void CreateDGNFile ();
    bool IsECInstance ();
}
public class Test
{
    static void CreateDgnFile () {}
    bool IsEcInstance () { return true; }
}
```

D. *AvoidUnusedPrivateFieldsRule*

Priklauso „Gendarme.Rules.Performance“ kategorijai. Taisyklė skirta nenaudojamiems privatiesiems klasės kintamiesiems aptikti. Patobulinta taisyklė neaptinka klaidos tik tuomet, kai privataus kintamojo reikšmė yra nuskaitoma kurioje nors klasės operacijoje. Taip pat patobulinta taisyklė aptinka ir nenaudojamas privačias konstantas.

Netinkamo programinio kodo pavyzdys, kai programa neaptikdavo klaidos:

```
public class Test
{
    private string m_unused;
    private const string m_unusedConst;
    public void AvoidUnusedPrivateFieldsRule ()
    {
        m_unused = "Unused private variable";
        m_unusedConst = "Unused private const";
        DoSomething ();
    }
}
```

Gero programinio kodo pavyzdys:

```
public class Test
{
    public void AvoidUnusedPrivateFieldsRule ()
    {
        DoSomething ();
    }
}
```

Patobulinta taisyklė tikrina visas klasės operacijas tol, kol aptinka tokią, kurioje yra nuskaitoma kintamojo reikšmė. Anksčiau buvo ieškoma operacijos, kur naudojamas kintamasis, neįvertinant, ar jam reikšmė priskiriama, ar nuskaitoma.

E. *RemoveUnusedLocalVariablesRule*

Taisyklė priklauso „Gendarme.Rules.Performance“ kategorijai. Taisyklė skirta nenaudojamiems metodo kintamiesiems aptikti. Patobulinta taisyklė neaptinka klaidos

tik tuomet, kai metodo kintamojo reikšmė yra nuskaitoma kurioje nors metodo operacijoje.

Netinkamo programinio kodo pavyzdys, kuriame programa neaptikdavo klaidos:

```
public void AvoidUnusedLocalVariablesRule (){
    String unused = "Unused local variable.";
    DoSomething ();
}
```

Gero programinio kodo pavyzdys:

```
public void AvoidUnusedLocalVariablesRule (){
    DoSomething ();
}
```

Taisyklė tikrina visas klasės operacijas tol, kol aptinka tokią, kurioje yra nuskaitoma kintamojo reikšmė. Anksčiau buvo ieškoma operacijos, kur naudojamas kintamasis, neįvertinant, ar jam reikšmė priskiriama, ar nuskaitoma.

F. ArrayFieldsShouldNotBeReadOnlyRule

Taisyklė priklauso „Gendarme.Rules.Security“ kategorijai. Ši taisyklė aptinka masyvus, kurie yra deklaruoti kaip `public readonly` ir praneša apie klaidą, nes tokiu atveju negalima pakeisti paties masyvo, bet galima pakeisti jo reikšmes. Patobulinta taisyklė aptinka ir deklaruotus sąrašus bei rinkinius (angl. *collection*), nes jiems galioja ta pati saugumo spraga.

Pavyzdys, kuriame programa neaptikdavo klaidos:

```
public class Test
{
     public readonly
System.Collections.ObjectModel.Collection<int>
readOnlyCollection = new Collection<int> ();
     public readonly
System.Collections.Generic.ICollection<int>
readOnlyICollection = new List<int> ();
 }
```

Gero programinio kodo pavyzdys:

```
public class Test{
     private readonly
System.Collections.ObjectModel.Collection<int>
ROCollection = new Collection<int> ();
     private readonly
System.Collections.Generic.ICollection<int>
ROICollection = new List<int> ();
    public int[] GetCollectionAsArray ()    {
    int[] array = new int[ROCollection.Count];
    ROCollection.CopyTo (array, 0);
    return array;
    }
    public int[] GetICollectionAsArray ()    {
    int[] array =new int[ROICollection.Count];
    ROICollection.CopyTo (array, 0);
    return array;
    }
}
```

Taisyklė tikrina visų klasių kintamuosius. Pirmiausiai patikrinama, ar kintamąjį galima tik skaityti ir ar jis yra viešas. Po to tikrinamas tipas ir, jei jis yra masyvas, sąrašas ar rinkinys, spausdinamas klaidos pranešimas.

V. IŠVADOS

Nemokamų statinės kodo analizės įrankių naudojimas – vienas pigiausių ir anksčiausiai saugumo spragas padedančių aptikti būdų. Kodo analizės įrankis „Gendarme“ skirtas aptikti klaidoms .NET programose. .NET platformos naudojimas vis labiau auga [13], taigi įrankis „Gendarme“ tampa vis aktualesnis. Programoje „Gendarme“ taisyklės ir aptinkamos klaidos aiškiai atskirtos, įrankis patogiai konfigūruojamas, o analizės rezultatai pateikiami aiškiai.

Esamos įrankio „Gendarme“ taisyklės yra gana išbaigtos, tačiau vis tiek buvo rasta likusių spragų, kurias verta ištaisyti. Įrankiui buvo patobulintos 6 esamos taisyklės. Patobulinimai turėtų padėti aptikti daugiau klaidų ir padidinti programų spartą, stabilumą, saugumą ir kodo skaitomumą.

Tolimesniuose darbuose numatoma pasiūlyti naujų taisyklių. Šios naujos taisyklės turėtų padėti aptikti daugiau perteklinio kodo, kurio neaptinka kompiliatorius, supaprastinti programos kodą ir pagerinti jo skaitomumą, aptikti daugiau kritinių programų klaidų.

LITERATŪRA

[1] D. Baca, K. Petersen, B. Carlsson, L. Lundberg, „Static Code Analysis to Detect Software Security Vulnerabilities – Does Experience Matter?“, 2009 International Conference on Availability, Reliability and Security, March 2009, pp. 804-810.

[2] M. Bradley, F. Cassez, A. Fehnker, T. Given-Wilson, and R. Huuck, „High Performance Static Analysis for Industry“, Tools for Automatic Program Analysis, vol. 289, December 2012, pp. 3-14.

[3] Dr. Paul E. Black, „Static Analyzers in Software Engineering“, CROSSTALK The Journal of Defence Software Engineering. March/April 2009.

[4] P. Anderson, „Measuring the Value of Static-Analysis Tool Deployments“, IEEE Security & Privacy, vol. 10, iss. 3, May/June 2012, pp. 40-47.

[5] A. G. Bardas, „Static Code Analysis“, Journal of Information Systems & Operation Management, vol. 4, no. 2, December 2010.

[6] B. Chess, G. McGraw, „Static Analysis for Security“, Security & Privacy, vol. 2, iss. 6, November/December 2004.

[7] U. Bayer, A. Moser, C. Kruegel, E. Kirda, „Dynamic analysis of malicious code“, Journal in Computer Virology, vol. 2, no. 1, August 2006.

[8] „FxCop“. Microsoft [interaktyvus]. [Žiūrėta 2013-04-06]. Prieiga per internetą: http://msdn.microsoft.com/en-us/library/bb429476%28v=vs.80%29.aspx.

[9] „Smokey“. Google Project Hosting [interaktyvus]. [Žiūrėta 2013-04-06]. Prieiga per internetą: http://code.google.com/p/smokey/.

[10] „Gendarme“. Mono [interaktyvus]. [Žiūrėta 2012-01-27]. Prieiga per internetą: http://www.mono-project.com/Gendarme.

[11] „Standard ECMA-335, Common Language Infrastructure (CLI)“. Ecma International. 6th Edition, June 2012.

[12] A. Bessey, K. Block, B. Chelf, A. Chou, B. Fulton, S. Hallem, C. Henri-Gros, A. Kamsky, S. McPeak, D. Engler. „A Few Billion Lines of Code Later: Using Static Code Analysis to Find Bugs in the Real World“, Communications of the ACM, vol. 53, iss. 2, February 2010, pp. 66-75.

[13] „Microsoft .Net Framework Support“. StatOwl [interaktyvus]. [Žiūrėta 2013-02-23]. Prieiga per internetą: http://www.statowl.com/microsoft_dotnet.php.

Karkaso „Zend Framework“ pagrindu sukurtų internetinių sistemų pritaikymo išmaniesiems įrenginiams technologinio sprendimo sudarymas ir tyrimas

Voldemaras Žitkus
Kauno technologijos universitetas
Informacijos sistemų katedra
Kaunas, Lietuva
voldes@gmail.com

Santrauka. Darbe sprendžiamas esamų internetinių sistemų pritaikymo išmaniesiems įrenginiams uždavinys. Pasiūlytas sprendimas, kuris palengvina internetinių sistemų pritaikymą ir palaikymą mobiliųjų įrenginių platformose. Pateiktame sprendime siūloma automatiškai generuoti dalį kodo struktūros panaudojant egzistuojančios internetinės sistemos kodo bazę. Pasiūlyto sprendimo koncepcija parengta programinio karkaso *„Zend Framework“* ir operacinės sistemos *„Android“* pagrindu. Darbe taip pat aptariamas ir duomenų sinchronizavimo tarp esamos internetinės sistemos ir išmaniųjų įrenginių posistemio aspketas. Pateikto sprendimo prototipas sukurtas, kaip programavimo kalbos aplinkos *„Eclipse“* įskiepio *„Android“* papildinys.

Reikšminiai žodžiai: Internetinė sistema, išmanusis įrenginis, „Android“ operacinė sistema, „Zend Framework“, programinis karkasas.

I. ĮVADAS

Kasdien augantis išmaniųjų įrenginių vartotojų skaičius ženkliai plečia ir siūlomų paslaugų, skirtų šiems įrenginiams skaičių. Šiuo metu ryškėja tendencija, kad kiekvieną naują, internetinę informacinę sistemą lydi ir išmaniesiems įrenginiams skirtas sprendimas. Tokia tendencija lemia poreikį pritaikyti internetines sistemas pritaikymą išmaniųjų įrenginių. Galima būtų teigti, kad naujausių išmaniųjų įrenginių interneto naršyklės gana gerai susidoroja su esamos internetinės sistemos turinio vaizdavimu ir funkcijų palaikymu. Arba galima internetinių sistemų vartotojo sąsają pritaikyti, pasinaudojant vienu iš galimų sprendimų [1] ir taip išspręsti panaudojamumo problemas. Tačiau reikia nepamiršti, kad didžioji dauguma šiuolaikinių išmaniųjų įrenginių pasižymi įvairiomis funkcinėmis savybėmis, tokiomis kaip giroskopinis sensorius, globalios pozicijos nustatymo sistema, liečiamas ekranas, kurios suteikia šiems įrenginiams panaudojamumo pranašumą, palyginti su asmeniniais kompiuteriais. Todėl kalbant apie esamų internetinių sistemų pritaikymą išmaniesiems įrenginiams reikia įvertinti ir išnaudoti vartotojo sąsajos kūrimo bei funkcines galimybes.

Pagrindinės problemos pritaikant internetines sistemas išmaniesiems įrenginiams kyla dėl skirtingų programinių platformų ir, programinio kodo palaikymo sudėtingumo. Šiame darbe pateikiamas sprendimas, kuris padeda automatizuotai išspręsti dalį minėtų problemų. Siekiant iki galo įgyvendinti darbe siūlomą konceptualų sprendimą buvo pasirinkti keli apribojimai: pritaikoma internetinė sistema turi būti sukurta programinio karkaso *„Zend Framework“* pagrindu, o pritaikomi – tie įrenginiai, kurių pagrindas yra operacinė sistema *„Android“*. Sukurto sprendimo eksperimentinis tyrimas buvo atliktas naudojant kolekcionieriams skirtą internetinę informacinę sistemą.

II. IŠMANIŲJŲ ĮRENGINIŲ IR JŲ PROGRAMINIŲ PLAFORMŲ IŠSKIRTINUMAS PALYGINTI SU ASMENINIAIS KOMPIUTERIAIS.

Lyginti išmaniuosius įrenginius su asmeniniais kompiuteriais galima dviem aspektais – pagal techninę charakteristiką ir ergonomiką.

Išmanieji įrenginiai techniškai tobulėja kiekvieną dieną, tad tampa vis patrauklesni vartotojams beto, vyksta ir pačių įrenginių inovacijų progresas. Bet vis dar galima teigti, jog vidutinis išmanusis įrenginys gerokai nusileidžia pajėgumu vidutinam stacionariam kompiuteriui. Šiuo metu asmeniniai kompiuteriai lengvai apdoroja internetinių informacinių sistemų pateikiamą turinį, o nemažai rinkoje esančių išmaniųjų įrenginių negali nepriekaištingai susidoroti su pažangių internetinių sistemų pateikiamu turiniu. Todėl esamas internetines informacines sistemas aktualu pritaikyti ir optimizuoti išmaniesiems įrenginiams, sukuriant taikomąsias programėles, ir taip siekiant išlaikyti sistemos patrauklumą bei išlikti tarp išmaniųjų įrenginių naudotojų.

Operacinę sistemą *„Android“* naudojantiems įrenginiams informacinės sistemos yra kuriamos programavimo kalba *„Java“*. Taip pat naudojamas programinis karkasas, pritaikytas kurti sistemas *„Android“*. Vienas iš pagrindinių operacinės sistemos *„Android“* išskirtinumų yra tai, jog komponentai ir įvairios sistemos savybės turi būti aprašytos faile *„androidManifest.xml“*, tai savotiškas turinys, kuriuo naudojasi operacinė sistema, valdydama sukurtą informacinę sistemą. Šis turinys užtikrina aplinkoje *„Android“* veikiančių sistemų komunikaciją ir suteikia galimybę pasinaudoti kitų platformos *„Android“* sistemų funkcionalumu [2], [3].

Mažas išmaniųjų įrenginių ekranas, palyginti su stacionarių kompiuterių monitoriais, yra vienas didžiausių skirtumų, vertinant ergonomikos aspektu. Dėl šios priežasties būtina optimizuoti informacijos kiekį, kuris pateikiamas išmaniajame įrenginyje – pateikti tik būtiniausią informaciją.

Taip pat svarbus aspektas yra naudojimo aplinka: išmaniaisiais įrenginiais vartotojai dažnai naudojasi eidami gatve, važiuodami viešuoju transportu ar kitoje, tokiai veiklai neįprastoje, aplinkoje. Atsižvelgiant į šiuos aspektus, sistema turi būti nesudėtinga, nereikalauti iš vartotojo didelio navigacinio tikslumo ir patogi naudotis viena ranka. Žinoma, kai sistema skirta naudoti namų aplinkoje (ar kitoje ramesnėje aplinkoje), anksčiau paminėti aspektai gali pasirodyti ne tokie aktualūs, bet tai – tik prielaidos, kurias reikia vertinti atsargiai.

Vienas iš pagrindinių išmaniųjų įrenginių privalumų yra jutiklinis ekranas, tačiau kartu tai – ir trūkumas, apsunkinantis vartotojo sąsajos kūrimą. Kadangi dauguma veiksmų vartotojas atlieka liesdamas ekraną, būtina sukurti vartotojo sąsają, pritaikytą būtent šiam techniniam sprendimui [4]. Internete yra daug oficialios medžiagos [5], kurioje pateikiama aibė rekomendacijų, kaip panaudoti įvairius sąsajos elementus (mygtukų dydis, tarpai tarp nuorodų ir t. t.) ir derinti su sistemos funkcionalumu.

Dėl šių ypatumų, nėra teisinga tiesiogiai perkelti egzistuojančios internetinės informacinės sistemos vartotojo sąsają į išmaniajam įrenginiui skirtą sprendimą. Galima perkelti dalį sąsajos elementų, o po to atlikti papildomas korekcijas, kad būtų priimtas optimalus sprendimas išmaniųjų įrenginių atžvilgiu.

III. Esamų internetių sistemų ir išmaniųjų įrenginių sąveika

Priklausomai nuo sistemos tipo bei paskirties, gali iškilti sąveikos tarp esamos internetinės informacinės sistemos ir išmaniųjų įrenginių platformoms pritaikytų sistemos versijų poreikis. Galima išskirti keletą pagrindinių sąveikų tipų.

A. *Duomenys labiau skaitomi, negu rašomi ar atnaujinami*

Naujajai platformai pritaikyta sistema daugiausiai naudojama informacijos skaitymui ir kito originaliosios sistemos turinio peržiūrai. Ši sąveika yra pati paprasčiausia, nes į pritaikytą išmaniojo įrenginio sistemą pakanka perduoti duomenis iš internetinės sistemos, naudojant *„web“* paslaugas ar kitas panašias technologijas.

B. *Dažnai naudojami tam tikros srities duomenys*

Priklausomai nuo sistemos paskirties, poreikis naudoti dalį duomenų išmaniojo įrenginio aplinkoje gali būti daug dažnesnis, pavyzdžiui, naudoti tam tikrą vartotojo registracijos informaciją. Siekiant paspartinti sistemų sąveikos našumą gali būti tikslinga, priklausomai nuo sistemos specifikos, dalį tokių dažnai naudojamų duomenų saugoti išmaniajame įrenginyje.

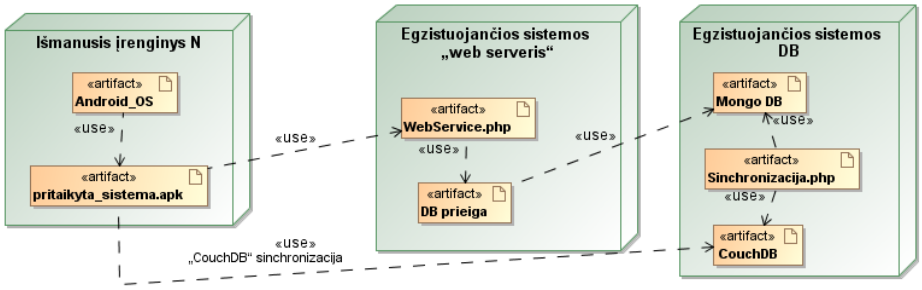
Tokius duomenys galima saugoti arba lokalioje duomenų bazėje, arba laikinojoje atmintyje, priklausomai nuo jų apimties ir panaudojimo dažnumo.

C. *Duomenys naudojami asmeniniams tikslams, ir vartotojai yra pagrindiniai sistemos turinio šaltiniai*

Šiuo atveju didžioji dalis informacinių sistemų turinio yra generuojama pačių vartotojų. Vartotojams gali būti aktualu pasiekti savo duomenis, kai išmaniajam įrenginiui nepasiekiamas interneto ryšys arba dėl tam tikrų priežasčių originaliosios sistemos serveriai nėra pasiekiami. Tokiu atveju reikalinga, jog pritaikyta sistema galėtų veikti nepriklausomai nuo originaliosios internetinės informacinės sistemos. Veikimui užtikrinti būtina, jog lokali sistema turėtų savo duomenų bazę su reikiamų duomenų kopijomis. Siekiant išvengti duomenų neatitikimų tarp internetinės ir išmaniojo įrenginio sistemos, reikalinga atlikti duomenų sinchronizaciją, kai tik internetinės sistemos duomenų bazių valdymo sistemos (DBVS) serveriai tampa pasiekiami.

Duomenims sinchronizuoti yra siūloma įvairių sprendimų [6], [7], [8], tačiau dauguma jų veikia korektiškai arba tik su identiškomis, arba to paties tipo DBVS (pavyzdžiui, sunku sinchronizuoti *„SQL“* ir *„NoSQL“* duomenų bazes). Taip pat reiktų įvertinti, kad išmaniųjų įrenginių platformos ne visas duomenų bazes, todėl priklausomai nuo sistemos specifikos gali reikėti atlikti skirtingų DBVS sinchronizaciją. Detaliau duomenų sinchronizavimo uždavinio sprendimas šiame darbe nenagrinėjamas.

Šiame tyrime buvo sinchronizuojami skirtingų duomenų bazių valdymo sistemų duomenys. Naudotas skirtingų DBVS duomenų sinchronizacijos architektūros modelis pateiktas 1 pav.



1 pav. Sistemų, turinčių skirtingas duomenų bazes, sąveika

Darbe eksperimentiniams tyrimams buvo naudojama kolekcinionierių internetinė sistema, sukurta panaudojant programinį karkasą *„Zend Framework“* ir duomenų bazių valdymo sistemą *„Mongo DB“*. Sistema buvo pritaikytą platformai *„Android“* naudojant duomenų bazę *„CouchDB“*. Ši duomenų bazė turi gerai parengtą atskirose platformose esančių sistemų sinchronizavimo mechanizmą. Kadangi abi duomenų bazės yra *„NoSQL“* tipo, jas galima sinchronizuoti atsižvelgiant į jų struktūrą.

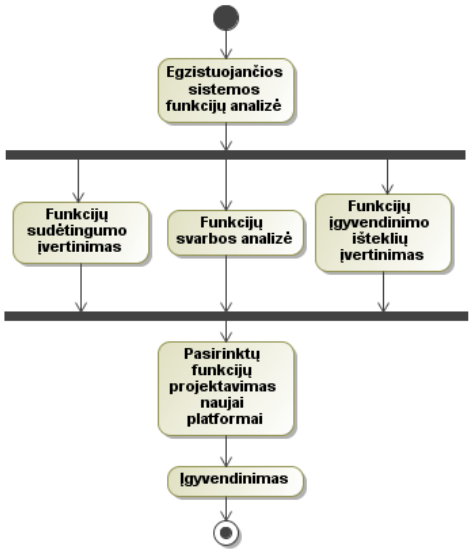
Bendro sprendimo, tinkančio visais atvejais nėra, tačiau pateiktą sinchronizavimo modulio architektūros principą su smulkiais pakeitimais galima pritaikyti įvairiems DBVS deriniams.

IV. „Zend Framework“ pagrindu sukurtos internetinės IS pritaikymo išmaniesiems įrenginiams su OS „Android“ koncepcija

Kuriant koncepciją buvo įvertinta ir jau esamų sprendimų patirtis bei logika [9], kurie buvo taikomi ir senesnės kartos mobiliesiems sprendimams [10]. Perkeliant sistemą į naują platformą būtina įvertinti, kurias funkcijas tikslinga pritaikyti naujajai sistemai, taip pat atsižvelgti į naujos programų kūrimo aplinkos programinio projekto struktūrą.

A. *Reikiamo funkcionalumo atrinkimas*

Pritaikant esamą sistemą išmaniesiems įrenginiams, reikėtų pirmiausia pasinaudoti sistemos turimomis reikalavimų ir projektinėmis specifikacijomis. Pagal šias specifikacijas paprasta nustatyti, kuris funkcionalumas turi būti perkeliamas į išmaniajam įrenginiui kuriamą taikomąją programą. 2 pav. pateikiamas koncepcinis sistemos reinžinerijos procesas, kuriame detalizuojamas perkeliamų į naują platformą funkcijų parinkimas.



2 pav. Perkeliamo funkcionalumo atrinkimas

Pirmiausia analizuojamas egzistuojančios sistemos funkcionalumas. Atsižvelgiama į tris pagrindinius aspektus:

- Reikia įvertinti kiekvieną sistemos funkcinį panaudojimo atvejį, ar šis funkcionalumas reikalingas išmaniajame įrenginyje. Nebūtino funkcionalumo pavyzdys gali būti sistemos administratoriaus funkcijos.
- Svarbu atsižvelgti į turimus, sistemai kurti skirtus, išteklius. Reikia atskirti funkcionalumą, kuris nėra būtinas naujajai sistemai, kuram įgyvendinti gali būti panaudota daug sistemai kurti skirtų išteklių.
- Taip pat svarbu nustatyti, ar pritaikomas labai sudėtingas funkcionalumas. Šis klausimas vertinamas tiek techniškai, tiek iš vartotojo pozicijos. Realizuoti techniškai sudėtingą funkcionalumą gali užtrukti daugiau laiko, negu planuojama, o įgyvendintas sudėtingas funkcionalumas vartotojui gali pasirodyti per daug sudėtingas ir nepatogus, todėl vartotojas jo nenaudos.

Remiantis šiais vertinimais sudaromas funkcijų prioritetinis sąrašas. Iš sąrašo empiriškai pašalinamos jo gale esančios funkcijos.

Po šios reikalavimų analizės turime pradinį sąrašą funkcijų reikalingų naujojoje sistemoje.

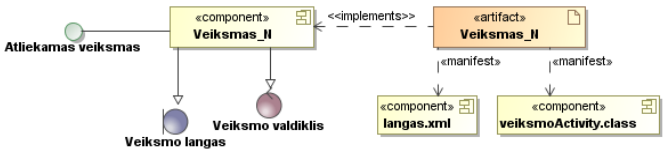
B. *Programinio kodo struktūros perkėlimas*

Žinant perkeliamų funkcijų aibę, galima atlikti programinio kodo analizę. Reikia pažymėti, jog kai kurie programinio karkaso *„Android“* komponentai turi atitikmenis kituose *„Model-view-controller“ (MVC)* tipo karkasuose. Komponentai *„Activity“* veikia panašiai kaip valdikliai, pakartotinio panaudojimo komponentai veikia kaip komponentai *„Fragments“*, o vaizdai aprašomi *„xml“* failuose. Šių komponentų įdiegimo klasių modelis matyti 3 pav.

Atsižvelgiant į šiuos panašumus galima automatizuotai sukurti išmaniojo įrenginio taikomosios programos išeities programinio kodo projektą, kurį sudarytų dalis reikalingų failų, kuriuose saugomos programinės klasės ir kiti programinio kodo epizodai. Turint jau dalinę programinio projekto struktūrą lengviau toliau kurti sistemą.

Išlaikant panašią programinę struktūrą tarp skirtingų sistemos versijų lengviau jas palaikyti.

Kadangi kiekvienas programavimo karkasas turi gana aiškiai apibrėžtą sistemos struktūrą, kuri parodo, kur ir kokio tipo komponentai saugomi, šį procesą galima iš dalies automatizuoti ir automatiškai sukurti reikiamus komponentus bei užpildyti jais failą *„androidManifest.xml“*.



3 pav. Komponentų įdiegimo klasių modelis

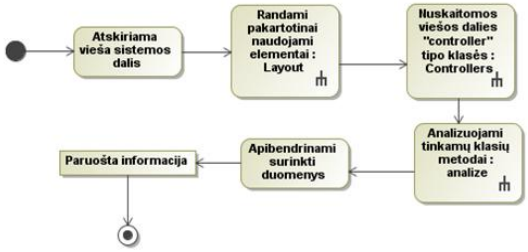
Detaliam koncepcijos įgyvendinimui buvo pasirinktas programinio karkaso *„Zend Framework“* pagrindu sukurtų internetinių informacinių sistemų pritaikymo platformai *„Android“* uždavinys. Analizuojant *„Zend Framework“* pagrindu kuriamų ir platformai *„Android“* skirtų informacinių sistemų programinio kodo projektų struktūrą buvo išskirti šie struktūrų sąsajų principai:

- *MVC* struktūra leidžia susieti kiekvieną veiksmą (puslapis bei jo logika) su jam priklausančia *„view“* klase (vaizdavimas). Tiesiogiai perkeliant šį ryšį į programinį karkasą *„Android“* galima iš karto sugeneruoti *„activity“* klases ir jų *„xml“* failus, kurie turi tą pačią paskirtį – puslapio logika ir jo vaizdavimas.
- Formoms kurti paprastai naudojama *„Zend_Form“* klasė, struktūrizuotai aprašanti formos elementus (įvedimo laukai, jų pavadinimai, mygtukai). Nustačius, jog atvaizdavimo klasėje (*„view“*) naudojama *„Zend_Form“* klasė, galima surinkti informaciją apie formos elementus ir sukurti atitinkamus elementus aplinkoje *„Android“*.
- Programinis karkasas *„Android“* kaip numatyta neturi *MVC „model“* klasių atitikmens, bet šiuo atveju šiame programiniame karkase nedraudžiama sukurti *„model“* tipo klasių ir jas naudoti.
- Internetinėse informacinėse sistemose logika dažnai atskyriama į viešą dalį (tai, ką mato eilinis vartotojas) ir administracinę dalį. Pritaikant sistemą išmaniesiems įrenginiams netikslinga perkelti administracinę dalį, todėl ši sistemos dalis dažniausiai turėtų būti neliečiama.
- Analizuojant *„Zend Framework“* pagrindu sukurtas sistemas, galima identifikuoti dažnai naudojamas klases (pavyzdžiui, šoninį meniu, kuris pateikiamas visuose puslapiuose). Atpažinimo požymis gali būti jog šios klasės yra įterpiamos (*„include“*) kitose klasėse. Jas atpažinus karkase *„Android“* galima sukurti *„fragment“* tipo klases, kurios taip pat skirtos naudoti pakartotinai.
- Sugeneruotų klasių informaciją galima automatiškai surašyti į failą *„androidManifest.xml“*.

Esamos internetinės informacinės sistemos duomenų surinkimo ir perkėlimo veiklos modelis, kuris grindžiamas anksčiau aprašytais principais, pateikiamas 4 pav.

Pirmiausia atskiriama vieša sistemos dalis nuo administracinės ir kitų sistemos dalių. Kadangi karkasas *„Zend Framework“*, kaip numatyta, neatskiria šių sistemos dalių, reikia papildomos informacijos norint nustatyti, kurias klases reikia perkelti, o kurių – ne. Šiuo atveju galima pasinaudoti egzistuojančios sistemos projektine specifikacija.

Pakartotinai naudojamus elementus galima rasti *„layout“* klasėse ieškant komponentų, kurie yra į juos įtraukiami.



4 pav. Esamos internetinės informacinės sistemos duomenų surinkimo ir perkėlimo veiklos modelis

Sistemos veiklos logiką aprašo *„controller“* klasės. Čia vartojama sąvoka „tinkama klasė“ – tai tokia klasė, kuri yra skirta sistemos viešajai daliai ir turi bent vieną viešą (*„public“*) metodą (*„controller“* klasė, neturinti nė vieno viešo metodo, neturi prasmės, bandant į ją kreiptis bus grąžinama klaida). Neprasminga perkelti tokias klases reinžinerijos metu.

Kitame žingsnyje analizuojamos surinktos *„controller“* klasės. Pirmiausia cikliniu būdu analizuojamas surinktų, perkelti tinkamų klasių rinkinys, kol pasiekiamas sąrašo galas. Po atliktos analizės gaunamas duomenų rinkinys pagal kurį bus kuriama sistemos struktūra aplinkoje *„Android“*.

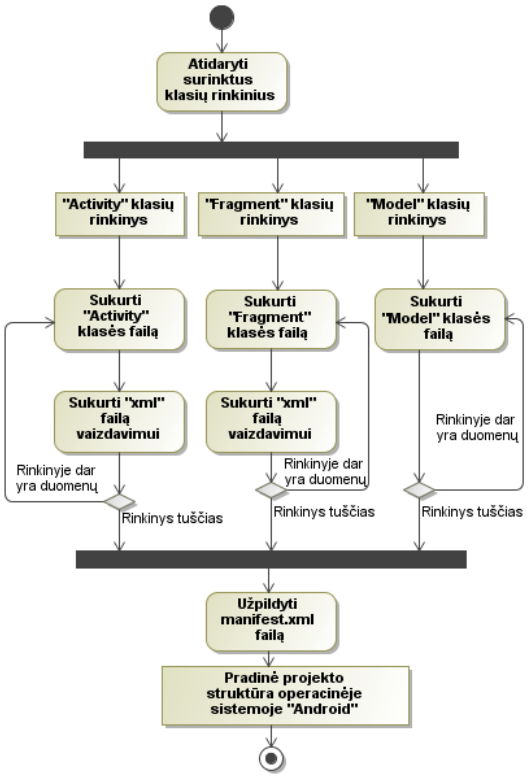
Ciklo metu yra analizuojami visi vieši *„controller“* klasės metodai – kiekvienas iš jų atitinka tam tikrą sistemos puslapį, kuris turi vaizdavimo būdą ir savo veiklos logiką.

Kiekvienas iš šių metodų gali kreiptis į modelio klasę (pavyzdžiui, reikia gauti duomenų iš duomenų klasės). Radus kreipinį, modelio klasės informacija yra įrašoma į *„models“* klasių rinkinį.

Kiekvienam *„controller“* klasės viešam metodui būtina turėti *„view“* klasę (kitokiu atveju bandant kreiptis į metodą bus grąžinama klaida), jeigu jame panaudota *„Zend_Form“* klasė (pavyzdžiui, kontaktų forma), tada galima nuskaityti elementus, esančius formoje, ir juos sukurti aplinkoje *„Android“*.

Atlikus šiuos veiksmus gaunamas duomenų rinkinys naudojamas paskutiniame šio reinžinerijos proceso žingsnyje.

Surinkus visą reikalingą informaciją iš egzistuojančios sistemos struktūros yra sukuriama nauja struktūra aplinkoje *„Android“*. Struktūros kūrimo modelis pateikiamas 5 pav. Lygiagrečiai kiekvienam duomenų rinkiniui yra kuriamos atitinkamos klasės ir *„xml“* failai aplinkoje *„Android“*.



5 pav. Naujos struktūros kūrimo veiklos modelis

Sukūrus visus failus, jų informacija surašoma į failą *„androidManifest.xml“*, kuriame saugoma pagrindinė informacija apie sistemos struktūrą.

Atlikus visus šiuos veiksmus gaunamas pradinis sistemos programinio kodo projektas aplinkoje *„Android“*, kurį toliau programuotojas papildo specifine veiklos logika, ir drauge suteikia išmaniesiems įrenginiams būdingų funkcinių elementų ir sprendimų.

V. Koncepcijos įgyvendinimas Aplinkoje „Eclipse“

Anksčiau aprašytai koncepcijai įgyvendinti sukurtas programos prototipas, kuris veikia kartu su programavimo aplinkos *„Eclipse“* įskiepiu, skirtu kurti aplinkos *„Android“* taikomosioms programoms. Pirmiausia naudojant įrankį *„Eclipse“* yra sukuriamas naujas *Android* projektas su numatyta failų struktūra ir konfigūracija, o antrojo žingsnio metu paleidžiama sukurta programa nurodant pritaikomos internetinės sistemos programinio kodo adresą ir tikslinį *Android* projektą, kuris užpildomas perkeliama informacija iš pasirinktos sistemos. Galiausiai naujajam projektui sugeneruojamos pagrindinės programinės klasės, dalis vartotojo sąsajos elementų ir užpildomas failas *„androidManifest.xml“*. Pastebėta, jog sukurtas ir eksperimentiškai išmėgintas kodo generavimas, paspartina esamų internetinių sistemų pritaikymo išmaniesiems įrenginiams procesą.

Išvados

Buvo atlikta išmaniųjų įrenginių funkcinių galimybių analizė, kuri parodė, kad identiškas funkcijų perkėlimas iš įprastinės internetinės sistemos į išmaniajam įrenginiui skirtą aplinką gali neišnaudoti išmaniesiems įrenginiams būdingų vartotojo sąsajos funkcinių elementų bei perkrauti pačią sistemą nereikalingomis funkcijomis.

Darbe buvo aptarti internetinės sistemos ir jos pagrindu sukurtos išmaniojo įrenginio informacinės sistemos sąveikos modeliai bei pateiktas sistemos architektūros modelis, kuris padeda spręsti duomenų sinchronizavimo problemas.

Darbe buvo pateiktas karkaso *„Zend Framework“* pagrindu sukurtos internetinės informacinės sistemos pritaikymo išmaniajam įrenginiui koncepcinis sprendimas, kuris padaro programuotojo darbą našesnį ir atliekamo sistemos pritaikymo proceso rezultatą kokybiškesni.

Pateikto koncepcinio sprendimo pagrindu buvo sukurtas programinis prototipas, kuris išbandytas su kolekcionierių internetine informacine sistema. Eksperimentinis tyrimas parodė, kad adaptavimo procesas gali būti automatizuotas ir gaunamas rezultatas yra korektiškas.

Tolimesniuose tyrimuose šį koncepcinį sprendimą norima integruoti į projektavimo įrankį *„MagicDraw“* siekiant automatizuoti perkeliamų funkcijų pasirinkimo ir perėjimo prie programinio kodo generavimo žingsnį.

Taip pat atsižvelgiant į įvairių *„PHP“* programavimo karkasų panašumus bus siekiama išplėsti sprendimą, kad jis apimtų daugiau negu vieną (*„Zend Framework“*) programinį karkasą.

Literatūra

[1] M. Hinz, Z. Fiala, F. Wehner, „Personalization-Based Optimization of Web Interfaces for Mobile Devices“, Mobile Human-Computer Interaction- MobileHCI 2004, pp. 204-215.

[2] J. Stark, Building Apps with HTML, CSS, and JavaScript. 1-oji laida. JAV, 2010.

[3] L. Vogel, Android Development Tutorial, 2011,[interaktyvus]. Prieiga per internetą: <http://www.vogella.de/articles/Android/article.html>.

[4] M. Song, H. Song, X. Fu, Methodology of User Interfaces Design Based on Android, Multimedia Technology (ICMT), 2011 International Conference, 2011 (July), pp. 408-411.

[5] Android API Guides [interaktyvus]. Prieiga per internetą: <http://developer.android.com/develop/index.html>.

[6] M.-Y. Choi, E.-A. Cho, D.-H. Park, C.-J. Moon, D.-K. Baik, A Database Synchronization Algorithm for Mobile Devices, Knowledge Creation Diffusion Utilization, 2010, vol. 56, pp. 392-398.

[7] S. Gansemer, U. Gröner, M. Maus, „Database Classification of Mobile Devices“, IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems Technology and Applications, 2007 (September), pp. 699-703.

[8] I. McDowall, Programming PC Connectivity Applications for Symbian OS. 1-oji laida. JAV, 2005.

[9] W. Höpken, M. Fuchs, M. Zanker, T. Beer, “Context-Based Adaptation of Mobile Applications in Tourism”, Information Technology & Tourism, 2010, vol. 12, no. 2, pp. 175-195.

[10] P. Cotter, B. Smyth, Content Personalisation for WAP-Enabled Devices, Adaptive Hypermedia and Adaptive Web-Based Systems, Lecture Notes in Computer Science, 2000, vol. 1892, pp. 98-108.

INFORMACINĖS TECHNOLOGIJOS. XVIII tarpuniversitetinės tarptautinės magistrantų ir doktorantų konferencijos „Informacinė visuomenė ir universitetinės studijos“ (IVUS 2013) medžiaga, 2013 m. balandžio 25 d., Kaunas, Lietuva / INFORMATION TECHNOLOGY. Proceedings of the XVIII International Master and PhD Conference “Information Society and University Studies” (IVUS 2013), 25th April, 2013, Kaunas, Lithuania. – Kaunas: Vytauto Didžiojo universiteto leidykla, 2013. – 210 p., iliustr.

ISSN 2029-249X
ISSN 2029-4824 (elektroninis leidinys)

INFORMACINĖS TECHNOLOGIJOS

XVIII tarpuniversitetinės tarptautinės magistrantų ir doktorantų konferencijos „Informacinė visuomenė ir universitetinės studijos“ (IVUS 2013) medžiaga, 2013 m. balandžio 25 d., Kaunas, Lietuva

INFORMATION TECHNOLOGY

Proceedings of the XVIII International Master and PhD Students Conference “Information Society and University Studies” (IVUS 2013), 25th April 2013, Kaunas, Lithuania

Editors Audrius Varoneckas and Aušra Mackutė-Varoneckienė
Lithuanian language editors Robertas Keturakis and Simona Grušaitė
Graphic designer Rasa Švobaitė
Approved for printing 2013 04 12. A run of 50 copies. (No of P. 210)
Order No. K13-031.
Published: Vytautas Magnus University Publishing Office, S. Daukanto g. 27, LT-44249 Kaunas. SL 1557

Redaktoriai Audrius Varoneckas ir Aušra Mackutė-Varoneckienė
Lietuvių kalbos redaktoriai Robertas Keturakis ir Simona Grušaitė
Leidinio dailininkė Rasa Švobaitė
Pasirašyta spausdinti 2013 04 12. Tiražas 50 egz. 13,25 leidyb. apsk. l. (210 p.)
Užsakymo Nr. K13-031
Išleido: Vytauto Didžiojo universiteto leidykla, S. Daukanto g. 27, LT-44249 Kaunas. SL 1557