

RĪGAS TEHNISKĀ UNIVERSITĀTE
Datorzinātnes un informācijas tehnoloģijas fakultāte
Informācijas tehnoloģijas institūts

Jurijs ČIŽOVS

Vadības informācijas tehnoloģijas doktora programmas doktorants

**DINAMISKĀS SISTĒMAS VADĀMĀ
MARKOVA LĒMUMU MODEĻA UZBŪVE UN
IZPĒTE AR DATU IEGUVES METODĒM**

Promocijas darba kopsavilkums

Zinātniskais vadītājs
Dr.habil.sc.comp., profesors
A. BORISOVS

Rīga 2012

UDK 519.857(043.2)

Či 958 d

Čižovs J. Dinamiskās sistēmas vadāmā Markova lēmumu modeļa uzbūve un izpēte ar datu ieguves metodēm. Promocijas darba kopsavilkums.-R.:RTU, 2012.-36 lpp.

Iespiests saskaņā ar 2012.gada 10.janvāra ITI padomes sēdes lēmumu, protokola Nr. 12-02.



Šis darbs izstrādāts ar Eiropas sociālā fonda atbalstu projektā “Atbalsts RTU doktora studiju īstenošanai”.

ISBN 978-9934-10-271-4

**PROMOCIJAS DARBS
IZVIRZĪTS INŽENIERZINĀTŅU DOKTORA GRĀDA IEGŪŠANAI
RĪGAS TEHNISKAJĀ UNIVERSITĀTĒ**

Promocijas darbs inženierzinātņu doktora grāda iegūšanai tiek publiski aizstāvēts 2012.gada 12. martā plkst. 14:30 Rīgas Tehniskās universitātes Datorzinātnes un informācijas tehnoloģijas fakultātē, Meža ielā 1, 3. korpusā, 202. auditorijā.

OFICIĀLIE RECENZENTI

Profesors, Dr.math. Kārlis Šadurskis
Rīgas Tehniskā universitāte

Profesors, Dr.habil.sc.ing. Jevgeņijs Kopitovs
Transporta un sakaru institūts, Latvija

Profesors, Dr. rer. nat. habil. Juri Tolujew
Magdeburga Otto-von-Guericke universitāte, Vācija.

APSTIPRINĀJUMS

Apstiprinu, ka esmu izstrādājis doto promocijas darbu, kas iesniegts izskatīšanai Rīgas Tehniskajā universitātē inženierzinātņu doktora grāda iegūšanai. Promocijas darbs nav iesniegts nevienā citā universitātē zinātniskā grāda iegūšanai.

Jurijs Čižovs
paraksts

Datums:

Promocijas darbs ir uzrakstīts latviešu valodā, satur ievadu, 6 nodaļas, secinājumus, literatūras sarakstu, 4 pielikumus, 15 tabulas, 70 zīmējumus un ilustrācijas, kopā 137 lappuses. Literatūras sarakstā ir 83 nosaukumi.

VISPĀRĒJS DARBA RAKSTUROJUMS

Ievads

Attīstoties ražošanas, tirdzniecības, finanšu utt. vadīšanas elektroniskajiem līdzekļiem, rodas jaunas iespējas uzglabāt uzņēmuma saimniecisko un ekonomisko darbību atspoguļojošus datus. Viens no mehānismiem, kas nosaka efektīvu uzņēmuma vadīšanu, ir pieejamo datu analīze par uzņēmuma darbību pagātnē, ar mērķi izstrādāt aktuālos vadīšanas risinājumus. Tā kā uzņēmuma darbības novērošana tiek īstenota laikā, dati ir izteikti ar daudzdimensiju laikrindām. Tādējādi rodas lēmumu pieņemšanas uzdevums nenoteiktības apstākļos ar datiem, kas ir daudzdimensiju laikrindas. Nenoteiktību nosaka tas, ka nav tehniski iespējams atspoguļot visus iekšējos un ārējos faktorus, kas ietekmē novērojamos parametrus.

Tēmas aktualitāte

Markova lēmumu pieņemšanas process (MLPP) tiek veiksmīgi pielietots optimālās vadīšanas stratēģijas atrašanas uzdevumos stohastiskajos diskrētajos procesos, kas attīstās laikā. Pastāv virkne tā modifikāciju un paplašinājumu, kuru mērķis ir atrisināt uzdevumus ar nepārtrauktu parametru skaitu, daļēji novērojamām vidēm, u.c. Tomēr, jautājumi, kas saistīti ar MLPP modeļa uzbūvi uzdevumos, kuros dati izteikti laikrindu veidā, paliek atvērti pētniecībai. Modeļa uzbūves sarežģītību nosaka MLPP aparāta prasības pētāmo datu struktūrai. Novērojamo parametru noteikta veida laikrindu realizācijas jāiegūst no relāciju datiem un jāpārveido atbilstoši MLPP struktūrai.

Standarta pieejas paplašināšana darbam ar laikrindām ļauj izmantot visas standarta MLPP modeļa priekšrocības lēmumu pieņemšanai ekonomiskajos uzdevumos reālā laika režīmā.

Darba mērķis

Promocijas darba mērķis ir izstrādāt lēmumu pieņemšanas pieeju, kas balstīta uz Markova lēmumu pieņemšanas procesu, dinamiskām sistēmām, kuru dati izteikti ar laikrindām. Izvirzītā mērķa sasniegšanai ir nepieciešams atrisināt zemāk minētus uzdevumus.

1. Izpildīt mūsdienīga stāvokļa pārskatu MLPP pielietošanas jomā uzdevumiem, kas izteikti daudzdimensiju laikrindās, izpētīt pastāvošās pieejas.
2. Izstrādāt uz datu iegūvi balstītu pieeju laikrindu konstruēšanai, to apstrādei un pārveidošanai struktūrās, kas atbilst MLPP prasībām.
3. Jaunajam aparātam izstrādāt programmnodrošinājumu, kas ir balstīts uz aģentu orientētu arhitektūru un ir intelektuāla atbalsta un lēmumu pieņemšanas sistēma.
4. Izpētīt iegūtās metodes tālākas attīstības iespēju, veicot lēmumu telpas aproksimāciju ar mākslīgo neironu tīklu līdzekļiem.
5. Izpildīt dinamiskās cenu politikas uzdevuma nostādni kā dinamiskās programmēšanas uzdevumu (MLPP kontekstā), lai iegūtu objektu praktiskiem eksperimentiem.
6. Lai novērtētu izstrādātā modeļa efektivitāti reālās pasaules uzdevumā, iegūto intelektuālo aģentu lēmumu pieņemšanas sistēmu aprobēt dinamiskās cenu politikas uzdevumā.

Pētījuma priekšmets

Pētījuma priekšmets ir paplašināta lēmumu pieņemšanas pieeja, kas balstīta uz Markova lēmumu pieņemšanas procesu. Pētāmā modeļa darbības joma ir dinamiskās programmēšanas uzdevumi, kuru dati izteikti ar laikrindām.

Pētījuma hipotēzes

Pētījuma gaitā izvirzītas šādas hipotēzes:

1. Uzdevums, kura dati izteikti ar laikrindām, var tikt uzskatīts par dinamiskās programmēšanas uzdevumu, kā arī izteikts Markova lēmumu pieņemšanas procesa terminos.
2. Markova procesa stāvokļu un lēmumu telpas aproksimācija var tikt īstenota, izmantojot mākslīgo neironu tīkla aparātu.

Pētījumu metodes

Promocijas darba centrālā pētāmā metode ir viens no dinamiskās programmēšanas aparātiem - Markova lēmumu pieņemšanas process. Varbūtības modeļa uzbūvei tiek izmantota nezināmo parametru novērtēšanas statistiskā metode – maksimālās ticamības metode. Tāpat tiek pielietotas datu ieguves metodes, tajā skaitā datu normalizācijas, klasterizācijas un klasifikācijas līdzekļi. Tiek izmantotas skaitļošanas intelekta jomas metodes: apmācība ar pastiprināšanu, mākslīgie neironu tīkli. Izstrādājamajām programmu sistēmām tiek pielietota aģenta orientēta arhitektūra.

Zinātniskā novitāte

Zinātnisku interesi izraisa izstrādātā lēmumu pieņemšanas pieeja, kas balstīta uz Markova lēmumu pieņemšanas procesu. Galvenā pazīme, ar ko tā atšķiras no standarta MLPP, ir pielietojuma iespējamība uzdevumos ar daudzdimensiju laikrindām.

Dinamiskās cenu politikas uzdevumā (ko raksturo daudzdimensiju laikrindas) nodemonstrēta risinājumu telpas aproksimācijas pieeja, izmantojot mākslīgos neironu tīklus.

Piedāvāta arī jauna pieeja aģentu sistēmas arhitektūras konstruēšanā. Tas ļauj izvairīties no konflikta aģenta mērķa un vides noteikšanā, ja aģentam liegta tieša mijiedarbība ar risināmā uzdevuma objektu.

Darba praktiskais pielietojums un aprobācija

Izstrādātā lēmumu pieņemšanas pieeja, kas balstīta uz Markova lēmumu pieņemšanas procesu, paredzēta uzdevumiem, kuros sistēmas stāvoklis aprakstīts nevis ar statistiskiem parametriem, bet ar parametriem, kas mainās laikā.

Iegūtās intelektuālās aģentu sistēmas praktiskais pielietojums nodemonstrēts dinamiskās cenu politikas uzdevumā. Par sākotnējiem datiem ir izmantoti reālās ražošanas un tirdzniecības vadības programmas „1C:Uzņēmums v7” ieraksti par pārdošanām. Datu periods aptver ražojamās pārtikas produkcijas divu gadu pārdošanas.

Uz MLPP balstīta intelektuālo aģentu programmas sistēma izstrādāta vidē „1C:Uzņēmums v7”. Tas ļāva tieši piekļūt datiem par pārdošanām.

Darbā parādīta atsevišķu eksperimentu sērija ar apakšsistēmām (mākslīgo neironu tīkla, Markova lēmumu pieņemšanas procesa), izmantojot modeļa uzdevumus. Veikta arī eksperimentu virkne, kā piemēru izmantojot dinamiskās cenu politikas uzdevumu, ar mērķi skaitliski novērtēt uzlabotā MLPP pieejas efektivitāti. Atsevišķi darba posmi un tās rezultāti ir prezentēti zinātniskajās konferencēs:

1. Chizhov Y. An agent-based approach to the Dynamic Price Problem, 5th International KES Symposium Agent and Multi-agent Systems, Agent-Based Optimization KES-AMSTA/ABO'2011, June 29-July 1, 2011, Manchester, United Kingdom.
2. Chizhov Y., Kuleshova G., Borisov a. Manufacturer – Wholesaler System Study Based on Markov Decision Process, 9th International Conference on Application of Fuzzy Systems and Soft Computing, ICAFS 2010, August 26-27, 2010, Prague, Czech Republic.

3. Chizhov Y., Kuleshova G., Borisov A. Time series clustering approach for decision support, 16th international multi-conference on Advanced Computer Systems ACS-AISBIS 2009, October 14-16, 2009, Miedzyzdroje, Poland.
4. Chizhov Y., Zmanovska T., Borisov A. Temporal Data Mining for Identifying Customer Behaviour Patterns, Data Mining in Marketing DMM' 2009, 9th Industrial Conference, ICDM 2009, 22-24 July, 2009., Leipzig, Germany.
5. Chizhov J., Borisov A. Applying Q-Learning To Non-Markovian Environments, First International Conference on Agents and Artificial Intelligence (ICAART 2009), 19-21 January, 2009, Porto, Portugal.
6. Chizhov J., Zmanovska T., Borisov A. Ambiguous states determining in non-markovian environments, RTU ISC 49: Starptautiskā zinātniskā konference, apakšsekcija „Informācijas tehnoloģija un vadības zinātne”. 13.10.2008. Rīga, Latvija.
7. Chizhov Y. Particulars of Neural Networks applying in Reinforcement Learning, 14th International Conference on Soft Computing „MENDEL 2008”, 18-20 June, 2008 Brno University of Technology, Brno, Czech Republic.
8. Chizhov Y. Reinforcement learning with function approximation: survey and practice experience, International Conference on Modeling of Business, Industrial and Transport Systems “MBITS’08”, 7-10 May, 2008, Transport and Telecommunication Institute, Riga, Latvia.
9. Chizhov J. Software agent developing: a practical experience, Scientific proceedings of Riga Technical University: RTU 48. rakstu krājums, 5. sērija, 31. sējums, 12 October, 2007, Riga Technical University, Riga, Latvia.
10. Chizhov J., Borisov A. Increasing the effectiveness of reinforcement learning by modifying the procedure of Q-table values update, Fourth International Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control „ICSCCW – 2007”, 27-28 August 2007, Antalya, Turkey.
11. Chizhov J., Agent Control in World with Non-Markovian Property, EWSCS’07: Estonian Winter School in Computer Science. 04 – 09 March, 2007, Palmse, Estonia.

Publikācijas

Atsevišķi promocijas darba fragmenti, kā arī tā rezultāti nopublicēti 11 zinātniskajos rakstos. Vairākums publikāciju indeksējušas starptautiskās elektroniskās bibliotēkas (Springer, ISI WEB, SCOPUS, DBLP, Io-port.net un citas). Ar publikāciju sarakstu var iepazīties pilna literatūras saraksta sadaļā, kas dots autoreferāta beigās.

Promocijas darba galvenie rezultāti

Promocijas darba ietvaros izstrādāta uz MLPP balstīta lēmumu pieņemšanas pieeja, kas nodrošina MLPP modeļa uzbūvi uzdevumos, kuru dati izteikti ar daudzdimensiju laikrindām. Veicot virkni eksperimentu, īstenota pieejas aprobācija. Rezultātā iegūti skaitliski lielumi, kas ļauj secināt, ka pieeja spēj būvēt MLPP modeli, kas adekvāti atspoguļo apmācības datu kopu. Atrisināti šādi uzdevumi un izdarīti secinājumi.

1. Uz Markova procesiem balstītu matemātisko aparātu apskats ļauj secināt, ka ir iespējams uzskatīt MLPP un apmācību ar pastiprināšanu par efektīvām dinamisko sistēmu modelēšanas pieejām; noteiktas galvenās to pielietošanas problēmas uzdevumos, kuru dati izteikti ar daudzdimensiju laikrindām.
2. Atsevišķu skaitļošanas intelekta metožu (MNT, AP, datu ieguves metodes un citas) analīze ļāva aprakstīt izstrādājamās pieejas galvenās iezīmes (plūsmas konveijera veida organizācija, lai pārveidotu datus, savlaicīga MLPP modeļa atjaunināšana utt.).

3. Izstrādāta speciāla aģentu arhitektūra, kas ļauj izvairīties no nekorekta intelektuālās sistēmas un vides mijiedarbības apraksta.
4. Realizēta stāvokļu telpas un MLPP lēmumu aproksimācija, izmantojot mākslīgā neironu tīkla palīdzību. Ar modeļa uzdevumu nodemonstrēta pieejas darbaspēja.
5. Noformulēta starpstruktūra (pētāmā mainīgā uzvedības profils) pētāmo laikrindu iegūto likumsakarību uzglabāšanai un apstrādei. Pētāmā mainīgā uzvedības profili tiek izmantoti MLPP modeļa uzbūvei.
6. Piedāvāta dažādu laikrindu klasterizācijas kritēriju pielietošanas pieeja (Eiklīda attālums un kontūru līdzība), atkarībā no katra pētāmā mainīgā jēdzieniskās noslodzes.
7. Pieejas aprobēšanai noformulēts dinamiskās cenu politikas uzdevums MLPP ietvaros, kā arī izstrādāts programmatūras nodrošinājums eksperimentu veikšanai.
8. Veikta eksperimentu sērija, kas ļauj skaitliski novērtēt MLPP modeļa uzbūvēšanas efektivitāti. Novērtējums balstīts uz iegūtā modeļa un apmācības datu kopas salīdzinājumu, kā arī modeļa pielietošanu datiem ārpus apmācības datu kopas.

Darba struktūra un apjoms

Darbs sastāv no ievada, 6 nodaļām, secinājumiem, literatūras saraksta un 4 pielikumiem. Promocijas darbs ir izklāstīts uz 137 lappusēm un paskaidrots ar 70 zīmējumiem un 15 tabulām. Literatūras sarakstā ir 83 nosaukumi.

Promocijas darba struktūra ir šāda:

IEVADS – dotas izmantotās definīcijas, noformulēts pētījuma priekšmets, darba mērķis un uzdevumi.

1. NODAĻA: VAIRĀKSOĻU LĒMUMU PIENĒMŠANAS PROCESS DINAMISKĀS SISTĒMĀS – šajā nodaļā dota vairāksolū lēmumu pieņemšanas pieeju, balstītu uz Markova procesiem, analīze. Aprakstītas to galvenās priekšrocības un trūkumi, kas nosaka pilnveidotas MLPP pieejas izstrādāšanas nepieciešamību.
2. NODAĻA: SKAITĻOŠANAS INTELEKTA METODOLOĢIJU DINAMISKAJOS UZDEVUMOS APSKATS – šajā nodaļā dota vairāku skaitļošanas intelekta jomas metožu analīze, to pielietošanas aspektā izstrādājamajā MLPP pieejā.
3. NODAĻA: UZ DATU IEGUVES BALSTĪTAS PIEEJAS IZSTRĀDE DINAMISKĀS SISTĒMAS MARKOVA LĒMUMU PROCESA MODEĻA IZVEIDEI – tā ir galvenā nodaļa, kas veltīta MLPP pieejas izstrādāšanai, kura spējīga uzbūvēt modeli balstoties uz daudzdimensiju laikrindu datiem.
4. NODAĻA: AR DATU IEGUVES PALĪDZĪBU IZVEIDOTA MLPP-MODEĻA PIEEJAS IZMANTOŠANA DINAMISKĀS CENU POLITIKAS UZDEVUMĀ – šajā nodaļā noformulēts dinamiskās cenu politikas uzdevums MLPP kontekstā eksperimentu veikšanai.
5. NODAĻA: MODEĻU IZVEIDES EKSPERIMENTU VEIKŠANA DINAMISKĀS CENU POLITIKAS UZDEVUMĀ – nodaļā aprakstīti eksperimenti, kuru mērķis ir jaunas pieejas efektivitātes skaitlisko vērtējumu iegūšana. Dots izstrādātā programmnodrošinājuma apraksts.
6. NODAĻA: REZULTĀTU ANALĪZE UN SECINĀJUMI – noslēdzošā nodaļa, kas veltīta iegūto rezultātu analīzei. Tajā vienlaikus noteikti turpmāko pētījumu virzieni.

PIELIKUMS – parādītas starpdata struktūras, MLPP-modeļa fragmenti XML formātā, pētījumā pielietotie algoritmi.

DARBA ATSEVIŠĶO NODAĻU ĪSS APRAKSTS

Pirmā nodaļa (Vairāksolū lēmumu pieņemšanas process dinamiskās sistēmās)

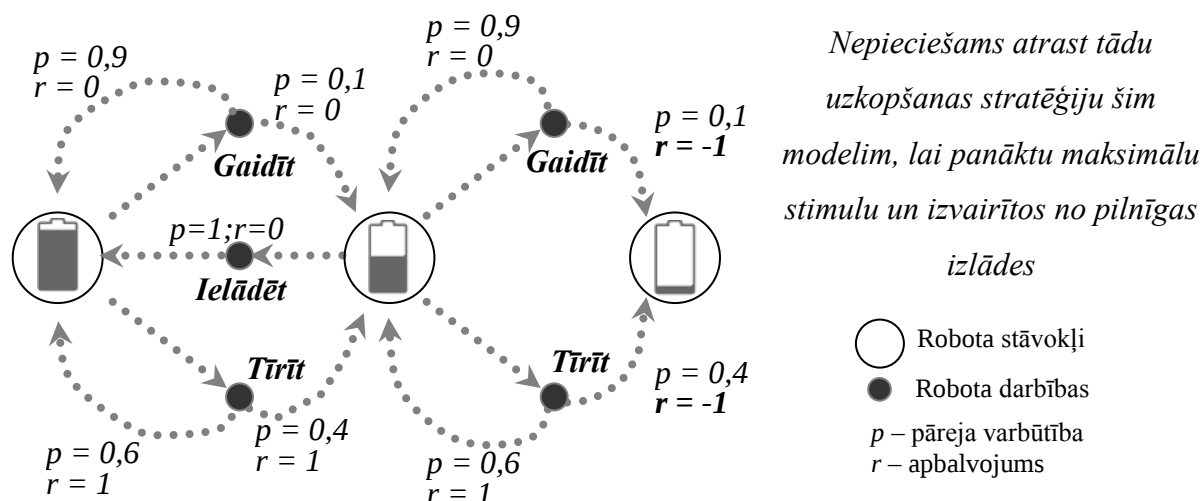
Nodaļā iztirzātas dinamisko uzdevumu risināšanas, kuru atrisināšanas princips ietverts secīgā, uz mērķa sasniegšanu vērstā, darbību izpildē, metodes. Dinamiskā programmēšana, būdama fundamentāls aparāts, ir Markova lēmumu pieņemšanas procesa pamatā, kas savukārt arī rada savu metožu kopu.

Vienlaicīgi parādīts, ka pastāv grūtības šādu metožu pielietošanai mūsdienīgos ekonomiskos uzdevumos. Pamatproblēma ir uzbūvēt modeli, kas būtu vairāksolū lēmumu pieņemšanas procesa metožu funkcionēšanas vide. Ar mērķi risināt modeļa uzbūves problēmu, tiek pētīta iespējamība pielietot likumsakarību ieguves procedūru, izstrādātu pielietošanai dinamiskajos uzdevumos.

Optimālu lēmumu pieņemšanu mūsdienu reālos vadības uzdevumos nevar aplūkot īslaicīgas vienreizējas peļņas gūšanas aspektā. Piemēram, ekonomikā ar efektīvu vadību saprot summāru apskatāmā parametra (piemēram, peļņas) maksimālo lielumu galīga vai neierobežota etapu skaita laikā. Tādējādi, ir uzdevumu veids, ko atrisina nevis uzreiz, bet pamazām, soli pa solim. Citiem vārdiem sakot, lēmumu pieņemšana uzskatāma nevis par vienības darbību, bet par procesu, kas sastāv no vairākiem etapiem [76].

Dinamiskā programmēšana ir matemātiska sistēma, kas ļauj īstenot daudzpakāpju vadības procesu un no ilguma atkarīgu procesu optimālu plānošanu [78]. Plānošana nozīmē pieņemamo lēmumu galīgo secību. DP fundamentāla īpašība ir tajā, ka izstrādājamie lēmumi nav izolēti cits no cita [40], bet ir saskaņoti mērķa sasniegšanai.

Markova lēmumu pieņemšanas process [17, 36, 54] paplašina Markova ķēdi, ieviešot vadības ietekmes jēdzienu. Pārejas varbūtību no viena stāvokļa otrā nosaka izvēlētais iedarbības vai lēmuma varbūtība. Markova lēmumu pieņemšanas procesa (MLPP) piemērs parādīts 1. attēlā, grafiski parādot pārejas atkritumu novākšanas uzdevumu robotam. Katrs melnais punkts ir pieņemtais lēmums, kas nosaka iespējamās sekojošās pārejas. Pieņemtais lēmums nosaka tālākās iespējamās pārejas.



1. att. Markova lēmumu pieņemšanas procesa piemērs

Faktiski pāreju P un apbalvojumu R varbūtības matricas apraksta procesa „likumus”, izejot no kuriem tiek aprēķināta kāda no novākšanas stratēģijām. Tā, pat pie nenozīmīgām atsevišķu matricas lielumu izmaiņām, starp rezultējošām stratēģijām var būt principiālas atšķirības.

Nodaļā aprakstīta formāla MLPP definīcija un tā korteža struktūra (1).

$$\langle S, A, P, R \rangle, \quad (1)$$

kur S – diskrēto stāvokļu galīgā kopa $S = \{s_1, s_2, \dots, s_{|S|}\}$. Katrs stāvoklis s_t parāda novērojamo parametru vektora pašreizējo nozīmi $s_t = (x_1, x_2, \dots, x_{|S|})$. Tādējādi, stāvoklis ir visa pieejamā informācija par dinamisko sistēmu noteiktā laika momentā.

A – vadības iedarbību galīgā kopa $A = \{a_1, a_2, \dots, a_{|A|}\}$. Visvienkāršākajā gadījumā, iedarbība a_i tieši izmainīs vienu vai vairākus stāvokļa s_t parametrus.

P – stāvokļu pārejas funkcija, kas nosaka varbūtību, ka darbība a , kas uzsākta s stāvoklī, t laika momentā, novedīs stāvoklī s' laika brīdī $t+1$. Tā ir formas $P : S \times A \times S \rightarrow \mathbf{P}$ attēlojums.

R – apbalvojumu funkcija, kura nosaka gaidāmo apbalvojumu (jeb pastiprinājumu), ko iegūst uzreiz pēc pārejas s' stāvoklī no s stāvokļa a darbības rezultātā. Tā ir formas $R : S \times A \times S \rightarrow \mathbf{R}$ attēlojums. Faktiski nosaka mērķa stāvokli, kuru nepieciešams sasniegt.

MLPP risinājums ir optimālas darbību politikas π^* atrašana, kas nosaka katram stāvoklim s_t atbilstošu darbību a_t . Tādā gadījumā ir patiens attēlojums $\pi : S \rightarrow A$.

MLPP pamatvērtība ir tā konverģence ar globāli optimālo politiku, kā arī vienkāršā modeļa struktūra. Iegūtā π^* politika ir cilvēkam viegli uztverama salīdzinot ar mākslīgo neironu tīkla risinājumiem. Kā arī to var izteikt, izmantojot dažādus zināšanu izpausmes veidus, piemēram, risinājumu kokus, risinājumu tabulas [67].

Būtisks MLPP trūkums ir automātiskās moduļa konstruēšanas mehānismu neesamība. Risinājums tādas politikas veidā, kas maksimizē sagaidāmo diskontēto apbalvojumu summu, var tikt iegūts, ja modeļa pamatu veidojošās pāreju matrica P un apbalvojumu funkcija R ir zināmas. Reālos uzdevumos to konstruēšana ir apgrūtināta.

Vēl viens MLPP trūkums ir politiku meklēšanas sarežģītība tā saucamajās ne-Markova sistēmās – dinamiskās sistēmās, kurām nepiemīt Markova īpašība. Ne-Markova sistēmās procesa attīstība ir atkarīga ne tikai no pašreizējā stāvokļa, bet arī no pagātnes stāvokļu secības. Ne-Markova uzdevumus iespējams risināt, aprakstot procesu, kas paredz mehānisma atmiņu. Pie tam, jāņem vērā arī, ka no pagātnes evolūcijas procesa rakstura ir atkarīgi arī tā statistiskie rādītāji nākotnē. Šāda pieeja sarežģīt risinājumu un faktiski liedz tā pielietošanas iespēju reālos uzdevumos. Pamatojoties uz MLPP un AP modeļu analīzi, 1. tabulā dotas to svarīgākās raksturlīknes stohastiskas vides izpētes un attiecīgas modeļu uzbūvēšanas uzdevumos.

1. tabula

Uz MLPP balstītu modeļu priekšrocības un trūkumi

Markova lēmumu pieņemšanas process	
<i>priekšrocības</i>	<i>trūkumi</i>
• globālā konverģence • politikas konstruēšana, ņemot vērā apbalvojuma aizkavēšanos • vienkāršas politikas izskaitļošanas metodes	• nepieciešamas aprioras sistēmas modeļa zināšanas; • pielietojuma sarežģītība ne-Markova sistēmās
Apmācība ar pastiprināšanu	
<i>priekšrocības</i>	<i>trūkumi</i>
• globālā konverģence • nav nepieciešams sistēmas modelis (<i>unsupervised learning</i>) • darba iespēja sistēmās, kurām nepiemīt Markova īpašība • politikas konstruēšana, ņemot vērā apbalvojuma aizkavējumu	• pētījuma-ekspluatācijas dilemma • modeļa konstruēšana pētījumu ceļā nav pieļaujama virknē praktisku uzdevumu • pielietojuma sarežģītība ne-Markova sistēmās

Standarta pieeja darbā ar nemarkova sistēmām ir atmiņas palielināšana operēšanai ar pāreju priekšvēsturi. Tāds pats princips tiek izmantots stāvokļa, kas darbojas ar laikrindām Markova modeļa uzbūvei, ieviešanas pieejā.

Uz Markova procesiem pamatotu matemātisko sistēmu apskats ļauj aplūkot Markova lēmumu pieņemšanas procesu un apmācību ar pastiprināšanu kā efektīvas pieejas dinamisko sistēmu modelēšanai.

Mūsdienu MLPP un AP sistēmu izpēte iezīmējusi galvenās problēmas to pielietojumam reālos ekonomikas, vadības un citos uzdevumos.

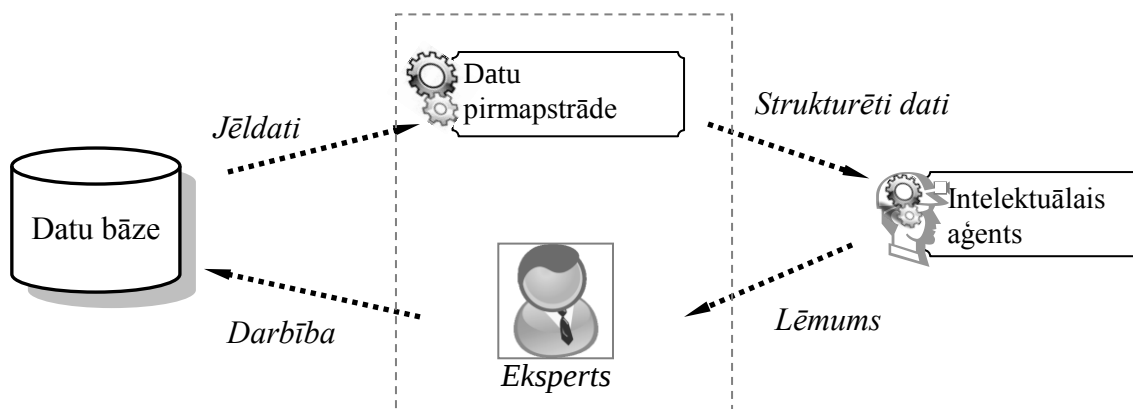
Atklāto problēmu analīze nodrošinājusi iespēju noformulēt uz datu ieguves metodēm balstītu pieeju, lai pilnveidotu MLPP aparātu, kas nodrošina tās pielietošanu mūsdienu uzdevumos, kuri aprakstīti ne-Markova dinamiskās sistēmas nostādnē.

Otrā nodaļa (Skaitļošanas intelekta metodoloģiju dinamiskajos uzdevumos apskats)

Šajā nodaļā sniegts pamatojums skaitļošanas intelekta metožu pielietošanas nepieciešamībai, lai izveidotu efektīvu lēmumu pieņemšanas sistēmu. Izanalizētas dažādas aģentu sistēmu arhitektūras specifiskai dinamiskās sistēmas aģentu arhitektūras uzbūvēšanai. Savukārt, uz aģentu balstīta pieeja ļauj cilvēkam dabiskā veidā izteikt programmatūras sistēmu un tās mijiedarbību ar risināmo reālās pasaules uzdevumu. Veikti arī eksperimenti ar modeļu dinamiskajiem uzdevumiem, kuros izmantoti mākslīgie neironu tīkli, lai aproksimētu risinājumu telpu.

Pastāv vairākas formālas skaitļošanas intelekta definīcijas. Darbā [40] SI jēdziens definēts kā skaitļošanas modeļu un instrumentu sakopojums, kuram ir intelektuāla piemērotība tiešai primāro sensoru datu uztverei, to apstrādei, iesaistot uzdevumu atpazīšanu un konveijerizāciju, izveidojot drošu un savlaicīgi reaģējošu, bojājumpieciefīgu sistēmu.

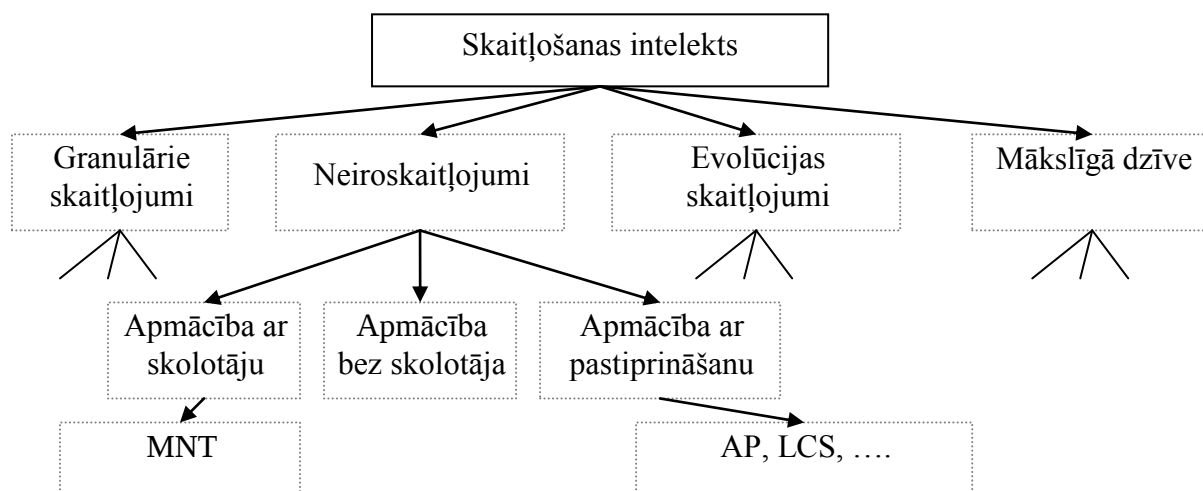
Parasti tieši apstrādāt „jēlus” datus, izmantojot intelektuālās programmēšanas instrumentus, nav iespējams. To nosaka algoritmu stingrās prasības pret datu struktūru. Piemēram, MLPP modeļi strādā ar stāvokli noteicošo, fiksētu datu struktūru. Tādā veidā, nepieciešams zināma veida starpnieks starp fizisko datu nesēju un to vai citu intelektuālo metodi. Tas, savukārt, nodrošina uzdevuma konveijerizāciju. Sistēmas paraugs, kas paredz pastarpinātu intelektuālo instrumentu un uzdevuma mijiedarbību, redzams 2. attēlā.



2. att. Pastarpinātas mijiedarbības pieeja

Šajā darbā tiek piedāvāts apsvērt arī intelektuālās sistēmas netiešo iedarbību uz uzdevuma fizisko avotu (šajā gadījumā – zināmu datu bāzi), kas piemīt uzdevumiem, kam raksturīgas augstas tāda vai cita veida izmaksas kļūdainu lēmumu gadījumā.

Šajā pētījumā izstrādājamajā pieejā dinamisko uzdevumu risināšanai iesaistīta virkne skaitļošanas intelekta metožu. Pamatojoties uz piedāvāto darbā [40] klasifikāciju, attēlotas šim pētījumam aktuālās metodes un to vieta skaitļošanas intelekta metožu saimē (sk. 3. att.).



3. att. Skaitļošanas intelekta metožu saime

Šajā darbā tika pieņemta aģenta definīcija, kas piedāvāta darbā [28]: „*Patstāvīgs aģents – sistēma, kas darbojas vidē un ir tās daļa, uztveroša šo vidi un iedarbojoties uz to laikā, ar mērķi iemiesot savas matērijas un atspoguļot parādīto iedarbību nākotnē uztveramajā vidē*”.

Pastāv liels skaits aģentu veidu, kas pilnībā vai daļēji atbilst dotajai definīcijai. Atkarībā no aģentiem piemītošajām īpašībām, izdala vairākas aģentu klases [11]. Starp tām raksturīgākie ir: programmējamie aģenti (reaktīvie aģenti, refleksīvie aģenti [58]), apmācāmie aģenti un plānojošie aģenti [37]. Īpašības, kādas var piemist tās vai citas klases aģentam (pamatojoties uz darbu [28]), sniegtas 2. tabulā.

2. tabula

Programmas aģentu īpašības

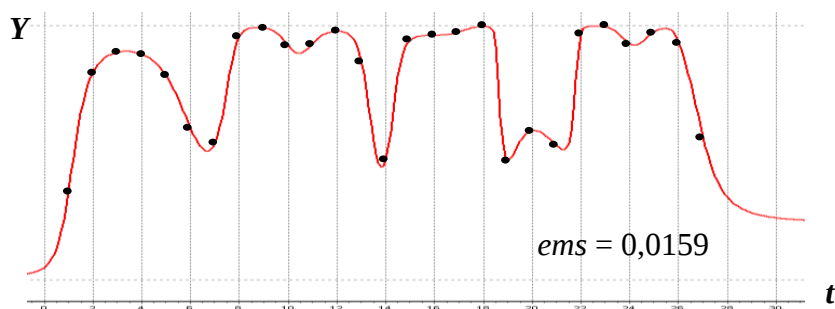
Īpašība	Apraksts
Reaktivitāte (uztvere un iedarbība)	Savlaicīgi reaģē uz izmaiņām ārējā vidē
Patstāvība	Piemīt meta kontrole pār savu rīcību
Mērķtiecība	Sarežģītu un neskaidru lēmumu pieņemšana atsevišķos apakšuzdevumos
Funkcionēšanas pastāvīgums	Funkcionēšanas process nav ierobežots laikā
Komunikabilitāte	Apmainās ar informāciju ar citiem aģentiem vai cilvēkiem
Apmācāmība (adaptējamība)	Maina savu uzvedību, pamatojoties uz pagātnes pieredzi
Mobilitāte	Var pārvietoties no vienas vides uz citu
Elastīgums	Darbību secība nav stingrs scenārijs
Specifiskums	Viendabīgu aģentu sabiedrībā piemīt raksturīga uzvedība, emocijas

Risinājumu telpas aproksimācija ar mākslīgo neironu tīklu. Viens no apmācības ar pastiprināšanu trūkumiem ir uzdevuma telpas eksponenciālā augšana ar katru jaunu dimensiju [58, 62]. Šajā nodaļā iztirzātas divas galvenās pieejas darbam ar lielu stāvokļu skaitu: gradientu politikas metodes un vērtības funkcijas aproksimācija. Viena no nelineārās aproksimācijas metodēm ir mākslīgo neironu tīklu aparāts (MNT).

Ar mērķi veikt MNT analīzi aproksimācijas funkcijas aspektā šajā darbā tiek realizēts daudzslāņu perceptrons ar kļūdu atgriezeniskās izplatīšanās apmācības metodi. Tāpat tiek apskatītas esošās komerciāli un brīvi izplatāmās programmas mākslīgo neironu tīklu realizācijai. Eksperimentu plāns paredz šādus uzdevumus:

1. Izstrādāt MNT programnodrošinājums. Izanalizēt tīkla darbības efektivitāti uz kāda stohastiskā procesa aproksimācijas piemēra;
2. Salīdzināt iegūtos aproksimācijas rezultātus ar esošo MNT pakešu rezultātiem;
3. Realizēt stāvokļu telpas aproksimāciju apmācīšanā ar pastiprināšanu metodē; izmantot modeļa uzdevuma piemēru, lai novērtētu metodes darbaspēju.

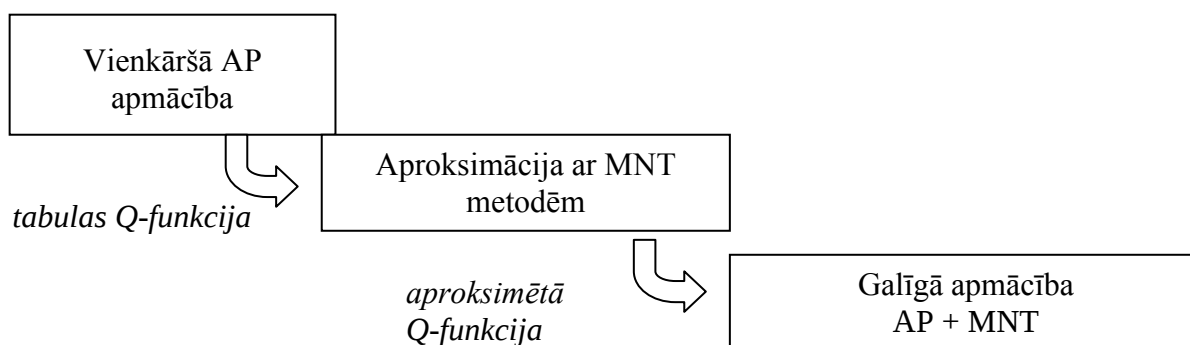
Pirmā eksperimenta ietvaros neironu tīkls, realizēts promocijas darbā, demonstrē augstus apmācības rezultātus. Trīs slēpto slāņu izmantošana un 70 neironu katrā slānī nodrošina vidējās kvadrātiskās kļūdas $mse = 0,0013$ vērtību (sk. 4. att.). Tāds kļūdas līmenis nodrošina pietiekamo precizitāti viendimensijas stohastiskā procesa modelēšanai, kas sastāv no 30 novērojumiem.



4. att. Aproksimējošā funkcija ar trim slēptajiem MNT slāņiem

Otrā eksperimenta ietvaros īstenota rezultātu salīdzināšana ar divām visvairāk izplatītajām MNT paketēm: *Neurosolutions* 6.0 un *Multiple Back-Propagation* (MBP) v.2.2.2. Programmnodrošinājums *Neurosolutions* 6.0 visvienkāršākajai tīkla arhitektūrai aproksimē ar vidējo kvadrātisko kļūdu $mse = 0,00943$. Savukārt MBP v.2.2.2. nodrošina konverģenci ar kļūdas lielumu $mse = 0.0012$. Salīdzinājums ļauj secināt, ka iegūtā precizitāte atbilst svešu pakešu precizitātei. Tas ļauj izmantot paša veiktu neironu tīkla realizāciju tālākajos eksperimentos.

Trešajā eksperimentā ar MNT palīdzību ir realizēta Q-funkcijas aproksimācijas ideja apmācībā ar pastiprināšanu (AP). Parādītas metodes realizācijas pamatproblēmas, kuru pārvarēšanai šajā darbā piedāvāta jauna pieeja MNT apmācībai, iebūvētai AP (sk. 5. att.).

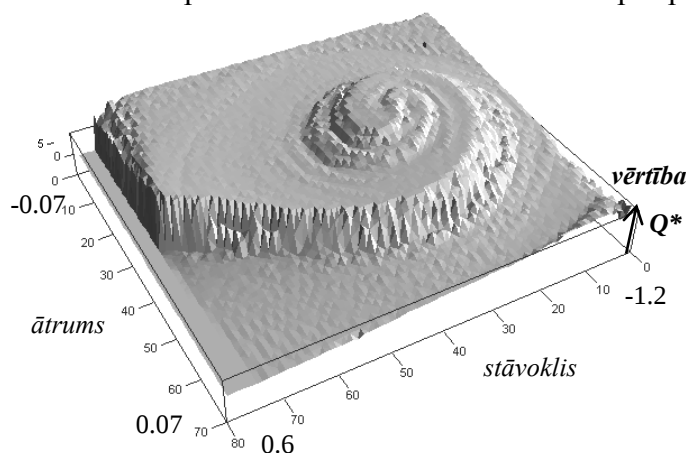


5. att. Trīssoļu modeļa uzbūves algoritms

Kā eksperimentālais uzdevums tiek izmantots *Mountain Car* uzdevums, jo tas nav triviāls: gravitācija ir stiprāka nekā mašīnas dzinējs, tādējādi pat pie pilnās dzinēja darbības, mašīna nevar paātrināties stāvā nogāzē, uzsākot kustību no apakšas ar nulles ātrumu [10]. Vienīgais risinājums – izveidot šūpošanās no vienas nogāzes otrā stratēģiju, ar mērķi uzkrāt papildus inerces spēku. Uzdevums demonstrē nepieciešamību vairākkārt attālināties no virsotnes,

lai sasniegtu to nākotnē. Mašīnas iespējamās darbības: pilnā gāze uz priekšu (+1), pilnā gāze atpakaļgaitā (-1) un dzinēja bezdarbība (0). Mašīna kustas saskaņā ar vienkāršotu fiziku.

Modeļu uzdevuma optimālā politika tiek meklēta, izmantojot diskrēto AP ar tabulas Q-funkciju. Q-funkcijas virsma stāvokļa telpā parādīta 6. attēlā. Darbību telpas dimensija izlaista, bet izmantota optimālās darbības vērtība katrā telpas punktā.



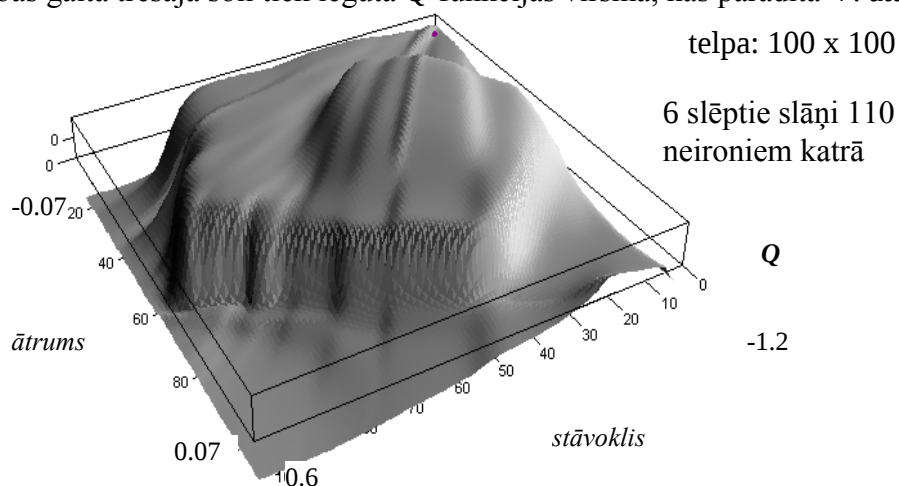
Q ass faktors: 0.1
Telpa: 70 x 80
Epizožu skaits: $\approx 8\,000$

ϵ : 0.1 *gadījuma komponente*
 γ : 0.99 *diskontēšanas parametrs*
 α : 0.3 *apmācības koeficients*
 λ : 0.92 *pēdu dzēšanas parametrs*

6. att. Virsmas $Q^* = \max_a Q_t(s)$ piemērs

Līdzīgu virsmu ir nepieciešams iegūt, izpildot apvienoto 3-soļu algoritmu. Eksperimenta pirmais solis ir uzbūvēt rupju politiku. Empīriski tika noteikts, ka šim mērķim var izmantot arī nelielu kopu: 20 reiz 20 elementi [9]. Nākamajā solī ir “jāielādē” starp-Q-funkcija neironu tīklā. Šajā gadījumā funkcijas precizitāte atkarīga no tīkla „kapacitātes”. Pamatojoties uz eksperimentu sēriju, noskaidrots, ka 6 slēptās kārtas ar 110 neironiem katrā, ir pietiekami Q-funkcijas virsmas iegūšanai [9]. Iegūtā politika izskatās gludāka nekā tās tabulveida analogs. No vienas puses, tas samazina precizitāti, bet no otras puses, ļauj strādāt vidē ar nepārtrauktiem parametriem.

Tīkla koeficientu izmaiņu tuvošanās nullei ($\Delta e_{ij} \rightarrow 0$) ļauj pāriet pie trešā soļa - tīkla apmācības mijiedarbības režīmā ar vidi, izmantojot AP aparātu. Eksperimenti demonstrē, ka tīkls jau pēc desmit apmācības iterācijām „aizmirst” iepriekš apmācītos apgabalus, ja tie nav bijuši pakļauti nepārtrauktai apmācībai reizē ar apmācāmajiem telpas apgabaliem. Tā rezultātā iepriekšējos soļos iegūtā aptuvenā politika kalpo kā reperu punktu matrica. Tāda matrica uztur neironu tīkla „atmiņu”, neļaujot aizmirst reakciju uz reti sastopamiem stāvokļiem vidē apmācības gaitā [9]. Apmācības gaitā trešajā solī tiek iegūta Q-funkcijas virsma, kas parādīta 7. att.



7. att. Iegūta 3. solī Q-funkcija

Eksperimenti ar modeļu uzdevumu parādījuši, ka algoritms ļauj izvairīties no sākotnējās apmācības datu kopas neesamības problēmām, pieļauj funkcionēšanu nepārtrauktās vidēs [9]. Vienlaicīgi pieaug apmācības laiks.

Var secināt, ka MNT ar atgriezenisku kļūdas izplatīšanu piešķir spēcīgu instrumentu nelineāro funkciju, dotu tabulas veidā, aproksimācijai [34]. Veiktie eksperimenti ar modeļu uzdevumu demonstrē, ka šī īpašība tiek veiksmīgi pielietota Q-funkcijas aproksimācijai apmācības ar pastiprināšanu algoritmā.

Skaitļošanas intelekta jēdziena un atbilstošo metožu iztīrāšana šajā sadaļā ļauj formulēt šādus secinājumus.

1. Atsevišķu skaitļošanas intelekta koncepciju, kā arī atsevišķu tā metožu (MNT, AP, aģentu pieeja, datu ieguves metode un citas) analīze ļāva aprakstīt izstrādājamās arhitektūras galvenās iezīmes (plūsmas konveijera organizācija ar datu pārvēršanas mērķi, savlaicīga reaģēšana uz notikumiem un tā tālāk), kā arī izmantojamās intelektuālās metodes.
2. Aplūkotā Markova lēmumu pieņemšanas procesu izteikšana ar programmu aģentiem demonstrē virkni pastāvošo metožu un plašu risināmo uzdevumu loku; vienlaikus, problēmām, kas saistītas ar uzdevuma konveijerizāciju, nepieciešama īpaša veida aģentu sistēmas arhitektūras izstrāde.
3. No metožu sinerģisma principa viedokļa skaitļošanas intelektā, šajā darbā mākslīgie neironu tīkli tiek aplūkoti kā MLPP stāvokļu telpas aproksimācija.

Trešā nodaļa (Uz datu ieguves balstītas pieejas izstrāde dinamiskās sistēmas Markova lēmumu procesa modeļa izveidei)

Šajā sadaļā tiek aplūkota intelektuālās sistēmas uzbūve lēmumu pieņemšana uzdevumā par datiem, izteiktiem ar daudzdimensiju laika struktūrām.

Risinājums tiek uzbūvēts uz slēptās likumsakarības meklēšanas tā saucamo laikrindu saimju starpā. Šiem mērķiem tiek izmantots Datu ieguves tehnoloģiju klāsts. Iegūtu datu operatīvai glābšanai tiek ieviesti klašu un uzvedības profilu jēdzieni. Uzvedības profili tiek interpretēti kā viensoļa Markova pārejas. Savukārt Markova pārejas tiek izmantotas pētāmā procesa modeļa uzbūvēšanai.

Šīs nodaļas ietvaros ieplānotā uzdevuma atrisināšanai jāīsteno sekojošais:

1. jāizpilda mūsdienīga stāvokļa apskats aparāta pielietošanas jomā uzdevumiem, izteiktiem daudzdimensiju laikrindās, kā arī pastāvošo iespējamo pieeju pētīšana;
2. jāizstrādā uz datu ieguves līdzekļiem balstītās metodes laikrindu uzbūvēšanai, to apstrādei un pārveidošanai datu struktūrās (sistēmas stāvokļi), kas savietojamas ar Markova lēmumu procesu.

Savukārt, Markova modeļa uzbūves metodes izpēte un izstrāde laikrindās izteiktiem datiem paredz sekojošu problēmu risināšanu:

- novērojamā procesa izejas datu notīrīšana, laikrindu veidošana;
- laikrindu kā vides, kas nepieciešama stāvokļu telpas uzbūvei, formalizācija;
- vispārinātas novērojamo mainīgo uzvedības dinamiku atspoguļojošā pāreju grafa uzbūve;
- optimālo politiku izstrādāšana un to pielietošana.

Tādējādi, pieeja balstīta uz atsevišķu novērojumu apkopojumu, stāvokļu telpas izveidošanu un attiecīgā modeļa uzbūvi. Kopumā sistēmas, pamatotas uz izstrādājamajām metodēm, funkcionēšanai jā sastāv no sekojošiem soļiem:

- „apgūt” esošos novērojumus par pētāmā procesa attīstību:
 - atklāt pāreju uz to vai citu stāvokli likumsakarības, pastāvot atbilstošai iedarbībai;
 - izveidot stāvokļu pāreju grafu, kas ir pētāmā procesa modelis;
- meklējamajam risinājumam, pastāvot pašreizējiem parametriem, pamatojoties uz modeli, izveidot Markova procesa partikulāru realizāciju;
- Markova procesa partikulārajā realizācijā izveidot darbību politiku π ;
- piedāvāt ekspertam rīkoties saskaņā ar izstrādāto politiku π ;
- īstenot nepārtrauktu vides, pāreju grafa un politikas atjaunināšanu, ienākot jaunām ziņām par faktisko procesa attīstību.

Šajā nodaļā sniegta informācija par MLPP modeļa uzbūves metodes izstrādi. Piedāvāta un aprakstīta laikrindās izteiktu datu pārveidošanas Markova lēmumu pieņemšanas procesā metode. Šajā nolūkā ieviests novērojamā mainīgā uzvedības profila jēdziens. Aplūkoti laikrindu salīdzināšanas kritēriji.

Pieejas apraksts. Viens no Markova lēmumu pieņemšanas modeļu pielietošanas novirzieniem ir izpētīt viena mainīgā atkarības no citiem dinamiku. Galamērķis ir iegūtā modeļa pielietošana savlaicīgai lēmuma pieņemšanai par veicamo darbību ar mērķi sasniegt vēlamās dinamiskās sistēmas rādītājus. Pamatproblēma, kas saistīta ar vairumu matemātisko modeļu, ir to uzbūvēšana.

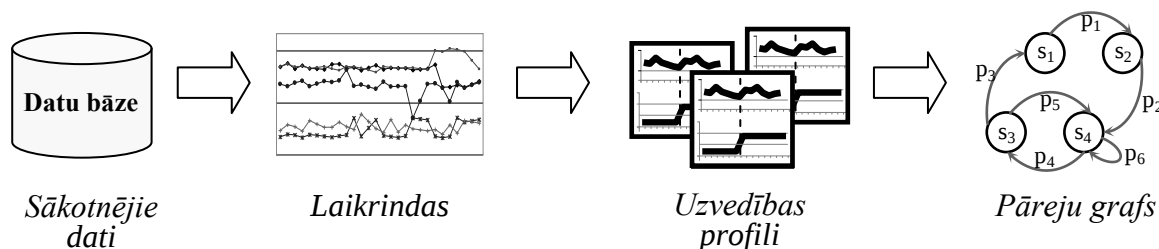
Līdzīgi [33] visa novērojumu kopa tiek aplūkota kā uzvedības šablonu (jeb profilu) avots, bet, atšķirībā no [33], šajā darbā modelis tiek uzbūvēts nevis izmantojot vienīgo laikrindas realizāciju, bet gan realizāciju kopu, atbilstošu dažādām laikrindu parametru vērtību kombinācijām. Laikrindu realizācijas klasterizācija ļauj aplūkot jauno operāciju ar pāreju modeļiem: minimāli nepieciešamo pāreju modeļa papildinājumu ar citu analogisku modeļu fragmentiem. Šī operācija dod iespēju pabeigt pastāvošā pāreju modeļa uzbūvi tā, lai modelis ar zināmu varbūtību ļautu novest sistēmu stāvoklī, kas nebija paredzēts apmācības (modeļa uzbūves) laikā.

Vispārējos vilcienos dinamiskās sistēmas modeļa uzbūve paredz partikulāru likumsakarību, kas notikušas pētāmo procesu attīstībā, izpēti, to apkopošanu un izteikšanu zināmā struktūrā [4, 7].

Izstrādājamā dinamiskās sistēmas uzbūves metode atbilst aprakstītajai pieejai. Kā modelis tiek aplūkots matemātiskais Markova lēmumu pieņemšanas aparāts. Kā modeļa grafiskās struktūras izteiksme tiek izmantots stāvokļu pārejas grafs.

Modeļa uzbūves metode (sk.8. att.) sadalīta sekojošos pamatposmos [6]:

1. apstrādāt izejas datus, izveidot laikrindas T ;
2. atklāt kopējās procesu attīstības likumsakarības, uzbūvēt novērojamo mainīgo uzvedības profilus Π ;
3. atrast kopējās pārejas dinamiku uzvedības profilos, uzbūvēt procesu attīstības modeli (pāreju grafu).

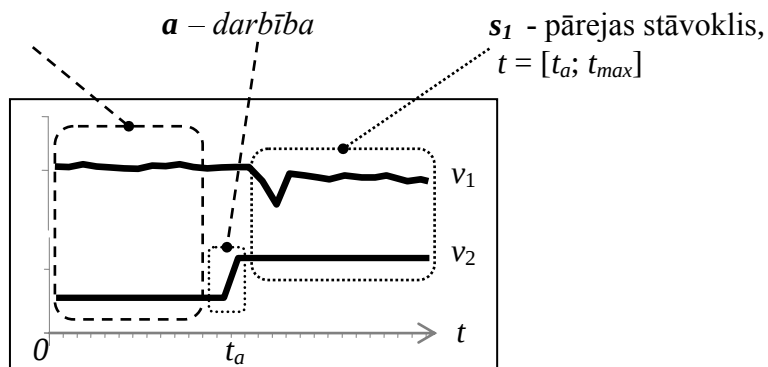


8. att. Galvenie Markova pāreju modeļa uzbūves posmi

Laikrindu pārveidošanas MLPP struktūrā atslēgas moments ir dinamiku uzvedības profila interpretācija stāvokļu matricas S veidošanās aspektā. Mūsdienās, aplūkojot Markova procesu, stāvokļa S jēdzienu saista ar programmas intelektuālo aģentu teoriju [62]. Par stāvokli sauc visu pieejamo informāciju, ko aģenta sensori iegūst no vides atsevišķā laika brīdī.

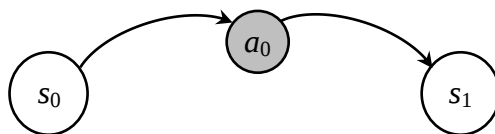
Uzdevumā, kura dati doti kā laikrindas, piedāvāta speciāla pieeja dinamiskās sistēmas stāvokļa definēšanai. Galvenā atšķirība ir nevis statistisku mainīgo, bet attiecīgo laikrindu izskatīšana, citiem vārdiem – tiek aplūkota pētāmo mainīgo attīstības dinamika. Šiem mērķiem atbilst sekojoša dinamiku uzvedības profila interpretācija. Lai laikrindu klasterizācijas rezultātā iegūtais uzvedības profils sastāv no diviem centroīdiem: v_1 un v_2 (sk. 9. att.).

s_0 - sākumstāvoklis,
 $t = [1; t_a)$



9. att. v_1 un v_2 uzvedības profila piemērs

Tad profilā tiek apzīmēta struktūra, kas satur trīs elementus [6]: a) mainīgo v_1 un v_2 dinamiku līdz notikumam e_i , b) notikumu e_i , kas noticis laika momentā t_a , c) mainīgo v_1 un v_2 dinamiku pēc notikuma e_i . Tādā veidā, mainīgo attīstības apgabalu, kas novērojams līdz notikumam, interpretēsim kā izejas stāvokli $s_0 \in S$ (sk. 9. att., apgabals $t = [1; t_a)$). Notikumu e_j – kā darbību $a \in A$. Mainīgo attīstības apgabalu, kas novērojams pēc notikuma, interpretēsim kā pārejas stāvokli $s_1 \in S$ (apgabals $t = [t_a; t_{\max}]$). Tad dinamiku uzvedības profilu Π var aplūkot kā viensoļu determinētu Markova procesa pāreju [6] (sk. 10. att.). Ņemot vērā, ka pāreja veidota, balstoties uz faktiskiem novērojumiem, pāreju varbūtības $p(s_1|s_0, a_0)$ vērtība pagaidām sastāda vieninieku



10. att. Pārdošanu profila interpretācija [6]

Šeit ar baltajiem apliem apzīmēti kopas stāvokļi S , ar pelēko apli – darbība a_0 , kuras rezultātā tiek īstenota determinētā pāreja. Pēc atbilstoša stāvokļa identifikatora piešķiršanas katra profila katram fragmentam kļūst iespējams strādāt nevis ar laikrindām, bet ar tām atbilstošiem stāvokļiem, kas ir nepieciešams nosacījums darbā ar Markova lēmumu pieņemšanas procesiem.

Mērķa stāvokļa un uzdevuma apraksts

Stāvokļu kopas veidošanas posmā mērķa stāvoklis nav aktuāls. Bet, tikko modelis uzbūvēts, jābūt noteiktiem sākuma un mērķa stāvokļiem, lai izveidotu darbības politiku. Attiecīgi, aprakstīsim kādā veidā tiek interpretēts Markova lēmumu pieņemšanas procesa mērķa stāvoklis uzdevumā, kas izteikts ar laikrindām.

Mērķa stāvoklis ir novērojamo mainīgo v_i , attiecība, kas apmierina eksperta prasības [6]. Piemēram, automašīnu pārdošanu uzdevumā, tiek izskatīti novērojamie mainīgie v_1 – pārdošanu apjoms un v_2 – pārdošanu cena. Tad, dažiem fiksētiem telpas Ψ parametriem mērķis var būt stāvoklis, kurā pārdošanu v_1 apjomu un pārdošanu cenu laikrinda atbilst norādītajam līmenim.

Darbību matrica A un apbalvojumu matrica R

Pētāmās sistēmas modeļa uzbūvē darbību matrica var tikt iegūta dažādos veidos, piemēram:

- 1) darbību kopa A tiek veidota tikai no darbībām, kas ir klātesošas profilu kopā Π ;
- 2) darbību kopa A tiek veidota, izmantojot visus pieļaujamās kontrolējamā mainīgā v_i izmaiņu lielumus zināmā diapazonā, neatkarīgi no tā, vai konkrētais lielums ir klātesošs profilos Π .

Analoģiski mērķa stāvokļiem, neskarot apbalvojumu matricas R uzbūves metodi, var aplūkot tās noteikšanu gadījumā, ja uzdevums izteikts laikrindās. Apbalvojumu matrica R ir

samērīga stāvokļu skaitam un katram stāvoklim glabā apbalvojuma vērtību, ko iegūs sistēma (vai aģents), sasniedzot kārtējo stāvokli s_i . Piemēram, lai apbalvojums sastāda lielumu 1, ja sasniegts mērķa stāvoklis s_{goal} , un lielumu -0,04 – pretējā gadījumā. Tad matrica $\mathbf{R}$ var tikt attēlota funkcijas veidā:

$$r_{s_i} = R(s_i) = \begin{cases} 1.0, & s_i \in S_{goal} \\ -0.04, & s_i \notin S_{goal} \end{cases} \quad (2)$$

Tādējādi, lai nodrošinātu sistēmu ar apbalvojumu matricu $\mathbf{R}$, pietiek noteikt mērķa stāvokļu kopu S_{goal} .

Pāreju varbūtības grafs $\mathbf{P}$

Par centrālu elementu MLPP tiek uzskatīts pāreju grafs, kurš prasa optimālas darbības politikas π^* atrašanu. Pāreju grafa uzbūve rada optimizācijas uzdevuma izteiksmes MLPP nostādnē pamatgrūtību.

Virknes ekonomisku uzdevumu gadījumā, pāreju modeli nevar būt izteikts analītiskā veidā un būt korekts visiem tā stāvokļiem, bet var tikt iegūts tabulas veidā, pamatojoties uz faktisko pāreju vispārinājumu. Tabulas veids ļauj aprakstīt jebkuras stāvokļa-darbības kombinācijas attēlojumu darbībā. Attiecīgi, pāreju grafs tiek uzbūvēts visai telpai Ψ (izejot no noteikta stāvokļa).

Pāreju grafs tiek būvēts atomāro pāreju (dinamiku uzvedības profilu) analīzes procesā. Vispārināšana nozīmē pāreju varbūtības aprēķinu no stāvokļa s stāvoklī s' , izpildot darbību a un pamatota uz faktisko šīs pārejas novērojumu skaitu (3). Tā kā vide ir pilnībā apskatāma (darbs notiek ar apmācības datu kopu), tad pāreju varbūtību matricas $\mathbf{P}$ aprēķinam ir lietderīgi izmantot nezināma parametra novērtēšanas statistisku pieeju. Šajā pētījumā izskatīsim vienu no vienkāršākajām pieejām – maksimālās ticamības metodi. Tad katras pārejas faktisko novērojumu daudzuma aprēķins attiecībā pret pāreju no izskatāmā stāvokļa kopējo daudzumu ir izteikts šādā veidā:

$$P(s'|s, a) = \frac{N(s, a, s')}{\sum_{s''} N(s, a, s'')}, \quad \sum_{s''} P(s, a) = 1, \quad (3)$$

kur s' – pārejas stāvoklis, s'' – jebkurš kopas S stāvoklis, kurā iespējama pāreja (tas ir, pastāv atbilstoši faktiskie novērojumi) no stāvokļa s , veicot darbību a . $N(s, a, s')$ - faktisko pāreju daudzums no stāvokļa s stāvoklī s' , veicot darbību a ; $\sum_{s''} N(s, a, s'')$ - pāreju visos iespējamajos stāvokļos no stāvokļa s , veicot darbību a , kopējais daudzums.

Klasterizācijas procedūru izstrāde

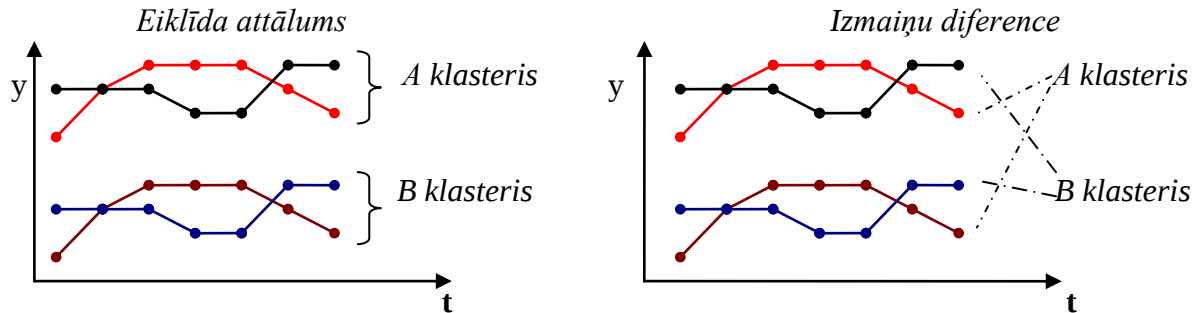
MLPP modeļa uzbūves procesā klasterizācijas procedūras ir centrālais faktisko novērojumu vispārināšanas mehānisms. No tā, cik lietpratīgi piemeklēti klasterizācijas kritēriji attiecībā pret klasterizējamiem datiem un kādas klasterizācijas parametru vērtības pieņemts, atkarīga prognozes precizitāte un topošā modeļa „adekvātums” apmācāmajiem datiem.

Šajā pētījumā tiek pielietota aglomeratīva hierarhiska laikrindu klasterizācija, kas prasa simetriskas attālumu matricas veidošanu. Tā kā klasterizācijas objekti ir laikrindas, vispirms jānosaka metrikas, kas ļauj skaitliski salīdzināt vienas laikrindas līdzību citām.

Distances aprēķināšanai izmantosim divus vienkāršus aprēķinu kritērijus: Eiklīda attālumu un izmaiņu diferenci (*Shape-based*), kas ļauj salīdzināt nevis absolūtus lielumus, bet atbilstošo laikrindu līkņu (kontūru) attīstības formas (4).

$$S(a, b) = \sum_{i=1}^{N-1} |\Delta_{a,i} - \Delta_{b,i}|, \quad \Delta_{a,i} = a_i - a_{i+1}, \quad \Delta_{b,i} = b_i - b_{i+1}. \quad (4)$$

Eiklīda attālums ļauj sagrupēt laikrindas, kas tuvas viena otrai attāluma ziņā, bet izmaiņas diferenci grupē pa grafika kontūru jeb profilu (sk. 11. att.).



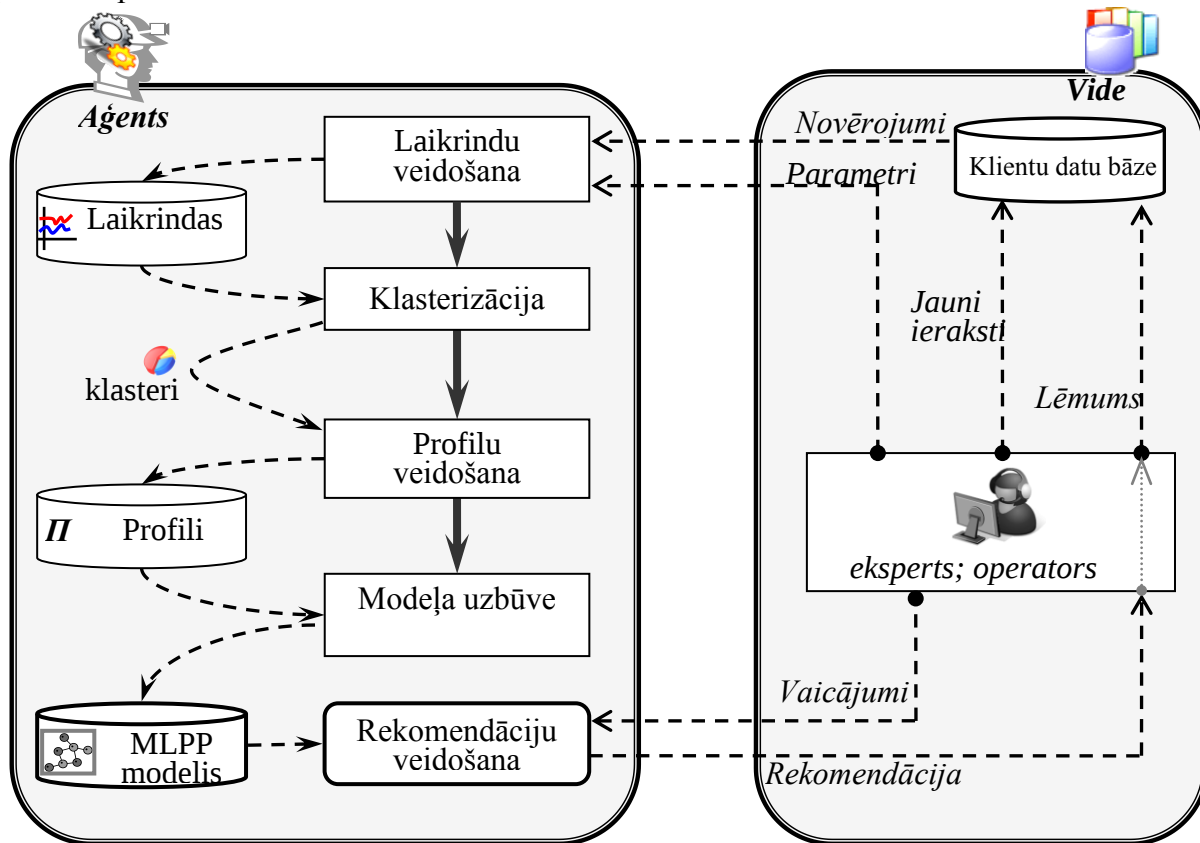
11. att. Dažādu metriku pielietošanas rezultāti

Tāda pieeja ļauj iegūt laikrindu klasterus ar līdzīgām kontūrām pēc noieta mēroga vienveidīgās grupās. Pārdošanu cenu laikrindām tiek veikta tikai klasterizācija pēc izmaiņu atšķirības kritērija.

MLPP modeļa uzbūves metodes attēlošana intelektuālās sistēmas veidā

Dinamiskās programmēšanas metodes ir balstītas uz secīgu starplēmumu, kas tuvina mērķim, pieņemšanas principu. Lēmuma meklēšanas tehnika iegūst iteratīvu raksturu. Tā rezultātā to vai citu metodi vislabāk realizēt programmas rīka veidā. Nepieciešams izstrādāt programmas rīka arhitektūru, kas priekšnoteiks intelektuālās sistēmas struktūru.

Sistēmas izveidošana nodrošina iepriekš minēto dinamiskās programmēšanas metožu savstarpēju mijiedarbību, kā arī mijiedarbību ar datu bāzi (uzdevuma datu avotu) un ar ekspertu (operatoru). Bez tam, intelektuālās sistēmas arhitektūrai jānodrošina mijiedarbību ar lietotāju datu bāzi, ekspertu un operatoriem. Sistēmas arhitektūra, kas atbilst visām iepriekš minētajām prasībām parādītā 12. attēlā.



12. att. Intelektuālās sistēmas funkcionēšanas diagramma un tās mijiedarbība ar apkārtējo vidi

Ar dubultu kontūru apzīmētajā apgabalā izvietoti intelektuālās sistēmas elementi, kas ietver apstrādes moduļus, laukrindu glabātuvi un profilu glabātuvi. Intelektuālajai sistēmai piemīt vizuālais interfeiss sadarbībai ar menedžeriem (operatoriem) un programmatūras interfeiss sadarbībai ar lietotāju datu bāzi. Ar punktētajām līnijām šeit apzīmēti datu plūsmu virzieni, ar nepārtrauktajām – vadības plūsmas virziens. Projektējamajai sistēmai piemīt tādas īpašības kā autonoma un nepārtraukta funkcionēšana, adaptācija datu izmaiņu gadījumā (citiem vārdiem – zināma veida apmācība), mērķtiecība (mērķis optimizēt tos vai citus pētāmā procesa parametrus modeļa uzbūves ceļā), komunikabilitāte (mijiedarbība ar lietotājiem un datu bāzi) un citas. Tas ļauj aplūkot programmēšanas sistēmu kā programmas intelektuālu aģentu.

Ceturrtā nodaļa (Ar datu ieguves palīdzību izveidota MLPP-modeļa pieejas izmantošana dinamiskās cenu politikas uzdevumā)

Pārdošanu vadības prakse norāda uz dinamiskās cenu politikas nepieciešamību konkurentsējas paaugstināšanai. Markova modeļi ir efektīvi, ja lēmumu pieņemšana notikumu hronikā iekļauj nenoteiktību, bet atslēgas notikumi norisinās vairāk kārt. Šīs nodaļas mērķis ir demonstrēt iespēju pielietot Markova lēmumu pieņemšanas procesa aparātu dinamiskās cenu politikas problēmas risināšanā, izmantojot testa piemēru. Attiecīgi, šīs nodaļas uzdevumi ir šādi:

- formulēt dinamiskās cenu politikas problēmu kā dinamiskās programmēšanas uzdevumus;
- formulēt un aprakstīt dinamiskās cenu politikas vadības metodi, pamatojoties uz Markova lēmumu pieņemšanas procesu;
- formulēt MLPP modeļa uzbūves metodi, pamatojoties uz likumsakarībām, kas iegūtas ar datu ieguves līdzekļiem no faktiskām ziņām par pārdošanām.

Galvenie dinamiskā cenu politikas kā eksperimentāla uzdevuma izvēles iemesli ir šādi:

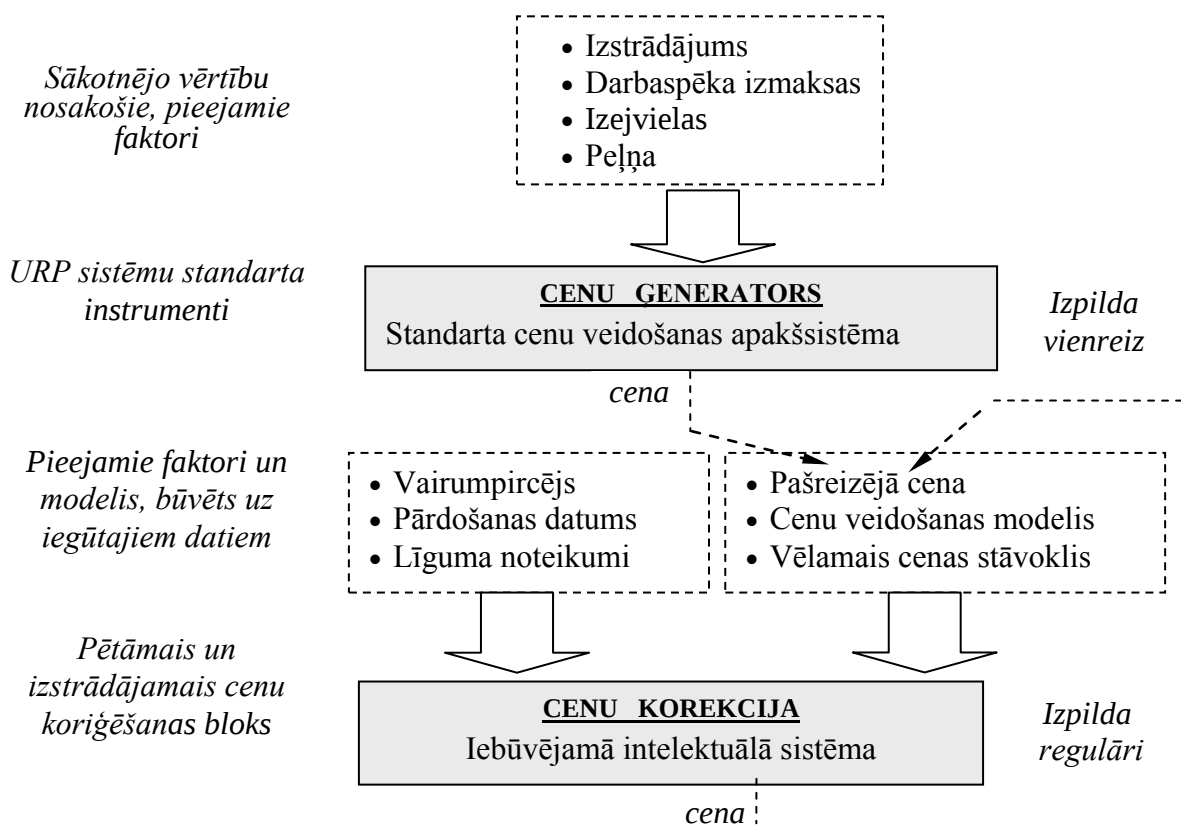
- laika dimensijas klātbūtne, kas nosaka iespēju attēlot datus kā laukrindas;
- vispārināšanas iespēja pēc novērojamo parametru virknes (piemēram, vairumtirdzniecības pircējs, produkcijas veids un citi) ar mērķi uzbūvēt modeli;
- iespēja aplūkot cenu korekcijas procesu kā sistēmu, kas katrā laika momentā atrodas zināmā stāvoklī, kurai piemīt vadības mehānismi (stāvokļu maiņas) un pastāv mērķa stāvokļa jēdziens.

Uzskaitītie iemesli svarīgi tādēļ, ka ļauj demonstrēt izstrādājamās Markova lēmumu pieņemšanas procesa dinamiskajā cenu vadīšanā pielietošanas pieejas īpatnības.

Visbeidzot, dinamiskās cenu politikas aktualitāte, ko diktē straujš interneta tehnoloģiju pieaugums mūsdienu biznesā, arī nosaka šī uzdevuma izvēli. Aplūkojamajā uzdevumā ziņu par pārdošanām avots, datu bāze un metožu izstrādes platforma ir resursu vadības sistēma „1C:Uzņēmums v7”.

Dinamiskās cenu politikas uzdevumam ir vairākas definīcijas. Šajā darbā tiek izmantota šāda: dinamiskā cenu veidošana ir dinamiska patērētāju cenu noregulēšana, atkarībā no vērtības, ar kādu šie pircēji korelē ar produkciju vai pakalpojumu [49] (definīcija, savukārt, pamatojas uz [56]). Kā *nozīmi*, ar kādu pircēji attiecas pret produkciju vai pakalpojumu, aplūkosim trīs cenu diferenciācijas formas, kas piedāvātas darbā [69].

Tiek piedāvāta lēmumu pieņemšanas atbalsta sistēma, kuras mērķis ir ilgtermiņa pārdošanu summu maksimizēšana, ko panāk pastāvošo cenu koriģēšanas ceļā, ņemot vērā pieejamos URP sistēmas faktorus (sk. 13. att.) un ekspertu noteikto mērķa stāvokli. Cenu korekcijas mehānisms tiek īstenots, pateicoties MLPP modeļa pielietošanai. Cenu attīstības modeļa uzbūve ietver pārdošanu procesa attīstības likumsakarību meklēšanu pagātnē un to apkopošanu (sk. 3. nodaļu). Citiem vārdiem sakot, lēmums noved pie pagātnes cenu izmaiņu analīzes un atbilstoša cenu attīstības procesa modeļa veidošanas.



13. att. Cenu ģenerēšanas moduļa un cenu koriģēšanas moduļa mijiedarbība

Kopējais vadāmo Markova procesu ar ienākumu diskontēšanu mērķis ir tāda sistēmas vadības vektora izvēle katrā solī, lai tā noteiktajā funkcionēšanas horizontā iegūtu maksimālo peļņu [80]. Pateicoties šai īpašībai, MLPP aparāts iederas dinamiskās cenu politikas uzdevumā.

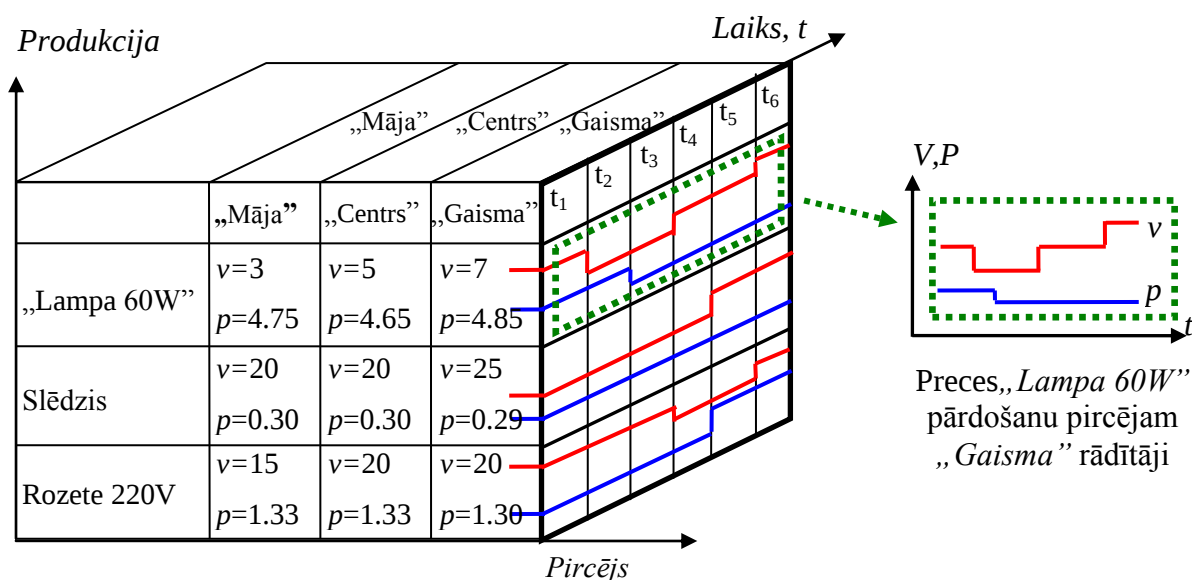
Šajā darbā sākotnēji paredzēts darbs ar daudzdimensiju datu telpu, kas iekļauj ziņas par vairumpircējiem un produkcijas nosaukumiem. Katra telpas dimensija noteikta vairāku simtu vērtību mērogā. Piesaistīti datu ieguves līdzekļi, lai atklātu likumsakarības, kas apraksta sekas, pie kādām novedušas darbības cenu politikā pagātnē. Realizēts programmas instruments (uz intelektuālā aģenta bāzes), lai reālā laikā izsekotu jaunās ziņas par pārdošanām, kas tiek saņemtas no menedžeriem un URP-sistēmas operatoriem. Tāda izsekošana ļauj savlaicīgi atjaunināt modeli, ievadot tajā jaunus datus par pārdošanām un veikto cenu korekciju sekām.

Datu modeļa apraksts.

Cenu korekcijas uzdevums ir diskrēts process, citiem vārdiem sakot, sistēmas „vairumpircējs-produkcija” uzvedība var būt izteikta ar galīgu stāvokļu skaitu. Jebkurā diskrētā laika momentā t_i sistēma atrodas vienā no iespējamiem stāvokļiem $s_j \in S$. Pārdošanu process laikā tiek novērots atsevišķu parametru (telpas dimensiju) lielumu fiksēšanai. Katru sistēmas stāvokli nosaka ar divām vektoru vērtībām (laikrindām):

- $\bar{p}$ - cenu laikrinda, atspoguļojoša cenu izmaiņu dinamiku;
- $\bar{v}$ - pārdošanu apjomu laikrinda, atspoguļojoša pārdošanu apjomu dinamiku.

Sistēmas piemērs (pa labi), kurā fiksēti lielumi ir „Gaisma”- dimensijai *Pircējs* un „Lampa 60W” dimensijai *Produkcija*, parādīts 14. attēlā. Sistēmas ietvaros novērojami rādītāji ir pārdošanu cena p_{ij} un pārdošanu apjomi v_{ij} . Attiecīgi, telpas dimensijas ir novērojamie mainīgie *Pircējs*, *Produkcija* un *Laiks*. Katrs hiperkuba (14. attēlā, pa kreisi) punkts noteikts ar diviem statistiskiem lielumiem: cenu un pārdošanu apjomu.



14. att. Daudzdimensiju datu kubs (pa kreisi) un stāvokļa laikrindas (pa labi)

Bez dimensijām „Produkcija” un „Pircējs” pastāv arī citas dimensijas. Šajā pētījumā pārējās dimensijas ir izslēgtas uzskatāmības nolūkā. Pieņemsim arī, ka dati jau izgājuši visus iepriekšējās apstrādes posmus. Datu notīrīšanas un to sagatavošanas procedūru apraksts izklāstīts darbā [4].

Cenu korekcijas uzdevums kā dinamiskās programmēšanas uzdevumi

Dinamiskās programmēšanas kontekstā uzdevumam tiek izvirzīts nosacījums, ka tā risināšanā jāizmanto zināma darbību secība. Citiem vārdiem sakot, risinājumam jābūt sadalītam apakšuzdevumos. Cenu korekcijas uzdevumam piemīt tāda īpašība. Piemēram, kādas produkcijas G_i cenu p var secīgi (noteiktā laika intervālā) novest līdz vēlajamai vērtībai bez būtiskas (uzdotā lieluma) noieta apjoma zaudēšanas. Tajā pašā laikā tieša cenas maiņa var novest pie vairumtirdzniecības pircēju zaudēšanas. Dinamiskās programmēšanas pielietošanu dinamiskās cenu politikas uzdevumā padara iederīgu vairāksolū cenu korekcijas pieeja.

Cenu korekcijas uzdevums MLPP terminos

Izteiksim iepriekš aprakstīto uzdevumu rekursīvā Belmana optimalitātes vienādojumā. MPD nostādnē vienādojums izsaka gaidāmā apbalvojuma lielumu, kas iegūts, pārejot no kāda pastāvošā stāvokļa s stāvoklī s' , saskaņā ar kādu politiku π [62]:

$$V^\pi(s) = R(s) + \gamma \sum_{s'} P(s'|s, \pi(s)) V^\pi(s'). \quad (5)$$

Optimālās politikas izteiksme pazīstama kā Belmana optimalitātes vienādojums MPD nostādnē. Tas apraksta pašreizējo apbalvojumu par veikto darbību, kas izraisa maksimālo sagaidāmo apbalvojumu *nākotnē* [62]:

$$V^*(s) = R(s) + \max_a \gamma \sum_{s'} P(s'|s, a) V^*(s'), \quad (6)$$

kur s – sistēmas stāvoklis, ko nosaka vektori $\bar{p}$ un $\bar{v}$: $s = \{ \bar{p}; \bar{v} \}$;
 $R(s)$ – pastiprinājums, kas iegūts pastāvošajā stāvoklī s .

Tāpat kā jebkurš stāvoklis, mērķa stāvoklis noteikts ar paša pastiprinājuma vērtību. Faktiski, mērķa stāvokļa būtību nosaka apbalvojumu funkcija. Vispārīgā gadījumā apbalvojuma funkcija nosaka, cik labi vai slikti ir atrasties kārtējā stāvoklī (analoģiski „labpatikai” vai „sāpēm bioloģiskā nostādnē”). Mērķis ir visu apbalvojumu summas maksimizēšana visa virzīšanās procesa uz mērķa stāvokli laikā. Cenu koriģēšanas uzdevuma gadījumā, pastiprinājums parāda lokālo summu, kas iegūta no pārdošanām, un zināmajam stāvoklim s tiek noteikta sekojošā veidā:

$$R(s) = \sum_{\tau=1}^{|s|} (p_{\tau} \cdot v_{\tau}), \quad (7)$$

kur $|s|$ - laikrindas lielums, kas raksturo stāvokli s ; tad apbalvojums ir visu dienu, kas novērojamas stāvokļa s ietvaros, ienākumu summa;

τ - pētāmās laikrindas laika mainīgais.

Turpināsim vienādojuma (6) mainīgo izskatīšanu: $P(s'|s, a)$ - apraksta sistēmas pāreju varbūtību stāvoklī s' no stāvokļa s , veicot darbības a (sk. formulu (3)).

Tā kā funkcijas $V^*(\cdot)$ izteiksmē (6) pastāv gan kreisajā, gan labajā daļā, tad aprēķins tiek veikts rekurentā veidā, tas ir, visa uzdevuma dekompozīcijas apakšuzdevumos ceļā. Vienādojuma (5) risināšanai pastāv divi pamatalgoritmi: *Value Iteration* un *Policy Iteration* [62].

Šajā nodaļā noformulēts dinamiskās cenu politikas uzdevums, izmantojot Markova lēmumu pieņemšanas procesa terminus. Piedāvāta modeļa, aplūkojama kā vides, kurā funkcionē MLPP aparāts, uzbūvēšanas metode. Metode balstīta uz likumsakarību atrašanu un to apkopošanu. Elementāras sistēmas ietvaros, kas satur vienu pircēju un vienu produkciju, nodemonstrēta modeļa uzbūves metodoloģija. Izpildīto uzdevumu analīze ļauj noformulēt sekojošus secinājumus:

1. dinamiska sistēma, kuras attīstība attēlota ar laikrindām, var tikt izteikta ar galīgu stāvokļu un darbību skaitu, tai var būt izveidoti pāreju modelis un apbalvojumu funkcija;
2. lēmumu pieņemšanas atbalsta moduļa un sistēmas „*vairumpircējs-produkcija*” mijiedarbība, kā arī nepārtrauktas modeļa atjaunināšanas nepieciešamība nosaka uz aģentu orientētas pieejas pielietojumu.

Piektā nodaļa (Modeļu izveides eksperimentu veikšana dinamiskās cenu politikas uzdevumā)

Piedāvātās modeļa uzbūves metodes aprobēšanas nolūkā šajā nodaļā aprakstīta programmatūras platformas izstrāde eksperimentu veikšanai, sastādīts eksperimentu plāns un aprakstīts to veikšanas process.

Programmnodrošinājuma izstrāde eksperimentu veikšanai

Programmatūras moduļi eksperimentu veikšanai īstenoti, izmantojot platformu „1C:Uzņēmums v7.7”. Izvēli nosaka tas, ka šai platformai pieejama datu bāze par reāla uzņēmuma produkcijas pārdošanām divu gadu laikā. Otrkārt, realizētie moduļi var kalpot par pamatu gala programmatūras nodrošinājuma radīšanai, kā uzdevums būtu funkcionēt fona režīmā un īstenot priekšlikumu izstrādi cenu korekcijai.

Bez tam, realizēti moduļi ar priekšmeta apgabalu tiešā veidā nesaistītu eksperimentu veikšanai. Tās ir programmas darbam ar mākslīgo neironu tīklu un Markova lēmumu pieņemšanas procesu modeļu uzdevumos.

Izstrādes vides Delphi 7 piesaistīšana paredzēta, lai nodrošinātu un paātrinātu datu masīvu apstrādi, kuru apjoms pārsniedz platformas „1C:Uzņēmums v7” iespējas. Tā, piemēram, distanču matricas veidošanu ierobežo 5000 vienības katrā dimensijā.

Eksperimentu plāns

Izstrādātās MLPP modeļa uzbūves metodes darbības un efektivitātes novērtēšanai izstrādāts eksperimentu plāns (sk. 3. tabulā). Plānā iekļautas divas eksperimentu sērijas. Modeļu un faktisko procesu salīdzināšanas ceļā pirmā eksperimentu sērija ļauj novērtēt, kādā mērā MLPP modelis atbilst apmācību datiem. Otrās eksperimentu sērijas mērķis ir MLPP modeļa, kas uzbūvēts telpas aproksimācijas ceļā, izmantojot mākslīgo neironu tīklu, kvalitātes novērtēšana.

Intelektuālā aģenta eksperimentu plāns MLPP modeļa uzbūvei un pielietošanai cenu korekcijas uzdevumā

Sērija	Eksperimenta apraksts	Mērķis
I.	izmantojot skaitliskus rādītājus, iegūt MLPP modeļa un faktisko procesu līdzības novērtējumu	novērtēt MLPP modeļa uzbūves algoritmu korektumu
	izmantojot skaitliskus rādītājus, salīdzināt MLPP modeļa un faktisko lēmumu efektivitāti apmācības procesā neiesaistītu datu apgabalā	novērtēt MLPP efektivitāti ekspluatācijas režīmā
II.	izmantojot skaitliskus rādītājus, veikt uzbūvētā MLPP+MNT modeļa līdzības novērtējumu attiecībā pret faktiskajiem procesiem	novērtēt MLPP modeļa uzbūves efektivitāti aproksimētā stāvokļu telpā

Kā modeļu kvalitātes novērtējuma kritērijs kalpo veiksmīgi uzmodelēto pāreju skaita attiecība pret faktisko pāreju skaitu, kā arī peļņas novērtējums, kas izteikts nosacītās naudas vienībās.

Cenu politikas modeļa uzbūvēšana un pielietošana

Šajā eksperimentā cenu politikas modeļa uzbūvei tiek izmantoti reāli dati par Latvijas pārtikas rūpniecības uzņēmuma saražotās produkcijas pārdošanām. Novērojamo datu periods ietver trīs mēnešus: maiju, jūniju, jūliju. Ziņas par pārdošanām attēlotas, izmantojot URP sistēmas „1C:Uzņēmums v7” elektroniskos dokumentus. Katrs elektroniskais dokuments satur visu nepieciešamo informāciju par darījumu, tajā skaitā darījuma datumu, vairumtirdzniecības pircēja nosaukumu, produkcijas uzskaitījumu, pārdošanu apjomus, cenas un pārējo. Pētāmā perioda ietvaros klātesoši apmēram 28,5 tūkstoši pārdošanu dokumentu, 1 725 vairumtirdzniecības pircēju un 476 realizējamās produkcijas nosaukumi.

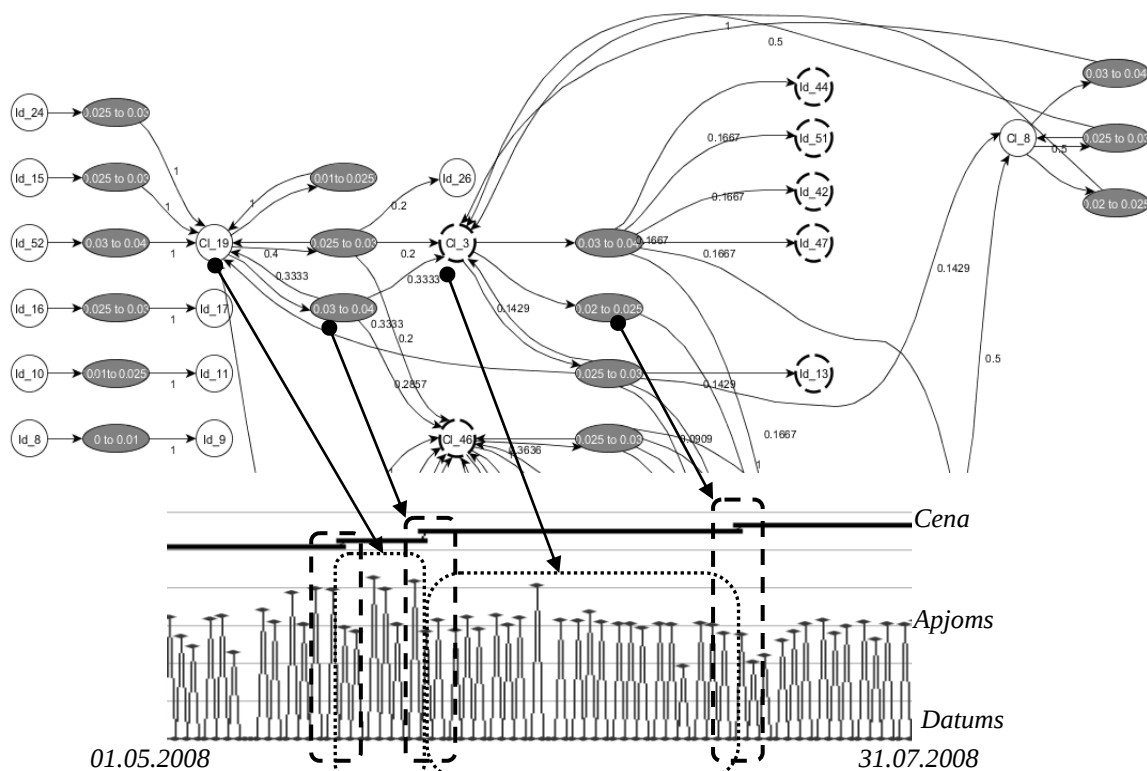
MLPP modeļa pāreju fragments

Izejas stāvoklis s	Darbība a	Pārejas stāvoklis s'	Pārejas varbūtība $P(s' s, a)$
Id_7	$\tilde{p}_2 = (-0.05, -0.0009]$	Cl_11	1,0
Cl_66	$\tilde{p}_2 = (-0.05, -0.0009]$	Cl_54	0,25
Cl_3	$\tilde{p}_1 = (-0.1, -0.05]$	Id_11	0,667
Id_14	$\tilde{p}_1 = (-0.1, -0.05]$	Id_15	0,4
Id_16	$\tilde{p}_3 = (0.0, 0.01]$	Id_17	0,667
...	...	...	...

Iegūtais pāreju modelis (sk. 4. tabulā) arī ir Markova lēmumu pieņemšanas process. Kopējais pāreju skaits sastādīja 1192, ko izveidoja 294 stāvokļi.

Ar grafiskās paketes „yEd Graph Editor” iespējams vizuāls modeļa attēlojums. Pamatojoties uz eksperimentālās platformas uzģenerēto XML faila modeli 15. att., augšā parādīts modeļa fragments.

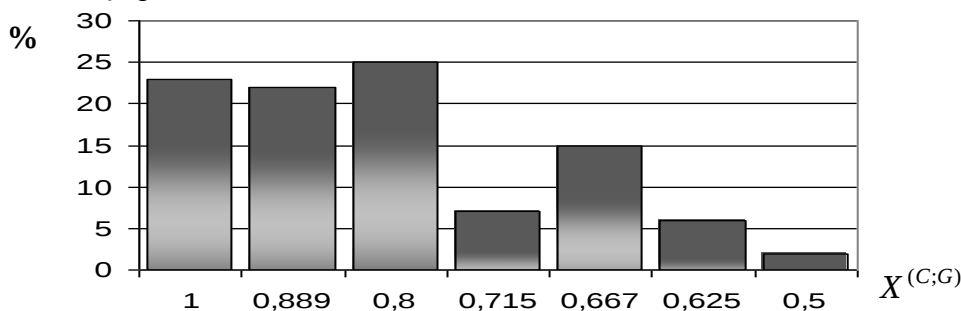
Gaišie ovāli apzīmē stāvokļus, tumšie – darbības, vērtības virs šķautnēm – atbilstošu pārejas varbūtību. Iegūtais grafs pārstāv cenu modeli noteiktā laika periodā.



15. att. Pāreju grafa fragmenta (augšā) atbilstība pārdošanas apjomu un cenu dinamikas fiksētai kombinācijai “vairumtirdzniecības pircējs – produkcija”

Pārdošanu apjomu un cenu attīstības fragments zināmai novērojamai vairumtirdzniecības pircēja un produkcijas kombinācijai parādīts 15. att. (apakšā). Ar punktēto līniju apvilkti stāvokļi, ko nepavada nekādas cenu izmaiņas (citiem vārdiem sakot, pārejas ar nulles cenas lielumu). Ar raustīto līniju apzīmētas pārejas, ko radījušas noteiktas cenu izmaiņas. Taisnās bultiņas norāda katra grafa elementa stāvokļus reālā pārdošanu procesa fragmentā.

Laikrindu skaits (procentuāli no kopējā skaita - 5000), kam piemīt atbilstoša skaitlisko lielumu vērtība $X^{(C;G)}$ parādīts 16. attēlā. Uzreiz jāatzīmē, ka izpaliek lielumi ar vērtību, kas mazāka par 0.5, tas nozīmē, ka visas laikrindas, vismaz pa pusei aprakstījis modelis un tām piemīt daļējs notikumu secīgums. 97% gadījumu zema (mazāka par 0.9) vērtība saistīta ar modeļa pārtrauktību (šo efektu var novērot 15. att., augšā, viensoļa izolētām pārejām). Tomēr, novērtējums $X^{(C;G)} > 0.8$ raksturīgs apmēram 70% novērojumu, kas raksturo modeli kā spējīgu reproducēt lielāko daļu procesa.



16. att. Vērtības $X^{(C;G)}$ sadalījums

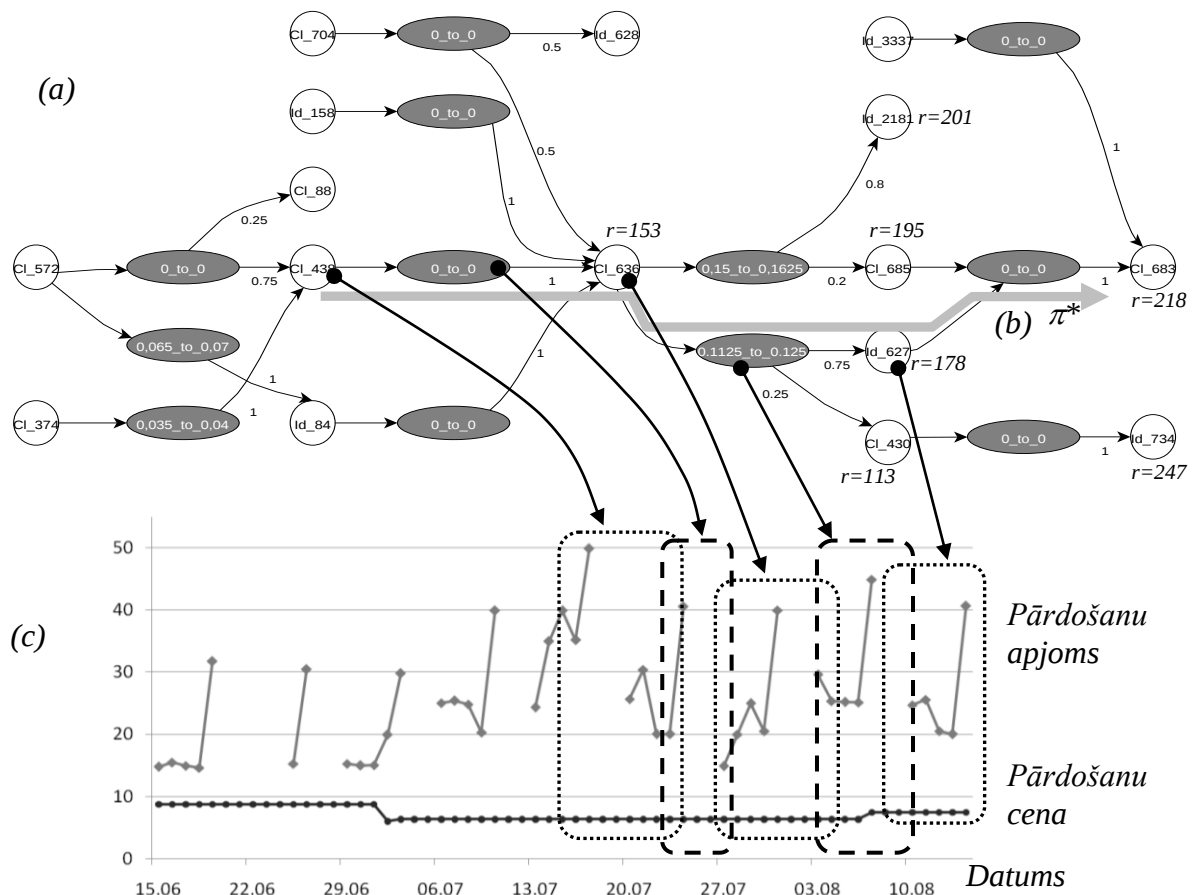
Ja ir pāreju modelis, kļūst iespējams izmantot labi izstrādātus politikas meklēšanas algoritmus Markova modelī (*Policy iteration*, *Value Iteration*), kam nepieciešams noteikt mērķa stāvokli. Kopumā modeļa pielietošana pētāmajā procesā ietver sekojošus etapus: (a) salīdzināt pašreizējo procesa stāvokli ar vienu no modeļa stāvokļiem, kas noteiks sākotnējo stāvokli s_0 ;

(b) noteikt mērķa stāvokli s_{target} ; (c) veidot cenu korekcijas politiku π^* ; (d) procesa attīstības gaitā uzraudzīt pašreizējo stāvokli un, nepieciešamības gadījumā koriģēt politiku.

Uzsāktās cenas izmaiņas stājas spēkā, un pēc noteikta laika perioda sistēma pāriet citā stāvoklī s_{t+1} . Pāreju raksturo izmainīts cenas lielums, kā arī pieprasījuma reakcija. Tāda lēmuma pieņemšanas moduļa un ārējās sistēmas „vairumtirdzniecības pircējs-produkcija” mijiedarbība pārstāv klasisku intelektuālās sistēmas un priekšmeta apgabala vides mijiedarbības klasisku shēmu.

Politikas pielietošana testu datiem

Šī eksperimenta mērķis ir izvērtēt iespēju pielietot modeli datiem, kas nebija iekļauti apmācāmajā datu kopā. Faktiski eksperiments atspoguļo sākotnējo uzdevumu, kas tiek izvirzīts intelektuālajai sistēmai. Atšķirībā no pirmā eksperimenta, šis eksperiments paredz fāzi optimālās politikas π^* izpildei. Tas nozīmē, ka kļūst iespējams skaitliskā veidā novērtēt modeļa efektivitāti peļņas aspektā. Aprēķinot politiku π^* , ar mērķa stāvokļiem saprot visus iespējamus stāvokļus. Stratēģija tiek aprēķināta, pamatojoties uz iegūstamo peļņu (7) un atbilstošām darbībām, kas nodrošina visu procesa attīstības variantu apskatu un optimālā varianta izvēli diskontētās peļņas ziņā. Kā zīmīgāks politikas pielietošanas gadījums datiem ārpus apmācāmās izlases kalpo 17. attēlā dotais piemērs. Saskaņā ar uzbūvēto MLPP modeli, pārejot no stāvokļa Cl_438 stāvoklī Cl_636 , klātesošas divas alternatīvas: vai nu uzsākt darbības cenas koriģēšanai uz lielumu $\tilde{p}_{27} = (0.15, 0.1625]$, vai arī darbības cenas koriģēšanai uz lielumu $\tilde{p}_{24} = (0.1125, 0.125]$.



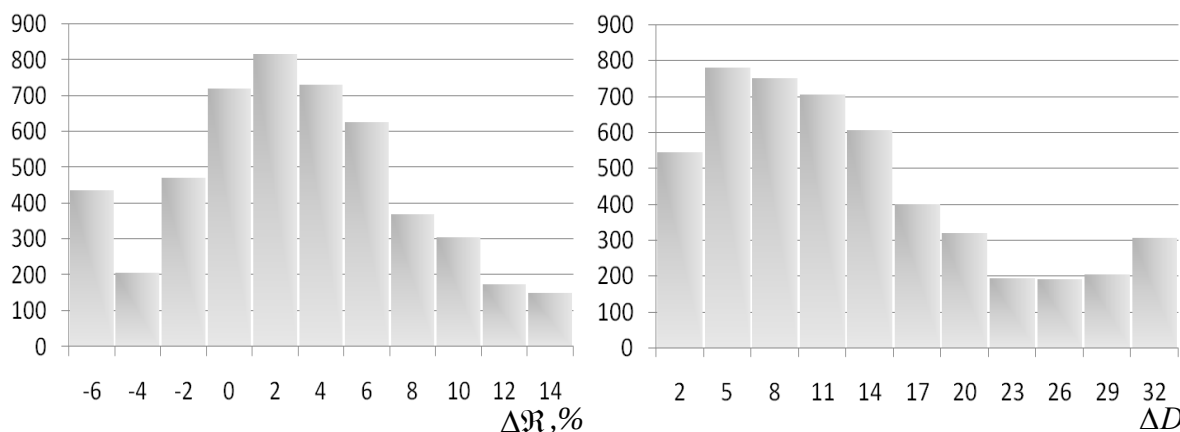
17. att. Pāreju grafa fragmenta (a) un optimālā politika π^* (b) saistība ar pārdošanu apjomu un cenu dinamiku uzdotai *pircējs-produkcija* kombinācijai

Atkarībā no izvēlētajās darbības, terminālie stāvokļi var būt stāvokļi Id_2181 , Cl_683 , Id_734 . Tā kā lokālā apbalvojuma vērtības ir zināmas (summārais ienākums stāvokļa iecirknī),

kļūst iespējams aprēķināt terminālo stāvokļu sasniegšanas politikas un izvēlēties optimālo politiku.

Neskatoties uz to, ka terminālajam stāvoklim CI_{683} nepiemīt maksimālā lokālā apbalvojuma vērtība, tas ir „pievilcīgāks” diskontētā apbalvojuma ziņā, sastādot lielumu $V(CI_{683} | \tilde{p}_{24}) = 405,585$. Maksimālā lieluma vērtība nosaka optimālo politiku π^* , kas attēlota 17. ar pelēku treknu bultiņu.

Galarezultātā precizitāte atkarīga no tā, kādas vispārējās pārdošanu procesa attīstības iezīmes tika noteiktas, pamatojoties uz atsevišķiem novērojamiem MLPP modeļa uzbūves gadījumiem. Lai novērtētu aģenta funkcionēšanas efektivitāti 5000 kombināciju ietvaros, 18. parādītas histogrammas, kas atspoguļo kombināciju sadalījumu pēc atbilstošām lielumu vērtībām modelējamā procesa salīdzināšanā ar faktisko procesu.



18. att. Kombināciju *pircējs-produkcija* skaita sadalījums pēc peļņas (pa kreisi), kombināciju *pircējs-produkcija* skaita sadalījums pēc distances (pa labi)

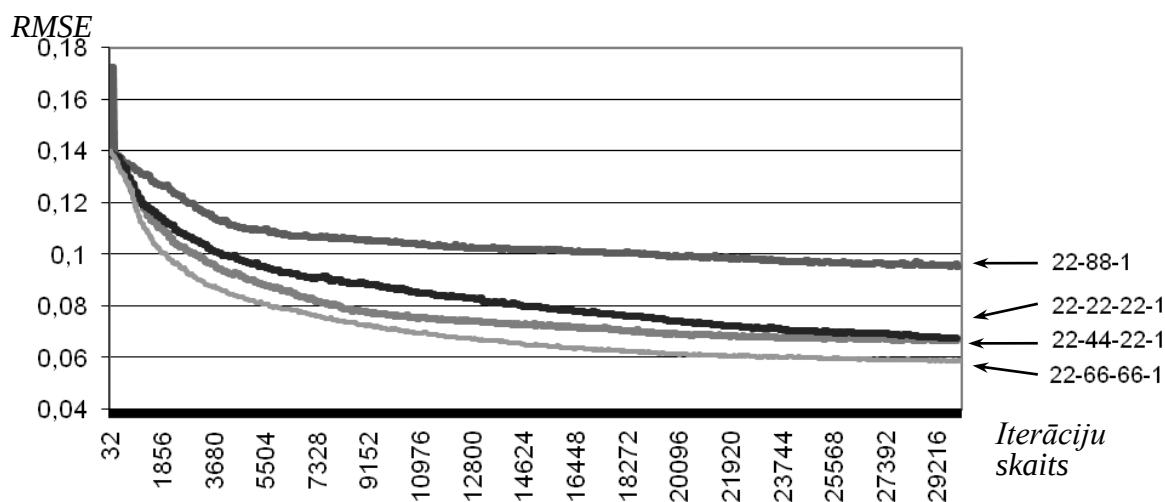
No 18. attēla var secināt, ka modelī pārsvarā ir kombinācijas (77,8%), ko cenu korekcija noved pie pozitīviem rezultātiem (t.i., peļņas vērtība lielāka par nulli). Savukārt negatīva lieluma ΔR vērtība nozīmē, ka modeļa piedāvātās cenu korekcijas izrādījušās mazāk efektīvas nekā eksperta lēmumi. Kombinācijas *pircējs-produkcija* esamība, kurai cenu korekcija rada zaudējumus, izskaidrojama tādējādi, ka tādām kombinācijām netika atrasts pietiekams skaits atsevišķu novērojumu, lai uzbūvētu pāreju grafu.

Telpas MLPP aproksimācijas ar MNT palīdzību eksperimenti

Šajā eksperimentā tiek īstenots praktisks pētījums, lai novērtētu telpas aproksimācijas iespējas dinamiskās cenu politikas uzdevumā, izmantojot mākslīgos neironu tīklus. Detalizēta informācija par metodi un tās pielietojumu kalnu automašīnas modeļu uzdevumā dota promocijas darba 2.3. apakšnodaļā.

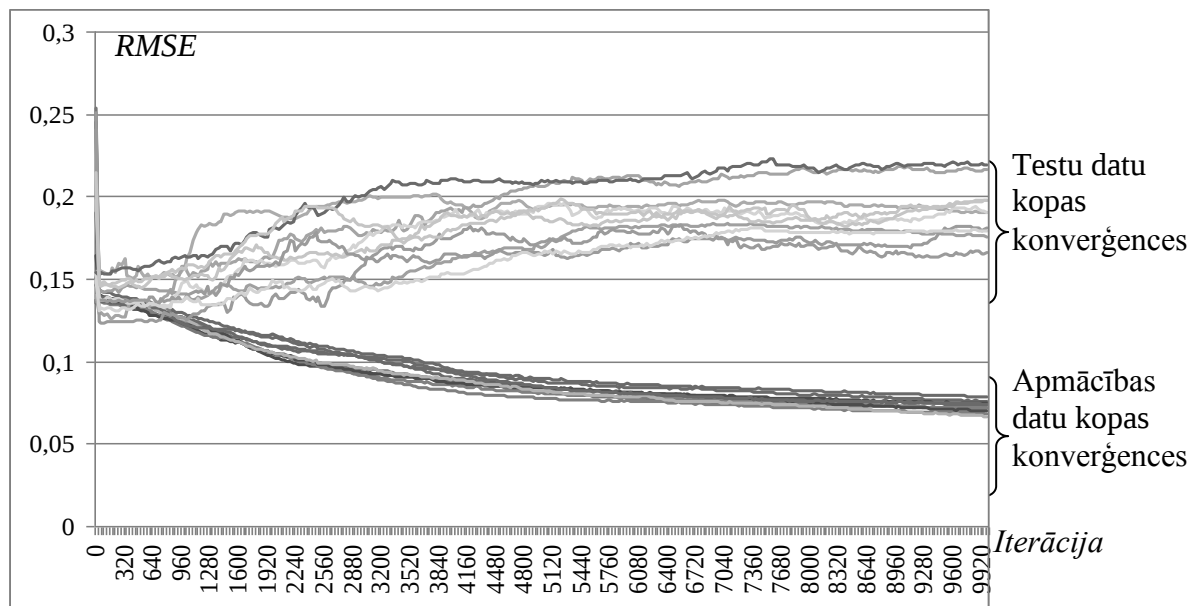
Pāreju telpas aproksimācija vajadzīga, lai iegūtu pāreju varbūtību stāvokļiem, kuros sistēma nekad nav atradusies, bet potenciāli varētu tajos nonākt. Tādā gadījumā ir lietderīgi, ja ir vismaz pāreju varbūtības aprēķinātais lielums.

Konverģences grafiki dažādam skaitam slēpto slāņu un neironu MNT slēptajā slānī parādīti 19. attēlā. Kā „visātrāko” (ieguldīto apmācības iterāciju skaita ziņā) tīkla konfigurāciju var minēt konfigurāciju „22-66-66-1” (divi slēptie slāņi pa 66 neironiem katrā), kā „vislēnāko” konfigurāciju „22-88-1”. Ja ņem vērā vienai apmācības iterācijai patērēto laiku, tad „ātra” izrādīsies konfigurācija „22-22-22-1”. Tā ļauj sasniegt rezultātu, kuru var salīdzināt ar konfigurāciju „22-66-66-1” ar lielāku skaitu iterāciju, bet īsākā laikā. Tādēļ turpmākajos eksperimentos tiks izmantota konfigurācija „22-22-22-1”.



19. att. Konverģence pie MNT dažādu konfigurāciju

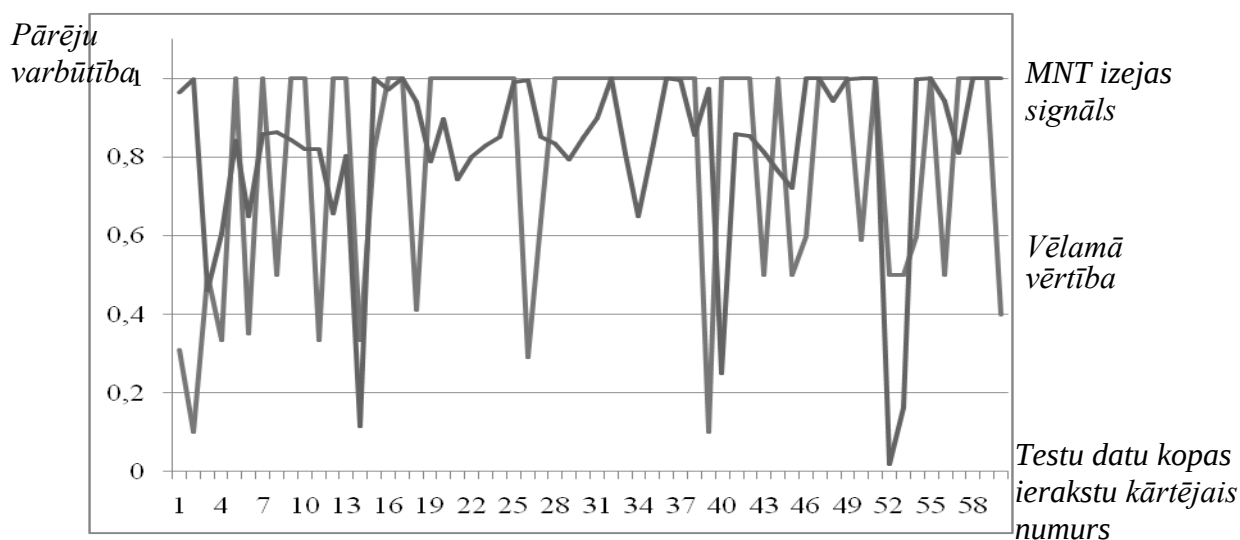
Aproximācijas kvalitātes novērtēšanai veiksīm šķērsvalidāciju. Šajā nolūkā visa apmācāmā izlase tiek sadalīta 10 blokos. Katrs bloks veidojas, samazinot ierakstus, kas izņemti no apmācāmās izlases pēc noteikta intervāla. Konverģences grafiki 10 šķērspārbaudēm parādīti 20. attēlā. Visas pārbaudes tiek īstenotas tīklā, kura konfigurācija „22-22-22-1”.



20. att. Šķērsvalidācijas konverģences grafiki

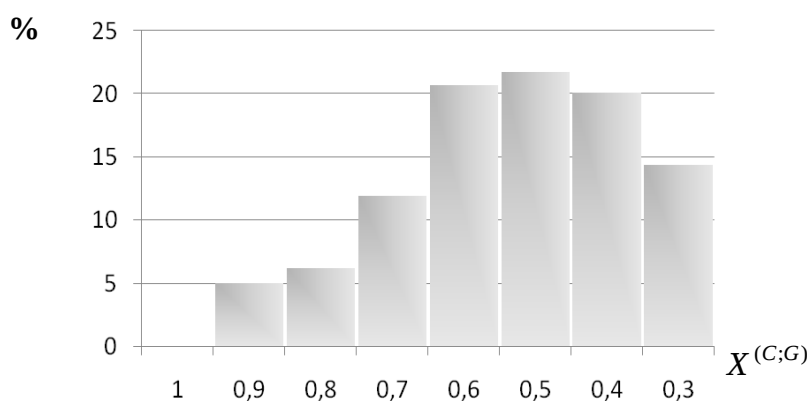
Jāatzīmē, ka testa izlases konverģencei nepiemīt asimptomātisks tuvinājums nullei (kā apmācāmai izlasei), bet tas noturas noteiktā līmenī. Vienlaicīgi, testa datu kopas vidējās kvadrātiskās kļūdas lielums (*rmse*), atrodas tajā pašā lielumu mērogā, kur apmācāmās izlases kļūda. Aplūkojot vienu no kontrolpārbaudžu bloku aproksimācijas rezultātu (sk. 21. att.), jāatzīmē, ka tīklam izdodas atkārtot galvenās testa funkcijas iezīmes (kļūda *rmse* sastāda lielumu 0,17).

Tāds kļūdas lielums ļauj izmantot aproksimēto modeli darbību politikas uzbūvē. Tagad politikas uzbūves algoritms pievēršas nevis pāreju varbūtības funkcijas attēlošanai tabulas veidā, bet tās aproksimācijai, ko veic, izmantojot MNT palīdzību. Lai novērtētu precizitāti, salīdzināmu ar iepriekšējo eksperimentu vērtībām, tiek izmantota izteiksme, kas dota promocijas darbā (sk. formulu (5.2)). Šī izteiksme ļauj katrai kombinācijai *vairumtirdzniecības pircējs-produkcija* ($C;G$) aprēķināt skaitlisko modeļa (vai tā atsevišķu daļu) atbilstības lielumu $X^{(C;G)}$ faktiskajām pārejām.



21. att. Testa datu kopas aproksimācijas fragments vienam atribūtam

Laikrindu skaits (procentos no kopējā skaita - 5000), kam piemīt atbilstošas skaitlisko lielumu vērtības $X^{(C;G)}$, parādīts 22. attēlā. Salīdzinājumā ar 16. attēlā parādītajiem rezultātiem, aproksimētajam modelim piemīt mazāka precizitāte. Svarīga pieejas priekšrocība ir tā, ka rodas iespēja pieņemt lēmumus stāvokļos, kuros sistēma nekad nav atradusies.



22. att. Aproksimētā modeļa vērtības lieluma $X^{(C;G)}$ sadalījums

Praksē, ņemot vērā relatīvi lielo kļūdu, tādi lēmumi var tikt piedāvāti izskatīšanai ekspertam, lai aplūkotu iespējamās cenu korekcijas stratēģijas. Viens no precizitātes paaugstināšanas veidiem ir katra stāvokļa aprakstošo neatkarīgo mainīgo skaita palielināšana.

Sestā nodaļa (rezultātu analīze un secinājumi)

Parādīts, ka dinamiska sistēma, ko raksturo laikrindu esamība, var tikt izteikta ar galīgu stāvokļu un darbību skaitu, tai var būt pāreju modelis un apbalvojumu funkcija. Galvenā uzmanība veltīta modeļa, kas tiek uzskatīts par vidi (kurā funkcionē MLPP aparāts), uzbūves metodei. Metode balstīta uz novērojamo lielumu attīstības likumsakarību atrašanu laikā un atklāto likumsakarību apkopošanu. Pielietotā uz aģentu orientētā arhitektūra nodrošinājusi tādu MLPP modeļa un vides mijiedarbību, ka var īstenot nepārtrauktu MLPP modeļa apmācību un politikas pārveidošanu dinamiskā vidē. Piedāvātā eksperimentālā platforma ļāva aprobēt izstrādāto modeļa uzbūves metodi reāla uzdevuma apstākļos. Platforma ļauj ekspertam izvēlēties vēlamā mērķa stāvokli. Pretējā gadījumā, kā mērķa stāvoklis var kalpot jebkurš stāvoklis no stāvokļu apakškopas ar maksimālu apbalvojuma lielumu vai jebkurš cits termināls stāvoklis, kura sasniegšana dod maksimālos diskontētos ienākumus.

Reālas sistēmas ietvaros, kas ietver simtiem pircēju un produkcijas nosaukumu, nodemonstrēta pilnībā automatizēta cenu attīstības modeļa uzbūves metode. Modeļa validācija parādīja pieņemamo precizitāti. Precizitātes paaugstināšanai nepieciešams veikt papildus eksperimentus: parametru meklēšana, datu struktūras pārskatīšana, kā arī atsevišķu algoritmu pilnveidošana, u.c. Šī promocijas darba pamatrezultāti ir šādi:

- izpētīts Markova modeļa uzbūves problēmas tekošais stāvoklis dinamiskās sistēmās, attēlotu ar laikrindām novērojamo mainīgo daudzdimensiju telpā;
- izpētītas tādas skaitļošanas intelekta nozares kā aģentu intelektuālās sistēmas, datu ieguves procedūras, mākslīgie neironu tīkli un citas, lai izstrādātu MLPP modeļa uzbūves metodi dinamiskā sistēmā un pielietotu to testa datu kopai;
- izstrādāta un aprobēta jauna vairāksoļu MLPP modeļa uzbūves pieeja, izmantojot mākslīgā neironu tīkla palīdzību, stāvokļu vērtības tabulu aproksimācijas gadījumā;
- izstrādāta MLPP modeļa uzbūves pieeja, kas balstīta uz novērojamo mainīgo attīstības procesu likumsakarību meklēšanu un to pārveidošanu stāvokļu kopā un pāreju varbūtībās;
- piedāvātas pieejas reāla uzdevuma datu pārveidošanai struktūrā, kas atbilst MLPP aparātam: izstrādātas metodes un datu struktūras (novērojamo mainīgo uzvedības profils), lai pārveidotu daudzdimensiju laikrindas Markova procesa stāvokļos, piedāvātas darbību kopas veidošanas metodes, aprakstīta mērķa stāvokļu meklēšanas metode;
- piedāvātās MLPP modeļa uzbūves metodes aprobēšanai izstrādāta programmas eksperimentālā platforma, kā arī attiecīgu programmatūras rīku virkne;
- eksperimentu gaitā, pamatojoties uz reālām ziņām par pārdošanām, iegūti skaitliski MLPP modeļa tuvuma novērtējumi faktiskajai pētāmo procesu attīstībai, kā arī aģentu sistēmas funkcionēšanas novērtējumi testu datu kopai.

PROMOCIJAS DARBA GALVENIE REZULTĀTI

Promocijas darba ietvaros izstrādāta uz MLPP balstīta lēmumu pieņemšanas pieeja, kas nodrošina MLPP modeļa uzbūvi uzdevumos, kuru dati izteikti ar daudzdimensiju laikrindām. Veicot eksperimentu virkni, īstenota pieejas aprobācija. Rezultātā iegūti skaitliski lielumi, kas ļauj secināt, ka pieeja spēj būvēt MLPP modeli, adekvāti atspoguļojošu apmācības datu kopu. Atrisināti sekojoši uzdevumi un iegūti rezultāti.

1. Uz Markova procesiem balstītu matemātisko aparātu apskats ļauj secināt par iespējamību uzskatīt MLPP un apmācību ar pastiprināšanu par efektīvām dinamisko sistēmu modelēšanas pieejām; noteiktas galvenās to pielietošanas problēmas uzdevumos, kuru dati izteikti ar daudzdimensiju laikrindām.
2. Atsevišķu skaitļošanas intelekta metožu (MNT, AP, datu ieguves metodes un citas) analīze ļāva aprakstīt izstrādājamās pieejas galvenās iezīmes (plūsmas konveijera veida organizācija, lai pārveidotu datus, savlaicīga MLPP modeļa atjaunināšana un tā tālāk).
3. Izstrādāta speciāla aģentu arhitektūra, kas ļauj izvairīties no nekorekta intelektuālās sistēmas un vides mijiedarbības apraksta.
4. Realizēta stāvokļu telpas un MLPP lēmumu aproksimācija, izmantojot mākslīgā neironu tīkla palīdzību. Ar modeļa uzdevumu nodemonstrēta pieejas darbība.
5. Noformulēta starpstruktūra (pētāmā mainīgā uzvedības profils) pētāmo laikrindu iegūto likumsakarību uzglabāšanai un apstrādei. Pētāmā mainīgā uzvedības profili tiek izmantoti MLPP modeļa uzbūvei.
6. Piedāvāta dažādu laikrindu klasterizācijas kritēriju pielietošanas pieeja (Eiklīda attālums un kontūru līdzība), atkarībā no katra pētāmā mainīgā jēdzieniskās noslodzes.
7. Pieejas aprobēšanai noformulēts dinamiskās cenu politikas uzdevums MLPP ietvaros, kā arī izstrādāts programmatūras nodrošinājums eksperimentu veikšanai.
8. Veikta eksperimentu sērija, kas ļauj skaitliski novērtēt MLPP modeļa uzbūvēšanas efektivitāti. Novērtējums balstīts uz iegūtā modeļa un apmācības datu kopas salīdzinājumu, kā arī modeļa pielietošanu datiem ārpus apmācības datu kopas.

IZMANTOTĀS LITERATŪRAS SARAKSTS

1. Carkova V., Šadurskis K. Gadījuma procesi. – Rīga : RTU, 2005. – 138 lpp.
2. Čižovs J., Borisovs A. Markov Decision Process in the Problem of Dynamic Pricing Policy // Automatic Control and Computer Sciences. - No. 6, Vol 45. (2011), pp 77-90.
➤ indexed in: SpringerLink, Ulrich's I.P.D., VINITI.
3. Čižovs J., Borisovs A., Zmanovska T. Ambiguous States Determination in Non-Markovian Environments // RTU zinātniskie raksti. 5. sēr., Datorzinātne. - 36. sēj. (2008), 140.-147. lpp.
➤ indexed in: EBSCO
4. Chizhov Y., Zmanovska T., Borisov A. Temporal Data Mining for Identifying Customer Behaviour Patterns // Workshop Proceedings Data Mining in Marketing DMM' 2009, 9th Industrial Conference, ICDM 2009, Leipzig, Germany, 22-24 July, 2009.– IBAI Publishing, 2009. – P. 22-32.
➤ indexed in: DBLP, Io-port.net.
5. Chizhov Y. An Agent-Based Approach to the Dynamic Price Problem // Proceedings of 5th International KES Symposium Agent and Multi-agent Systems, Agent-Based Optimization KES-AMSTA/ABO'2011, Manchester, U.K., June 29-July 1, 2011.– Heidelberg: Springer-Verlag Berlin, 2011. – P. 446-455.
➤ indexed in: SpringerLink, Scopus, ACM DL, DBLP, Io-Port.
6. Chizhov Y., Kuleshova G., Borisov A. Manufacturer – Wholesaler System Study Based on Markov Decision Process // Proceedings of 9th International Conference on Application of Fuzzy Systems and Soft Computing, ICAFS 2010, Prague, Czech Republic, August 26-27, 2010.– b-Quadrat Verlag, 2010. – P. 79-89.
7. Chizhov Y., Kuleshova G., Borisov A. Time series clustering approach for decision support // Polish Journal of Environmental Studies. – Vol.18, N4A (16th International Multi-Conference ACS-AISBIS, Miedzyzdroje, Poland, 16-18 October, 2009), pp.12-17.
➤ indexed in: Scopus, Web of Science
8. Chizhov J., Borisov A. Applying Q-Learning To Non-Markovian Environments // Proceedings of the International Conference on Agents and Artificial Intelligence (ICAART 2009), Porto, Portugal, January 19 - 21, 2009. – INSTICC Press, 2009.- P. 306-311.
➤ indexed in: Engineering Village2, ISI WEB of KNOWLEDGE, SCOPUS, DBLP, Io-port.net.
9. Chizhov Y. Particulars of Neural Networks applying in Reinforcement Learning // Proceedings of 14th International Conference on Soft Computing „MENDEL 2008”, Czech Republic, Brno, 18.-20. June, 2008. – Brno: BUT, 2008. – P. 154-160.
➤ indexed in: ISI Web of Knowledge, INSPEC
10. Chizhov Y. Reinforcement learning with function approximation: survey and practice experience // Proceedings of International Conference on Modelling of Business, Industrial and Transport Systems, Latvija, Rīga, May 7-10, 2008.- Riga: TSI, 2008.- P. 204-210.
➤ indexed in: ISI Web of Knowledge
11. Chizhov J. Software agent developing: a practical experience, Scientific proceedings of Riga Technical University: RTU 48. rakstu krājums, 5. sērija, 31. sējums, 12 October, 2007, Riga Technical University, Riga, Latvia.
12. Chizhov J., Borisov A. Increasing the effectiveness of reinforcement learning by modifying the procedure of Q-table values update // Proceedings of Fourth International Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control „ICSCCW-2007”, Antalya, Turkey, 27-28 August 2007.– b-Quadrat Verlag, 2007.– P. 19-27.

13. Chizhov J., Agent Control In World With Non-Markovian Property // Poster presentation in EWSCS'07: Estonian Winter School in Computer Science, Palmse, Estonia, March 4-09, 2007.
14. Athanasiadis I.N., Mitkas P.A. An agent-based intelligent environmental monitoring system // Management of Environmental Quality.–Vol.15 (2004), P. 229-237.
15. Baxter J., Bartlett P.L. Infinite-horizon policy-gradient estimation // Journal of Artificial Intelligence Research. –Vol.15 (2001), P. 319–350.
16. Beitelspacher J., Fager J., Henriques G., ...[etc]. Policy Gradient vs. Value Function Approximation: A Reinforcement Learning Shootout. Technical Report No.CS-TR-06-001, School of Computer Science University of Oklahoma Norman, OK 73019, 2006.
17. Bellman R. Dynamic Programming. – New Jersey: Princeton University Press, 1957.
18. Bertsekas D.P., Tsitsiklis J. Neuro-Dynamic Programming.-Athena Scientific, 1996.– 512 p.
19. Butz M.V. Rule-based evolutionary online learning systems: learning bounds, classification, and prediction. – Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Computer Science in the Graduate College of the University of Illinois at Urbana-Champaign, 2004.
20. Carreras M. and other. Application of SONQL for real-time learning of robot behaviours // Robotics and Autonomous systems. – Vol. 55, Issue 8 [2007], P. 628-642.
21. Cervenka R., Trencansky I. AML. The Agent Modeling Language. A comprehensive Approach to Modeling MAS. – Berlin: Springer, 2007. – 355 p.
22. Chakraborty D., Stone P. Online Model Learning in Adversarial Markov Decision Processes // Proceedings of 9th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2010), Toronto, Canada, May 10-14, 2010. – International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 2010. –P. 1583-1584.
23. Cheung T., Okamoto K., Maker F., ... [etc]. Markov Decision Process Framework for Optimizing Software on Mobile Phones // Proceedings of the 9th ACM IEEE International conference on Embedded software, EMSOFT 2009, Grenoble, France, October 12-16, 2009. – New York: ACM, 2009. – P. 11-20.
24. Cotofrei P., Stoffel K. Rule extraction from time series databases using classification trees // Proceedings of the 20th IASTED Conference on Applied Informatics, Innsbruck, Austria, February 18-21, 2002. – Calgary, Canada: ACTA Press, 2002. – P. 327-332.
25. Crespo F., Weber R. A methodology for dynamic data mining based on fuzzy clustering // Fuzzy Sets and Systems. – Vol.150 (2005), P. 267-284.
26. Das G., Lin K.-I., Mannila H., ... [etc]. Rule discovery from time series // Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining (KDD-98), New York City, USA, August 27-31, 1998. – NY: AAAI Press, 1998. – P. 16-22.
27. Fager J. Online Policy-Gradient Reinforcement Learning using OLGARB for SpaceWar. // Technical report, University of Oklahoma, 660 Parrington Oval, Norman, OK 73019 USA. 2006. –P.5.
28. Franklin S., Graesser A. Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents // Proceedings of the Workshop on Intelligent Agents III, Agent Theories, Architectures, and Languages, ECAI '96, Hungary, Budabest, August 11-16, 1996. – London: Springer-Verlag, 1997. – P.21-35.
29. Ganzhorn D., de Beaumont W. Learning Algorithms and Quake // Technical report, University of Rochester, March 19th, 2004. –P.13.

30. Gearhart C. Genetic Programming as Policy Search in Markov Decision Processes // Genetic Algorithms and Genetic Programming at Stanford.– (2003), P. 61-67.
31. Goto J., Lewis M.E., Puterman M.L. A Markov Decision Process Model for Airline Meal Provisioning // Transportation Science. – Vol. 38, No. 1 (2004), P. 107-118.
32. Guestrin C., Koller D., Gearhart C., ... [etc]. Generalizing Plans to New Environments in Relational MDPs // Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence, Acapulco, Mexico, August 9-15, 2003. – San Francisco, USA: Morgan Kaufmann, 2003. – P. 1003-1010.
33. Hassan Md.R., Nath B. Stock market forecasting using hidden Markov model: A new approach // Proceedings of the 5th International Conference on Intelligent Systems Design and Applications, ISDA'05, Wroclaw, Poland, 8-10 September, 2005.– Washington, USA: IEEE Computer Society, 2005. – P. 192-196.
34. Haykin S. Neural Networks and Learning Machines (3rd Edition).– New Jersey: Prentice Hall, 2008.– 936 p.
35. Hewitt C. Viewing Control Structures as Patterns of Passing Messages // Artificial Intelligence. – Vol. 8(3) (1977), P. 323-364.
36. Howard R.A. Dynamic Programming and Markov Processes.– Cambridge, MA: MIT Press, 1960.– 136 p.
37. Jacobs S. Applying ReadyLog to Agent Programming in Interactive Computer Games // Diplomarbeit, Fakultät für Mathematik, Informatik und Naturwissenschaften der Rheinisch-Westfälischen Technischen Hochschule Aachen, 2005.
38. Kampen van N.G. Remarks on Non-Markov Processes // Brazilian Journal of Physics.–Vol. 28, Nr 2 (1998), P. 90-96.
39. Keogh E., Lin J., Truppel W. Clustering of Time Series Subsequences is Meaningless: Implications for Previous and Future Research // Proceedings of the 3rd IEEE International Conference on Data Mining (ICDM 2003), Melbourne, Florida, USA November 19-22, 2003.– Melbourne: IEEE Computer Society, 2003.– P. 115-122.
40. Konar A. Computational Intelligence: Principles, Techniques and Applications.– Berlin Heidelberg: Springer-Verlag, 2005.– 732 p.
41. Krzysztof L. Markov Decision Processes in Finance // Master's Thesis, Department of Mathematics Vrije Universiteit Amsterdam, 2006.
42. Lazaric A., Taylor M.E. Transfer Learning in Reinforcement Learning Domains // Lecture material of European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases '09, Bled, Slovenia, September 7-11, 2009.
43. Lee S.-I., Chun S.-J., Kim D.-H., Lee J.-H., ... [etc]. Similarity Search for Multidimensional Data Sequences // Proceedings of IEEE 16th International Conference on Data Engineering, San Diego, USA, 28 February - 3 March, 2000.– IEEE Computer Society, 2000.–P.599-608.
44. Li C., Wang H., Zhang Y. Dynamic Pricing Decision in a Duopolistic Retailing Market // Proceedings of the 6th World Congress on Intelligent Control and Automation, Dalian, China, 21-23 June 2006.– IEEE, 2006.– P. 6993-6997.
45. Lin L-J. Reinforcement learning for Robots Using Neural Networks // PhD thesis, Carnegie Mellon University, Pittsburgh, CMU-CS-93-103, 1993.
46. Lind J. Issues in Agent-Oriented Software Engineering // Agent-Oriented Software Engineering: First International Workshop, AOSE 2000. Lectures Notes in Artificial Intelligence. – Vol. 1957 (2001), P. 45-58.

47. Melo F.S., Ribeiro M.I. Coordinated Learning in Multiagent MDPs with Infinite State-Space // *Autonomous agents and multi-agent systems.* – Vol. 21, Number 3 (2010), P. 321-367.
48. Mitkus S., Trinkūnienė E. Reasoned Decisions In Construction Contracts Evaluation // *Baltic Journal on Sustainability.* – Vol. 14, Nr3 (2008), P. 402-416.
49. Narahari Y., Raju C., Ravikumar K., ... [etc]. Dynamic Pricing Models for Electronic Business. *Sadhana* // *Sadhana.* – Vol. 30, Part 2 & 3 (2005), P. 231–256.
50. Palit A.K., Popovic D. Computational Intelligence in Time Series Forecasting. Theory and Engineering Applications. – London: Springer, 2005. – 372 p.
51. Povinelli R.J., Xin F. Temporal pattern identification of time series data using pattern wavelets and genetic algorithms // *Artificial Neural Networks in Engineering.* – New York: ASME Press, 1998. – P.691-696.
52. Powell W.B. Approximate Dynamic Programming I: Modeling // *Encyclopedia of Operations Research and Management Science.* – John Wiley and Sons, 2011. – P. 1-11.
53. Pranevičius H., Budnikas G. PLA-Based Formalization Of Business Rules And Their Analysis By Means Of Knowledge-Based Techniques // *Baltic Journal on Sustainability.* – Vol. 14, Nr 3 (2008), P. 328-343.
54. Puterman M.L. Markov Decision Processes: Discrete Stochastic Dynamic Programming. – New York: John Wiley & Sons, 1994. – 649 p.
55. Pyeatt L.D., Howe A.E. Decision Tree Function Approximation in Reinforcement Learning // *Proceedings of the Third International Symposium on Adaptive Systems: Evolutionary Computation and Probabilistic Graphical Models*, Vol. 2, Issue: 1/2 (2001). – Citeseer 2001. – P. 70-77.
56. Reinartz W. Customising Prices in Online Markets // *European Business Forum.* – Issue 6 (2001).– P. 35-41.
57. Rosu I. The Bellman Principle of Optimality // *Pieejas veids: tīmeklis WWW.* URL: <http://appli8.hec.fr/rosu/research/notes/bellman.pdf>. – Resurss apraksts 2011.g. 12.dec.
58. Russel S., Norvig P. Artificial Intelligence: A modern approach, 2nd edition.– New Jersey: Prentice Hall, 2003. – 1132 p.
59. Song H., Liu C.-C. Optimal Electricity Supply Bidding by Markov Decision Process // *IEEE transactions on power systems.* – Vol. 15, Nr. 2 (2000). – P. 618-624.
60. Sonnenberg F.A., Beck J.R. Markov models in medical decision making: a practical guide // *Medical Decision Making.* – Vol. 13, Nr 4 (1993). – P. 322-338.
61. Sunderejan R., Kumar A.P.N., Badri K.K.N., ... [etc]. Stock market trend prediction using Markov models // *Electron.* – Vol.1, Issue 1 (2009). – P. 285-289.
62. Sutton R.S., Barto A.R. Reinforcement learning. An Introduction. – Cambridge, MA: MIT Press, 1998. – 342 p.
63. Sutton R.S., McAllester D., Singh S., ... [etc]. Policy Gradient Methods for Reinforcement Learning with Function Approximation // *Advances in Neural Information Processing Systems.* – Vol. 12 (2000). – P. 1057-1063.
64. Symeonidis A.L., Kehagias D., Mitkas P.A. Intelligent policy recommendations on enterprise resource planning by the use of agent technology and data mining techniques // *Expert Systems with Applications.* – N 25 (2003). – P. 589-602.
65. Taylor M.E. Transfer in Reinforcement Learning Domains. *Studies in Computational Intelligence.* – Berlin: Springer-Verlag, 2009. – 244 p.

66. Tokic M. Exploration and Exploitation Techniques in Reinforcement Learning // Invited lecture, Ravensburg-Weingarten University of Applied Sciences, Germany, Nov. 2008.
67. Vanthienen J. Ruling the business: About business rules and decision tables // New Directions in Software Engineering. – (2001) pp. 103-120.
68. Varges S., Riccardi G., Quarteroni S., ... [etc]. The exploration/exploitation trade-off in Reinforcement Learning for dialogue management // Proceedings of IEEE Workshop on Automatic Speech Recognition & Understanding ASRU'09, Merano, Italy, December 13-17, 2009. – IEEE Signal Processing Society, 2009. – P. 479-484.
69. Varian H.R. Differential Pricing and Efficiency // First Monday.–Vol.1, Nr.2(1996).–P.1-10.
70. Vengerov D. A Gradient-Based Reinforcement Learning Approach to Dynamic Pricing in Partially-Observable Environment // Future Generation Comp. Syst. – Vol. 24/7 (2008). –P. 687-693.
71. Wang R., Sun L., Ruan X.-G., ... [etc]. Control of Inverted Pendulum Based on Reinforcement Learning and Internally Recurrent Net // Proceedings of International Conference on Intelligent Computing, HeFei, China, August 23-26, 2005. – IEEE Computational Intelligence Society, 2005. – P. 2133-2142.
72. Weld M., Weld D. Solving Concurrent Markov Decision Processes // Proceedings of the 19th national conference on Artificial intelligence, San Jose, California, S.J. Convention Center, July 25-29, 2004. – San Jose, California: AAAI Press, 2004. – P. 716-722.
73. Witten I.H., Frank E. Data Mining: Practical Machine Learning Tools and Techniques, Second Edition – Morgan Kaufman, 2005. – 560 p.
74. Беллман Р., Дрейфус С. Прикладные задачи динамического программирования – Москва: «Наука», 1965.– 460 с.
75. Бережная Е.В., Бережной В.И. Математические методы моделирования экономических систем. – Москва: «Финансы и статистика», 2006. – 432 с.
76. Воробьев Н.Н. Предисловие редактора перевода к Беллман Р. Динамическое программирование. – Москва: «Издательство иностранной литературы», 1960. – 400 с.
77. Горбань А.Н. Обобщенная аппроксимационная теорема и вычислительные возможности нейронных сетей // Сибирский журнал вычислительной математики. – Т.1, Nr.1 (1998), с. 12-24.
78. Кузнецов Ю.Н., Кузубов В.И., Волощенко А.Б. Математическое программирование: Учеб. пособие. – 2-е изд., перераб. и доп. – Москва: «Высшая школа», 1980. – 300 с.
79. Растринин Л.А. Современные принципы управления сложными объектами.– Москва: «Советское радио», 1980. – 232 с.
80. Таха Х.А. Введение в исследование операций.– 7-е. изд. – М.: «Вильямс», 2005.–912 с.
81. Тёрнер Д. Вероятность, статистика и исследование операций. – Москва: «Статистика», 1976. – 431 с.
82. Фомин Г.П. Математические методы и модели в коммерческой деятельности. – Москва: «Финансы и статистика», 2005. – 616 с.
83. Черноушко Ф.Л. Динамическое программирование // Соросовский образовательный журнал. – № 2 (1998), с. 139-144.

Jurijs ČIŽOVS

**DINAMISKĀS SISTĒMAS VADĀMĀ MARKOVA
LĒMUMU MODEĻA UZBŪVE UN IZPĒTE AR
DATU IEGUVES METODĒM**

Promocijas darba kopsavilkums

Parakstīts iespiešanai 2011. 28.12. Reģ. apl. Nr. 2-0282.
Formāts 60x84/16. Ofseta papīrs. Ofseta papīrs. 11,25 iesp.l.,
8,12 uzsk.izd.l. Metiens 100 eks. Pasūt. Nr. 164.
Iespiests RTU tipogrāfijā, Rīga LV-1658, Kaļķu ielā 1.